

## References

- [1] F. Aarts, H. Kuppens, J. Tretmans, F. Vaandrager, and S. Verwer. Learning and testing the bounded retransmission protocol. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 4–18, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [2] H. Abasi, A. Z. Abdi, and N. H. Bshouty. Learning Boolean halfspaces with small weights from membership queries. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 96–110, Berlin, Heidelberg, New York, 2014. Springer.
- [3] H. Abasi, A. Z. Abdi, and N. H. Bshouty. Learning boolean halfspaces with small weights from membership queries. *Theoret. Comput. Sci.*, 650:45–56, 2016. Special Issue: Algorithmic Learning Theory.
- [4] H. Abasi, N. H. Bshouty, and H. Mazzawi. On exact learning monotone DNF from membership queries. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 111–124, Berlin, Heidelberg, New York, 2014. Springer.
- [5] H. Abasi, N. H. Bshouty, and H. Mazzawi. Non-adaptive learning of a hidden hypergraph. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 89–101, Berlin, Heidelberg, New York, 2015. Springer.
- [6] H. Abasi, N. H. Bshouty, and H. Mazzawi. Non-adaptive learning of a hidden hypergraph. *Theoret. Comput. Sci.*, 716:15–27, 2018. Special Issue on ALT 2015.
- [7] Y. Abbasi-Yadkori and C. Szepesvári. Regret bounds for the adaptive control of linear quadratic systems. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 1–26, Budapest, Hungary, 09–11 June 2011. PMLR.
- [8] E. Abboud, N. Agha, N. H. Bshouty, N. Radwan, and F. Saleh. Learning threshold functions with small weights using membership queries. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 318–322, New York, NY, 1999. ACM Press.
- [9] S. Abdallah, M. Sandler, C. Rhodes, and M. Casey. Using duration models to reduce fragmentation in audio segmentation. *Machine Learning*, 65(2-3):485–515, 2006.
- [10] H. Abe and T. Yamaguchi. Constructing inductive applications by meta-learning with method repositories. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 576–585. Springer, 2001.
- [11] N. Abe. *Polynomial Learnability as a Formal Model of Natural Language Acquisition*. PhD thesis, Department of Computer and Information Science, University of Pennsylvania, 1989.

- [12] N. Abe. Polynomial learning of semilinear sets. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 25–40, San Mateo, CA, 1989. Morgan Kaufmann.
- [13] N. Abe. Learning commutative deterministic finite state automata in polynomial time. *New Generation Computing*, 8(4):319–336, 1991.
- [14] N. Abe. Preservation of predictability under polynomially sparse variations and its applications. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 163–174. IOS Press, 1992.
- [15] N. Abe. On the sample complexity of various learning strategies in the probabilistic PAC learning paradigms. In *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 89–106. Springer-Verlag, 1993.
- [16] N. Abe. Characterizing PAC-learnability of semilinear sets. *Inform. Comput.*, 116(1):81–102, 1995.
- [17] N. Abe. Towards realistic theories of learning. *New Generation Computing*, 15(1):3–26, 1997. Special issue for AII '94 and ALT '94.
- [18] N. Abe and R. Khardon. Foreword. *Theoret. Comput. Sci.*, 313(2):173–174, 2004. Special issue for ALT 2001.
- [19] N. Abe, R. Khardon, and T. Zeugmann, editors. *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Nov. 2001. Springer.
- [20] N. Abe, R. Khardon, and T. Zeugmann. Editors' introduction. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 1–7. Springer, 2001.
- [21] N. Abe and H. Li. Learning word association norms using tree cut pair models. In *Proc. 13th International Conference on Machine Learning*, pages 3–11. Morgan Kaufmann, 1996.
- [22] N. Abe, H. Li, and A. Nakamura. On-line learning of binary lexical relations using two-dimensional weighted majority algorithms. In *Proc. 12th International Conference on Machine Learning*, pages 3–11. Morgan Kaufmann, 1995.
- [23] N. Abe and P. M. Long. Associative reinforcement learning using linear probabilistic concepts. In *Proc. 16th International Conf. on Machine Learning*, pages 3–11. Morgan Kaufmann, San Francisco, CA, 1999.
- [24] N. Abe and H. Mamitsuka. A new method for predicting protein secondary structures based on stochastic tree grammars. In *Proc. 11th International Conference on Machine Learning*, pages 3–11. Morgan Kaufmann, 1994.
- [25] N. Abe and H. Mamitsuka. Predicting protein secondary structure using stochastic tree grammars. *Machine Learning*, 29(2-3):275–301, 1997.
- [26] N. Abe and H. Mamitsuka. Query learning strategies using boosting and bagging. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 1–9. Morgan Kaufmann, San Francisco, CA, 1998.

- [27] N. Abe and A. Nakamura. Learning to optimally schedule internet banner advertisements. In *Proc. 16th International Conf. on Machine Learning*, pages 12–21. Morgan Kaufmann, San Francisco, CA, 1999.
- [28] N. Abe and J. Takeuchi. The ‘lob-pass’ problem and an on-line learning model of rational choice. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 422–428, New York, NY, 1993. ACM Press.
- [29] N. Abe, J. Takeuchi, and M. K. Warmuth. Polynomial learnability of probabilistic concepts with respect to the Kullback-Leibler divergence. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 277–289, San Mateo, CA, 1991. Morgan Kaufmann.
- [30] N. Abe, J. Takeuchi, and M. K. Warmuth. Polynomial learnability of stochastic rules with respect to the KL-divergence and quadratic distance. *IEICE Transactions on Information and Systems*, E84-D(3):299–316, 2001.
- [31] N. Abe and M. K. Warmuth. On the computational complexity of approximating distributions by probabilistic automata. *Machine Learning*, 9:205–260, 1992.
- [32] N. Abe and O. Watanabe. Polynomially sparse variations and reducibility among prediction problems. *IEICE Trans. Inf. and Syst.*, E75-D(4):449–458, July 1992.
- [33] J. Abela, F. Coste, and S. Spina. Mutually compatible and incompatible merges for the search of the smallest consistent DFA. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 28–39. Springer, 2004.
- [34] J. Abernethy, P. Bartlett, and A. Rakhlin. Multitask learning with expert advice. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 484–498, Berlin, 2007. Springer.
- [35] J. Abernethy, P. L. Bartlett, N. Buchbinder, and I. Stanton. A regularization approach to metrical task systems. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 270–284, Berlin, Heidelberg, New York, 2010. Springer.
- [36] J. Abernethy, P. L. Bartlett, and E. Hazan. Blackwell approachability and no-regret learning are equivalent. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 27–46, Budapest, Hungary, 09–11 June 2011. PMLR.
- [37] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari. Optimal strategies and minimax lower bounds for online convex games. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 415–424. Omnipress, 2008.
- [38] J. Abernethy, E. Hazan, and A. Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 263–274. Omnipress, 2008.

- [39] J. Abernethy, J. Langford, and M. K. Warmuth. Continuous experts and the binning algorithm. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 544–558, Berlin, 2006. Springer.
- [40] J. Abernethy, C. Lee, A. Sinha, and A. Tewari. Online linear optimization via smoothing. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 807–823, Barcelona, Spain, 13–15 June 2014. PMLR.
- [41] J. Abernethy and S. Mannor. Does an efficient calibrated forecasting strategy exist? In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 809–812, Budapest, Hungary, 09–11 June 2011. PMLR.
- [42] J. Abernethy, M. K. Warmuth, and J. Yellin. When random play is optimal against an adversary. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 437–446. Omnipress, 2008.
- [43] J. D. Abernethy and R. M. Frongillo. A characterization of scoring rules for linear properties. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 27.1–27.13, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [44] P. J. Abisha, D. G. Thomas, and S. J. Kumaar. Learning subclasses of pure pattern languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 280–282. Springer, 2008.
- [45] I. Abraham, O. Alonso, V. Kandylas, and A. Slivkins. Adaptive crowdsourcing algorithms for the bandit survey problem. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 882–910. JMLR.org, 2013.
- [46] Y. S. Abu-Mostafa. The Vapnik-Chervonenkis dimension: Information verses complexity in learning. *Neural Computation*, 1(3):312–317, 1989.
- [47] J. Acharya, H. Das, A. Jafarpour, A. Orlitsky, and S. Pan. Competitive closeness testing. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 47–68, Budapest, Hungary, 09–11 June 2011. PMLR.
- [48] J. Acharya, H. Das, A. Jafarpour, A. Orlitsky, S. Pan, and A. Suresh. Competitive classification and closeness testing. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 22.1–22.18, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [49] J. Acharya, A. Jafarpour, A. Orlitsky, and A. T. Suresh. Optimal probability estimation with applications to prediction and classification. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 764–796. JMLR.org, 2013.

- [50] D. Achlioptas and F. McSherry. On spectral learning of mixtures of distributions. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 458–469. Springer, 2005.
- [51] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive Science*, 9:147–169, 1985.
- [52] R. T. Adams. Inference of LISP programs from examples. Bachelor’s thesis, MIT EECS Department, June 1990.
- [53] D. Adamskiy, W. M. Koolen, A. Chernov, and V. Vovk. A closer look at adaptive regret. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 290–304, Berlin, Heidelberg, New York, 2012. Springer.
- [54] H. Adé, L. De Raedt, and M. Bruynooghe. Declarative bias for specific-to-general ILP systems. *Machine Learning*, 20:119–154, 1995.
- [55] L. M. Adleman and M. Blum. Inductive inference and unsolvability. *J. Symbolic Logic*, 56(3):891–900, 1991.
- [56] P. Adriaans. ML techniques and text analysis. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 460–469. Springer-Verlag, 1993.
- [57] P. Adriaans. Learning deterministic DEC grammars is learning rational numbers. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 320–326, Berlin, 2006. Springer.
- [58] P. Adriaans and C. Jacobs. Using MDL for grammar induction. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 293–306, Berlin, 2006. Springer.
- [59] P. Adriaans and W. Mulder. MDL in the limit. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 258–261. Springer, 2010.
- [60] P. Adriaans and M. Vervoort. The EMILE 4.1 grammar induction toolbox. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 293–295. Springer, 2002.
- [61] P. W. Adriaans, H. Fernau, and M. van Zaanen, editors. *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, Berlin, 2002. Springer.
- [62] A. Agarwal and S. Agarwal. On consistent surrogate risk minimization and property elicitation. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 4–22, Paris, France, 03–06 July 2015. PMLR.

- [63] A. Agarwal, S. Agarwal, S. Assadi, and S. Khanna. Learning with limited rounds of adaptivity: Coin tossing, multi-armed bandits, and ranking from pairwise comparisons. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 39–75, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [64] A. Agarwal, A. Anandkumar, P. Jain, P. Netrapalli, and R. Tandon. Learning sparsely used overcomplete dictionaries. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 123–137, Barcelona, Spain, 13–15 June 2014. PMLR.
- [65] A. Agarwal, A. Badanidiyuru, M. Dudk, R. E. Schapire, and A. Slivkins. Robust multi-objective learning with mentor feedback. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 726–741, Barcelona, Spain, 13–15 June 2014. PMLR.
- [66] A. Agarwal, J. C. Duchi, P. L. Bartlett, and C. Levrard. Oracle inequalities for computationally budgeted model selection. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 69–86, Budapest, Hungary, 09–11 June 2011. PMLR.
- [67] A. Agarwal, A. Krishnamurthy, J. Langford, H. Luo, and S. R. E. Open problem: First-order regret bounds for contextual bandits. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 4–7, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [68] A. Agarwal, H. Luo, B. Neyshabur, and R. E. Schapire. Corraling a band of bandit algorithms. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 12–38, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [69] S. Agarwal. Generalization bounds for some ordinal regression algorithms. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 7–21, Berlin, Oct. 2008. Springer.
- [70] S. Agarwal. Surrogate regret bounds for the area under the ROC curve via strongly proper losses. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 338–353. JMLR.org, 2013.
- [71] S. Agarwal, T. Graepel, R. Herbrich, S. Har-Peled, and D. Roth. Generalization bounds for the area under the ROC curve. *Journal of Machine Learning Research*, 6:393–425, 2005.
- [72] S. Agarwal and P. Niyogi. Stability and generalization of bipartite ranking algorithms. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 32–47. Springer, 2005.
- [73] S. Agarwal and D. Roth. Learnability of bipartite ranking functions. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 16–31. Springer, 2005.

- [74] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo. Fast discovery of association rules. In *Advances in Knowledge Discovery and Data Mining*, pages 307–328. American Association for Artificial Intelligence, Menlo Park, CA, USA, 1996.
- [75] S. Agrawal, V. Avadhanula, V. Goyal, and A. Zeevi. Thompson sampling for the MNL-bandit. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 76–78, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [76] S. Agrawal, N. R. Devanur, and L. Li. An efficient algorithm for contextual bandits with knapsacks, and an extension to concave objectives. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 4–18, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [77] S. Agrawal and N. Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 39.1–39.26, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [78] D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.
- [79] W. Aha, David, S. Lapointe, C. X. Ling, and S. Matwin. Learning recursive relations with randomly selected small training sets. In *Proc. 11th International Conference on Machine Learning*, pages 12–18. Morgan Kaufmann, 1994.
- [80] S. Ahmad. A study of scaling and generalization in neural networks. Technical Report UIUCDCS-R-88-1454, University of Illinois at Urbana-Champaign, Sept. 1988.
- [81] W. Aiello and M. Mihail. Learning the fourier spectrum of probabilistic lists and trees. In *Proc. SODA 91*, pages 291–299, Boston, MA, Jan. 1991. AAAI Press/The MIT Press.
- [82] N. Ailon. Learning and optimizing with preferences. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 13–21, Berlin, Heidelberg, New York, 2013. Springer.
- [83] N. Ailon, R. Begleiter, and E. Ezra. Active learning using smooth relative regret approximations with applications. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 19.1–19.20, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [84] N. Ailon, K. Hatano, and E. Takimoto. Bandit online optimization over the permutahedron. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 215–229, Berlin, Heidelberg, New York, 2014. Springer.
- [85] N. Ailon, K. Hatano, and E. Takimoto. Bandit online optimization over the permutahedron. *Theoret. Comput. Sci.*, 650:92–108, 2016. Special Issue: Algorithmic Learning Theory.
- [86] N. Ailon and M. Mohri. An efficient reduction of ranking to classification. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 87–98. Omnipress, 2008.

- [87] E. Aïmeur and M. Vézeau. Short-term profiling for a case-based reasoning. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 23–30. Springer, 2000.
- [88] F. Aiolli and A. Sperduti. Multiclass classification with multi-prototype support vector machines. *Journal of Machine Learning Research*, 6:817–850, 2005.
- [89] H. Aït-Kaci and Y. Sasaki. An axiomatic approach to feature term generalization. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 1–12. Springer, 2001.
- [90] H. Aizenstein. *On the Learnability of Disjunctive Normal Form Formulas and Decision Trees*. PhD thesis, Dept of Computer Science, University of Illinois at Urbana-Champaign, 1993. Technical report UIUCDCS-R-93-1813, June, 1993, Urbana, Illinois.
- [91] H. Aizenstein, A. Blum, R. Khardon, E. Kushilevitz, L. Pitt, and D. Roth. On learning read- $k$ -satisfy- $j$  DNF. *SIAM J. Comput.*, 27(6):1515–1530, 1998.
- [92] H. Aizenstein, L. Hellerstein, and L. Pitt. Read-thrice DNF is hard to learn with membership and equivalence queries. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 523–532. IEEE Computer Society Press, Los Alamitos, CA, 1992.
- [93] H. Aizenstein and L. Pitt. Exact learning of read-twice DNF formulas. In *Proc. 32th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 170–179. IEEE Computer Society Press, 1991.
- [94] H. Aizenstein and L. Pitt. Exact learning of read- $k$  disjoint DNF and not-so-disjoint DNF. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 71–76, New York, NY, 1992. ACM Press.
- [95] H. Aizenstein and L. Pitt. On the learnability of disjunctive normal form formulas. *Machine Learning*, 19(3):183–208, 1995.
- [96] S. Akaho. Regularization learning of neural networks for generalization. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 99–110. Springer-Verlag, 1993.
- [97] Y. Akama. Commutative regular shuffle closed languages, noetherian property, and learning theory. In A. H. Dediu, A. M. Ionescu, and C. Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2009, Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 93–104. Springer, 2009.
- [98] Y. Akama. Set systems: Order types, continuous nondeterministic deformations, and quasi-orders. *Theoret. Comput. Sci.*, 412(45):6235–6251, 2011.
- [99] Y. Akama and T. Zeugmann. Consistency conditions for inductive inference of recursive functions. In *New Frontiers in Artificial Intelligence, JSAI 2006 Conference and Workshops, Tokyo, Japan, June 2006, Revised Selected Papers*, volume 4384 of *Lecture Notes in Artificial Intelligence*, pages 251–264, Berlin, 2007. Springer.
- [100] Y. Akama and T. Zeugmann. Consistent and coherent learning with  $\delta$ -delay. Technical Report TCS-TR-A-07-29, Division of Computer Science, Hokkaido University, Oct. 2007.



- [101] Y. Akama and T. Zeugmann. Consistent and coherent learning with  $\delta$ -delay. *Inform. Comput.*, 206(11):1362–1374, 2008.
- [102] A. Akkus and H. A. Güvenir. K nearest neighbor classification on feature projections. In *Proc. 13th International Conference on Machine Learning*, pages 12–19. Morgan Kaufmann, 1996.
- [103] H. I. Akram, C. de la Higuera, H. Xiao, and C. Eckert. Grammatical inference algorithms in MATLAB. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 262–266. Springer, 2010.
- [104] H. I. Akram, C. de la Higuera, and C. Eckert. Actively learning probabilistic subsequential transducers. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 19–33, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [105] T. Akutsu, S. Kuhara, O. Maruyama, and S. Miyano. Identification of genetic networks by strategic gene disruptions and gene overexpressions under a boolean model. *Theoret. Comput. Sci.*, 298(1):235–251, 2003.
- [106] T. Akutsu, S. Miyano, and S. Kuhara. A simple greedy algorithm for finding functional relations: Efficient implementation and average case analysis. In *Discovery Science, Third International Conference, DS 2000, Kyoto, Japan, December 2000, Proceedings*, volume 1967 of *Lecture Notes in Artificial Intelligence*, pages 86–98. Springer, 2000.
- [107] T. Akutsu, S. Miyano, and S. Kuhara. A simple greedy algorithm for finding functional relations: efficient implementation and average case analysis,. *Theoret. Comput. Sci.*, 292(2):481–495, 2003.
- [108] T. Akutsu and A. Takasu. On PAC learnability of functional dependencies. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 229–239. Springer-Verlag, 1993.
- [109] T. Akutsu, T. Tamura, and K. Horimoto. Completing networks using observed data. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 126–140, Berlin/Heidelberg, 2009. Springer.
- [110] M. Alamgir, G. Lugosi, and U. Luxburg. Density-preserving quantization with application to graph downsampling. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 543–559, Barcelona, Spain, 13–15 June 2014. PMLR.
- [111] A. E. Alaoui, X. Cheng, A. Ramdas, M. J. Wainwright, and M. I. Jordan. Asymptotic behavior of  $\ell_p$ -based Laplacian regularization in semi-supervised learning. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 879–906, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [112] S. Albers and M. R. Henzinger. Exploring unknown environments. *SIAM J. Comput.*, 29(4):1164–1188, 1999.

- [113] A. Albrecht. On lower bounds for the depth of threshold circuits with weights from  $\{-1, 0, +1\}$ . In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 404–416. Springer, 1995.
- [114] A. A. Albrecht and K. Steinhöfel. A simulated annealing-based learning algorithm for Boolean DNF. In *Advanced Topics in Artificial Intelligence, 12th Australian Joint Conference on Artificial Intelligence, AI '99, Sydney, Australia, December 6-10, 1999, Proceedings*, volume 1747 of *Lecture Notes in Artificial Intelligence*, pages 193–204. Springer, 1999.
- [115] D. Aldous and U. Vazirani. A Markovian extension of Valiant’s learning model. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 392–396. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [116] D. Aldous and U. Vazirani. A Markovian extension of valiant’s learning model. *Inform. Comput.*, 117(2):181–186, 1995.
- [117] M. Alekhnovich, M. Braverman, V. Feldman, A. R. Klivans, and T. Pitassi. Learnability and automatizability. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 621–630. IEEE Computer Society, 2004.
- [118] M. Alekhnovich, M. Braverman, V. Feldman, A. R. Klivans, and T. Pitassi. The complexity of properly learning simple concept classes. *J. Comput. Syst. Sci.*, 74(1):16–34, 2008.
- [119] R. Aler, D. Borrajo, and P. Isasi. Genetic programming and deductive-inductive learning: a multi-strategy approach. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 10–18. Morgan Kaufmann, San Francisco, CA, 1998.
- [120] R. Aler, D. Borrajo, and P. Isasi. Knowledge representation issues in control knowledge learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1–8. Morgan Kaufmann, San Francisco, CA, 2000.
- [121] K. M. Ali and M. J. Pazzani. Reducing the small disjuncts problem by learning probabilistic concept descriptions. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 10, pages 183–199. MIT Press, 1995.
- [122] K. M. Ali and M. J. Pazzani. Error reduction through learning multiple descriptions. *Machine Learning*, 24:173–202, 1996.
- [123] M. Aliakbarpour, E. Blais, and R. Rubinfeld. Learning and testing junta distributions. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 19–46, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [124] T. Van Allen and R. Greiner. Model selection criteria for learning belief nets: An empirical comparison. In *Proc. 17th International Conf. on Machine Learning*, pages 1047–1054. Morgan Kaufmann, San Francisco, CA, 2000.
- [125] C. Allenberg. Individual sequence prediction - upper bounds and application for complexity. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 233–242, New York, NY, 1999. ACM Press.
- [126] C. Allenberg, P. Auer, L. Györfi, and G. Ottucsák. Hannan consistency in on-line learning in case of unbounded losses under partial monitoring. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 229–243. Springer, Oct. 2006.

- [127] C. Allenberg-Neeman and B. Neeman. Full information game with gains and losses. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 264–278. Springer, 2004.
- [128] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of Machine Learning Research*, 1:113–141, 2000.
- [129] L. B. Almeida. MISEP – linear and nonlinear ICA based on mutual information. *Journal of Machine Learning Research*, 4:1297–1318, 2003.
- [130] H. Almuallim. An efficient algorithm for optimal pruning of decision trees. *Artificial Intelligence*, 83(2):347–362, 1996.
- [131] H. Almuallim, Y. Akiba, and S. Kaneda. On handling tree-structured attributes in decision tree learning. In *Proc. 12th International Conference on Machine Learning*, pages 12–20. Morgan Kaufmann, 1995.
- [132] H. Almuallim, Y. Akiba, T. Yamazaki, and S. Kaneda. Learning verb translation rules from ambiguous examples and a large semantic hierarchy. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 18, pages 323–336. MIT Press, 1997.
- [133] H. Almuallim and T. G. Dietterich. Learning with many irrelevant features. In *Proc. Ninth National Conference on Artificial Intelligence*, pages 547–552. AAAI Press, 1991.
- [134] N. Alon and V. Asodi. Learning a hidden subgraph. In *Automata, Languages and Programming: 31st International Colloquium, ICALP 2004, Turku, Finland, July 12-16, 2004. Proceedings*, volume 3142 of *Lecture Notes in Computer Science*, pages 110–121. Springer, 2004.
- [135] N. Alon, R. Beigel, S. Kasif, S. Rudich, and B. Sudakov. Learning a hidden matching. In *43rd Annual Symposium on Foundations of Computer Science, 6-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 197–206. IEEE Computer Society, 2002.
- [136] N. Alon, R. Beigel, S. Kasif, S. Rudich, and B. Sudakov. Learning a hidden matching. *SIAM J. Comput.*, 33(2):487–501, 2004.
- [137] N. Alon, S. Ben-David, N. Cesa-Bianchi, and D. Haussler. Scale-sensitive dimensions, uniform convergence, and learnability. *J. ACM*, 44(4):616–631, 1997.
- [138] N. Alon, N. Cesa-Bianchi, O. Dekel, and T. Koren. Online learning with feedback graphs: Beyond bandits. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 23–35, Paris, France, 03–06 July 2015. PMLR.
- [139] N. Alon, S. Dar, M. Parnas, and D. Ron. Testing of clustering. In *Proc. 41st Annual Symposium on Foundations of Computer Science*, pages 240–250. IEEE Computer Society Press, Los Alamitos, CA, 2000.
- [140] N. Alon, S. Moran, and A. Yehudayoff. Sign rank versus VC dimension. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 47–80, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [141] C. L. Alonso and J. L. Montaña. Vapnik-chervonenkis dimension of parallel arithmetic computations. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 107–119, Berlin, Oct. 2007. Springer.
- [142] R. Alquézar, A. Sanfeliu, and J. Cueva. Learning of context-sensitive language acceptors through regular inference and constrained induction. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 134–145. Springer, 1996.
- [143] P. Alquier. An algorithm for iterative selection of blocks of features. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 35–49, Berlin, Heidelberg, New York, 2010. Springer.
- [144] P. Alquier. Bayesian methods for low-rank matrix estimation: Short survey and theoretical study. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 309–323, Berlin, Heidelberg, New York, 2013. Springer.
- [145] J. Alspector and R. B. Allen. Performance of a stochastic learning chip. In *Proc. 1988 Neural Information Processing Conference*, pages 748–760, 1989.
- [146] Y. Altun and A. Smola. Unifying divergence minimization and statistical inference via convex duality. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 139–153, Berlin, 2006. Springer.
- [147] G. Alvarez, P. García, and J. Ruiz. A merging states algorithm for inference of RFSAs. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 340–341, Berlin, 2006. Springer.
- [148] G. I. Alvarez, J. H. Victoria, E. Bravo, and P. García. A non-deterministic grammar inference algorithm applied to the cleavage site prediction problem in bioinformatics. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 267–270. Springer, 2010.
- [149] K. Amakawa. Self-learning reaching motion of a multi-joint arm using a trial-and-error heuristic and a neural network. Master’s thesis, UC Santa Cruz, 1991.
- [150] K. Amano and A. Maruoka. On learning monotone Boolean functions under the uniform distribution. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 57–68. Springer, 2002.
- [151] K. Amano and A. Maruoka. On learning monotone Boolean functions under the uniform distribution. *Theoret. Comput. Sci.*, 350(1):3–12, 2006. Special issue for ALT 2002.

- [152] R. A. Amar, D. R. Dooly, S. A. Goldman, and Q. Zhang. Multiple-instance learning of real-valued data. In *Proc. 18th International Conf. on Machine Learning*, pages 3–10. Morgan Kaufmann, San Francisco, CA, 2001.
- [153] S. Amari. Mathematical theory of neural learning. *New Generation Computing*, 8(4):281–294, 1991.
- [154] S. Amari. The EM algorithm and information geometry in neural network learning. *Neural Computation*, 7(1):13–18, Jan. 1995.
- [155] S. Amari, N. Fujita, and S. Shinomoto. Four types of learning curves. *Neural Computation*, 4(4):605–618, 1992.
- [156] S. Amari and S. Wu. Improving support vector machine classifiers by modifying kernel functions. *Neural Networks*, 12(6):783–789, 1999.
- [157] S. ichi Amari. Information geometry of the EM and em algorithms for neural networks. *Neural Networks*, 8(9):1379–1408, 1995.
- [158] S. ichi Amari, T. ping Chen, and A. Cichocki. Stability analysis of learning algorithms for blind source separation. *Neural Networks*, 10(8):1345–1351, 1997.
- [159] A. Ambainis. Application of Kolmogorov complexity to inductive inference with limited memory. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 313–318. Springer, 1995.
- [160] A. Ambainis. Optimization problem in inductive inference. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 96–107. Springer, 1995.
- [161] A. Ambainis. The power of procrastination in inductive inference: how it depends on used ordinal notations. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 99–111. Springer, 1995.
- [162] A. Ambainis. Probabilistic and team PFIN-type learning: general properties. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 157–168, New York, NY, 1996. ACM Press.
- [163] A. Ambainis. How rich is the structure of the intrinsic complexity of learning. *Information Processing Letters*, 75(3):109–112, 2000.
- [164] A. Ambainis. On learning formulas in the limit and with assurance. *Information Processing Letters*, 77(1):9–11, Jan. 2001.
- [165] A. Ambainis. Probabilistic inductive inference: A survey. *Theoret. Comput. Sci.*, 264(1):155–167, 2001.
- [166] A. Ambainis. Probabilistic and team PFIN-type learning: General properties. *J. Comput. Syst. Sci.*, 74(4):457–489, 2008.
- [167] A. Ambainis, K. Apsītis, C. Calude, R. Freivalds, M. Karpinski, T. Larfeldt, I. Sala, and J. Smotrovs. Effects of Kolmogorov complexity present in inductive inference as well. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 244–259. Springer, 1997.

- [168] A. Ambainis, K. Apsītis, R. Freivalds, W. Gasarch, and C. H. Smith. Team learning as a game. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 2–17. Springer, 1997.
- [169] A. Ambainis, K. Apsītis, R. Freivalds, and C. H. Smith. Hierarchies of probabilistic and team FIN-learning. *Theoret. Comput. Sci.*, 261(1):91–117, 2001. Special issue for ALT '97.
- [170] A. Ambainis, R. Desper, M. Farach, and S. Kannan. Nearly tight bounds on the learnability of evolution. In *Proc. 38th Annu. IEEE Symposium on Foundations of Computer Science*, pages 524–533. IEEE Computer Society Press, Los Alamitos, CA, 1997.
- [171] A. Ambainis and R. Freivalds. Transformations that preserve learnability. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 299–311. Springer, 1996.
- [172] A. Ambainis, R. Freivalds, and C. H. Smith. General inductive inference types based on linearly-ordered sets. In *Proc. 13th Annual Symposium on Theoretical Aspects of Computer Science: STACS'96*, volume 1046 of *Lecture Notes in Computer Science*, pages 243–253. Springer, 1996.
- [173] A. Ambainis, R. Freivalds, and C. H. Smith. Inductive inference with procrastination: Back to definitions. *Fundamenta Informaticae*, 40:1–16, 1999.
- [174] A. Ambainis, S. Jain, and A. Sharma. Ordinal mind change complexity of language identification. *Theoret. Comput. Sci.*, 220(2):323–343, 1999.
- [175] A. Ambainis and J. Smotrovs. Enumerable classes of total recursive functions: Complexity of inductive inference. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 10–25. Springer-Verlag, 1994.
- [176] K. Ambos-Spies. Learning finite variants of single languages from informant. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 163–173. Springer, 2016.
- [177] K. Ambos-Spies, S. Badaev, and S. Goncharov. Inductive inference and computable numberings. *Theoret. Comput. Sci.*, 412(18):1652–1668, 2011.
- [178] K. Ambos-Spies, S. Badaev, and S. Gontcharov. On a question of frank stephan. In M. Agrawal, D. Du, Z. Duan, and A. Li, editors, *Theory and Applications of Models of Computation, 5th International Conference, TAMC 2008, Xi'an, China, April 2008, Proceedings*, pages 423–432. Springer, 2008.
- [179] A. Ambroladze, E. Parrado-Hernández, and J. Shawe-Taylor. Complexity of pattern classes and the Lipschitz property. *Theoret. Comput. Sci.*, 382(3):232–246, 2007. Special issue for ALT 2004.
- [180] A. Ambroladze and J. Shawe-Taylor. When is small beautiful? In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 729–730. Springer, 2003.

- [181] A. Ambroladze and J. Shawe-Taylor. Complexity of pattern classes and lipschitz property. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 181–193. Springer, 2004.
- [182] J.-C. Amengual, A. Sanchis, E. Vidal, and J.-M. Benedi. Language simplification through error-correcting and grammatical inference techniques. *Machine Learning*, 44(1/2):143–159, 2001.
- [183] F. Ameur. A space-bounded learning algorithm for axis-parallel rectangles. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 313–321. Springer, 1995.
- [184] F. Ameur, P. Fischer, K. U. Höffgen, and F. M. auf der Heide. Trial and error: a new approach to space-bounded learning. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 133–144, Oxford, 1994. Oxford University Press.
- [185] K. Amin, M. Kearns, and U. Syed. Bandits, query learning, and the haystack dimension. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 87–106, Budapest, Hungary, 09–11 June 2011. PMLR.
- [186] K. Ammon. Some experiments with a learning procedure. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 87–98. Springer-Verlag, 1992.
- [187] T. R. Amoth, P. Cull, and P. Tadepalli. Exact learning of tree patterns from queries and counterexamples. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 175–186, New York, NY, 1998. ACM Press.
- [188] T. R. Amoth, P. Cull, and P. Tadepalli. Exact learning of unordered tree patterns from queries. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 323–332, New York, NY, 1999. ACM Press.
- [189] T. R. Amoth, P. Cull, and P. Tadepalli. On exact learning of unordered tree patterns. *Machine Learning*, 44(3):211–243, 2001.
- [190] J. Amsterdam. Extending the Valiant learning model. In *Proc. of the Fifth International Conference on Machine Learning*, pages 381–394, San Mateo, CA, 1988. Morgan Kaufmann.
- [191] J. Amsterdam. Some philosophical problems with formal learning theory. In *Proc. AAAI-88*, pages 580–584, Saint Paul, Minn., Aug. 1988. American Association for Artificial Intelligence.
- [192] J. Amsterdam. The valiant learning model: Extensions and assessment. Master’s thesis, MIT Department of Electrical Engineering and Computer Science, Jan. 1988.
- [193] A. Anandkumar, Y. Deng, R. Ge, and H. Mobahi. Homotopy analysis for tensor PCA. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 79–104, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [194] A. Anandkumar and R. Ge. Efficient approaches for escaping higher order saddle points in non-convex optimization. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 81–102, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [195] A. Anandkumar, R. Ge, D. Hsu, and S. Kakade. A tensor spectral approach to learning mixed membership community models. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 867–881. JMLR.org, 2013.
- [196] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky. Tensor decompositions for learning latent variable models (A survey for ALT). In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 19–38, Berlin, Heidelberg, New York, 2015. Springer.
- [197] A. Anandkumar, R. Ge, and M. Janzamin. Learning overcomplete latent variable models through tensor methods. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 36–112, Paris, France, 03–06 July 2015. PMLR.
- [198] A. Anandkumar, D. Hsu, and S. M. Kakade. A method of moments for mixture models and hidden markov models. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 33.1–33.34, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [199] N. Anari, S. O. Gharan, and A. Rezaei. Monte carlo markov chain algorithms for sampling strongly rayleigh distributions and determinantal point processes. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 103–115, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [200] O. Anava, E. Hazan, S. Mannor, and O. Shamir. Online learning for time series prediction. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 172–184. JMLR.org, 2013.
- [201] B. S. Anderson, A. W. Moore, and D. Cohn. A nonparametric approach to noisy and costly optimization. In *Proc. 17th International Conf. on Machine Learning*, pages 17–24. Morgan Kaufmann, San Francisco, CA, 2000.
- [202] C. W. Anderson, B. A. Draper, and D. A. Peterson. Behavioral cloning of student pilots with modular neural networks. In *Proc. 17th International Conf. on Machine Learning*, pages 25–32. Morgan Kaufmann, San Francisco, CA, 2000.
- [203] J. Anderson, M. Belkin, N. Goyal, L. Rademacher, and J. Voss. The more, the merrier: the blessing of dimensionality for learning large gaussian mixtures. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1135–1164, Barcelona, Spain, 13–15 June 2014. PMLR.
- [204] J. Anderson, N. Goyal, and L. Rademacher. Efficient learning of simplices. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1020–1045. JMLR.org, 2013.



- [205] J. A. Anderson and E. Rosenfeld, editors. *Neurocomputing: Foundations of Research*. MIT Press, Cambridge, Massachusetts, 1988.
- [206] J. R. Anderson. ACT-R and learning. In *Proc. 16th International Conf. on Machine Learning*, page 522. Morgan Kaufmann, San Francisco, CA, 1999.
- [207] J. R. Anderson and M. Matessa. Explorations of an incremental, Bayesian algorithm for categorization. *Machine Learning*, 9:275–308, 1992.
- [208] S. S. H. Ando and T. Fujita. Unsupervised visual learning of three-dimensional objects using a modular network architecture. *Neural Networks*, 12(7-8):1037–1051, 1999.
- [209] A. Andoni, D. Hsu, K. Shi, and X. Sun. Correspondence retrieval. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 105–126, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [210] P. M. Andreae. Justified generalization: Acquiring procedures from examples. Technical Report AI-TR-834, MIT Artificial Intelligence Laboratory, Jan. 1985.
- [211] L. Andrew, S. Barman, K. Ligett, M. Lin, A. Meyerson, A. Roytman, and A. Wierman. A tale of two metrics: Simultaneous bounds on competitiveness and regret. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 741–763. JMLR.org, 2013.
- [212] C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan. An introduction to MCMC for machine learning. *Machine Learning*, 50(1-2):5–43, 2003.
- [213] C. Anglano, A. Giordana, G. L. Bello, and L. Saitta. An experimental evaluation of co-evolutionary concept learning. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 19–27. Morgan Kaufmann, San Francisco, CA, 1998.
- [214] D. Angluin. *An Application of the Theory of Computational Complexity to the Study of Inductive Inference*. PhD thesis, Electrical Engineering and Computer Science Department, University of California, Berkeley, 1976.
- [215] D. Angluin. On the complexity of minimum inference of regular sets. *Inform. Control*, 39(3):337–350, 1978.
- [216] D. Angluin. Finding patterns common to a set of strings. *J. of Comput. Syst. Sci.*, 21(1):46–62, 1980.
- [217] D. Angluin. Inductive inference of formal languages from positive data. *Inform. Control*, 45(2):117–135, May 1980.
- [218] D. Angluin. Inference of reversible languages. *J. ACM*, 29(3):741–765, July 1982.
- [219] D. Angluin. A note on the number of queries needed to identify regular languages. *Inform. Control*, 51:76–87, 1982.
- [220] D. Angluin. Types of queries for concept learning. Technical Report YALEU/DCS/TR-479, Yale University, Department of Computer Science, June 1986.
- [221] D. Angluin. Debugging a DAG efficiently. Technical Report 591, Yale University, 1987.

- [222] D. Angluin. Learning  $k$ -bounded context-free grammars. Technical Report YALEU/DCS/RR-557, Department of Computer Science, Yale University, Aug. 1987.
- [223] D. Angluin. Learning  $k$ -term DNF formulas using queries and counterexamples. Technical Report YALEU/DCS/RR-559, Department of Computer Science, Yale University, Aug. 1987.
- [224] D. Angluin. Learning propositional Horn sentences with hints. Technical Report YALEU/DCS/RR-590, Department of Computer Science, Yale University, Dec. 1987.
- [225] D. Angluin. Learning regular sets from queries and counterexamples. *Inform. Comput.*, 75(2):87–106, Nov. 1987.
- [226] D. Angluin. Equivalence queries and DNF formulas. Technical Report YALEU/DCS/RR-659, Yale University Department of Computer Science, Nov. 1988.
- [227] D. Angluin. Identifying languages from stochastic examples. Technical Report YALEU/DCS/RR-614, Yale University, Dept. of Computer Science, New Haven, CT, 1988.
- [228] D. Angluin. Learning with hints. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 167–181, San Mateo, CA, 1988. Morgan Kaufmann.
- [229] D. Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.
- [230] D. Angluin. Requests for hints that return no hints. Technical report, Yale University, YALE/DCS/RR-647, 1988.
- [231] D. Angluin. Equivalence queries and approximate fingerprints. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 134–145, San Mateo, CA, 1989. Morgan Kaufmann.
- [232] D. Angluin. Using queries to identify  $\mu$ -formulas. Technical report, Yale University, YALE/DCS/RR-694, 1989.
- [233] D. Angluin. Negative results for equivalence queries. *Machine Learning*, 5:121–150, 1990.
- [234] D. Angluin. Computational learning theory: survey and selected bibliography. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 351–369, New York, NY, 1992. ACM Press.
- [235] D. Angluin. Exact learning of  $\mu$ -DNF formulas with malicious membership queries. Technical Report YALEU/DCS/TR-1020, Yale University, Mar. 1994.
- [236] D. Angluin. Queries revisited. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 12–31. Springer, 2001.
- [237] D. Angluin. Learning and mathematics. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 1–2. Springer, 2004.
- [238] D. Angluin. Queries revisited. *Theoret. Comput. Sci.*, 313(2):175–194, 2004. Special issue for ALT 2001.
- [239] D. Angluin, J. Aspnes, J. Chen, D. Eisenstat, and L. Reyzin. Learning acyclic probabilistic circuits using test paths. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9–12, 2008*, pages 169–180. Omnipress, 2008.

- [240] D. Angluin, J. Aspnes, J. Chen, and L. Reyzin. Learning large-alphabet and analog circuits with value injection queries. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 51–65, Berlin, 2007. Springer.
- [241] D. Angluin, J. Aspnes, J. Chen, and Y. Wu. Learning a circuit by injecting values. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 584–593. ACM Press, 2006.
- [242] D. Angluin, J. Aspnes, and A. Kontorovich. On the learnability of shuffle ideals. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 111–123, Berlin, Heidelberg, New York, 2012. Springer.
- [243] D. Angluin, J. Aspnes, and L. Reyzin. Optimally learning social networks with activations and suppressions. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 272–286, Berlin, Oct. 2008. Springer.
- [244] D. Angluin, J. Aspnes, and L. Reyzin. Inferring social networks from outbreaks. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 104–118, Berlin, Heidelberg, New York, 2010. Springer.
- [245] D. Angluin, J. Aspnes, and L. Reyzin. Optimally learning social networks with activations and suppressions. *Theoret. Comput. Sci.*, 411(29-30):2729–2740, 2010. Special issue for ALT 2008.
- [246] D. Angluin and L. Becerra-Bonache. Learning meaning before syntax. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 1–14. Springer, 2008.
- [247] D. Angluin, L. Becerra-Bonache, A. H. Dediu, and L. Reyzin. Learning finite automata using label queries. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 171–185, Berlin/Heidelberg, 2009. Springer.
- [248] D. Angluin, L. Birnbaum, J. Feldman, R. Rivest, and L. Valiant. Summary of the panel discussion. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 398–401, San Mateo, CA, 1988. Morgan Kaufmann.
- [249] D. Angluin and D. Chen. Learning a random DFA from uniform strings and state information. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 119–133, Berlin, Heidelberg, New York, 2015. Springer.
- [250] D. Angluin and J. Chen. Learning a hidden graph using  $O(\log n)$  queries per edge. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 210–223. Springer, 2004.

- [251] D. Angluin and J. Chen. Learning a hidden hypergraph. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 561–575. Springer, 2005.
- [252] D. Angluin and J. Chen. Learning a hidden graph using  $O(\log n)$  queries per edge. *J. Comput. Syst. Sci.*, 74(4):546–556, 2008.
- [253] D. Angluin and M. Cs ur os. Learning Markov chains with variable length memory from noisy output. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 298–308, New York, NY, 1997. ACM Press.
- [254] D. Angluin and T. Dohrn. The power of random counterexamples. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 452–465, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [255] D. Angluin, D. Eisenstat, L. A. Kontorovich, and L. Reyzin. Lower bounds on learning random structures with statistical queries. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 194–208, Berlin, Heidelberg, New York, 2010. Springer.
- [256] D. Angluin and D. Fisman. Learning regular omega languages. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 125–139, Berlin, Heidelberg, New York, 2014. Springer.
- [257] D. Angluin and D. Fisman. Learning regular omega languages. *Theoret. Comput. Sci.*, 650:57–72, 2016. Special Issue: Algorithmic Learning Theory.
- [258] D. Angluin, M. Frazier, and L. Pitt. Learning conjunctions of Horn clauses. *Machine Learning*, 9:147–164, 1992.
- [259] D. Angluin, W. I. Gasarch, and C. H. Smith. Training sequences. *Theoret. Comput. Sci.*, 66(3):255–272, 1989.
- [260] D. Angluin, L. Hellerstein, and M. Karpinski. Learning read-once formulas with queries. *J. ACM*, 40(1):185–210, 1993.
- [261] D. Angluin and M. Kharitonov. When won’t membership queries help? *J. of Comput. Syst. Sci.*, 50(2):336–355, 1995. Earlier version appeared in 23rd STOC, 1991.
- [262] D. Angluin and M. Kri kis. Learning with malicious membership queries and exceptions. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 57–66, New York, NY, 1994. ACM Press.
- [263] D. Angluin and M. Kri kis. Teachers, learners and black boxes. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 285–297, New York, NY, 1997. ACM Press.
- [264] D. Angluin and M. Kri kis. Learning from different teachers. *Machine Learning*, 51(2):137–163, 2003.
- [265] D. Angluin, M. Kri kis, R. H. Sloan, and G. Tur an. Malicious omissions and errors in answers to membership queries. *Machine Learning*, 28(2-3):211–255, 1997.

- [266] D. Angluin and P. Laird. Learning from noisy examples. *Machine Learning*, 2(4):343–370, 1988.
- [267] D. Angluin and D. K. Slonim. Learning monotone DNF with an incomplete membership oracle. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 139–146, San Mateo, CA, 1991. Morgan Kaufmann.
- [268] D. Angluin and D. K. Slonim. Randomly fallible teachers: Learning monotone DNF with an incomplete membership oracle. *Machine Learning*, 14:7–26, 1994.
- [269] D. Angluin and C. H. Smith. A survey of inductive inference: Theory and methods. *ACM Comput. Surv.*, 15(3):237–269, Sept. 1983.
- [270] D. Angluin and C. H. Smith. Inductive inference. In *Encyclopedia of Artificial Intelligence*, pages 409–418. J. Wiley and Sons, New York, 1987.
- [271] D. Angluin and L. G. Valiant. Fast probabilistic algorithms for hamiltonian circuits and matchings. *J. of Comput. Syst. Sci.*, 18(2):155–193, 1979.
- [272] D. Angluin, J. Westbrook, and W. Zhu. Robot navigation with distance queries. *SIAM J. Comput.*, 30(1):110–144, 2000.
- [273] A. C. C. Angulo.  $K$ -SVRC. A multi-class support vector machine. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 31–38. Springer, 2000.
- [274] S. Anoulova, P. Fischer, S. Pölt, and H. U. Simon. PAB-decisions for Boolean and real-valued features. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 353–362, New York, NY, 1992. ACM Press.
- [275] S. Anoulova, P. Fischer, S. Pölt, and H. U. Simon. Probably almost Bayes decisions. *Inform. Comput.*, 129(1):63–71, Aug. 1996.
- [276] S. Anoulova and S. Pölt. Using Kullback-Leibler divergence in learning theory. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 145–156, Oxford, 1994. Oxford University Press.
- [277] M. Anthony. Generalization error bounds for threshold decision lists. *Journal of Machine Learning Research*, 5:189–217, 2004.
- [278] M. Anthony and P. Bartlett. Function learning from interpolation. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 211–221. Springer, 1995.
- [279] M. Anthony and N. Biggs. *Computational Learning Theory*. Cambridge Tracts in Theoretical Computer Science (30). Cambridge University Press, Cambridge, 1992.
- [280] M. Anthony, N. Biggs, and J. Shawe-Taylor. The learnability of formal concepts. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 246–257, San Mateo, CA, 1990. Morgan Kaufmann.
- [281] M. Anthony, G. Brightwell, D. Cohen, and J. Shawe-Taylor. On exact specification by examples. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 311–318, New York, NY, 1992. ACM Press.

- [282] M. Anthony, G. Brightwell, and J. Shawe-Taylor. On specifying Boolean functions by labelled examples. *Discrete Applied Mathematics*, 61(1):1–25, 1995.
- [283] M. Anthony and S. Holden. On the power of polynomial discriminators and radial basis function networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 158–164, New York, NY, 1993. ACM Press.
- [284] M. Anthony and S. B. Holden. Cross-validation for binary classification by real-valued functions: theoretical analysis. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 218–229, New York, NY, 1998. ACM Press.
- [285] M. Anthony and J. Shawe-Taylor. A result of Vapnik with applications. Technical Report CSD-TR-628, University of London, Surrey, England, 1990.
- [286] M. Anthony and J. Shawe-Taylor. Valid generalisation of functions from close approximations on a sample. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 94–108, Oxford, 1994. Oxford University Press.
- [287] S. Anthony and A. M. Frisch. Cautious induction: An alternative to clause-at-a-time hypothesis construction in inductive logic programming. *New Generation Computing*, 17(1):25–52, 1999.
- [288] A. Antos. Lower bounds on the rate of convergence of nonparametric pattern recognition. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 241–252. Springer, 1999.
- [289] A. Antos. Lower bounds for the rate of convergence in nonparametric pattern recognition. *Theoret. Comput. Sci.*, 284(1):3–24, 2002.
- [290] A. Antos. Improved minimax bounds on the test and training distortion of empirically designed vector quantizers. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 531–544. Springer, 2005.
- [291] A. Antos, G. Bartók, and C. Szepesvári. Toward a classification of finite partial-monitoring games. *Theoret. Comput. Sci.*, 473:77–99, 2013. Special Issue on Learning Theory.
- [292] A. Antos, V. Grover, and C. Szepesvári. Active learning in multi-armed bandits. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 287–302, Berlin, Oct. 2008. Springer.
- [293] A. Antos, V. Grover, and C. Szepesvári. Active learning in heteroscedastic noise. *Theoret. Comput. Sci.*, 411(29-30):2712–2728, 2010. Special issue for ALT 2008.
- [294] A. Antos, B. Kégl, T. Linder, and G. Lugosi. Data-dependent margin-based generalization bounds for classification. *Journal of Machine Learning Research*, 3:73–98, 2002.
- [295] A. Antos and G. Lugosi. Strong minimax lower bounds for learning. *Machine Learning*, 30(1):31–56, 1998. Earlier version in COLT96.
- [296] A. Antos, C. Szepesvári, and R. Munos. Learning near-optimal policies with Bellman-residual minimization based fitted policy iteration and a single sample path. In *Learning*

*Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 574–588, Berlin, 2006. Springer.

- [297] C. M. Antunes and A. L. Oliveira. Inference of sequential association rules guided by context-free grammars. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 1–1. Springer, 2002.
- [298] S. Anulova, J. R. Cuellar, K.-U. Höffgen, and H.-U. Simon. Bayes decisions in a neural network-PAC setting. In *Computational Learning Theory and Natural Learning Systems II: Intersections Between Theory and Experiment*, chapter 1, pages 3–14. MIT Press, 1994.
- [299] B. Apolloni and S. Chiaravalli. PAC learning of concept classes through the boundaries of their items. *Theoret. Comput. Sci.*, 172(1-2):91–120, 1997.
- [300] B. Apolloni and C. Gentile. P-sufficient statistics for PAC learning  $k$ -term-DNF formulas through enumeration. *Theoret. Comput. Sci.*, 230(1-2):1–37, 2000.
- [301] B. Apolloni and D. Malchiodi. Gaining degrees of freedom in subsymbolic learning. *Theoret. Comput. Sci.*, 255(1-2):295–321, 2001.
- [302] B. Apolloni, D. Malchiodi, and S. Gaito. *Algorithmic Inference in Machine Learning*, volume 5 of *International Series on Advanced Intelligence*. Advanced Knowledge International Pty Ltd, Magill, Adelaide, Australia, 2003.
- [303] B. Apolloni, G. Zamponi, and A. M. Zanaboni. Learning fuzzy decision trees. *Neural Networks*, 11(5):885–895, 1998.
- [304] K. Apsītis. Derived sets and inductive inference. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 26–39. Springer-Verlag, 1994.
- [305] K. Apsītis. Topological considerations in composing teams of learning machines. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 146–154. Springer, 1995.
- [306] K. Apsītis, S. Arikawa, R. Freivalds, E. Hirowatari, and C. H. Smith. On the inductive inference of recursive real-valued functions. *Theoret. Comput. Sci.*, 219(1/2):3–17, 1999.
- [307] K. Apsītis, R. Freivalds, M. Kriķis, R. Simanovskis, and J. Smotrovs. Unions of identifiable classes of total recursive functions. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 99–107. Springer-Verlag, 1992.
- [308] K. Apsītis, R. Freivalds, R. Simanovskis, and J. Smotrovs. Unions of identifiable families of languages. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 48–58. Springer, 1996.
- [309] K. Apsītis, R. Freivalds, R. Simanovskis, and J. Smotrovs. Closedness properties in EX-identification of recursive functions. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 46–60. Springer, 1998.

- [310] K. Apsītis, R. Freivalds, R. Simanovskis, and J. Smotrovs. Closedness properties in identification. *Theoret. Comput. Sci.*, 268(2):367–393, 2001. Special issue for ALT '98.
- [311] K. Apsītis, R. Freivalds, and C. Smith. Choosing a learning team: a topological approach. In *Proc. of Twenty-sixth ACM Symposium on Theory of Computing*, 1994.
- [312] K. Apsītis, R. Freivalds, and C. H. Smith. On the inductive inference of real valued functions. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 170–177, New York, NY, 1995. ACM Press.
- [313] K. Apsītis, R. Freivalds, and C. H. Smith. On duality in learning and the selection of learning teams. *Inform. Comput.*, 129(1):53–62, Aug. 1996.
- [314] K. Apsītis, R. Freivalds, and C. H. Smith. Asymmetric team learning. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 90–95, New York, NY, 1997. ACM Press.
- [315] S. Ar, R. J. Lipton, R. Rubinfeld, and M. Sudan. Reconstructing algebraic functions from mixed data. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 503–512. IEEE Computer Society Press, Los Alamitos, CA, 1992.
- [316] J. Aracena and E. Goles. Complexity of perceptron recognition for a class of geometric patterns. *Theoret. Comput. Sci.*, 299(1-3):65–79, 2003.
- [317] A. Arapostathis and S. I. Marcus. Analysis of an identification algorithm arising in the adaptive estimation of Markov chains. *Math. Control Signals Systems*, 3:1–29, 1990.
- [318] J.-L. Arcos and R. L. de Mántaras. A case based approach to expressivity-aware tempo transformation. *Machine Learning*, 65(2-3):411–437, 2006.
- [319] S. Argamon-Engelson, M. Koppel, and H. Walters. Maximizing theory accuracy through selective reinterpretation. *Machine Learning*, 41(2):123–152, 2000.
- [320] A. Argyriou, C. A. Micchelli, and M. Pontil. Learning convex combinations of continuously parameterized basic kernels. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 338–352. Springer, 2005.
- [321] M. Arias and J. L. Balcázar. Query learning and certificates in lattices. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 303–315, Berlin, Oct. 2008. Springer.
- [322] M. Arias and J. L. Balcázar. Canonical Horn representations and query learning. In R. Gavalda, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 156–170, Berlin/Heidelberg, 2009. Springer.
- [323] M. Arias and R. Khardon. Learning closed horn expressions. *Inform. Comput.*, 178(1):214–240, 2002.
- [324] M. Arias and R. Khardon. The subsumption lattice and query learning. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 410–424. Springer, 2004.



- [325] M. Arias, R. Khardon, and R. A. Servedio. Polynomial certificates for propositional classes. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 537–551. Springer, 2003.
- [326] S. Arikawa. The discovery science project in japan. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 9–11. Springer, 2001.
- [327] S. Arikawa, K. Furukawa, S. Morishita, and H. Motoda. Preface. *Theoret. Comput. Sci.*, 292(2):343–344, 2003.
- [328] S. Arikawa, S. Goto, S. Ohsuga, and T. Yokomori, editors. *Proceedings of the First International Workshop on Algorithmic Learning Theory*, Tokyo, 1990. Japanese Society for Artificial Intelligence, Ohmsha, LTD.
- [329] S. Arikawa and K. P. Jantke, editors. *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Oct. 1994.
- [330] S. Arikawa, S. Kuhara, S. Miyano, Y. Mukouchi, A. Shinohara, and T. Shinohara. A machine discovery from amino acid sequences by decision trees over regular patterns. In *Proceedings of the International Conference on Fifth Generation Computer Systems. Volume 2*, pages 618–625, Burke, VA, USA, 1992. IOS Press.
- [331] S. Arikawa, A. Maruoka, and T. Sato, editors. *Proceedings of the Second International Workshop on Algorithmic Learning Theory*, Tokyo, 1991. Japanese Society for Artificial Intelligence, Ohmsha, LTD.
- [332] S. Arikawa, S. Miyano, A. Shinohara, T. Shinohara, and A. Yamamoto. Algorithmic learning theory with elementary formal systems. *IEICE Trans. Inf. & Syst.*, E75-D(4):405–414, 1992.
- [333] S. Arikawa and S. Morishita, editors. *Discovery Science, Third International Conference, DS 2000, Kyoto, Japan, December 2000, Proceedings*, volume 1967 of *Lecture Notes in Artificial Intelligence*, Berlin, 2000. Springer.
- [334] S. Arikawa and M. M. Richter. Preface. *New Generation Computing*, 15(1):1–2, 1997. Special issue for AII '94 and ALT '94.
- [335] S. Arikawa and A. K. Sharma, editors. *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 1996. Springer-Verlag.
- [336] S. Arikawa, T. Shinohara, S. Miyano, and A. Shinohara. More about learning elementary formal systems. In *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardtsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 107–117. Springer-Verlag, 1993.
- [337] S. Arikawa, T. Shinohara, and A. Yamamoto. Elementary formal system as a unifying framework for language learning. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 312–327, San Mateo, CA, 1989. Morgan Kaufmann.
- [338] S. Arikawa, T. Shinohara, and A. Yamamoto. Learning elementary formal systems. *Theoret. Comput. Sci.*, 95(1):97–113, 1992.

- [339] J. Arima. A logical analysis of relevance in analogy. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 255–265. IOS Press, 1992.
- [340] J. Arima and H. Sawamura. Reformulation of explanation by linear logic — Toward logic for explanation —. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 45–57. Springer-Verlag, 1993.
- [341] H. Arimura. Learning acyclic first-order Horn sentences from entailment. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 432–445. Springer, 1997.
- [342] H. Arimura and K. Hirata. Learning of first-order formulas and inductive logic programming. *Journal of Japanese Society for Artificial Intelligence*, 14(5):790–799, 1999.
- [343] H. Arimura, H. Ishizaka, and T. Shinohara. Polynomial time inference of a subclass of context-free transformations. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 136–143, New York, NY, 1992. ACM Press.
- [344] H. Arimura, H. Ishizaka, and T. Shinohara. Learning unions of tree patterns using queries. *Theoret. Comput. Sci.*, 185(1):47–62, 1997. Special issue for ALT '95.
- [345] H. Arimura and S. Jain. Preface. *Theoret. Comput. Sci.*, 348(1):1–2, 2005. Special issue for ALT 2000.
- [346] H. Arimura, S. Jain, and A. Sharma, editors. *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Dec. 2000. Springer.
- [347] H. Arimura, H. Sakamoto, and S. Arikawa. Efficient data mining from large text databases. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 123–139. Springer, 2001.
- [348] H. Arimura, H. Sakamoto, and S. Arikawa. Efficient learning of semi-structured data from queries. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 315–331. Springer, 2001.
- [349] H. Arimura and T. Shinohara. Inductive inference of Prolog programs with linear dependency from positive data. In *Information Modelling and Knowledge Bases V*, pages 365–375. IOS Press, 1994.
- [350] H. Arimura, T. Shinohara, and S. Otsuki. Polynomial time inference of unions of tree pattern languages. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 105–114. IOS Press, 1992.
- [351] H. Arimura, T. Shinohara, and S. Otsuki. A polynomial time algorithm for finding finite unions of tree pattern languages. In *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardtsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 118–131. Springer-Verlag, 1993.

- [352] H. Arimura, T. Shinohara, and S. Otsuki. Finding minimal generalizations for unions of pattern languages and its application to inductive inference from positive data. In *11th Annual Symposium on Theoretical Aspects of Computer Science, Proceedings*, volume 775 of *Lecture Notes in Computer Science*, pages 649–660. Springer-Verlag, 1994.
- [353] H. Arimura and T. Uno. Mining maximal flexible patterns in a sequence. In *JSAI 2007 Conference and Workshops, Miyazaki, Japan, June 18-22, 2007, Revised Selected Papers*, volume 4914 of *Lecture Notes in Artificial Intelligence*, pages 307–317. Springer, Berlin, 2008.
- [354] H. Arimura, A. Wataki, R. Fujino, and S. Arikawa. A fast algorithm for discovering optimal string patterns in large text databases. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 247–261. Springer, 1998.
- [355] H. Arimura and A. Yamamoto. Inductive logic programming: From logic of discovery to machine learning. *IEICE Transactions on Information and Systems*, E83-D(1):10–18, 2000.
- [356] S. Arlot, G. Blanchard, and É. Roquain. Resampling-based confidence regions and multiple tests for a correlated random vector. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 127–141, Berlin, 2007. Springer.
- [357] E. Armengol and E. Plaza. Bottom-up induction of feature terms. *Machine Learning*, 41(3):259–294, 2000.
- [358] E. Armengol and E. Plaza. Lazy induction of descriptions for relational case-based learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 13–24. Springer, 2001.
- [359] T. Armstrong and T. Oates. Which came first, the grammar or the lexicon? In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 283–285. Springer, 2008.
- [360] O. Arnold and K. P. Jantke. Therapy plan generation as program synthesis. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 40–55. Springer-Verlag, 1994.
- [361] O. Arnold and K. P. Jantke. Inductive program synthesis for therapy plan generation. *New Generation Computing*, 15(1):27–58, 1997. Special issue for AII '94 and ALT '94.
- [362] S. Arora, R. Ge, T. Ma, and A. Moitra. Simple, efficient, and neural algorithms for sparse coding. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 113–149, Paris, France, 03–06 July 2015. PMLR.
- [363] S. Arora, R. Ge, and A. Moitra. New algorithms for learning incoherent and overcomplete dictionaries. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 779–806, Barcelona, Spain, 13–15 June 2014. PMLR.

- [364] S. Arora and R. Kannan. Learning mixtures of arbitrary gaussians. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 247–257. ACM Press, 2001.
- [365] J. Arpe and R. Reischuk. Robust inference of relevant attributes. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 99–113. Springer, 2003.
- [366] J. Arpe and R. Reischuk. Learning juntas in the presence of noise. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 387–398, Berlin, 2006. Springer.
- [367] J. Arpe and R. Reischuk. Learning juntas in the presence of noise. *Theoret. Comput. Sci.*, 384(1):2–21, 2007. Special issue for TAMC 2006.
- [368] D. Arrivault, D. Benielli, F. Denis, and R. Eyraud. Sp2Learn: A toolbox for the spectral learning of weighted automata. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 105–119, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [369] V. Arvind, J. Köbler, and W. Lindner. Parameterized learnability of  $k$ -juntas and related problems. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 120–134, Berlin, Oct. 2007. Springer.
- [370] V. Arvind and N. V. Vinodchandran. The complexity of exactly learning algebraic concepts. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 100–112. Springer, 1996.
- [371] V. Arvind and N. V. Vinodchandran. Exact learning via teaching assistants. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 291–306. Springer, 1997.
- [372] V. Arvind and N. V. Vinodchandran. Exact learning via teaching assistants. *Theoret. Comput. Sci.*, 241(1-2):51–81, 2000. Special issue for ALT '96.
- [373] M. Asada, S. Noda, S. Tawaratsumida, and K. Hosoda. Purposive behavior acquisition for a real robot by vision-based reinforcement learning. *Machine Learning*, 23:279–303, 1996.
- [374] L. Asker. Improving accuracy of incorrect domain theories. In *Proc. 11th International Conference on Machine Learning*, pages 19–27. Morgan Kaufmann, 1994.
- [375] L. Asker and R. Maclin. Feature engineering and classifier selection: A case study in Venusian volcano detection. In *Proc. 14th International Conference on Machine Learning*, pages 3–11. Morgan Kaufmann, 1997.
- [376] J. A. Aslam. Inferring graphs from walks. Master’s thesis, Massachusetts Institute of Technology, Jan. 1992.
- [377] J. A. Aslam. Improving algorithms for boosting. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 200–207. Morgan Kaufmann, San Francisco, 2000.

- [378] J. A. Aslam and S. E. Decatur. General bounds on statistical query learning and PAC learning with noise via hypothesis boosting. In *Proc. of the 34th Annual Symposium on Foundations of Computer Science*, pages 282–291. IEEE Computer Society Press, Los Alamitos, CA, 1993.
- [379] J. A. Aslam and S. E. Decatur. Specification and simulation of statistical query algorithms for efficiency and noise tolerance. *J. of Comput. Syst. Sci.*, 56(2):191–208, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [380] J. A. Aslam and A. Dhagat. Searching in the presence of linearly bounded errors. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 486–493, New York, NY, 1991. ACM Press.
- [381] J. A. Aslam and R. L. Rivest. Inferring graphs from walks. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 359–370, San Mateo, CA, 1990. Morgan Kaufmann.
- [382] P. Assouad. Densité et dimension. *Ann. Inst. Fourier*, pages 233–282, 1983.
- [383] A. Atıcı and R. A. Servedio. Learning unions of  $\omega(1)$ -dimensional rectangles. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 32–47. Springer, Oct. 2006.
- [384] A. Atıcı and R. A. Servedio. Learning unions of  $\omega(1)$ -dimensional rectangles. *Theoret. Comput. Sci.*, 405(3):209–222, 2008. Special issue for ALT 2006.
- [385] C. G. Atkeson. Learning arm kinematics and dynamics. In *Annual Review of Neuroscience*, volume 12, pages 157–183. Annual Reviews, 1989.
- [386] C. G. Atkeson and S. Schaal. Robot learning from demonstration. In *Proc. 14th International Conference on Machine Learning*, pages 12–20. Morgan Kaufmann, 1997.
- [387] J.-Y. Audibert. A randomized online learning algorithm for better variance control. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 392–407, Berlin, 2006. Springer.
- [388] J.-Y. Audibert, S. Bubeck, and G. Lugosi. Minimax policies for combinatorial prediction games. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 107–132, Budapest, Hungary, 09–11 June 2011. PMLR.
- [389] J.-Y. Audibert, R. Munos, and C. Szepesvári. Tuning bandit algorithms in stochastic environments. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 150–165, Berlin, Oct. 2007. Springer.
- [390] J.-Y. Audibert, R. Munos, and C. Szepesvári. Exploration-exploitation tradeoff using variance estimates in multi-armed bandits. *Theoret. Comput. Sci.*, 410(9):1876–1902, 2009. Special issue for ALT 2007.
- [391] P. Auer. On-line learning of rectangles in noisy environments. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 253–261, New York, NY, 1993. ACM Press.

- [392] P. Auer. Learning nested differences in the presence of malicious noise. *Theoret. Comput. Sci.*, 185(1):159–175, 1997. Special issue for ALT '95.
- [393] P. Auer. On learning from multi-instance examples: Empirical evaluation of a theoretical approach. In *Proc. 14th International Conference on Machine Learning*, pages 21–29. Morgan Kaufmann, 1997.
- [394] P. Auer. An improved on-line algorithm for learning linear evaluation functions. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 118–125. Morgan Kaufmann, San Francisco, 2000.
- [395] P. Auer. Using upper confidence bounds for online learning. In *Proc. 41st Annual Symposium on Foundations of Computer Science*, pages 270–279. IEEE Computer Society Press, Los Alamitos, CA, 2000.
- [396] P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3:397–422, 2002.
- [397] P. Auer and N. Cesa-Bianchi. On-line learning with malicious noise and the closure algorithm. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):83–99, 1998. Special issue for AII '94 and ALT '94.
- [398] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2/3):235–256, 2002.
- [399] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. Gambling in a rigged casino: the adversarial multi-armed bandit problem. In *Proc. of the 36th Annual Symposium on Foundations of Computer Science*, pages 322–331. IEEE Computer Society Press, Los Alamitos, CA, 1995.
- [400] P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *J. Comput. Syst. Sci.*, 64(1):48–75, 2002. Special issue for COLT 2000.
- [401] P. Auer and C.-K. Chiang. An algorithm with nearly optimal pseudo-regret for both stochastic and adversarial bandits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 116–120, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [402] P. Auer, A. Clark, and T. Zeugmann. Guest Editors' foreword. *Theoret. Comput. Sci.*, 650:1–3, 2016. Special Issue: Algorithmic Learning Theory.
- [403] P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors. *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, Berlin, Heidelberg, New York, Oct. 2014. Springer.
- [404] P. Auer, A. Clark, T. Zeugmann, and S. Zilles. Editors' introduction. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 1–7, Berlin, Heidelberg, New York, 2014. Springer.
- [405] P. Auer and C. Gentile. Adaptive and self-confident on-line learning algorithms. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 107–117. Morgan Kaufmann, San Francisco, 2000.

- [406] P. Auer, M. Herbster, and M. K. Warmuth. Exponentially many local minima for single neurons. In *Advances in Neural Information Processing Systems*, volume 8, pages 316–322. The MIT Press, 1996.
- [407] P. Auer, R. C. Holte, and W. Maass. Theory and applications of agnostic PAC-learning with small decision trees. In *Proc. 12th International Conference on Machine Learning*, pages 21–29. Morgan Kaufmann, 1995.
- [408] P. Auer, S. Kwek, W. Maass, and M. K. Warmuth. Learning of depth two neural networks with constant fan-in at the hidden nodes. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 333–343, New York, NY, 1996. ACM Press.
- [409] P. Auer, S. H. Lim, and C. Watkins. Models for autonomously motivated exploration in reinforcement learning (extended abstract). In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 14–17, Berlin, Heidelberg, New York, 2011. Springer.
- [410] P. Auer and P. Long. Simulating access to hidden information while learning. In *Proc. of Twenty-sixth ACM Symposium on Theory of Computing*, pages 263–272, 1994.
- [411] P. Auer and P. M. Long. Structural results about on-line learning models with and without queries. *Machine Learning*, 36(3):147–181, 1999.
- [412] P. Auer, P. M. Long, W. Maass, and G. J. Woeginger. On the complexity of function learning. *Machine Learning*, 18(2/3):187–230, 1995.
- [413] P. Auer, P. M. Long, and A. Srinivasan. Approximating hyper-rectangles: Learning and pseudorandom sets. *J. Comput. Syst. Sci.*, 57(3):376–388, 1998.
- [414] P. Auer and R. Meir, editors. *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, Berlin, 2005. Springer.
- [415] P. Auer and R. Ortner. A new PAC bound for intersection-closed concept classes. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 408–414. Springer, 2004.
- [416] P. Auer and R. Ortner. A new PAC bound for intersection-closed concept classes. *Machine Learning*, 66(2-3):151–163, 2007.
- [417] P. Auer, R. Ortner, and C. Szepesvári. Improved rates for the stochastic continuum-armed bandit problem. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 454–468, Berlin, 2007. Springer.
- [418] P. Auer and M. K. Warmuth. Tracking the best disjunction. *Machine Learning*, 32(2):127–150, 1998. Earlier version in 36th FOCS, 1995.
- [419] Authorless. Errata to extending. *Machine Learning*, 5:229–229, 1990. This is a correction to the article *Extending domain theories: two case studies in student Modeling*, by D. Slemman, et al., which appeared in volume 5, number 1, of *Machine Learning*.

- [420] Authorless. Erratum to discovery. *Machine Learning*, 13:153–153, 1993. This is a correction to the article *Discovery by minimal length encoding: A case study in molecular evolution*, by Aleksandar Milosavljevic and Jerzy Jurka, which appeared in volume 12, number 1,2,3 of *Machine Learning*.
- [421] P. Awasthi, M.-F. Balcan, N. Haghtalab, and R. Uner. Efficient learning of linear separators under bounded noise. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 167–190, Paris, France, 03–06 July 2015. PMLR.
- [422] P. Awasthi, M.-F. Balcan, N. Haghtalab, and H. Zhang. Learning and 1-bit compressed sensing under asymmetric noise. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 152–192, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [423] P. Awasthi, A. Blum, N. Haghtalab, and Y. Mansour. Efficient PAC learning from the crowd. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 127–150, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [424] P. Awasthi, M. Charikar, K. A. Lai, and A. Risteski. Label optimal regret bounds for online local learning. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 150–166, Paris, France, 03–06 July 2015. PMLR.
- [425] P. Awasthi, V. Feldman, and V. Kanade. Learning using local membership queries. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 398–431. JMLR.org, 2013.
- [426] B. Awerbuch, M. Betke, R. L. Rivest, and M. Singh. Piecemeal graph exploration by a mobile robot. *Inform. Comput.*, 152(2):155–172, 1999.
- [427] B. Awerbuch and R. Kleinberg. Online linear optimization and adaptive routing. *J. Comput. Syst. Sci.*, 74(1):97–114, 2008.
- [428] B. Awerbuch and R. D. Kleinberg. Adaptive routing with end-to-end feedback: distributed learning and geometric approaches. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 45–53. ACM Press, 2004.
- [429] B. Awerbuch and R. D. Kleinberg. Competitive collaborative learning. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 233–248. Springer, 2005.
- [430] B. Awerbuch and S. G. Kobourov. Polylogarithmic-overhead piecemeal graph exploration. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 280–286, New York, NY, 1998. ACM Press.
- [431] M. Aziz, J. Anderton, E. Kaufmann, and J. Aslam. Pure exploration in infinitely-armed bandit models with fixed-confidence. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 3–24. PMLR, 07–09 Apr. 2018.



- [432] K. Azizzadenesheli, A. Lazaric, and A. Anandkumar. Open problem: Approximate planning of POMDPs in the class of memoryless policies. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1639–1642, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [433] K. Azizzadenesheli, A. Lazaric, and A. Anandkumar. Reinforcement learning of POMDPs using spectral methods. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 193–256, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [434] K. S. Azoury and M. K. Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine Learning*, 43(3):211–246, 2001.
- [435] A. Azran and R. Meir. Data dependent risk bounds for hierarchical mixture of experts classifiers. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 427–441. Springer, 2004.
- [436] F. Bach. Sharp analysis of low-rank kernel matrix approximations. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 185–209. JMLR.org, 2013.
- [437] F. Bach and V. Perchet. Highly-smooth zero-th order online optimization. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 257–283, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [438] F. R. Bach and M. I. Jordan. Kernel independent component analysis. *Journal of Machine Learning Research*, 3:1–48, 2002.
- [439] F. R. Bach and M. I. Jordan. Beyond independent components: Trees and clusters. *Journal of Machine Learning Research*, 4:1205–1233, 2003.
- [440] J. R. Bachrach. *Connectionist Modeling and Control of Finite-State Environments*. PhD thesis, University of Massachusetts, Feb. 1992.
- [441] R. Bachrach, S. Fine, and E. Shamir. Query by committee, linear separation and random walks. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 34–49. Springer, 1999.
- [442] A. Badanidiyuru, J. Langford, and A. Slivkins. Resourceful contextual bandits. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1109–1134, Barcelona, Spain, 13–15 June 2014. PMLR.
- [443] L. Badea. Learning trading rules with inductive logic programming. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 39–46. Springer, 2000.

- [444] M. Bafna and J. Ullman. The price of selection in differential privacy. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 151–168, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [445] P. G. Bagos, T. D. Liakopoulos, and S. J. Hamodrakas. Faster gradient descent training of hidden Markov models, using individual learning rate adaptation. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 40–52. Springer, 2004.
- [446] S. Bai. A calculus for logical clustering. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 56–63. Springer-Verlag, 1994.
- [447] T. L. Bailey and C. Elkan. Cross-validation and modal theories. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 18, pages 345–359. MIT Press, 1995.
- [448] T. L. Bailey and C. Elkan. Unsupervised learning of multiple motifs in biopolymers using expectation maximization. *Machine Learning*, 21(1/2):51–80, 1995.
- [449] A. H. R. Baily and F. Denis. A spectral approach for probabilistic grammatical inference on trees. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 74–88, Berlin, Heidelberg, New York, 2010. Springer.
- [450] L. Baird. Residual algorithms: reinforcement learning with function approximation. In *Proc. 12th International Conference on Machine Learning*, pages 30–37. Morgan Kaufmann, 1995.
- [451] B. Bakker and T. Heskes. Task clustering and gating for Bayesian multitask learning. *Journal of Machine Learning Research*, 4:83–99, 2003.
- [452] S. Balakrishnan, S. S. Du, J. Li, and A. Singh. Computationally efficient robust sparse estimation in high dimensions. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 169–212, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [453] F. Balbach. Changing the inference type – keeping the hypothesis space. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 84–98. Springer, 2003.
- [454] F. J. Balbach. Teaching classes with high teaching dimension using few examples. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 668–683. Springer, 2005.
- [455] F. J. Balbach. Measuring teachability using variants of the teaching dimension. *Theoret. Comput. Sci.*, 397(1-3):94–113, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.

- [456] F. J. Balbach and T. Zeugmann. Teaching learners that can only perform restricted mind changes. Technical Report TCS-TR-A-05-5, Division of Computer Science, Hokkaido University, July 2005.
- [457] F. J. Balbach and T. Zeugmann. Teaching learners with restricted mind changes. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 474–489. Springer, Oct. 2005.
- [458] F. J. Balbach and T. Zeugmann. On the teachability of randomized learners. Technical Report TCS-TR-A-06-13, Division of Computer Science, Hokkaido University, Apr. 2006.
- [459] F. J. Balbach and T. Zeugmann. Teaching memoryless randomized learners without feedback. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 93–108. Springer, Oct. 2006.
- [460] F. J. Balbach and T. Zeugmann. Teaching randomized learners. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 229–243, Berlin, 2006. Springer.
- [461] F. J. Balbach and T. Zeugmann. Recent developments in algorithmic teaching. In A. H. Dediu, A. M. Ionescu, and C. Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2009, Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2009. Invited Paper.
- [462] F. J. Balbach and T. Zeugmann. Teaching randomized learners with feedback. *Inform. Comput.*, 209(3):296–319, 2011. Special Issue: 3rd International Conference on Language and Automata Theory and Applications (LATA 2009).
- [463] M.-F. Balcan, N. Bansal, A. Beygelzimer, D. Coppersmith, J. Langford, and G. B. Sorkin. Robust reductions from ranking to classification. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 604–619, Berlin, 2007. Springer.
- [464] M.-F. Balcan and A. Blum. A PAC-style model for learning from labeled and unlabeled data. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 111–126. Springer, 2005.
- [465] M.-F. Balcan and A. Blum. Clustering with interactive feedback. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 316–328, Berlin, Oct. 2008. Springer.
- [466] M. F. Balcan, A. Blum, S. Fine, and Y. Mansour. Distributed learning, communication complexity and privacy. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 26.1–26.22, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [467] M.-F. Balcan, A. Blum, and V. Nagarajan. Lifelong learning in costly feature spaces. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on*

- Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 250–287, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [468] M.-F. Balcan, A. Blum, and N. Srebro. Improved guarantees for learning via similarity functions. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 287–298. Omnipress, 2008.
- [469] M.-F. Balcan, A. Blum, and S. Vempala. On kernels, margins, and low-dimensional mappings. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 194–205. Springer, 2004.
- [470] M.-F. Balcan, A. Blum, and S. Vempala. Kernels as features: On kernels, margins, and low-dimensional mappings. *Machine Learning*, 65(1):79–94, 2006.
- [471] M.-F. Balcan, A. Blum, and S. Vempala. Efficient representations for lifelong learning and autoencoding. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 191–210, Paris, France, 03–06 July 2015. PMLR.
- [472] M.-F. Balcan, A. Broder, and T. Zhang. Margin based active learning. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 35–50, Berlin, 2007. Springer.
- [473] M. F. Balcan, F. Constantin, S. Iwata, and L. Wang. Learning valuation functions. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 4.1–4.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [474] M.-F. Balcan, S. S. Du, Y. Wang, and A. W. Yu. An improved gap-dependency analysis of the noisy power method. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 284–309, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [475] M. F. Balcan, V. Feldman, and C. Szepesvári, editors. *Conference on Learning Theory, 13-15 June 2014, Barcelona, Spain*, volume 35 of *Proceedings of Machine Learning Research*. PMLR, 2015.
- [476] M. F. Balcan and S. Hanneke. Robust interactive learning. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 20.1–20.34, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [477] M.-F. Balcan, S. Hanneke, and J. Wortman. The true sample complexity of active learning. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 45–56. Omnipress, 2008.
- [478] M.-F. Balcan and P. Long. Active and passive learning of linear separators under log-concave distributions. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 288–316. JMLR.org, 2013.

- [479] M.-F. Balcan, V. Nagarajan, E. Vitercik, and C. White. Learning-theoretic foundations of algorithm configuration for combinatorial partitioning problems. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 213–274, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [480] M. F. Balcan, H. Röglin, and S.-H. Teng. Agnostic clustering. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 384–398, Berlin/Heidelberg, 2009. Springer.
- [481] M. F. Balcan and C. Szepesvári. Preface. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1–2, Barcelona, Spain, 13–15 June 2014. PMLR.
- [482] M.-F. Balcan, E. Vitercik, and C. White. Learning combinatorial functions from pairwise comparisons. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 310–335, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [483] J. Balcázar, J. Castro, and D. Guijarro. A new abstract combinatorial dimension for exact learning via queries. *J. Comput. Syst. Sci.*, 64(1):2–21, 2002. Special issue for COLT 2000.
- [484] J. Balcázar, Y. Dai, and O. Watanabe. A random sampling technique for training support vector machines. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 119–134. Springer, 2001.
- [485] J. L. Balcázar, H. Buhrman, and M. Hermo. Learnability of Kolmogorov-easy circuit expressions via queries. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 112–124. Springer, 1995.
- [486] J. L. Balcázar, J. Castro, and D. Guijarro. Abstract combinatorial characterizations of exact learning via queries. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 248–254. Morgan Kaufmann, San Francisco, 2000.
- [487] J. L. Balcázar, J. Castro, and D. Guijarro. A general dimension for exact learning. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 354–367. Springer, 2001.
- [488] J. L. Balcázar, J. Castro, D. Guijarro, and H.-U. Simon. The consistency dimension and distribution-dependent learning from queries. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 77–92. Springer, 1999.
- [489] J. L. Balcázar, J. Castro, D. Guijarro, and H.-U. Simon. The consistency dimension and distribution-dependent learning from queries. *Theoret. Comput. Sci.*, 288(2):197–215, 2002. Special issue for ALT '99.
- [490] J. L. Balcázar, J. Díaz, R. Gavaldà, and O. Watanabe. A note on the query complexity of learning DFA. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan*,

- October 1992, *Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 53–62. Springer-Verlag, 1993.
- [491] J. L. Balcázar, J. Díaz, R. Gavaldà, and O. Watanabe. An optimal parallel algorithm for learning DFA. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 208–217, New York, NY, 1994. ACM Press.
- [492] J. L. Balcázar, P. M. Long, and F. Stephan, editors. *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2006. Springer.
- [493] J. L. Balcázar, P. M. Long, and F. Stephan. Editors’ introduction. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 1–9. Springer, Oct. 2006.
- [494] P. Baldi and G. Pollastri. The principled design of large-scale recursive neural network architectures—DAG-RNNs and the protein structure prediction problem. *Journal of Machine Learning Research*, 4:575–602, 2003.
- [495] G. Baliga and J. Case. Learning with higher order additional information. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 64–75. Springer-Verlag, 1994.
- [496] G. Baliga and J. Case. Learnability: Admissible, co-finite, and hypersimple languages. *J. Comput. Syst. Sci.*, 53(1):26–32, 1996. First appeared in ICALP ’93.
- [497] G. Baliga, J. Case, and S. Jain. Language learning with some negative information. *J. of Comput. Syst. Sci.*, 51(2):273–285, 1995.
- [498] G. Baliga, J. Case, S. Jain, and M. Suraj. Machine learning of higher order programs. *J. Symbolic Logic*, 59(2):486–500, 1994.
- [499] G. Baliga, J. Case, W. Merkle, and F. Stephan. Unlearning helps. In *Automata, Languages and Programming, 27th International Colloquium, ICALP 2000, Geneva, Switzerland, July 9-15, 2000, Proceedings*, volume 1853 of *Lecture Notes in Computer Science*, pages 844–855. Springer, 2000.
- [500] G. Baliga, S. Jain, and A. Sharma. Learning from multiple sources of inaccurate data. In *Analogical and Inductive Inference, International Workshop AII ’92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 108–128. Springer-Verlag, 1992.
- [501] G. Baliga, S. Jain, and A. Sharma. Learning from multiple sources of inaccurate data. *SIAM J. Comput.*, 26(4):961–990, 1997.
- [502] G. R. Baliga, J. Case, and S. Jain. Synthesizing enumeration techniques for language learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 169–180, New York, NY, 1996. ACM Press.
- [503] G. R. Baliga, J. Case, and S. Jain. The synthesis of language learners. *Inform. Comput.*, 152(1):16–43, 1999.

- [504] E. Balkanski and Y. Singer. The sample complexity of optimizing a convex function. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 275–301, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [505] B. Balle, J. Castro, and R. Gavaldà. Learning PDFA with asynchronous transitions. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 271–275. Springer, 2010.
- [506] B. Balle, J. Castro, and R. Gavaldà. A lower bound for learning distributions generated by probabilistic automata. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 179–193, Berlin, Heidelberg, New York, 2010. Springer.
- [507] B. Balle, J. Castro, and R. Gavaldà. Bootstrapping and learning PDFA in data streams. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 34–48, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [508] B. Balle, J. Castro, and R. Gavaldà. Learning probabilistic automata: A study in state distinguishability. *Theoret. Comput. Sci.*, 473:46–60, 2013. Special Issue on Learning Theory.
- [509] B. Balle, R. Eyraud, F. M. Luque, A. Quattoni, and S. Verwer. Results of the sequence prediction challenge (SPiCe): a competition on learning the next symbol in a sequence. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 132–136, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [510] B. Balle and M. Mohri. On the rademacher complexity of weighted automata. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 179–193, Berlin, Heidelberg, New York, 2015. Springer.
- [511] B. Balle and M. Mohri. Generalization bounds for learning weighted automata. *Theoret. Comput. Sci.*, 716:89–106, 2018. Special Issue on ALT 2015.
- [512] A. Balsubramani and Y. Freund. Optimally combining classifiers using unlabeled data. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 211–225, Paris, France, 03–06 July 2015. PMLR.
- [513] A. Balsubramani, Z. Karnin, R. E. Schapire, and M. Zoghi. Instance-dependent regret bounds for dueling bandits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 336–360, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [514] S. Baluja and R. Caruana. Removing the genetics from the standard genetic algorithm. In *Proc. 12th International Conference on Machine Learning*, pages 38–46. Morgan Kaufmann, 1995.

- [515] S. Baluja and S. Davies. Using optimal dependency-trees for combinatorial optimization: Learning the structure of the search space. In *Proc. 14th International Conference on Machine Learning*, pages 30–38. Morgan Kaufmann, 1997.
- [516] A. S. Bandeira, N. Boumal, and V. Voroninski. On the low-rank approach for semidefinite programs arising in synchronization and community detection. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 361–382, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [517] A. S. Bandeira, Y. Khoo, and A. Singer. Open problem: Tightness of maximum likelihood semidefinite relaxations. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1265–1267, Barcelona, Spain, 13–15 June 2014. PMLR.
- [518] C. Bandera, F. J. Vico, J. M. Bravo, M. E. Harmon, and L. C. B. III. Residual Q-learning applied to visual attention. In *Proc. 13th International Conference on Machine Learning*, pages 20–27. Morgan Kaufmann, 1996.
- [519] A. Banerjee, S. Chen, and V. Sivakumar. Open problem: Restricted eigenvalue condition for heavy tailed designs. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1752–1755, Paris, France, 03–06 July 2015. PMLR.
- [520] B. Banerjee, S. Debnath, and S. Sen. Combining multiple perspectives. In *Proc. 17th International Conf. on Machine Learning*, pages 33–40. Morgan Kaufmann, San Francisco, CA, 2000.
- [521] A. Banerji. Initializing neural networks using decision trees. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 1, pages 3–15. MIT Press, 1997.
- [522] R. B. Banerji. The logic of learning: a basis for pattern recognition and for improvement of performance. *Advances in Computers*, 24:177–216, 1985.
- [523] R. B. Banerji. Learning theories in a subset of a polyadic logic. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 267–278, San Mateo, CA, 1988. Morgan Kaufmann.
- [524] R. B. Banerji. Learning theoretical terms. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 2, pages 25–47. MIT Press, 1994.
- [525] J. Banks, C. Moore, J. Neeman, and P. Netrapalli. Information-theoretic thresholds for community detection in sparse networks. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 383–416, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [526] A. Bar-Hillel, T. Hertz, N. Shental, and D. Weinshall. Learning a mahalanobis metric from equivalence constraints. *Journal of Machine Learning Research*, 6:937–965, 2005.
- [527] A. Bar-Hillel and D. Weinshall. Learning with equivalence constraints and the relation to multiclass learning. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 640–654. Springer, 2003.



- [528] B. Barak and A. Moitra. Noisy tensor completion via the sum-of-squares hierarchy. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 417–445, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [529] Y. Baram, R. E. Yaniv, and K. Luz. Online choice of active learning algorithms. *Journal of Machine Learning Research*, 5:255–291, 2004.
- [530] R. Bareiss. A reply to Reich’s book review of exemplar-based knowledge acquisition. *Machine Learning*, 6:105–106, 1991.
- [531] R. Bareiss, B. W. Porter, and K. S. Murray. Supporting start-to-finish development of knowledge bases. *Machine Learning*, 4:259–283, 1989.
- [532] I. Barland. Some ideas on learning with directional feedback. Master’s thesis, UC Santa Cruz, June 1992.
- [533] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3:1107–1135, 2003.
- [534] C. Baroglio, A. Giordana, M. Kaiser, M. Nuttin, and R. Piola. Learning controllers for industrial robots. *Machine Learning*, 23:221–249, 1996.
- [535] A. Barron, B. Clarke, and D. Haussler. Information bounds for the risk of Bayesian predictions and the redundancy of universal codes. In *Proc. International Symposium on Information Theory*, Jan. 1993.
- [536] A. R. Barron. Approximation and estimation bounds for artificial neural networks. *Machine Learning*, 14(1):115–133, 1994.
- [537] A. R. Barron. Information theory in probability, statistics, learning, and neural nets. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, page 1, New York, NY, 1997. ACM Press.
- [538] P. Bartlett. Lower bounds on the Vapnik-Chervonenkis dimension of multi-layer threshold networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 144–150, New York, NY, 1993. ACM Press.
- [539] P. Bartlett, P. Grünwald, P. Harremoës, F. Hedayati, and W. Kotłowski. Horizon-independent optimal prediction with log-loss in exponential families. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 639–661. JMLR.org, 2013.
- [540] P. Bartlett, T. Linder, and G. Lugosi. A minimax lower bound for empirical quantizer design. In *Computational Learning Theory, Third European Conference, EuroCOLT ’97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 210–222. Springer, 1997.
- [541] P. Bartlett and Y. Mansour, editors. *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, New York, N.Y., 1998. ACM Press.
- [542] P. L. Bartlett. Learning with a slowly changing distribution. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 243–252, New York, NY, 1992. ACM Press.

- [543] P. L. Bartlett. Optimal online prediction in adversarial environments. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, page 34, Berlin, Heidelberg, New York, 2010. Springer.
- [544] P. L. Bartlett and J. Baxter. Estimation and approximation bounds for gradient-based reinforcement learning. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 133–141. Morgan Kaufmann, San Francisco, 2000.
- [545] P. L. Bartlett and J. Baxter. Estimation and approximation bounds for gradient-based reinforcement learning. *J. Comput. Syst. Sci.*, 64(1):133–150, 2002. Special issue for COLT 2000.
- [546] P. L. Bartlett and S. Ben-David. Hardness results for neural network approximation problems. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 50–62. Springer, 1999.
- [547] P. L. Bartlett and S. Ben-David. Hardness results for neural network approximation problems. *Theoret. Comput. Sci.*, 284(1):53–66, 2002.
- [548] P. L. Bartlett, S. Ben-David, and S. R. Kulkarni. Learning changing concepts by exploiting the structure of change. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 131–139, New York, NY, 1996. ACM Press.
- [549] P. L. Bartlett, S. Ben-David, and S. R. Kulkarni. Learning changing concepts by exploiting the structure of change. *Machine Learning*, 41(2):153–174, 2000. first in COLT '96.
- [550] P. L. Bartlett, S. Boucheron, and G. Lugosi. Model selection and error estimation. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 286–297. Morgan Kaufmann, San Francisco, 2000.
- [551] P. L. Bartlett, S. Boucheron, and G. Lugosi. Model selection and error estimation. *Machine Learning*, 48(1/3):85–113, 2002.
- [552] P. L. Bartlett, O. Bousquet, and S. Mendelson. Localized Rademacher complexities. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 44–58. Springer, 2002.
- [553] P. L. Bartlett, V. Dani, T. P. Hayes, S. Kakade, A. Rakhlin, and A. Tewari. High-probability regret bounds for bandit online linear optimization. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 335–342. Omnipress, 2008.
- [554] P. L. Bartlett, P. Fischer, and K.-U. Hoffgen. Exploiting random walks for learning. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 318–327, New York, NY, 1994. ACM Press.
- [555] P. L. Bartlett, P. Fischer, and K.-U. Hoffgen. Exploiting random walks for learning. *Inform. Comput.*, 176(2):121–135, 2002.
- [556] P. L. Bartlett, W. M. Koolen, A. Malek, E. Takimoto, and M. K. Warmuth. Minimax fixed-design linear regression. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 226–239, Paris, France, 03–06 July 2015. PMLR.

- [557] P. L. Bartlett and P. M. Long. More theorems about scale-sensitive dimensions and learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 392–401, New York, NY, 1995. ACM Press.
- [558] P. L. Bartlett and P. M. Long. Prediction, learning, uniform convergence, and scale-sensitive dimensions. *J. of Comput. Syst. Sci.*, 56(2):174–190, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [559] P. L. Bartlett, P. M. Long, and R. C. Williamson. Fat-shattering and the learnability of real-valued functions. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 299–310, New York, NY, 1994. ACM Press.
- [560] P. L. Bartlett, P. M. Long, and R. C. Williamson. Fat-shattering and the learnability of real-valued functions. *J. of Comput. Syst. Sci.*, 52(3):434–452, 1996.
- [561] P. L. Bartlett and S. Mendelson. Rademacher and Gaussian complexities: Risk bounds and structural results. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 224–240. Springer, 2001.
- [562] P. L. Bartlett and S. Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3:463–482, 2002.
- [563] P. L. Bartlett, S. Mendelson, and P. Philips. Local complexities for empirical risk minimization. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 270–284. Springer, 2004.
- [564] P. L. Bartlett and A. Tewari. Sparseness versus estimating conditional probabilities: Some asymptotic results. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 564–578. Springer, 2004.
- [565] P. L. Bartlett and R. C. Williamson. Investigating the distribution assumptions in the PAC learning model. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 24–32, San Mateo, CA, 1991. Morgan Kaufmann.
- [566] G. Bartók. A near-optimal algorithm for finite partial-monitoring games against adversarial opponents. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 696–710. JMLR.org, 2013.
- [567] G. Bartók, D. Pál, and C. Szepesvári. Toward a classification of finite partial-monitoring games. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 224–238, Berlin, Heidelberg, New York, 2010. Springer.
- [568] G. Bartók, D. Pál, and C. Szepesvári. Minimax regret of finite partial-monitoring games in stochastic environments. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 133–154, Budapest, Hungary, 09–11 June 2011. PMLR.

- [569] G. Bartók and C. Szepesvári. Partial monitoring with side information. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 305–319, Berlin, Heidelberg, New York, 2012. Springer.
- [570] G. Bartók, C. Szepesvári, and S. Zilles. Active learning of group-structured environments. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 329–343, Berlin, Oct. 2008. Springer.
- [571] G. Bartók, C. Szepesvári, and S. Zilles. Models of active learning in group-structured state spaces. *Inform. Comput.*, 208(4):364–384, 2010.
- [572] R. D. Barve and P. M. Long. On the complexity of learning from drifting distributions. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 122–130, New York, NY, 1996. ACM Press.
- [573] J. Bārzdiņš and G. Bārzdiņš. Towards efficient inductive synthesis: Rapid construction of local regularities. In *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardtsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 132–140. Springer-Verlag, 1993.
- [574] J. Bārzdiņš, R. Freivalds, and C. H. Smith. Learning with belief levels. *J. Comput. Syst. Sci.*, 74(4):527–545, 2008.
- [575] J. Barzdin. On synthesizing programs given by examples. In *International Symposium on Theoretical Programming*, volume 5 of *Lecture Notes in Computer Science*, pages 53–63. Springer-Verlag, 1974.
- [576] J. M. Barzdin. On decoding automata in the absence of an upper bound on the number of states. *Dokl. Akad. Nauk. SSSR*, pages 1048–1051, 1970.
- [577] J. M. Barzdin. Prognostication of automata and functions. In C. V. Freiman, J. E. Griffith, and J. L. Rosenfeld, editors, *Information Processing 71, Proceedings of IFIP Congress 71, Volume 1 - Foundations and Systems, Ljubljana, Yugoslavia, August 23-28, 1971*, pages 81–84. North-Holland, 1972.
- [578] J. M. Barzdin. Limiting synthesis of  $t$  numbers. *Latvijas Valsts Univ. Zinatm. Raksti*, 210:112–116, 1974.
- [579] J. M. Barzdin. Two theorems on the limiting synthesis of functions. In *Theory of Algorithms and Programs*, volume 1, pages 82–88. Latvian State University, Riga, 1974. In Russian.
- [580] J. M. Barzdin. Über eine Eigenschaft limes berechenbarer Funktionale. In *Theory of Algorithms and Programs*, volume 1, pages 20–24. Latvian State University, Riga, 1974.
- [581] J. M. Barzdin. Inductive inference of automata, functions and programs. In *Amer. Math. Soc. Transl.*, pages 107–122, 1977.
- [582] J. M. Barzdin and G. J. Barzdin. Rapid construction of algebraic axioms from samples. *Theoret. Comput. Sci.*, 90:199–208, 1991.
- [583] J. M. Barzdin and R. V. Freivald. On the prediction of general recursive functions. *Soviet Math. Doklady*, 13:1224–1228, 1972.

- [584] J. M. Barzdin and R. Freivalds. Prediction and limiting synthesis of recursively enumerable classes of functions. *Latvijas Valsts Univ. Zinatm. Raksti*, 210:101–111, 1974.
- [585] J. M. Barzdin, E. B. Kinber, and K. M. Podnieks. Concerning synthesis and prediction of functions. In *Theory of Algorithms and Programs*, volume 1, pages 117–128. Latvian State University, Riga, 1974.
- [586] J. M. Barzdin and K. M. Podnieks. Towards a theory of inductive inference. In *Proc. 2nd Conf. Mathematical Foundations of Computer Science*, pages 9–15. Math. Inst. of the Slovak Academy of Sciences, 1973.
- [587] G. Barzdins. Inductive synthesis of encoding for algebraic abstract data types. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 328–338. Springer-Verlag, 1989.
- [588] G. Barzdins. Inductive synthesis of term rewriting systems. In *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 253–285. Springer, Berlin, 1991.
- [589] J. Barzdins. Towards efficient inductive synthesis from input/output examples. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, page 1. Springer-Verlag, 1994.
- [590] J. Bārzdiņš, G. Bārzdiņš, K. Apsītis, and U. Sarkans. Towards efficient inductive synthesis of expressions from input/output examples. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 59–72. Springer-Verlag, 1993.
- [591] J. Bārzdiņš, R. Freivalds, and C. H. Smith. Learning with confidence. In *Proc. 13th Annual Symposium on Theoretical Aspects of Computer Science: STACS '96*, volume 1046 of *Lecture Notes in Computer Science*, pages 207–218. Springer, 1996.
- [592] J. Bārzdiņš, R. Freivalds, and C. H. Smith. Learning formulae from elementary facts. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 272–285. Springer, 1997.
- [593] J. Bārzdiņš and U. Sarkans. Incorporating hypothetical knowledge into the process of inductive synthesis. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 156–168. Springer, 1996.
- [594] J. Basak, A. Sudarshan, D. Trivedi, and M. S. Santhanam. Weather data mining using independent component analysis. *Journal of Machine Learning Research*, 5:239–253, 2004.
- [595] G. Basalyga and M. Rattray. Statistical dynamics of on-line independent component analysis. *Journal of Machine Learning Research*, 4:1393–1410, 2003.
- [596] R. Bassily, S. Moran, I. Nachum, J. Shafer, and A. Yehudayoff. Learners that use little information. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 25–55. PMLR, 07–09 Apr. 2018.

- [597] K. Basye, T. Dean, and J. S. Vitter. Coping with uncertainty in map learning. *Machine Learning*, 29(1):65–88, 1997.
- [598] R. Battiti. Accelerated backpropagation learning: Two optimization methods. *Complex Systems*, 3(4):331–342, Aug. 1989.
- [599] T. Batu, S. Guha, and S. Kannan. Inferring mixtures of markov chains. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 186–199. Springer, 2004.
- [600] T. Batu and P. Taptagaporn. Competitive portfolio selection using stochastic predictions. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 288–302. Springer, 2016.
- [601] V. Bauche. Using the algorithm of analogy for generation of robot programs. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 19–22. Springer-Verlag, 1986.
- [602] E. Bauer and R. Kohavi. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, 36(1/2):105–139, 1999.
- [603] E. Baum. A polynomial time algorithm that learns two hidden net units. *Neural Computation*, 2:510–522, 1990.
- [604] E. Baum. Neural net algorithms that learn in polynomial time from examples and queries. *IEEE Transactions on Neural Networks*, 2:5–19, 1991.
- [605] E. Baum and D. Haussler. What size net gives valid generalization? *Neural Computation*, 1(1):151–160, 1989.
- [606] E. Baum and F. Wilczek. Supervised learning of probability distributions by neural networks. In D. Z. Anderson, editor, *Neural Information Processing Systems*, pages 52–61, New York, 1988. American Insitute of Physics.
- [607] E. B. Baum. On learning a union of half spaces. *Journal of Complexity*, 6(1):67–101, Mar. 1990.
- [608] E. B. Baum. The perceptron algorithm is fast for nonmalicious distributions. *Neural Computation*, 2:248–260, 1990.
- [609] E. B. Baum. Polynomial time algorithms for learning neural nets. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 258–272, San Mateo, CA, 1990. Morgan Kaufmann.
- [610] E. B. Baum. When are k-nearest neighbor and backpropagation accurate for feasible-sized sets of examples? In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 14, pages 415–442. MIT Press, 1994.
- [611] E. B. Baum. Toward a model of mind as a laissez-faire economy of idiots. In *Proc. 13th International Conference on Machine Learning*, pages 28–36. Morgan Kaufmann, 1996.

- [612] E. B. Baum. Toward a model of intelligence as an economy of agents. *Machine Learning*, 35(2):155–185, 1999.
- [613] E. B. Baum, D. Boneh, and C. Garrett. On genetic algorithms. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 230–239, New York, NY, 1995. ACM Press.
- [614] J. Baxter. Learning internal representations. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 311–320, New York, NY, 1995. ACM Press.
- [615] J. Baxter. A Bayesian/information theoretic model of bias learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 77–88, New York, NY, 1996. ACM Press.
- [616] J. Baxter. A Bayesian/information theoretic model of learning to learn via multiple task sampling. *Machine Learning*, 28(1):7–39, 1997.
- [617] J. Baxter. The canonical distortion measure for vector quantization and function approximation. In *Proc. 14th International Conference on Machine Learning*, pages 39–47. Morgan Kaufmann, 1997.
- [618] J. Baxter. A model of inductive bias learning. *Journal of Artificial Intelligence Research*, 12:149–198, 2000.
- [619] J. Baxter and P. Bartlett. A result relating convex  $n$ -widths to covering numbers with some applications to neural networks. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 251–259. Springer, 1997.
- [620] J. Baxter and P. L. Bartlett. Reinforcement learning in POMDP’s via direct gradient ascent. In *Proc. 17th International Conf. on Machine Learning*, pages 41–48. Morgan Kaufmann, San Francisco, CA, 2000.
- [621] J. Baxter and P. L. Bartlett. Infinite-horizon policy-gradient estimation. *Journal of Artificial Intelligence Research*, 15:319–350, 2001.
- [622] J. Baxter and N. Cesa-Bianchi. Guest editors’ introduction. *Machine Learning*, 37(3):239–240, 1999.
- [623] J. Baxter, A. Tridgell, and L. Weaver. Learning to play chess using temporal differences. *Machine Learning*, 40(3):243–263, 2000.
- [624] J. Baxter, A. Tridgell, and L. Weaver. Knightcap: a chess program that learns by combining TD( $\lambda$ ) with game-tree search. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 28–36. Morgan Kaufmann, San Francisco, CA, 1998.
- [625] R. A. Baxter and J. J. Oliver. The kindest cut: minimum message length segmentation. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 83–90. Springer, 1996.
- [626] S. D. Bay. Combining nearest neighbor classifiers through multiple feature subsets. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 37–45. Morgan Kaufmann, San Francisco, CA, 1998.
- [627] S. D. Bay and M. J. Pazdani. Characterizing model errors and differences. In *Proc. 17th International Conf. on Machine Learning*, pages 49–56. Morgan Kaufmann, San Francisco, CA, 2000.

- [628] F. Bayeh, Z. Gao, and S. Zilles. Erasing pattern languages distinguishable by a finite number of strings. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 72–108, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [629] M. F. Bear, L. N. Cooper, and F. F. Ebner. A physiological basis for a theory of synapse modification. *Science*, 237:42–48, July 3 1987.
- [630] D. Beaver. Polynomially sized Boolean circuits are not learnable. Technical Report TR-13-87, Aiken Computation Laboratory, Harvard University, Dec. 1987.
- [631] L. Becerra-Bonache, J. Case, S. Jain, and F. Stephan. Iterative learning of simple external contextual languages. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 359–373, Berlin, Oct. 2008. Springer.
- [632] L. Becerra-Bonache, J. Case, S. Jain, and F. Stephan. Iterative learning of simple external contextual languages. *Theoret. Comput. Sci.*, 411(29-30):2741–2756, 2010. Special issue for ALT 2008.
- [633] L. Becerra-Bonache, A. H. Dediú, and C. Tîrnăuică. Learning DFA from correction and equivalence queries. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 281–292, Berlin, 2006. Springer.
- [634] L. Becerra-Bonache and T. Yokomori. Learning mild context-sensitiveness: Toward understanding children’s language learning. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 53–64. Springer, 2004.
- [635] D. Béchet and A. Foret. Simple K-star categorial dependency grammars and their inference. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 3–14, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [636] D. Béchet, A. Foret, and I. Tellier. Learnability of pregroup grammars. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 65–76. Springer, 2004.
- [637] D. Beeferman, A. Berger, and J. D. Lafferty. Statistical models for text segmentation. *Machine Learning*, 34:177–210, 1999.
- [638] H. R. Beick. Zur Konvergenzgeschwindigkeit von Strategien der induktiven Inferenz. *Elektronische Informationsverarbeitung und Kybernetik*, 18:163–172, 1982.
- [639] H.-R. Beick and V. Stankov. GoslerP – A logic programming tool for inductive inference. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 482–510. Springer, 1995.
- [640] A. Beimel, F. Bergadano, N. H. Bshouty, E. Kushilevitz, and S. Varricchio. Learning functions represented as multiplicity automata. *J. ACM*, 47(3):506–530, 2000.



- [641] A. Beimel, F. Geller, and E. Kushilevitz. The query complexity of finding local minima in the lattice. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 294–302, New York, NY, 1998. ACM Press.
- [642] A. Beimel, F. Geller, and E. Kushilevitz. The query complexity of finding local minima in the lattice. *Inform. Comput.*, 171(1):69–83, Nov. 2001.
- [643] A. Beimel and E. Kushilevitz. Learning boxes in high dimension. *Algorithmica*, 22(1/2):76–90, 1998.
- [644] A. Beimel and E. Kushilevitz. Learning unions of high-dimensional boxes over the reals. *Information Processing Letters*, 73(5-6):213–220, 2000.
- [645] J. Béjar. Improving knowledge discovery using domain knowledge in unsupervised learning. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 47–54. Springer, 2000.
- [646] R. Bekkerman, R. El-Yaniv, N. Tishby, and Y. Winter. Distributional word clusters vs. words for text categorization. *Journal of Machine Learning Research*, 3:1183–1208, 2003.
- [647] R. K. Belew and S. Forrest. Learning and programming in classifier systems. *Machine Learning*, 3:193–223, 1988.
- [648] M. Belkin, I. Matveeva, and P. Niyogi. Regularization and semi-supervised learning on large graphs. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 624–638. Springer, 2004.
- [649] M. Belkin and P. Niyogi. Towards a theoretical foundation for Laplacian-based manifold methods. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 486–500. Springer, 2005.
- [650] M. Belkin, Q. Que, Y. Wang, and X. Zhou. Toward understanding complex spaces: Graph laplacians on manifolds with singularities and boundaries. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 36.1–36.26, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [651] M. Belkin, L. Rademacher, and J. Voss. Blind signal separation in the presence of gaussian noise. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 270–287. JMLR.org, 2013.
- [652] M. Belkin, L. Rademacher, and J. Voss. Basis learning as an algorithmic primitive. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 446–487, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [653] D. A. Bell and H. Wang. A formalism for relevance and its application in feature subset selection. *Machine Learning*, 41(2):175–195, 2000.
- [654] M. Bellare. A technique for upper bounding the spectral norm with applications to learning. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 62–70, New York, NY, 1992. ACM Press.

- [655] P. C. Bellec. Aggregation of supports along the lasso path. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 488–529, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [656] A. Belloni, R. M. Freund, and S. S. Vempala. An efficient re-scaled perceptron algorithm for conic systems. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 393–408, Berlin, 2007. Springer.
- [657] A. Belloni, T. Liang, H. Narayanan, and A. Rakhlin. Escaping the local minima via simulated annealing: Optimization of approximately convex functions. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 240–265, Paris, France, 03–06 July 2015. PMLR.
- [658] A. Belz. PCFG learning by nonterminal partition search. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 14–27. Springer, 2002.
- [659] A. Ben-David. Monotonicity maintenance in information-theoretic machine learning algorithms. *Machine Learning*, 19(1):29–43, 1995.
- [660] A. Ben-David and J. Mandel. (research note) classification accuracy: Machine learning vs. explicit knowledge acquisition. *Machine Learning*, 18(1):109–114, 1995.
- [661] S. Ben-David. Can finite samples detect singularities of real-valued functions? *Algorithmica*, 22(1/2):3–17, 1998.
- [662] S. Ben-David. A framework for statistical clustering with a constant time approximation algorithms for  $K$ -median clustering. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 415–426. Springer, 2004.
- [663] S. Ben-David. Alternative measures of computational complexity with applications to agnostic learning. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 231–235, Berlin, 2006. Springer.
- [664] S. Ben-David. A framework for statistical clustering with constant time approximation algorithms for  $K$ -median and  $K$ -means clustering. *Machine Learning*, 66(2-3):243–257, 2007.
- [665] S. Ben-David and S. Ben-David. Learning a classifier when the labeling is known. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 440–451, Berlin, Heidelberg, New York, 2011. Springer.
- [666] S. Ben-David and G. M. Benedek. Measurability constraints on PAC learnability. Technical report, Technion, Haifa, 1991.
- [667] S. Ben-David, G. M. Benedek, and Y. Mansour. A parametrization scheme for classifying models of learnability. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 285–302, San Mateo, CA, 1989. Morgan Kaufmann.

- [668] S. Ben-David, J. Case, and A. Maruoka, editors. *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2004. Springer.
- [669] S. Ben-David, J. Case, and T. Zeugmann. Foreword. *Theoret. Comput. Sci.*, 382(3):167–169, 2007. Special issue for ALT 2004.
- [670] S. Ben-David, N. Cesa-Bianchi, D. Haussler, and P. M. Long. Characterizations of learnability for classes of  $\{0, \dots, n\}$ -valued functions. *J. of Comput. Syst. Sci.*, 50(1):74–86, 1995.
- [671] S. Ben-David, N. Cesa-Bianchi, and P. M. Long. Characterizations of learnability for classes of  $\{0, \dots, n\}$ -valued functions. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 333–340, New York, NY, 1992. ACM Press.
- [672] S. Ben-David and E. Dichterman. Learning with restricted focus of attention. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 287–296, New York, NY, 1993. ACM Press.
- [673] S. Ben-David and E. Dichterman. Learnability with restricted focus of attention guarantees noise-tolerance. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 248–259. Springer-Verlag, 1994.
- [674] S. Ben-David and E. Dichterman. Learning with restricted focus of attention. *J. of Comput. Syst. Sci.*, 56(3):277–298, 1998.
- [675] S. Ben-David and N. Eiron. Self-directed learning and its relation to the VC-dimension and to teacher-directed learning. *Machine Learning*, 33(1):87–104, 1998.
- [676] S. Ben-David, N. Eiron, and E. Kushilevitz. On self-directed learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 136–143, New York, NY, 1995. ACM Press.
- [677] S. Ben-David, N. Eiron, and P. M. Long. On the difficulty of approximately maximizing agreements. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 266–274. Morgan Kaufmann, San Francisco, 2000.
- [678] S. Ben-David, N. Eiron, and H. U. Simon. The complexity of densest region detection. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 255–265. Morgan Kaufmann, San Francisco, 2000.
- [679] S. Ben-David, N. Eiron, and H. U. Simon. Limitations of learning via embeddings in euclidean half-spaces. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 385–401. Springer, 2001.
- [680] S. Ben-David, N. Eiron, and H.-U. Simon. The computational complexity of densest region detection. *J. Comput. Syst. Sci.*, 64(1):22–47, 2002. Special issue for COLT 2000.
- [681] S. Ben-David, N. Eiron, and H. U. Simon. Limitations of learning via embeddings in euclidean half spaces. *Journal of Machine Learning Research*, 3:441–461, 2002.
- [682] S. Ben-David and L. Gurvits. A note on VC-dimension and measure of sets of reals. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 454–462, New York, NY, 1995. ACM Press.

- [683] S. Ben-David, A. Itai, and E. Kushilevitz. Learning by distances. *Inform. Comput.*, 117(2):240–250, 1995.
- [684] S. Ben-David and M. Jacovi. On learning in the limit and non-uniform  $(\epsilon, \delta)$ - learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 209–217, New York, NY, 1993. ACM Press.
- [685] S. Ben-David, E. Kushilevitz, and Y. Mansour. Online learning versus offline learning. *Machine Learning*, 29(1):45–63, 1997.
- [686] S. Ben-David and M. Lindenbaum. Learning distributions by their density levels: A paradigm for learning without a teacher. *J. Comput. Syst. Sci.*, 55(1):171–182, 1997. Special issue for EuroCOLT '95.
- [687] S. Ben-David and M. Lindenbaum. Localization vs. identification of semi-algebraic sets. *Machine Learning*, 32(3):207–224, 1998. Earlier version in 6th COLT, 1993.
- [688] S. Ben-David and P. Long, editors. *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, New York, N.Y., 1999. ACM Press.
- [689] S. Ben-David, P. M. Long, and Y. Mansour. Agnostic boosting. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 507–516. Springer, 2001.
- [690] S. Ben-David, T. Lu, and D. Pál. Does unlabeled data provably help? Worst-case analysis of the sample complexity of semi-supervised learning. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 33–44. Omnipress, 2008.
- [691] S. Ben-David and U. von Luxburg. Relating clustering stability to properties of cluster boundaries. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 379–390. Omnipress, 2008.
- [692] S. Ben-David, U. von Luxburg, and D. Pál. A sober look at clustering stability. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 5–19, Berlin, 2006. Springer.
- [693] S. Ben-David, D. Pál, and H. U. Simon. Stability of  $k$ -means clustering. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 20–34, Berlin, 2007. Springer.
- [694] S. Ben-David and L. Reyzin. Data stability in clustering: A closer look. *Theoret. Comput. Sci.*, 558:51–61, 2014. Special Issue: Algorithmic Learning Theory.
- [695] S. Ben-David and R. Schuller. Exploiting task relatedness for multiple task learning. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 567–580. Springer, 2003.

- [696] S. Ben-David and R. Urner. On the hardness of domain adaptation and the utility of unlabeled target samples. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 139–153, Berlin, Heidelberg, New York, 2012. Springer.
- [697] S. Ben-David and R. Urner. The sample complexity of agnostic learning under deterministic labels. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 527–542, Barcelona, Spain, 13–15 June 2014. PMLR.
- [698] S. Ben-David and R. Urner. On version space compression. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 50–64. Springer, 2016.
- [699] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik. Support vector clustering. *Journal of Machine Learning Research*, 2:125–137, 2001.
- [700] M. A. Bender and D. K. Slonim. The power of team exploration: two robots can learn unlabeled directed graphs. In *Proc. of the 35rd Annual Symposium on Foundations of Computer Science*, pages 75–85. IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [701] G. M. Benedek and A. Itai. Learnability by fixed distributions. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 80–90, San Mateo, CA, 1988. Morgan Kaufmann.
- [702] G. M. Benedek and A. Itai. Nonuniform learnability. In *ICALP*, pages 82–92, July 1988.
- [703] G. M. Benedek and A. Itai. Learnability with respect to fixed distributions. *Theoret. Comput. Sci.*, 86(2):377–389, 1991.
- [704] G. M. Benedek and A. Itai. Dominating distributions and learnability. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 253–264, New York, NY, 1992. ACM Press.
- [705] G. M. Benedek and A. Itai. Nonuniform learnability. *J. of Comput. Syst. Sci.*, 48(2):311–323, 1994.
- [706] Y. Bengio and N. Chapados. Extensions to metric-based model selection. *Journal of Machine Learning Research*, 3:1209–1227, 2003.
- [707] Y. Bengio and O. Delalleau. On the expressive power of deep architectures. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 18–36, Berlin, Heidelberg, New York, 2011. Springer.
- [708] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin. A neural probabilistic language model. *Journal of Machine Learning Research*, 3:1137–1155, 2003.
- [709] Y. Bengio and D. Schuurmans. Guest introduction: Special issue on new methods for model selection and model combination. *Machine Learning*, 48(1/3):5–7, 2002.
- [710] R. Bennet and N. H. Bshout. Learning attribute-efficiently with corrupt oracles. *Theoret. Comput. Sci.*, 387(1):32–50, 2007. Special issue for ALT 2005.

- [711] R. Bennet and N. H. Bshouty. Learning attribute-efficiently with corrupt oracles. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 183–197. Springer, Oct. 2005.
- [712] K. P. Bennett and E. J. Bredensteiner. Duality and geometry in SVM classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 57–64. Morgan Kaufmann, San Francisco, CA, 2000.
- [713] K. P. Bennett, N. Cristianini, J. Shawe-Taylor, and D. Wu. Enlarging the margins in perceptron decision trees. *Machine Learning*, 41(3):295–313, 2000.
- [714] K. P. Bennett, A. Demiriz, and J. Shawe-Taylor. A column generation algorithm for boosting. In *Proc. 17th International Conf. on Machine Learning*, pages 65–72. Morgan Kaufmann, San Francisco, CA, 2000.
- [715] S. W. Bennett and G. F. DeJong. Real-world robotics: Learning to plan for robust execution. *Machine Learning*, 23:121–161, 1996.
- [716] S. Benson. Inductive learning of reactive action models. In *Proc. 12th International Conference on Machine Learning*, pages 47–54. Morgan Kaufmann, 1995.
- [717] H. Bensusan and A. Kalousis. Estimating the predictive accuracy of a classifier. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 25–36. Springer, 2001.
- [718] J. van den Berg and J. C. Bioch. Some theorems concerning the free energy of (un)constrained stochastic Hopfield neural networks. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 298–312. Springer, 1995.
- [719] F. Bergadano, N. H. Bshouty, C. Tamon, and S. Varricchio. On learning branching programs and small depth circuits. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 150–161. Springer, 1997.
- [720] F. Bergadano and D. Gunetti. Funtional inductive logic programming with queries to the user. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 323–328. Springer-Verlag, 1993.
- [721] F. Bergadano, S. Matwin, R. S. Michalski, and J. Zhang. Learning two-tiered descriptions of flexible concepts: The POSEIDON system. *Machine Learning*, 8:5–43, 1992.
- [722] F. Bergadano and L. Saitta. On the error probabily of boolean concept descriptions. In *Proc. of the 1989 European Working Session on Learning*, pages 25–35, 1989.
- [723] F. Bergadano and S. Varricchio. Learning behaviors of automata from shortest counterexamples. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 380–391. Springer, 1995.
- [724] F. Bergadano and S. Varricchio. Learning behaviors of automata from multiplicity and equivalence queries. *SIAM J. Comput.*, 25(6):1268–1280, 1996.

- [725] B. Berger, J. Rompel, and P. W. Shor. Efficient NC algorithms for set cover with applications to learning and geometry. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 54–59. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [726] B. Berger, J. Rompel, and P. W. Shor. Efficient NC algorithms for set cover with applications to learning and geometry. *J. of Comput. Syst. Sci.*, 49(3):454–477, 1994.
- [727] J. Berger and C. Pair. Inference for regular bilanguages. *J. of Comput. Syst. Sci.*, 16:100–122, 1978.
- [728] U. Berggren. Linear time deterministic learning of  $k$ -term DNF. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 37–40, New York, NY, 1993. ACM Press.
- [729] J. Bergstra, N. Casagrande, D. Erhan, D. Eck, and B. Kégl. Aggregate features and ADABOOST for music classification. *Machine Learning*, 65(2-3):473–484, 2006.
- [730] P. Berman and R. Roos. Learning one-counter languages in polynomial time. In *Proc. 28th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 61–67. IEEE Computer Society Press, Los Alamitos, CA, 1987.
- [731] M. Bernard, J.-C. Janodet, and M. Sebban. A discriminative model of stochastic edit distance in the form of a conditional transducer. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 240–252, Berlin, 2006. Springer.
- [732] A. Bernstein, S. Mannor, and N. Shimkin. Opportunistic strategies for generalized no-regret problems. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 158–171. JMLR.org, 2013.
- [733] A. Bernstein and N. Shimkin. Adaptive aggregation for reinforcement learning with efficient exploration: Deterministic domains. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 323–334. Omnipress, 2008.
- [734] E. J. Bernstein. Absolute error bounds for learning linear functions online. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 160–163, New York, NY, 1992. ACM Press.
- [735] A. Beros and C. de la Higuera. A canonical semi-deterministic transducer. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 33–48, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [736] A. A. Beros, Z. Gao, and S. Zilles. Classifying the arithmetical complexity of teaching models. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 145–160. Springer, 2016.
- [737] Q. Berthet and P. Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1046–1066. JMLR.org, 2013.

- [738] A. Bertoni, P. Campadelli, A. Morpurgo, and S. Panizza. Polynomial uniform convergence and polynomial-sample learnability. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 265–271, New York, NY, 1992. ACM Press.
- [739] A. Bertoni, C. Choffrut, M. Goldwurm, and V. Lonati. Local limit distributions in pattern statistics: Beyond the markovian models. In *STACS 2004, 21st Annual Symposium on Theoretical Aspects of Computer Science, Montpellier, France, March 25-27, 2004, Proceedings*, volume 2996 of *Lecture Notes in Computer Science*, pages 117–128. Springer, 2004.
- [740] D. Bertsimas, D. Gamarnik, and J. N. Tsitsiklis. Estimation of time-varying parameters in statistical models: An optimization approach. *Machine Learning*, 35(3):225–245, 1999.
- [741] R. Berwick. *The Acquisition of Syntactic Knowledge*. MIT Press, Cambridge, Massachusetts, 1985.
- [742] R. Berwick. Learning from positive-only examples. In *Machine Learning, Vol. II*, pages 625–645. Morgan Kaufmann, 1986.
- [743] R. C. Berwick. Language learning by automata induction. *Machine Learning*, 2(3):9–38, 1987.
- [744] R. C. Berwick and S. Pilato. Learning syntax by automata induction. *Machine Learning*, 2(1):9–38, 1987.
- [745] J. Besombes and J.-Y. Marion. Learning tree languages from positive examples and membership queries. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 440–453. Springer, 2004.
- [746] J. Besombes and J.-Y. Marion. Learning tree languages from positive examples and membership queries. *Theoret. Comput. Sci.*, 382(3):183–197, 2007. Special issue for ALT 2004.
- [747] L. Besson and E. Kaufmann. Multi-Player Bandits Revisited. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 56–92. PMLR, 07–09 Apr. 2018.
- [748] M. Betke. Algorithms for exploring an unknown graph. Master’s thesis, MIT Department of Electrical Engineering and Computer Science, Feb. 1992. (Published as MIT Laboratory for Computer Science Technical Report MIT/LCS/TR-536 (March, 1992)).
- [749] M. Betke, R. L. Rivest, and M. Singh. Piecemeal learning of an unknown environment. *Machine Learning*, 18(2/3):231–254, 1995. Special issue for COLT ’93.
- [750] A. Beygelzimer, J. Langford, and P. Ravikumar. Error-correcting tournaments. In R. Gavalda, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 247–262, Berlin/Heidelberg, 2009. Springer.
- [751] S. Bhansali and M. T. Harandi. Synthesis of UNIX programs using derivational analogy. *Machine Learning*, 10:7–55, 1993.
- [752] S. Bhansali and M. T. Harandi. What makes derivational analogy work: an experience report using APU. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 17, pages 303–321. MIT Press, 1997.



- [753] B. Bhanu and A. Dong. Concepts learning with fuzzy clustering and relevance feedback. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 102–116. Springer, 2001.
- [754] A. Bhaskara, M. Charikar, A. Moitra, and A. Vijayaraghavan. Open problem: Tensor decompositions: Algorithms up to the uniqueness threshold? In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1280–1282, Barcelona, Spain, 13–15 June 2014. PMLR.
- [755] A. Bhaskara, M. Charikar, and A. Vijayaraghavan. Uniqueness of tensor decompositions with applications to polynomial identifiability. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 742–778, Barcelona, Spain, 13–15 June 2014. PMLR.
- [756] N. Bhatnagar and J. Mostow. On-line learning from search failures. *Machine Learning*, 15(1):69–117, 1994.
- [757] S. Bhojanapalli, A. Kyrillidis, and S. Sanghavi. Dropping convexity for faster semi-definite optimization. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 530–582, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [758] J. Bi, K. Bennett, M. Embrechts, C. Breneman, and M. Song. Dimensionality reduction via sparse support vector machines. *Journal of Machine Learning Research*, 3:1229–1243, 2003.
- [759] J. Bi and V. N. Vapnik. Learning with rigorous support vector machines. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 243–257. Springer, 2003.
- [760] G. Biau and L. Devroye. Cellular tree classifiers. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 8–17, Berlin, Heidelberg, New York, 2014. Springer.
- [761] W. Bibel and M. Thielscher. Deductive plan generation. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 2–5. Springer-Verlag, 1994.
- [762] T. D. Bie, M. Momma, and N. Cristianini. Efficiently learning the metric with side-information. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 175–189. Springer, 2003.
- [763] L. Bienvenu, B. Monin, and A. Shen. Algorithmic identification of probabilities is hard. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 85–95, Berlin, Heidelberg, New York, 2014. Springer.

- [764] A. W. Biermann. On the inference of Turing machines from sample computations. *Art. Int.*, 3:181–198, 1972.
- [765] A. W. Biermann. The inference of regular LISP programs from examples. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-8:585–600, 1978.
- [766] A. W. Biermann, R. I. Baum, and F. E. Petry. Speeding up the synthesis of programs from traces. *IEEE Trans. on Computers*, C-24:122–136, 1975.
- [767] A. W. Biermann and J. Feldman. On the synthesis of finite-state machines from samples of their behavior. *IEEE Trans. Computers*, 21:592–597, 1972.
- [768] A. W. Biermann and J. A. Feldman. A survey of results in grammatical inference. In S. Watanabe, editor, *Frontiers of Pattern Recognition*, pages 31–54. Academic Press, N. Y., 1972.
- [769] A. W. Biermann and R. Krishnaswamy. Constructing programs from example computations. *IEEE Trans. on Software Engineering*, SE-2:141–153, 1976.
- [770] A. W. Biermann and D. R. Smith. The hierarchical synthesis of LISP scanning functions. In *Information Processing 77*, pages 41–45. North-Holland, Amsterdam, 1977.
- [771] A. Bifet and R. Gavaldà. Kalman filters and adaptive windows for learning in data streams. In *Discovery Science, 9th International Conference, DS 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4265 of *Lecture Notes in Artificial Intelligence*, pages 29–40. Springer, Oct. 2006.
- [772] D. M. Bikel, R. Schwartz, and R. M. Weischedel. An algorithm that learns what’s in a name. *Machine Learning*, 34(1-3):211–231, 1999.
- [773] P. Billingsley. *Probability and Measure*. Wiley, New York, NY, 1986. Second Edition.
- [774] D. Billsus and M. J. Pazzani. Learning collaborative information filters. In *Proc. 15th International Conf. on Machine Learning*, pages 46–54. Morgan Kaufmann, San Francisco, CA, 1998.
- [775] J. Binder, D. Koller, S. Russell, and K. Kanazawa. Adaptive probabilistic networks with hidden variables. *Machine Learning*, 29(2-3):213–244, 1997.
- [776] J. C. Bioch and V. Popova. Rough sets and ordinal classification. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 291–305. Springer, 2000.
- [777] L. Birge and P. Massart. Rates of convergence for minimum contrast estimators. *Probability Theory and Related Fields*, 97:113–150, 1993.
- [778] A. Birkendorf. On fast and simple algorithms for finding maximal subarrays and applications in learning theory. In *Computational Learning Theory, Third European Conference, EuroCOLT ’97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 198–209. Springer, 1997.
- [779] A. Birkendorf, E. Dichterman, J. Jackson, N. Klasner, and H. U. Simon. On restricted-focus-of-attention learnability of Boolean functions. *Machine Learning*, 30(1):89–123, 1998. Earlier version in COLT ’96.

- [780] A. Birkendorf, N. Klasner, C. Kuhlmann, and H. U. Simon. Structural results about exact learning with unspecified attribute values. *J. Comput. Syst. Sci.*, 60(2):258–277, 2000. Special issue for COLT '98.
- [781] A. Birkendorf and H.-U. Simon. Using computational learning strategies as a tool for combinatorial optimization. *Annals of Mathematics and Artificial Intelligence*, 22(3/4):237–257, 1998.
- [782] L. Bisht, N. H. Bshouty, and L. Khoury. Learning with errors in answers to membership queries. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 611–620. IEEE Computer Society, 2004.
- [783] L. Bisht, N. H. Bshouty, and L. Khoury. Learning with errors in answers to membership queries. *J. Comput. Syst. Sci.*, 74(1):2–15, 2008.
- [784] L. Bisht, N. H. Bshouty, and H. Mazzawi. On optimal learning algorithms for multiplicity automata. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 184–198, Berlin, 2006. Springer.
- [785] J. Björklund, H. Fernau, and A. Kasprzik. MAT learning of universal automata. In A.-H. Dediu, C. Martín-Vide, and B. Truthe, editors, *Language and Automata Theory and Applications, 7th International Conference, LATA 2013, Bilbao, Spain, April 2013, Proceedings*, volume 7810 of *Lecture Notes in Computer Science*, pages 141–152. Springer, 2013.
- [786] J. Björklund, H. Fernau, and A. Kasprzik. Polynomial inference of universal automata from membership and equivalence queries. *Inform. Comput.*, 246:3–19, 2016. Special Issue: 7th International Conference on Language and Automata Theory and Applications (LATA 2013).
- [787] K. Blackmond Laskey and J. W. Myers. Population Markov chain Monte Carlo. *Machine Learning*, 50(1-2):175–196, 2003.
- [788] J. Blackmore and R. Miikkulainen. Visualizing high-dimensional structure with the incremental grid-growing neural network. In *Proc. 12th International Conference on Machine Learning*, pages 55–63. Morgan Kaufmann, 1995.
- [789] K. L. Blackmore, R. C. Williamson, I. M. Y. Mareels, and W. A. Sethares. Online learning via congregational gradient descent. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 265–272, New York, NY, 1995. ACM Press.
- [790] E. Blais, R. O’Donnell, and K. Wimmer. Polynomial regression under arbitrary product distributions. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 193–204. Omnipress, 2008.
- [791] A. Blanca, Z. Chen, D. Štefankovič, and E. Vigoda. Structure learning of  $H$ -colorings. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 152–185. PMLR, 07–09 Apr. 2018.
- [792] G. Blanchard, O. Bousquet, and L. Zwald. Statistical properties of kernel principal component analysis. *Machine Learning*, 66(2-3):259–294, 2007.
- [793] G. Blanchard and F. Fleuret. Occam’s hammer. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 112–126, Berlin, 2007. Springer.

- [794] G. Blanchard, G. Lugosi, and N. Vayatis. On the rate of convergence of regularized boosting classifiers. *Journal of Machine Learning Research*, 4:861–894, 2003.
- [795] G. Blanchard, C. Schäfer, and Y. Rozenholc. Oracle bounds and exact algorithm for dyadic classification trees. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 378–392. Springer, 2004.
- [796] G. Blanchard, C. Schäfer, Y. Rozenholc, and K.-R. Müller. Optimal dyadic decision trees. *Machine Learning*, 66(2-3):209–241, 2007.
- [797] E. Blanzieri, C. Eccher, S. Forti, and A. Sboner. Exploiting classifier combination for early melanoma diagnosis support. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 55–62. Springer, 2000.
- [798] E. Blanzieri and P. Katenkamp. Learning radial basis function networks on-line. In *Proc. 13th International Conference on Machine Learning*, pages 37–45. Morgan Kaufmann, 1996.
- [799] E. Blanzieri and F. Ricci. A minimum risk metric for nearest neighbor classification. In *Proc. 16th International Conf. on Machine Learning*, pages 22–31. Morgan Kaufmann, San Francisco, CA, 1999.
- [800] R. Blasig. Discrete sequence prediction with commented Markov models. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 191–202. Springer, 1996.
- [801] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.
- [802] H. Blockeel, L. De Raedt, and J. Ramon. Top-down induction of clustering trees. In *Proc. 15th International Conf. on Machine Learning*, pages 55–63. Morgan Kaufmann, San Francisco, CA, 1998.
- [803] H. Blockeel and J. Struyf. Efficient algorithms for decision tree cross-validation. In *Proc. 18th International Conf. on Machine Learning*, pages 11–18. Morgan Kaufmann, San Francisco, CA, 2001.
- [804] H. Blockeel and J. Struyf. Efficient algorithms for decision tree cross-validation. *Journal of Machine Learning Research*, 3:621–650, 2002.
- [805] P. Bloem, F. Mota, S. de Rooij, L. Antunes, and P. Adriaans. A safe approximation for Kolmogorov complexity. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 336–350. Berlin, Heidelberg, New York, 2014. Springer.
- [806] P. Bloem, S. de Rooij, and P. Adriaans. Two problems for sophistication. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 379–394. Berlin, Heidelberg, New York, 2015. Springer.

- [807] A. Blum. On the computational complexity of training simple neural networks. Master's thesis, MIT Department of Electrical Engineering and Computer Science, May 1989. (Published as Laboratory for Computer Science Technical Report MIT/LCS/TR-445 (May, 1989)).
- [808] A. Blum. Separating distribution-free and mistake-bound learning models over the Boolean domain. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 211–218. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [809] A. Blum. Separating PAC and mistake-bound learning models over the Boolean domain. *Proc. 31th Annu. IEEE Sympos. Found. Comput. Sci.*, 1990.
- [810] A. Blum. Learning Boolean functions in an infinite attribute space. *Machine Learning*, 9(4):373–386, Oct. 1992.
- [811] A. Blum. Empirical support for Winnow and weighted-majority based algorithms: results on a calendar scheduling domain. In *Proc. 12th International Conference on Machine Learning*, pages 64–72, San Francisco, CA., 1995. Morgan Kaufmann.
- [812] A. Blum. Empirical support for Winnow and weighted-majority algorithms: Results on a calendar scheduling domain. *Machine Learning*, 26(1):5–23, 1997.
- [813] A. Blum. Learning a function of  $r$  relevant variables. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 731–733. Springer, 2003.
- [814] A. Blum. A theory of similarity functions for learning and clustering. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, page 9, Berlin, Oct. 2007. Springer.
- [815] A. Blum and M.-F. Balcan. Open problems in efficient semi-supervised PAC learning. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 622–624, Berlin, 2007. Springer.
- [816] A. Blum and C. Burch. On-line learning and the metrical task system problem. *Machine Learning*, 39(1):35–58, 2000.
- [817] A. Blum and P. Chalasani. Learning switching concepts. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 231–242, New York, NY, 1992. ACM Press.
- [818] A. Blum and P. Chalasani. An on-line algorithm for improving performance in navigation. In *Proc. 34th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 2–11. IEEE Computer Society Press, Los Alamitos, CA, 1993.
- [819] A. Blum, P. Chalasani, S. A. Goldman, and D. K. Slonim. Learning with unreliable boundary queries. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 98–107, New York, NY, 1995. ACM Press.
- [820] A. Blum, P. Chalasani, S. A. Goldman, and D. K. Slonim. Learning with unreliable boundary queries. *J. of Comput. Syst. Sci.*, 56(2):209–222, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.

- [821] A. Blum, P. Chalasani, and J. Jackson. On learning embedded symmetric concepts. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 337–346, New York, NY, 1993. ACM Press.
- [822] A. Blum and S. Chawla. Learning from labeled and unlabeled data using graph mincuts. In *Proc. 18th International Conf. on Machine Learning*, pages 19–26. Morgan Kaufmann, San Francisco, CA, 2001.
- [823] A. Blum, A. M. Frieze, R. Kannan, and S. Vempala. A polynomial-time algorithm for learning noisy linear threshold functions. *Algorithmica*, 22(1/2):35–52, 1998.
- [824] A. Blum, M. Furst, J. Jackson, M. Kearns, Y. Mansour, and S. Rudich. Weakly learning DNF and characterizing statistical query learning using fourier analysis. In *Proc. of Twenty-sixth ACM Symposium on Theory of Computing*, 1994.
- [825] A. Blum, L. Hellerstein, and N. Littlestone. Learning in the presence of finitely or infinitely many irrelevant attributes. *J. of Comput. Syst. Sci.*, 50(1):32–40, 1995. Earlier version in 4th COLT, 1991.
- [826] A. Blum, J. Jackson, T. Sandholm, and M. Zinkevich. Preference elicitation and query learning. *Journal of Machine Learning Research*, 5:649–667, 2004.
- [827] A. Blum, J. C. Jackson, T. Sandholm, and M. Zinkevich. Preference elicitation and query and learning. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 13–25. Springer, 2003.
- [828] A. Blum and A. Kalai. A note on learning from multiple-instance examples. *Machine Learning*, 30(1):23–29, 1998.
- [829] A. Blum and A. Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999. Special issue for COLT '97.
- [830] A. Blum, A. Kalai, and J. Langford. Beating the hold-out: bounds for k-fold and progressive cross-validation. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 203–208, New York, NY, 1999. ACM Press.
- [831] A. Blum, A. Kalai, and H. Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 435–440. ACM Press, 2000.
- [832] A. Blum, A. Kalai, and H. Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *Journal of the ACM*, 50(4):506–519, 2003.
- [833] A. Blum and R. Kannan. Learning an intersection of a constant number of halfspaces over a uniform distribution. *J. Comput. Syst. Sci.*, 54(2):371–380, 1997. First appeared in FOCS '93.
- [834] A. Blum and M. Kearns, editors. *Proceedings of the Ninth Annual Conference on Computational Learning Theory, June 28th-July 1st, 1996, Desenzano del Garda, Italy*, New York, N.Y., 1996. ACM Press.

- [835] A. Blum, R. Khairon, E. Kushilevitz, L. Pitt, and D. Roth. On learning read- $k$ -satisfy- $j$  DNF. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 110–117, New York, NY, 1994. ACM Press.
- [836] A. Blum and J. Langford. PAC-MDL bounds. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 344–357. Springer, 2003.
- [837] A. Blum, G. Lugosi, and H. U. Simon. Introduction to the special issue on COLT 2006. *Machine Learning*, 69(2-3):75–77, 2007.
- [838] A. Blum and Y. Mansour. From external to internal regret. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 621–636. Springer, 2005.
- [839] A. Blum and Y. Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8:1307–1324, 2007.
- [840] A. Blum and Y. Mansour. Efficient co-training of linear separators under weak dependence. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 302–318, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [841] A. Blum and T. Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 92–100, New York, NY, 1998. ACM Press.
- [842] A. Blum, P. Raghavan, and B. Schieber. Navigating in unfamiliar geometric terrain. In *Proc. 23th Annu. ACM Sympos. Theory Comput.*, pages 494–504. ACM, 1991.
- [843] A. Blum and R. L. Rivest. Training a 3-node neural net is NP-Complete. In *Advances in Neural Information Processing Systems I*, pages 494–501. Morgan Kaufmann, 1989.
- [844] A. Blum and S. Rudich. Fast learning of  $k$ -term DNF formulas with queries. In *Proc. of Twenty-Fourth ACM Symposium on Theory of Computing*, pages 382–389. ACM, 1992.
- [845] A. Blum and S. Rudich. Fast learning of  $k$ -term DNF formulas with queries. *J. of Comput. Syst. Sci.*, 51(3):367–373, 1995.
- [846] A. Blum and M. Singh. Learning functions of  $k$  terms. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 144–153, San Mateo, CA, 1990. Morgan Kaufmann.
- [847] E. K. Blum. Approximation of Boolean functions by sigmoidal networks: Part I: XOR and other two-variable functions. *Neural Computation*, 1:532–540, Feb. 1989.
- [848] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Inform. Control*, 28(2):125–155, June 1975.
- [849] M. Blum. Program error detection/correction: Turning PAC learning into PERFECT learning. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 1–1. Springer, 1997.

- [850] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Occam’s razor. *Inform. Proc. Lett.*, 24:377–380, Apr. 1987.
- [851] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. ACM*, 36(4):929–965, 1989.
- [852] A. Blumer and N. Littlestone. Learning faster than promised by the Vapnik-Chervonenkis dimension. *Discrete Applied Mathematics*, 24:47–53, 1989.
- [853] R. Board and L. Pitt. On the necessity of Occam algorithms. *Theoret. Comput. Sci.*, 100(1):157–184, 1992.
- [854] R. A. Board and L. Pitt. Semi-supervised learning. *Machine Learning*, 4(1):41–65, 1989.
- [855] M. Bohanec and I. Bratko. Trading accuracy for simplicity in decision trees. *Machine Learning*, 15(3):223–250, 1994.
- [856] M. Boicu, G. Tecuci, D. Marcu, M. Bowman, P. Shyr, F. Ciucu, and C. Levcovici. Disciple-COA: From agent programming to agent teaching. In *Proc. 17th International Conf. on Machine Learning*, pages 73–80. Morgan Kaufmann, San Francisco, CA, 2000.
- [857] A. Boiret, A. Lemay, and J. Niehren. Learning top-down tree transducers with regular domain inspection. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 54–65, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [858] K. D. Bollacker and J. Ghosh. A supra-classifier architecture for scalable knowledge reuse. In *Proc. 15th International Conf. on Machine Learning*, pages 64–72. Morgan Kaufmann, San Francisco, CA, 1998.
- [859] L. B. Bonache, E. Fromont, A. Habrard, M. Perrot, and M. Sebban. Speeding up syntactic learning using contextual information. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 49–53, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [860] D. Boneh. Learning using group representations. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 418–426, New York, NY, 1995. ACM Press.
- [861] D. Boneh and R. Lipton. Amplification of weak learning under the uniform distribution. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 347–351, New York, NY, 1993. ACM Press.
- [862] B. Bonet and H. Geffner. Learning sorting and decision trees with POMDPs. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML ’98)*, pages 73–81. Morgan Kaufmann, San Francisco, CA, 1998.
- [863] M. Bonnell Harries, C. Sammut, and K. Horn. Extracting hidden context. *Machine Learning*, 32(2):101–126, 1998.
- [864] G. Bontempi, M. Birattari, and H. Bersini. Local learning for iterated time series prediction. In *Proc. 16th International Conf. on Machine Learning*, pages 32–38. Morgan Kaufmann, San Francisco, CA, 1999.
- [865] L. B. Booker. Classifier systems that learn internal world models. *Machine Learning*, 3:161–192, 1988.



- [866] J. H. Boose and B. R. Gaines. Knowledge acquisition for knowledge-based systems: Notes on the state-of-the-art. *Machine Learning*, 4:377–394, 1989.
- [867] E. Boros, Y. Crama, and P. L. Hammer. Polynomial-time inference of all valid implications for Horn and related formulae. *Annals of Mathematics and Artificial Intelligence*, 1:21–32, 1990.
- [868] E. Boros, T. Ibaraki, and K. Makino. Monotone extensions of Boolean data sets. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 161–175. Springer, 1997.
- [869] A. van den Bosch. Instance-family abstraction in memory-based language learning. In *Proc. 16th International Conf. on Machine Learning*, pages 39–48. Morgan Kaufmann, San Francisco, CA, 1999.
- [870] A. van den Bosch and J. Zavrel. Unpacking multi-valued symbolic features and classes in memory-based language learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1055–1062. Morgan Kaufmann, San Francisco, CA, 2000.
- [871] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 144–152, New York, NY, 1992. ACM Press.
- [872] H. Boström. Improving example-guided unfolding. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 124–135. Springer-Verlag, 1993.
- [873] H. Boström. Theory-guided induction of logic programs by inference of regular languages. In *Proc. 13th International Conference on Machine Learning*, pages 46–53. Morgan Kaufmann, 1996.
- [874] M. Botta, A. Giordana, and R. Piola. FONN: Combining first order logic with connectionist learning. In *Proc. 14th International Conference on Machine Learning*, pages 48–56. Morgan Kaufmann, 1997.
- [875] M. Botta, A. Giordana, L. Saitta, and M. Sebag. Relational learning as search in a critical region. *Journal of Machine Learning Research*, 4:431–463, 2003.
- [876] M. Botta and R. Piola. Refining numerical constants in first order logic theories. *Machine Learning*, 38(1/2):109–131, 2000.
- [877] S. Boucheron and J. Sallantin. Some remarks about space-complexity of learning, and circuit complexity of recognizing. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 125–138, San Mateo, CA, 1988. Morgan Kaufmann.
- [878] M. Boullé. MODL: A Bayes optimal discretization method for continuous attributes. *Machine Learning*, 65(1):131–165, 2006.
- [879] S. Bounkong, B. Toch, D. Saad, and D. Lowe. ICA for watermarking digital images. *Journal of Machine Learning Research*, 4:1471–1498, 2003.
- [880] A. E. O. Bousquet. Stability and generalization. *Journal of Machine Learning Research*, 2:499–526, 2002.

- [881] O. Bousquet and A. Elisseeff. Guest editorial: Learning theory. *Machine Learning*, 66(2-3):115–118, 2007.
- [882] O. Bousquet, V. Koltchinskii, and D. Panchenko. Some local measures of complexity of convex hulls and generalization bounds. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 59–73. Springer, 2002.
- [883] O. Bousquet and M. K. Warmuth. Tracking a small set of experts by mixing past posteriors. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 31–47. Springer, 2001.
- [884] O. Bousquet and M. K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3:363–396, 2002.
- [885] C. Boutilier and R. Dearden. Approximating value trees in structured dynamic programming. In *Proc. 13th International Conference on Machine Learning*, pages 54–62. Morgan Kaufmann, 1996.
- [886] A. F. Bowers, C. Giraud-Carrier, and J. W. Lloyd. Classification of individuals with complex structure. In *Proc. 17th International Conf. on Machine Learning*, pages 81–88. Morgan Kaufmann, San Francisco, CA, 2000.
- [887] M. Bowling. Convergence problems of general-sum multiagent reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 89–94. Morgan Kaufmann, San Francisco, CA, 2000.
- [888] M. Bowling and M. Veloso. Convergence of gradient dynamics with a variable learning rate. In *Proc. 18th International Conf. on Machine Learning*, pages 27–34. Morgan Kaufmann, San Francisco, CA, 2001.
- [889] J. Boyan and A. W. Moore. Learning evaluation functions to improve optimization by local search. *Journal of Machine Learning Research*, 1:77–112, 2000.
- [890] J. A. Boyan. Least-squares temporal difference learning. In *Proc. 16th International Conf. on Machine Learning*, pages 49–56. Morgan Kaufmann, San Francisco, CA, 1999.
- [891] J. A. Boyan. Technical update: Least-squares temporal difference learning. *Machine Learning*, 49(2-3):233–246, 2002.
- [892] J. A. Boyan and A. W. Moore. Learning evaluation functions for large acyclic domains. In *Proc. 13th International Conference on Machine Learning*, pages 63–70. Morgan Kaufmann, 1996.
- [893] S. Boykin and A. Merlino. Machine learning of event segmentation for news on demand. *Communications of the ACM*, 43(2):35–41, 2000.
- [894] J. P. Bradford and C. E. Brodley. The effect of instance-space partition on significance. *Machine Learning*, 42(3):269–286, 2001.
- [895] P. S. Bradley and U. M. Fayyad. Refining initial points for K-Means clustering. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 91–99. Morgan Kaufmann, San Francisco, CA, 1998.

- [896] P. S. Bradley and O. L. Mangasarian. Feature selection via concave minimization and support vector machines. In *Proc. 15th International Conf. on Machine Learning*, pages 82–90. Morgan Kaufmann, San Francisco, CA, 1998.
- [897] G. Bradshaw. Invention and artificial intelligence. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 10–10. Springer, Oct. 2005.
- [898] S. J. Bradtke and A. G. Barto. Linear least-squares algorithms for temporal difference learning. *Machine Learning*, 22(1/2/3):33–57, 1996.
- [899] M. L. Brady, R. Raghavan, and J. Slawny. Back propagation fails to separate where perceptrons succeed. *IEEE Transactions On Circuits and Systems*, 36(5):665–674, May 1989.
- [900] D. Braess, J. Forster, T. Sauer, and H. U. Simon. How to achieve minimax expected Kullback-Leibler distance from an unknown finite distribution. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 380–394. Springer, 2002.
- [901] R. I. Brafman and M. Tennenholtz. R-MAX - A general polynomial time algorithm for near-optimal reinforcement learning. *Journal of Machine Learning Research*, 3:213–231, 2002.
- [902] M. Brand. Finding variational structure in data by cross-entropy optimization. In *Proc. 17th International Conf. on Machine Learning*, pages 95–101. Morgan Kaufmann, San Francisco, CA, 2000.
- [903] Y. Brandman, J. Hennessy, and A. Orlitsky. A spectral lower bound technique for the size of decision trees and two level circuits. *IEEE Trans. Computers*, 39(2):282–287, 1990.
- [904] U. Brandt. The position of index sets of identifiable sets in the arithmetical hierarchy. *Inform. Control*, 68(1-3):185–195, 1986.
- [905] I. Bratko. Machine learning and qualitative reasoning. *Machine Learning*, 14(3):305–312, 1994.
- [906] I. Bratko. Discovery of abstract concepts by a robot. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, page 32, Berlin, Heidelberg, New York, 2010. Springer.
- [907] I. Bratko and S. Muggleton. Applications of inductive logic programming. *Commun. ACM*, 38(11):65–70, 1995.
- [908] J. M. Brayer and K. S. Fu. A note on the k-tail method of tree grammar inference. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-7:293–300, 1977.
- [909] P. Brazdil. Learning in multi-agent environments. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 5–34. IOS Press, 1992.
- [910] P. Brazdil and C. Soares. A comparison of ranking methods for classification algorithm selection. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 63–74. Springer, 2000.

- [911] P. B. Brazdil, C. Soares, and J. Pinto da Costa. Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. *Machine Learning*, 50(3):251–277, 2003.
- [912] A. Brāzma. Inductive synthesis of dot expressions. In *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 156–212. Springer, Berlin, 1991.
- [913] A. Brāzma. Efficient identification of regular expressions from representative examples. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 236–242, New York, NY, 1993. ACM Press.
- [914] A. Brāzma. Efficient algorithm for learning simple regular expressions from noisy examples. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 260–271. Springer-Verlag, 1994.
- [915] A. Brāzma. Learning of regular expressions by pattern matching. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 392–403. Springer, 1995.
- [916] A. Brāzma. Efficient learning of regular expressions from approximate examples. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 19, pages 337–352. MIT Press, 1997.
- [917] A. Brāzma and Čerāns. Noise-tolerant efficient inductive synthesis of regular expressions from good examples. *New Generation Computing*, 15(1):105–140, 1997. Special issue for AII '94 and ALT '94.
- [918] A. Brāzma and K. Čerāns. Efficient learning of regular expressions from good examples. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 76–90. Springer-Verlag, 1994.
- [919] A. Brāzma, I. Jonassen, J. Vilo, and E. Ukkonen. Pattern discovery in biosequences. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 257–270. Springer, 1998.
- [920] A. Brāzma, E. Ukkonen, and J. Vilo. Discovering unbounded unions of regular pattern languages from positive examples (extended abstract). In *Algorithms and Computation, 7th International Symposium, ISAAC '96, Osaka, Japan, December 16-18, 1996, Proceedings*, volume 1178 of *Lecture Notes in Computer Science*, pages 95–104. Springer, 1996.
- [921] M. de Brecht, M. Kobayashi, H. Tokunaga, and A. Yamamoto. Inferability of closed set systems from positive data. In *New Frontiers in Artificial Intelligence, JSAI 2006 Conference and Workshops, Tokyo, Japan, June 2006, Revised Selected Papers*, volume 4384 of *Lecture Notes in Artificial Intelligence*, pages 265–275, Berlin, 2007. Springer.
- [922] M. de Brecht and A. Yamamoto. Mind change complexity of inferring unbounded unions of pattern languages from positive data. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 124–138. Springer, Oct. 2006.

- [923] M. de Brecht and A. Yamamoto. Topological properties of concept spaces. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 374–388, Berlin, Oct. 2008. Springer.
- [924] M. de Brecht and A. Yamamoto. Mind change complexity of inferring unbounded unions of restricted pattern languages from positive data. *Theoret. Comput. Sci.*, 411(7-9):976–985, 2010.
- [925] M. de Brecht and A. Yamamoto. Topological properties of concept spaces (full version). *Inform. Comput.*, 208(4):327–340, 2010.
- [926] L. Bréhélin, O. Gascuel, and G. Caraux. Hidden Markov models with patterns and their application to integrated circuit testing. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 75–87. Springer, 2000.
- [927] L. Breiman. Stacked regressions. *Machine Learning*, 24:49–64, 1996.
- [928] L. Breiman. Technical note: Some properties of splitting criteria. *Machine Learning*, 24:41–47, 1996.
- [929] L. Breiman. Pasting small votes for classification in large databases and on-line. *Machine Learning*, 36(1/2):85–103, 1999.
- [930] L. Breiman. Randomizing outputs to increase prediction accuracy. *Machine Learning*, 40(3):229–242, 2000.
- [931] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [932] L. Breiman. Using iterated bagging to Debias regressions. *Machine Learning*, 45(3):261–277, 2001.
- [933] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Wadsworth International Group, 1984.
- [934] M. R. Brent. An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, 34:71–105, 1999.
- [935] M. R. Brent and T. A. Cartwright. Lexical categorization: fitting template grammars by incremental MDL optimization. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 84–94. Springer, 1996.
- [936] G. Brewka. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge, 1991.
- [937] P. Brézelle and H. Soldano. SAMIA: A bottom-up learning method using a simulated annealing algorithm. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 297–312. Springer-Verlag, 1993.
- [938] T. Brijs, G. Swinnen, K. Vanhoof, and G. Wets. Comparing complete and partial classification for identifying latently dissatisfied customers. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June*

- 2, 2000, *Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 88–95. Springer, 2000.
- [939] M. Brodie and G. DeJong. Learning to ride a bicycle using iterated phantom induction. In *Proc. 16th International Conf. on Machine Learning*, pages 57–66. Morgan Kaufmann, San Francisco, CA, 1999.
- [940] M. Brodie and G. DeJong. Iterated phantom induction: A knowledge-based approach to learning control. *Machine Learning*, 45(1):45–76, 2001.
- [941] C. E. Brodley. Automatic selection of split criterion during tree growing based on node location. In *Proc. 12th International Conference on Machine Learning*, pages 73–80. Morgan Kaufmann, 1995.
- [942] C. E. Brodley. Recursive automatic bias selection for classifier construction. *Machine Learning*, 20(1/2):63–94, 1995.
- [943] C. E. Brodley and A. P. Danyluk. Special issue on the eighteenth international conference on machine learning (ICML2001). *Journal of Machine Learning Research*, 3:619–620, 2002.
- [944] C. E. Brodley and M. A. Friedl. Identifying mislabeled training data. *Journal of Artificial Intelligence Research*, 11:131–167, 1999.
- [945] C. E. Brodley and P. E. Utgoff. Multivariate decision trees. *Machine Learning*, 19(1):45–77, 1995.
- [946] D. Brooks and M. Lee. Learning syntax from function words. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 273–274. Springer, 2004.
- [947] R. Brooks. The role of learning in autonomous robots. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 5–10, San Mateo, CA, 1991. Morgan Kaufmann.
- [948] N. Brosse, A. Durmus, É. Moulines, and M. Pereyra. Sampling from a log-concave distribution with compact support with proximal langevin monte carlo. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 319–342, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [949] M. P. Brown, R. Hughey, A. Krogh, I. S. Mian, K. Sjölander, and D. Haussler. Using Dirichlet mixture priors to derive hidden Markov models for protein families. In L. Hunter, D. Searls, and J. Shavlik, editors, *Proc. of First Int. Conf. on Intelligent Systems for Molecular Biology*, pages 47–55, Menlo Park, CA, July 1993. AAAI/MIT Press.
- [950] V.-E. Brunel, A. Moitra, P. Rigollet, and J. Urschel. Rates of estimation for determinantal point processes. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 343–345, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [951] C. Brunk and M. Pazzani. A lexically based semantic bias for theory revision. In *Proc. 12th International Conference on Machine Learning*, pages 81–89. Morgan Kaufmann, 1995.

- [952] J. D. Brutlag and C. Meek. Challenges of the email domain for text classification. In *Proc. 17th International Conf. on Machine Learning*, pages 103–110. Morgan Kaufmann, San Francisco, CA, 2000.
- [953] M. Bruynooghe and D. De Schreye. Some thoughts on the role of examples in program transformation and its relevance for explanation-based learning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 60–77. Springer-Verlag, 1989.
- [954] B. R. Bryant, M. Mernik, D. Hrnčič, F. Javed, Q. Liu, and A. Sprague. Grammar inference technology applications in software engineering. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 276–279. Springer, 2010.
- [955] D. Bshouty and N. H. Bshouty. On learning arithmetic read-once formulas with exponentiation. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 311–317, New York, NY, 1994. ACM Press.
- [956] N. Bshouty. On learning discretized geometric concepts. In *Proc. of the 35rd Symposium on the Foundations of Comp. Sci.* IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [957] N. Bshouty, T. Hancock, L. Hellerstein, and M. Karpinski. An algorithm to learn read-once threshold formulas, and transformations between learning models. *Computational Complexity*, 4:37–61, 1994.
- [958] N. Bshouty and L. Hellerstein. Attribute-efficient learning in query and mistake-bound models. *J. of Comput. Syst. Sci.*, 56(3):310–319, 1998.
- [959] N. Bshouty and C. Tamon. On the Fourier spectrum of monotone functions. In *Proc. 27th Annual ACM Symposium on Theory of Computing*, pages 219–228, New York, NY, 1995. ACM Press.
- [960] N. H. Bshouty. Exact learning via the monotone theory. In *Proc. of the 34rd Annual Symposium on Foundations of Computer Science*, pages 302–311. IEEE Computer Society Press, Los Alamitos, CA, 1993.
- [961] N. H. Bshouty. A note on learning multivariate polynomials under the uniform distribution. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 79–82, New York, NY, 1995. ACM Press.
- [962] N. H. Bshouty. Simple learning algorithms using divide and conquer. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 447–453, New York, NY, 1995. ACM Press.
- [963] N. H. Bshouty. Exact learning of formulas in parallel. *Machine Learning*, 26(1):25–41, 1997.
- [964] N. H. Bshouty. The monotone theory for the PAC-model. *Inform. Comput.*, 186(1):20–35, 2003.
- [965] N. H. Bshouty. Polynomial time prediction strategy with almost optimal mistake probability. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 64–76. Springer, 2004.

- [966] N. H. Bshouty. Exact learning from membership queries: Some techniques, results and new directions. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 33–52, Berlin, Heidelberg, New York, 2013. Springer.
- [967] N. H. Bshouty. Exact learning from an honest teacher that answers membership queries. *Theoret. Comput. Sci.*, 733:4–43, 2018. Special Issue on Learning Theory and Complexity.
- [968] N. H. Bshouty and I. Bentov. On exact learning from random walk. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 184–198. Springer, Oct. 2006.
- [969] N. H. Bshouty, V. E. Bshouty-Hurani, G. Haddad, T. Hashem, F. Khoury, and O. Sharafy. Adaptive group testing algorithms to estimate the number of defectives. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 93–110. PMLR, 07–09 Apr. 2018.
- [970] N. H. Bshouty and L. Burroughs. Bounds for the minimum disagreement problem with applications to learning theory. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 271–286. Springer, 2002.
- [971] N. H. Bshouty and L. Burroughs. Maximizing agreements and coagnostic learning. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 83–97. Springer, 2002.
- [972] N. H. Bshouty and L. Burroughs. On the proper learning of axis parallel concepts. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 287–302. Springer, 2002.
- [973] N. H. Bshouty and L. Burroughs. On the proper learning of axis-parallel concepts. *Journal of Machine Learning Research*, 4:157–176, 2003.
- [974] N. H. Bshouty and L. Burroughs. Maximizing agreements and coagnostic learning. *Theoret. Comput. Sci.*, 350(1):24–39, 2006. Special issue for ALT 2002.
- [975] N. H. Bshouty, Z. Chen, S. E. Decatur, and S. Homer. On the learnability of  $Z_N$ -DNF formulas. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 198–205, New York, NY, 1995. ACM Press.
- [976] N. H. Bshouty, Z. Chen, and S. Homer. On learning discretized geometric concepts. In *Proc. of the 35rd Annual Symposium on Foundations of Computer Science*, pages 54–63. IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [977] N. H. Bshouty and R. Cleve. On the exact learning of formulas in parallel. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 513–522. IEEE Computer Society Press, Los Alamitos, CA, 1992.
- [978] N. H. Bshouty, R. Cleve, R. Gavaldà, S. Kannan, and C. Tamon. Oracles and queries that are sufficient for exact learning. *J. of Comput. Syst. Sci.*, 52(3):421–433, 1996.



- [979] N. H. Bshouty, R. Cleve, S. Kannan, and C. Tamon. Oracles and queries that are sufficient for exact learning. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 130–139, New York, NY, 1994. ACM Press.
- [980] N. H. Bshouty and A. Costa. Exact learning of juntas from membership queries. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 115–129. Springer, 2016.
- [981] N. H. Bshouty, N. Diab, S. R. Kawar, and R. J. Shahla. Non-adaptive randomized algorithm for group testing. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 109–128, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [982] N. H. Bshouty, D. Drachslor-Cohen, M. Vechev, and E. Yahav. Learning disjunctions of predicates. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 346–369, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [983] N. H. Bshouty and N. Eiron. Learning monotone DNF from a teacher that almost does not answer membership queries. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 546–557. Springer, 2001.
- [984] N. H. Bshouty and N. Eiron. Learning monotone DNF from a teacher that almost does not answer membership queries. *Journal of Machine Learning Research*, 3:49–57, 2002.
- [985] N. H. Bshouty, N. Eiron, and E. Kushilevitz. PAC learning with nasty noise. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 206–218. Springer, 1999.
- [986] N. H. Bshouty, N. Eiron, and E. Kushilevitz. PAC learning with nasty noise. *Theoret. Comput. Sci.*, 288(2):255–275, 2002. Special issue for ALT '99.
- [987] N. H. Bshouty and V. Feldman. On using extended statistical queries to avoid membership queries. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 529–545. Springer, 2001.
- [988] N. H. Bshouty and V. Feldman. On using extended statistical queries to avoid membership queries. *Journal of Machine Learning Research*, 2:359–395, 2002.
- [989] N. H. Bshouty and D. Gavinsky. On boosting with optimal poly-bounded distributions. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 490–506. Springer, 2001.
- [990] N. H. Bshouty and D. Gavinsky. On boosting with polynomially bounded distributions. *Journal of Machine Learning Research*, 3:483–506, 2002.

- [991] N. H. Bshouty and D. Gavinsky.  $\text{pac} = \text{paexact}$  and other equivalent models in learning. In *43rd Annual Symposium on Foundations of Computer Science, 6-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 167–176. IEEE Computer Society, 2002.
- [992] N. H. Bshouty and C. Gentile, editors. *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, Berlin, 2007. Springer.
- [993] N. H. Bshouty, P. W. Goldberg, S. A. Goldman, and H. D. Mathias. Exact learning of discretized geometric concepts. *SIAM J. Comput.*, 28(2):674–699, 1998.
- [994] N. H. Bshouty, S. A. Goldman, T. R. Hancock, and S. Matar. Asking questions to minimize errors. *J. of Comput. Syst. Sci.*, 52(2):268–286, 1996. Earlier version in 6th COLT, 1993.
- [995] N. H. Bshouty, S. A. Goldman, and D. H. Mathias. Noise-tolerant parallel learning of geometric concepts. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 345–352, New York, NY, 1995. ACM Press.
- [996] N. H. Bshouty, S. A. Goldman, H. D. Mathias, S. Suri, and H. Tamaki. Noise-tolerant distribution-free learning of general geometric concepts. In *Proc. 28th Annu. ACM Symposium on the Theory of Computing*, pages 151–160, New York, NY, 1996. ACM Press.
- [997] N. H. Bshouty, S. A. Goldman, H. D. Mathias, S. Suri, and H. Tamaki. Noise-tolerant distribution-free learning of general geometric concepts. *J. ACM*, 45(5):863–890, 1998.
- [998] N. H. Bshouty, T. R. Hancock, and L. Hellerstein. Learning arithmetic read-once formulas. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 370–381, New York, NY, 1992. ACM Press.
- [999] N. H. Bshouty, T. R. Hancock, and L. Hellerstein. Learning Boolean read-once formulas with arbitrary symmetric and constant fan-in gates. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 1–15, New York, NY, 1992. ACM Press.
- [1000] N. H. Bshouty and L. Hellerstein. Attribute-efficient learning in query and mistake-bound models. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 235–243, New York, NY, 1996. ACM Press.
- [1001] N. H. Bshouty and J. C. Jackson. Learning DNF over the uniform distribution using a quantum example oracle. *SIAM J. Comput.*, 28(3):1136–1153, 1999.
- [1002] N. H. Bshouty, J. C. Jackson, and C. Tamon. More efficient PAC-learning of DNF with membership queries under the uniform distribution. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 286–295, New York, NY, 1999. ACM Press.
- [1003] N. H. Bshouty, J. C. Jackson, and C. Tamon. Uniform-distribution attribute noise learnability. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 75–80, New York, NY, 1999. ACM Press.
- [1004] N. H. Bshouty, J. C. Jackson, and C. Tamon. Exploring learnability between Exact and PAC. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 244–254. Springer, 2002.

- [1005] N. H. Bshouty, J. C. Jackson, and C. Tamon. More efficient PAC-learning of DNF with membership queries under the uniform distribution. *J. Comput. Syst. Sci.*, 68(1):205–234, 2004.
- [1006] N. H. Bshouty, J. C. Jackson, and C. Tamon. Uniform-distribution attribute noise learnability. *Inform. Comput.*, 187(2):277–290, 2004.
- [1007] N. H. Bshouty, J. C. Jackson, and C. Tamon. Exploring learnability between exact and PAC. *J. Comput. Syst. Sci.*, 70(4):471–484, 2005. Special issue on COLT 2002.
- [1008] N. H. Bshouty and Y. Mansour. Simple learning algorithms for decision trees and multivariate polynomials. In *Proc. of the 36th Annual Symposium on Foundations of Computer Science*, pages 304–311. IEEE Computer Society Press, Los Alamitos, CA, 1995.
- [1009] N. H. Bshouty and H. Mazzawi. Exact learning composed classes with a small number of mistakes. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 199–213, Berlin, 2006. Springer.
- [1010] N. H. Bshouty and H. Mazzawi. Reconstructing weighted graphs with minimal query complexity. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 97–109, Berlin/Heidelberg, 2009. Springer.
- [1011] N. H. Bshouty and H. Mazzawi. Reconstructing weighted graphs with minimal query complexity. *Theoret. Comput. Sci.*, 412(19):1782–1790, 2011.
- [1012] N. H. Bshouty, E. Mossel, R. O’Donnell, and R. A. Servedio. Learning DNF from random walks. In *44th Annual Symposium on Foundations of Computer Science, 1-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 189–198. IEEE Computer Society, 2003.
- [1013] N. H. Bshouty, E. Mossel, R. O’Donnell, and R. A. Servedio. Learning DNF from random walks. *J. Comput. Syst. Sci.*, 71(3):250–265, 2005. Special issue Learning Theory 2003.
- [1014] N. H. Bshouty and A. Owshanko. Learning regular sets with an incomplete membership oracle. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 574–588. Springer, 2001.
- [1015] N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors. *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, Berlin, Heidelberg, New York, Oct. 2012. Springer.
- [1016] N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann. Editors’ introduction. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 1–11, Berlin, Heidelberg, New York, 2012. Springer.
- [1017] N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann. Guest Editors’ foreword. *Theoret. Comput. Sci.*, 558:1–4, 2014. Special Issue: Algorithmic Learning Theory.

- [1018] N. H. Bshouty, C. Tamon, and D. K. Wilson. On learning decision trees with large output domains. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 190–197, New York, NY, 1995. ACM Press.
- [1019] N. H. Bshouty, C. Tamon, and D. K. Wilson. On learning width two branching programs. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 224–227, New York, NY, 1996. ACM Press.
- [1020] N. H. Bshouty, C. Tamon, and D. K. Wilson. Learning matrix functions over rings. *Algorithmica*, 22(1/2):91–111, 1998.
- [1021] N. H. Bshouty and E. Wattad. On exact learning halfspaces with random consistent hypothesis oracle. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 48–62. Springer, Oct. 2006.
- [1022] N. H. Bshouty and D. K. Wilson. On learning in the presence of unspecified attribute values. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 81–87, New York, NY, 1999. ACM Press.
- [1023] S. Bubeck, N. Cesa-Bianchi, and S. M. Kakade. Towards minimax policies for online linear optimization with bandit feedback. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 41.1–41.14, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1024] S. Bubeck, M. Cohen, and Y. Li. Sparsity, variance and curvature in multi-armed bandits. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 111–127. PMLR, 07–09 Apr. 2018.
- [1025] S. Bubeck, O. Dekel, T. Koren, and Y. Peres. Bandit convex optimization:  $\sqrt{T}$  regret in one dimension. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 266–278, Paris, France, 03–06 July 2015. PMLR.
- [1026] S. Bubeck and R. Eldan. The entropic barrier: a simple and optimal universal self-concordant barrier. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 279–279, Paris, France, 03–06 July 2015. PMLR.
- [1027] S. Bubeck and R. Eldan. Multi-scale exploration of convex functions and bandit convex optimization. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 583–589, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1028] S. Bubeck, R. Munos, and G. Stoltz. Pure exploration in multi-armed bandits problems. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 23–37, Berlin/Heidelberg, 2009. Springer.
- [1029] S. Bubeck, R. Munos, and G. Stoltz. Pure exploration in finitely-armed and continuous-armed bandits. *Theoret. Comput. Sci.*, 412(19):1832–1852, 2011.

- [1030] S. Bubeck, V. Perchet, and P. Rigollet. Bounded regret in stochastic multi-armed bandits. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 122–134. JMLR.org, 2013.
- [1031] S. Bubeck and A. Slivkins. The best of both worlds: Stochastic and adversarial bandits. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 42.1–42.23, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1032] S. Bubeck, G. Stoltz, and J. Y. Yu. Lipschitz bandits without the Lipschitz constant. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 144–158, Berlin, Heidelberg, New York, 2011. Springer.
- [1033] B. G. Buchanan. Can machine learning offer anything to expert systems? *Machine Learning*, 4:251–254, 1989.
- [1034] N. Buchbinder, S. Chen, J. S. Naor, and O. Shamir. Unified algorithms for online learning and competitive analysis. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 5.1–5.18, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1035] C. Buckley, G. Salton, and J. Allan. The effect of adding relevance information in a relevance feedback environment. In *Proc. 17th Annu. Internat. ACM-SIGIR conference on Research and Development in Information Retrieval*, pages 292–301. Springer-Verlag, 1994.
- [1036] K. Buescher and P. R. Kumar. Simultaneous learning of concepts and simultaneous estimation of probabilities. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 33–42, San Mateo, CA, 1991. Morgan Kaufmann.
- [1037] K. L. Buescher and P. R. Kumar. Learning stochastic functions by smooth simultaneous estimation. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 272–279, New York, NY, 1992. ACM Press.
- [1038] A. Bulatov, H. Chen, and V. Dalmau. Learnability of relatively quantified generalized formulas. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 365–379. Springer, 2004.
- [1039] A. Bulatov, H. Chen, and V. Dalmau. Learning intersection-closed classes with signatures. *Theoret. Comput. Sci.*, 382(3):209–220, 2007. Special issue for ALT 2004.
- [1040] W. J. Bultman and W. Maass. Fast identification of geometric objects with membership queries. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 337–353, San Mateo, CA, 1991. Morgan Kaufmann.
- [1041] A. Bundy and B. Silver. A critical survey of rule learning programs. Technical report, Edinburgh Univ. Dept. of AI, No. 169, 1981.
- [1042] F. Bunea, A. B. Tsybakov, and M. H. Wegkamp. Aggregation and sparsity via  $\ell_1$  penalized least squares. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 379–391, Berlin, 2006. Springer.

- [1043] F. Bunea, A. B. Tsybakov, and M. H. Wegkamp. Sparse density estimation with  $\ell_1$  penalties. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 530–543, Berlin, 2007. Springer.
- [1044] H. K. Büning and T. Lettmann. Learning a representation for optimizable formulas. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 51–58. Springer, 1996.
- [1045] W. Buntine. A critique of the Valiant model. In *Proc. of the 11th International Joint Conference on Artificial Intelligence*, pages 837–842, 1989.
- [1046] W. Buntine and T. Niblett. A further comparison of splitting rules for decision-tree induction. *Machine Learning*, 8:75–85, 1992.
- [1047] W. L. Buntine. *A Theory of Learning Classification Rules*. PhD thesis, University of Technology, Sydney, 1990.
- [1048] A. Burago. Learning structurally reversible context-free grammars from queries and counterexamples in polynomial time. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 140–146, New York, NY, 1994. ACM Press.
- [1049] W. Burgard, D. Fox, H. Jans, C. Matenar, and S. Thrun. Sonar-based mapping with mobile robots using EM. In *Proc. 16th International Conf. on Machine Learning*, pages 67–76. Morgan Kaufmann, San Francisco, CA, 1999.
- [1050] C. J. C. Burges. Simplified support vector decision rules. In *Proc. 13th International Conference on Machine Learning*, pages 71–77. Morgan Kaufmann, 1996.
- [1051] M. C. Burl, L. Asker, P. Smyth, U. Fayyad, P. Perona, L. Crumpler, and J. Aubele. Learning to recognize volcanoes on Venus. *Machine Learning*, 30:165–194, 1998.
- [1052] E. Burnaev and V. Vovk. Efficiency of conformalized ridge regression. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 605–622, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1053] B. D. Burns and A. Pohoreckyj-Danyluk. Feature selection vs theory reformulation: A study of genetic refinement of knowledge-based neural networks. *Machine Learning*, 38(1/2):89–107, 2000.
- [1054] R. Busa-Fekete and E. Hüllermeier. A survey of preference-based online learning with bandit algorithms. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 18–39, Berlin, Heidelberg, New York, 2014. Springer.
- [1055] S. Busuttill and Y. Kalnishkan. Online regression competitive with changing predictors. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 181–195, Berlin, Oct. 2007. Springer.
- [1056] T. Bylander. Polynomial learnability of linear threshold approximations. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 297–302, New York, NY, 1993. ACM Press.

- [1057] T. Bylander. Learning linear threshold functions in the presence of classification noise. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 340–347, New York, NY, 1994. ACM Press.
- [1058] T. Bylander. The binary exponentiated gradient algorithm for learning linear functions. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 184–192, New York, NY, 1997. ACM Press.
- [1059] T. Bylander. Learning probabilistically consistent linear threshold functions. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 62–71, New York, NY, 1997. ACM Press.
- [1060] T. Bylander. Estimating generalization error on two-class datasets using out-of-bag estimates. *Machine Learning*, 48(1/3):287–297, 2002.
- [1061] I. V. Cadez, C. E. McLaren, P. Smyth, and G. J. McLachlan. Hierarchical models for screening iron deficiency anemia. In *Proc. 16th International Conf. on Machine Learning*, pages 77–86. Morgan Kaufmann, San Francisco, CA, 1999.
- [1062] I. V. Cadez, P. Smyth, G. J. McLachlan, and C. E. McLaren. Maximum likelihood estimation of mixture densities for binned and truncated multivariate data. *Machine Learning*, 47(1):7–34, 2002.
- [1063] Y. Cai, C. Daskalakis, and C. Papadimitriou. Optimum statistical estimation with strategic data sources. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 280–296, Paris, France, 03–06 July 2015. PMLR.
- [1064] T. Calders. Recent developments in pattern mining. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 34–34, Berlin, Heidelberg, New York, 2012. Springer.
- [1065] J. Calera-Rubio and J. Oncina. Identifying left-right deterministic linear languages. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 283–284. Springer, 2004.
- [1066] M. E. Califf and R. J. Mooney. Bottom-up relational learning of pattern matching rules for information extraction. *Journal of Machine Learning Research*, 4:177–210, 2003.
- [1067] J. Callut and P. Dupont. A Markovian approach to the induction of regular string distributions. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 77–90. Springer, 2004.
- [1068] R. Camacho and P. Brazdil. Improving the robustness and encoding complexity of behavioural clones. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 37–48. Springer, 2001.
- [1069] C. Campbell, N. Cristianini, and A. Smola. Query learning with large margin classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 111–118. Morgan Kaufmann, San Francisco, CA, 2000.

- [1070] W. M. Campbell, K. Torkkola, and S. V. Balakrishnan. Dimension reduction techniques for training polynomial networks. In *Proc. 17th International Conf. on Machine Learning*, pages 119–126. Morgan Kaufmann, San Francisco, CA, 2000.
- [1071] N. Cancedda, E. Gaussier, C. Goutte, and J.-M. Renders. Word-sequence kernels. *Journal of Machine Learning Research*, 3:1059–1082, 2003.
- [1072] A. Cannon, J. M. Ettinger, D. Hush, and C. Scovel. Machine learning with data dependent hypothesis classes. *Journal of Machine Learning Research*, 2:335–358, 2002.
- [1073] A. Cano, J. Ruiz, and P. García. Inferring subclasses of regular languages faster using RPNI and forbidden configurations. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 28–36. Springer, 2002.
- [1074] A. Cano, J. Ruiz, and P. Garca. Running FCRPNI in efficient time for piecewise and right piecewise testable languages. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 275–276. Springer, 2004.
- [1075] C. L. Canonne, I. Diakonikolas, D. M. Kane, and A. Stewart. Testing Bayesian networks. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 370–448, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1076] D. Caragea, J. Zhang, J. Bao, J. Pathak, and V. Honavar. Algorithms and software for collaborative discovery from autonomous, semantically heterogeneous, distributed, information sources. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 13–44. Springer, Oct. 2005.
- [1077] C. Caramanis and S. Mannor. An inequality for nearly log-concave distributions with applications to learning. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 534–548. Springer, 2004.
- [1078] C. Caramanis and S. Mannor. Learning in the limit with adversarial disturbances. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 467–478. Omnipress, 2008.
- [1079] L. Carbonara and D. Sleeman. Improving the efficiency of knowledge base refinement. In *Proc. 13th International Conference on Machine Learning*, pages 78–86. Morgan Kaufmann, 1996.
- [1080] L. Carbonara and D. Sleeman. Effective and efficient knowledge base refinement. *Machine Learning*, 37(2):143–181, 1999.
- [1081] J. Carbonell, editor. *Machine Learning: Paradigms and Methods*. MIT Press, 1990.
- [1082] J. Carbonell. Machine learning: A maturing field. *Machine Learning*, 9:5–7, 1992.
- [1083] J. Carbonell, Y. Yang, and W. Cohen. Special issue of machine learning on information retrieval introduction. *Machine Learning*, 39(2/3):99–101, 2000.



- [1084] J. G. Carbonell and Y. Gil. Learning by experimentation. In *Proc. of the Fourth International Workshop on Machine Learning*, pages 256–266. Morgan Kaufmann, June 1987.
- [1085] C. Cardie. Integrating case-based learning and cognitive biases for machine learning of natural language. *J. of Experimental and Theoret. Artif. Intell.*, 11(3):297–337, 1999.
- [1086] C. Cardie. A cognitive bias approach to feature selection and weighting for case-based learners. *Machine Learning*, 41(1):85–116, 2000.
- [1087] C. Cardie. Successes, failures, and new directions in natural language learning. In *Proc. 18th International Conf. on Machine Learning*, pages 641–641. Morgan Kaufmann, San Francisco, CA, 2001.
- [1088] C. Cardie and N. Howe. Improving minority class prediction using case-specific feature weights. In *Proc. 14th International Conference on Machine Learning*, pages 57–65. Morgan Kaufmann, 1997.
- [1089] C. Cardie, S. Mardis, and D. Pierce. Combining error-driven pruning and classification for partial parsing. In *Proc. 16th International Conf. on Machine Learning*, pages 87–96. Morgan Kaufmann, San Francisco, CA, 1999.
- [1090] C. Cardie and R. J. Mooney. Guest editors’ introduction: Machine learning and natural language. *Machine Learning*, 34:5–9, 1999.
- [1091] J.-F. Cardoso. Dependence, correlation and gaussianity in independent component analysis. *Journal of Machine Learning Research*, 4:1177–1203, 2003.
- [1092] L. Carlucci. Incremental learning with ordinal bounded example memory. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 323–337, Berlin/Heidelberg, 2009. Springer.
- [1093] L. Carlucci, J. Case, and S. Jain. Learning correction grammars. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 203–217, Berlin, 2007. Springer.
- [1094] L. Carlucci, J. Case, S. Jain, and F. Stephan. Non U-shaped vacillatory and team learning. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 241–255. Springer, Oct. 2005.
- [1095] L. Carlucci, J. Case, S. Jain, and F. Stephan. Memory-limited U-shaped learning. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 244–258, Berlin, 2006. Springer.
- [1096] L. Carlucci, J. Case, S. Jain, and F. Stephan. Results on memory-limited U-shaped learning. *Inform. Comput.*, 205(10):1551–1573, 2007.
- [1097] L. Carlucci, J. Case, S. Jain, and F. Stephan. Non-U-shaped vacillatory and team learning. *J. Comput. Syst. Sci.*, 74(4):409–430, 2008.
- [1098] L. Carlucci, S. Jain, E. Kinber, and F. Stephan. Variations on U-shaped learning. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 382–397. Springer, 2005.

- [1099] L. Carlucci, S. Jain, E. Kinber, and F. Stephan. Variations on U-shaped learning. *Inform. Comput.*, 204(8):1264–1294, 2006.
- [1100] L. Carlucci, S. Jain, and F. Stephan. Learning with ordinal-bounded memory from positive data. *J. Comput. Syst. Sci.*, 78(5):1623–1636, 2012.
- [1101] J. Carme, M. Ceresna, and M. Goebel. Query-based learning of Xpath expressions. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 342–343, Berlin, 2006. Springer.
- [1102] J. Carme, R. Gilleron, A. Lemay, and J. Niehren. Interactive learning of node selecting tree transducer. *Machine Learning*, 66(1):33–67, 2007.
- [1103] J. Carme, A. Lemay, and J. Niehren. Learning node selecting tree transducer from completely annotated examples. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 91–102. Springer, 2004.
- [1104] A. Carpentier, A. Lazaric, M. Ghavamzadeh, R. Munos, and P. Auer. Upper-confidence-bound algorithms for active learning in multi-armed bandits. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 189–203, Berlin, Heidelberg, New York, 2011. Springer.
- [1105] A. Carpentier and A. Locatelli. Tight (lower) bounds for the fixed budget best arm identification bandit problem. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 590–604, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1106] A. Carpentier and R. Munos. Minimax number of strata for online stratified sampling given noisy samples. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 229–244, Berlin, Heidelberg, New York, 2012. Springer.
- [1107] A. Carpentier and R. Munos. Minimax number of strata for online stratified sampling: The case of noisy samples. *Theoret. Comput. Sci.*, 558:77–106, 2014. Special Issue: Algorithmic Learning Theory.
- [1108] R. Carrasco and J. Oncina. Learning stochastic regular grammars by means of a state merging method. In *Proc. 2nd International Colloquium on Grammatical Inference - ICGI '94*, volume 862 of *Lecture Notes in Artificial Intelligence*, pages 139–150. Springer-Verlag, 1994.
- [1109] R. C. Carrasco, M. L. Forcada, and L. Santamaría. Inferring stochastic regular grammars with recurrent neural networks. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 274–281. Springer, 1996.
- [1110] R. C. Carrasco and J. Oncina. Learning deterministic regular grammars from stochastic samples in polynomial time. *Theoretical Informatics and Applications*, 33(1):1–20, 1999.

- [1111] R. C. Carrasco, J. Oncina, and J. Calera. Stochastic inference of regular tree languages. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 187–198. Springer, 1998.
- [1112] R. C. Carrasco, J. Oncina, and J. Calera-Rubio. Stochastic inference of regular tree languages. *Machine Learning*, 44(1/2):185–197, 2001.
- [1113] R. Caruana. Algorithms and applications for multitask learning. In *Proc. 13th International Conference on Machine Learning*, pages 87–95. Morgan Kaufmann, 1996.
- [1114] R. Caruana. Multitask learning. *Machine Learning*, 28(1):41–75, 1997.
- [1115] R. Caruana and D. Freitag. Greedy attribute selection. In *Proc. 11th International Conference on Machine Learning*, pages 28–36. Morgan Kaufmann, 1994.
- [1116] R. Caruana and V. R. de Sa. Benefitting from the variables that variable selection discards. *Journal of Machine Learning Research*, 3:1245–1264, 2003.
- [1117] F. Casacuberta. Maximum mutual information and conditional maximum likelihood estimation of stochastic regular syntax-directed translation schemes. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 282–291. Springer, 1996.
- [1118] F. Casacuberta. Inference of finite-state transducers by using regular grammars and morphisms. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 1–14. Springer, 2000.
- [1119] F. Casacuberta and C. le la Higuera. Computational complexity of problems on probabilistic grammars and transducers. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 15–24. Springer, 2000.
- [1120] F. Casacuberta and E. Vidal. Learning finite-state models for machine translation. *Machine Learning*, 66(1):69–91, 2007.
- [1121] S. Casalaina-Martin, R. Frongillo, T. Morgan, and B. Waggoner. Multi-observation elicitation. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 449–464, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1122] J. Case. Periodicity in generations of automata. *Math. Syst. Theory*, 8(1):15–32, 1974.
- [1123] J. Case. Problems in inductive inference. *Bulletin of the European Association for Theoretical Computer Science*, 18:117–118, 1982.
- [1124] J. Case. Learning machines. In *Language Learning and Concept Acquisition*. Ablex Publishing Company, 1986.
- [1125] J. Case. The power of vacillation. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 196–205, San Mateo, CA, 1988. Morgan Kaufmann.
- [1126] J. Case. Infinitary self-reference in learning theory. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):3–36, 1994. Special issue for AII '92.

- [1127] J. Case. The power of vacillation in language learning. *SIAM J. Comput.*, 28(6):1941–1969, 1999.
- [1128] J. Case, K. J. Chen, and S. Jain. Strong separation of learning classes. *J. of Experimental and Theoret. Artif. Intell.*, 4:281–293, 1992.
- [1129] J. Case, K.-J. Chen, and S. Jain. Costs of general purpose learning. In *STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science, Trier, Germany, March 1999, Proceedings*, volume 1563 of *Lecture Notes in Computer Science*, pages 424–433. Springer, 1999.
- [1130] J. Case, K.-J. Chen, and S. Jain. Costs of general purpose learning. *Theoret. Comput. Sci.*, 259(1-2):455–473, 2001.
- [1131] J. Case, K.-J. Chen, S. Jain, W. Merkle, and J. S. Royer. Generality’s price: Inescapable deficiencies in machine-learned programs. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 684–698. Springer, 2003.
- [1132] J. Case, K.-J. Chen, S. Jain, W. Merkle, and J. S. Royer. Generality’s price: Inescapable deficiencies in machine-learned programs. *Annals of Pure and Applied Logic*, 139(1-3):303–326, 2006.
- [1133] J. Case and H. Chi. Machine learning of nearly minimal size grammars. Unpublished Manuscript, 1986.
- [1134] J. Case and M. Fulk. Solution to inductive inference problem P3. *Bulletin of the European Association for Theoretical Computer Science*, 19:8, 1983.
- [1135] J. Case and M. A. Fulk. Maximal machine learnable classes. *J. of Comput. Syst. Sci.*, 58(1):211–214, 1999.
- [1136] J. Case and S. Jain. Synthesizing learners tolerating computable noisy data. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 205–219. Springer, 1998.
- [1137] J. Case and S. Jain. Synthesizing learners tolerating computable noisy data. *J. Comput. Syst. Sci.*, 62(3):413–441, 2001.
- [1138] J. Case and S. Jain. Connections between inductive inference and machine learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 210–219. Springer, New York, 2010.
- [1139] J. Case, S. Jain, S. Kaufmann, A. Sharma, and F. Stephan. Predictive learning models for concept drift. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 276–290. Springer, 1998.
- [1140] J. Case, S. Jain, S. Kaufmann, A. Sharma, and F. Stephan. Predictive learning models for concept drift. *Theoret. Comput. Sci.*, 268(2):323–349, 2001. Special issue for ALT '98.
- [1141] J. Case, S. Jain, S. Lange, and T. Zeugmann. Incremental concept learning for bounded data mining. *Inform. Comput.*, 152(1):74–110, 1999.

- [1142] J. Case, S. Jain, T. D. Le, Y. S. Ong, P. Semukhin, and F. Stephan. Automatic learning of subclasses of pattern languages. *Inform. Comput.*, 218:17–35, 2012.
- [1143] J. Case, S. Jain, E. Martin, A. Sharma, and F. Stephan. Identifying clusters from positive data. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 103–114. Springer, 2004.
- [1144] J. Case, S. Jain, E. Martin, A. Sharma, and F. Stephan. Identifying clusters from positive data. *SIAM J. Comput.*, 36(1):28–55, 2006.
- [1145] J. Case, S. Jain, F. Montagna, G. Simi, and A. Sorbi. On learning to coordinate: Random bits help, insightful normal forms, and competency isomorphisms. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 699–713. Springer, 2003.
- [1146] J. Case, S. Jain, F. Montagna, G. Simi, and A. Sorbi. On learning to coordinate: random bits help, insightful normal forms, and competency isomorphisms. *J. Comput. Syst. Sci.*, 71(3):308–332, 2005. Special issue Learning Theory 2003.
- [1147] J. Case, S. Jain, and S. Ngo-Manguelle. Refinements of inductive inference by popperian and reliable machines. *Kybernetika*, 30(1):23–52, 1994.
- [1148] J. Case, S. Jain, Y. S. Ong, P. Semukhin, and F. Stephan. Automatic learners with feedback queries. *J. Comput. Syst. Sci.*, 80(4):806–820, 2014.
- [1149] J. Case, S. Jain, M. Ott, A. Sharma, and F. Stephan. Robust learning aided by context. *J. Comput. Syst. Sci.*, 60(2):234–257, 2000. Special issue for COLT '98.
- [1150] J. Case, S. Jain, R. Reischuk, F. Stephan, and T. Zeugmann. Learning a subclass of regular patterns in polynomial time. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 234–246. Springer, 2003.
- [1151] J. Case, S. Jain, R. Reischuk, F. Stephan, and T. Zeugmann. A polynomial time learner for a subclass of regular patterns. Technical Report Report TR04-038, Electronic Colloquium on Computational Complexity, Apr. 2004.
- [1152] J. Case, S. Jain, R. Reischuk, F. Stephan, and T. Zeugmann. Learning a subclass of regular patterns in polynomial time. *Theoret. Comput. Sci.*, 364(1):115–131, 2006. Special issue for ALT 2003.
- [1153] J. Case, S. Jain, and A. Sharma. Anomalous learning helps succinctness. In *Proc. of the First Workshop on Algorithmic Learning Theory, Tokyo, Japan*, pages 282–288. Japanese Society for Artificial Intelligence, Oct. 1990.
- [1154] J. Case, S. Jain, and A. Sharma. Convergence to nearly minimal size grammars by vacillating learning machines. Technical Report 90-10, University of Delaware, Newark, Delaware, 1990.
- [1155] J. Case, S. Jain, and A. Sharma. On learning limiting programs. *Internat. J. of Found. of Comput. Sci.*, 3(1):93–115, 1992.
- [1156] J. Case, S. Jain, and A. Sharma. Vacillatory learning of nearly-minimal size grammars. *J. of Comput. Syst. Sci.*, 49(2):189–207, 1994.

- [1157] J. Case, S. Jain, and A. Sharma. Complexity issues for vacillatory function identification. *Inform. Comput.*, 116(2):174–192, 1995.
- [1158] J. Case, S. Jain, and A. Sharma. Machine induction without revolutionary paradigm shifts. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 153–168. Springer, 1995.
- [1159] J. Case, S. Jain, and A. Sharma. Anomalous learning helps succinctness. *Theoret. Comput. Sci.*, 164(1-2):13–28, 1996.
- [1160] J. Case, S. Jain, and A. Sharma. Machine induction without revolutionary changes in hypothesis size. *Inform. Comput.*, 128(2):73–86, Aug. 1996.
- [1161] J. Case, S. Jain, and A. Sharma. Synthesizing noise-tolerant language learners. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 228–243. Springer, 1997.
- [1162] J. Case, S. Jain, and A. Sharma. Synthesizing noise-tolerant language learners. *Theoret. Comput. Sci.*, 261(1):31–56, 2001. Special issue for ALT '97.
- [1163] J. Case, S. Jain, and F. Stephan. Vacillatory and BC learning on noisy data. *Theoret. Comput. Sci.*, 241(1-2):115–141, 2000. Special issue for ALT '96.
- [1164] J. Case, S. Jain, and F. Stephan. Effectivity questions for kleene’s recursion theorem. *Theoret. Comput. Sci.*, 733:55–70, 2018. Special Issue on Learning Theory and Complexity.
- [1165] J. Case, S. Jain, F. Stephan, and R. Wiehagen. Robust learning — rich and poor. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 143–159. Springer, 2001.
- [1166] J. Case, S. Jain, F. Stephan, and R. Wiehagen. Robust learning – rich and poor. *J. Comput. Syst. Sci.*, 69(2):123–165, 2004.
- [1167] J. Case, S. Jain, and M. Suraj. Control structures in hypothesis spaces: the influence on learning. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 286–300. Springer, 1997.
- [1168] J. Case, S. Jain, and M. Suraj. Control structures in hypothesis spaces: the influence on learning. *Theoret. Comput. Sci.*, 270(1-2):287–308, 2002.
- [1169] J. Case, S. Kaufmann, E. B. Kinber, and M. Kummer. Learning recursive functions from approximations. *J. Comput. Syst. Sci.*, 55(1):183–196, 1997. Special issue for EuroCOLT '95.
- [1170] J. Case, E. Kinber, A. Sharma, and F. Stephan. On the classification of computable languages. In *Proc. 14th Annual Symposium on Theoretical Aspects of Computer Science: STACS'97*, volume 1200 of *Lecture Notes in Computer Science*, pages 225–236. Springer, 1997.
- [1171] J. Case, E. Kinber, A. Sharma, and F. Stephan. On the classification of recursive languages. *Inform. Comput.*, 192(1):15–40, 2004.

- [1172] J. Case and T. Kötzing. Dynamic modeling in inductive inference. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 404–418, Berlin, Oct. 2008. Springer.
- [1173] J. Case and T. Kötzing. Dynamically delayed postdictive completeness and consistency in learning. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 389–403, Berlin, Oct. 2008. Springer.
- [1174] J. Case and T. Kötzing. Difficulties in forcing fairness of polynomial time inductive inference. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 263–277, Berlin/Heidelberg, 2009. Springer.
- [1175] J. Case and T. Kötzing. Solutions to open questions for non-U-shaped learning with memory limitations. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 285–299, Berlin, Heidelberg, New York, 2010. Springer.
- [1176] J. Case and T. Kötzing. Learning secrets interactively. dynamic modeling in inductive inference. *Inform. Comput.*, 220-221:60–73, 2012.
- [1177] J. Case and T. Kötzing. Memory-limited non-U-shaped learning with solved open problems. *Theoret. Comput. Sci.*, 473:100–123, 2013. Special Issue on Learning Theory.
- [1178] J. Case and T. Kötzing. Topological separations in inductive inference. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 128–142, Berlin, Heidelberg, New York, 2013. Springer.
- [1179] J. Case and T. Kötzing. Topological separations in inductive inference. *Theoret. Comput. Sci.*, 620:33–45, 2016. Special Issue: Algorithmic Learning Theory.
- [1180] J. Case, T. Kötzing, and T. Paddock. Feasible iteration of feasible learning functionals. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 34–48, Berlin, Oct. 2007. Springer.
- [1181] J. Case and C. Lynes. Machine inductive inference and language identification. In *Automata, Languages and Programming, 9th Colloquium, Proceedings*, volume 140, pages 107–115. Springer-Verlag, 1982.
- [1182] J. Case and S. E. Moelius III. Parallelism increases iterative learning power. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 49–63, Berlin, Oct. 2007. Springer.
- [1183] J. Case and S. E. Moelius III. U-shaped, iterative, and iterative-with-counter learning. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 172–186, Berlin, 2007. Springer.

- [1184] J. Case and S. E. Moelius III. Optimal language learning. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 419–433, Berlin, Oct. 2008. Springer.
- [1185] J. Case and S. E. Moelius III. Parallelism increases iterative learning power. *Theoret. Comput. Sci.*, 410(9):1863–1875, 2009. Special issue for ALT 2007.
- [1186] J. Case and S. E. Moelius III. Optimal language learning from positive data. *Inform. Comput.*, 209(10):1293–1311, 2011.
- [1187] J. Case and S. Ngo-Manguelle. Refinements of inductive inference by Popperian machines. Technical Report 152, SUNY Buffalo, Dept. of Computer Science, 1979.
- [1188] J. Case, M. Ott, A. Sharma, and F. Stephan. Learning to win process-control games watching game-masters. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 8-10, 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 31–45. Springer, 1998.
- [1189] J. Case, M. Ott, A. Sharma, and F. Stephan. Learning to win process-control games watching game-masters. *Inform. Comput.*, 174(1):1–19, 2002.
- [1190] J. Case, D. S. Rajan, and A. M. Shende. Representing the spatial/kinematic domain and lattice computers. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 1–25. Springer-Verlag, 1992.
- [1191] J. Case, T. Shinohara, T. Zeugmann, and S. Zilles. Foreword. *Theoret. Comput. Sci.*, 397(1-3):1–3, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [1192] J. Case and C. Smith. Comparison of identification criteria for machine inductive inference. *Theoret. Comput. Sci.*, 25(2):193–220, 1983.
- [1193] J. Case, M. Suraj, and S. Jain. Not-so-nearly-minimal-size program inference. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 76–95. Springer, 1995.
- [1194] A. Castellanos. Approximate learning of random subsequential transducers. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 67–78. Springer, 1998.
- [1195] T. K. R. Castelo. On inclusion-driven learning of Bayesian networks. *Journal of Machine Learning Research*, 4:527–574, 2003.
- [1196] E. Castillo, A. S. Hadi, and C. Solares. Learning and updating of uncertainty in dirichlet models. *Machine Learning*, 26(1):43–63, 1997.
- [1197] J. Castro. How many query superpositions are needed to learn? In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 78–92. Springer, Oct. 2006.
- [1198] J. Castro and J. L. Balcázar. Simple PAC learning of simple decision lists. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 239–248. Springer, 1995.



- [1199] J. Castro and R. Gavaldà. Towards feasible PAC-learning of probabilistic deterministic finite automata. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 163–174. Springer, 2008.
- [1200] J. Castro and D. Guijarro. PACS, simple-PAC and query learning. *Information Processing Letters*, 73(1-2):11–16, 2000.
- [1201] J. Castro, D. Guijarro, and V. Lavin. Learning nearly monotone  $k$ -term DNF. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 162–170. Springer, 1997.
- [1202] R. M. Castro and R. D. Nowak. Minimax bounds for active learning. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 5–19, Berlin, 2007. Springer.
- [1203] G. Cavallanti, N. Cesa-Bianchi, and C. Gentile. Tracking the best hyperplane with a simple budget perceptron. *Machine Learning*, 69(2-3):143–167, 2007.
- [1204] G. Cavallanti, N. Cesa-Bianchi, and C. Gentile. Linear algorithms for online multitask classification. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 251–262. Omnipress, 2008.
- [1205] D. A. Cezner and W. R. Moser. Inductive inference of functions on the rationals. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 178–181, New York, NY, 1995. ACM Press.
- [1206] D. A. Cezner and W. R. Moser. A good oracle is hard to beat. *Algorithmica*, 22(1/2):18–34, 1998.
- [1207] P. Černo. Clearing restarting automata and grammatical inference. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 54–68, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1208] N. Cesa-Bianchi. Learning the distribution in the extended PAC model. In *Proc. of the First International Workshop on Algorithmic Learning Theory*, pages 236–246. Japanese Society for Artificial Intelligence, 1990.
- [1209] N. Cesa-Bianchi. Analysis of two gradient-based algorithms for on-line regression. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 163–170, New York, NY, 1997. ACM Press.
- [1210] N. Cesa-Bianchi. Applications of regularized least squares to classification problems. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 14–18. Springer, 2004.
- [1211] N. Cesa-Bianchi. Applications of regularized least squares to pattern classification. *Theoret. Comput. Sci.*, 382(3):221–231, 2007. Special issue for ALT 2004.

- [1212] N. Cesa-Bianchi, A. Conconi, and C. Gentile. A second-order perceptron algorithm. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 121–137. Springer, 2002.
- [1213] N. Cesa-Bianchi, A. Conconi, and C. Gentile. Learning probabilistic linear-threshold classifiers via selective sampling. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 373–387. Springer, 2003.
- [1214] N. Cesa-Bianchi, A. Conconi, and C. Gentile. Regret bounds for hierarchical classification with linear-threshold functions. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 93–108. Springer, 2004.
- [1215] N. Cesa-Bianchi, A. Conconi, and C. Gentile. A second-order perceptron algorithm. *SIAM J. Comput.*, 34(3):640–668, 2005.
- [1216] N. Cesa-Bianchi, E. Dichterman, P. Fischer, E. Shamir, and H. U. Simon. Sample-efficient strategies for learning in the presence of noise. *J. ACM*, 46(5):684–719, 1999.
- [1217] N. Cesa-Bianchi and P. Fischer. Finite-time regret bounds for the multiarmed bandit problem. In *Proc. 15th International Conf. on Machine Learning*, pages 100–108. Morgan Kaufmann, San Francisco, CA, 1998.
- [1218] N. Cesa-Bianchi, P. Fischer, E. Shamir, and H. U. Simon. Randomized hypotheses and minimum disagreement hypotheses for learning with noise. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 119–133. Springer, 1997.
- [1219] N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to use expert advice. *J. ACM*, 44(3):427–485, 1997.
- [1220] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, D. Haussler, R. E. Schapire, and M. K. Warmuth. How to use expert advice. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 382–391, New York, NY, 1993. ACM Press. Expanded version in Univ. of Calif. Computer Research Lab TR UCSC-CRL-94-33, From Santa Cruz, CA.
- [1221] N. Cesa-Bianchi, Y. Freund, D. P. Helmbold, and M. K. Warmuth. On-line prediction and conversion strategies. *Machine Learning*, 25:71–110, 1996. An extended abstract appeared in *EuroColt '93*.
- [1222] N. Cesa-Bianchi, P. Gaillard, C. Gentile, and S. Gerchinovitz. Algorithmic chaining and the role of partial feedback in online nonparametric learning. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 465–481, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1223] N. Cesa-Bianchi and C. Gentile. Tracking the best hyperplane with a simple budget perceptron. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 483–498, Berlin, 2006. Springer.

- [1224] N. Cesa-Bianchi, C. Gentile, Y. Mansour, and A. Minora. Delay and cooperation in non-stochastic bandits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 605–622, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1225] N. Cesa-Bianchi, C. Gentile, and F. Vitale. Learning unknown graphs. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 110–125, Berlin/Heidelberg, 2009. Springer.
- [1226] N. Cesa-Bianchi, C. Gentile, and F. Vitale. Predicting the labels of an unknown graph via adaptive exploration. *Theoret. Comput. Sci.*, 412(19):1791–1804, 2011.
- [1227] N. Cesa-Bianchi, C. Gentile, F. Vitale, and G. Zappella. A correlation clustering approach to link classification in signed networks. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 34.1–34.20, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1228] N. Cesa-Bianchi and S. Goldman, editors. *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory*, San Francisco, CA, 2000. Morgan Kaufmann, Inc.
- [1229] N. Cesa-Bianchi, D. P. Helmbold, and S. Panizza. On Bayes methods for on-line Boolean prediction. *Algorithmica*, 22(1/2):112–137, 1998.
- [1230] N. Cesa-Bianchi, A. Krogh, and M. K. Warmuth. Bounds on approximate steepest descent for likelihood maximization in exponential families. *IEEE Transaction on Information Processing*, 40(4):1215–1220, July 1994.
- [1231] N. Cesa-Bianchi, P. Long, and M. Warmuth. Worst-case quadratic loss bounds for a generalization of the Widrow-Hoff rule. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 429–438, New York, NY, 1993. ACM Press.
- [1232] N. Cesa-Bianchi, P. M. Long, and M. K. Warmuth. Worst-case quadratic loss bounds for on-line prediction of linear functions by gradient descent. *IEEE Transactions on Neural Networks*, 7(3):604–619, 1996. Earlier version in 6th COLT, 1993.
- [1233] N. Cesa-Bianchi and G. Lugosi. On sequential prediction of individual sequences relative to a set of experts. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 1–11, New York, NY, 1998. ACM Press.
- [1234] N. Cesa-Bianchi and G. Lugosi. Minimax regret under log loss for general classes of experts. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 12–18, New York, NY, 1999. ACM Press.
- [1235] N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 48–64. Springer, 2001.
- [1236] N. Cesa-Bianchi and G. Lugosi. Worst-case bounds for the logarithmic loss of predictors. *Machine Learning*, 43(3):247–264, 2001.
- [1237] N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory. *Machine Learning*, 51(3):239–261, 2003.

- [1238] N. Cesa-Bianchi and G. Lugosi. Combinatorial bandits. *J. Comput. Syst. Sci.*, 78(5):1404–1422, 2012.
- [1239] N. Cesa-Bianchi, G. Lugosi, and G. Stoltz. Minimizing regret with label efficient prediction. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 77–92. Springer, 2004.
- [1240] N. Cesa-Bianchi, Y. Mansour, and O. Shamir. On the complexity of learning with kernels. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 297–325, Paris, France, 03–06 July 2015. PMLR.
- [1241] N. Cesa-Bianchi, Y. Mansour, and G. Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.
- [1242] N. Cesa-Bianchi, M. Numao, and R. Reischuk, editors. *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany, November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Nov. 2002. Springer.
- [1243] N. Cesa-Bianchi, M. Numao, and R. Reischuk. Editors’ introduction. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 1–6. Springer, 2002.
- [1244] N. Cesa-Bianchi, R. Reischuk, and T. Zeugmann. Foreword. *Theoret. Comput. Sci.*, 350(1):1–2, 2006. Special issue for ALT 2002.
- [1245] N. Cesa-Bianchi and O. Shamir. Bandit regret scaling with the effective loss range. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 128–151. PMLR, 07–09 Apr. 2018.
- [1246] I. K. Cestnik, B and I. Bratko. Assistant 86: A knowledge-elicitation tool for sophisticated users. In *Progress in Machine Learning—Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 31–45, Bled, Yugoslavia, May 1987.
- [1247] U. Chajewska, D. Koller, and D. Ormoneit. Learning an agent’s utility function by observing behavior. In *Proc. 18th International Conf. on Machine Learning*, pages 35–42. Morgan Kaufmann, San Francisco, CA, 2001.
- [1248] S. Chakrabartty, G. Cauwenberghs, and Jayadeva. Sparse probability regression by label partitioning. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 231–242. Springer, 2003.
- [1249] S. K. Chalup and A. D. Blair. Software for analysing recurrent neural nets that learn to predict non-regular languages. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 296–298. Springer, 2002.

- [1250] R. G. J. Champavère, A. Lemay, and J. Niehren. Schema-guided induction of monadic queries. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 15–28. Springer, 2008.
- [1251] K. Chan, T.-W. Lee, and T. J. Sejnowski. Variational learning of clusters of undercomplete nonsymmetric independent components. *Journal of Machine Learning Research*, 3:99–114, 2002.
- [1252] P. K. Chan and S. J. Stolfo. A comparative evaluation of voting and meta-learning on partitioned data. In *Proc. 12th International Conference on Machine Learning*, pages 90–98. Morgan Kaufmann, 1995.
- [1253] P. K. Chan, S. J. Stolfo, and D. Wolpert. Guest editors’ introduction. *Machine Learning*, 36(1/2):5–7, 1999.
- [1254] S. O. Chan, D. Papailiopoulos, and A. Rubinstein. On the approximability of sparse PCA. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 623–646, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1255] J. Chandlee, J. Fu, K. Karydis, C. Koirala, J. Heinz, and H. Tanner. Integrating grammatical inference into robotic planning. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 69–83, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1256] B. Chandrasekaran. Task-structures, knowledge acquisition and learning. *Machine Learning*, 4:339–345, 1989.
- [1257] K. Chandrasekaran and R. Karp. Finding a most biased coin with fewest flips. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 394–407, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1258] H. Chang, D. Cohn, and A. K. McCallum. Learning to create customized authority lists. In *Proc. 17th International Conf. on Machine Learning*, pages 127–134. Morgan Kaufmann, San Francisco, CA, 2000.
- [1259] K. L. Chang. Multiple pass streaming algorithms for learning mixtures of distributions in  $\mathbb{R}^d$ . In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 211–226, Berlin, Oct. 2007. Springer.
- [1260] K. L. Chang. Multiple pass streaming algorithms for learning mixtures of distributions in  $\mathbb{R}^d$ . *Theoret. Comput. Sci.*, 410(9):1765–1780, 2009. Special issue for ALT 2007.
- [1261] K. L. Chang and R. Kannan. Pass-efficient algorithms for learning mixtures of uniform distributions. *SIAM J. Comput.*, 39(3):783–812, 2009.
- [1262] R. Chang and W. I. Gasarch. On bounded queries and approximation. In *Proc. of the 34rd Annual Symposium on Foundations of Computer Science*, pages 547–556. IEEE Computer Society Press, Los Alamitos, CA, 1993.

- [1263] O. Chapelle, V. Vapnik, and Y. Bengio. Model selection for small sample regression. *Machine Learning*, 48(1/3):9–23, 2002.
- [1264] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee. Choosing multiple parameters for support vector machines. *Machine Learning*, 46(1-3):131–159, 2002.
- [1265] M. Charikar, S. Guha, É. Tardos, and D. B. Shmoys. A constant-factor approximation algorithm for the  $k$ -median problem (extended abstract). In *Proc. 31st Annu. ACM Symposium on Theory of Comput.*, pages 1–10, New York, NY, 1999. ACM Press.
- [1266] M. Charikar, V. Guruswami, and A. Wirth. Clustering with qualitative information. *J. Comput. Syst. Sci.*, 71(3):360–383, 2005. Special issue Learning Theory 2003.
- [1267] K. Chaudhuri, F. Chung, and A. Tsiatas. Spectral clustering of graphs with general degrees in the extended planted partition model. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 35.1–35.23, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1268] K. Chaudhuri and C. Gentile. Guest editors’ introduction. *Theoret. Comput. Sci.*, 716:1–3, 2018. Special Issue on ALT 2015.
- [1269] K. Chaudhuri, C. Gentile, and S. Zilles, editors. *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*. Springer, 2015.
- [1270] K. Chaudhuri and D. Hsu. Sample complexity bounds for differentially private learning. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 155–186, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1271] K. Chaudhuri and A. McGregor. Finding metric structure in information theoretic clustering. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 391–402. Omnipress, 2008.
- [1272] K. Chaudhuri and S. Rao. Beyond gaussians: Spectral methods for learning mixtures of heavy-tailed distributions. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 21–32. Omnipress, 2008.
- [1273] K. Chaudhuri and S. Rao. Learning mixtures of product distributions using correlations and independence. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 9–20. Omnipress, 2008.
- [1274] D. Chawla, L. Li, and S. Scott. Efficiently approximating weighted sums with exponentially many terms. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 82–98. Springer, 2001.
- [1275] N. V. Chawla, L. O. Hall, K. W. Bowyer, and W. P. Kegelmeyer. Learning ensembles from bites: A scalable and accurate approach. *Journal of Machine Learning Research*, 5:421–451, 2004.
- [1276] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss. Information bottleneck for gaussian variables. *Journal of Machine Learning Research*, 6:165–188, 2005.

- [1277] P. C. Cheeseman. A method of computing generalized Bayesian probability values for expert systems. In *Proc. Eighth International Conference on Artificial Intelligence (Karlsruhe, West Germany)*, pages 198–202, Aug. 1983.
- [1278] P. C. Cheeseman. Learning of expert systems from data. In *Proc. of the Workshop on Principles of Knowledge-Based Systems*, pages 115–122, Dec. 1984.
- [1279] C. Chen, D. Dong, and Z. Chen. Grey reinforcement learning for incomplete information processing. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 399–407, Berlin, 2006. Springer.
- [1280] D. Chen, C. L. Giles, G. Sun, M. W. Goudreau, H.-H. Chen, and Y.-C. Lee. A method for constructive learning of recurrent neural networks. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 6, pages 95–114. MIT Press, 1995.
- [1281] D. Chen and J. M. Phillips. Relative error embeddings of the gaussian kernel distance. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 560–576, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [1282] H. Chen and M. Valeriote. Learnability of solutions to conjunctive queries: The full dichotomy. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 326–337, Paris, France, 03–06 July 2015. PMLR.
- [1283] K.-J. Chen. *Tradeoffs in Machine Inductive Inference*. PhD thesis, Department of Computer Science, State University of New York at Buffalo, 1981.
- [1284] K.-J. Chen. Tradeoffs in the inductive inference of nearly minimal size programs. *Inform. Control*, 52(1):68–86, 1982.
- [1285] L. Chen, A. Gupta, and J. Li. Pure exploration of multi-armed bandit under matroid constraints. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 647–669, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1286] L. Chen, A. Gupta, J. Li, M. Qiao, and R. Wang. Nearly optimal sampling algorithms for combinatorial pure exploration. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 482–534, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1287] L. Chen and J. Li. Open problem: Best arm identification: Almost instance-wise optimality and the gap entropy conjecture. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1643–1646, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1288] L. Chen, J. Li, and M. Qiao. Towards instance optimal bounds for best arm identification. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 535–592, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [1289] P. C. Y. Chen, J. K. Mills, and K. C. Smith. Performance improvement of robot continuous-path operation through iterative learning using neural networks. *Machine Learning*, 23:191–220, 1996.
- [1290] P.-H. Chen, R.-E. Fan, and C.-J. Lin. Training support vector machines via SMO-type decomposition methods. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 45–62. Springer, Oct. 2005.
- [1291] Y. Chen, S. H. Hassani, A. Karbasi, and A. Krause. Sequential information maximization: When is greedy near-optimal? In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 338–363, Paris, France, 03–06 July 2015. PMLR.
- [1292] Y. Chen, X. Yi, and C. Caramanis. A convex formulation for mixed regression with two components: Minimax optimal rates. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 560–604, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1293] Z. Chen. Learning unions of two rectangles in the plane with equivalence queries. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 243–252, New York, NY, 1993. ACM Press.
- [1294] Z. Chen. A note on learning DNF formulas using equivalence and incomplete membership queries. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 272–281. Springer-Verlag, 1994.
- [1295] Z. Chen and F. Ameur. The learnability of unions of two rectangles in the two-dimensional discretized space. *J. of Comput. Syst. Sci.*, 59(1):70–83, 1999.
- [1296] Z. Chen and S. Homer. The bounded injury priority method and the learnability of unions of rectangles. *Ann. Pure and Applied Logic*, 77(2):143–168, 1996.
- [1297] Z. Chen and S. Homer. Learning counting functions with queries. *Theoret. Comput. Sci.*, 180(1-2):155–168, 1997. first in COLT '94.
- [1298] Z. Chen and W. Maass. On-line learning of rectangles. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 16–28, New York, NY, 1992. ACM Press.
- [1299] Z. Chen and W. Maass. A solution of the credit assignment problem in the case of learning rectangles (abstract). In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 26–34. Springer-Verlag, 1992.
- [1300] Z. Chen and W. Maass. On-line learning of rectangles and unions of rectangles. *Machine Learning*, 17(2/3):23–50, 1994. Special issue for COLT '92.
- [1301] D. Cheng, Y. Cheng, Y. Liu, R. Peng, and S.-H. Teng. Efficient sampling for gaussian graphical models via spectral sparsification. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 364–390, Paris, France, 03–06 July 2015. PMLR.



- [1302] X. Cheng and P. Bartlett. Convergence of langevin MCMC in KL-divergence. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 186–211. PMLR, 07–09 Apr. 2018.
- [1303] Y. Cheng and Z. Wan. Analysis of the blurring process. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 14, pages 257–276. MIT Press, 1995.
- [1304] Y. Cherapanamjeri, P. Jain, and P. Netrapalli. Thresholding based outlier robust PCA. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 593–628, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1305] J. Cherniavsky and C. Smith. Using telltales in developing program test sets. Technical report, Georgetown University, Dept. of Computer Science, TR-4, 1986.
- [1306] J. Cherniavsky and R. Statman. Testing and inductive inference: abstract approaches. Preprint, Georgetown University, 1987.
- [1307] J. C. Cherniavsky, M. Velauthapillai, and R. Statman. Inductive inference: an abstract approach. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 251–266, San Mateo, CA, 1988. Morgan Kaufmann.
- [1308] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Math. Stat.*, 23:493–509, 1952.
- [1309] A. Chernov and M. Hutter. Monotone conditional complexity bounds on future prediction errors. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 414–428. Springer, Oct. 2005.
- [1310] A. Chernov, Y. Kalnishkan, and V. Vovk. Supermartingales in prediction with expert advice. *Theoret. Comput. Sci.*, 411(29-30):2647–2669, 2010. Special issue for ALT 2008.
- [1311] A. Chernov, Y. Kalnishkan, F. Zhdanov, and V. Vovk. Supermartingales in prediction with expert advice. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 199–213, Berlin, Oct. 2008. Springer.
- [1312] A. Chernov, A. Shen, N. Vereshchagin, and V. Vovk. On-line probability, complexity and randomness. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 138–153, Berlin, Oct. 2008. Springer.
- [1313] A. Chernov and V. Vovk. Prediction with expert evaluators’ advice. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 8–22, Berlin/Heidelberg, 2009. Springer.
- [1314] A. Chernov and F. Zhdanov. Prediction with expert advice under discounted loss. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 255–269, Berlin, Heidelberg, New York, 2010. Springer.

- [1315] D. Cheung. Learning real polynomials with a Turing machine. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 231–240. Springer, 1999.
- [1316] D. Cheung and F. Cucker. Learning from rounded-off data. *Inform. Comput.*, 182(1):1–13, 2003.
- [1317] S. Cheung. Learning from approximate data. In *Computing and Combinatorics, 6th Annual International Conference, COCOON 2000, Sydney, Australia, July 26–28, 2000, Proceedings*, volume 1858 of *Lecture Notes in Computer Science*, pages 407–415. Springer, 2000.
- [1318] Y. Chevaleyre and J.-D. Zucker. A framework for learning rules from multiple instance data. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 49–60. Springer, 2001.
- [1319] C.-K. Chiang, C.-J. Lee, and C.-J. Lu. Beating bandits in gradually evolving worlds. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 210–227. JMLR.org, 2013.
- [1320] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 6.1–6.20, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1321] D. M. Chickering. Learning equivalence classes of Bayesian-network structures. *Journal of Machine Learning Research*, 2:445–498, 2002.
- [1322] D. M. Chickering. Optimal structure identification with greedy search. *Journal of Machine Learning Research*, 3:507–554, 2002.
- [1323] D. M. Chickering and D. Heckerman. Efficient approximations for the marginal likelihood of Bayesian networks with hidden variables. *Machine Learning*, 29(2-3):181–212, 1997.
- [1324] B. Chidlovskii. Wrapping web information providers by transducer induction. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 61–72. Springer, 2001.
- [1325] B. Chidlovskii, J. Ragetli, and M. de Rijke. Wrapper generation via grammar induction. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 96–108. Springer, 2000.
- [1326] P. Chin, A. Rao, and V. Vu. Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 391–423, Paris, France, 03–06 July 2015. PMLR.
- [1327] H. A. Chipman, E. I. George, and R. E. McCulloch. Bayesian treed models. *Machine Learning*, 48(1/3):299–320, 2002.

- [1328] J. Chodorowski and L. Miclet. Applying grammatical inference in learning a language model for oral dialogue. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 102–113. Springer, 1998.
- [1329] D. Choi and B. Van Roy. A generalized Kalman filter for fixed point approximation and efficient temporal-difference learning. In *Proc. 18th International Conf. on Machine Learning*, pages 43–50. Morgan Kaufmann, San Francisco, CA, 2001.
- [1330] S.-S. Choi. Polynomial time optimal query algorithms for finding graphs with arbitrary real weights. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 797–818. JMLR.org, 2013.
- [1331] S.-S. Choi, K. Jung, and J. H. Kim. Almost tight upper bound for finding fourier coefficients of bounded pseudo-Boolean functions. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 123–134. Omnipress, 2008.
- [1332] Y. S. Choi and S. I. Yoo. Learning to select text databases with neural nets. In *Proc. 17th International Conf. on Machine Learning*, pages 135–142. Morgan Kaufmann, San Francisco, CA, 2000.
- [1333] P. Cholak, R. Downey, L. Fortnow, W. Gasarch, E. Kinber, M. Kummer, S. Kurtz, and T. A. Slaman. Degrees of inferability. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 180–192, New York, NY, 1992. ACM Press.
- [1334] A. Choromanska, K. Choromanski, G. Jagannathan, and C. Monteleoni. Differentially-private learning of low dimensional manifolds. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 249–263, Berlin, Heidelberg, New York, 2013. Springer.
- [1335] A. Choromanska, K. Choromanski, G. Jagannathan, and C. Monteleoni. Differentially-private learning of low dimensional manifolds. *Theoret. Comput. Sci.*, 620:91–104, 2016. Special Issue: Algorithmic Learning Theory.
- [1336] A. Choromanska, T. Jebara, H. Kim, M. Mohan, and C. Monteleoni. Fast spectral clustering via the nyström method. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 367–381, Berlin, Heidelberg, New York, 2013. Springer.
- [1337] A. Choromanska, Y. LeCun, and G. B. Arous. Open problem: The landscape of the loss surfaces of multilayer networks. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1756–1760, Paris, France, 03–06 July 2015. PMLR.
- [1338] E. Chown and T. G. Dietterich. A divide and conquer approach to learning from prior knowledge. In *Proc. 17th International Conf. on Machine Learning*, pages 143–150. Morgan Kaufmann, San Francisco, CA, 2000.
- [1339] S. W. Christensen, I. Sinclair, and P. A. S. Reed. Designing committees of models through deliberate weighting of data points. *Journal of Machine Learning Research*, 4:39–66, 2003.

- [1340] P. Christiano. Open problem: Online local learning. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1290–1294, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1341] P. Christiano. Provably manipulation-resistant reputation systems. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 670–697, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1342] W. Chu, S. S. Keerthi, and C. J. Ong. A unified loss function in Bayesian framework for support vector regression. In *Proc. 18th International Conf. on Machine Learning*, pages 51–58. Morgan Kaufmann, San Francisco, CA, 2001.
- [1343] T. H. Chung. Approximate methods for sequential decision making using expert advice. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 183–189, New York, NY, 1994. ACM Press.
- [1344] N. A. Chuzhanova. Inductive synthesis of programs for symbolic sequences processing. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 317–327. Springer-Verlag, 1989.
- [1345] O. Cicchello and S. C. Kremer. Beyond EDSM. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 37–48. Springer, 2002.
- [1346] O. Cicchello and S. C. Kremer. Inducing grammars from sparse data sets: A survey of algorithms and results. *Journal of Machine Learning Research*, 4:603–632, 2003.
- [1347] P. Cichosz and J. J. Mulawka. Fast and efficient reinforcement learning with truncated temporal differences. In *Proc. 12th International Conference on Machine Learning*, pages 99–107. Morgan Kaufmann, 1995.
- [1348] R. Cilibrasi and P. Vitanyi. Similarity of objects and the meaning of words. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 21–45, Berlin, 2006. Springer.
- [1349] L. Ciortuz. A framework for inductive learning of typed-unification grammars. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 299–302. Springer, 2002.
- [1350] W. J. Clancey. The knowledge level reinterpreted: Modeling how systems interact. *Machine Learning*, 4:285–291, 1989.
- [1351] A. Clark. Large scale inference of deterministic transductions: Tenjinno problem 1. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 227–239, Berlin, 2006. Springer.

- [1352] A. Clark. PAC-learning unambiguous NTS languages. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 59–71, Berlin, 2006. Springer.
- [1353] A. Clark. Learning deterministic context free grammars: The Omphalos competition. *Machine Learning*, 66(1):93–110, 2007.
- [1354] A. Clark. Distributional learning of some context-free languages with a minimally adequate teacher. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 24–37. Springer, 2010.
- [1355] A. Clark. Learning context free grammars with the syntactic concept lattice. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 38–51. Springer, 2010.
- [1356] A. Clark. Towards general algorithms for grammatical inference. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 11–30, Berlin, Heidelberg, New York, 2010. Springer.
- [1357] A. Clark. Testing distributional properties of context-free grammars. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 42–53, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [1358] A. Clark and R. Eyraud. Identification in the limit of substitutable context-free languages. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 283–296. Springer, Oct. 2005.
- [1359] A. Clark, R. Eyraud, and A. Habrard. A polynomial algorithm for the inference of context free languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 29–42. Springer, 2008.
- [1360] A. Clark, C. C. Florêncio, C. Watkins, and M. Serayet. Planar languages and learnability. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 148–160, Berlin, 2006. Springer.
- [1361] A. Clark, M. Kanazawa, and R. Yoshinaka, editors. *International Conference on Grammatical Inference, 17-19 September 2014, Kyoto, Japan*, volume 57 of *Proceedings of Machine Learning Research*. PMLR, 2014.
- [1362] A. Clark, M. Kanazawa, and R. Yoshinaka. Preface. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 1–2, Kyoto, Japan, 17–19 Sept. 2014. PMLR.

- [1363] A. Clark and F. Thollard. PAC-learnability of probabilistic deterministic finite state automata. *Journal of Machine Learning Research*, 5:473–497, 2004.
- [1364] A. Clark and R. Yoshinaka. Beyond semilinearity: Distributional learning of parallel multiple context-free grammars. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 84–96, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1365] P. Clark and S. Matwin. Learning domain theories using abstract background knowledge. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 360–365. Springer-Verlag, 1993.
- [1366] P. Clark and T. Niblett. Induction in noisy domains. In *Progress in Machine Learning—Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 11–30, Bled, Yugoslavia, May 1987.
- [1367] P. Clark and T. Niblett. The CN2 induction algorithm. *Machine Learning*, 3:261–283, 1989.
- [1368] B. Clarke. Comparing Bayes model averaging and stacking when model approximation error cannot be ignored. *Journal of Machine Learning Research*, 4:683–712, 2003.
- [1369] V. Claveau, P. Sébillot, C. Fabre, and P. Bouillon. Learning semantic lexicons from a part-of-speech and semantically tagged corpus using inductive logic programming. *Journal of Machine Learning Research*, 4:493–525, 2003.
- [1370] S. H. Clearwater and Y. Lee. Sensitivity constraints in learning. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 6, pages 81–95. MIT Press, 1994.
- [1371] J. G. Cleary and L. E. Trigg. K\*: an instance-based learner using an entropic distance measure. In *Proc. 12th International Conference on Machine Learning*, pages 108–114. Morgan Kaufmann, 1995.
- [1372] S. Cléménçon, A. Korba, and E. Sibony. Ranking median regression: Learning to order through local consensus. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 212–245. PMLR, 07–09 Apr. 2018.
- [1373] S. Cléménçon and N. Vayatis. Approximation of the optimal ROC curve and a tree-based ranking algorithm. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 22–37, Berlin, Oct. 2008. Springer.
- [1374] S. Cléménçon, G. Lugosi, and N. Vayatis. Ranking and scoring using empirical risk minimization. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 1–15. Springer, 2005.
- [1375] S. Cléménçon and N. Vayatis. Adaptive estimation of the optimal ROC curve and a bipartite ranking algorithm. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 216–231, Berlin/Heidelberg, 2009. Springer.

- [1376] C. W. Codrington. Boosting with confidence information. In *Proc. 18th International Conf. on Machine Learning*, pages 59–65. Morgan Kaufmann, San Francisco, CA, 2001.
- [1377] J. Coelho and R. Grupen. Learning in non-stationary conditions: A control theoretic approach. In *Proc. 17th International Conf. on Machine Learning*, pages 151–158. Morgan Kaufmann, San Francisco, CA, 2000.
- [1378] A. Cohen, T. Hazan, and T. Koren. Tight bounds for bandit combinatorial optimization. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 629–642, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1379] D. M. Cohen, C. Kulikowski, and H. Berman. DEXTER: A system that experiments with choices of training data using expert knowledge in the domain of DNA hydration. *Machine Learning*, 21(1/2):81–101, 1995.
- [1380] E. Cohen. Learning noisy perceptrons by a perceptron in polynomial time. In *Proc. 38th Annu. IEEE Symposium on Foundations of Computer Science*, pages 514–523. IEEE Computer Society Press, Los Alamitos, CA, 1997.
- [1381] N. Cohen, O. Sharir, and A. Shashua. On the expressive power of deep learning: A tensor analysis. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 698–728, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1382] P. R. Cohen and E. A. Feigenbaum, editors. *The Handbook of Artificial Intelligence*, volume 3, chapter XIV: *Learning and Inductive Inference*, pages 324–511. William Kaufman, Inc., Los Altos, California, 1968.
- [1383] P. R. Cohen, T. Oates, N. Adams, and C. R. Beal. Robot baby 2001. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 32–56. Springer, 2001.
- [1384] W. Cohen. Cryptographic limitations on learning one-clause logic programs. In *Proc. of the Tenth National Conference on Artificial Intelligence*, Washington, D.C., 1993.
- [1385] W. Cohen. Pac-learning a restricted class of recursive logic programs. In *Proc. of the Tenth National Conference on Artificial Intelligence*, Washington, D.C., 1993.
- [1386] W. W. Cohen. Abductive explanation-based learning: A solution to the multiple inconsistent explanation problem. *Machine Learning*, 8:167–219, 1992.
- [1387] W. W. Cohen. Desiderata for generalization-to-N algorithms. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 140–150. Springer-Verlag, 1992.
- [1388] W. W. Cohen. Creating a memory of casual relationships. *Machine Learning*, 10:179–183, 1993.
- [1389] W. W. Cohen. Incremental abductive EBL. *Machine Learning*, 15(1):5–24, 1994.
- [1390] W. W. Cohen. Fast effective rule induction. In *Proc. 12th International Conference on Machine Learning*, pages 115–123. Morgan Kaufmann, 1995.

- [1391] W. W. Cohen. Pac-learning recursive logic programs: Efficient algorithms. *Journal of Artificial Intelligence Research*, 2:501–539, 1995.
- [1392] W. W. Cohen. Pac-learning recursive logic programs: Negative results. *Journal of Artificial Intelligence Research*, 2:541–573, 1995.
- [1393] W. W. Cohen. Text categorization and relational learning. In *Proc. 12th International Conference on Machine Learning*, pages 124–132. Morgan Kaufmann, 1995.
- [1394] W. W. Cohen. The dual DFA learning problem: hardness results for programming by demonstration and learning first-order representations. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 29–40, New York, NY, 1996. ACM Press.
- [1395] W. W. Cohen. Hardness results for learning first-order representations and programming by demonstration. *Machine Learning*, 30:57–87, 1998.
- [1396] W. W. Cohen. What can we learn from the web? In *Proc. 16th International Conf. on Machine Learning*, pages 515–521. Morgan Kaufmann, San Francisco, CA, 1999.
- [1397] W. W. Cohen. Automatically extracting features for concept learning from the web. In *Proc. 17th International Conf. on Machine Learning*, pages 159–166. Morgan Kaufmann, San Francisco, CA, 2000.
- [1398] W. W. Cohen. Extracting information from the web for concept learning and collaborative filtering. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 1–12. Springer, 2000.
- [1399] W. W. Cohen and P. Devanbu. A comparative study of inductive logic programming methods for software fault prediction. In *Proc. 14th International Conference on Machine Learning*, pages 66–74. Morgan Kaufmann, 1997.
- [1400] W. W. Cohen, R. Greiner, and D. Schuurmans. Probabilistic hill-climbing. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 11, pages 171–181. MIT Press, 1994.
- [1401] W. W. Cohen and H. Hirsh. Learnability of description logics. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 116–127, New York, NY, 1992. ACM Press.
- [1402] W. W. Cohen and H. Hirsh. The learnability of description logics with equality constraints. *Machine Learning*, 17(2/3):169–199, 1994. Special issue for COLT '92.
- [1403] W. W. Cohen and H. Hirsh. Learning the CLASSIC description logic: Theoretical and experimental results. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth International Conference (KR94)*. Morgan Kaufmann, 1994.
- [1404] W. W. Cohen and H. Hirsh. Corrigendum for: Learnability of description logics. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, page 463, New York, NY, 1995. ACM Press.
- [1405] W. W. Cohen, R. E. Schapire, and Y. Singer. Learning to order things. *Journal of Artificial Intelligence Research*, 10:243–270, 1999.
- [1406] D. Cohn, L. Atlas, and R. Ladner. Improving generalization with active learning. *Machine Learning*, 15(2):201–221, 1994.



- [1407] D. Cohn and H. Chang. Learning to probabilistically identify authoritative documents. In *Proc. 17th International Conf. on Machine Learning*, pages 167–174. Morgan Kaufmann, San Francisco, CA, 2000.
- [1408] M. Collins. Discriminative reranking for natural language parsing. In *Proc. 17th International Conf. on Machine Learning*, pages 175–182. Morgan Kaufmann, San Francisco, CA, 2000.
- [1409] M. Collins. Tutorial: Machine learning methods in natural language processing. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 655–655. Springer, 2003.
- [1410] M. Collins, R. E. Schapire, and Y. Singer. Logistic regression, AdaBoost and Bregman distances. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 158–169. Morgan Kaufmann, San Francisco, 2000.
- [1411] M. Collins, R. E. Schapire, and Y. Singer. Logistic regression, AdaBoost and Bregman distances. *Machine Learning*, 48(1/3):253–285, 2002.
- [1412] S. Colton, A. Bundy, and T. Walsh. Automatic identification of mathematical concepts. In *Proc. 17th International Conf. on Machine Learning*, pages 183–190. Morgan Kaufmann, San Francisco, CA, 2000.
- [1413] F. D. Comité, F. Denis, R. Gilleron, and F. Letouzey. Positive and unlabeled examples help learning. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 219–230. Springer, 1999.
- [1414] L. Comminges and A. S. Dalalyan. Tight conditions for consistent variable selection in high dimensional nonparametric regression. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 187–206, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1415] V. Conitzer and T. Sandholm. AWESOME: A general multiagent learning algorithm that converges in self-play and learns a best response against stationary opponents. *Machine Learning*, 67(1-2):23–43, 2007.
- [1416] D. Conklin. Machine discovery of protein motifs. *Machine Learning*, 21(1/2):125–150, 1995.
- [1417] D. Conklin. Melodic analysis with segment classes. *Machine Learning*, 65(2-3):349–360, 2006.
- [1418] D. Conklin and I. H. Witten. Complexity-based induction. *Machine Learning*, 16(3):203–226, 1994.
- [1419] J. Conradt, G. Tevatia, S. Vijayakumar, and S. Schaal. On-line learning for humanoid robot systems. In *Proc. 17th International Conf. on Machine Learning*, pages 191–198. Morgan Kaufmann, San Francisco, CA, 2000.
- [1420] C. M. Cook, A. Rosenfeld, and A. R. Aronson. Grammatical inference by hill-climbing. *Inform. Sci.*, 10:59–80, 1976.
- [1421] D. J. Cook. Defining the limits of analogical planning. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 4, pages 65–80. MIT Press, 1994.

- [1422] G. F. Cooper and E. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347, 1992.
- [1423] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 1990.
- [1424] A. Cornuéjols. Getting order independence in incremental learning. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 196–212. Springer-Verlag, 1993.
- [1425] C. Cortes, G. DeSalvo, and M. Mohri. Learning with rejection. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 67–82. Springer, 2016.
- [1426] C. Cortes, P. Haffner, and M. Mohri. Positive definite rational kernels. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 41–56. Springer, 2003.
- [1427] C. Cortes, L. Kontorovich, and M. Mohri. Learning languages with rational kernels. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13–15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 349–364, Berlin, 2007. Springer.
- [1428] C. Cortes, V. Kuznetsov, M. Mohri, and M. Warmuth. On-line learning algorithms for path experts with non-additive losses. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 424–447, Paris, France, 03–06 July 2015. PMLR.
- [1429] C. Cortes and M. Mohri. Domain adaptation in regression. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 308–323, Berlin, Heidelberg, New York, 2011. Springer.
- [1430] C. Cortes and M. Mohri. Domain adaptation and sample bias correction theory and algorithm for regression. *Theoret. Comput. Sci.*, 519:103–126, 2014. Special Issue: Algorithmic Learning Theory.
- [1431] C. Cortes, M. Mohri, M. Riley, and A. Rostamizadeh. Sample selection bias correction theory. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 38–53, Berlin, Oct. 2008. Springer.
- [1432] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.
- [1433] D. Cossock and T. Zhang. Subset ranking using regression. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 605–619, Berlin, 2006. Springer.
- [1434] S. Cost and S. Salzberg. A weighted nearest neighbor algorithm for learning with symbolic features. *Machine Learning*, 10:57–78, 1993.

- [1435] V. S. Costa, A. Srinivasan, R. Camacho, H. Blockeel, B. Demoen, G. Janssens, J. Struyf, H. Vandecasteele, and W. V. Laer. Query transformations for improving the efficiency of ILP systems. *Journal of Machine Learning Research*, 4:465–491, 2003.
- [1436] F. Coste and D. Fredouille. Efficient ambiguity detection in C-NFA, a step towards the inference on non deterministic automata. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 25–38. Springer, 2000.
- [1437] F. Coste, D. Fredouille, C. Kermorvant, and C. de la Higuera. Introducing domain and typing bias in automata inference. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 115–126. Springer, 2004.
- [1438] F. Coste, G. Garet, and J. Nicolas. Locally substitutable languages for enhanced inductive leaps. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 97–111, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1439] F. Coste, G. Garet, and J. Nicolas. A bottom-up efficient algorithm learning substitutable languages from positive examples. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 49–63, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [1440] F. Coste and J. Nicolas. How considering incompatible state mergins may reduce the DFA induction search tree. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 199–210. Springer, 1998.
- [1441] A. Cotter, M. Gupta, and J. Pfeifer. A light touch for heavily constrained SGD. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 729–771, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1442] R. Coulom. Feedforward neural networks in reinforcement learning applied to high-dimensional motor control. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 403–413. Springer, 2002.
- [1443] D. Coulon and D. Kayser. Construction of natural language sentence acceptors by a supervised learning technique. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-1:94–99, 1979.
- [1444] T. Cover. Behavior of sequential predictors of binary sequences. In *Proc. of the 4th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 263–272. Publishing House of the Czechoslovak Academy of Sciences, 1965.
- [1445] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [1446] T. M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. on Electronic Computers*, EC-14:326–334, 1965.

- [1447] T. M. Cover. Learning in pattern recognition. In *Methodologies of Pattern Recognition*, pages 111–132. Academic Press, 1969.
- [1448] T. M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- [1449] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions in Information Theory*, IT-13(1):21–27, Jan. 1967.
- [1450] T. M. Cover and T. J. Wagner. Topics in statistical pattern recognition. In *Communication and Cybernetics: Digital Pattern Recognition*, volume 10, chapter 2. Springer-Verlag, 1976.
- [1451] I. J. Cox and J. J. Leonard. Unsupervised learning for mobile robot navigation using probabilistic data association. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 18, pages 297–319. MIT Press, 1994.
- [1452] K. Cramer and Y. Singer. On the learnability and design of output codes for multiclass problems. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 35–46. Morgan Kaufmann, San Francisco, 2000.
- [1453] K. Crammer. Online tracking of linear subspaces. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 438–452, Berlin, 2006. Springer.
- [1454] K. Crammer and Y. Singer. On the algorithmic implementation of multiclass kernel-based vector machines. *Journal of Machine Learning Research*, 2:265–292, 2001.
- [1455] K. Crammer and Y. Singer. Ultraconservative online algorithms for multiclass problems. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 99–115. Springer, 2001.
- [1456] K. Crammer and Y. Singer. On the learnability and design of output codes for multiclass problems. *Machine Learning*, 47(2/3):201–233, 2002.
- [1457] K. Crammer and Y. Singer. A family of additive online algorithms for category ranking. *Journal of Machine Learning Research*, 3:1025–1058, 2003.
- [1458] K. Crammer and Y. Singer. Learning algorithms for enclosing points in Bregmanian spheres. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 388–402. Springer, 2003.
- [1459] K. Crammer and Y. Singer. Ultraconservative online algorithms for multiclass problems. *Journal of Machine Learning Research*, 3:951–991, 2003.
- [1460] K. Crammer and Y. Singer. Loss bounds for online category ranking. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 48–62. Springer, 2005.
- [1461] M. Craven, D. Page, J. Shavlik, J. Bockhorst, and J. Glasner. Using multiple levels of learning and diverse evidence sources to uncover coordinately controlled genes. In *Proc. 17th International Conf. on Machine Learning*, pages 199–206. Morgan Kaufmann, San Francisco, CA, 2000.

- [1462] M. Craven and S. Slattery. Relational learning with statistical predicate invention: Better models for hypertext. *Machine Learning*, 43(1/2):97–119, 2001.
- [1463] M. W. Craven and J. W. Shavlik. Using sampling and queries to extract rules from trained neural networks. In *Proc. 11th International Conference on Machine Learning*, pages 37–45. Morgan Kaufmann, 1994.
- [1464] M. W. Craven and J. W. Shavlik. Investigating the value of a good input representation. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 16, pages 307–321. MIT Press, 1995.
- [1465] S. Craw and P. Hutton. Protein folding: symbolic refinement competes with neural networks. In *Proc. 12th International Conference on Machine Learning*, pages 133–141. Morgan Kaufmann, 1995.
- [1466] S. Crespi-Reghizzi. An effective model for grammar inference. In *Information Processing 71*, pages 524–529. North-Holland Publishing Co., 1972.
- [1467] S. Crespi-Reghizzi and D. Mandrioli. Inferring grammars by means of profiles: a unifying view. Internal report, Istituto di Elettrotecnica ed Elettronica del Politecnico di Milano, 1980.
- [1468] M. M. A. Crespi-Reghizzi, S and L. Lichten. The use of grammatical inference for designing programming languages. *Commun. ACM*, 16:83–90, 1973.
- [1469] N. Cristianini, C. Campbell, and C. Burges. Editorial: Kernel methods: Current research and future directions. *Machine Learning*, 46(1-3):5–9, 2002.
- [1470] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press, Cambridge, U.K., 2000.
- [1471] N. Cristianini, J. Shawe-Taylor, and H. Lodhi. Latent semantic kernels. In *Proc. 18th International Conf. on Machine Learning*, pages 66–73. Morgan Kaufmann, San Francisco, CA, 2001.
- [1472] N. Cristianini, J. Shawe-Taylor, and P. Sykacek. Bayesian classifiers are large margin hyperplanes in a Hilbert space. In *Proc. 15th International Conf. on Machine Learning*, pages 109–117. Morgan Kaufmann, San Francisco, CA, 1998.
- [1473] N. Cristianini, J. Shawe-Taylor, and R. C. Williamson. Introduction to the special issue on kernel methods. *Journal of Machine Learning Research*, 2:95–96, 2001.
- [1474] R. H. Crites and A. G. Barto. Elevator group control using multiple reinforcement learning agents. *Machine Learning*, 33(2/3):235–262, 1998.
- [1475] P. P. Cruz-Alcázar and E. Vidal-Ruiz. Learning regular grammars to model music style: comparing different coding schemes. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 211–222. Springer, 1998.
- [1476] M. Cryan, L. A. Goldberg, and P. W. Goldberg. Evolutionary trees can be learned in polynomial time in the two-state general Markov model. *SIAM J. Comput.*, 31(2):375–397, 2001.

- [1477] B. C. Csáji, R. M. Jungers, and V. D. Blondel. Pagerank optimization in polynomial time by stochastic shortest path reformulation. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 89–103, Berlin, Heidelberg, New York, 2010. Springer.
- [1478] D. Csiba and P. Richtárik. Coordinate descent faceoff: Primal or dual? In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 246–267. PMLR, 07–09 Apr. 2018.
- [1479] I. Csiszár. On iterative algorithms with an information geometry background. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, page 1, Berlin, Oct. 2008. Springer.
- [1480] J. R. Cuellar and H. U. Simon. Neural discriminant analysis. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 223–236. Springer-Verlag, 1993.
- [1481] R. Cummings, S. Ioannidis, and K. Ligett. Truthful linear regression. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 448–483, Paris, France, 03–06 July 2015. PMLR.
- [1482] R. Cummings, K. Ligett, K. Nissim, A. Roth, and Z. S. Wu. Adaptive learning with robust generalization guarantees. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 772–814, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1483] P. Cunningham and J. Carney. Diversity versus quality in classification ensembles based on feature selection. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 109–116. Springer, 2000.
- [1484] J. Cussens. Bayes and Pseudo-Bayes estimates of conditional probabilities and their reliability. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 136–152. Springer-Verlag, 1993.
- [1485] J. Cussens. A Bayesian analysis of algorithms for learning finite functions. In *Proc. 12th International Conference on Machine Learning*, pages 142–149. Morgan Kaufmann, 1995.
- [1486] J. Cussens. Parameter estimation in stochastic logic programs. *Machine Learning*, 44(3):245–271, 2001.
- [1487] J. Cussens and A. M. Frisch. Introduction to the special issue on inductive logic programming. *Journal of Machine Learning Research*, 4:413–414, 2003.
- [1488] A. Cutkosky and K. Boahen. Online learning without prior information. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 643–677, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [1489] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control*, 2:303–314, 1989.
- [1490] W. Daelemans, A. van den Bosch, and J. Zavrel. Forgetting exceptions is harmful in language learning. *Machine Learning*, 34:11–41, 1999.
- [1491] I. Dagan and S. P. Engelson. Committee-based sampling for training probabilistic classifiers. In *Proc. 12th International Conference on Machine Learning*, pages 150–157. Morgan Kaufmann, 1995.
- [1492] I. Dagan, L. Lee, and F. C. N. Pereira. Similarity-based models of word cooccurrence probabilities. *Machine Learning*, 34:43–69, 1999.
- [1493] Y. Dagan and C. Koby. A better resource allocation algorithm with semi-bandit feedback. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 268–320. PMLR, 07–09 Apr. 2018.
- [1494] F. A. Dahl. A reinforcement learning algorithm applied to simplified two-player texas hold'em poker. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 85–96. Springer, 2001.
- [1495] F. A. Dahl. The lagging anchor algorithm: Reinforcement learning in two-player zero-sum games with imperfect information. *Machine Learning*, 49(1):5–37, 2002.
- [1496] F. A. Dahl and O. M. Halck. Minimax TD-learning with neural nets in a Markov game. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 117–128. Springer, 2000.
- [1497] F. A. Dahl and O. M. Halck. Learning while exploring: Bridging the gaps in the eligibility traces. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 73–84. Springer, 2001.
- [1498] A. Dalalyan. Further and stronger analogy between sampling and optimization: Langevin monte carlo and gradient descent. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 678–689, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1499] A. S. Dalalyan and J. Salmon. Competing against the best nearest neighbor filter in regression. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 129–143, Berlin, Heidelberg, New York, 2011. Springer.
- [1500] A. S. Dalalyan and A. B. Tsybakov. Aggregation by exponential weighting and sharp oracle inequalities. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 97–111, Berlin, 2007. Springer.
- [1501] A. S. Dalalyan and A. B. Tsybakov. Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. Syst. Sci.*, 78(5):1423–1443, 2012.

- [1502] R. Daley. On the inference of optimal descriptions. *Theoret. Comput. Sci.*, 4(3):301–319, 1977.
- [1503] R. Daley. Transformation of probabilistic learning strategies into deterministic learning strategies. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 220–226, San Mateo, CA, 1988. Morgan Kaufmann.
- [1504] R. Daley. Multi-agent learning: Theoretical and empirical studies. In *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardtsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 175–183. Springer-Verlag, 1993.
- [1505] R. Daley and B. Kalyanasundaram. Capabilities of probabilistic learners with bounded mind changes. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 182–191, New York, NY, 1993. ACM Press.
- [1506] R. Daley and B. Kalyanasundaram. Use of reduction arguments in determining Popperian FIN-type learning capabilities. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 173–186. Springer-Verlag, 1993.
- [1507] R. Daley and B. Kalyanasundaram. Towards reduction arguments for FINite learning. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 63–75. Springer, 1995.
- [1508] R. Daley and B. Kalyanasundaram. FINite learning capabilities and their limits. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 81–89, New York, NY, 1997. ACM Press.
- [1509] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. The power of probabilism in Popperian FINite learning (extended abstract). In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 151–169. Springer-Verlag, 1992.
- [1510] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. Capabilities of fallible FINite learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 199–208, New York, NY, 1993. ACM Press.
- [1511] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. The power of probabilism in Popperian FINite learning. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):41–62, 1994. Special issue for AII '92.
- [1512] R. Daley, B. Kalyanasundaram, and M. Velauthapillai. Breaking the probability 1/2 barrier in FIN-type learning. *J. of Comput. Syst. Sci.*, 50(3):574–599, 1995. Earlier version in 5th COLT, 1992.
- [1513] R. Daley, L. Pitt, M. Velauthapillia, and T. Will. Relations between probabilistic and team one-shot learners. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 228–239, San Mateo, CA, 1991. Morgan Kaufmann.
- [1514] R. P. Daley. On the error correcting power of pluralism in BC-type inductive inference. *Theoret. Comput. Sci.*, 24(1):95–104, 1983.



- [1515] R. P. Daley. Inductive inference hierarchies: Probabilistic VS pluralistic strategies. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 73–82. Springer-Verlag, 1986.
- [1516] R. P. Daley. Towards the development of an analysis of learning algorithms. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 1–18. Springer-Verlag, 1986.
- [1517] R. P. Daley and C. H. Smith. On the complexity of inductive inference. *Inform. Control*, 69(1-3):12–40, 1986.
- [1518] V. Dalmau. Boolean formulas are hard to learn for most gate bases. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 301–312. Springer, 1999.
- [1519] V. Dalmau. A dichotomy theorem for learning quantified Boolean formulas. *Machine Learning*, 35(3):207–224, 1999.
- [1520] V. Dalmau and P. Jeavons. Learnability of quantified formula. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 63–78. Springer, 1999.
- [1521] P. Damaschke. Computational aspects of parallel attribute-efficient learning. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 103–111. Springer, 1998.
- [1522] P. Damaschke. Adaptive versus nonadaptive attribute-efficient learning. *Machine Learning*, 41(2):197–215, 2000.
- [1523] P. Damaschke. Parallel attribute-efficient learning of monotone Boolean functions. In *Algorithm Theory - SWAT 2000, 7th Scandinavian Workshop on Algorithm Theory, Bergen, Norway, July 5–7, 2000, Proceedings*, volume 1851 of *Lecture Notes in Computer Science*, pages 504–512. Springer, 2000.
- [1524] P. Damaschke. On parallel attribute-efficient learning. *J. Comput. Syst. Sci.*, 67(1):46–62, 2003.
- [1525] V. Dani, T. P. Hayes, and S. M. Kakade. Stochastic linear optimization under bandit feedback. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 355–366. Omnipress, 2008.
- [1526] A. Daniely. A PTAS for agnostically learning halfspaces. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 484–502, Paris, France, 03–06 July 2015. PMLR.
- [1527] A. Daniely. Depth separation for neural networks. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 690–696, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [1528] A. Daniely and T. Helbertal. The price of bandit information in multiclass online classification. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 93–104. JMLR.org, 2013.
- [1529] A. Daniely, N. Linial, and S. Shalev-Shwartz. The complexity of learning halfspaces using generalized linear methods. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 244–286, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1530] A. Daniely, S. Sabato, S. Ben-David, and S. Shalev-Shwartz. Multiclass learnability and the ERM principle. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 207–232, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1531] A. Daniely and S. Shalev-Shwartz. Optimal learners for multiclass problems. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 287–316, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1532] A. Daniely and S. Shalev-Shwartz. Complexity theoretic limitations on learning DNF’s. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 815–830, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1533] E. E. Dar and Y. Mansour;. Learning rates for  $Q$ -learning. *Journal of Machine Learning Research*, 5:1–25, 2003.
- [1534] L. Darden. Discovering mechanisms: A computational philosophy of science perspective. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, page 57. Springer, 2001.
- [1535] C. Darken, M. Donahue, L. Gurvits, and E. Sontag. Rate of approximation results motivated by robust neural network learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 303–309, New York, NY, 1993. ACM Press.
- [1536] M. Darnstädt, T. Doliwa, H. U. Simon, and S. Zilles. Order compression schemes. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 173–187, Berlin, Heidelberg, New York, 2013. Springer.
- [1537] M. Darnstädt, T. Kiss, H. U. Simon, and S. Zilles. Order compression schemes. *Theoret. Comput. Sci.*, 620:73–90, 2016. Special Issue: Algorithmic Learning Theory.
- [1538] M. Darnstädt, C. Ries, and H. U. Simon. Hierarchical design of fast minimum disagreement algorithms. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 134–148, Berlin, Heidelberg, New York, 2015. Springer.
- [1539] M. Darnstädt, C. Ries, and H. U. Simon. Hierarchical design of fast Minimum Disagreement algorithms. *Theoret. Comput. Sci.*, 716:70–88, 2018. Special Issue on ALT 2015.

- [1540] M. Darnstädt and H. U. Simon. Smart PAC-learners. *Theoret. Comput. Sci.*, 412(19):1756–1766, 2011.
- [1541] M. Darnstädt, H. U. Simon, and B. Szörényi. Supervised learning and co-training. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 425–439, Berlin, Heidelberg, New York, 2011. Springer.
- [1542] M. Darnstädt, H. U. Simon, and B. Szörényi. Supervised learning and co-training. *Theoret. Comput. Sci.*, 519:68–87, 2014. Special Issue: Algorithmic Learning Theory.
- [1543] S. Das. Filters, wrappers and a boosting-based hybrid for feature selection. In *Proc. 18th International Conf. on Machine Learning*, pages 74–81. Morgan Kaufmann, San Francisco, CA, 2001.
- [1544] G. Dasarathy, R. Nowak, and X. Zhu. S2: An efficient graph based active learning algorithm with application to nonparametric classification. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 503–522, Paris, France, 03–06 July 2015. PMLR.
- [1545] B. DasGupta and B. Hammer. On approximate learning by multi-layered feedforward circuits. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 264–278. Springer, 2000.
- [1546] B. DasGupta and B. Hammer. On approximate learning by multi-layered feedforward circuits. *Theoret. Comput. Sci.*, 348(1):95–127, 2005. Special issue for ALT 2000.
- [1547] B. DasGupta, H. T. Siegelmann, and E. Sontag. On a learnability question associated to neural networks with continuous activations. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 47–56, New York, NY, 1994. ACM Press.
- [1548] B. DasGupta, H. T. Siegelmann, and E. Sontag. On the complexity of training neural networks with continuous activation functions. *IEEE Transactions on Neural Networks*, 6(6):1490–1504, 1995.
- [1549] S. Dasgupta. The sample complexity of learning fixed-structure Bayesian networks. *Machine Learning*, 29(2-3):165–180, 1997.
- [1550] S. Dasgupta. Performance guarantees for hierarchical clustering. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 351–363. Springer, 2002.
- [1551] S. Dasgupta. How fast is k-means? In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 735–735. Springer, 2003.
- [1552] S. Dasgupta. Subspace detection: A robust statistics formulation. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 734–734. Springer, 2003.
- [1553] S. Dasgupta. Foreword, special issue on learning theory. *J. Comput. Syst. Sci.*, 74(1):1–1, 2008.

- [1554] S. Dasgupta. Two faces of active learning. *Theoret. Comput. Sci.*, 412(19):1767–1781, 2011.
- [1555] S. Dasgupta. Consistency of nearest neighbor classification under selective sampling. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 18.1–18.15, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1556] S. Dasgupta and D. Hsu. On-line estimation with the multivariate gaussian distribution. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 278–292, Berlin, 2007. Springer.
- [1557] S. Dasgupta, A. T. Kalai, and C. Monteleoni. Analysis of perceptron-based active learning. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 249–263. Springer, 2005.
- [1558] S. Dasgupta, W. S. Lee, and P. M. Long. A theoretical analysis of query selection for collaborative filtering. *Machine Learning*, 51(3):283–298, 2003.
- [1559] S. Dasgupta and P. M. Long. Boosting with diverse base classifiers. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 273–287. Springer, 2003.
- [1560] S. Dasgupta and P. M. Long. Performance guarantees for hierarchical clustering. *J. Comput. Syst. Sci.*, 70(4):555–569, 2005. Special issue on COLT 2002.
- [1561] S. Dasgupta, E. Pavlov, and Y. Singer. An efficient PAC algorithm for reconstructing a mixture of lines. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 351–364. Springer, 2002.
- [1562] S. Dasgupta and K. Sinha. Randomized partition trees for exact nearest neighbor search. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 317–337. JMLR.org, 2013.
- [1563] M. Dashevskiy. Aggregating algorithm for a space of analytic functions. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 214–226, Berlin, Oct. 2008. Springer.
- [1564] M. Dashevskiy and Z. Luo. Prediction of long-range dependent time series data with performance guarantee. In O. Watanabe and T. Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo, Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, pages 31–45, Berlin/Heidelberg, 2009. Springer.
- [1565] C. Daskalakis and G. Kamath. Faster and sample near-optimal algorithms for proper learning mixtures of gaussians. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1183–1213, Barcelona, Spain, 13–15 June 2014. PMLR.

- [1566] C. Daskalakis and Q. Pan. Square hellinger subadditivity for Bayesian networks and its applications to identity testing. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 697–703, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1567] C. Daskalakis, C. Tzamos, and M. Zampetakis. Ten steps of EM suffice for mixtures of two gaussians. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 704–710, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1568] A. Datta, A. Choudhary, M. L. Bittner, and E. R. Dougherty. External control in markovian genetic regulatory networks. *Machine Learning*, 52(1-2):169–191, 2003.
- [1569] P. Datta and D. Kibler. Learning prototypical concept descriptions. In *Proc. 12th International Conference on Machine Learning*, pages 158–166. Morgan Kaufmann, 1995.
- [1570] P. Datta and D. Kibler. Learning symbolic prototypes. In *Proc. 14th International Conference on Machine Learning*, pages 75–82. Morgan Kaufmann, 1997.
- [1571] H. Daumé III, J. M. Phillips, A. Saha, and S. Venkatasubramanian. Efficient protocols for distributed classification and optimization. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 154–168, Berlin, Heidelberg, New York, 2012. Springer.
- [1572] G. David and Z. Ilias. High dimensional regression with binary coefficients. estimating squared error and a phase transtition. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 948–953, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1573] C. d’Avignon and D. Geman. Tree-structured neural decoding. *Journal of Machine Learning Research*, 4:743–754, 2003.
- [1574] A. P. Dawid. Prequential analysis, stochastic complexity and Bayesian inference. *Bayesian Statistics*, 109:109–125, 1992.
- [1575] P. Dayan. The convergence of TD( $\lambda$ ) for general  $\lambda$ . *Machine Learning*, 8:341–362, 1992.
- [1576] P. Dayan and T. J. Sejnowski. TD( $\lambda$ ) converges with probability 1. *Machine Learning*, 14(3):295–301, 1994.
- [1577] P. Dayan and T. J. Sejnowski. Exploration bonuses and dual control. *Machine Learning*, 25:5–22, 1996.
- [1578] L. De Raedt. Declarative modeling for machine learning and data mining. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 12–12, Berlin, Heidelberg, New York, 2012. Springer.
- [1579] L. De Raedt and M. Bruynooghe. Interactive concept-learning and constructive induction by analogy. *Machine Learning*, 8:107–150, 1992.
- [1580] L. De Raedt and L. Dehaspe. Clausal discovery. *Machine Learning*, 26(2-3):99–146, 1997.

- [1581] L. De Raedt and P. Flach, editors. *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*. Springer, 2001.
- [1582] L. De Raedt and W. Van Laer. Inductive constraint logic. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 80–94. Springer, 1995.
- [1583] T. Dean, D. Angluin, K. Basye, S. Engelson, L. Kaelbling, E. Kokkevis, and O. Maron. Inferring finite automata with stochastic output functions and an application to map learning. *Machine Learning*, 18(1):81–108, 1995. First appeared in Proc. of AAAI-92.
- [1584] R. Dearden. Structured prioritized sweeping. In *Proc. 18th International Conf. on Machine Learning*, pages 82–89. Morgan Kaufmann, San Francisco, CA, 2001.
- [1585] L. Dębowski. Computable Bayesian compression for uniformly discretizable statistical models. In R. Gavalda, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 53–67, Berlin/Heidelberg, 2009. Springer.
- [1586] S. Decatur. PAC learning with constant-partition classification noise and applications to decision tree induction. In *Proc. 14th International Conference on Machine Learning*, pages 83–91. Morgan Kaufmann, 1997.
- [1587] S. E. Decatur. Statistical queries and faulty PAC oracles. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 262–268, New York, NY, 1993. ACM Press.
- [1588] S. E. Decatur and R. Gennaro. On learning from noisy and incomplete examples. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 353–360, New York, NY, 1995. ACM Press.
- [1589] S. E. Decatur, O. Goldreich, and D. Ron. Computational sample complexity. *SIAM J. Comput.*, 29(3):854–879, 1999.
- [1590] R. Dechter. On the design of networks with hidden variables. Technical report, Technion, Haifa, July 1990.
- [1591] G. Deco and B. Schürmann. Dynamic modeling of chaotic time series by neural networks. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 9, pages 137–153. MIT Press, 1997.
- [1592] D. Decoste and B. Schölkopf. Training invariant support vector machines. *Machine Learning*, 46(1-3):161–190, 2002.
- [1593] A.-H. Dediu, J. M. Matos, and C. Moraga. Query learning automata with helpful labels. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 15–29, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [1594] B. Defourny, D. Ernst, and L. Wehenkel. Bounds for multistage stochastic programs using supervised learning strategies. In O. Watanabe and T. Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo*,

- Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, pages 61–73, Berlin/Heidelberg, 2009. Springer.
- [1595] C. Dégremont and N. Gierasimczuk. Finite identification from the viewpoint of epistemic update. *Inform. Comput.*, 209(3):383–396, 2011. Special Issue: 3rd International Conference on Language and Automata Theory and Applications (LATA 2009).
- [1596] H. Déjean. Learning rules and their exceptions. *Journal of Machine Learning Research*, 2:669–693, 2002.
- [1597] G. DeJong. A case study of explanation-based control. In *Proc. 12th International Conference on Machine Learning*, pages 167–175. Morgan Kaufmann, 1995.
- [1598] G. DeJong. Hidden strengths and limitations: An empirical investigation of reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 215–222. Morgan Kaufmann, San Francisco, CA, 2000.
- [1599] G. Dejong and R. Mooney. Explanation-based learning: An alternative view. *Machine Learning*, 1(2):145–176, 1986.
- [1600] O. Dekel, J. Ding, T. Koren, and Y. Peres. Online learning with composite loss functions. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1214–1231, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1601] O. Dekel and Y. S. P. M. Long. Online multitask learning. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 453–467, Berlin, 2006. Springer.
- [1602] O. Dekel, S. Shalev-Shwartz, and Y. Singer. Smooth  $\epsilon$ -insensitive regression by loss symmetrization. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 433–447. Springer, 2003.
- [1603] O. Dekel, S. Shalev-Shwartz, and Y. Singer. Smooth  $\epsilon$ -insensitive regression by loss symmetrization. *Journal of Machine Learning Research*, 6:711–741, 2005.
- [1604] A. Delhay and L. Miclet. Analogical equations in sequences: Definition and resolution. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 127–138. Springer, 2004.
- [1605] S. Deligne, F. Yvon, and F. Bimbot. Introducing statistical dependencies and structural constraints in variable-length sequence models. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25–27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 156–167. Springer, 1996.
- [1606] F. Dellaert, S. M. Seitz, C. E. Thorpe, and S. Thrun. EM, MCMC, and chain flipping for structure from motion with unknown correspondence. *Machine Learning*, 50(1-2):45–71, 2003.
- [1607] A. Demiriz, K. P. Bennett, and J. Shawe-Taylor. Linear programming boosting via column generation. *Machine Learning*, 46(1-3):225–254, 2002.

- [1608] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum-likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, B39:1–38, 1977.
- [1609] X. Deng, T. Kameda, and C. Papadimitriou. How to learn in an unknown environment. In *Proc. of the 32nd Symposium on the Foundations of Comp. Sci.*, pages 298–303. IEEE Computer Society Press, Los Alamitos, CA, 1991.
- [1610] X. Deng, T. Kameda, and C. Papadimitriou. How to learn an unknown environment. I: the rectilinear case. *J. ACM*, 45(2):215–245, 1998.
- [1611] X. Deng and C. H. Papadimitriou. Exploring an unknown graph. In *Proc. 31th Annu. IEEE Sympos. Found. Comput. Sci.*, volume I, pages 355–361, 1990.
- [1612] F. Denis. PAC learning from positive statistical queries. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 112–126. Springer, 1998.
- [1613] F. Denis. Finding a minimal 1-DNF consistent with a positive sample is LOGSNP-complete. *Inform. Process. Lett.*, 69(1):1–5, 1999.
- [1614] F. Denis. Learning regular languages from simple positive examples. *Machine Learning*, 44(1/2):37–66, 2001.
- [1615] F. Denis, C. D’Halluin, and R. Gilleron. PAC learning with simple examples. In *Proc. 13th Annual Symposium on Theoretical Aspects of Computer Science: STACS’96*, volume 1046 of *Lecture Notes in Computer Science*, pages 231–242. Springer, 1996.
- [1616] F. Denis and Y. Esposito. Learning classes of probabilistic automata. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 124–139. Springer, 2004.
- [1617] F. Denis, Y. Esposito, and A. Habrard. Learning rational stochastic languages. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 274–288, Berlin, 2006. Springer.
- [1618] F. Denis, É. Gilbert, A. Habrard, F. Ouardi, and M. Tommasi. Relevant representations for the inference of rational stochastic tree languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 57–70. Springer, 2008.
- [1619] F. Denis and R. Gilleron. PAC learning under helpful distributions. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 132–145. Springer, 1997.
- [1620] F. Denis and R. Gilleron. PAC learning under helpful distributions. *Theoretical Informatics and Applications*, 35(2):129–148, 2001.
- [1621] F. Denis, R. Gilleron, and F. Letouzey. Learning from positive and unlabeled examples. *Theoret. Comput. Sci.*, 348(1):70–83, 2005. Special issue for ALT 2000.
- [1622] F. Denis and A. Habrard. Learning rational stochastic tree languages. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 242–256, Berlin, Oct. 2007. Springer.



- [1623] F. Denis, A. Lemay, and A. Terlutte. Learning regular languages using non deterministic finite automata. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 39–50. Springer, 2000.
- [1624] F. Denis, A. Lemay, and A. Terlutte. Learning regular languages using RFSA. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 348–363. Springer, 2001.
- [1625] F. Denis, A. Lemay, and A. Terlutte. Some classes of regular languages identifiable in the limit from positive data. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 63–76. Springer, 2002.
- [1626] F. Denis, A. Lemay, and A. Terlutte. Learning regular languages using RFSA. *Theoret. Comput. Sci.*, 313(2):267–294, 2004. Special issue for ALT 2001.
- [1627] J. Denker, D. Schwartz, B. Wittner, S. Solla, R. Howard, L. Jackel, and J. Hopfield. Automatic learning, rule extraction and generalization. *Complex Systems*, 1:877–922, 1987.
- [1628] T. M. C. F. for Deriving Perceptron Based Linear Classifiers. Nader H. bshouty and catherine A. haddad-zaknoon. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 207–222. Springer, 2016.
- [1629] N. Dershowitz. From specifications to programs: Induction in the service of synthesis. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 6–7. Springer-Verlag, 1994.
- [1630] G. DeSalvo, M. Mohri, and U. Syed. Learning with deep cascades. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 254–269. Berlin, Heidelberg, New York, 2015. Springer.
- [1631] A. DeSantis, G. Markowsky, and M. N. Wegman. Learning probabilistic prediction functions. In *Proc. 29th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 110–119. IEEE Computer Society Press, Los Alamitos, CA, 1988.
- [1632] Y. Deshpande and A. Montanari. Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 523–562, Paris, France, 03–06 July 2015. PMLR.
- [1633] M. desJardins. Evaluation of learning biases using probabilistic domain knowledge. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 7, pages 97–112. MIT Press, 1994.

- [1634] I. Desyatnikov and R. Meir. Data-dependent bounds for multi-category classification based on convex losses. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 159–172. Springer, 2003.
- [1635] M. Devaney and A. Ram. Efficient feature selection in conceptual clustering. In *Proc. 14th International Conference on Machine Learning*, pages 92–97. Morgan Kaufmann, 1997.
- [1636] L. Devroye. Automatic pattern recognition: A study of the probability of error. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10(4):530–543, July 1988.
- [1637] L. Devroye. Random multivariate search trees. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 1–1, Berlin, 2006. Springer.
- [1638] L. Devroye, G. Lugosi, and G. Neu. Prediction by random-walk perturbation. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 460–473. JMLR.org, 2013.
- [1639] L. P. Devroye and T. J. Wagner. A distribution-free performance bound in error estimation. *IEEE Trans. Inform. Theory*, IT-22:586–587, 1976.
- [1640] L. P. Devroye and T. J. Wagner. Distribution-free performance bounds with the resubstitution error estimate. *IEEE Trans. Inform. Theory*, IT-25:208–210, 1979.
- [1641] A. Dhagat and L. Hellerstein. PAC learning with irrelevant attributes. In *Proc. of the 35rd Annual Symposium on Foundations of Computer Science*, pages 64–74. IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [1642] I. S. Dhillon, S. Mallela, and R. Kumar. A divisive information-theoretic feature clustering algorithm for text classification. *Journal of Machine Learning Research*, 3:1265–1287, 2003.
- [1643] I. S. Dhillon and D. S. Modha. Concept decompositions for large sparse text data using clustering. *Machine Learning*, 42(1/2):143–175, 2001.
- [1644] I. Diakonikolas, D. M. Kane, and A. Stewart. Optimal learning via the Fourier transform for sums of independent integer random variables. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 831–849, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1645] I. Diakonikolas, D. M. Kane, and A. Stewart. Properly learning Poisson binomial distributions in almost polynomial time. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 850–878, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1646] I. Diakonikolas, D. M. Kane, and A. Stewart. Learning multivariate log-concave distributions. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 711–727, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [1647] T. Dietterich, M. Kearns, and Y. Mansour. Applying the weak learning framework to understand and improve C4.5. In *Proc. 13th International Conference on Machine Learning*, pages 96–104. Morgan Kaufmann, 1996.
- [1648] T. G. Dietterich. Learning about systems that contain state variables. In *Proc. of the National Conference on Artificial Intelligence*, pages 96–100, Aug. 1984.
- [1649] T. G. Dietterich. Learning at the knowledge level. *Machine Learning*, 1(3):287–315, 1986.
- [1650] T. G. Dietterich. Exploratory research in machine learning. *Machine Learning*, 5:5–9, 1990.
- [1651] T. G. Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7):1895–1924, 1998.
- [1652] T. G. Dietterich. The MAXQ method for hierarchical reinforcement learning. In *Proc. 15th International Conf. on Machine Learning*, pages 118–126. Morgan Kaufmann, San Francisco, CA, 1998.
- [1653] T. G. Dietterich. Machine learning. In R. Wilson and F. Keil, editors, *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press, 1999.
- [1654] T. G. Dietterich. The divide-and-conquer manifesto. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 13–26. Springer, 2000.
- [1655] T. G. Dietterich. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, Boosting, and randomization. *Machine Learning*, 40(2):139–157, 2000.
- [1656] T. G. Dietterich. Hierarchical reinforcement learning with the MAXQ value function decomposition. *Journal of Artificial Intelligence Research*, 13:227–303, 2000.
- [1657] T. G. Dietterich. Machine learning in ecosystem informatics. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 10–11, Berlin, Oct. 2007. Springer.
- [1658] T. G. Dietterich and G. Bakiri. Error-correcting output codes: A general method for improving multiclass inductive learning programs. In *Proc. Ninth National Conference on Artificial Intelligence*, pages 572–577. AAAI Press, 1991.
- [1659] T. G. Dietterich and G. Bakiri. Solving multiclass learning problems via error-correcting output codes. *Journal of Artificial Intelligence Research*, 2:263–286, 1995.
- [1660] T. G. Dietterich and N. S. Flann. Explanation-based learning and reinforcement learning: A unified view. *Machine Learning*, 28(2-3):169–210, 1997.
- [1661] T. G. Dietterich, H. Hild, and G. Bakiri. A comparison of ID3 and backpropagation for English text-to-speech mapping. *Machine Learning*, 18(1):51–80, 1995.
- [1662] T. G. Dietterich, R. L. London, K. Clarkson, and G. Dromey. Learning and inductive inference. In *The Handbook of Artificial Intelligence*, volume 3, pages 323–512. William Kaufman, Inc., 1982.
- [1663] T. G. Dietterich and R. S. Michalski. Learning and generalization of characteristic descriptions: evaluation criteria and comparative review of selected methods. In *Proc. 6th International Joint Conference on Artificial Intelligence*, pages 223–231. IJCAI, 1979.

- [1664] T. G. Dietterich and R. S. Michalski. Discovering patterns in sequences of events. *Artificial Intelligence*, 25:187–232, 1985.
- [1665] T. G. Dietterich and X. Wang. Support vectors for reinforcement learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, page 600. Springer, 2001.
- [1666] S. Dietzen and F. Pfenning. Higher-order and modal logic as a framework for explanation-based generalization. *Machine Learning*, 9:23–55, 1992.
- [1667] C. Dimitrakakis, B. Nelson, A. Mitrokotsa, and B. I. P. Rubinstein. Robust and private Bayesian inference. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 291–305, Berlin, Heidelberg, New York, 2014. Springer.
- [1668] N. Ding, Y. Ren, and D. Gu. PAC learning depth-3  $AC^0$  circuits of bounded top fanin. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 667–680, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [1669] D. I. Diochnos. On the evolution of monotone conjunctions: Drilling for best approximations. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 98–112. Springer, 2016.
- [1670] D. I. Diochnos and G. Turán. On evolvability: The swapping algorithm, product distributions, and covariance. In O. Watanabe and T. Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo, Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, Berlin/Heidelberg, 2009. Springer.
- [1671] A. Doan, P. Domingos, and A. Halevy. Learning to match the schemas of data sources: A multistrategy approach. *Machine Learning*, 50(3):279–301, 2003.
- [1672] D. P. Dobkin and D. Gunopulos. Concept learning with geometric hypotheses. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 329–336, New York, NY, 1995. ACM Press.
- [1673] D. P. Dobkin and D. Gunopulos. Computing the maximum bichromatic discrepancy, with applications to computer graphics and machine learning. *J. of Comput. Syst. Sci.*, 52(3):453–470, 1996.
- [1674] A. Dobra and J. Gehrke. Bias correction in classification tree construction. In *Proc. 18th International Conf. on Machine Learning*, pages 90–97. Morgan Kaufmann, San Francisco, CA, 2001.
- [1675] C. P. Dolan. *Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning, and Planning*. PhD thesis, UCLA Computer Science Department, June 1989.
- [1676] T. Doliwa, M. Kallweit, and H.-U. Simon. Dimension and margin bounds for reflection-invariant kernels. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9–12, 2008*, pages 157–168. Omnipress, 2008.

- [1677] T. Doliwa, H. U. Simon, and S. Zilles. Recursive teaching dimension, learning complexity, and maximum classes. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 209–223, Berlin, Heidelberg, New York, 2010. Springer.
- [1678] C. Domeniconi and D. Gunopulos. An efficient approach for approximating multi-dimensional range queries and nearest neighbor classification in large datasets. In *Proc. 18th International Conf. on Machine Learning*, pages 98–105. Morgan Kaufmann, San Francisco, CA, 2001.
- [1679] C. Domingo. Faster near-optimal reinforcement learning: Adding adaptiveness to the  $E^3$  algorithm. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 241–251. Springer, 1999.
- [1680] C. Domingo, R. Gavaldà, and O. Watanabe. Practical algorithms for on-line sampling. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 150–161. Springer, 1998.
- [1681] C. Domingo and V. Lavín. The query complexity of learning some subclasses of context-free grammars. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 404–414. Springer, 1995.
- [1682] C. Domingo, N. Mishra, and L. Pitt. Efficient read-restricted monotone CNF/DNF dualization by learning with membership queries. *Machine Learning*, 37(1):89–110, 1999.
- [1683] C. Domingo and J. Shawe-Taylor. The complexity of learning minor closed graph classes. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 249–260. Springer, 1995.
- [1684] C. Domingo, T. Tsukiji, and O. Watanabe. Partial Occam’s razor and its applications. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 85–99. Springer, 1997.
- [1685] C. Domingo and O. Watanabe. MadaBoost: A modification of AdaBoost. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 180–189. Morgan Kaufmann, San Francisco, 2000.
- [1686] P. Domingos. Knowledge acquisition from examples via multiple models. In *Proc. 14th International Conference on Machine Learning*, pages 98–106. Morgan Kaufmann, 1997.
- [1687] P. Domingos. A process-oriented heuristic for model selection. In *Proc. 15th International Conf. on Machine Learning*, pages 127–135. Morgan Kaufmann, San Francisco, CA, 1998.
- [1688] P. Domingos. Bayesian averaging of classifiers and the overfitting problem. In *Proc. 17th International Conf. on Machine Learning*, pages 223–230. Morgan Kaufmann, San Francisco, CA, 2000.
- [1689] P. Domingos. Beyond occam’s razor: Process-oriented evaluation. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 3–3. Springer, 2000.

- [1690] P. Domingos. A unified bias-variance decomposition and its applications. In *Proc. 17th International Conf. on Machine Learning*, pages 231–238. Morgan Kaufmann, San Francisco, CA, 2000.
- [1691] P. Domingos. Learning, logic, and probability: A unified view. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 53–53. Springer, 2004.
- [1692] P. Domingos and G. Hulten. A general method for scaling up machine learning algorithms and its application to clustering. In *Proc. 18th International Conf. on Machine Learning*, pages 106–113. Morgan Kaufmann, San Francisco, CA, 2001.
- [1693] P. Domingos and M. Pazzani. Beyond independence: conditions for the optimality of the simple Bayesian classifier. In *Proc. 13th International Conference on Machine Learning*, pages 105–112. Morgan Kaufmann, 1996.
- [1694] P. Domingos and M. Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29(2-3):103–130, 1997.
- [1695] M. Dominguez and R. A. Jacobs. Visual development and the acquisition of binocular disparity sensitivities. In *Proc. 18th International Conf. on Machine Learning*, pages 114–121. Morgan Kaufmann, San Francisco, CA, 2001.
- [1696] J. Dong, A. Krzyżak, and C. Y. Suen. Local learning framework for recognition of lowercase handwritten characters. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 226–238. Springer, 2001.
- [1697] S. Donoho and L. Rendell. Constructive induction using fragmentary knowledge. In *Proc. 13th International Conference on Machine Learning*, pages 113–121. Morgan Kaufmann, 1996.
- [1698] S. K. Donoho and L. A. Rendell. Lessons from theory revision applied to constructive induction. In *Proc. 12th International Conference on Machine Learning*, pages 185–193. Morgan Kaufmann, 1995.
- [1699] D. R. Dooly, S. A. Goldman, and S. S. Kwek. Real-valued multiple-instance learning with queries. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 167–180. Springer, 2001.
- [1700] D. R. Dooly, S. A. Goldman, and S. S. Kwek. Real-valued multiple-instance learning with queries. *J. Comput. Syst. Sci.*, 72(1):1–15, 2006.
- [1701] D. R. Dooly, Q. Zhang, S. A. Goldman, and R. A. Amar. Multiple-instance learning of real-valued data. *Journal of Machine Learning Research*, 3:651–678, 2002.
- [1702] M. Dorigo. ALECSYS and the AutonoMouse: learning to control a real robot by distributed classifier systems. *Machine Learning*, 19(3):209–240, 1995.
- [1703] S. Doshita, K. Furukawa, K. P. Jantke, and T. Nishida, editors. *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Oct. 1993.
- [1704] V. Dötsch, G. Grieser, and S. Lange. Applying formal concepts to learning systems validation. *J. of Experimental and Theoret. Artif. Intell.*, 12(4):433–445, 2000.

- [1705] P. G. Doucet. The syntactic inference problem for D0L sequences. In *L-systems*, pages 146–161. Springer-Verlag, 1974.
- [1706] J. Dougherty, R. Kohavi, and M. Sahami. Supervised and unsupervised discretization of continuous features. In *Proc. 12th International Conference on Machine Learning*, pages 194–202. Morgan Kaufmann, 1995.
- [1707] D. L. Dowe, J. J. Oliver, and C. S. Wallace. MML estimation of the parameters of the spherical Fisher distribution. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 213–227. Springer, 1996.
- [1708] R. Downey, P. Evans, and M. Fellows. Parameterized learning complexity. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 51–57, New York, NY, 1993. ACM Press.
- [1709] T. Downs, K. E. Gates, and A. Masters;. Exact simplification of support vector solutions. *Journal of Machine Learning Research*, 2:293–297, 2001.
- [1710] J. Doyle. On rationality and learning. Technical report, Carnegie Mellon University, 1988.
- [1711] J. A. Drakopoulos. Bounds on the classification error of the nearest neighbor rule. In *Proc. 12th International Conference on Machine Learning*, pages 203–208. Morgan Kaufmann, 1995.
- [1712] G. Drastal. Learning in abstraction space. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 7, pages 163–211. MIT Press, 1994.
- [1713] G. L. Drescher. Suggestions for genetic A.I. Technical Report 198, MIT Artificial Intelligence Laboratory, Feb. 1980.
- [1714] G. L. Drescher. The schema mechanism: A conception of constructivist intelligence. Master’s thesis, MIT Department of Electrical Engineering and Computer Science, Feb. 1985.
- [1715] G. L. Drescher. Genetic AI—translating Piaget into Lisp. Technical Report 890, MIT Artificial Intelligence Laboratory, Feb. 1986.
- [1716] G. L. Drescher. A mechanism for early Piagetian learning. In *Proc. of AAAI-87: Sixth National Conference on Artificial Intelligence*, pages 290–294, Seattle, Washington, July 1987.
- [1717] G. L. Drescher. *Made-up Minds: A Constructivist Approach to Artificial Intelligence*. PhD thesis, MIT EECS Department, Sept. 1989. (Reprinted by MIT Press.).
- [1718] K. Driessens, J. Ramon, and H. Blockeel. Speeding up relational reinforcement learning through the use of an incremental first order decision tree learner. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 97–108. Springer, 2001.
- [1719] P. Drineas and M. W. Mahoney. Approximating a gram matrix for improved kernel-based learning. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 323–337. Springer, 2005.

- [1720] H. Drucker. Improving regressors using boosting techniques. In *Proc. 14th International Conference on Machine Learning*, pages 107–115. Morgan Kaufmann, 1997.
- [1721] H. Drucker, C. Cortes, L. D. Jackel, Y. LeCun, and V. Vapnik. Boosting and other machine learning algorithms. In *Proc. 11th International Conference on Machine Learning*, pages 53–61. Morgan Kaufmann, 1994.
- [1722] H. Drucker, R. Schapire, and P. Simard. Improving performance in neural networks using a boosting algorithm. In *Advances in Neural Information Processing Systems*. Morgan Kaufmann, 1992.
- [1723] H. Drucker, B. Shahrory, and D. C. Gibbon. Relevance feedback using support vector machines. In *Proc. 18th International Conf. on Machine Learning*, pages 122–129. Morgan Kaufmann, San Francisco, CA, 2001.
- [1724] E. Drukh and Y. Mansour. Concentration bounds for unigrams language model. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 170–185. Springer, 2004.
- [1725] C. Drummond and R. C. Holte. Exploiting the cost of (in)sensitivity of decision tree splitting criteria. In *Proc. 17th International Conf. on Machine Learning*, pages 239–246. Morgan Kaufmann, San Francisco, CA, 2000.
- [1726] C. Drummond and R. C. Holte. Cost curves: An improved method for visualizing classifier performance. *Machine Learning*, 65(1):95–130, 2006.
- [1727] M. J. Druzdzel and F. J. Díez. Combining knowledge from different sources in causal probabilistic models. *Journal of Machine Learning Research*, 4:295–316, 2003.
- [1728] A. Dubey, P. Jalote, and S. K. Aggarwal. Inferring grammar rules of programming language dialects. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 201–213, Berlin, 2006. Springer.
- [1729] S. Dubnov, R. El-Yaniv, Y. Gdalyahu, E. Schneidman, N. Tishby, and G. Yona. A new non-parametric pairwise clustering algorithm based on iterative estimation of distance profiles. *Machine Learning*, 47(1):35–61, 2002.
- [1730] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. Wiley, 1973.
- [1731] M. Dudík, S. J. Phillips, and R. E. Schapire. Performance guarantees for regularized maximum entropy density estimation. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 472–486. Springer, 2004.
- [1732] M. Dudík and R. E. Schapire. Maximum entropy distribution estimation with generalized regularization. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 123–138, Berlin, 2006. Springer.
- [1733] R. Dudley, S. Kulkarni, T. Richardson, and O. Zeituni. A metric entropy bound is not sufficient for learnability. Technical Report LIDS-P-2144, MIT LIDS, Oct. 1992.
- [1734] R. M. Dudley. Central limit theorems for empirical measures. *Annals of Probability*, 6(6):899–929, 1978.



- [1735] M. Dudk, K. Hofmann, R. E. Schapire, A. Slivkins, and M. Zoghi. Contextual dueling bandits. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 563–587, Paris, France, 03–06 July 2015. PMLR.
- [1736] M. O. Duff. Q-learning for bandit problems. In *Proc. 12th International Conference on Machine Learning*, pages 209–217. Morgan Kaufmann, 1995.
- [1737] N. Duffy and D. Helmbold. Leveraging for regression. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 208–219. Morgan Kaufmann, San Francisco, 2000.
- [1738] N. Duffy and D. Helmbold. Boosting methods for regression. *Machine Learning*, 47(2/3):153–200, 2002.
- [1739] N. Duffy and D. Helmbold. A geometric approach to leveraging weak learners. *Theoret. Comput. Sci.*, 284(1):67–108, 2002.
- [1740] N. Duffy and D. P. Helmbold. A geometric approach to leveraging weak learners. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 18–33. Springer, 1999.
- [1741] P. Dupon. Regular grammatical inference from positive and negative samples by genetic search: The GIG method. In *Proc. 2nd International Colloquium on Grammatical Inference - ICGI '94*, volume 862 of *Lecture Notes in Artificial Intelligence*, pages 236–245. Springer-Verlag, 1994.
- [1742] P. Dupont. Incremental regular inference. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 222–237. Springer, 1996.
- [1743] P. Dupont and J.-C. Amengual. Smoothing probabilistic automata: An error-correcting approach. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 51–64. Springer, 2000.
- [1744] P. Dupont and L. Chase. Using symbol clustering to improve probabilistic automaton inference. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 232–243. Springer, 1998.
- [1745] S. Džeroski, S. Muggleton, and S. Russell. PAC-learnability of determinate logic programs. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 128–135, New York, NY, 1992. ACM Press.
- [1746] S. Džeroski, S. Muggleton, and S. Russell. Learnability of constrained logic programs. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 342–347. Springer-Verlag, 1993.
- [1747] J. G. Dy and C. E. Brodley. Feature subset selection and order identification for unsupervised learning. In *Proc. 17th International Conf. on Machine Learning*, pages 247–254. Morgan Kaufmann, San Francisco, CA, 2000.

- [1748] R. Dybowski, K. B. Laskey, J. W. Myers, and S. Parsons. Introduction to the special issue on the fusion of domain knowledge with data for decision support. *Journal of Machine Learning Research*, 4:293–294, 2003.
- [1749] S. Džeroski, L. De Raedt, and H. Blockeel. Relational reinforcement learning. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 136–143. Morgan Kaufmann, San Francisco, CA, 1998.
- [1750] S. Džeroski, L. De Raedt, and K. Driessens. Relational reinforcement learning. *Machine Learning*, 43(1/2):7–52, 2001.
- [1751] S. Džeroski, S. Muggleton, and S. Russell. PAC-learnability of constrained nonrecursive logic programs. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 13, pages 243–255. MIT Press, 1995.
- [1752] G. M. Edelman. *Neural Darwinism – The Theory of Neuronal Group Selection*. Basic Books, 1987.
- [1753] H. Edelsbrunner. *Algorithms in Combinatorial Geometry*. Springer-Verlag, 1987.
- [1754] A. Ehrenfeucht and D. Haussler. Learning decision trees from random examples. *Inform. Comput.*, 82(3):231–246, 1989. First appeared in COLT 88.
- [1755] A. Ehrenfeucht, D. Haussler, M. Kearns, and L. Valiant. A general lower bound on the number of examples needed for learning. *Inform. Comput.*, 82(3):247–261, 1989. First appeared in Proc. 1st Annu. Workshop on Comput. Learning Theory, 1988.
- [1756] G. Eibl and K. P. Pfeiffer. Analysis of the performance of AdaBoost.M2 for the simulated digit-recognition-example. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 109–120. Springer, 2001.
- [1757] G. Eibl and K.-P. Pfeiffer. Multiclass boosting for weak classifiers. *Journal of Machine Learning Research*, 6:189–210, 2005.
- [1758] B. Eisenberg and R. L. Rivest. On the sample complexity of PAC-learning using random and chosen examples. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 154–162, San Mateo, CA, 1990. Morgan Kaufmann.
- [1759] B. B. Eisenberg. On the sample complexity of PAC-learning using random and chosen examples. Master’s thesis, MIT EECS Department, 1992. (LCS Technical report MIT/LCS/TR-535, March 1992.).
- [1760] J. Eisner. Grammar induction: Beyond local search. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 112–113, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1761] T. Eiter. Abduction and the dualization problem. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 1–2. Springer, 2003.
- [1762] P. Ejdys and G. Gára. The more we learn the less we know? on inductive learning from examples. *Fundamenta Informaticae*, 39(4):359–374, 1999.

- [1763] R. El-Yaniv and D. Pechyony. Stable transductive learning. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 35–49, Berlin, 2006. Springer.
- [1764] R. El-Yaniv and D. Pechyony. Transductive rademacher complexity and its applications. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 157–171, Berlin, 2007. Springer.
- [1765] R. El-Yaniv and O. Souroujon. Iterative double clustering for unsupervised and semi-supervised learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 121–132. Springer, 2001.
- [1766] A. Elbaz, H. K. Lee, R. A. Servedio, and A. Wan. Separating models of learning from correlated and uncorrelated data. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 637–651. Springer, 2005.
- [1767] A. Elbaz, H. K. Lee, R. A. Servedio, and A. Wan. Separating models of learning from correlated and uncorrelated data. *Journal of Machine Learning Research*, 8:277–290, 2007.
- [1768] R. Eldan and O. Shamir. The power of depth for feedforward neural networks. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 907–940, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1769] J. Eldridge, M. Belkin, and Y. Wang. Beyond hartigan consistency: Merge distortion metric for hierarchical clustering. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 588–606, Paris, France, 03–06 July 2015. PMLR.
- [1770] J. Eldridge, M. Belkin, and Y. Wang. Unperturbed: spectral analysis beyond Davis-Kahan. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 321–358. PMLR, 07–09 Apr. 2018.
- [1771] T. Eliassi-Rad and J. Shavlik. A theory-refinement approach to information extraction. In *Proc. 18th International Conf. on Machine Learning*, pages 130–137. Morgan Kaufmann, San Francisco, CA, 2001.
- [1772] G. Elidan and N. Friedman. Learning hidden variable networks: The information bottleneck approach. *Journal of Machine Learning Research*, 6:81–127, 2005.
- [1773] R. Elio and L. Watanabe. An incremental deductive strategy for controlling constructive induction in learning from examples. *Machine Learning*, 7:7–44, 1991.
- [1774] A. Elisseeff, T. Evgeniou, and M. Pontil. Stability of randomized learning algorithms. *Journal of Machine Learning Research*, 6:55–79, 2005.
- [1775] D. P. W. Ellis and G. E. Poliner. Classification-based melody transcription. *Machine Learning*, 65(2-3):439–456, 2006.
- [1776] J. L. Elman. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7:195–225, 1991.

- [1777] T. Elomaa. In defense of C4.5: notes on learning one-level decision trees. In *Proc. 11th International Conference on Machine Learning*, pages 62–69. Morgan Kaufmann, 1994.
- [1778] T. Elomaa and M. Kääriäinen. On the practice of branching program boosting. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 133–144. Springer, 2001.
- [1779] T. Elomaa and J. Rousu. General and efficient multisplitting of numerical attributes. *Machine Learning*, 36(3):201–244, 1999.
- [1780] A.-M. Emde and B. Schinzel. Aggregating inductive expertise on partial recursive functions. *Inform. Comput.*, 96(2):139–167, 1992.
- [1781] W. Emde and D. Wettschereck. Relational instance-based learning. In *Proc. 13th International Conference on Machine Learning*, pages 122–130. Morgan Kaufmann, 1996.
- [1782] J. D. Emerald, K. G. Subramanian, and D. G. Thomas. Learning code regular and code linear languages. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 211–221. Springer, 1996.
- [1783] J. D. Emerald, K. G. Subramanian, and D. G. Thomas. Learning a subclass of context-free languages. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 223–231. Springer, 1998.
- [1784] J. D. Emerald, K. G. Subramanian, and D. G. Thomas. Inferring subclasses of contextual languages. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 65–74. Springer, 2000.
- [1785] A. Engel. Complexity of learning in artificial neural networks. *Theoret. Comput. Sci.*, 265(1-2):285–306, 2001.
- [1786] Y. Engel and S. Mannor. Learning embedded maps of Markov processes. In *Proc. 18th International Conf. on Machine Learning*, pages 138–145. Morgan Kaufmann, San Francisco, CA, 2001.
- [1787] S. P. Engelson and M. Koppel. Distilling reliable information from unreliable theories. In *Proc. 12th International Conference on Machine Learning*, pages 218–225. Morgan Kaufmann, 1995.
- [1788] S. P. Engelson and M. Koppel. Identifying the information contained in a flawed theory. In *Proc. 13th International Conference on Machine Learning*, pages 131–138. Morgan Kaufmann, 1996.
- [1789] D. Ensign, S. Neville, A. Paul, and S. Venkatasubramanian. The complexity of explaining neural networks through (group) invariants. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory, volume 76 of Proceedings of Machine Learning Research*, pages 341–359, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.

- [1790] D. Ensign, F. Sorelle, N. Scott, S. Carlos, and V. Suresh. Decision making with limited feedback. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 359–367. PMLR, 07–09 Apr. 2018.
- [1791] S. L. Epstein. Hard questions about easy tasks: issues from learning to play games. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 18, pages 487–521. MIT Press, 1994.
- [1792] S. L. Epstein. Toward an ideal trainer. *Machine Learning*, 15(3):251–278, 1994.
- [1793] F. Ergün, R. S. Kumar, and R. Rubinfeld. On learning bounded-width branching programs. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 361–368, New York, NY, 1995. ACM Press.
- [1794] F. Ergün, S. R. Kumar, and R. Rubinfeld. Learning distributions from random walks. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 243–249, New York, NY, 1997. ACM Press.
- [1795] T. Erlebach, P. Rossmanith, H. Stadtherr, A. Steger, and T. Zeugmann. Efficient learning of one-variable pattern languages from positive examples. Technical Report DOI-TR-128, Department of Informatics, Kyushu University, 1996.
- [1796] T. Erlebach, P. Rossmanith, H. Stadtherr, A. Steger, and T. Zeugmann. Learning one-variable pattern languages very efficiently on average, in parallel, and by asking queries. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 260–276. Springer, 1997.
- [1797] T. Erlebach, P. Rossmanith, H. Stadtherr, A. Steger, and T. Zeugmann. Learning one-variable pattern languages very efficiently on average, in parallel, and by asking queries. *Theoret. Comput. Sci.*, 261(1):119–156, 2001. Special issue for ALT '97.
- [1798] D. Ernst, P. Geurts, and L. Wehenkel. Tree-based batch mode reinforcement learning. *Journal of Machine Learning Research*, 6:503–556, 2005.
- [1799] M. Ernst, M. Liśkiewicz, and R. Reischuk. Algorithmic learning for steganography: Proper learning of k-term DNF formulas from positive samples. In K. Elbassioni and K. Makino, editors, *Algorithms and Computation, 26th International Symposium, ISAAC 2015, Nagoya, Japan, December 9-11, 2015, Proceedings*, volume 9472 of *Lecture Notes in Computer Science*, pages 151–162, Berlin, Heidelberg, New York, 2015. Springer.
- [1800] T. Erven, M. D. Reid, and R. C. Williamson. Mixability is Bayes risk curvature relative to log loss. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 233–252, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1801] T. V. Erven, W. Kotłowski, and M. K. Warmuth. Follow the leader with dropout perturbations. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 949–974, Barcelona, Spain, 13–15 June 2014. PMLR.

- [1802] G. Escudero, L. Màrquez, and G. Rigau. Boosting applied to word sense disambiguation. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 129–141. Springer, 2000.
- [1803] E. Eskin. Anomaly detection over noisy data using learned probability distributions. In *Proc. 17th International Conf. on Machine Learning*, pages 255–262. Morgan Kaufmann, San Francisco, CA, 2000.
- [1804] F. Esposito, N. Fanizzi, S. Ferilli, and G. Semeraro. Ideal theory refinement under object identity. In *Proc. 17th International Conf. on Machine Learning*, pages 263–270. Morgan Kaufmann, San Francisco, CA, 2000.
- [1805] F. Esposito, D. Malerba, and G. Semeraro. Decision tree pruning as a search in the state space. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 165–184. Springer-Verlag, 1993.
- [1806] F. Esposito, G. Semeraro, N. Fanizzi, and S. Ferilli. Multistrategy theory revision: Induction and abduction in INTHELEX. *Machine Learning*, 38(1/2):133–156, 2000.
- [1807] Y. Esposito, A. Lemay, F. Denis, and P. Dupont. Learning probabilistic residual finite state automata. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 77–91. Springer, 2002.
- [1808] O. Etzioni and R. Etzioni. Statistical methods for analyzing speedup learning experiments. *Machine Learning*, 14:333–347, 1994.
- [1809] O. Etzioni and R. Etzioni. Technical note: statistical methods for analyzing speedup learning experiments. *Machine Learning*, 14(3):333–347, 1994.
- [1810] W. Evans, S. Rajagopalan, and U. Vazirani. Choosing a reliable hypothesis. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 269–276, New York, NY, 1993. ACM Press.
- [1811] E. Even-Dar, M. Kearns, Y. Mansour, and J. Wortman. Regret to the best vs. regret to the average. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 233–247, Berlin, 2007. Springer.
- [1812] E. Even-Dar, M. Kearns, and J. Wortman. Risk-sensitive online learning. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 199–213. Springer, Oct. 2006.
- [1813] E. Even-Dar, S. Mannor, and Y. Mansour. PAC bounds for multi-armed bandit and Markov decision processes. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 255–270. Springer, 2002.
- [1814] E. Even-Dar and Y. Mansour. Learning rates for Q-learning. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 589–604. Springer, 2001.

- [1815] E. Even-Dar and Y. Mansour. Approximate equivalence of Markov decision processes. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 581–594. Springer, 2003.
- [1816] T. Evgeniou, C. A. Micchelli, and M. Pontil. Learning multiple tasks with kernel methods. *Journal of Machine Learning Research*, 6:615–637, 2005.
- [1817] T. Evgeniou, L. Perez-Breva, M. Pontil, and T. Poggio. Bounds on the generalization performance of kernel machine ensembles. In *Proc. 17th International Conf. on Machine Learning*, pages 271–278. Morgan Kaufmann, San Francisco, CA, 2000.
- [1818] T. Evgeniou and M. Pontil. On the  $V_\gamma$  dimension for regression in reproducing kernel hilbert spaces. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 106–117. Springer, 1999.
- [1819] T. Evgeniou and M. Pontil. A note on the generalization performance of kernel classifiers with margin. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 306–315. Springer, 2000.
- [1820] R. Eyraud, C. de la Higuera, and J.-C. Janodet. Representing languages by learnable rewriting systems. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 139–150. Springer, 2004.
- [1821] R. Eyraud, C. de la Higuera, and J.-C. Janodet. LARS: A learning algorithm for rewriting systems. *Machine Learning*, 66(1):7–31, 2007.
- [1822] R. Eyraud, J.-C. Janodet, and T. Oates. Learning substitutable binary plane graph grammars. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 114–128, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1823] K. J. Ezawa, M. Singh, and S. W. Norton. Learning goal oriented Bayesian networks for telecommunications risk management. In *Proc. 13th International Conference on Machine Learning*, pages 139–147. Morgan Kaufmann, 1996.
- [1824] V. Faber and J. Mycielski. Applications of learning theorems. *Fundamenta Informaticae*, 15(2):145–167, 1991.
- [1825] S. E. Fahlman and C. Lebiere. The cascade-correlation learning architecture. In *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann, 1990.
- [1826] A. F. Fahmy and A. W. Biermann. Synthesis of real time acceptors. *J. Symbolic Computation*, 15:807–842, 1993.
- [1827] A. F. Fahmy and R. S. Roos. Efficient learning of real time one-counter automata. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 25–40. Springer, 1995.

- [1828] A. F. Fahmy and R. S. Roos. Efficient learning of real time two-counter automata. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 113–126. Springer, 1996.
- [1829] J. Fakcharoenphol and B. Kijssirikul. Constructing multiclass learners from binary learners: A simple black-box analysis of the generalization errors. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 135–147. Springer, Oct. 2005.
- [1830] J. Fakcharoenphol and B. Kijssirikul. Erratum: Constructing multiclass learners from binary learners: A simple black-box analysis of the generalization errors. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 464–466, Berlin, Oct. 2008. Springer.
- [1831] M. Falahatgar, A. Jafarpour, A. Orlitsky, V. Pichapati, and A. T. Suresh. Faster algorithms for testing under conditional sampling. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 607–636, Paris, France, 03–06 July 2015. PMLR.
- [1832] B. C. Falkenhainer and R. S. Michalski. Integrating quantitative and qualitative discovery: The ABACUS system. *Machine Learning*, 1(4):367–401, 1986.
- [1833] W. Fan, W. Lee, S. J. Stolfo, and M. Miller. A multiple model cost-sensitive approach for intrusion detection. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 142–153. Springer, 2000.
- [1834] W. Fan, S. J. Stolfo, J. Zhang, and P. K. Chan. AdaCost: misclassification cost-sensitive boosting. In *Proc. 16th International Conf. on Machine Learning*, pages 97–105. Morgan Kaufmann, San Francisco, CA, 1999.
- [1835] S. Fang and S. Venkatesh. On the average tractability of binary integer programming and the curious transition to perfect generalization in learning majority functions. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 310–316, New York, NY, 1993. ACM Press.
- [1836] S. C. Fang and S. S. Venkatesh. Learning binary perceptrons perfectly efficiently. *J. of Comput. Syst. Sci.*, 52(2):374–389, 1996.
- [1837] J. D. Farmer and J. J. Sidorowich. Exploiting chaos to predict the future and reduce noise. Technical Report LA-UR-88-901, Los Alamos National Laboratory, Mar. 1988.
- [1838] L. Fass. Inference of skeletal automata. Technical report, Georgetown University Computer Science Dept. TR-2, 1984.
- [1839] Y. E. Fattah and P. O’Rorke. Explanation-based learning for diagnosis. *Machine Learning*, 13:35–70, 1993.
- [1840] U. M. Fayyad and K. B. Irani. On the handling of continuous-valued attributes in decision tree generation. *Machine Learning*, 8:87–102, 1992.
- [1841] M. Feder. Learning to communicate via unknown channel. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, page 79, New York, NY, 1998. ACM Press.



- [1842] M. Feder, N. Merhav, and M. Gutman. Universal prediction of individual sequences. *IEEE Transactions on Information Theory*, 38:1258–1270, 1992.
- [1843] U. Feige, Y. Mansour, and R. Schapire. Learning and inference in the presence of corrupted inputs. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 637–657, Paris, France, 03–06 July 2015. PMLR.
- [1844] U. Feige, Y. Mansour, and R. E. Schapire. Robust inference for multiclass classification. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 368–386. PMLR, 07–09 Apr. 2018.
- [1845] U. Feige and A. Shamir. Learning in permutation groups (extended abstract). Preprint, Applied Mathematics Dept., The Weizmann Institute of Science., 1987.
- [1846] A. Feigelson and L. Hellerstein. Learning conjunctions of two unate DNF formulas: computational and informational results. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 255–265, New York, NY, 1996. ACM Press.
- [1847] J. Feldman. Some decidability results on grammatical inference and complexity. *Inform. Control*, 20(3):244–262, 1972.
- [1848] J. Feldman, R. O’Donnell, and R. A. Servedio. Learning mixtures of product distributions over discrete domains. *SIAM J. Comput.*, 37(5):1536–1564, 2008.
- [1849] J. Feldman, R. A. Servedio, and R. O’Donnell. PAC learning axis-aligned mixtures of Gaussians with no separation assumption. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 20–34, Berlin, 2006. Springer.
- [1850] J. A. Feldman. First thoughts on grammatical inference. Technical report, Stanford University Artificial Intelligence Memo 55, 1967.
- [1851] J. A. Feldman. Real language learning. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 114–125. Springer, 1998.
- [1852] J. A. Feldman and S. Reder. Grammatical complexity and inference. Technical report, Computer Science Dept., Stanford University, 1969.
- [1853] J. A. Feldman and P. Shields. Total complexity and the inference of best programs. *Math. Syst. Theory*, 10:181–191, 1977.
- [1854] M. Feldman, C. Harshaw, and A. Karbasi. Greed is good: Near-optimal submodular maximization via greedy optimization. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 758–784, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1855] V. Feldman. On attribute efficient and non-adaptive learning of parities and DNF expressions. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 576–590. Springer, 2005.

- [1856] V. Feldman. Hardness of approximate two-level logic minimization and PAC learning with membership queries. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 363–372. ACM Press, 2006.
- [1857] V. Feldman. Attribute-efficient and non-adaptive learning of parities and DNF expressions. *Journal of Machine Learning Research*, 8:1431–1460, 2007.
- [1858] V. Feldman. On the power of membership queries in agnostic learning. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 147–156. Omnipress, 2008.
- [1859] V. Feldman. Distribution-independent evolvability of linear threshold functions. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 253–272, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1860] V. Feldman. A complete characterization of statistical query learning with applications to evolvability. *J. Comput. Syst. Sci.*, 78(5):1444–1459, 2012.
- [1861] V. Feldman. Learning DNF expressions from fourier spectrum. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 17.1–17.19, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1862] V. Feldman. Open problem: The statistical query complexity of learning sparse halfspaces. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1283–1289, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1863] V. Feldman. Dealing with range anxiety in mean estimation via statistical queries. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 629–640, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [1864] V. Feldman. A general characterization of the statistical query complexity. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 785–830, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1865] V. Feldman, P. Gopalan, S. Khot, and A. K. Ponnuswami. On agnostic learning of parities, monomials, and halfspaces. *SIAM J. Comput.*, 39(2):606–645, 2009.
- [1866] V. Feldman, V. Guruswami, P. Raghavendra, and Y. Wu. Agnostic learning of monomials by halfspaces is hard. *SIAM J. Comput.*, 41(6):1558–1590, 2011.
- [1867] V. Feldman and V. Kanade. Computational bounds on statistical query learning. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 16.1–16.22, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [1868] V. Feldman and P. Kothari. Learning coverage functions and private release of marginals. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 679–702, Barcelona, Spain, 13–15 June 2014. PMLR.

- [1869] V. Feldman, P. Kothari, and J. Vondrák. Representation, approximation and learning of submodular functions using low-rank decision trees. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 711–740. JMLR.org, 2013.
- [1870] V. Feldman, P. Kothari, and J. Vondrák. Tight bounds on  $\ell_1$  approximation and learning of self-bounding functions. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 540–559, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [1871] V. Feldman, H. K. Lee, and R. A. Servedio. Lower bounds and hardness amplification for learning shallow monotone formulas. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 273–292, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1872] V. Feldman and A. Rakhlin. Conference on learning theory 2016: Preface. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1–3, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1873] V. Feldman, A. Rakhlin, and O. Shamir, editors. *Conference on Learning Theory, 23-26 June 2016, Columbia University, New York, New York, USA*, volume 49 of *Proceedings of Machine Learning Research*. PMLR, 2016.
- [1874] V. Feldman and S. Shah. Separating models of learning with faulty teachers. *Theoret. Comput. Sci.*, 410(9):1903–1912, 2009. Special issue for ALT 2007.
- [1875] V. Feldman, S. Shah, and N. Wadhwa. Separating models of learning with faulty teachers. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 94–106, Berlin, Oct. 2007. Springer.
- [1876] V. Feldman and T. Steinke. Generalization for adaptively-chosen estimators via stable median. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 728–757, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1877] V. Feldman and L. G. Valiant. The learning power of evolution. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 513–514. Omnipress, 2008.
- [1878] V. Feldman and D. Xiao. Sample complexity bounds on differentially private learning via communication complexity. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1000–1019, Barcelona, Spain, 13–15 June 2014. PMLR.
- [1879] H. Feliciangeli and G. Herman. Algorithms for producing grammars from sample derivations: a common problem of formal language theory and developmental biology. *J. of Comput. Syst. Sci.*, 7:97–118, 1977.
- [1880] J. S. Felix A. Gers, Nicol N. Schraudolph. Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*, 3:115–143, 2002.

- [1881] W. Feller. *An Introduction to Probability and its Applications*, volume 2. John Wiley and Sons, second edition, 1971.
- [1882] S. Fenner and W. Gasarch. The complexity of learning  $\text{SUBSEQ}(A)$ . In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 109–123. Springer, Oct. 2006.
- [1883] D. Fensel and M. Wiese. Refinement of rule sets with JoJo. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 378–383. Springer-Verlag, 1993.
- [1884] A. Feragen and S. Hauberg. Open problem: Kernel methods on manifolds and metric spaces. what is the probability of a positive definite geodesic exponential kernel? In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1647–1650, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1885] A. Fern and R. Givan. Online ensemble learning: An empirical study. In *Proc. 17th International Conf. on Machine Learning*, pages 279–286. Morgan Kaufmann, San Francisco, CA, 2000.
- [1886] A. Fern and R. Givan. Online ensemble learning: An empirical study. *Machine Learning*, 53(1-2):71–109, 2003.
- [1887] A. Fern, R. Goetschalckx, M. Hamidi-Haines, and P. Tadepalli. Adaptive submodularity with varying query sets: An application to active multi-label learning. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 577–592, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [1888] H. Fernau. Identification of function distinguishable languages. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 116–130. Springer, 2000.
- [1889] H. Fernau. Learning XML grammars. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 73–87. Springer, 2001.
- [1890] H. Fernau. Even linear simple matrix languages: formal language properties and grammatical inference. *Theoret. Comput. Sci.*, 289(1):425–456, 2002.
- [1891] H. Fernau. Fragmentation: Enhancing identifiability. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 92–105. Springer, 2002.
- [1892] H. Fernau. Learning tree languages from text. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 153–168. Springer, 2002.
- [1893] H. Fernau. Identification of function distinguishable languages. *Theoret. Comput. Sci.*, 290(3):1679–1711, 2003.

- [1894] H. Fernau. Extracting minimum length document type definitions is NP-hard: (Abstract). In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 277–278. Springer, 2004.
- [1895] H. Fernau. Identifying terminal distinguishable languages. *Annals of Mathematics and Artificial Intelligence*, 40(3-4):263–281, 2004.
- [1896] H. Fernau. Algorithms for learning regular expressions. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 297–311. Springer, Oct. 2005.
- [1897] H. Fernau. Algorithms for learning regular expressions from positive data. *Inform. Comput.*, 207(4):521–541, 2009.
- [1898] H. Fernau, F. Manea, R. Mercas, and M. L. Schmid. Revisiting shinohara’s algorithm for computing descriptive patterns. *Theoret. Comput. Sci.*, 733:44–54, 2018. Special Issue on Learning Theory and Complexity.
- [1899] H. Fernau and J. M. Sempere. Permutations and control sets for learning non-regular language families. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 75–88. Springer, 2000.
- [1900] J. Ferreira, J. Correia, T. Jamet, and E. Costa. An application of machine learning in the domain of loan analysis. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 414–419. Springer-Verlag, 1993.
- [1901] C. Ferretti and G. Mauri. Identifying regular languages over partially-commutative monoids. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 282–289. Springer-Verlag, 1994.
- [1902] A. Fiat, S. Moses, A. Shamir, I. Shimshoni, and G. Tardos. Planning and learning in permutation groups. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 274–279. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [1903] A. Fiat and D. Pechyony. Decision trees: More theoretical justification for practical algorithms. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 156–170. Springer, 2004.
- [1904] C. N. Fiechter. Efficient reinforcement learning. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 88–97, New York, NY, 1994. ACM Press.
- [1905] C.-N. Fiechter. PAC adaptive control of linear systems. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 72–80, 1997.
- [1906] C.-N. Fiechter and S. Rogers. Learning subjective functions with large margins. In *Proc. 17th International Conf. on Machine Learning*, pages 287–294. Morgan Kaufmann, San Francisco, CA, 2000.

- [1907] S. Fine, R. Gilad-Bachrach, and E. Shamir. Query by committee, linear separation and random walks. *Theoret. Comput. Sci.*, 284(1):25–51, 2002.
- [1908] S. Fine and Y. Mansour. Active sampling for multiple output identification. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 620–634, Berlin, 2006. Springer.
- [1909] S. Fine and Y. Mansour. Active sampling for multiple output identification. *Machine Learning*, 69(2-3):213–228, 2007.
- [1910] S. Fine and K. Scheinberg. Efficient SVM training using low-rank kernel representations. *Journal of Machine Learning Research*, 2:243–264, 2001.
- [1911] P. Finn, S. Muggleton, D. Page, and A. Srinivasan. Pharmacophore discovery using the inductive logic programming system PROGOL. *Machine Learning*, 30:241–270, 1998.
- [1912] W. Finnoff and H. G. Zimmermann. Detecting structure in small datasets by network fitting under complexity constraints. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 8, pages 113–131. MIT Press, 1994.
- [1913] D. J. Finton and Y. H. Hu. Importance-based feature extraction for reinforcement learning. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 5, pages 77–94. MIT Press, 1995.
- [1914] S. Fiori. Quasi-geodesic neural learning algorithms over the orthogonal group: A tutorial. *Journal of Machine Learning Research*, 6:743–781, 2005.
- [1915] L. Firoiu and P. R. Cohen. Abstracting from robot sensor data using hidden Markov models. In *Proc. 16th International Conf. on Machine Learning*, pages 106–114. Morgan Kaufmann, San Francisco, CA, 1999.
- [1916] L. Firoiu, T. Oates, and P. R. Cohen. Learning deterministic finite automaton with a recurrent neural network. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 90–101. Springer, 1998.
- [1917] P. Fischer. Learning unions of convex polygons. In *Computational Learning Theory: Euro-Colt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 61–67, Oxford, 1994. Oxford University Press.
- [1918] P. Fischer. More or less efficient agnostic learning of convex polygons. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 337–344, New York, NY, 1995. ACM Press.
- [1919] P. Fischer, K.-U. Höffgen, and H. Lefmann. PAC learning from general examples. *Theoret. Comput. Sci.*, 172(1-2):43–65, 1997.
- [1920] P. Fischer, S. Pölt, and H. U. Simon. Probably almost Bayes decisions. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 88–94, San Mateo, CA, 1991. Morgan Kaufman.
- [1921] P. Fischer and H.-U. Simon. On learning ring-sum expansions. *SIAM J. Comput.*, 21(1):181–192, 1992.

- [1922] P. Fischer and H.-U. Simon, editors. *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, Berlin, Mar. 1999. Springer.
- [1923] S. Fischer and H. Bunke. Automatic identification of diatoms using decision forests. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 173–183. Springer, 2001.
- [1924] M. Fischlin. Cryptographic limitations on parallelizing membership and equivalence queries with applications to random self-reductions. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 72–86. Springer, 1998.
- [1925] M. Fischlin. Cryptographic limitations on parallelizing membership and equivalence queries with applications to random-self-reductions. *Theoret. Comput. Sci.*, 268(2):199–219, 2001. Special issue for ALT '98.
- [1926] B. Fish and L. Reyzin. Open problem: Meeting times for learning random automata. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 8–11, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1927] D. Fisher. Editorial. *Machine Learning*, 47(1):5–6, 2002.
- [1928] D. H. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2(2):139–172, 1987.
- [1929] L. J. Fitzgibbon, L. Allison, and D. L. Dowe. Minimum message length grouping of ordered data. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 56–70. Springer, 2000.
- [1930] J. M. Fitzpatrick and J. J. Grefenstette. Genetic algorithms in noisy environments. *Machine Learning*, 3:101–120, 1988.
- [1931] P. Flach and S. Dzeroski. Editorial: Inductive logic programming is coming of age. *Machine Learning*, 44(3):207–209, 2001.
- [1932] P. A. Flach. Second-order inductive learning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 202–216. Springer-Verlag, 1989.
- [1933] P. A. Flach. An analysis of various forms of ‘jumping to conclusions’. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 170–186. Springer-Verlag, 1992.
- [1934] P. A. Flach. Predicate invention in inductive data engineering. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 83–94. Springer-Verlag, 1993.
- [1935] P. A. Flach and N. Lachiche. Confirmation-guided discovery of first-order rules with tertius. *Machine Learning*, 42(1/2):61–95, 2001.

- [1936] G. W. Flake and S. Lawrence. Efficient SVM regression training with SMO. *Machine Learning*, 46(1-3):271–290, 2002.
- [1937] N. Flammarion and F. Bach. From averaging to acceleration, there is only a step-size. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 658–695, Paris, France, 03–06 July 2015. PMLR.
- [1938] N. Flammarion and F. Bach. Stochastic composite least-squares regression with convergence rate  $O(1/n)$ . In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 831–875, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1939] M. Flammini, A. Marchetti-Spaccamela, and L. Kučera. Learning DNF formulae under classes of probability distributions. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 85–92, New York, NY, 1992. ACM Press.
- [1940] P. A. Flanagan. A discrete model of semantic learning. Technical report, Brown University Computer Science Dept., CS-79, 1981.
- [1941] N. S. Flann and T. G. Dietterich. A study of explanation-based methods for inductive learning. *Machine Learning*, 4(2):187–226, 1989.
- [1942] M. J. Flanagan, L. S. Fried, and K. J. Holyoak. Distributional expectations and the induction of category structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(2):241–256, 1986.
- [1943] R. Fleischer. Decision trees: Old and new results. *Inform. Comput.*, 152(1):44–61, 1999.
- [1944] J. Flesch, R. Laraki, and V. Perchet. Online learning and blackwell approachability in quitting games. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 941–942, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1945] C. C. Florêncio. Learning node label controlled graph grammars (extended abstract). In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 286–288. Springer, 2008.
- [1946] C. C. Florêncio. Learning tree adjoining grammars from structures and strings. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 129–132, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [1947] C. C. Florêncio and H. Fernau. Hölder norms and a hierarchy theorem for parameterized classes of CCG. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 280–283. Springer, 2010.
- [1948] C. C. Florêncio and H. Fernau. On families of categorial grammars of bounded value, their learnability and related complexity questions. *Theoret. Comput. Sci.*, 452:21–38, 2012.



- [1949] C. C. Florêncio and S. Verwer. Regular inference as vertex coloring. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 81–95, Berlin, Heidelberg, New York, 2012. Springer.
- [1950] C. C. Florêncio and S. Verwer. Regular inference as vertex coloring. *Theoret. Comput. Sci.*, 558:18–34, 2014. Special Issue: Algorithmic Learning Theory.
- [1951] L. Florescu and W. Perkins. Spectral thresholds in the bipartite stochastic block model. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 943–959, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1952] C. C. Florêncio. On the complexity of consistent identification of some classes of structure languages. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 89–102. Springer, 2000.
- [1953] C. C. Florêncio. Consistent identification in the limit of rigid grammars from strings is NP-hard. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 49–62. Springer, 2002.
- [1954] S. Floyd. Space-bounded learning and the Vapnik-Chervonenkis dimension. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 349–364, San Mateo, CA, 1989. Morgan Kaufmann.
- [1955] S. Floyd. *Space-bounded learning and the Vapnik-Chervonenkis Dimension (Ph.D)*. PhD thesis, U.C. Berkeley, Dec. 1989. ICSI Tech Report TR-89-061.
- [1956] S. Floyd and M. Warmuth. Sample compression, learnability, and the Vapnik-Chervonenkis dimension. *Machine Learning*, 21(3):269–304, 1995.
- [1957] T. C. Fogarty. First nearest neighbor classification on frey and slate’s letter recognition problem. *Machine Learning*, 9:387–388, 1992.
- [1958] E. Fokoué and D. M. Titterington. Mixtures of factor analysers. Bayesian estimation and inference by stochastic simulation. *Machine Learning*, 50(1-2):73–94, 2003.
- [1959] P. W. L. Fong. A quantitative study of hypothesis selection. In *Proc. 12th International Conference on Machine Learning*, pages 226–234. Morgan Kaufmann, 1995.
- [1960] A. Foret and Y. L. Nir. On limit points for some variants of rigid lambek grammars. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 106–119. Springer, 2002.
- [1961] G. Forman. An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 3:1289–1305, 2003.
- [1962] S. Forrest and M. Mitchell. What makes a problem hard for a genetic algorithm? some anomalous results and their explanation. *Machine Learning*, 13:285–319, 1993.

- [1963] J. Forster. Learning multiplicity automata from smallest counterexamples. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 79–90. Springer, 1999.
- [1964] J. Forster, N. Schmitt, and H. U. Simon. Estimating the optimal margins of embeddings in euclidean half spaces. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 402–415. Springer, 2001.
- [1965] J. Forster, N. Schmitt, H. U. Simon, and T. Suttorp. Estimating the optimal margins of embeddings in euclidean half spaces. *Machine Learning*, 51(3):263–281, 2003.
- [1966] J. Forster and H. U. Simon. On the smallest possible dimension and the largest possible margin of linear arrangements representing given concept classes uniform distribution. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 128–138. Springer, 2002.
- [1967] J. Forster and H. U. Simon. On the smallest possible dimension and the largest possible margin of linear arrangements representing given concept classes. *Theoret. Comput. Sci.*, 350(1):40–48, 2006. Special issue for ALT 2002.
- [1968] J. Forster and M. Warmuth. Relative expected instantaneous loss bounds. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 90–99. Morgan Kaufmann, San Francisco, 2000.
- [1969] J. Forster and M. Warmuth. Relative loss bounds for temporal-difference learning. In *Proc. 17th International Conf. on Machine Learning*, pages 295–302. Morgan Kaufmann, San Francisco, CA, 2000.
- [1970] J. Forster and M. K. Warmuth. Relative expected instantaneous loss bounds. *J. Comput. Syst. Sci.*, 64(1):76–102, 2002. Special issue for COLT 2000.
- [1971] J. Forster and M. K. Warmuth. Relative loss bounds for temporal-difference learning. *Machine Learning*, 51(1):23–50, 2003.
- [1972] L. Fortnow, R. Freivalds, W. I. Gasarch, M. Kummer, S. A. Kurtz, C. H. Smith, and F. Stephan. Relative sizes of learnable sets. *Theoret. Comput. Sci.*, 197(1/2):139–156, 1998.
- [1973] L. Fortnow, W. Gasarch, S. Jain, E. Kinber, M. Kummer, S. Kurtz, M. Pleszkoch, T. Slaman, R. Solovay, and F. Stephan. Extremes in the degrees of inferability. *Annals of Pure and Applied Logic*, 66(3):231–276, 1994.
- [1974] L. Fortnow and A. R. Klivans. Efficient learning algorithms yield circuit lower bounds. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 350–363, Berlin, 2006. Springer.
- [1975] L. Fortnow and J. H. Lutz. Prediction and dimension. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 380–395. Springer, 2002.
- [1976] L. Fortnow and J. H. Lutz. Prediction and dimension. *J. Comput. Syst. Sci.*, 70(4):570–589, 2005. Special issue on COLT 2002.

- [1977] D. Foster and P. Dayan. Structure in the space of value functions. *Machine Learning*, 49(2-3):325–346, 2002.
- [1978] D. Foster, S. Kale, and H. Karloff. Online sparse linear regression. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 960–970, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [1979] D. Foster, H. Karloff, and J. Thaler. Variable selection is hard. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 696–709, Paris, France, 03–06 July 2015. PMLR.
- [1980] D. J. Foster, A. Rakhlin, and K. Sridharan. Zigzag: A new approach to adaptive online learning. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 876–924, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [1981] D. P. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. Complexity-based approach to calibration with checking rules. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 293–314, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1982] R. Foygel and N. Srebro. Concentration-based guarantees for low-rank matrix reconstruction. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 315–340, Budapest, Hungary, 09–11 June 2011. PMLR.
- [1983] E. Frank and M. Hall. A simple approach to ordinal classification. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 145–156. Springer, 2001.
- [1984] E. Frank, L. Trigg, G. Holmes, and I. H. Witten. Technical note: Naive Bayes for regression. *Machine Learning*, 41(1):5–25, 2000.
- [1985] E. Frank and I. H. Witten. Generating accurate rule sets without global optimization. In *Proc. 15th International Conf. on Machine Learning*, pages 144–151. Morgan Kaufmann, San Francisco, CA, 1998.
- [1986] E. Frank and I. H. Witten. Using a permutation test for attribute selection in decision trees. In *Proc. 15th International Conf. on Machine Learning*, pages 152–160. Morgan Kaufmann, San Francisco, CA, 1998.
- [1987] E. Frank and I. H. Witten. Making better use of global discretization. In *Proc. 16th International Conf. on Machine Learning*, pages 115–123. Morgan Kaufmann, San Francisco, CA, 1999.
- [1988] J. Franklin, T. Mitchell, and S. Thrun, editors. *Recent Advances in Robot Learning*. Number 368 in Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, 1996.
- [1989] J. A. Franklin, T. M. Mitchell, and S. Thrun. Introduction. *Machine Learning*, 23:117–119, 1996.

- [1990] P. Frasconi, M. Gori, M. Maggini, and G. Soda. Representation of finite state automata in recurrent radial basis function networks. *Machine Learning*, 23(1):5–32, 1996.
- [1991] M. Frazier. *Matters Horn and Other Features in the Computational Learning Theory Landscape: The Notion of Membership*. PhD thesis, Dept. of Computer Science, University of Illinois at Urbana-Champaign, 1994. Technical report UIUCDCS-R-94-1858.
- [1992] M. Frazier, S. Goldman, N. Mishra, and L. Pitt. Learning from a consistently ignorant teacher. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 328–339, New York, NY, 1994. ACM Press.
- [1993] M. Frazier, S. Goldman, N. Mishra, and L. Pitt. Learning from a consistently ignorant teacher. *J. of Comput. Syst. Sci.*, 52(3):471–492, 1996.
- [1994] M. Frazier and C. D. Page. Learnability of recursive, non-determinate theories: Some basic results and techniques. In *Third International Workshop on Inductive Logic Programming*, 1993.
- [1995] M. Frazier and L. Pitt. Learning from entailment: An application to propositional Horn sentences. In *Proc. Tenth International Conf. Machine Learning*, pages 120–127. Morgan Kaufmann, June 1993.
- [1996] M. Frazier and L. Pitt. Some new directions in computational learning theory. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 19–32, Oxford, 1994. Oxford University Press.
- [1997] M. Frazier and L. Pitt. CLASSIC learning. *Machine Learning*, 25:151–193, 1996. Was in COLT 1994.
- [1998] A. L. N. Fred. Clustering of sequences using minimum grammar complexity criterion. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 107–116. Springer, 1996.
- [1999] A. L. N. Fred. Computation of substring probabilities in stochastic grammars. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 103–114. Springer, 2000.
- [2000] E. Fredkin. Techniques using LISP for automatically discovering interesting relations in data. In *The Programming Language LISP*, pages 108–124. MIT Press, 1964.
- [2001] J. A. Freeman. Neural networks in mathematica. *AI Expert: the Magazine of Artificial Intelligence in Practice*, 7(11):26–35, Nov. 1992.
- [2002] D. Freitag. Multistrategy learning for information extraction. In *Proc. 15th International Conf. on Machine Learning*, pages 161–169. Morgan Kaufmann, San Francisco, CA, 1998.
- [2003] D. Freitag. Machine learning for information extraction in informal domains. *Machine Learning*, 39(2/3):169–202, 2000.
- [2004] R. V. Freivald. Minimal Gödel numbers and their identification in the limit. In *Mathematical Foundations of Computer Science, 1975, 4th Symposium, Mariánské Lázně, September 1-5, 1975, Proceedings*, volume 32 of *Lecture Notes in Computer Science*, pages 219–225. Springer-Verlag, 1975.

- [2005] R. V. Freivalds and R. Wiehagen. Inductive inference with additional information. *Elektronische Informationsverarbeitung und Kybernetik*, 15(4):179–185, 1979.
- [2006] R. Freivalds. On the complexity and optimality of computation in the limit. In *Theory of Algorithms and Programs*, volume 1, pages 155–173. Latvian State University, Riga, 1974. (in Russian).
- [2007] R. Freivalds. Uniform and non uniform predictability. In *Theory of Algorithms and Programs*, volume 1, pages 89–100. Latvian State University, Riga, 1974.
- [2008] R. Freivalds. On the principle capabilities of probabilistic algorithms in inductive inference. *Semiotika Inform*, 12:137–140, 1979.
- [2009] R. Freivalds. An answer to an open problem. *Bulletin of the EATCS*, 23:31–32, 1984.
- [2010] R. Freivalds. Recursiveness of the enumerating functions increases the inferrability of recursively enumerable sets. *Bulletin of the European Association for Theoretical Computer Science*, 27:35–40, 1985.
- [2011] R. Freivalds. Inductive inference of minimal programs. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 3–20, San Mateo, CA, 1990. Morgan Kaufmann.
- [2012] R. Freivalds. Inductive inference of recursive functions: Qualitative theory. In *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 77–110. Springer, Berlin, 1991.
- [2013] R. Freivalds, J. Bārzdīņš, and K. Podnieks. Inductive inference of recursive functions: Complexity bounds. In *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 111–155. Springer, Berlin, 1991.
- [2014] R. Freivalds and R. F. Bonner. Quantum inductive inference by finite automata. *Theoret. Comput. Sci.*, 397(1-3):70–76, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [2015] R. Freivalds, O. Botuscharov, and R. Wiehagen. Identifying nearly minimal Gödel numbers from additional information. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):199–209, 1998. Special issue for AII '94 and ALT '94.
- [2016] R. Freivalds, D. Gobleja, M. Karpinski, and C. H. Smith. Co-learnability and FIN-identifiability of enumerable classes of total recursive functions. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 100–105. Springer-Verlag, 1994.
- [2017] R. Freivalds and A. G. Hoffmann. An inductive inference approach to classification. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 187–196. Springer-Verlag, 1992.
- [2018] R. Freivalds and A. G. Hoffmann. An inductive inference approach to classification. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):63–72, 1994. Special issue for AII '92.
- [2019] R. Freivalds and S. Jain. Kolmogorov numberings and minimal identification. *Theoret. Comput. Sci.*, 188(1-2):175–194, 1997. first in EuroColt '95.

- [2020] R. Freivalds, M. Karpinski, and C. H. Smith. Co-learning of total recursive functions. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 190–197, New York, NY, 1994. ACM Press.
- [2021] R. Freivalds, M. Karpinski, C. H. Smith, and R. Wiehagen. Learning by the process of elimination. *Inform. Comput.*, 176(1):37–50, 2002.
- [2022] R. Freivalds, E. Kinber, and C. H. Smith. On the impact of forgetting on learning machines. *J. ACM*, 42(6):1146–1168, 1995. Earlier version in 6th COLT, 1993.
- [2023] R. Freivalds, E. Kinber, and C. H. Smith. On the intrinsic complexity of learning. *Inform. Comput.*, 123(1):64–71, 1995. Earlier version in EUROCOLT 1995.
- [2024] R. Freivalds, E. Kinber, and C. H. Smith. The functions of finite support: a canonical learning problem. *J. of Experimental and Theoret. Artif. Intell.*, 11(4):543–552, 1999.
- [2025] R. Freivalds and E. B. Kinber. Limit identification of minimal Gödel numbers. In *Theory of Algorithms and Programs*, volume 3, pages 3–34. Latvian State University, Riga, 1977.
- [2026] R. Freivalds, E. B. Kinber, and C. H. Smith. Probabilistic versus deterministic memory limited learning. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 155–161. Springer, 1995.
- [2027] R. Freivalds, E. B. Kinber, and R. Wiehagen. Inductive inference and computable one-one numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 28:463–479, 1982.
- [2028] R. Freivalds, E. B. Kinber, and R. Wiehagen. Connections between identifying functionals, standardizing operations, and computable numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 30(9–11):145–164, 1984.
- [2029] R. Freivalds, E. B. Kinber, and R. Wiehagen. Probabilistic versus deterministic inductive inference in nonstandard numberings. *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 34:531–539, 1988.
- [2030] R. Freivalds, E. B. Kinber, and R. Wiehagen. Inductive inference from good examples. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 1–17. Springer-Verlag, 1989.
- [2031] R. Freivalds, E. B. Kinber, and R. Wiehagen. On the power of inductive inference from good examples. *Theoret. Comput. Sci.*, 110(1):131–144, 1993.
- [2032] R. Freivalds, E. B. Kinber, and R. Wiehagen. Error detecting in inductive inference. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 25–48. Springer, 1995.
- [2033] R. Freivalds, E. B. Kinber, and R. Wiehagen. How inductive inference strategies discover their errors. *Inform. Comput.*, 118(2):208–226, 1995.
- [2034] R. Freivalds, E. B. Kinber, and R. Wiehagen. Learning from good examples. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 49–62. Springer, 1995.
- [2035] R. Freivalds and C. Smith. Memory limited inductive inference machines. In *Proc. 3rd Scandinavian Workshop on Algorithm Theory, SWAT '92*, volume 621 of *Lecture Notes in Computer Science*, pages 19–29. Springer-Verlag, 1992.

- [2036] R. Freivalds and C. H. Smith. On the duality between mechanistic learners and what it is they learn. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 137–149. Springer-Verlag, 1993.
- [2037] R. Freivalds and C. H. Smith. On the role of procrastination for machine learning. *Inform. Comput.*, 107(2):237–271, 1993.
- [2038] R. Freivalds and C. H. Smith. Category, measure, inductive inference: A triality theorem and its applications. In *LATIN 2002: Theoretical Informatics, 5th Latin American Symposium, Cancun, Mexico, April 3-6, 2002. Proceedings*, volume 2286 of *Lecture Notes in Computer Science*, pages 386–399. Springer, 2002.
- [2039] R. Freivalds, C. H. Smith, and M. Velauthapillai. Trade-off among parameters affecting inductive inference. *Inform. Comput.*, 82(3):323–349, 1989.
- [2040] R. Freivalds and J. Vīksna. Inductive inference up to immune sets. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 138–147. Springer-Verlag, 1989.
- [2041] R. Freivalds and T. Zeugmann. Co-learning of recursive languages from positive data. In *Perspectives of System Informatics, Second International Andrei Ershov Memorial Conference, Akademgorodok, Novosibirsk, Russia, June 1996, Proceedings*, volume 1181 of *Lecture Notes in Computer Science*, pages 122–133. Springer, 1996.
- [2042] R. Freivalds and T. Zeugmann. On the amount of nonconstructivity in learning recursive functions. In M. Ogihara and J. Tarui, editors, *Theory and Applications of Models of Computation, 8th Annual Conference, TAMC 2011, Tokyo, Japan, May 23-25, 2011, Proceedings*, volume 6648 of *Lecture Notes in Computer Science*, pages 332–343. Springer, 2011.
- [2043] R. Freivalds and T. Zeugmann. Active learning of classes of recursive functions by ultrametric algorithms. Technical Report TCS-TR-A-13-68, Division of Computer Science, Hokkaido University, 2013.
- [2044] R. Freivalds and T. Zeugmann. Active learning of recursive functions by ultrametric algorithms. In V. Geffert, B. Preneel, B. Rován, J. Štuller, and A. M. Tjoa, editors, *SOFSEM 2014: Theory and Practice of Computer Science, 40th International Conference on Current Trends in Theory and Practice of Computer Science, Nový Smokovec, Slovakia, January 26–29, 2014, Proceedings*, volume 8327 of *Lecture Notes in Computer Science*, pages 246–257. Cham, Heidelberg, New York, Dordrecht, London, 2014. Springer International Publishing Switzerland.
- [2045] R. V. Freivalds. On the limit synthesis of numbers of general recursive functions in various computable numerations. *Soviet Math. Doklady*, 15:1681–1683, 1974.
- [2046] R. V. Freivalds. Functions computable in the limit by probabilistic machines. In *Proc. 3rd Symposium on Mathematical Foundations of Computer Science, June 17-22, 1974*, volume 28 of *Lecture Notes in Computer Science*, pages 77–87. Springer-Verlag, 1975.
- [2047] R. V. Freivalds. Finite identification of general recursive functions by probabilistic strategies. In *Fundamentals of Computations Theory, FCT' 79, Proc. of the Conference on Algebraic, Arithmetic, and Categorical Methods in Computation Theory*, pages 138–145, Berlin, 1979. Akademie-Verlag.

- [2048] Y. Freund. An improved boosting algorithm and its implications on learning complexity. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 391–398, New York, NY, 1992. ACM Press.
- [2049] Y. Freund. Boosting a weak learning algorithm by majority. *Inform. Comput.*, 121(2):256–285, Sept. 1995. Also appeared in COLT90.
- [2050] Y. Freund. Predicting a binary sequence almost as well as the optimal biased coin. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 89–98, New York, NY, 1996. ACM Press.
- [2051] Y. Freund. Self bounding learning algorithms. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 247–258, New York, NY, 1998. ACM Press.
- [2052] Y. Freund. An adaptive version of the boost by majority algorithm. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 102–113, New York, NY, 1999. ACM Press.
- [2053] Y. Freund. An adaptive version of the boost by majority algorithm. *Machine Learning*, 43(3):293–318, 2001.
- [2054] Y. Freund. Predicting a binary sequence almost as well as the optimal biased coin. *Inform. Comput.*, 182(1):73–94, 2003.
- [2055] Y. Freund. Open problem: Second order regret bounds based on scaling time. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1651–1654, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2056] Y. Freund, L. Györfi, G. Turán, and T. Zeugmann, editors. *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2008. Springer.
- [2057] Y. Freund and D. Haussler. Unsupervised learning of distributions on binary vectors using two layer networks. Technical Report UCSC-CRL-91-20, Univ. of Calif. Computer Research Lab, Santa Cruz, CA, 1991.
- [2058] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. In *Proc. 15th International Conf. on Machine Learning*, pages 170–178. Morgan Kaufmann, San Francisco, CA, 1998.
- [2059] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. *Journal of Machine Learning Research*, 4:933–969, 2003.
- [2060] Y. Freund, M. Kearns, Y. Mansour, D. Ron, and R. Rubinfeld. Efficient algorithms for learning to play repeated games against computationally bounded adversaries. In *Proc. of the 36th Annual Symposium on Foundations of Computer Science*, pages 332–341. IEEE Computer Society Press, Los Alamitos, CA, 1995.
- [2061] Y. Freund, M. Kearns, D. Ron, R. Rubinfeld, R. Schapire, and L. Sellie. Efficient learning of typical finite automata from random walks. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 315–324, New York, NY, 1993. ACM Press.



- [2062] Y. Freund and Y. Mansour. Learning under persistent drift. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 109–118. Springer, 1997.
- [2063] Y. Freund and Y. Mansour. Estimating a mixture of two product distributions. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 53–62, New York, NY, 1999. ACM Press.
- [2064] Y. Freund and L. Mason. The alternating decision tree learning algorithm,. In *Proc. 16th International Conf. on Machine Learning*, pages 124–133. Morgan Kaufmann, San Francisco, CA, 1999.
- [2065] Y. Freund and M. Opper. Continuous drifting games. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 126–132. Morgan Kaufmann, San Francisco, 2000.
- [2066] Y. Freund and M. Opper. Drifting games and brownian motion. *J. Comput. Syst. Sci.*, 64(1):113–132, 2002. Special issue for COLT 2000.
- [2067] Y. Freund, A. Orlitsky, P. Santhanam, and J. Zhang. Universal coding of zipf distributions. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 736–737. Springer, 2003.
- [2068] Y. Freund and D. Ron. Learning to model sequences generated by switching distributions. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 41–50, New York, NY, 1995. ACM Press.
- [2069] Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *Proc. 13th International Conference on Machine Learning*, pages 148–156. Morgan Kaufmann, 1996.
- [2070] Y. Freund and R. E. Schapire. Game theory, on-line prediction and boosting. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 325–332, New York, NY, 1996. ACM Press.
- [2071] Y. Freund and R. E. Schapire:. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 55(1):119–139, 1997. Special issue for EuroCOLT '95.
- [2072] Y. Freund and R. E. Schapire. Large margin classification using the perceptron algorithm. *Machine Learning*, 37(3):277–296, 1999.
- [2073] Y. Freund, H. S. Seung, E. Shamir, and N. Tishby. Selective sampling using the query by committee algorithm. *Machine Learning*, 28(2-3):133–168, 1997.
- [2074] Y. Freund and R. Shapire, editors. *Proceedings of the Tenth Annual Conference on Computational Learning Theory*, New York, N.Y., 1997. ACM Press.
- [2075] P. W. Frey and D. J. Slate. Letter recognition using holland-style adaptive classifiers. *Machine Learning*, 6:161–182, 1991.
- [2076] D. D. Freydenberger and D. Reidenbach. Bad news on decision problems for patterns. In *Developments in Language Theory, 12th International Conference, DLT 2008, Kyoto, Japan, September 16-19, 2008, Proceedings*, volume 5257 of *Lecture Notes in Computer Science*, pages 327–338, Berlin, 2008. Springer.

- [2077] D. D. Freydenberger and D. Reidenbach. Inferring descriptive generalisations of formal languages. *J. Comput. Syst. Sci.*, 79(5):622–639, 2013.
- [2078] C. Friedman and S. Sandow. Learning probabilistic models: An expected utility maximization approach. *Journal of Machine Learning Research*, 4:257–291, 2003.
- [2079] J. H. Friedman and W. Stuetzle. Projection pursuit regression. *J. American Statistical Association*, 76(376):817–823, Dec. 1981.
- [2080] N. Friedman. Learning belief networks in the presence of missing values and hidden variables. In *Proc. 14th International Conference on Machine Learning*, pages 125–133. Morgan Kaufmann, 1997.
- [2081] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29(2-3):131–163, 1997.
- [2082] N. Friedman and M. Goldszmidt. Discretizing continuous attributes while learning Bayesian networks. In *Proc. 13th International Conference on Machine Learning*, pages 157–165. Morgan Kaufmann, 1996.
- [2083] N. Friedman, M. Goldszmidt, and T. J. Lee. Bayesian network classification with continuous attributes: getting the best of both discretization and parametric fitting. In *Proc. 15th International Conf. on Machine Learning*, pages 179–187. Morgan Kaufmann, San Francisco, CA, 1998.
- [2084] N. Friedman and D. Koller. Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian networks. *Machine Learning*, 50(1-2):95–125, 2003.
- [2085] H. Friedrich, S. Münch, R. Dillman, S. Bocionek, and M. Sassin. Robot programming by demonstration (RPD): Supporting the induction by human interaction. *Machine Learning*, 23:163–189, 1996.
- [2086] T.-T. Frieß, N. Cristianini, and C. Campbell. The Kernel-Adatron algorithm: a fast and simple learning procedure for Support Vector machines. In *Proc. 15th International Conf. on Machine Learning*, pages 188–196. Morgan Kaufmann, San Francisco, CA, 1998.
- [2087] M. Fromont. Model selection by bootstrap penalization for classification. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 285–299. Springer, 2004.
- [2088] M. Fromont. Model selection by bootstrap penalization for classification. *Machine Learning*, 66(2-3):165–207, 2007.
- [2089] M. Fromont, B. Laurent, M. Lerasle, and P. Reynaud-Bouret. Kernels based tests with non-asymptotic bootstrap approaches for two-sample problems. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 23.1–23.23, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2090] M. Fromont and C. Tuleau. Functional classification with margin conditions. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 94–108, Berlin, 2006. Springer.

- [2091] R. Frongillo, I. Kash, and S. Becker. Open problem: Property elicitation and elicitation complexity. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1655–1658, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2092] R. Frongillo and I. A. Kash. Vector-valued property elicitation. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 710–727, Paris, France, 03–06 July 2015. PMLR.
- [2093] R. Frongillo and A. Nobel. Memoryless sequences for differentiable losses. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 925–939, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2094] B. Fronhöfer and A. Yamamoto. Minimised residue hypotheses in relevant logic. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 278–292. Springer, 2002.
- [2095] B. Fronhöfer and A. Yamamoto. Hypothesis finding with proof theoretical appropriateness criteria. *Theoret. Comput. Sci.*, 350(1):140–162, 2006. Special issue for ALT 2002.
- [2096] R. Frostig, R. Ge, S. M. Kakade, and A. Sidford. Competing with the empirical risk minimizer in a single pass. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 728–763, Paris, France, 03–06 July 2015. PMLR.
- [2097] K. Fu and T. L. Booth. Grammatical inference: Introduction and survey – part 1. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-5(1):85–111, Jan. 1975.
- [2098] K. Fu and T. L. Booth. Grammatical inference: Introduction and survey – part 2. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-5(4):409–423, July 1975.
- [2099] K. S. Fu. *Syntactic Methods in Pattern Recognition*. Academic Press, N. Y., 1974.
- [2100] K. S. Fu. *Syntactic Pattern Recognition, Applications*. Springer-Verlag, 1977.
- [2101] M. Fuchs. Learning proof heuristics by adapting parameters. In *Proc. 12th International Conference on Machine Learning*, pages 235–243. Morgan Kaufmann, 1995.
- [2102] K. Fujita and M. Harao. Proving based on similarity. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 213–223. IOS Press, 1992.
- [2103] T. Fujita, K. Hatano, S. Kijima, and E. Takimoto. Online linear optimization for job scheduling under precedence constraints. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 332–346, Berlin, Heidelberg, New York, 2015. Springer.
- [2104] T. Fujita, K. Hatano, and E. Takimoto. Combinatorial online prediction via metarounding. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 68–82, Berlin, Heidelberg, New York, 2013. Springer.

- [2105] K. Fukumizu. Generalization error of linear neural networks in unidentifiable cases. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 51–62. Springer, 1999.
- [2106] K. Fukumizu, F. R. Bach, and M. I. Jordan. Dimensionality reduction for supervised learning with reproducing kernel hilbert spaces. *Journal of Machine Learning Research*, 5:73–99, 2004.
- [2107] S. Fukunaga, Y. Takabatake, T. I., and H. Sakamoto. Online grammar compression for frequent pattern discovery. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 93–104, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [2108] M. Fukushima and S. Goto. Packet analysis in congested networks. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 600–615. Springer, 2001.
- [2109] M. Fulk. *A Study of Inductive Inference machines*. PhD thesis, SUNY at Buffalo, 1985.
- [2110] M. Fulk. The disjunctive learning problem. *20th Asilomar Conference on Signals, Systems and Computers*, pages 343–345, 1986.
- [2111] M. Fulk. Regressiveness. unpublished, 1989.
- [2112] M. Fulk and S. Jain. Learning in the presence of inaccurate information. *Theoret. Comput. Sci.*, 161(1-2):235–261, 1996.
- [2113] M. Fulk, S. Jain, and D. N. Osherson. Open problems in *systems that learn*. *J. of Comput. Syst. Sci.*, 49(3):589–604, 1994.
- [2114] M. A. Fulk. Saving the phenomenon: Requirements that inductive inference machines not contradict known data. *Inform. Comput.*, 79(3):193–209, 1988.
- [2115] M. A. Fulk. Prudence and other conditions on formal language learning. *Inform. Comput.*, 85(1):1–11, 1990.
- [2116] M. A. Fulk. Robust separations in inductive inference. In *Proc. of the 31st Annual Symposium on Foundations of Computer Science*, pages 405–410. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [2117] M. A. Fulk. Inductive inference with additional information. *J. Comput. Syst. Sci.*, 64(2):153–159, 2002.
- [2118] M. A. Fulk and J. Case, editors. *Proceedings of the Third Annual Workshop on Computational Learning Theory*, San Mateo, CA, 1990. Morgan Kaufmann, Inc.
- [2119] M. A. Fulk and S. Jain. Learning in the presence of inaccurate information. In *Proc. of the Second Annual Workshop on Comput. Learning Theory, Santa Cruz, California*, pages 175–188. Morgan Kaufmann, Inc., Aug. 1989.
- [2120] M. A. Fulk and S. Jain. Approximate inference and scientific method. *Inform. Comput.*, 114(2):179–191, 1994.

- [2121] T. Fulton, S. Kasif, and S. Salzberg. Efficient algorithms for finding multi-way splits for decision trees. In *Proc. 12th International Conference on Machine Learning*, pages 244–251. Morgan Kaufmann, 1995.
- [2122] G. M. Fung, O. L. Mangasarian, and J. W. Shavlik. Knowledge-based nonlinear kernel classifiers. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 102–113. Springer, 2003.
- [2123] G. M. Fung, O. L. Mangasarian, and A. J. Smola;. Minimal kernel classifiers. *Journal of Machine Learning Research*, 3:303–321, 2002.
- [2124] J. Fürnkranz. Pruning algorithms for rule learning. *Machine Learning*, 27(2):139–171, 1997.
- [2125] J. Fürnkranz. Round robin rule learning. In *Proc. 18th International Conf. on Machine Learning*, pages 146–153. Morgan Kaufmann, San Francisco, CA, 2001.
- [2126] J. Fürnkranz. A pathology of bottom-up hill-climbing in inductive rule learning. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 263–277. Springer, 2002.
- [2127] J. Fürnkranz. Round robin classification. *Journal of Machine Learning Research*, 2:721–747, 2002.
- [2128] J. Fürnkranz and G. Widmer. Incremental reduced error pruning. In *Proc. 11th International Conference on Machine Learning*, pages 70–77. Morgan Kaufmann, 1994.
- [2129] M. L. Furst, J. C. Jackson, and S. W. Smith. Improved learning of  $AC^0$  functions. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 317–325. Morgan Kaufmann, 1991.
- [2130] K. Furukawa. Preliminary study on program synthesis based on induction and verification. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 299–306. IOS Press, 1992.
- [2131] S. Furuya and S. Miyano. Analogy is NP-hard. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 207–212. IOS Press, 1992.
- [2132] Z. Gábor, Z. Kalmár, and C. Szepesvári. Multi-criteria reinforcement learning. In *Proc. 15th International Conf. on Machine Learning*, pages 197–205. Morgan Kaufmann, San Francisco, CA, 1998.
- [2133] P. Gács, J. Tromp, and P. Vitányi. Towards an algorithmic statistics. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 41–55. Springer, 2000.
- [2134] S. C. Gadohan. Learning behavior-selection by emotions and cognition in a multi-goal robot task. *Journal of Machine Learning Research*, 4:385–412, 2003.
- [2135] L. Gaga, V. Moustakis, G. Charissis, and S. Orphanoudakis. IDDD: An inductive, domain dependent decision algorithm. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 408–413. Springer-Verlag, 1993.

- [2136] P. Gaillard and S. Gerchinovitz. A chaining algorithm for online nonparametric regression. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 764–796, Paris, France, 03–06 July 2015. PMLR.
- [2137] P. Gaillard, G. Stoltz, and T. van Erven. A second-order bound with excess losses. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 176–196, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2138] B. R. Gaines. Maryanski’s grammatical inferencer. *IEEE Trans. Computers*, C-27(1):62–64, Jan. 1979.
- [2139] P. Gajane, T. Urvoy, and E. Kaufmann. Corrupt bandits for preserving local privacy. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 387–412. PMLR, 07–09 Apr. 2018.
- [2140] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*. R. E. Krieger, Malabar, FL, 1987. Second Edition.
- [2141] M. Gallagher. Fitness distance correlation of neural network error surfaces: A scalable, continuous optimization problem. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 157–166. Springer, 2001.
- [2142] I. Galperin. Analysis of greedy expert hiring and an application to memory-based learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 217–223, New York, NY, 1996. ACM Press.
- [2143] J. Gama. Probabilistic linear tree. In *Proc. 14th International Conference on Machine Learning*, pages 134–142. Morgan Kaufmann, 1997.
- [2144] J. Gama. Local cascade generalization. In *Proc. 15th International Conf. on Machine Learning*, pages 206–214. Morgan Kaufmann, San Francisco, CA, 1998.
- [2145] J. Gama. Discriminant trees. In *Proc. 16th International Conf. on Machine Learning*, pages 134–142. Morgan Kaufmann, San Francisco, CA, 1999.
- [2146] J. Gama. Iterative Bayes. *Theoret. Comput. Sci.*, 292(2):417–430, 2003.
- [2147] J. Gama and P. Brazdil. Cascade generalization. *Machine Learning*, 41(3):315–343, 2000.
- [2148] P. Gamallo, G. P. Lopes, and J. F. D. Silva. A divide-and-conquer approach to acquire syntactic categories. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 151–162. Springer, 2004.
- [2149] D. Gamarnik. Efficient learning of monotone concepts via quadratic optimization. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 134–143, New York, NY, 1998. ACM Press.
- [2150] D. Gamarnik. Extension of the PAC framework to finite and countable Markov chains. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 308–317, New York, NY, 1999. ACM Press.

- [2151] D. Gamarnik, Q. Li, and H. Zhang. Matrix completion from  $O(n)$  samples in linear time. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 940–947, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2152] L. M. Gambardella and M. Dorigo. Ant-Q:a reinforcement learning approach to the traveling salesman problem. In *Proc. 12th International Conference on Machine Learning*, pages 252–260. Morgan Kaufmann, 1995.
- [2153] D. Gamberger, N. Lavrač, and S. Džeroski. Noise elimination in inductive concept learning: a case study in medical diagnosis. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 199–212. Springer, 1996.
- [2154] D. Gamberger, N. Lavrač, and C. Grošelj. Experiments with noise filtering in a medical domain. In *Proc. 16th International Conf. on Machine Learning*, pages 143–151. Morgan Kaufmann, San Francisco, CA, 1999.
- [2155] A. Gammerman and V. Vovk. Prediction algorithms and confidence measures based on algorithmic randomness theory. *Theoret. Comput. Sci.*, 287(1):209–217, 2002.
- [2156] M. Gams, M. Bohanec, and B. Cestnik. A schema for using multiple knowledge. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 10, pages 157–170. MIT Press, 1994.
- [2157] M. Gams and N. Lavrac. Review of five empirical learning systems within a proposed schemata. In *Progress in Machine Learning, Proceedings of EWSL 87: 2nd European Working Session on Learning, Bled, Yugoslavia, May 1987*, pages 46–66, Wilmslow, May 1987. Sigma Press.
- [2158] J.-G. Ganascia. Algebraic structure of some learning systems. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 398–409. Springer-Verlag, 1993.
- [2159] J.-G. Ganascia. Extraction of recurrent patterns from stratified ordered trees. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 167–178. Springer, 2001.
- [2160] J.-G. Ganascia, J. Thomas, and P. Laublet. Integrating models of knowledge and machine learning. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 396–401. Springer-Verlag, 1993.
- [2161] B. Gangly. The degenerate science of machine learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1213–1216. Morgan Kaufmann, San Francisco, CA, 2000.
- [2162] J. B. Gao, S. R. Gunn, C. J. Harris, and M. Brown. A probabilistic framework for SVM regression and error bar estimation. *Machine Learning*, 46(1-3):71–89, 2002.
- [2163] Q. Gao and M. Li. An application of minimum description length principle to online recognition of handprinted alphanumerals. In *Proc. 11th International Joint Conference on Artificial Intelligence*, pages 843–848. Morgan Kaufmann, 1989.

- [2164] W. Gao and Z.-H. Zhou. Approximation stability and boosting. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 59–73, Berlin, Heidelberg, New York, 2010. Springer.
- [2165] W. Gao and Z.-H. Zhou. On the consistency of multi-label learning. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 341–358, Budapest, Hungary, 09–11 June 2011. PMLR.
- [2166] Z. Gao, S. Jain, F. Stephan, and T. Zeugmann. On the help of bounded shot verifiers, comparers, and standardisers in inductive inference. Technical Report TCS-TR-A-17-82, Division of Computer Science, Hokkaido University, 2017.
- [2167] Z. Gao, S. Jain, F. Stephan, and T. Zeugmann. On the help of bounded shot verifiers, comparators and standardisers for learnability in inductive inference. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 413–437. PMLR, 07–09 Apr. 2018.
- [2168] Z. Gao, D. Kirkpatrick, C. Ries, H. Simon, and S. Zilles. Preference-based teaching of unions of geometric objects. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 185–207, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2169] Z. Gao, C. Ries, H. Simon, and S. Zilles. Preference-based teaching. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 971–997, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2170] Z. Gao, H. U. Simon, and S. Zilles. On the teaching complexity of linear sets. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 102–116, Berlin, Heidelberg, New York, 2015. Springer.
- [2171] Z. Gao, H. U. Simon, and S. Zilles. On the teaching complexity of linear sets. *Theoret. Comput. Sci.*, 716:107–123, 2018. Special Issue on ALT 2015.
- [2172] Z. Gao and F. Stephan. Confident and consistent partial learning of recursive functions. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 51–65, Berlin, Heidelberg, New York, 2012. Springer.
- [2173] Z. Gao and F. Stephan. Confident and consistent partial learning of recursive functions. *Theoret. Comput. Sci.*, 558:5–17, 2014. Special Issue: Algorithmic Learning Theory.
- [2174] Z. Gao, F. Stephan, and S. Zilles. Partial learning of recursively enumerable languages. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 113–127, Berlin, Heidelberg, New York, 2013. Springer.



- [2175] Z. Gao, F. Stephan, and S. Zilles. Combining models of approximation with partial learning. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 56–70, Berlin, Heidelberg, New York, 2015. Springer.
- [2176] Z. Gao, F. Stephan, and S. Zilles. Partial learning of recursively enumerable languages. *Theoret. Comput. Sci.*, 620:15–32, 2016. Special Issue: Algorithmic Learning Theory.
- [2177] D. García, A. González, and R. Pérez. A feature construction approach for genetic iterative rule learning algorithm. *J. Comput. Syst. Sci.*, 80(1):101–117, 2014.
- [2178] F. Garcia and S. M. Ndiaye. A learning rate analysis of reinforcement learning algorithms in finite-horizon. In *Proc. 15th International Conf. on Machine Learning*, pages 215–223. Morgan Kaufmann, San Francisco, CA, 1998.
- [2179] P. Garcia, A. Cano, and J. Ruiz. A comparative study of two algorithms for automata identification. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 115–126. Springer, 2000.
- [2180] P. García, D. López, and M. Vázquez de Parga. Polynomial characteristic sets for DFA identification. *Theoret. Comput. Sci.*, 448:41–46, 2012.
- [2181] P. García, M. Vázquez de Parga, G. I. Álvarez, and J. Ruiz. Learning regular languages using nondeterministic finite automata. In *Implementation and Applications of Automata, 13th International Conference, CIAA 2008, San Francisco, California, USA, July 21-24, 2008. Proceedings*, volume 5148 of *Lecture Notes in Computer Science*, pages 92–101. Springer, 2008.
- [2182] P. García, M. Vázquez de Parga, D. López, and J. Ruiz. Learning automata teams. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 52–65. Springer, 2010.
- [2183] D. García-García and R. C. Williamson. Divergences and risks for multiclass experiments. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 28.1–28.20, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2184] A. Garg and D. Roth. Learning coherent concepts. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 135–150. Springer, 2001.
- [2185] A. Garg and D. Roth. Understanding probabilistic classifiers. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 179–191. Springer, 2001.
- [2186] A. Garivier and O. Cappé. The KL-UCB algorithm for bounded stochastic bandits and beyond. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 359–376, Budapest, Hungary, 09–11 June 2011. PMLR.

- [2187] A. Garivier and E. Kaufmann. Optimal best arm identification with fixed confidence. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 998–1027, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2188] A. Garivier, E. Kaufmann, and W. M. Koolen. Maximin action identification: A new bandit framework for games. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1028–1050, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2189] A. Garivier and E. Moulines. On upper-confidence bound policies for switching bandit problems. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 174–188, Berlin, Heidelberg, New York, 2011. Springer.
- [2190] T. Gärtner, P. Flach, and S. Wrobel. On graph kernels: Hardness results and efficient alternatives. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 129–143. Springer, 2003.
- [2191] T. Gärtner and P. A. Flach.  $WBC_{SVM}$ : Weighted Bayesian classification based on support vector machines. In *Proc. 18th International Conf. on Machine Learning*, pages 154–161. Morgan Kaufmann, San Francisco, CA, 2001.
- [2192] D. Garvin and P. Rayner. Probabilistic self-structuring and learning. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 15, pages 245–278. MIT Press, 1997.
- [2193] W. Gasarch. Memorial issue for carl smith. *J. Comput. Syst. Sci.*, 74(4):408–408, 2008.
- [2194] W. Gasarch, J. Glenn, and C. P. Kruskal. Finding large 3-free sets I: The small  $n$  case. *J. Comput. Syst. Sci.*, 74(4):628–655, 2008.
- [2195] W. Gasarch and G. R. Hird. Reductions for learning via queries. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 152–161, New York, NY, 1995. ACM Press.
- [2196] W. Gasarch and G. R. Hird. Automata techniques for query inference machines. *Annals of Pure and Applied Logic*, 117(1-3):169–201, 2002.
- [2197] W. Gasarch, E. Kinber, M. Pleszkoch, C. Smith, and T. Zeugmann. Learning via queries, teams, and anomalies. *Fundamenta Informaticae*, 23(1):67–89, 1995. Shorter version in Third Annual Workshop on Computational Learning Theory, 1990, pages 327–337, published by Morgan Kaufman.
- [2198] W. Gasarch and M. Li. Editors’ foreword. *J. of Comput. Syst. Sci.*, 52(3):391, 1996.
- [2199] W. Gasarch, M. Pleszkoch, and R. Solovay. Learning via queries to  $[+, <]$ . *J. Symbolic Logic*, 57(1):53–81, Mar. 1992.
- [2200] W. Gasarch, M. G. Pleszkoch, F. Stephan, and M. Velauthapillai. Classification using information. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):147–168, 1998. Special issue for AII ’94 and ALT ’94.

- [2201] W. Gasarch and C. H. Smith. A survey of inductive inference with an emphasis on learning via queries. In A. Sorbi, editor, *Complexity, Logic, and Recursion Theory*, volume 187 of *Lecture Notes in Pure and Applied Mathematics*, pages 225–260. Marcel Dekker, Inc., 1997.
- [2202] W. I. Gasarch and A. C. Y. Lee. Inferring answers to queries. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 275–284, New York, NY, 1997. ACM Press.
- [2203] W. I. Gasarch and A. C. Y. Lee. Inferring answers to queries. *J. Comput. Syst. Sci.*, 74(4):490–512, 2008.
- [2204] W. I. Gasarch, M. G. Pleszkoch, and M. Velauthapillai. Classification using information. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 290–300. Springer-Verlag, 1994.
- [2205] W. I. Gasarch, M. G. Pleszkoch, and M. Velauthapillai. Classification using information. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 162–173. Springer, 1995.
- [2206] W. I. Gasarch, R. K. Sitaraman, C. H. Smith, and M. Velauthapillai. Learning programs with an easy to calculate set of errors. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 124–137. Springer-Verlag, 1989.
- [2207] W. I. Gasarch, R. K. Sitaraman, C. H. Smith, and M. Velauthapillai. Learning programs with an easy to calculate set of errors. *Fundamenta Informaticae*, pages 355–370, 1992.
- [2208] W. I. Gasarch and C. H. Smith. On the inference of sequences of functions. In *Analogical and Inductive Inference, International Workshop AII '86, Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 23–41. Springer-Verlag, 1986.
- [2209] W. I. Gasarch and C. H. Smith. Learning via queries. *J. ACM*, 39(3):649–675, July 1992. A shorter version is in 29th FOCS conference, 1988, pp. 130-137.
- [2210] W. I. Gasarch and M. Velauthapillai. Asking questions versus verifiability. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 197–213. Springer-Verlag, 1992.
- [2211] E. Gassiat and E. Gautherat. Identification of noisy linear systems with discrete random input. *IEEE Trans. on Information Theory*, 44(5):1941–1952, 1998.
- [2212] Y. Gat. A learning generalization bound with an application to sparse-representation classifiers. *Machine Learning*, 42(3):233–239, 2001.
- [2213] Y. Gat. A microchoice bound for continuous-space classification algorithms. *Machine Learning*, 53(1-2):5–21, 2003.
- [2214] D. R. Gaur and R. Krishnamurti. Self-duality of bounded monotone Boolean functions and related problems. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 209–223. Springer, 2000.

- [2215] R. Gavaldà. On the power of equivalence queries. In *Computational Learning Theory: Euro-Colt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 193–203, Oxford, 1994. Oxford University Press.
- [2216] R. Gavaldà and D. Guijarro. Learning ordered binary decision diagrams. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 228–238. Springer, 1995.
- [2217] R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors. *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2009. Springer.
- [2218] R. Gavaldà and E. Takimoto. Foreword. *Theoret. Comput. Sci.*, 364(1):1–2, 2006. Special issue for ALT 2003.
- [2219] R. Gavaldà and K. P. J. E. Takimoto, editors. *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2003. Springer.
- [2220] R. Gavaldà, P. Tesson, and D. Thérien. Learning expressions and programs over monoids. *Inform. Comput.*, 204(2):177–209, 2006.
- [2221] R. Gavaldà and D. Thérien. Learning expressions over monoids. In *STACS 2001, 18th Annual Symposium on Theoretical Aspects of Computer Science, Dresden, Germany, February 15–17, 2001, Proceedings*, volume 2010 of *Lecture Notes in Computer Science*, pages 283–293. Springer, 2001.
- [2222] R. Gavaldà and D. Thérien. An algebraic perspective on Boolean function learning. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 201–215, Berlin/Heidelberg, 2009. Springer.
- [2223] D. Gavinsky. Optimally-smooth adaptive boosting and application to agnostic learning. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 98–112. Springer, 2002.
- [2224] D. Gavinsky. Optimally-smooth adaptive boosting and application to agnostic learning. *Journal of Machine Learning Research*, 4:101–117, 2003.
- [2225] D. Gawronska, B. Schürmann, and J. Hollatz. Learning properties of multi-layer perceptrons with and without feedback. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 17, pages 279–294. MIT Press, 1994.
- [2226] R. Ge, F. Huang, C. Jin, and Y. Yuan. Escaping from saddle points — online stochastic gradient for tensor decomposition. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 797–842, Paris, France, 03–06 July 2015. PMLR.
- [2227] J. Geertzen and M. van Zaanen. Grammatical inference using suffix trees. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 163–174. Springer, 2004.

- [2228] P. Geibel. Reinforcement learning with bounded risk. In *Proc. 18th International Conf. on Machine Learning*, pages 162–169. Morgan Kaufmann, San Francisco, CA, 2001.
- [2229] P. Geibel and F. Wysotzki. Learning relational concepts with decision trees. In *Proc. 13th International Conference on Machine Learning*, pages 166–174. Morgan Kaufmann, 1996.
- [2230] M. Geilke and S. Zilles. Learning relational patterns. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 84–98, Berlin, Heidelberg, New York, 2011. Springer.
- [2231] S. Geisser. The inferential use of predictive distributions. In *Foundations of Statistical Inference*. Holt, Rinehart, and Winston, 1970.
- [2232] K. Gelbrich and R. Nakhaeizadeh. Value miner: A data mining environment for the calculation of the customer lifetime value with application to the automotive industry. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 154–161. Springer, 2000.
- [2233] S. Geman. Stochastic relaxation methods for image restoration and expert systems. In *Automated Image Analysis: Theory and Experiments*. Academic Press, 1986.
- [2234] R. Gemello, F. Mana, and L. Saitta. Rigel: An inductive learning system. *Machine Learning*, 6:7–35, 1991.
- [2235] J. H. Gennari, P. Langley, and D. Fisher. Models of incremental concept formation. *Artificial Intelligence*, 40(1–3):11–61, 1989.
- [2236] C. Gentile. A new approximate maximal margin classification algorithm. *Journal of Machine Learning Research*, 2:213–242, 2001.
- [2237] C. Gentile. Guest editor’s introduction. *Machine Learning*, 51(3):215–216, 2003.
- [2238] C. Gentile. The robustness of the  $p$ -norm algorithms. *Machine Learning*, 53(3):265–299, 2003.
- [2239] C. Gentile and D. P. Helmbold. Improved lower bounds for learning from noisy examples: an information-theoretic approach. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 104–115, New York, NY, 1998. ACM Press.
- [2240] C. Gentile and D. P. Helmbold. Improved lower bounds for learning from noisy examples: An information-theoretic approach. *Inform. Comput.*, 166(2):133–155, May 2001.
- [2241] C. Gentile, M. Herbster, and S. Pasteris. Online similarity prediction of networked data from known and unknown graphs. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 662–695. JMLR.org, 2013.
- [2242] C. Gentile and N. Littlestone. The robustness of the  $p$ -norm algorithms. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th–9th, 1999, Santa Cruz, California*, pages 1–11, New York, NY, 1999. ACM Press.
- [2243] M. G. Genton. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, 2:299–312, 2001.

- [2244] A. P. George and W. B. Powell. Adaptive stepsizes for recursive estimation with applications in approximate dynamic programming. *Machine Learning*, 65(1):167–198, 2006.
- [2245] F. H. George. Logical networks and probability. *Bulletin of Mathematical Biophysics*, 19:187–199, 1957.
- [2246] F. H. George. Inductive machines and the problem of learning. *Cybernetica*, II:109–126, 1959.
- [2247] M. P. Georgeff and C. S. Wallace. A general selection criterion for inductive inference. In *ECAI 84: Advances in Artificial Intelligence*, pages 473–482. Elsevier Science, 1984.
- [2248] S. Gerchinovitz. Sparsity regret bounds for individual sequences in online linear regression. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 377–396, Budapest, Hungary, 09–11 June 2011. PMLR.
- [2249] S. Gerchinovitz and J. Y. Yu. Adaptive and optimal online linear regression on  $\ell^1$ -balls. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 99–113, Berlin, Heidelberg, New York, 2011. Springer.
- [2250] S. Gerchinovitz and J. Y. Yu. Adaptive and optimal online linear regression on  $\ell^1$ -balls. *Theoret. Comput. Sci.*, 519:4–28, 2014. Special Issue: Algorithmic Learning Theory.
- [2251] T. Gergely and Z. Szabó. Fixed point equations as hypotheses in inductive reasoning. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 42–54. Springer-Verlag, 1986.
- [2252] M. T. Gervasio and G. F. DeJong. An incremental learning approach for completable planning. In *Proc. 11th International Conference on Machine Learning*, pages 78–86. Morgan Kaufmann, 1994.
- [2253] M. T. Gervasio, W. Iba, and P. Langley. Learning user evaluation functions for adaptive scheduling assistance. In *Proc. 16th International Conf. on Machine Learning*, pages 152–161. Morgan Kaufmann, San Francisco, CA, 1999.
- [2254] L. Getoor, N. Friedman, D. Koller, and B. Taskar. Learning probabilistic models of relational structure. In *Proc. 18th International Conf. on Machine Learning*, pages 170–177. Morgan Kaufmann, San Francisco, CA, 2001.
- [2255] L. Getoor, N. Friedman, D. Koller, and B. Taskar. Learning probabilistic models of link structure. *Journal of Machine Learning Research*, 3:679–707, 2002.
- [2256] P. Geurts and L. Wehenkel. Investigation and reduction of discretization variance in decision tree induction. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 162–170. Springer, 2000.
- [2257] Z. Ghahramani and M. I. Jordan. Factorial hidden Markov models. *Machine Learning*, 29(2-3):245–273, 1997.

- [2258] Z. Ghahramani and M. I. Jordan. Mixture models for learning from incomplete data. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 5, pages 67–85. MIT Press, 1997.
- [2259] R. Ghani. Using error-correcting codes for text classification. In *Proc. 17th International Conf. on Machine Learning*, pages 303–310. Morgan Kaufmann, San Francisco, CA, 2000.
- [2260] R. Ghani, S. Slattery, and Y. Yang. Hypertext categorization using hyperlink patterns and meta data. In *Proc. 18th International Conf. on Machine Learning*, pages 178–185. Morgan Kaufmann, San Francisco, CA, 2001.
- [2261] M. Ghavamzadeh and S. Mahadevan. Continuous-time hierarchical reinforcement learning. In *Proc. 18th International Conf. on Machine Learning*, pages 186–193. Morgan Kaufmann, San Francisco, CA, 2001.
- [2262] M. Ghosh and S. Nandakumar. Predictive complexity and generalized entropy rate of stationary ergodic processes. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 365–379, Berlin, Heidelberg, New York, 2012. Springer.
- [2263] D. Ghoshdastidar, M. Gutzeit, A. Carpentier, and U. von Luxburg. Two-sample tests for large random graphs using network statistics. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 954–977, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2264] G. Giacinto, F. Roli, and G. Fumera. Adaptive query shifting for content-based image retrieval. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 337–346. Springer, 2001.
- [2265] G. J. Gibson. Exact classification with two-layer neural nets. *J. of Comput. Syst. Sci.*, 52(2):349–356, 1996.
- [2266] N. Gierasimczuk. Bridging learning theory and dynamic epistemic logic. *Synthese*, 169(2):371–384, 2009.
- [2267] N. Gierasimczuk. Learning by erasing in dynamic epistemic logic. In A. H. Dediu, A. M. Ionescu, and C. Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2009, Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 362–373. Springer, 2009.
- [2268] N. Gierasimczuk and D. de Jongh. On the complexity of conclusive update. *The Computer Journal*, 56(3):365–377, 2013.
- [2269] N. Gierasimczuk, L. Kurzen, and F. R. Velázquez-Quesada. Learning and teaching as a game: A sabotage approach. In *Logic, Rationality, and Interaction, Second International Workshop, LORI 2009, Chongqing, China, October 8–11, 2009. Proceedings*, volume 5834 of *Lecture Notes in Artificial Intelligence*, pages 119–132, Heidelberg, 2009. Springer.
- [2270] Y. Gil. Learning by experimentation: incremental refinement of incomplete planning domains. In *Proc. 11th International Conference on Machine Learning*, pages 87–95. Morgan Kaufmann, 1994.

- [2271] R. Gilad-Bachrach, A. Navot, and N. Tishby. An information theoretic tradeoff between complexity and accuracy. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 595–609. Springer, 2003.
- [2272] R. Gilad-Bachrach, A. Navot, and N. Tishby. Bayes and tukey meet at the center point. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 549–563. Springer, 2004.
- [2273] C. L. Giles, S. Lawrence, and A. C. Tsoi. Noisy time series prediction using recurrent neural networks and grammatical inference. *Machine Learning*, 44(1/2):161–183, 2001.
- [2274] D. Gillman and M. Sipser. Inference and minimization of hidden Markov chains. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 147–158, New York, NY, 1994. ACM Press.
- [2275] A. Giordana and G. L. Bello. Learning classification programs: The genetic algorithm approach. *Fundamenta Informaticae*, 35(1/4):163–177, 1998.
- [2276] A. Giordana, F. Neri, L. Saitta, and M. Botta. Integrating multiple learning strategies in first order logics. *Machine Learning*, 27(3):209–240, 1997.
- [2277] A. Giordana and R. Piola. On some misbehaviour of back-propagation with non-normalized RBFNs and a solution. In *Proc. 16th International Conf. on Machine Learning*, pages 162–170. Morgan Kaufmann, San Francisco, CA, 1999.
- [2278] A. Giordana and L. Saitta. Phase transitions in relational learning. *Machine Learning*, 41(2):217–251, 2000.
- [2279] A. Giordana, L. Saitta, M. Sebag, and M. Botta. Analyzing relational learning in the phase transition framework. In *Proc. 17th International Conf. on Machine Learning*, pages 311–318. Morgan Kaufmann, San Francisco, CA, 2000.
- [2280] A. Giordana, L. Saitta, and F. Zini. Learning disjunctive concepts by means of genetic algorithms. In *Proc. 11th International Conference on Machine Learning*, pages 96–104. Morgan Kaufmann, 1994.
- [2281] J. Y. Giordano. Inference of context-free grammars by enumeration: Structural containment as an ordering bias. In *Proc. 2nd International Colloquium on Grammatical Inference - ICGI '94*, volume 862 of *Lecture Notes in Artificial Intelligence*, pages 212–221. Springer-Verlag, 1994.
- [2282] J.-Y. Giordano. Grammatical inference using tabu search. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 292–300. Springer, 1996.
- [2283] P. Giudici and R. Castelo. Improving Markov chain Monte Carlo model search for data mining. *Machine Learning*, 50(1-2):127–158, 2003.
- [2284] M. R. Glickman and K. Sycara. Evolutionary search, stochastic policies with memory, and reinforcement learning with hidden state. In *Proc. 18th International Conf. on Machine Learning*, pages 194–201. Morgan Kaufmann, San Francisco, CA, 2001.



- [2285] A. Globerson, R. Livni, and S. Shalev-Shwartz. Effective semisupervised learning on manifolds. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 978–1003, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2286] A. Globerson and N. Tishby. Sufficient dimensionality reduction. *Journal of Machine Learning Research*, 3:1307–1331, 2003.
- [2287] C. Globig, K. P. Jantke, S. Lange, and Y. Sakakibara. On case-based learnability of languages. *New Generation Computing*, 15(1):59–85, 1997. Special issue for AII '94 and ALT '94.
- [2288] C. Globig and S. Lange. On case-based representability and learnability of languages. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 106–120. Springer-Verlag, 1994.
- [2289] C. Globig and S. Wess. Learning in case-based classification algorithms. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 340–362. Springer, 1995.
- [2290] C. Goble, O. Corcho, P. Alper, and D. D. Roure. e-science and the semantic web: A symbiotic relationship. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, page 12. Springer, Oct. 2006.
- [2291] R. Godin and R. Missaoui. An incremental concept formation approach for learning from databases. *Theoret. Comput. Sci.*, 133(2):387–419, 1994.
- [2292] R. Goebel. A sketch of analogy as reasoning with equality hypotheses. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 243–253. Springer-Verlag, 1989.
- [2293] S. Goel, V. Kanade, A. Klivans, and J. Thaler. Reliably learning the reLU in polynomial time. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1004–1042, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2294] G. J. Goetsch. CONSENSUS: A statistical learning procedure in a connectionist network. Technical Report CMU-CS-86-131, Carnegie-Mellon Computer Science Department, May 1986.
- [2295] P. Goetz, S. Kumar, and R. Miikkulainen. On-line adaptation of a signal predistorter through dual reinforcement learning. In *Proc. 13th International Conference on Machine Learning*, pages 175–181. Morgan Kaufmann, 1996.
- [2296] E. Gofer. Higher-order regret bounds with switching costs. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 210–243, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2297] E. Gofer, N. Cesa-Bianchi, C. Gentile, and Y. Mansour. Regret minimization for branching experts. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 618–638. JMLR.org, 2013.

- [2298] E. Gofer and Y. Mansour. Regret minimization algorithms for pricing lookback options. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 234–248, Berlin, Heidelberg, New York, 2011. Springer.
- [2299] E. Gofer and Y. Mansour. Lower bounds on individual sequence regret. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 275–289, Berlin, Heidelberg, New York, 2012. Springer.
- [2300] N. Goix, A. Sabourin, and S. Cléménçon. Learning the dependence structure of rare events: a non-asymptotic study. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 843–860, Paris, France, 03–06 July 2015. PMLR.
- [2301] E. M. Gold. Limiting recursion. *J. Symbolic Logic*, 30:28–48, 1965.
- [2302] E. M. Gold. Language identification in the limit. *Inform. Control*, 10(5):447–474, 1967.
- [2303] E. M. Gold. Universal goal seekers. *Inform. Control*, 18:395–403, 1971.
- [2304] E. M. Gold. System identification via state characterization. *Automatica*, 8:621–636, 1972.
- [2305] E. M. Gold. Complexity of automaton identification from given data. *Inform. Control*, 37:302–320, 1978.
- [2306] M. Gold. *Models of Goal Seeking and Learning*. PhD thesis, University of California, Los Angeles, 1965.
- [2307] U. Goldammer. Inductive synthesis of rewrite programs. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 435–466. Springer, 1995.
- [2308] D. Goldberg and M. J. Matarić. Learning multiple models for reward maximization. In *Proc. 17th International Conf. on Machine Learning*, pages 319–326. Morgan Kaufmann, San Francisco, CA, 2000.
- [2309] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
- [2310] D. E. Goldberg. Probability matching, the magnitude of reinforcement, and classifier system bidding. *Machine Learning*, 5:407–425, 1990.
- [2311] D. E. Goldberg and J. H. Holland. Genetic algorithms and machine learning. *Machine Learning*, 3:95–99, 1988.
- [2312] L. A. Goldberg, W. E. Hart, and D. B. Wilson. Analysis of a simple learning algorithm: learning foraging thresholds for lizards. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 2–9, New York, NY, 1996. ACM Press.
- [2313] P. Goldberg and S. Kwek. The precision of query points as a resource for learning convex polytopes with membership queries. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 225–235. Morgan Kaufmann, San Francisco, 2000.

- [2314] P. W. Goldberg. Learning fixed-dimension linear thresholds from fragmented data. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 88–99, New York, NY, 1999. ACM Press.
- [2315] P. W. Goldberg. Estimating a Boolean perceptron from its average satisfying assignment: A bound on the precision required. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 116–127. Springer, 2001.
- [2316] P. W. Goldberg. Learning fixed-dimension linear thresholds from fragmented data. *Inform. Comput.*, 171(1):98–122, Nov. 2001.
- [2317] P. W. Goldberg. When can two unsupervised learners achieve PAC separation? In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 303–319. Springer, 2001.
- [2318] P. W. Goldberg and S. A. Goldman. Learning one-dimensional geometric patterns under one-sided random misclassification noise. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 246–255, New York, NY, 1994. ACM Press.
- [2319] P. W. Goldberg, S. A. Goldman, and H. D. Mathais. Learning unions of boxes with membership and equivalence queries. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 198–207, New York, NY, 1994. ACM Press.
- [2320] P. W. Goldberg, S. A. Goldman, and S. D. Scott. PAC learning of one-dimensional patterns. *Machine Learning*, 25:51–70, 1996.
- [2321] P. W. Goldberg and M. R. Jerrum. Bounding the Vapnik-Chervonenkis dimension of concept classes parametrized by real numbers. *Machine Learning*, 18(2/3):131–148, 1995. Special issue for COLT '93.
- [2322] A. R. Golding and D. Roth. Applying Winnow to context-sensitive spelling correction. In *Proc. 13th International Conference on Machine Learning*, pages 182–190. Morgan Kaufmann, 1996.
- [2323] A. R. Golding and D. Roth. A Winnow-based approach to context-sensitive spelling correction. *Machine Learning*, 34:107–130, 1999.
- [2324] S. Goldman. *Learning Binary Relations, Total Orders, and Read-Once Formulas*. PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, Sept. 1990. (MIT Laboratory for Computer Science Technical Report MIT/LCS/TR-483, July 1990).
- [2325] S. Goldman. Special issue: Learning theory 2003. *J. Comput. Syst. Sci.*, 71(3):249, 2005. Special issue Learning Theory 2003.
- [2326] S. Goldman and Y. Zhou. Enhancing supervised learning with unlabeled data. In *Proc. 17th International Conf. on Machine Learning*, pages 327–334. Morgan Kaufmann, San Francisco, CA, 2000.
- [2327] S. A. Goldman. Guest editor's introduction. *Machine Learning*, 18(2/3):127–129, 1995.
- [2328] S. A. Goldman and M. J. Kearns. On the complexity of teaching. *J. of Comput. Syst. Sci.*, 50(1):20–31, 1995. Earlier version in 4th COLT, 1991.

- [2329] S. A. Goldman, M. J. Kearns, and R. E. Schapire. Exact identification of circuits using fixed points of amplification functions. *SIAM J. Comput.*, 22(4):705–726, Aug. 1993.
- [2330] S. A. Goldman, M. J. Kearns, and R. E. Schapire. On the sample complexity of weak learning. *Inform. Comput.*, 117(2):276–287, 1995. Earlier version in 3rd COLT, 1990.
- [2331] S. A. Goldman, S. Kwek, and S. D. Scott. Learning from examples with unspecified attribute values. *Inform. Comput.*, 180(2):82–100, 2003.
- [2332] S. A. Goldman and S. S. Kwek. On learning unions of pattern languages and tree patterns. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 347–363. Springer, 1999.
- [2333] S. A. Goldman and S. S. Kwek. On learning unions of pattern languages and tree patterns in the mistake bound model. *Theoret. Comput. Sci.*, 288(2):237–254, 2002. Special issue for ALT '99.
- [2334] S. A. Goldman, S. S. Kwek, and S. D. Scott. Agnostic learning of geometric patterns. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 325–333, New York, NY, 1997. ACM Press.
- [2335] S. A. Goldman, S. S. Kwek, and S. D. Scott. Learning from examples with unspecified attribute values. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 231–242, New York, NY, 1997. ACM Press.
- [2336] S. A. Goldman, S. S. Kwek, and S. D. Scott. Agnostic learning of geometric patterns. *J. Comput. Syst. Sci.*, 62(1):123–151, 2001.
- [2337] S. A. Goldman and H. D. Mathias. Learning  $k$ -term DNF formulas with an incomplete membership oracle. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 77–84, New York, NY, 1992. ACM Press.
- [2338] S. A. Goldman and H. D. Mathias. Teaching a smarter learner. *J. of Comput. Syst. Sci.*, 52(2):255–267, 1996. Earlier version in 6th COLT, 1993.
- [2339] S. A. Goldman and R. L. Rivest. A non-iterative maximum entropy algorithm. In *Uncertainty in Artificial Intelligence*, volume 2, pages 133–148. Elsevier Science B.V. North-Holland, 1988.
- [2340] S. A. Goldman, R. L. Rivest, and R. E. Schapire. Learning binary relations and total orders. *SIAM J. Comput.*, 22(5):1006–1034, Oct. 1993.
- [2341] S. A. Goldman and S. D. Scott. A theoretical and empirical study of a noise-tolerant algorithm to learn geometric patterns. *Machine Learning*, 37(1):5–49, 1999.
- [2342] S. A. Goldman and R. H. Sloan. The power of self-directed learning. *Machine Learning*, 14(3):271–294, 1994.
- [2343] S. A. Goldman and R. H. Sloan. Can PAC learning algorithms tolerate random attribute noise? *Algorithmica*, 14:70–84, 1995.
- [2344] S. A. Goldman and M. K. Warmuth. Learning binary relations using weighted majority voting. *Machine Learning*, 20(3):245–271, 1995. Earlier version in 6th COLT, 1993.

- [2345] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.
- [2346] O. Goldreich, R. Rubinfeld, and M. Sudan. Learning polynomials with queries: the highly noisy case. In *Proc. of the 36th Annual Symposium on Foundations of Computer Science*, pages 294–303. IEEE Computer Society Press, Los Alamitos, CA, 1995.
- [2347] J. Goldsmith and R. H. Sloan. More theory revision with queries (extended abstract). In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 441–448. ACM Press, 2000.
- [2348] J. Goldsmith, R. H. Sloan, B. Szörényi, and G. Turán. Improved algorithms for theory revision with queries. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 236–247. Morgan Kaufmann, San Francisco, 2000.
- [2349] J. Goldsmith, R. H. Sloan, B. Szörényi, and G. Turán. New revision algorithms. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 395–409. Springer, 2004.
- [2350] J. Goldsmith, R. H. Sloan, B. Szörényi, and G. Turán. Theory revision with queries: Horn, read-once, and parity formulas. *Artificial Intelligence*, 156:139–176, 2004.
- [2351] J. Goldsmith, R. H. Sloan, and G. Turán. Theory revision with queries: DNF formulas. *Machine Learning*, 47(2/3):257–295, 2002.
- [2352] M. Golea. Average case analysis of a learning algorithm for  $\mu$ -DNF expressions. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 342–356. Springer, 1995.
- [2353] M. Golea. On learning the neural network architecture: a case study. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 14, pages 227–243. MIT Press, 1997.
- [2354] M. Golea and M. Marchand. Average case analysis of the clipped Hebb rule for nonoverlapping perceptron networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 151–157, New York, NY, 1993. ACM Press.
- [2355] M. Golea and M. Marchand. On learning simple deterministic and probabilistic neural concepts. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 47–60, Oxford, 1994. Oxford University Press.
- [2356] M. Golea, M. Matsuoka, and Y. Sakakibara. Stochastic simple recurrent neural networks. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 262–273. Springer, 1996.
- [2357] P. Golland, F. Liang, S. Mukherjee, and D. Panchenko. Permutation tests for classification. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 501–515. Springer, 2005.

- [2358] A. C. Gómez. Inferring regular trace languages from positive and negative samples. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 11–23. Springer, 2010.
- [2359] A. C. Gómez and G. I. Alvarez. Learning commutative regular languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 71–83. Springer, 2008.
- [2360] M. Gomez-Rodriguez, L. Song, and B. Schoelkopf. Open problem: Finding good cascade sampling processes for the network inference problem. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1276–1279, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2361] A. Gonem and R. Gilad-Bachrach. Smooth sensitivity based approach for differentially private PCA. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 438–450. PMLR, 07–09 Apr. 2018.
- [2362] A. Gonen and S. Shalev-Shwartz. Fast rates for empirical risk minimization of strict saddle problems. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1043–1063, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2363] R. C. Gonzalez, J. J. Edwards, and M. G. Thomason. An algorithm for the inference of tree grammars. *Internat. J. Comput. Info. Sci.*, 5(2):145–164, 1976.
- [2364] R. C. Gonzalez and M. G. Thomason. *Syntactic Pattern Recognition, An Introduction*. Addison-Wesley, Reading, Mass., 1978.
- [2365] A. Gopalan and S. Mannor. Thompson Sampling for Learning Parameterized Markov Decision Processes. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 861–898, Paris, France, 03–06 July 2015. PMLR.
- [2366] P. Gopalan, A. Kalai, and A. R. Klivans. A query algorithm for agnostically learning DNF?. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 515–516. Omnipress, 2008.
- [2367] P. Gopalan, A. R. Klivans, and R. Meka. Learning functions of halfspaces using prefix covers. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 15.1–15.10, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2368] D. F. Gordon. Well-behaved Borgs, Bolos, and Berserkers. In *Proc. 15th International Conf. on Machine Learning*, pages 224–232. Morgan Kaufmann, San Francisco, CA, 1998.
- [2369] D. F. Gordon and M. desJardins. Evaluation and selection of biases in machine learning. *Machine Learning*, 20(1/2):5–22, 1995.
- [2370] G. J. Gordon. Stable function approximation in dynamic programming. In *Proc. 12th International Conference on Machine Learning*, pages 261–268. Morgan Kaufmann, 1995.

- [2371] G. J. Gordon. Regret bounds for prediction problems. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 29–40, New York, NY, 1999. ACM Press.
- [2372] G. J. Gordon and A. Moore. Learning filaments. In *Proc. 17th International Conf. on Machine Learning*, pages 335–342. Morgan Kaufmann, San Francisco, CA, 2000.
- [2373] G. J. Gordon and A. M. Segre. Nonparametric statistical methods for experimental evaluations of speedup learning. In *Proc. 13th International Conference on Machine Learning*, pages 200–206. Morgan Kaufmann, 1996.
- [2374] J. Gordon and E. H. Shortliffe. A method for managing evidential reasoning in a hierarchical hypothesis space. *Artificial Intelligence*, 26(3):323–357, July 1985.
- [2375] M. Gori and A. Tesi. On the problem of local minima in backpropagation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(1):76–86, 1992.
- [2376] R. P. Gorman and T. J. Sejnowski. Analysis of hidden units in a layered network trained to classify sonar targets. *Neural Networks*, 1:75–89, 1988.
- [2377] Y. Goto, Y. Kasahara, and T. Sato. A flexible modeling of global plasma profile deduced from wave data. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 438–448. Springer, 2001.
- [2378] L.-A. Gottlieb, A. Kontorovich, and R. Krauthgamer. Adaptive metric dimensionality reduction. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 279–293, Berlin, Heidelberg, New York, 2013. Springer.
- [2379] L.-A. Gottlieb, A. Kontorovich, and R. Krauthgamer. Adaptive metric dimensionality reduction. *Theoret. Comput. Sci.*, 620:105–118, 2016. Special Issue: Algorithmic Learning Theory.
- [2380] G. Gottlob, N. Leone, and F. Scarcello. On the complexity of some inductive logic programming problems. *New Generation Computing*, 17(1):53–75, 1999.
- [2381] J. Grabowski. Inductive inference of functions from noised observations. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 55–60. Springer-Verlag, 1986.
- [2382] J. Grabowski. Starke Erkennung. In R. Linder and H. Thiele, editors, *Strukturerkennung diskreter kybernetischer Systeme*, volume 82, pages 168–184. Seminarberichte der Sektion Mathematik der Humboldt-Universität zu Berlin, 1986.
- [2383] T. Graepel, R. Herbrich, and J. Shawe-Taylor. Generalisation error bounds for sparse linear classifiers. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 298–303. Morgan Kaufmann, San Francisco, 2000.
- [2384] H. P. Graf, S. Satyanarayana, and Y. Tsividis. A reconfigurable analog (VLSI neural network chip. In *Proc. 1989 Neural Information Processing Conference*, pages 758–768, 1989.

- [2385] L. Grate, M. Herbster, R. Hughey, I. S. Mian, H. Noller, and D. Haussler. RNA modeling using Gibbs sampling and stochastic context free grammars. In *Proc. of Second Int. Conf. on Intelligent Systems for Molecular Biology*, pages 138–146, Menlo Park, CA, Aug. 1994. AAAI/MIT Press.
- [2386] D. P. Greene and S. F. Smith. Competition-based induction of decision models from examples. *Machine Learning*, 13:229–257, 1993.
- [2387] A. Greenwald and A. Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 2–12. Springer, 2003.
- [2388] A. Greenwald and M. L. Littman. Introduction to the special issue on learning and computational game theory. *Machine Learning*, 67(1-2):3–6, 2007.
- [2389] A. R. Greenwald, Z. Li, and W. Schudy. More efficient internal-regret-minimizing algorithms. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 239–250. Omnipress, 2008.
- [2390] J. Grefenstette. Genetic algorithms and machine learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 3–4, New York, NY, 1993. ACM Press.
- [2391] J. J. Grefenstette. Credit assignment in rule discovery systems based on genetic algorithms. *Machine Learning*, 3:225–245, 1988.
- [2392] J. J. Grefenstette. Learning decision strategies with genetic algorithms. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 35–50. Springer-Verlag, 1992.
- [2393] J. J. Grefenstette, C. L. Ramsey, and A. C. Schultz. Learning sequential decision rules using simulation models and competition. *Machine Learning*, 5:355–381, 1990.
- [2394] J. Gregor. Data-driven inductive inference of finite-state automata. *International Journal of Pattern Recognition and Artificial Intelligence*, 8(1):305–322, 1994.
- [2395] J. Gregor and M. G. Thomason. A disagreement count scheme for inference of constrained Markov networks. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 168–178. Springer, 1996.
- [2396] R. Greiner. Probabilistic hill-climbing: Theory and applications. In *Proc. of the Ninth Canadian Conference on Artificial Intelligence*, pages 60–67, Vancouver, June 1992. Morgan Kaufmann. (Awarded “AIJ Best Paper Prize”).
- [2397] R. Greiner. The challenge of revising an impure theory. In *Proc. 12th International Conference on Machine Learning*, pages 269–277. Morgan Kaufmann, 1995.
- [2398] R. Greiner. The complexity of theory revision. In *Proc. of IJCAI-95*, 1995.
- [2399] R. Greiner. PALO: A probabilistic hill-climbing algorithm. *Artificial Intelligence*, 84(1–2):177–208, July 1996.



- [2400] R. Greiner, A. J. Grove, and A. Kogan. Exploiting the omission of irrelevant data. In *Proc. 13th International Conference on Machine Learning*, pages 216–224. Morgan Kaufmann, 1996.
- [2401] R. Greiner, A. J. Grove, and A. Kogan. Knowing what doesn’t matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, 97(1–2):345–380, 1997.
- [2402] R. Greiner, A. J. Grove, and D. Roth. Learning active classifiers. In *Proc. 13th International Conference on Machine Learning*, pages 207–215. Morgan Kaufmann, 1996.
- [2403] R. Greiner and R. Isukapalli. Learning to select useful landmarks. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 26(3), June 1996. Special issue on Learning Approaches to Autonomous Robots Control (ed., Marco Dorigo). Prior version in AAAI-94.
- [2404] R. Greiner and I. Jurišica. A statistical approach to solving the EBL utility problem. In *Proc. AAAI-92*, San Jose, 1992.
- [2405] R. Greiner and P. Orponen. Probably approximately optimal derivation strategies. In J. A. Allen, R. Fikes, and E. Sandewall, editors, *Proc. of the Second International Joint Conference on Knowledge Representation*, Rochester, NJ, Apr. 1991. Morgan Kaufmann.
- [2406] R. Greiner and P. Orponen. Probably approximately optimal satisficing strategies. *Artificial Intelligence*, 82(1–2):21–44, Apr. 1996.
- [2407] R. Greiner and D. Schuurmans. Learning useful Horn approximations. In *Proc. of the Third International Joint Conference on Knowledge Representation*, Cambridge, MA, Oct. 1992.
- [2408] R. Greiner and D. Schuurmans. Learning an optimally accurate representation system. In G. Lakemeyer and B. Nebel, editors, *Foundations of Knowledge Representation and Reasoning*, volume 810, pages 145–59. Springer-Verlag, June 1994. Extended from article in *ECAI Workshop on Theoretical Foundations of Knowledge Representation and Reasoning*, Vienna, August 1992.
- [2409] R. Greiner, B. Silver, S. Becker, and M. Grüninger. A review of machine learning at AAAI-87. *Machine Learning*, 3(1):79–92, 1988.
- [2410] I. Greitāne. Probabilistic inductive inference of indices in enumerable classes of total recursive functions. In *Analogical and Inductive Inference, International Workshop AII ’89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 277–287. Springer-Verlag, 1989.
- [2411] U. Grenander. *Abstract Inference*. John Wiley and Sons, Inc., 1981.
- [2412] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 63–77. Springer, Oct. 2005.
- [2413] A. Gretton and L. Györfi. Nonparametric independence tests: Space partitioning and kernel approaches. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 183–198, Berlin, Oct. 2008. Springer.
- [2414] G. Grieser.  $\text{TP}^{\text{S}}$  – a term rewriting laboratory (not only) for experiments in automatic program synthesis. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 467–481. Springer, 1995.

- [2415] G. Grieser. Reflecting inductive inference machines and its improvement by therapy. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 325–336. Springer, 1996.
- [2416] G. Grieser. Reflective inductive inference of recursive functions. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 203–217. Springer, 2002.
- [2417] G. Grieser. Reflective inductive inference of recursive functions. *Theoret. Comput. Sci.*, 397(1-3):57–69, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [2418] G. Grieser, K. P. Jantke, and S. Lange. Characterizing sufficient expertise for learning system validation. In *Proc. 11th International Florida Artificial Intelligence Research Symposium Conference - FLAIRS '98*, pages 452–456. AAAI Press, Menlo Park, 1998.
- [2419] G. Grieser, K. P. Jantke, and S. Lange. Towards the validation of inductive learning systems. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 409–423. Springer, 1998.
- [2420] G. Grieser, K. P. Jantke, and S. Lange. Consistency queries in information extraction. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 173–187. Springer, 2002.
- [2421] G. Grieser, K. P. Jantke, S. Lange, and B. Thomas. A unifying approach to HTML wrapper representation and learning. In *Discovery Science, Third International Conference, DS 2000, Kyoto, Japan, December 2000, Proceedings*, volume 1967 of *Lecture Notes in Artificial Intelligence*, pages 50–64. Springer, 2000.
- [2422] G. Grieser, S. Lange, and T. Zeugmann. Learning recursive concepts with anomalies. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 101–115. Springer, 2000.
- [2423] E. Grigorescu, L. Reyzin, and S. Vempala. On noise-tolerant learning of sparse parities and related problems. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 413–424, Berlin, Heidelberg, New York, 2011. Springer.
- [2424] O. Grinchtein, B. Jonsson, and M. Leucker. Learning of event-recording automata. *Theoret. Comput. Sci.*, 411(47):4029–4054, 2010.
- [2425] O. Grinchtein and M. Leucker. Learning finite-state machines from inexperienced teachers. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 344–345, Berlin, 2006. Springer.
- [2426] G. Grindlay and D. Helmbold. Modeling, analyzing, and synthesizing expressive piano performance with graphical models. *Machine Learning*, 65(2-3):361–387, 2006.

- [2427] M. Grohe, C. Löding, and M. Ritzert. Learning MSO-definable hypotheses on strings. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 434–451, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2428] M. Grohe and G. Turán. Learnability and definability in trees and similar structures. In *STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings*, volume 2285 of *Lecture Notes in Computer Science*, pages 645–657. Springer, 2002.
- [2429] S. Grolimund and J.-G. Ganascia. Speeding-up nearest neighbour memories: the template tree case memory organisation. In *Proc. 13th International Conference on Machine Learning*, pages 225–233. Morgan Kaufmann, 1996.
- [2430] A. J. Grove, N. Littlestone, and D. Schuurmans. General convergence results for linear discriminant updates. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 171–183, New York, NY, 1997. ACM Press.
- [2431] A. J. Grove, N. Littlestone, and D. Schuurmans. General convergence results for linear discriminant updates. *Machine Learning*, 43(3):173–210, 2001.
- [2432] A. J. Grove and D. Roth. Linear concepts and hidden variables. *Machine Learning*, 42(1/2):123–141, 2001.
- [2433] R. Groz, C. Oriat, and N. Brémont. Inferring non-resettable Mealy machines with  $n$  states. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 30–41, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [2434] T. R. Gruber. Automated knowledge acquisition for strategic knowledge. *Machine Learning*, 4:293–336, 1989.
- [2435] G. Z. Grudic and L. H. Ungar. Localizing policy gradient estimates to action transitions. In *Proc. 17th International Conf. on Machine Learning*, pages 343–350. Morgan Kaufmann, San Francisco, CA, 2000.
- [2436] P. Grünwald. Viewing all models as ‘probabilistic’. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 171–182, New York, NY, 1999. ACM Press.
- [2437] P. Grünwald. Strong entropy concentration, game theory and algorithmic randomness. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 320–336. Springer, 2001.
- [2438] P. Grünwald. The catch-up phenomenon in Bayesian inference. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 1–2. Omnipress, 2008.
- [2439] P. Grünwald. Safe learning: bridging the gap between Bayes, MDL and statistical learning theory via empirical convexity. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 397–420, Budapest, Hungary, 09–11 June 2011. PMLR.

- [2440] P. Grünwald. The safe Bayesian: Learning the learning rate via the mixability gap. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 169–183, Berlin, Heidelberg, New York, 2012. Springer.
- [2441] P. Grünwald and E. Hazan. Conference on learning theory 2015: Preface. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1–3, Paris, France, 03–06 July 2015. PMLR.
- [2442] P. Grünwald, E. Hazan, and S. Kale, editors. *Conference on Learning Theory, 3-6 July 2015, Paris, France*, volume 40 of *Proceedings of Machine Learning Research*. PMLR, 2015.
- [2443] P. Grünwald and J. Langford. Suboptimal behavior of Bayes and MDL in classification under misspecification. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 331–347. Springer, 2004.
- [2444] P. Grünwald and J. Langford. Suboptimal behavior of Bayes and MDL in classification under misspecification. *Machine Learning*, 66(2-3):119–149, 2007.
- [2445] P. Grünwald and S. de Rooij. Asymptotic log-loss of prequential maximum likelihood codes. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 652–667. Springer, 2005.
- [2446] P. D. Grünwald and W. Kotłowski. Bounds on individual risk for log-loss predictors. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 813–816, Budapest, Hungary, 09–11 June 2011. PMLR.
- [2447] B. Gu, B. Liu, F. Hu, and H. Liu. Efficiently determining the starting sample size for progressive sampling. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 192–202. Springer, 2001.
- [2448] S. Guha, P. Indyk, and A. McGregor. Sketching information divergences. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 424–438, Berlin, 2007. Springer.
- [2449] S. Guha and K. Munagala. Stochastic regret minimization via thompson sampling. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 317–338, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2450] D. Guijarro, V. Lavin, and V. Raghavan. Learning monotone term decision lists. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 16–26. Springer, 1997.
- [2451] D. Guijarro, V. Lavin, and V. Raghavan. Exact learning when irrelevant variables abound. *Inform. Process. Lett.*, 70(5):233–239, 1999.

- [2452] D. Guijarro, V. Lavin, and V. Raghavan. Monotone term decision lists. *Theoret. Comput. Sci.*, 259(1-2):549–575, 2001.
- [2453] D. Guijarro, J. Tarui, and T. Tsukiji. Finding relevant variables in PAC model with membership queries. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 313–322. Springer, 1999.
- [2454] A. Guillory and J. Bilmes. Average-case active learning with costs. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 141–155, Berlin/Heidelberg, 2009. Springer.
- [2455] G. Guimarães. The induction of temporal grammatical rules from multivariate time series. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 127–140. Springer, 2000.
- [2456] S. R. Gunn and J. S. Kandola. Structural modelling with sparse kernels. *Machine Learning*, 48(1/3):137–163, 2002.
- [2457] A. Guntuboyina and B. Sen. L1 covering numbers for uniformly bounded convex functions. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 12.1–12.13, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2458] Y. Guo, P. L. Bartlett, J. Shawe-Taylor, and R. C. Williamson. Covering numbers for support vector machines. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 267–277, New York, NY, 1999. ACM Press.
- [2459] A. Gupta and D. Hsu. Parameter identification in markov chain choice models. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 330–340, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2460] S. Gupta and R. K. P. Zia. Quantum neural networks. *J. Comput. Syst. Sci.*, 63(3):355–383, 2001.
- [2461] V. Guruswami and P. Raghavendra. Hardness of learning halfspaces with noise. *SIAM J. Comput.*, 39(2):742–765, 2009.
- [2462] V. Guruswami and A. Sahai. Multiclass learning, boosting, and error-correcting codes. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 145–155, New York, NY, 1999. ACM Press.
- [2463] L. Gurvits. Linear algebraic proofs of VC-dimension based inequalities. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 238–250. Springer, 1997.
- [2464] L. Gurvits. A note on a scale-sensitive dimension of linear bounded functionals in Banach spaces. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 352–363. Springer, 1997.

- [2465] L. Gurvits. A note on a scale-sensitive dimension of linear bounded functionals in Banach spaces. *Theoret. Comput. Sci.*, 261(1):81–90, 2001. Special issue for ALT '97.
- [2466] L. Gurvits and P. Koiran. Approximation and learning of convex superpositions. *J. Comput. Syst. Sci.*, 55(1):161–170, 1997. Special issue for EuroCOLT '95.
- [2467] O. Guttman, S. V. N. Vishwanathan, and R. C. Williamson. Learnability of probabilistic automata via oracles. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 171–182. Springer, Oct. 2005.
- [2468] H. A. Güvenir and I. Sirin. Classification by feature partitioning. *Machine Learning*, 23(1):47–67, 1996.
- [2469] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
- [2470] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3):389–422, 2002.
- [2471] C. Guzmán. Open problem: The oracle complexity of smooth convex optimization in non-standard settings. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1761–1763, Paris, France, 03–06 July 2015. PMLR.
- [2472] M. Gybels, F. Denis, and A. Habrard. Some improvements of the spectral learning approach for probabilistic grammatical inference. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 64–78, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [2473] L. Györfi and P. Kevei. St. Petersburg portfolio games. In R. Gavalda, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 83–96, Berlin/Heidelberg, 2009. Springer.
- [2474] L. Györfi, G. Turán, and T. Zeugmann. Guest editors' foreword. *Theoret. Comput. Sci.*, 411(29-30):2629–2631, 2010. Special issue for ALT 2008.
- [2475] L. Györfi and I. Vajda. Growth optimal investment with transaction costs. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 108–122, Berlin, Oct. 2008. Springer.
- [2476] G. Györfi and N. Tishby. Statistical theory of learning a rule. In W. K. Theumann and R. Köberle, editors, *Proc. of the STATPHYS-17 Workshop on Neural Networks and Spin Glasses*, Singapore, Teaneck NJ, London, Hong Kong, 1990. World Scientific Publishing Co.
- [2477] A. Györfi, T. Linder, and G. Lugosi. Tracking the best of many experts. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 204–216. Springer, 2005.
- [2478] A. Györfi, T. Linder, and G. Ottucsák. The shortest path problem under partial monitoring. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 468–482, Berlin, 2006. Springer.

- [2479] A. György, G. Lugosi, and G. Ottucsák. On-line sequential bin packing. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 447–454. Omnipress, 2008.
- [2480] G. Györgyi. Inference of a rule by a neural network with thermal noise. *Phys. Rev. Lett.*, 64:2957, 1990.
- [2481] A. Habrard, M. Bernard, and F. Jacquenet. Generalized stochastic tree automata for multi-relational data mining. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 120–133. Springer, 2002.
- [2482] A. Habrard, F. Denis, and Y. Esposito. Using pseudo-stochastic rational languages in probabilistic grammatical inference. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 112–124, Berlin, 2006. Springer.
- [2483] A. Habrard and J. Oncina. Learning multiplicity tree automata. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 268–280, Berlin, 2006. Springer.
- [2484] P. Haddawy, V. Ha, A. Restificar, B. Geisler, and J. Miyamoto. Preference elicitation via theory refinement. *Journal of Machine Learning Research*, 4:317–337, 2003.
- [2485] T. Hagerup and C. Rub. A guided tour of Chernov bounds. *Inform. Proc. Lett.*, 33:305–308, 1990.
- [2486] M. Hagiya. Synthesis of rewrite programs by higher-order and semantic unification. *New Generation Computing*, 8(4):403–420, 1991.
- [2487] M. Hagiya. An iterative and bottom-up procedure for proving-by-example. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 336–341. Springer-Verlag, 1993.
- [2488] M. Hagiya. A typed  $\lambda$ -calculus for proving-by-example and bottom-up generalization procedure. *Theoret. Comput. Sci.*, 137(1):3–23, 1995. Special issue for ALT '93.
- [2489] M. Hagiya and K. Takahashi. Discovery and deduction. In *Discovery Science, Third International Conference, DS 2000, Kyoto, Japan, December 2000, Proceedings*, volume 1967 of *Lecture Notes in Artificial Intelligence*, pages 17–37. Springer, 2000.
- [2490] B. Hajek, Y. Wu, and J. Xu. Computational lower bounds for community detection on random graphs. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 899–928, Paris, France, 03–06 July 2015. PMLR.
- [2491] B. Hajek, Y. Wu, and J. Xu. Semidefinite programs for exact recovery of a hidden community. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1051–1095, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [2492] P. Hájek and M. Holeňa. Formal logics of discovery and hypothesis formation by machine. *Theoret. Comput. Sci.*, 292(2):345–357, 2003.
- [2493] O. M. Halck and F. A. Dahl. Asymmetric co-evolution for imperfect-information zero-sum games. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 171–182. Springer, 2000.
- [2494] K. Hall and T. Hofmann. Learning curved multinomial subfamilies for natural language processing and information retrieval. In *Proc. 17th International Conf. on Machine Learning*, pages 351–358. Morgan Kaufmann, San Francisco, CA, 2000.
- [2495] M. A. Hall. Correlation-based feature selection for discrete and numeric class machine learning. In *Proc. 17th International Conf. on Machine Learning*, pages 359–366. Morgan Kaufmann, San Francisco, CA, 2000.
- [2496] R. Hall, E. Wenger, D. Kibler, and P. Langley. The effect of multiple knowledge sources on learning and teaching. Technical Report 85-11, University of California, Irvine, 1985.
- [2497] R. J. Hall. Learning by failing to explain: Using partial explanations to learn in incomplete or intractable domains. *Machine Learning*, 3(1):45–77, 1988.
- [2498] R. P. Hall, B. Falkenhainer, N. Flann, S. Hampson, R. Reinke, J. Shrager, M. H. Sims, and P. Tadepalli. A review of the fourth international workshop on machine learning. *Machine Learning*, 2(2):173–190, 1987.
- [2499] H. Hamburger and K. Wexler. Identifiability of a class of transformational grammars. In *Approaches to Natural Language*. Dordrecht, D. reidel, 1973.
- [2500] H. Hamburger and K. N. Wexler. A mathematical theory of learning transformational grammar. *Journal of Mathematical Psychology*, 12:137–177, 1975.
- [2501] G. Hamerly and C. Elkan. Bayesian approaches to failure prediction for disk drives. In *Proc. 18th International Conf. on Machine Learning*, pages 202–209. Morgan Kaufmann, San Francisco, CA, 2001.
- [2502] C. Hammerschmidt, B. Loos, R. State, and T. Engel. Flexible state-merging for learning (P)DFAs in python. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 154–159, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [2503] J. Hammerton, M. Osborne, S. Armstrong, and W. Daelemans. Introduction to special issue on machine learning approaches to shallow parsing. *Journal of Machine Learning Research*, 2:551–558, 2002.
- [2504] K. Hammond, T. Converse, M. Marks, and C. M. Seifert. Opportunism and learning. *Machine Learning*, 10:279–309, 1993.
- [2505] S. Hampson and D. Kibler. A Boolean complete neural model of adaptive behavior. *Biol. Cybern.*, 49:9–19, 1983.
- [2506] S. Hampson and D. Kibler. Minimum generalization via reflection: A fast linear threshold learner. *Machine Learning*, 37(1):51–73, 1999.



- [2507] S. E. Hampson and D. J. Volper. Linear function neurons: Structure and training. *Biol. Cybern.*, 53:203–217, 1986.
- [2508] Y. Hamuro, H. Kawata, N. Katoh, and K. Yada. A machine learning algorithm for analyzing string patterns helps to discover simple and interpretable business rules from purchase history. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 565–575. Springer, 2001.
- [2509] W. Han, A. Rakhlin, and K. Sridharan. Competing with strategies. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 966–992. JMLR.org, 2013.
- [2510] T. Hancock. Identifying decision trees with equivalence queries. Manuscript, Harvard University, 1989.
- [2511] T. Hancock. Identifying  $\mu$ -decision trees and  $\mu$ -formulas with constrained instance queries. Manuscript, Harvard University, 1989.
- [2512] T. Hancock. Learning  $k\mu$  decision trees on the uniform distribution. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 352–360, New York, NY, 1993. ACM Press.
- [2513] T. Hancock. Guest editor’s introduction. *Machine Learning*, 25(2/3):115–116, 1996.
- [2514] T. Hancock and L. Hellerstein. Learning read-once formulas over fields and extended bases. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 326–336. Morgan Kaufmann, San Mateo, CA, 1991.
- [2515] T. Hancock, T. Jiang, M. Li, and J. Tromp. Lower bound on learning decision lists and trees. *Inform. Comput.*, 126(2):114–122, 1996.
- [2516] T. Hancock and Y. Mansour. Learning monotone  $k\mu$ -DNF formulas on product distributions. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 179–183, San Mateo, CA, 1991. Morgan Kaufmann.
- [2517] T. R. Hancock. Identifying  $\mu$ -formula decision trees with queries. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 23–37, San Mateo, CA, 1990. Morgan Kaufmann.
- [2518] T. R. Hancock. Learning  $2\mu$ -DNF formulas and  $k\mu$  decision trees. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 199–209, San Mateo, CA, 1991. Morgan Kaufmann.
- [2519] T. R. Hancock, M. Golea, and M. Marchand. Learning nonoverlapping perceptron networks from examples and membership queries. *Machine Learning*, 16(3):161–183, 1994.
- [2520] D. J. Hand and R. J. Till. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning*, 45(2):171–186, 2001.
- [2521] S. Hanneke. Teaching dimension and the complexity of active learning. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13–15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 66–81, Berlin, 2007. Springer.

- [2522] S. Hanneke, V. Kanade, and L. Yang. Learning with a drifting target concept. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 149–164, Berlin, Heidelberg, New York, 2015. Springer.
- [2523] S. Hanneke and L. Reyzin. Algorithmic Learning Theory (ALT) 2017: Preface. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 1–2, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2524] S. Hanneke and L. Reyzin, editors. *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*. PMLR, 2017.
- [2525] D. Hansel and H. Sompolinsky. Learning from examples in a single-layer neural network. *Europhys. Lett.*, 11(7):687–692, 1990.
- [2526] R. Hanson. Combinatorial prediction markets. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 3–4. Omnipress, 2008.
- [2527] S. J. Hanson and M. Bauer. Conceptual clustering, categorization, and polymorphy. *Machine Learning*, 3:343–372, 1989.
- [2528] S. J. Hanson and D. J. Burr. What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences*, 13:471–518, 1990.
- [2529] S. J. Hanson, W. Remmele, and R. L. Rivest, editors. *Machine Learning: From Theory to Applications; Cooperative Research at Siemens and MIT*. Springer-Verlag, 1993.
- [2530] S. Har-Peled, D. Roth, and D. Zimak. Constraint classification: A new approach to multiclass classification. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 365–379. Springer, 2002.
- [2531] S. Harada, E. Takimoto, and A. Maruoka. Online allocation with risk information. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 343–355. Springer, Oct. 2005.
- [2532] K. Haraguchi, S.-H. Hong, and H. Nagamochi. Bipartite graph representation of multiple decision table classifiers. In O. Watanabe and T. Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo, Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, pages 46–60, Berlin/Heidelberg, 2009. Springer.
- [2533] M. Haraguchi. A form of analogy as an abductive inference. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 266–276. IOS Press, 1992.
- [2534] M. Haraguchi and S. Arikawa. Reasoning by analogy as a partial identity between models. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 61–87. Springer-Verlag, 1986.

- [2535] M. Haraguchi and Y. Kudoh. Some criteria for selecting the best data abstractions. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 156–167. Springer, 2001.
- [2536] M. Haraguchi and Y. Okubo. An extended branch and bound search algorithm for finding top- $N$  formal concepts of documents. In *New Frontiers in Artificial Intelligence, JSAI 2006 Conference and Workshops, Tokyo, Japan, June 2006, Revised Selected Papers*, volume 4384 of *Lecture Notes in Artificial Intelligence*, pages 276–288, Berlin, 2007. Springer.
- [2537] Z. Harchaoui, A. Juditsky, A. Nemirovski, and D. Ostrovsky. Adaptive recovery of signals by convex optimization. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 929–955, Paris, France, 03–06 July 2015. PMLR.
- [2538] M. Hardt, R. Meka, P. Raghavendra, and B. Weitz. Computational limits for matrix completion. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 703–725, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2539] M. Hardt and A. Moitra. Algorithms and hardness for robust subspace recovery. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 354–375. JMLR.org, 2013.
- [2540] M. Hardt and M. Wootters. Fast matrix completion without the condition number. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 638–678, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2541] S. Hardy. Synthesis of LISP programs from examples. In *Proc. 4th IJCAI*, pages 268–273. IJCAI, 1975.
- [2542] V. S. Harizanov and F. Stephan. On the learnability of vector spaces. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 233–247. Springer, 2002.
- [2543] V. S. Harizanov and F. Stephan. On the learnability of vector spaces. *J. Comput. Syst. Sci.*, 73(1):109–122, 2007.
- [2544] M. Harries. Boosting a strong learner: evidence against the minimum margin. In *Proc. 16th International Conf. on Machine Learning*, pages 171–180. Morgan Kaufmann, San Francisco, CA, 1999.
- [2545] G. W. Hart. Prototype nonintrusive appliance load monitor. Technical Report Progress Report 2, MIT Energy Laboratory, Sept. 1985.
- [2546] G. W. Hart. *Minimum Information Estimation of Structure*. PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, Apr. 1987. Appears as LIDS-TH-1664.
- [2547] M. Haruno, S. Shirai, and Y. Ooyama. Using decision trees to construct a practical parser. *Machine Learning*, 34:131–149, 1999.

- [2548] A. Harutyunyan, M. G. Bellemare, T. Stepleton, and R. Munos.  $Q(\lambda)$  with off-policy corrections. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 305–320. Springer, 2016.
- [2549] N. Harvey, C. Liaw, and A. Mehrabian. Nearly-tight VC-dimension bounds for piecewise linear neural networks. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1064–1068, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2550] N. Harvey and S. Samadi. Near-optimal herding. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1165–1182, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2551] H. H. Hasegawa, T. Washio, and Y. Ishimiya. Inductive thermodynamics from time series data analysis. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 384–394. Springer, 2001.
- [2552] S. Hasegawa, H. Imai, and M. Ishiguro.  $\varepsilon$ -approximations of  $k$ -label spaces. *Theoret. Comput. Sci.*, 137(1):145–157, 1995. Special issue for ALT '93.
- [2553] B. Hassibi and T. Kailath.  $H^\infty$  bounds for the recursive-least-squares algorithm. In *Proc. of the Thirtythird IEEE Conference on Decision and Control*, pages 3927–3929. IEEE, Dec. 1994.
- [2554] B. Hassibi, A. H. Sayed, and T. Kailath.  $H^\infty$  optimality of the LMS algorithm. *IEEE Trans. on Signal Processing*, 44, 1996.
- [2555] B. Hassibi, A. H. Sayed, and T. Kailath. Linear estimation in Krein spaces - part I: Theory. *IEEE Trans. on Automatic Control*, 41(1):18–33, 1996.
- [2556] A. Hassidim and Y. Singer. Submodular optimization under noise. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1069–1122, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2557] K. Hatano. A simpler analysis of the multi-way branching decision tree boosting algorithm. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 77–91. Springer, 2001.
- [2558] K. Hatano. Smooth boosting using an information-based criterion. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 304–318. Springer, Oct. 2006.
- [2559] K. Hatano and O. Watanabe. Learning  $r$ -of- $k$  functions by boosting. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 114–126. Springer, 2004.
- [2560] D. T. Hau and E. W. Coiera. Learning qualitative models of dynamic systems. *Machine Learning*, 26(2-3):177–211, 1997.

- [2561] D. Haussler. Applying valiant’s learning framework to AI concept learning problems. Technical Report UCSC-CRL-87-11, UCSC, 1987.
- [2562] D. Haussler. Bias, version spaces, and Valiant’s learning framework. In *Proc. of the 4th International Workshop on Machine Learning*, pages 324–336, San Mateo, CA, June 1987. Morgan Kaufmann.
- [2563] D. Haussler. New theoretical directions in machine learning. *Machine Learning*, 2(4):281–284, 1988.
- [2564] D. Haussler. Quantifying inductive bias: AI learning algorithms and Valiant’s learning framework. *Artificial Intelligence*, 36:177–221, 1988.
- [2565] D. Haussler. Space efficient learning algorithms. Technical Report UCSC-CRL-88-2, University of Calif. Computer Research Laboratory, Santa Cruz, CA, 1988.
- [2566] D. Haussler. Generalizing the PAC model for neural net and other learning applications. Technical Report UCSC-CRL-89-30, University of California Santa Cruz, Sept. 1989.
- [2567] D. Haussler. Generalizing the PAC model: sample size bounds from metric dimension-based uniform convergence results. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 40–45. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [2568] D. Haussler. Learning conjunctive concepts in structural domains. *Machine Learning*, 4(1):7–40, 1989.
- [2569] D. Haussler. Probably approximately correct learning. In *Proc. of the 8th National Conference on Artificial Intelligence*, pages 1101–1108. Morgan Kaufmann, 1990.
- [2570] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. *Inform. Comput.*, 100(1):78–150, Sept. 1992.
- [2571] D. Haussler, editor. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, July 27-29, 1992, Pittsburgh, Pennsylvania*, New York, N.Y., 1992. ACM Press.
- [2572] D. Haussler. Sphere packing numbers for subsets of the Boolean  $n$ -cube with bounded Vapnik-Chervonenkis dimension. *J. Combinatorial Theory (A)*, 69(2):217–232, 1995. Was University of CA at Santa Cruz TR UCSC-CRL-91-41.
- [2573] D. Haussler. A brief look at some machine learning problems in genomics. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 109–113, New York, NY, 1997. ACM Press.
- [2574] D. Haussler, M. Kearns, N. Littlestone, and M. K. Warmuth. Equivalence of models for polynomial learnability. *Inform. Comput.*, 95(2):129–161, Dec. 1991.
- [2575] D. Haussler, M. Kearns, and R. E. Schapire. Bounds on the sample complexity of Bayesian learning using information theory and the VC dimension. *Machine Learning*, 14(1):83–113, 1994.
- [2576] D. Haussler, M. Kearns, H. S. Seung, and N. Tishby. Rigorous learning curve bounds from statistical mechanics. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 76–87, 1994.

- [2577] D. Haussler, J. Kivinen, and M. K. Warmuth. Tight worst-case loss bounds for predicting with expert advice. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 69–83. Springer, 1995. Early version in U. of CA, Santa Cruz, TR UCSC-CRL-94-36.
- [2578] D. Haussler, J. Kivinen, and M. K. Warmuth. Sequential prediction of individual sequences under general loss functions. *IEEE Trans. on Information Theory*, 44(5):1906–1925, 1998.
- [2579] D. Haussler, N. Littlestone, and M. K. Warmuth. Predicting  $\{0, 1\}$  functions on randomly drawn points. *Inform. Comput.*, 115(2):248–292, 1994. Was in FOCS88, COLT88, and Univ. of California at Santa Cruz TR UCSC-CRL-90-54.
- [2580] D. Haussler and P. Long. A generalization of sauer’s lemma. *J. Comb. Theory (A)*, 71(2):219–240, 1995. Was Univ. of CA. at Santa Cruz TR UCS-CRL-90-15.
- [2581] D. Haussler and M. Opper. General bounds on the mutual information between a parameter and  $n$  conditionally independent observations. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 402–411, New York, NY, 1995. ACM Press.
- [2582] D. Haussler and M. Opper. Mutual information and Bayes methods for learning a distribution. In *Proc. Workshop on the Theory of Neural Networks: The Statistical Mechanics Perspective*. World Scientific, 1995.
- [2583] D. Haussler and L. Pitt, editors. *Proceedings of the First Annual Workshop on Computational Learning Theory*, San Mateo, CA, 1988. Morgan Kaufmann, Inc.
- [2584] D. Haussler and M. K. Warmuth. The probably approximately correct (PAC) and other learning models. In A. Meyrowitz and S. Chipman, editors, *Machine Learning: Induction, Analogy and Discovery*. Kluwer Academic Publishers, Boston, MA, 1993.
- [2585] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Computational Geometry*, 2:127–151, 1987.
- [2586] S. Hautaniemi, O. Yli-Harja, J. Astola, P. Kauraniemi, A. Kallioniemi, M. Wolf, J. Ruiz, S. Mousses, and O.-P. Kallioniemi. Analysis and visualization of gene expression microarray data in human cancer using self-organizing maps. *Machine Learning*, 52(1-2):45–66, 2003.
- [2587] S. Hayashi. Mathematics based on learning. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 7–21. Springer, 2002.
- [2588] S. Hayashi. Mathematics based on incremental learning—excluded middle and inductive inference. *Theoret. Comput. Sci.*, 350(1):125–139, 2006. Special issue for ALT 2002.
- [2589] Y. Hayashi, S. Matsumoto, A. Shinohara, and M. Takeda. Uniform characterizations of polynomial-query learnabilities. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 84–92. Springer, 1998.
- [2590] Y. Hayashi, S. Matsumoto, A. Shinohara, and M. Takeda. Uniform characterizations of polynomial-query learnabilities. *Theoret. Comput. Sci.*, 292(2):377–385, 2003.
- [2591] F. Hayes-Roth and J. McDermott. An interference matching technique for inducing abstractions. *Commun. ACM*, 21:401–410, 1978.

- [2592] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1991. Second edition.
- [2593] S. Haykin. *Neural Networks: a Comprehensive Foundation*. Macmillan, New York, NY, 1994.
- [2594] E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- [2595] E. Hazan and S. M. Kakade. (weak) calibration is computationally hard. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 3.1–3.10, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2596] E. Hazan, A. Kalai, S. Kale, and A. Agarwal. Logarithmic regret algorithms for online convex optimization. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 499–513, Berlin, 2006. Springer.
- [2597] E. Hazan and S. Kale. Extracting certainty from uncertainty: Regret bounded by variation in costs. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 57–68. Omnipress, 2008.
- [2598] E. Hazan and S. Kale. Beyond the regret minimization barrier: an optimal algorithm for stochastic strongly-convex optimization. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 421–436, Budapest, Hungary, 09–11 June 2011. PMLR.
- [2599] E. Hazan and S. Kale. A simple multi-armed bandit algorithm with optimal variation-bounded regret. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 817–820, Budapest, Hungary, 09–11 June 2011. PMLR.
- [2600] E. Hazan, S. Kale, and S. Shalev-Shwartz. Near-optimal algorithms for online matrix prediction. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 38.1–38.13, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2601] E. Hazan, Z. Karnin, and R. Meka. Volumetric spanners: an efficient exploration basis for learning. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 408–422, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2602] E. Hazan, T. Koren, and K. Y. Levy. Logistic regression: Tight bounds for stochastic and online optimization. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 197–209, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2603] E. Hazan, T. Koren, R. Livni, and Y. Mansour. Online learning with low rank experts. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1096–1114, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [2604] E. Hazan and N. Megiddo. Online learning with prior knowledge. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 499–513, Berlin, 2007. Springer.
- [2605] Y.-B. He, Z. Geng, and X. Liang. Learning causal structures based on markov equivalence class. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 92–106. Springer, Oct. 2005.
- [2606] T. Head, S. Kobayashi, and T. Yokomori. Locality, reversibility, and beyond: Learning languages from positive data. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 191–204. Springer, 1998.
- [2607] D. Heckerman, D. M. Chickering, C. Meek, R. Rounthwaite, and C. Kadie. Dependency networks for inference, collaborative filtering, and data visualization. *Journal of Machine Learning Research*, 1:49–75, 2000.
- [2608] D. Heckerman, D. Geiger, and D. M. Chickering. Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, 20(3):197–243, 1995.
- [2609] F. Hedayati and P. L. Bartlett. The optimality of jeffreys prior for online density estimation and the asymptotic normality of maximum likelihood estimators. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 7.1–7.13, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2610] T. Hegedűs. Combinatorial results on the complexity of teaching and learning. In *Mathematical Foundations of Computer Science 1994, 19th International Symposium, MFCS '94, Košice, Slovakia, August 22-26, 1994, Proceedings*, volume 841 of *Lecture Notes in Computer Science*, pages 393–402. Springer-Verlag, 1994.
- [2611] T. Hegedűs. Can complexity theory benefit from learning theory? In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 354–359. Springer-Verlag, 1993.
- [2612] T. Hegedűs. Geometrical concept learning and convex polytopes. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 228–236, New York, NY, 1994. ACM Press.
- [2613] T. Hegedűs. On training simple neural networks and small-weight neurons. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 69–82, Oxford, 1994. Oxford University Press.
- [2614] T. Hegedűs. Generalized teaching dimensions and the query complexity of learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 108–117, New York, NY, 1995. ACM Press.
- [2615] T. Hegedűs and P. Indyk. On learning disjunctions of zero-one threshold functions with queries. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 446–460. Springer, 1997.



- [2616] M. Heger. Consideration of risk in reinforcement learning. In *Proc. 11th International Conference on Machine Learning*, pages 105–111. Morgan Kaufmann, 1994.
- [2617] M. Heger. The loss from imperfect value functions in expectation-based and minimax-based tasks. *Machine Learning*, 22(1/2/3):197–225, 1996.
- [2618] M. Hein. Uniform convergence of adaptive graph-based regularization. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 50–64, Berlin, 2006. Springer.
- [2619] M. Hein, J.-Y. Audibert, and U. von Luxburg. From graphs to manifolds – weak and strong pointwise consistency of graph Laplacians. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 470–485. Springer, 2005.
- [2620] M. Hein, J.-Y. Audibert, and U. von Luxburg. Graph laplacians and their convergence on random neighborhood graphs. *Journal of Machine Learning Research*, 8:1325–1368, 2007.
- [2621] M. Hein and O. Bousquet. Maximal margin classification for metric spaces. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 72–86. Springer, 2003.
- [2622] M. Hein, O. Bousquet, and B. Schölkopf. Maximal margin classification for metric spaces. *J. Comput. Syst. Sci.*, 71(3):333–359, 2005. Special issue Learning Theory 2003.
- [2623] J. Heinz. Learning left-to-right and right-to-left iterative languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 84–97. Springer, 2008.
- [2624] J. Heinz, C. de la Higuera, and T. Oates. Preface. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 1–3, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [2625] J. Heinz, C. de la Higuera, and T. Oates, editors. *International Conference on Grammatical Inference, 5-8 September 2012, University of Maryland, College Park, MD, USA*, volume 57 of *Proceedings of Machine Learning Research*. PMLR, 2014.
- [2626] J. Heinz, A. Kasprzik, and T. Kötzing. Learning in the limit with lattice-structured hypothesis spaces. *Theoret. Comput. Sci.*, 457(1):111–127, 2012.
- [2627] J. Hekanaho. Symbiosis in multimodal concept learning. In *Proc. 12th International Conference on Machine Learning*, pages 278–285. Morgan Kaufmann, 1995.
- [2628] J. Hekanaho. Background knowledge in GA-based concept learning. In *Proc. 13th International Conference on Machine Learning*, pages 234–242. Morgan Kaufmann, 1996.
- [2629] L. Hellerstein. *On Characterizing and Learning Some Classes of Read-once Functions*. PhD thesis, University of California, Berkeley, 1989.
- [2630] L. Hellerstein. Book Review: Machine learning: A theoretical approach. *Machine Learning*, 13:145–149, 1993.

- [2631] L. Hellerstein. Guest editor’s introduction. *Machine Learning*, 17(2/3):111–114, 1994.
- [2632] L. Hellerstein and M. Karpinski. Learning read-once formulas using membership queries. In *Proc. of the Second Annual Workshop on Computational Learning Theory*, pages 146–161. Morgan Kaufmann, 1989.
- [2633] L. Hellerstein and M. Karpinski. Computational complexity of learning read-once formulas over different bases. Technical report, International Computer Science Institute, Berkeley, CA, TR-91-014, 1991.
- [2634] L. Hellerstein, D. Kletenik, L. Sellie, and R. Servedio. Tight bounds on proper equivalence query learning of DNF. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 31.1–31.18, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2635] L. Hellerstein, K. Pillaipakkamnatt, V. Raghavan, and D. Wilkins. How many queries are needed to learn? *J. ACM*, 43(5):840–862, 1996.
- [2636] L. Hellerstein and V. Raghavan. Exact learning of DNF formulas using DNF hypotheses. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing, May 19–21, 2002, Montréal, Québec, Canada*, pages 465–473. ACM Press, 2002.
- [2637] L. Hellerstein and V. Raghavan. Exact learning of DNF formulas using DNF hypotheses. *J. Comput. Syst. Sci.*, 70(4):435–470, 2005. Special issue on COLT 2002.
- [2638] L. Hellerstein and R. A. Servedio. On PAC learning algorithms for rich Boolean function classes. *Theoret. Comput. Sci.*, 384(1):66–76, 2007. Special issue for TAMC 2006.
- [2639] D. Helmbold, S. Kwek, and L. Pitt. Learning when to trust which experts. In *Computational Learning Theory, Third European Conference, EuroCOLT ’97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 134–149. Springer, 1997.
- [2640] D. Helmbold and S. Panizza. Some label efficient learning results. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 218–230, New York, NY, 1997. ACM Press.
- [2641] D. Helmbold, R. Sloan, and M. K. Warmuth. Learning nested differences of intersection-closed concept classes. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 41–56, San Mateo, CA, 1989. Morgan Kaufmann.
- [2642] D. Helmbold, R. Sloan, and M. K. Warmuth. Learning nested differences of intersection closed concept classes. *Machine Learning*, 5(2):165–196, 1990. Special issue on Computational Learning Theory; first appeared in 2nd COLT conference (1989).
- [2643] D. Helmbold, R. Sloan, and M. K. Warmuth. Learning integer lattices. *SIAM J. Comput.*, 21(2):240–266, 1992.
- [2644] D. Helmbold and B. Williamson, editors. *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, Berlin, 2001. Springer.
- [2645] D. P. Helmbold, J. Kivinen, and M. K. Warmuth. Worst-case loss bounds for single neurons. In *Advances in Neural Information Processing Systems*, volume 8, pages 309–315. The MIT Press, 1996.

- [2646] D. P. Helmbold, N. Littlestone, and P. M. Long. Apple tasting and nearly one-sided learning. In *Proc. of the 33rd Symposium on the Foundations of Comp. Sci.*, pages 493–502. IEEE Computer Society Press, Los Alamitos, CA, 1992.
- [2647] D. P. Helmbold, N. Littlestone, and P. M. Long. Apple tasting. *Inform. Comput.*, 161(2):85–139, Sept. 2000.
- [2648] D. P. Helmbold, N. Littlestone, and P. M. Long. On-line learning with linear loss constraints. *Inform. Comput.*, 161(2):140–171, Sept. 2000.
- [2649] D. P. Helmbold and P. M. Long. Tracking drifting concepts using random examples. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 13–23, San Mateo, CA, 1991. Morgan Kaufmann.
- [2650] D. P. Helmbold and P. M. Long. Tracking drifting concepts by minimizing disagreements. *Machine Learning*, 14(1):27–45, 1994.
- [2651] D. P. Helmbold and P. M. Long. New bounds for learning intervals with implications for semi-supervised learning. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 30.1–30.15, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2652] D. P. Helmbold and P. M. Long. Surprising properties of dropout in deep networks. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1123–1146, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2653] D. P. Helmbold, S. Panizza, and M. K. Warmuth. Direct and indirect algorithms or on-line learning of disjunctions. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 138–152. Springer, 1999.
- [2654] D. P. Helmbold, S. Panizza, and M. K. Warmuth. Direct and indirect algorithms for on-line learning of disjunctions. *Theoret. Comput. Sci.*, 284(1):109–142, 2002.
- [2655] D. P. Helmbold and R. E. Schapire. Predicting nearly as well as the best pruning of a decision tree. *Machine Learning*, 27(1):51–68, 1997.
- [2656] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. On-line portfolio selection using multiplicative updates. In *Proc. 13th International Conference on Machine Learning*, pages 243–251. Morgan Kaufmann, 1996.
- [2657] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. A comparison of new and old algorithms for a mixture estimation problem. *Machine Learning*, 27(1):97–119, 1997.
- [2658] D. P. Helmbold and M. K. Warmuth. Some weak learning results. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 399–412, New York, NY, 1992. ACM Press.
- [2659] D. P. Helmbold and M. K. Warmuth. On weak learning. *J. of Comput. Syst. Sci.*, 50(3):551–573, 1995.
- [2660] D. P. Helmbold and M. K. Warmuth. Learning permutations with exponential weights. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 469–483, Berlin, 2007. Springer.

- [2661] R. J. Hendel. Mathematical learning theory: a formalized, axiomatic, abstract approach. *Inform. Control*, 41:67–117, 1979.
- [2662] M. Henderson, J. Shawe-Taylor, and J. Žerovnik. Mixture of vector experts. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 386–398. Springer, Oct. 2005.
- [2663] R. Herbrich and T. Graepel. Introduction to the special issue on learning theory. *Journal of Machine Learning Research*, 4:755–757, 2003.
- [2664] R. Herbrich, T. Graepel, and C. Campbell. Bayes point machines. *Journal of Machine Learning Research*, 1:245–279, 2001.
- [2665] R. Herbrich, T. Graepel, and J. Shawe-Taylor. Sparsity vs. large margins for linear classifiers. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 304–308. Morgan Kaufmann, San Francisco, 2000.
- [2666] R. Herbrich and R. C. Williamson. Algorithmic luckiness. *Journal of Machine Learning Research*, 3:175–212, 2002.
- [2667] J. Herbst. A machine learning approach to workflow management. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 183–194. Springer, 2000.
- [2668] M. Herbster. Learning additive models online with fast evaluating kernels. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 444–460. Springer, 2001.
- [2669] M. Herbster. Relative loss bounds and polynomial-time predictions for the K-LMS-NET algorithm. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 309–323. Springer, 2004.
- [2670] M. Herbster. Exploiting cluster-structure to predict the labeling of a graph. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 54–69, Berlin, Oct. 2008. Springer.
- [2671] M. Herbster and M. Warmuth. Tracking the best expert. *Machine Learning*, 32(2):151–178, 1998. Earlier version in 12th ICML, 1995.
- [2672] M. Herbster and M. K. Warmuth. Tracking the best regressor. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 24–31, New York, NY, 1998. ACM Press.
- [2673] M. Herbster and M. K. Warmuth. Tracking the best linear predictor. *Journal of Machine Learning Research*, 1:281–309, 2001.
- [2674] F. Hergert, W. Finnoff, and H.-G. Zimmermann. Improving model selection by dynamic regularization methods. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 17, pages 323–343. MIT Press, 1995.

- [2675] M. Hermo and V. Lavin. Learning minimal covers of functional dependencies with queries. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 291–300. Springer, 1999.
- [2676] M. Hermo and A. Ozaki. Exact learning of multivalued dependencies. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 73–88, Berlin, Heidelberg, New York, 2015. Springer.
- [2677] M. Hermo and A. Ozaki. Exact learning of multivalued dependency formulas. *Theoret. Comput. Sci.*, 716:4–14, 2018. Special Issue on ALT 2015.
- [2678] T. Heskes. Solving a huge number of similar tasks: a combination of multi-task learning and a hierarchical Bayesian approach. In *Proc. 15th International Conf. on Machine Learning*, pages 233–241. Morgan Kaufmann, San Francisco, CA, 1998.
- [2679] T. Heskes. Empirical Bayes for learning to learn. In *Proc. 17th International Conf. on Machine Learning*, pages 367–374. Morgan Kaufmann, San Francisco, CA, 2000.
- [2680] T. Heskes, B. Bakker, and B. Kappen. Approximate algorithms for neural-Bayesian approaches. *Theoret. Comput. Sci.*, 287(1):219–238, 2002.
- [2681] M. J. H. Heule and S. Verwer. Exact DFA identification using SAT solvers. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13–16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 66–79. Springer, 2010.
- [2682] Y. Higa, S. Inenaga, H. Bannai, and M. Takeda. A new family of string classifiers based on local relatedness. In *Discovery Science, 9th International Conference, DS 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4265 of *Lecture Notes in Artificial Intelligence*, pages 114–124. Springer, Oct. 2006.
- [2683] M. Higgs and J. Shawe-Taylor. A PAC-Bayes bound for tailored density estimation. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6–8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 148–162, Berlin, Heidelberg, New York, 2010. Springer.
- [2684] I. Higuchi and S. Eguchi. Robust principal component analysis with adaptive selection for tuning parameters. *Journal of Machine Learning Research*, 5:453–471, 2004.
- [2685] C. de la Higuera. Characteristic sets for polynomial grammatical inference. *Machine Learning*, 27(2):125–138, 1997.
- [2686] C. de la Higuera. Learning stochastic finite automata from experts. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 79–89. Springer, 1998.
- [2687] C. de la Higuera. A bibliographical study of grammatical inference. *Pattern Recognition*, 38(9):1332–1348, 2005.

- [2688] C. de la Higuera. Ten open problems in grammatical inference. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 32–44, Berlin, 2006. Springer.
- [2689] C. de la Higuera and J.-C. Janodet. Inference of  $\omega$ -languages from prefixes. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 364–377. Springer, 2001.
- [2690] C. de la Higuera and J. C. Janodet. Inference of  $\omega$ -languages from prefixes. *Theoret. Comput. Sci.*, 313(2):295–312, 2004. Special issue for ALT 2001.
- [2691] C. de la Higuera, J.-C. Janodet, and F. Tantini. Learning languages from bounded resources: The case of the DFA and the balls of strings. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 43–56. Springer, 2008.
- [2692] C. de la Higuera and J. Oncina. Inferring deterministic linear languages. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 185–200. Springer, 2002.
- [2693] C. de la Higuera and J. Oncina. On sufficient conditions to identify in the limit classes of grammars from polynomial time and data. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 134–148. Springer, 2002.
- [2694] C. de la Higuera and J. Oncina. Identification with probability one of stochastic deterministic linear languages. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 247–258. Springer, 2003.
- [2695] C. de la Higuera and J. Oncina. Learning stochastic finite automata. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 175–186. Springer, 2004.
- [2696] C. de la Higuera. Characteristic sets for polynomial grammatical inference. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 59–71. Springer, 1996.
- [2697] C. de la Higuera, J. Oncina, and E. Vidal. Identification of DFA: data-dependent vs data-independent algorithms. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 313–325. Springer, 1996.
- [2698] C. de la Higuera and F. Thollard. Identification in the limit with probability one of stochastic deterministic finite automata. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 141–156. Springer, 2000.

- [2699] G. Hinton and D. van Camp. Keeping neural networks simple by minimizing the description length of the weights. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 5–13, New York, NY, 1993. ACM Press.
- [2700] G. E. Hinton. Learning distributed representations of concepts. In *Proc. 8th Annual Conf. of the Cognitive Science Society*, pages 1–12, Hillsdale, 1986. Erlbaum.
- [2701] G. E. Hinton. Deterministic Boltzmann learning performs steepest descent in weight-space. *Neural Computation*, 1(1):143–150, 1989.
- [2702] G. E. Hinton, T. J. Sejnowski, and D. H. Ackley. Boltzmann machines: Constraint satisfaction networks that learn. Technical Report CMS-CS-84-119, CMU Computer Science Department, May 1984.
- [2703] H. Hiraishi and F. Mizoguchi. WWW visualization tools for discovering interesting web pages. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 650–660. Springer, 2001.
- [2704] M. Hirao, H. Hoshino, A. Shinohara, M. Takeda, and S. Arikawa. A practical algorithm to find the best subsequence patterns. *Theoret. Comput. Sci.*, 292(2):465–479, 2003.
- [2705] K. Hiraoka and S. Amari. Strategy under the unknown stochastic environment: The non-parametric lob-pass problem. *Algorithmica*, 22(1/2):138–156, 1998.
- [2706] K. Hirata. Rule-generating abduction for recursive Prolog. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 121–136. Springer-Verlag, 1994.
- [2707] K. Hirata. Flattening and implication. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 157–168. Springer, 1999.
- [2708] K. Hirata. On the hardness of learning acyclic conjunctive queries. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 238–251. Springer, 2000.
- [2709] K. Hirata. Prediction-hardness of acyclic conjunctive queries. *Theoret. Comput. Sci.*, 348(1):84–94, 2005. Special issue for ALT 2000.
- [2710] S. Hirokawa and D. Ikeda. Visualization and analysis of web graphs. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 616–627. Springer, 2001.
- [2711] E. Hirowatari and S. Arikawa. Inferability of recursive real-valued functions. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 18–31. Springer, 1997.

- [2712] E. Hirowatari and S. Arikawa. A comparison of identification criteria for inductive inference of recursive real-valued functions. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 262–275. Springer, 1998.
- [2713] E. Hirowatari and S. Arikawa. A comparison of identification criteria for inductive inference of recursive real-valued functions. *Theoret. Comput. Sci.*, 268(2):351–366, 2001. Special issue for ALT '98.
- [2714] D. S. Hirschberg, M. J. Pazzani, and K. M. Ali. Average case analysis of k-CNF and k-DNF learning algorithms. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 2, pages 15–28. MIT Press, 1994.
- [2715] H. Hirsh. Polynomial time learning with version spaces. In *Proc. of AAAI-92*, 1992.
- [2716] H. Hirsh. Generalizing version spaces. *Machine Learning*, 17(1):5–46, 1994.
- [2717] H. Hirsh and W. W. Cohen. Learning from data with bounded inconsistency: theoretical and experimental results. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 12, pages 355–380. MIT Press, 1994.
- [2718] J. M. Hitchcock. Online learning and resource-bounded dimension: Winnow yields new lower bounds for hard sets. In *STACS 2006, 23rd Annual Symposium on Theoretical Aspects of Computer Science, Marseille, France, February 23-25, 2006, Proceedings*, volume 3884 of *Lecture Notes in Computer Science*, pages 408–419. Springer, 2006.
- [2719] F. Ho and M. Kamel. Learning coordination strategies for cooperative multiagent systems. *Machine Learning*, 33(2/3):155–177, 1998.
- [2720] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, Mar. 1963.
- [2721] K. Höffgen. Learning and robust learning of product distributions. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 77–83, New York, NY, 1993. ACM Press.
- [2722] K.-U. Höffgen and H.-U. Simon. Robust trainability of single neurons. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 428–439, New York, New York, 1992. ACM Press.
- [2723] K.-U. Höffgen, H.-U. Simon, and K. S. Van Horn. Robust trainability of single neurons. *J. of Comput. Syst. Sci.*, 50(1):114–125, 1995.
- [2724] A. Hoffmann, R. Kwok, and P. Compton. Using subclasses to improve classification learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 203–213. Springer, 2001.
- [2725] B. Hoffmeister and T. Zeugmann. Text mining using markov chains of variable length. In *Federation over the Web: International Workshop, Dagstuhl Castle, Germany, May 1-6, 2005. Revised Selected Papers*, volume 3847 of *Lecture Notes in Artificial Intelligence*, pages 1–24. Springer, Berlin, 2006.



- [2726] T. Hofmann. Learning what people (don't) want. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 214–225. Springer, 2001.
- [2727] T. Hofmann. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1/2):177–196, 2001.
- [2728] T. Hofmeister. An application of codes to attribute-efficient learning. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29–31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 101–110. Springer, 1999.
- [2729] S. B. Holden. PAC-like upper bounds for the sample complexity of leave-one-out cross-validation. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 41–50, New York, NY, 1996. ACM Press.
- [2730] J. H. Holland. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, 1975.
- [2731] C. C. Holmes and D. G. T. Denison. Classification with Bayesian MARS. *Machine Learning*, 50(1-2):159–173, 2003.
- [2732] J. H. Holmes, P. L. Lanzi, W. Stolzmann, and S. W. Wilson. Learning classifier systems: New models, successful applications. *Information Processing Letters*, 82(1):23–30, 2002.
- [2733] R. C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11:63–91, 1993.
- [2734] R. Hözl, S. Jain, P. Schlicht, K. Seidel, and F. Stephan. Automatic learning from repetitive texts. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 129–150, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2735] R. Hözl, S. Jain, and F. Stephan. Learning pattern languages over groups. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 189–203. Springer, 2016.
- [2736] V. Honavar. Neural network design and the complexity of learning. *Machine Learning*, 9:95–98, 1992.
- [2737] V. Honavar and C. de la Higuera. Introduction. *Machine Learning*, 44(1/2):5–7, 2001.
- [2738] R. Honda, Y. Iijima, and O. Konishi. Mining of topographic feature from heterogeneous imagery and its application to lunar craters. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 395–407. Springer, 2001.
- [2739] Y.-T. Hong and C.-J. Lu. Online learning in Markov decision processes with continuous actions. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 302–316, Berlin, Heidelberg, New York, 2015. Springer.

- [2740] S. Hoory and O. Margalit. Finding the rare cube. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 344–358, Berlin, Oct. 2008. Springer.
- [2741] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79:2554–2558, Apr. 1982.
- [2742] S. B. Hopkins, J. Shi, and D. Steurer. Tensor principal component analysis via sum-of-square proofs. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 956–1006, Paris, France, 03–06 July 2015. PMLR.
- [2743] K. Hornik, M. Stinchcombe, and H. White. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3:551–560, 1990.
- [2744] J. J. Horning. *A Study of Grammatical Inference*. PhD thesis, Stanford University, 1969.
- [2745] M. Horváth. On error estimation for the partitioning classification rule. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 253–262. Springer, 1999.
- [2746] T. Horváth, R. H. Sloan, and G. Turán. Learning logic programs by using the product homomorphism method. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 10–20, New York, NY, 1997. ACM Press.
- [2747] T. Horváth, S. Wrobel, and U. Bohnebeck. Relational instance-based learning with lists and terms. *Machine Learning*, 43(1/2):53–80, 2001.
- [2748] V. Hoste, W. Daelemans, E. T. K. Sang, and S. Gillis. Meta-learning for phonemic annotation of corpora. In *Proc. 17th International Conf. on Machine Learning*, pages 375–382. Morgan Kaufmann, San Francisco, CA, 2000.
- [2749] D. F. Hougen, M. Gini, and J. Slagle. An integrated connectionist approach to reinforcement learning for robotic control: The advantages of indexed partitioning. In *Proc. 17th International Conf. on Machine Learning*, pages 383–390. Morgan Kaufmann, San Francisco, CA, 2000.
- [2750] N. R. Howe. Data as ensembles of records: Representation and comparison. In *Proc. 17th International Conf. on Machine Learning*, pages 391–398. Morgan Kaufmann, San Francisco, CA, 2000.
- [2751] D. C. Hoyle and M. Rattray. A statistical mechanics analysis of gram matrix eigenvalue spectra. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 579–593. Springer, 2004.
- [2752] Y. ping Hsieh and V. Cevher. Dimension-free information concentration via exp-concavity. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 451–469. PMLR, 07–09 Apr. 2018.
- [2753] T. Hsing, S. Attoor, and E. Dougherty. Relation between permutation-test P values and classifier error estimates. *Machine Learning*, 52(1-2):11–30, 2003.

- [2754] C.-N. Hsu, H.-J. Huang, and T.-T. Wong. Why discretization works for naïve Bayesian classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 399–406. Morgan Kaufmann, San Francisco, CA, 2000.
- [2755] C.-N. Hsu, H.-J. Huang, and T.-T. Wong. Implications of the Dirichlet assumption for discretization of continuous variables in naive Bayesian classifiers. *Machine Learning*, 53(3):235–263, 2003.
- [2756] C.-N. Hsu and C. A. Knoblock. Rule induction for semantic query optimization. In *Proc. 11th International Conference on Machine Learning*, pages 112–120. Morgan Kaufmann, 1994.
- [2757] C.-W. Hsu and C.-J. Lin. A simple decomposition method for support vector machines. *Machine Learning*, 46(1-3):291–314, 2002.
- [2758] D. Hsu, S. M. Kakade, and T. Zhang. Random design analysis of ridge regression. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 9.1–9.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2759] D. Hsu, S. M. Kakade, and T. Zhang. A spectral algorithm for learning Hidden Markov Models. *J. Comput. Syst. Sci.*, 78(5):1460–1480, 2012.
- [2760] W. H. Hsu, S. R. Ray, and D. C. Wilkins. A multistrategy approach to classifier learning from time series. *Machine Learning*, 38(1/2):213–236, 2000.
- [2761] B. Hu, P. Seiler, and A. Rantzer. A unified analysis of stochastic optimization methods using jump system theory and quadratic constraints. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1157–1189, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2762] J. Hu and M. P. Wellman. Multiagent reinforcement learning: theoretical framework and an algorithm. In *Proc. 15th International Conf. on Machine Learning*, pages 242–250. Morgan Kaufmann, San Francisco, CA, 1998.
- [2763] J. Hu and M. P. Wellman. Experimental results on Q-learning for general-sum stochastic games. In *Proc. 17th International Conf. on Machine Learning*, pages 407–414. Morgan Kaufmann, San Francisco, CA, 2000.
- [2764] J. Hu and M. P. Wellman. Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, 4:1039–1069, 2003.
- [2765] L. Hu, R. Wu, T. Li, and L. Wang. Quadratic upper bound for recursive teaching dimension of finite VC classes. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1147–1156, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [2766] N. Hu and R. B. Dannenberg. Bootstrap learning for accurate onset detection. *Machine Learning*, 65(2-3):457–471, 2006.
- [2767] Y.-J. Hu, S. Sandmeyer, and D. Kibler. Detecting motifs from sequences. In *Proc. 16th International Conf. on Machine Learning*, pages 181–190. Morgan Kaufmann, San Francisco, CA, 1999.
- [2768] C. Hua and K. Ko. A note on the pattern-finding problem. Technical Report UH-CS-84-4, Department of Computer Science, University of Houston, 1984.

- [2769] R. Huang, M. M. Ajallooeian, C. Szepesvári, and M. Müller. Structured best arm identification with fixed confidence. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 593–616, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2770] Y.-C. Huang, B. Selman, and H. Kautz. Learning declarative control rules for constraint-based planning. In *Proc. 17th International Conf. on Machine Learning*, pages 415–422. Morgan Kaufmann, San Francisco, CA, 2000.
- [2771] P. J. Huber. Projection pursuit. *Ann. of Statist.*, 13(2):435–475, 1985.
- [2772] M. Hulden. Treba: Efficient numerically stable EM for PFA. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 249–253, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [2773] E. Hüllermeier and J. Fürnkranz. Learning from label preferences. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 38–38, Berlin, Heidelberg, New York, 2011. Springer.
- [2774] D. Hush and C. Scovel. On the VC dimension of bounded margin classifiers. *Machine Learning*, 45(1):33–44, 2001.
- [2775] D. Hush and C. Scovel. Polynomial-time decomposition algorithms for support vector machines. *Machine Learning*, 51(1):51–71, 2003.
- [2776] J.-C. Hütter and P. Rigollet. Optimal rates for total variation denoising. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1115–1146, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2777] M. Hutter. Convergence and error bounds for universal prediction of nonbinary sequences. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 239–250. Springer, 2001.
- [2778] M. Hutter. General loss bounds for universal sequence prediction. In *Proc. 18th International Conf. on Machine Learning*, pages 210–217. Morgan Kaufmann, San Francisco, CA, 2001.
- [2779] M. Hutter. New error bounds for Solomonoff prediction. *J. Comput. Syst. Sci.*, 62(4):653–667, 2001.
- [2780] M. Hutter. Towards a universal theory of artificial intelligence based on algorithmic probability and sequential decisions. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 226–238. Springer, 2001.
- [2781] M. Hutter. Self-optimizing and pareto-optimal policies in general environments based on Bayes-mixtures. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 364–379. Springer, 2002.

- [2782] M. Hutter. On the existence and convergence of computable universal priors. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 298–312. Springer, 2003.
- [2783] M. Hutter. An open problem regarding the convergence of universal A priori probability. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 738–740. Springer, 2003.
- [2784] M. Hutter. Optimality of universal Bayesian sequence prediction for general loss and alphabet. *Journal of Machine Learning Research*, 4:971–1000, 2003.
- [2785] M. Hutter. Sequence prediction based on monotone complexity. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 506–521. Springer, 2003.
- [2786] M. Hutter. General discounting versus average reward. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 244–258. Springer, Oct. 2006.
- [2787] M. Hutter. On generalized computable universal priors and their convergence. *Theoret. Comput. Sci.*, 364(1):27–41, 2006. Special issue for ALT 2003.
- [2788] M. Hutter. On the foundations of universal sequence prediction. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 408–420, Berlin, 2006. Springer.
- [2789] M. Hutter. Sequential predictions based on algorithmic complexity. *J. Comput. Syst. Sci.*, 72(1):95–117, 2006.
- [2790] M. Hutter. The loss rank principle for model selection. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 589–603, Berlin, 2007. Springer.
- [2791] M. Hutter. On universal prediction and Bayesian confirmation. *Theoret. Comput. Sci.*, 384(1):33–48, 2007. Special issue for TAMC 2006.
- [2792] M. Hutter. Universal learning theory. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 1001–1008. Springer, New York, 2010.
- [2793] M. Hutter. Sparse adaptive Dirichlet-multinomial-like processes. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 432–459. JMLR.org, 2013.
- [2794] M. Hutter. Extreme state aggregation beyond MDPs. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 185–199, Berlin, Heidelberg, New York, 2014. Springer.

- [2795] M. Hutter. Offline to online conversion. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 230–244, Berlin, Heidelberg, New York, 2014. Springer.
- [2796] M. Hutter. Extreme state aggregation beyond Markov decision processes. *Theoret. Comput. Sci.*, 650:73–91, 2016. Special Issue: Algorithmic Learning Theory.
- [2797] M. Hutter. Tractability of batch to sequential conversion. *Theoret. Comput. Sci.*, 733:71–82, 2018. Special Issue on Learning Theory and Complexity.
- [2798] M. Hutter and A. Muchnik. Universal convergence of semimeasures on individual random sequences. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 234–248. Springer, 2004.
- [2799] M. Hutter and A. Muchnik. On semimeasures predicting Martin-Löf random sequences. *Theoret. Comput. Sci.*, 382(3):247–261, 2007. Special issue for ALT 2004.
- [2800] M. Hutter and J. Poland. Prediction with expert advice by following the perturbed leader for general weights. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 279–293. Springer, 2004.
- [2801] M. Hutter and J. Poland. Adaptive online prediction by following the perturbed leader. *Journal of Machine Learning Research*, 6:639–660, 2005.
- [2802] M. Hutter and R. A. Servedio. Preface. *Theoret. Comput. Sci.*, 410(9):1747–1748, 2009. Special issue for ALT 2007.
- [2803] M. Hutter, R. A. Servedio, and E. Takimoto, editors. *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2007. Springer.
- [2804] M. Hutter, R. A. Servedio, and E. Takimoto. Editors’ introduction. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 1–8, Berlin, Oct. 2007. Springer.
- [2805] M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors. *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6–8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, Berlin, Heidelberg, New York, Oct. 2010. Springer.
- [2806] M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann. Editors’ introduction. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6–8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 1–10, Berlin, Heidelberg, New York, 2010. Springer.
- [2807] M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann. Guest Editors’ foreword. *Theoret. Comput. Sci.*, 473:1–3, 2013. Special Issue on Learning Theory.
- [2808] A. Hyvärinen. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6:695–709, 2005.

- [2809] G. A. Iba. A heuristic approach to the discovery of macro-operators. *Machine Learning*, 3:285–317, 1989.
- [2810] O. H. Ibarra and T. Jiang. Learning regular languages from counterexamples. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 371–385, San Mateo, CA, 1988. Morgan Kaufmann.
- [2811] O. H. Ibarra and T. Jiang. Learning regular languages from counterexamples. *J. of Comput. Syst. Sci.*, 43(2):299–316, 1991.
- [2812] T. Ida, M. Marin, and T. Suzuki. Reducing search space in solving higher-order equations. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 19–30. Springer, 2001.
- [2813] P. Idestam-Almquist. Generalization under implication by using Or-introduction. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 56–64. Springer-Verlag, 1993.
- [2814] P. Idestam-Almquist. Generalization of clauses relative to a theory. *Machine Learning*, 26(2-3):213–226, 1997.
- [2815] P. M. Idziak, P. Markovic, R. McKenzie, M. Valeriote, and R. Willard. Tractability and learnability arising from algebras with few subpowers. *SIAM J. Comput.*, 39(7):3023–3037, 2010.
- [2816] A. Iglesias, P. Martínez, R. Aler, and F. Fernández. Learning content sequencing in an educational environment according to student needs. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 454–463. Springer, 2004.
- [2817] A. T. Ihler, J. W. F. III, and A. S. Willsky. Loopy belief propagation: Convergence and effects of message errors. *Journal of Machine Learning Research*, 6:905–936, 2005.
- [2818] D. Iijima, W. Yu, H. Yokoi, and Y. Kakazu. Distributed robotic learning: adaptive behavior acquisition for distributed autonomous swimming robot in real-world. In *Proc. 16th International Conf. on Machine Learning*, pages 191–199. Morgan Kaufmann, San Francisco, CA, 1999.
- [2819] Y. Ikeda and D. Mazurkiewicz. Application of neural network technique to combustion spray dynamics analysis. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 408–425. Springer, 2001.
- [2820] A. Imiya and K. Iwawaki. Statistics of flow vectors and its application to the voting method for the detection of flow fields. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 293–306. Springer, 2001.
- [2821] A. Imiya and H. Ootani. PCA-based model selection and fitting for linear manifolds. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 278–292. Springer, 2001.

- [2822] B. Indurkha. Modes of analogy. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 217–230. Springer-Verlag, 1989.
- [2823] B. Indurkha. On the role of interpretive analogy in learning. *New Generation Computing*, 8(4):385–402, 1991.
- [2824] B. Indurkha. Predictive analogy and cognition. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 214–231. Springer-Verlag, 1992.
- [2825] N. Indurkha and S. M. Weiss. Rule-based ensemble solutions for regression. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 62–72. Springer, 2001.
- [2826] P. Indyk. A sublinear time approximation scheme for clustering in metric spaces. In *Proc. 40th Annu. IEEE Symposium on Foundations of Computer Science*, pages 154–159. IEEE Computer Society Press, Los Alamitos, CA, 1999.
- [2827] A. Inokuchi, T. Washio, and H. Motoda. Complete mining of frequent patterns from graphs: Mining graph data. *Machine Learning*, 50(3):321–354, 2003.
- [2828] N. Intrator. Feature extraction using an unsupervised neural network. In *Proc. of the 1990 Connectionist Models Summer School*, pages 310–318. Morgan Kaufmann., 1990.
- [2829] T. R. Ioerger, L. A. Rendell, and S. Subramaniam. Searching for representations to improve protein sequence fold-class prediction. *Machine Learning*, 21:151–175, 1995.
- [2830] M. N. Irfan, R. Groz, and C. Oriat. Improving model inference of black box components having large input test set. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 133–138, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [2831] M. Isberner and B. Steffen. An abstract framework for counterexample analysis in active automata learning. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 79–93, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [2832] K. Ishibashi, K. Hatano, and M. Takeda. Online learning of maximum  $p$ -norm margin classifiers with bias. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 69–80. Omnipress, 2008.
- [2833] Y. Ishigami and S. Tani. The VC-dimensions of finite automata with  $n$  states. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 328–341. Springer-Verlag, 1993.
- [2834] A. Ishino and A. Yamamoto. Learning from examples with typed equational programming. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 301–316. Springer-Verlag, 1994.



- [2835] A. Ishino and A. Yamamoto. Generalizations in typed equational programming and their application to learning functions. *New Generation Computing*, 15(1):85–104, 1997. Special issue for AII '94 and ALT '94.
- [2836] H. Ishizaka. Learning simple deterministic languages. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 162–174, San Mateo, CA, 1989. Morgan Kaufmann.
- [2837] H. Ishizaka. Polynomial time learnability of simple deterministic languages. *Machine Learning*, 5(2):151–164, 1990. Special issue on Computational Learning Theory; first appeared in 2nd COLT conference (1989).
- [2838] H. Ishizaka, H. Arimura, and T. Shinohara. Efficient inductive inference of primitive prologs from positive data. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 135–146. Springer-Verlag, 1993.
- [2839] H. Ishizaka, H. Arimura, and T. Shinohara. Finding tree patterns consistent with positive and negative examples using queries. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):101–115, 1998. Special issue for AII '94 and ALT '94.
- [2840] H. Ito, S. Amari, and K. Kobayashi. Identifiability of hidden Markov information sources and their minimum degrees of freedom. *IEEE Trans. Information Theory*, 38(2):324–333, Mar. 1992.
- [2841] K. Ito and A. Yamamoto. Polynomial-time MAT learning of multilinear logic programs. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 63–74. Springer-Verlag, 1993.
- [2842] K. Ito, T. Zeugmann, and Y. Zhu. Clustering the normalized compression distance for influenza virus data. In T. Elomaa, H. Mannila, and P. Orponen, editors, *Algorithms and Applications, Essays Dedicated to Esko Ukkonen on the Occasion of His 60th Birthday*, volume 6060 of *Lecture Notes in Computer Science*, pages 130–146. Springer, Heidelberg, 2010.
- [2843] K. Ito, T. Zeugmann, and Y. Zhu. Recent experiences in parameter-free data mining. In E. Gelenbe, R. Lent, G. Sakellari, A. Sacan, H. Toroslu, and A. Yazici, editors, *Computer and Information Science, Proceedings of the 25th International Symposium on Computer and Information Sciences*, volume 62 of *Lecture Notes in Electrical Engineering*, pages 365–371, Heidelberg, 2010. Springer.
- [2844] S. Y. Itoga. A new heuristic for inferring regular grammars. *IEEE Trans. Pattern Anal. Mach. Intell.*, TPAMI-3:191–197, 1981.
- [2845] Y. Itoh and M. Haraguchi. Conceptual classifications guided by a concept hierarchy. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 166–178. Springer, 2000.
- [2846] A. Ittner and M. Schlosser. Non-linear decision trees – NDT. In *Proc. 13th International Conference on Machine Learning*, pages 252–258. Morgan Kaufmann, 1996.
- [2847] Y. Ivanov, B. Blumberg, and A. Pentland. Expectation maximization for weakly labeled data. In *Proc. 18th International Conf. on Machine Learning*, pages 218–225. Morgan Kaufmann, San Francisco, CA, 2001.

- [2848] T. Iwatani, S. Tano, A. Inoue, and W. Okamoto. Fuzzy analogy based reasoning and classification of fuzzy analogies. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 137–148. Springer-Verlag, 1994.
- [2849] M. Iwayama, N. Indurkha, and H. Motoda. A new algorithm for automatic configuration of hidden Markov models. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 237–250. Springer-Verlag, 1993.
- [2850] A. Izworski, R. Tadeusiewicz, and A. Paslawski. The utilization of context signals in the analysis of ABR potentials by application of neural networks. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 195–202. Springer, 2000.
- [2851] E. Jääsaari, J. Leppä-aho, T. Silander, and T. Roos. Minimax optimal bayes mixtures for memoryless sources over large alphabets. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 470–488. PMLR, 07–09 Apr. 2018.
- [2852] J. Jackson. On the efficiency of noise-tolerant PAC algorithms derived from statistical queries. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 7–15. Morgan Kaufmann, San Francisco, 2000.
- [2853] J. Jackson and A. Tomkins. A computational model of teaching. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 319–326, New York, NY, 1992. ACM Press.
- [2854] J. C. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. *J. Comput. Syst. Sci.*, 55(3):414–440, 1997. Earlier version appeared in FOCS '94.
- [2855] J. C. Jackson. Uniform-distribution learnability of noisy linear threshold functions with restricted focus of attention. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 304–318, Berlin, 2006. Springer.
- [2856] J. C. Jackson, A. Klivans, and R. A. Servedio. Learnability beyond  $AC^0$ . In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 776–784. ACM Press, 2002.
- [2857] J. C. Jackson and R. A. Servedio. Learning random log-depth decision trees under the uniform distribution. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 610–624. Springer, 2003.
- [2858] J. C. Jackson and R. A. Servedio. Learning random log-depth decision trees under uniform distribution. *SIAM J. Comput.*, 34(5):1107–1128, 2005.
- [2859] R. A. Jacobs, M. A. Jordan, and A. G. Barto. Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. Technical Report 90-27, COINS, Mar. 1990.

- [2860] S. A. Jaeger. Generalization bounds and complexities based on sparsity and clustering for convex combinations of functions from random classes. *Journal of Machine Learning Research*, 6:307–340, 2005.
- [2861] A. Jafari, A. Greenwald, D. Gondek, and G. Ercal. On no-regret learning, fictitious play, and nash equilibrium. In *Proc. 18th International Conf. on Machine Learning*, pages 226–233. Morgan Kaufmann, San Francisco, CA, 2001.
- [2862] M. Jagadish and A. Sen. Learning a bounded-degree tree using separator queries. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 188–202, Berlin, Heidelberg, New York, 2013. Springer.
- [2863] P. Jain, C. Jin, S. M. Kakade, P. Netrapalli, and A. Sidford. Streaming PCA: Matching matrix bernstein and near-optimal finite sample guarantees for oja’s algorithm. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1147–1164, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [2864] P. Jain, P. Kothari, and A. Thakurta. Differentially private online learning. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 24.1–24.34, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [2865] P. Jain and P. Netrapalli. Fast exact matrix completion with finite samples. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1007–1034, Paris, France, 03–06 July 2015. PMLR.
- [2866] P. Jain and S. Oh. Learning mixtures of discrete product distributions using spectral decompositions. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 824–856, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2867] S. Jain. *Learning in the Presence of Additional Information and Inaccurate Information*. PhD thesis, University of Rochester, 1990.
- [2868] S. Jain. An infinite class of functions identifiable using minimal programs in all Kolmogorov numberings. *International Journal of Foundations of Computer Science*, 6(1):89–94, 1995.
- [2869] S. Jain. On a question about learning nearly minimal programs. *Inform. Proc. Lett.*, 53(1):1–4, 1995.
- [2870] S. Jain. Program synthesis in the presence of infinite number of inaccuracies. *J. Comput. Syst. Sci.*, 53(3):583–591, 1996. First appeared in ALT ’94.
- [2871] S. Jain. Strong monotonic and set-driven inductive inference. *J. of Experimental and Theoret. Artif. Intell.*, 9(1):137–143, 1997.
- [2872] S. Jain. Learning with refutation. *J. Comput. Syst. Sci.*, 57(3):356–365, 1998. First appeared in ALT ’98.
- [2873] S. Jain. Minimal concept identification and reliability. *International Journal of Foundations of Computer Science*, 9(3):315–320, 1998.

- [2874] S. Jain. On a question of nearly minimal identification of functions. *Inform. Process. Lett.*, 70(3):113–117, 1999.
- [2875] S. Jain. Robust behaviorally correct learning. *Inform. Comput.*, 153(2):238–248, 1999.
- [2876] S. Jain. On an open problem in classification of languages. *Journal of Experimental and Theoretical Artificial Intelligence*, 13(2):113–118, Apr. 2001.
- [2877] S. Jain. Hypothesis spaces for learning. In A. H. Dediu, A. M. Ionescu, and C. Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2009, Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 43–58. Springer, 2009. Invited Paper.
- [2878] S. Jain. On some open problems in reflective inductive inference. *Information Processing Letters*, 109(3):208–211, 2009.
- [2879] S. Jain. Hypothesis spaces for learning. *Inform. Comput.*, 209(3):513–527, 2011. Special Issue: 3rd International Conference on Language and Automata Theory and Applications (LATA 2009).
- [2880] S. Jain and E. Kinber. Intrinsic complexity of learning geometrical concepts from positive data. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 177–193. Springer, 2001.
- [2881] S. Jain and E. Kinber. Learning languages from positive data and negative counterexamples. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 54–68. Springer, 2004.
- [2882] S. Jain and E. Kinber. Learning multiple languages in groups. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 256–268. Springer, Oct. 2005.
- [2883] S. Jain and E. Kinber. Iterative learning from positive data and negative counterexamples. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 154–168. Springer, Oct. 2006.
- [2884] S. Jain and E. Kinber. Learning and extending sublanguages. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 139–153. Springer, Oct. 2006.
- [2885] S. Jain and E. Kinber. Learning languages from positive data and a finite number of queries. *Inform. Comput.*, 204(1):123–175, 2006.
- [2886] S. Jain and E. Kinber. On learning languages from positive data and a limited number of short counterexamples. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 259–273, Berlin, 2006. Springer.
- [2887] S. Jain and E. Kinber. Iterative learning from positive data and negative counterexamples. *Inform. Comput.*, 205(12):1777–1805, 2007.

- [2888] S. Jain and E. Kinber. Learning multiple languages in groups. *Theoret. Comput. Sci.*, 387(1):67–76, 2007. Special issue for ALT 2005.
- [2889] S. Jain and E. Kinber. One-shot learners using negative counterexamples and nearest positive examples. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 257–271, Berlin, Oct. 2007. Springer.
- [2890] S. Jain and E. Kinber. Learning and extending sublanguages. *Theoret. Comput. Sci.*, 397(1-3):233–246, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [2891] S. Jain and E. Kinber. Learning languages from positive data and negative counterexamples. *J. Comput. Syst. Sci.*, 74(4):431–456, 2008.
- [2892] S. Jain and E. Kinber. Iterative learning from texts and counterexamples using additional information. In R. Gavalda, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 308–322, Berlin/Heidelberg, 2009. Springer.
- [2893] S. Jain and E. Kinber. Inductive inference of languages from samplings. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 330–344, Berlin, Heidelberg, New York, 2010. Springer.
- [2894] S. Jain and E. Kinber. Automatic learning from positive data and negative counterexamples. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 66–80, Berlin, Heidelberg, New York, 2012. Springer.
- [2895] S. Jain and E. Kinber. Mind change speed-up for learning languages from positive data. *Theoret. Comput. Sci.*, 489-490:37–47, 2013.
- [2896] S. Jain and E. Kinber. Parallel learning of automatic classes of languages. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 70–84, Berlin, Heidelberg, New York, 2014. Springer.
- [2897] S. Jain and E. Kinber. Intrinsic complexity of partial learning. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 174–188. Springer, 2016.
- [2898] S. Jain and E. Kinber. Parallel learning of automatic classes of languages. *Theoret. Comput. Sci.*, 650:25–44, 2016. Special Issue: Algorithmic Learning Theory.
- [2899] S. Jain, E. Kinber, S. Lange, R. Wiehagen, and T. Zeugmann. Learning languages and functions by erasing. *Theoret. Comput. Sci.*, 241(1-2):143–189, 2000. Special issue for ALT '96.

- [2900] S. Jain, E. Kinber, C. Papazian, C. Smith, and R. Wiehagen. On the intrinsic complexity of learning recursive functions. *Inform. Comput.*, 184(1):45–70, 2003.
- [2901] S. Jain, E. Kinber, and R. Wiehagen. On learning and co-learning of minimal programs. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 242–255. Springer, 1996.
- [2902] S. Jain, E. Kinber, and R. Wiehagen. Language learning from texts: Degrees of intrinsic complexity and their characterizations. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 47–58. Morgan Kaufmann, San Francisco, 2000.
- [2903] S. Jain, E. Kinber, and R. Wiehagen. Language learning from texts: Degrees of intrinsic complexity and their characterizations. *J. Comput. Syst. Sci.*, 63(3):305–354, 2001.
- [2904] S. Jain, E. Kinber, and R. Wiehagen. Learning all subfunctions of a function. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 714–728. Springer, 2003.
- [2905] S. Jain, E. Kinber, and R. Wiehagen. Learning all subfunctions of a function. *Inform. Comput.*, 192(2):185–215, 2004.
- [2906] S. Jain, E. Kinber, R. Wiehagen, and T. Zeugmann. Learning recursive functions refutably. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 283–298. Springer, 2001.
- [2907] S. Jain, E. Kinber, R. Wiehagen, and T. Zeugmann. On learning of functions refutably. *Theoret. Comput. Sci.*, 298(1):111–143, 2003.
- [2908] S. Jain and E. B. Kinber. Intrinsic complexity of learning geometrical concepts from positive data. *J. Comput. Syst. Sci.*, 67(3):546–607, 2003.
- [2909] S. Jain, T. Kötzing, J. Ma, and F. Stephan. On the role of update constraints and text-types in iterative learning. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 55–69. Berlin, Heidelberg, New York, 2014. Springer.
- [2910] S. Jain, T. Kötzing, and F. Stephan. Enlarging learnable classes. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 36–50. Berlin, Heidelberg, New York, 2012. Springer.
- [2911] S. Jain, T. Kötzing, and F. Stephan. Enlarging learnable classes. *Inform. Comput.*, 251:194–207, 2016.
- [2912] S. Jain, S. Lange, S. E. Moelius III, and S. Zilles. Incremental learning with temporary memory. *Theoret. Comput. Sci.*, 411(29-30):2757–2772, 2010. Special issue for ALT 2008.
- [2913] S. Jain, S. Lange, and J. Nessel. Learning of r.e. languages from good examples. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 32–47. Springer, 1997.

- [2914] S. Jain, S. Lange, and J. Nessel. On the learnability of recursively enumerable languages from good examples. *Theoret. Comput. Sci.*, 261(1):3–29, 2001. Special issue for ALT '97.
- [2915] S. Jain, S. Lange, and S. Zilles. Gold-style and query learning under various constraints on the target class. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 226–240. Springer, Oct. 2005.
- [2916] S. Jain, S. Lange, and S. Zilles. Towards a better understanding of incremental learning. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 169–183. Springer, Oct. 2006.
- [2917] S. Jain, S. Lange, and S. Zilles. A general comparison of language learning from examples and from queries. *Theoret. Comput. Sci.*, 387(1):51–66, 2007. Special issue for ALT 2005.
- [2918] S. Jain, S. Lange, and S. Zilles. Some natural conditions on incremental learning. *Inform. Comput.*, 205(11):1671–1684, 2007.
- [2919] S. Jain, Q. Luo, P. Semukhin, and F. Stephan. Uncountable automatic classes and learning. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 293–307, Berlin/Heidelberg, 2009. Springer.
- [2920] S. Jain, Q. Luo, P. Semukhin, and F. Stephan. Uncountable automatic classes and learning. *Theoret. Comput. Sci.*, 412(19):1805–1820, 2011.
- [2921] S. Jain, Q. Luo, and F. Stephan. Learnability of automatic classes. *J. Comput. Syst. Sci.*, 78(6):1910–1927, 2012.
- [2922] S. Jain, J. Ma, and F. Stephan. Priced learning. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 41–55, Berlin, Heidelberg, New York, 2015. Springer.
- [2923] S. Jain, E. Martin, and F. Stephan. Absolute versus probabilistic classification in a logical setting. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 327–342. Springer, Oct. 2005.
- [2924] S. Jain, E. Martin, and F. Stephan. Absolute versus probabilistic classification in a logical setting. *Theoret. Comput. Sci.*, 397(1-3):114–128, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [2925] S. Jain, E. Martin, and F. Stephan. Learning and classifying. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 70–83, Berlin, Heidelberg, New York, 2011. Springer.
- [2926] S. Jain, E. Martin, and F. Stephan. Robust learning of automatic classes of languages. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 55–69, Berlin, Heidelberg, New York, 2011. Springer.

- [2927] S. Jain, E. Martin, and F. Stephan. Learning and classifying. *Theoret. Comput. Sci.*, 482:73–85, 2013.
- [2928] S. Jain, E. Martin, and F. Stephan. Robust learning of automatic classes of languages. *J. Comput. Syst. Sci.*, 80(4):777–795, 2014.
- [2929] S. Jain, W. Menzel, and F. Stephan. Classes with easily learnable subclasses. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 218–232. Springer, 2002.
- [2930] S. Jain, W. Menzel, and F. Stephan. Classes with easily learnable subclasses. *Inform. Comput.*, 190(1):81–99, 2004.
- [2931] S. Jain, S. E. Moelius III, and S. Zilles. Learning without coding. *Theoret. Comput. Sci.*, 473:124–148, 2013. Special Issue on Learning Theory.
- [2932] S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors. *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, Berlin, Heidelberg, New York, Oct. 2013. Springer.
- [2933] S. Jain, R. Munos, F. Stephan, and T. Zeugmann. Editors’ introduction. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 1–12, Berlin, Heidelberg, New York, 2013. Springer.
- [2934] S. Jain, R. Munos, F. Stephan, and T. Zeugmann. Guest Editors’ foreword. *Theoret. Comput. Sci.*, 620:1–3, 2016. Special Issue: Algorithmic Learning Theory.
- [2935] S. Jain and J. Nessel. Some independence results for control structures in complete numberings. *J. Symbolic Logic*, 66(1):357–382, 2001.
- [2936] S. Jain, Y. K. Ng, and T. S. Tay. Learning languages in a union. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 235–250. Springer, 2001.
- [2937] S. Jain, Y. K. Ng, and T. S. Tay. Learning languages in a union. *J. Comput. Syst. Sci.*, 73(1):89–108, 2007.
- [2938] S. Jain, Y. S. Ong, and F. Stephan. Regular patterns, regular languages and context-free languages. *Information Processing Letters*, 110(24):1114–1119, 2010.
- [2939] S. Jain, D. Osherson, J. S. Royer, and A. Sharma. *Systems that Learn: An Introduction to Learning Theory, second edition*. MIT Press, Cambridge, Massachusetts, 1999.
- [2940] S. Jain and A. Sharma. Knowledge of an upper bound on grammar size helps language learning. Technical Report 283, University of Rochester, 1989.
- [2941] S. Jain and A. Sharma. Recursion theoretic characterizations of language learning. Technical Report 281, University of Rochester, 1989.
- [2942] S. Jain and A. Sharma. Finite learning by a team. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 163–177, San Mateo, CA, 1990. Morgan Kaufmann.



- [2943] S. Jain and A. Sharma. Hypothesis formation and language acquisition with an infinitely often correct teacher. In *Proc. of the Third Conference on Theoretical Aspects of Reasoning About Knowledge*, pages 225–239. Morgan Kaufmann, Inc., Mar. 1990.
- [2944] S. Jain and A. Sharma. Language learning by a “team” (extended abstract). In *Automata, Languages and Programming, 17th International Colloquium Warwick University, England, July 16-20, 1990 Proceedings*, volume 443 of *Lecture Notes in Computer Science*, pages 153–166. Springer-Verlag, 1990.
- [2945] S. Jain and A. Sharma. Program size restrictions in inductive learning. Technical Report 90-06, University of Delaware, Newark, Delaware, 1990.
- [2946] S. Jain and A. Sharma. Learning in the presence of partial explanations. *Inform. Comput.*, 95(2):162–191, 1991.
- [2947] S. Jain and A. Sharma. Restrictions on grammar size in language identification. In *Proc. MLNLO’91, Machine Learning of Natural Language and Ontology, Stanford University, California. Document D91-09, DFKI: Kaiserslautern FRG, 1991*, pages 87–92, Mar. 1991.
- [2948] S. Jain and A. Sharma. Computational limits on team identification of languages. Technical Report 9301, School of Computer Science and Engineering; University of New South Wales, 1993.
- [2949] S. Jain and A. Sharma. Learning with the knowledge of an upper bound on program size. *Inform. Comput.*, 102(1):118–166, 1993.
- [2950] S. Jain and A. Sharma. On the non-existence of maximal inference degrees for language identification. *Inform. Proc. Lett.*, 47(2):81–88, 1993.
- [2951] S. Jain and A. Sharma. Probability is more powerful than team for language identification from positive data. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 192–198, New York, New York, 1993. ACM Press.
- [2952] S. Jain and A. Sharma. Characterizing language identification by standardizing operations. *J. of Comput. Syst. Sci.*, 49(1):96–107, 1994.
- [2953] S. Jain and A. Sharma. On monotonic strategies for learning r.e. languages. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 349–364. Springer-Verlag, 1994.
- [2954] S. Jain and A. Sharma. On the intrinsic complexity of language identification. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 278–286, New York, NY, 1994. ACM Press.
- [2955] S. Jain and A. Sharma. Program size restrictions in computational learning. *Theoret. Comput. Sci.*, 127(2):351–386, 1994.
- [2956] S. Jain and A. Sharma. On aggregating teams of learning machines. *Theoret. Comput. Sci.*, 137(1):85–108, 1995. Special issue for ALT ’93.
- [2957] S. Jain and A. Sharma. On identification by teams and probabilistic machines. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 108–145. Springer, 1995.

- [2958] S. Jain and A. Sharma. Prudence in vacillatory language identification. *Math. Syst. Theory*, 28(3):267–279, 1995.
- [2959] S. Jain and A. Sharma. The structure of intrinsic complexity of learning. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 169–181. Springer, 1995.
- [2960] S. Jain and A. Sharma. Computational limits on team identification of languages. *Inform. Comput.*, 130(1):19–60, 1996.
- [2961] S. Jain and A. Sharma. The intrinsic complexity of language identification. *J. of Comput. Syst. Sci.*, 52(3):393–402, 1996.
- [2962] S. Jain and A. Sharma. Team learning of recursive languages. In *Proceedings of the Fourth Pacific Rim International Conference on Artificial Intelligence, Cairns, Australia*, volume 1114 of *Lecture Notes in Artificial Intelligence*, pages 324–335. Springer, 1996.
- [2963] S. Jain and A. Sharma. Characterizing language learning in terms of computable numberings. *Annals of Pure and Applied Logic*, 84(1):51–72, 1997.
- [2964] S. Jain and A. Sharma. Elementary formal systems, intrinsic complexity, and procrastination. *Inform. Comput.*, 132(1):65–84, 1997. first in Colt '96.
- [2965] S. Jain and A. Sharma. The structure of intrinsic complexity of learning. *Journal of Symbolic Logic*, 62(4):1187–1201, 1997.
- [2966] S. Jain and A. Sharma. Generalization and specialization strategies for learning r.e. languages. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):1–26, 1998. Special issue for AII '94 and ALT '94.
- [2967] S. Jain and A. Sharma. Mind change complexity of learning logic programs. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 198–213. Springer, 1999.
- [2968] S. Jain and A. Sharma. On a generalized notion of mistake bounds. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 249–256, New York, NY, 1999. ACM Press.
- [2969] S. Jain and A. Sharma. Team learning of computable languages. *Theory of Computing Systems*, 33(1):35–58, 2000.
- [2970] S. Jain and A. Sharma. On a generalized notion of mistake bounds. *Inform. Comput.*, 166(2):156–166, May 2001.
- [2971] S. Jain and A. Sharma. Mind change complexity of learning logic programs. *Theoret. Comput. Sci.*, 284(1):143–160, 2002.
- [2972] S. Jain, A. Sharma, and J. Case. Convergence to nearly minimal size grammars by vacillating learning machines. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 189–199, San Mateo, CA, 1989. Morgan Kaufmann.
- [2973] S. Jain, A. Sharma, and M. Velauthapillai. Finite identification of functions by teams with success ratio  $\frac{1}{2}$  and above. *Inform. Comput.*, 121(2):201–213, 1995.

- [2974] S. Jain, H. U. Simon, and E. Tomita, editors. *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 2005. Springer.
- [2975] S. Jain, H. U. Simon, and E. Tomita. Editors' introduction. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 1–9. Springer, Oct. 2005.
- [2976] S. Jain, C. Smith, and R. Wiehagen. On the power of learning robustly. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 187–197, New York, NY, 1998. ACM Press.
- [2977] S. Jain, C. Smith, and R. Wiehagen. Robust learning is rich. *J. Comput. Syst. Sci.*, 62(1):178–212, 2001.
- [2978] S. Jain and F. Stephan. Learning by switching type of information. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 205–218. Springer, 2001.
- [2979] S. Jain and F. Stephan. Learning how to separate. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 219–234. Springer, 2001.
- [2980] S. Jain and F. Stephan. Learning by switching type of information. *Inform. Comput.*, 185(1):89–104, 2003.
- [2981] S. Jain and F. Stephan. Learning how to separate. *Theoret. Comput. Sci.*, 313(2):209–228, 2004. Special issue for ALT 2001.
- [2982] S. Jain and F. Stephan. Some recent results in U-shaped learning. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 421–431, Berlin, 2006. Springer.
- [2983] S. Jain and F. Stephan. Learning in friedberg numberings. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 79–93, Berlin, Oct. 2007. Springer.
- [2984] S. Jain and F. Stephan. Mitotic classes. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 218–232, Berlin, 2007. Springer.
- [2985] S. Jain and F. Stephan. Mitotic classes in inductive inference. *SIAM J. Comput.*, 38(4):1283–1299, 2008.
- [2986] S. Jain and F. Stephan. Numberings optimal for learning. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 434–448, Berlin, Oct. 2008. Springer.
- [2987] S. Jain and F. Stephan. Complexity of inductive inference. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 198–201. Springer, New York, 2010.

- [2988] S. Jain and F. Stephan. Inductive inference. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 523–528. Springer, New York, 2010.
- [2989] S. Jain and F. Stephan. Query-based learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 820–822. Springer, New York, 2010.
- [2990] S. Jain, F. Stephan, and S. Terwijn. Counting extensional differences in BC-learning. *Inform. Comput.*, 188(1):127–142, 2004.
- [2991] S. Jain, F. Stephan, and N. Ye. Prescribed learning of R.E. classes. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 64–78, Berlin, Oct. 2007. Springer.
- [2992] S. Jain, F. Stephan, and N. Ye. Learning from streams. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 338–352, Berlin/Heidelberg, 2009. Springer.
- [2993] S. Jain, F. Stephan, and N. Ye. Prescribed learning of r.e. classes. *Theoret. Comput. Sci.*, 410(9):1796–1806, 2009. Special issue for ALT 2007.
- [2994] S. Jain, F. Stephan, and T. Zeugmann. On the amount of nonconstructivity in learning formal languages from positive data. In M. Agrawal, S. B. Cooper, and A. Li, editors, *Theory and Applications of Models of Computation, 9th Annual Conference, TAMC 2012, Beijing, China, May 16-21, 2012, Proceedings*, volume 7287 of *Lecture Notes in Computer Science*, pages 423–434. Springer, 2012.
- [2995] G. M. James. Variance and bias for general loss functions. *Machine Learning*, 51(2):115–135, 2003.
- [2996] K. Jamieson, M. Malloy, R. Nowak, and S. Bubeck. lil’ UCB : An optimal exploration algorithm for multi-armed bandits. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 423–439, Barcelona, Spain, 13–15 June 2014. PMLR.
- [2997] M. V. Janardhanan. Graph verification with a betweenness oracle. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 238–249, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [2998] G.-J. Jang and T.-W. Lee. A maximum likelihood approach to single-channel source separation. *Journal of Machine Learning Research*, 4:365–392, 2003.
- [2999] S.-W. Jang, G. young Kim, and H.-I. Choi. FAM-based fuzzy inference for detecting shot transitions. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 52–61. Springer, 2001.
- [3000] C. Z. Janikow. A knowledge-intensive genetic algorithm for supervised learning. *Machine Learning*, 13:189–228, 1993.
- [3001] F. Janoos, M. Mohri, and K. Sridharan, editors. *Algorithmic Learning Theory, 7–9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*. PMLR, 2018. Proceedings of the 29th International Conference on Algorithmic Learning Theory.

- [3002] K. P. Jantke. Automatic synthesis of programs and inductive inference of functions. In *Fundamentals of Computations Theory, FCT '79, Proc. of the Conference on Algebraic, Arithmetic, and Categorical Methods in Computation Theory, Berlin/Wendisch-Rietz (GDR), September 17-21, 1979*, pages 219–225, Berlin, 1979. Akademie-Verlag.
- [3003] K. P. Jantke. *Leistungsfähigkeit und Kompliziertheit universeller Verfahren zur Erkennung allgemein-rekursiver Funktionen*. PhD thesis, Humboldt-Universität, Berlin, 1979.
- [3004] K. P. Jantke. Natural properties of strategies identifying recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 15:487–496, 1979.
- [3005] K. P. Jantke. Polynomial-time inference of general pattern languages. In *Proc. of the Symposium of Theoretical Aspects of Computer Science*, volume 166 of *Lecture Notes in Computer Science*, pages 314–325. Springer-Verlag, 1984.
- [3006] K. P. Jantke. An algebraic framework for inductive program synthesis. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 101–122. Springer-Verlag, 1986.
- [3007] K. P. Jantke, editor. *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*. Springer-Verlag, 1987.
- [3008] K. P. Jantke, editor. *Analogical and Inductive Inference, International Workshop AII '89. Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Oct. 1989.
- [3009] K. P. Jantke. Monotonic and non-monotonic inductive inference. *New Generation Computing*, 8(4):349–360, 1991.
- [3010] K. P. Jantke. Monotonic and nonmonotonic inductive inference of functions and patterns. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany, December 1990, Proceedings*, volume 543 of *Lecture Notes in Computer Science*, pages 161–177. Springer-Verlag, 1991.
- [3011] K. P. Jantke, editor. *Analogical and Inductive Inference, International Workshop AII '92. Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Oct. 1992.
- [3012] K. P. Jantke. Case-based learning in inductive inference. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 218–223, New York, NY, 1992. ACM Press.
- [3013] K. P. Jantke. Reflecting and self-confident inductive inference machines. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 282–297. Springer, 1995.
- [3014] K. P. Jantke and H.-R. Beick. Combining postulates of naturalness in inductive inference. *Elektronische Informationsverarbeitung und Kybernetik*, 17(8/9):465–484, 1981.
- [3015] K. P. Jantke, S. Kobayashi, E. Tomita, and T. Yokomori, editors. *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Nov. 1993.

- [3016] K. P. Jantke and S. Lange. Case-based representation and learning of pattern languages. *Theoret. Comput. Sci.*, 137(1):25–51, 1995. Special issue for ALT '93.
- [3017] K. P. Jantke, T. Shinohara, and T. Zeugmann, editors. *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 1995. Springer-Verlag.
- [3018] K. P. Jantke, T. Shinohara, and T. Zeugmann. Editors' introduction. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages ix–xv. Springer, 1995.
- [3019] N. Japkowicz. Supervised versus unsupervised binary-learning by feedforward neural networks. *Machine Learning*, 42(1/2):97–122, 2001.
- [3020] P. Jappy, R. Nock, and O. Gascuel. Negative robust learning results for Horn clause programs. In *Proc. 13th International Conference on Machine Learning*, pages 258–265. Morgan Kaufmann, 1996.
- [3021] A. Jardine, J. Chandlee, R. Eyraud, and J. Heinz. Very efficient learning of structured classes of subsequential functions from positive data. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 94–108, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [3022] M. Jayasrirani, M. H. Begam, and D. G. Thomas. Learning of regular omega-tree languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22–24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 295–297. Springer, 2008.
- [3023] M. Jayasrirani, M. H. Begam, D. G. Thomas, and J. D. Emerald. Learning of bi- $\omega$  languages from factors. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 139–144, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [3024] M. Jayasrirani, D. G. Thomas, A. K. Nagar, and T. Robinson. Learning of Church-Rosser tree rewriting systems. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13–16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 284–287. Springer, 2010.
- [3025] T. Jebara. Kernelizing sorting, permutation, and alignment for minimum volume PCA. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 609–623. Springer, 2004.
- [3026] T. Jebara and R. Kondor. Bhattacharyya and expected likelihood kernels. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 57–71. Springer, 2003.

- [3027] S. Jegelka, S. Sra, and A. Banerjee. Approximation algorithms for tensor clustering. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 368–383, Berlin/Heidelberg, 2009. Springer.
- [3028] D. D. Jensen and P. R. Cohen. Multiple comparisons in induction algorithms. *Machine Learning*, 38(3):309–338, 2000.
- [3029] M. Jerrum. Simple translation-invariant concepts are hard to learn. *Inform. Comput.*, 113(2):300–311, 1994.
- [3030] C. Ji and D. Psaltis. The VC-dimension vs. the statistical capacity for two layer networks with binary weights. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 250–256, San Mateo, CA, 1991. Morgan Kaufmann.
- [3031] A. X. Jiang and K. Leyton-Brown. Bidding agents for online auctions with hidden bids. *Machine Learning*, 67(1-2):117–143, 2007.
- [3032] F. Jiang and M. L. Littman. Approximate dimension equalization in vector-based information retrieval. In *Proc. 17th International Conf. on Machine Learning*, pages 423–430. Morgan Kaufmann, San Francisco, CA, 2000.
- [3033] T. Jiang, P. Kearney, and M. Li. A polynomial time approximation scheme for inferring evolutionary trees from quartet topologies and its application. *SIAM J. Comput.*, 30(6):1942–1961, 2000.
- [3034] T. Jiang and M. Li. On the complexity of learning strings and sequences. *Theoret. Comput. Sci.*, 119(2):363–371, 1993.
- [3035] T. Jiang, A. Salomaa, K. Salomaa, and S. Yu. Inclusion is undecidable for pattern languages. In *Automata, Languages and Programming, 20nd International Colloquium, ICALP '93, Proceedings*, volume 700 of *Lecture Notes in Computer Science*, pages 301–312. Springer-Verlag, 1993.
- [3036] W. Jiang. Some theoretical aspects of boosting in the presence of noisy data. In *Proc. 18th International Conf. on Machine Learning*, pages 234–241. Morgan Kaufmann, San Francisco, CA, 2001.
- [3037] R. Jin and A. G. Hauptmann. Learning to select good title words: A new approach based on reverse information retrieval. In *Proc. 18th International Conf. on Machine Learning*, pages 242–249. Morgan Kaufmann, San Francisco, CA, 2001.
- [3038] R. Jin, S. C. H. Hoi, and T. Yang. Online multiple kernel learning: Algorithms and mistake bounds. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 390–404, Berlin, Heidelberg, New York, 2010. Springer.
- [3039] T. Joachims. A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization. In *Proc. 14th International Conference on Machine Learning*, pages 143–151. Morgan Kaufmann, 1997.
- [3040] T. Joachims. Transductive inference for text classification using support vector machines. In *Proc. 16th International Conf. on Machine Learning*, pages 200–209. Morgan Kaufmann, San Francisco, CA, 1999.

- [3041] T. Joachims. Estimating the generalization performance of an SVM efficiently. In *Proc. 17th International Conf. on Machine Learning*, pages 431–438. Morgan Kaufmann, San Francisco, CA, 2000.
- [3042] T. Joachims, N. Cristianini, and J. Shawe-Taylor. Composite kernels for hypertext categorisation. In *Proc. 18th International Conf. on Machine Learning*, pages 250–257. Morgan Kaufmann, San Francisco, CA, 2001.
- [3043] G. H. John, R. Kohavi, and K. Pfleger. Irrelevant features and the subset selection problem. In *Proc. 11th International Conference on Machine Learning*, pages 121–129. Morgan Kaufmann, 1994.
- [3044] G. H. John and P. Langley. Estimating continuous distributions in Bayesian classifiers. In *Proc. 11th Conference on Uncertainty in Artificial Intelligence*, pages 338–345. Morgan Kaufmann, 1995.
- [3045] M. Johnson. Learning and parsing stochastic unification-based grammars. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 671–683. Springer, 2003.
- [3046] J. Jones and T. Oates. Learning deterministic finite automata from interleaved strings. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 80–93. Springer, 2010.
- [3047] L. K. Jones. A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training. *Annals of Statistics*, 20:608–613, 1992.
- [3048] R. M. Jones and K. Vanlehn. Acquisition of children’s addition strategies: A model of impasse-free, knowledge-level learning. *Machine Learning*, 16(1/2):11–36, 1994.
- [3049] K. A. D. Jong, W. M. Spears, and D. F. Gordon. Using genetic algorithms for concept learning. *Machine Learning*, 13:161–188, 1993.
- [3050] K. D. Jong. Learning with genetic algorithms: An overview. *Machine Learning*, 3:121–138, 1988.
- [3051] K. D. Jong. Special issue on genetic algorithms. *Machine Learning*, 5(4), Oct. 1990.
- [3052] D. de Jongh and M. Kanazawa. Angluin’s theorem for indexed families of r.e. sets and applications. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 193–204, New York, NY, 1996. ACM Press.
- [3053] I. Jonyer, D. J. Cook, and L. B. Holder. Graph-based hierarchical conceptual clustering. *Journal of Machine Learning Research*, 2:19–43, 2001.
- [3054] M. I. Jordan. Supervised learning and systems with excess degrees of freedom. Technical Report 88-27, COINS, May 1988.
- [3055] M. I. Jordan. A statistical approach to decision tree modeling. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 13–20, New York, NY, 1994. ACM Press.



- [3056] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37(2):183–233, 1999.
- [3057] M. I. Jordan and D. E. Rumelhart. Forward models: Supervised learning with a distal teacher. *Cognitive Science*, 16:307–354, 1992.
- [3058] M. I. Jordan and L. Xu. Convergence results for the EM approach to mixtures of experts architectures. *Neural Networks*, 8(9):1409–1431, 1995.
- [3059] J. P. Jouannaud and G. Guiho. Inference of functions with an interactive system. In *Machine Intelligence 9*, pages 227–250. John Wiley and Sons, N. Y., 1979.
- [3060] J. P. Jouannaud and Y. Kodratoff. Characterization of a class of functions synthesized by a summers-like method using a B.M.W. matching technique. In *Proc. 6th IJCAI*, pages 440–447. IJCAI, 1979.
- [3061] J. P. Jouannaud and Y. Kodratoff. An automatic construction of LISP programs by transformations of functions synthesized from their input-output behavior. *International J. of Policy Analysis and Information Systems*, 4:331–358, 1980.
- [3062] J. P. Jouannaud and T. P. Treuil. SISP/1, an interactive system able to synthesize functions from examples. In *Proc. 5th IJCAI*, pages 412–418. IJCAI, 1977.
- [3063] P. Joulani, A. György, and C. Szepesvári. A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, and variational bounds. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 681–720, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3064] P. Ju, L. P. Kaelbling, and Y. Singer. State-based classification of finger gestures from electromyographic signals. In *Proc. 17th International Conf. on Machine Learning*, pages 439–446. Morgan Kaufmann, San Francisco, CA, 2000.
- [3065] B. Juang. On the hidden Markov model and dynamic time warping for speech recognition – A unified view. *AT&T Bell Laboratories Technical Journal*, 63(7):1213–1242, Sept. 1985.
- [3066] B. Juba and S. Vempala. Semantic communication for simple goals is equivalent to on-line learning. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 277–291, Berlin, Heidelberg, New York, 2011. Springer.
- [3067] J. S. Judd. Complexity of connectionist learning with various node functions. Technical Report 87-60, Department of Computer and Information Science, University of Massachusetts at Amherst, July 1987. Also presented at the First IEEE International Conference on Neural Networks, June 21–24, 1987, San Diego, California.
- [3068] J. S. Judd. *Neural Network Design and the Complexity of Learning*. MIT Press, Cambridge, Massachusetts, 1990.
- [3069] J. S. Judd. A reply to Honavar’s book review of neural network design and the complexity of learning. *Machine Learning*, 9:99–100, 1992.
- [3070] S. Judd. Learning in neural networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 2–8, San Mateo, CA, 1988. Morgan Kaufmann.

- [3071] S. Judd. How loading complexity is affected by node function sets. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 3, pages 49–64. MIT Press, 1994.
- [3072] A. Juhos. Can missing information be also useful? In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 88–104. Springer-Verlag, 1986.
- [3073] H. Juille and J. B. Pollack. Coevolutionary learning: a case study. In *Proc. 15th International Conf. on Machine Learning*, pages 251–259. Morgan Kaufmann, San Francisco, CA, 1998.
- [3074] H. Juillé and J. B. Pollack. A stochastic search approach to grammar induction. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 126–137. Springer, 1998.
- [3075] G. Jumarie. *Relative information*. Springer-Verlag, 1990.
- [3076] H. Jung. *Zur Untersuchung von Abstrakten Interaktiven Erkennungssystemen*. PhD thesis, Humbolt University, Berlin, 1977.
- [3077] T. Jurgenson and Y. Mansour. Learning decision trees with stochastic linear classifiers. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 489–528. PMLR, 07–09 Apr. 2018.
- [3078] M. Kääriäinen. Generalization error bounds using unlabeled data. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 127–142. Springer, 2005.
- [3079] M. Kääriäinen. Active learning in the non-realizable case. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 63–77. Springer, Oct. 2006.
- [3080] A. Kabán. On compressive ensemble induced regularisation: How close is the finite ensemble precision matrix to the infinite ensemble? In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 617–628, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3081] A. Kabán and R. J. Durrant. Dimension-adaptive bounds on compressive FLD classification. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 294–308, Berlin, Heidelberg, New York, 2013. Springer.
- [3082] Y. Kabashima and S. Shinomoto. Acceleration of learning in binary choice problems. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 446–452, New York, NY, 1993. ACM Press.
- [3083] Y. Kabashima and S. Uda. A BP-based algorithm for performing Bayesian inference in large perceptron-type networks. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 479–493. Springer, 2004.

- [3084] M. W. Kadous. Learning comprehensible descriptions of multivariate time series. In *Proc. 16th International Conf. on Machine Learning*, pages 454–463. Morgan Kaufmann, San Francisco, CA, 1999.
- [3085] L. P. Kaelbling. Associative methods in reinforcement learning: an empirical study. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 9, pages 133–153. MIT Press, 1994.
- [3086] L. P. Kaelbling. Associative reinforcement learning: A generate and test algorithm. *Machine Learning*, 15(3):299–319, 1994.
- [3087] L. P. Kaelbling. Associative reinforcement learning: Functions in  $k$ -DNF. *Machine Learning*, 15(3):279–298, 1994.
- [3088] L. P. Kaelbling. Introduction. *Machine Learning*, 22:7–9, 1996.
- [3089] S. Kakade. Optimizing average reward using discounted rewards. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 605–615. Springer, 2001.
- [3090] S. M. Kakade and D. P. Foster. Deterministic calibration and nash equilibrium. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 33–48. Springer, 2004.
- [3091] S. M. Kakade and D. P. Foster. Multi-view regression via canonical correlation analysis. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 82–96, Berlin, 2007. Springer.
- [3092] S. M. Kakade and D. P. Foster. Deterministic calibration and nash equilibrium. *J. Comput. Syst. Sci.*, 74(1):115–130, 2008.
- [3093] S. M. Kakade and M. Kearns. Trading in Markovian price models. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 606–620. Springer, 2005.
- [3094] S. M. Kakade, M. Kearns, and L. E. Ortiz. Graphical economics. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 17–32. Springer, 2004.
- [3095] S. M. Kakade and U. von Luxburg. Preface. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages i–i, Budapest, Hungary, 09–11 June 2011. PMLR.
- [3096] S. M. Kakade and U. von Luxburg, editors. *Proceedings of the 24th Annual Conference on Learning Theory, 9-11 June 2011, Budapest, Hungary*, volume 19 of *Proceedings of Machine Learning Research*. PMLR, 2011.

- [3097] K. ichiro Kakihara and H. Imai. Notes on the PAC learning of geometric concepts with additional information. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 252–259. Springer-Verlag, 1993.
- [3098] A. Kalai. Learning monotonic linear functions. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 487–501. Springer, 2004.
- [3099] A. Kalai and R. A. Servedio. Boosting in the presence of noise. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 9-11, 2003*, pages 196–205. ACM Press, 2003.
- [3100] A. Kalai and S. Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, 3:423–440, 2002.
- [3101] A. Kalai and S. Vempala. Efficient algorithms for online decision problems. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 26–40. Springer, 2003.
- [3102] A. Kalai and S. Vempala. Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.*, 71(3):291–307, 2005. Special issue Learning Theory 2003.
- [3103] A. T. Kalai. Learning nested halfspaces and uphill decision trees. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 378–392, Berlin, 2007. Springer.
- [3104] A. T. Kalai, V. Kanade, and Y. Mansour. Reliable agnostic learning. *J. Comput. Syst. Sci.*, 78(5):1481–1495, 2012.
- [3105] A. T. Kalai, A. R. Klivans, Y. Mansour, and R. A. Servedio. Agnostically learning halfspaces. *SIAM J. Comput.*, 37(6):1777–1805, 2008.
- [3106] A. T. Kalai and R. A. Servedio. Boosting in the presence of noise. *J. Comput. Syst. Sci.*, 71(3):266–290, 2005. Special issue Learning Theory 2003.
- [3107] S. Kale. Multiarmed bandits with limited expert advice. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 107–122, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3108] S. Kale. Open problem: Efficient online sparse regression. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1299–1301, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3109] S. Kale and O. Shamir. Preface: Conference on learning theory (COLT), 2017. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1–3, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3110] S. Kale and O. Shamir, editors. *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*. PMLR, 2017.

- [3111] D. Kalles and T. Morris. Efficient incremental induction of decision trees. *Machine Learning*, 24:231–242, 1996.
- [3112] N. Kallus. Instrument-armed bandits. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 529–546. PMLR, 07–09 Apr. 2018.
- [3113] M. Kallweit and H. U. Simon. A close look to margin complexity and related parameters. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 437–456, Budapest, Hungary, 09–11 June 2011. PMLR.
- [3114] Y. Kalnishkan. General linear relations among different types of predictive complexity. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 323–334. Springer, 1999.
- [3115] Y. Kalnishkan. Linear relations between square-loss and Kolmogorov complexity. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 226–232, New York, NY, 1999. ACM Press.
- [3116] Y. Kalnishkan. Complexity approximation principle and Rissanen’s approach to real-valued parameters. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 203–210. Springer, 2000.
- [3117] Y. Kalnishkan. An upper bound for aggregating algorithm for regression with changing dependencies. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 238–252. Springer, 2016.
- [3118] Y. Kalnishkan, V. Vovk, and M. V. Vyugin. A criterion for the existence of predictive complexity for binary games. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 249–263. Springer, 2004.
- [3119] Y. Kalnishkan, V. Vovk, and M. V. Vyugin. Loss functions, complexities, and the Legendre transformation. *Theoret. Comput. Sci.*, 313(2):195–207, 2004. Special issue for ALT 2001.
- [3120] Y. Kalnishkan, V. Vovk, and M. V. Vyugin. How many strings are easy to predict? *Inform. Comput.*, 201(1):55–71, 2005.
- [3121] Y. Kalnishkan, V. Vovk, and M. V. Vyugin. Generalised entropy and asymptotic complexities of languages. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 293–307, Berlin, 2007. Springer.
- [3122] Y. Kalnishkan and M. V. Vyugin. Mixability and the existence of weak complexities. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 105–120. Springer, 2002.
- [3123] Y. Kalnishkan and M. V. Vyugin. On the absence of predictive complexity for some games. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 164–172. Springer, 2002.

- [3124] Y. Kalnishkan and M. V. Vyugin. The weak aggregating algorithm and weak mixability. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 188–203. Springer, 2005.
- [3125] Y. Kalnishkan, M. V. Vyugin, and V. Vovk. Loss functions, complexities, and the Legendre transformation. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 181–189. Springer, 2001.
- [3126] B. Kalyanasundaram and M. Velauthapillai. Simulating teams with many conjectures. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 201–214. Springer, 1995.
- [3127] B. Kalyanasundaram and M. Velauthapillai. Taming teams with mind changes. *J. Comput. Syst. Sci.*, 74(4):513–526, 2008.
- [3128] T. Kalyani, V. R. Dare, D. G. Thomas, and T. Robinson. Iso-array acceptors and learning. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 327–339, Berlin, 2006. Springer.
- [3129] P. Kamalaruban, R. Williamson, and X. Zhang. Exp-concavity of proper composite losses. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1035–1065, Paris, France, 03–06 July 2015. PMLR.
- [3130] S. Kamath, A. Orlitsky, D. Pichapati, and A. T. Suresh. On learning distributions from their samples. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1066–1100, Paris, France, 03–06 July 2015. PMLR.
- [3131] Y. Kameda, H. Tokunaga, and A. Yamamoto. Learning bounded unions of noetherian closed set systems via characteristic sets. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 98–110. Springer, 2008.
- [3132] R. Kamimura. Competitive learning by entropy minimization. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 111–122. Springer-Verlag, 1993.
- [3133] R. Kamimura. Minimizing alpha-information for generalization and interpretation. *Algorithmica*, 22(1/2):173–197, 1998.
- [3134] T. Kamishima and F. Motoyoshi. Learning from cluster examples. *Machine Learning*, 53(3):199–233, 2003.
- [3135] V. Kanade and E. Mossel. MCMC learning. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1101–1128, Paris, France, 03–06 July 2015. PMLR.

- [3136] V. Kanade and J. Thaler. Distribution-independent reliable learning. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 3–24, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3137] T. Kanamori. Multiclass boosting algorithms for shrinkage estimators of class probability. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 358–372, Berlin, Oct. 2007. Springer.
- [3138] T. Kanamori, A. Takeda, and T. Suzuki. A conjugate property between loss functions and uncertainty sets in classification problems. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 29.1–29.23, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [3139] P. Kanani and A. McCallum. Resource-bounded information gathering for correlation clustering. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 625–627, Berlin, 2007. Springer.
- [3140] J. Kandola, T. Graepel, and J. Shawe-Taylor. Reducing kernel matrix diagonal dominance using semi-definite programming. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 288–302. Springer, 2003.
- [3141] J. Kandola, T. Hofmann, T. Poggio, and J. Shawe-Taylor. Introduction to the special issue on machine learning methods for text and images. *Journal of Machine Learning Research*, 3:1023–1024, 2003.
- [3142] D. Kane, A. Klivans, and R. Meka. Learning halfspaces under log-concave densities: Polynomial approximations and moment matching. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 522–545. JMLR.org, 2013.
- [3143] R. Kane, I. Tchoumatchenko, and M. Milgram. Extraction of knowledge from data using constrained neural networks. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 420–425. Springer-Verlag, 1993.
- [3144] S. Kaneda, H. Almuallim, Y. Akiba, and M. Ishi. Learning from expert hypotheses and training examples. *IEICE Transactions on Information and Systems*, E80-D(12):1205–1214, 1998.
- [3145] P. Kanerva. *Sparse Distributed Memory*. MIT Press, Cambridge, Massachusetts, 1988.
- [3146] K. Kang and J.-H. Oh. Learning by a population of perceptrons. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 297–300, New York, NY, 1995. ACM Press.
- [3147] S. Kannamma, D. G. Thomas, and K. Rangarajan. Inference of uniquely terminating EML. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 289–291. Springer, 2008.

- [3148] R. Kannan, H. Salmasian, and S. Vempala. The spectral method for general mixture models. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 444–457. Springer, 2005.
- [3149] R. Kannan and S. Vempala. The hidden hubs problem. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1190–1213, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3150] R. Kannan, S. Vempala, and D. Woodruff. Principal component analysis and higher correlations for distributed data. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1040–1057, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3151] S. Kannan. On the query complexity of learning. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 58–66, New York, NY, 1993. ACM Press.
- [3152] M. Kao, J. H. Reif, and S. R. Tate. Searching in an unknown environment: An optimal randomized algorithm for the cow-path problem. In *Proc. SODA 93*, pages 441–447, Jan. 1993.
- [3153] H. Kaplan, E. Kushilevitz, and Y. Mansour. Learning with attribute costs. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 356–365. ACM Press, 2005.
- [3154] J. N. Kapur and H. K. Kesavan. *Entropy Optimization Principles with Applications*. Academic Press, Inc., 1992.
- [3155] S. Kapur. *Computational Learning of Languages*. PhD thesis, Department of Computer Science, Cornell University, Ithaca, New York, 1991.
- [3156] S. Kapur. Monotonic language learning. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 147–158. Springer-Verlag, 1993.
- [3157] S. Kapur. Uniform characterizations of various kinds of language learning. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 197–208. Springer-Verlag, 1993.
- [3158] S. Kapur. Language learning under various types of constraint combinations. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 365–378. Springer-Verlag, 1994.
- [3159] S. Kapur and G. Bilardi. Language learning from stochastic input. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 303–310, New York, NY, 1992. ACM Press.
- [3160] S. Kapur and G. Bilardi. On uniform learnability of language families. *Information Processing Letters*, 44(1):35–38, 1992.
- [3161] S. Kapur and G. Bilardi. Language learning without overgeneralization. *Theoret. Comput. Sci.*, 141(1/2):151–162, 1995. first in STACS '92.



- [3162] S. Kapur, B. Lust, W. Harbert, and G. Motohardjono. Universal grammar and learnability theory: The case of binding domains and the ‘subset principle’. In E. Reuland and W. Abraham, editors, *Knowledge and Language*, volume 1. Kluwer Academic Publishers, 1993.
- [3163] A. Karalič and I. Bratko. First order regression. *Machine Learning*, 26(2-3):147–176, 1997.
- [3164] N. Karampatziakis, G. Paliouras, D. Pierrakos, and P. Stamatopoulos. Navigation pattern discovery using grammatical inference. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 187–198. Springer, 2004.
- [3165] Z. Karnin and E. Liberty. Online PCA with spectral bounds. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1129–1140, Paris, France, 03–06 July 2015. PMLR.
- [3166] Z. Karnin, E. Liberty, S. Lovett, R. Schwartz, and O. Weinstein. Unsupervised SVMs: On the complexity of the furthest hyperplane problem. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 2.1–2.17, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [3167] P. D. Karp. Design methods for scientific hypothesis formation and their application to molecular biology. *Machine Learning*, 12:89–116, 1993.
- [3168] M. Karpinski and A. Macintyre. Bounding VC-dimension for neural networks: progress and prospects. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 337–341. Springer, 1995.
- [3169] M. Karpinski and A. Macintyre. Polynomial bounds for VC dimension of sigmoidal neural networks. In *Proc. 27th Annual ACM Symposium on Theory of Computing*, pages 200–208, New York, NY, 1995. ACM Press.
- [3170] M. Karpinski and A. Macintyre. Polynomial bounds for VC dimension of sigmoidal and general pfaffian neural networks. *J. Comput. Syst. Sci.*, 54(1):169–176, 1997.
- [3171] M. Karpinski and T. Werther. VC dimension and sampling complexity of learning sparse polynomials and rational functions. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 11, pages 331–354. MIT Press, 1994.
- [3172] Y. Kasahara, R. Niitsu, and T. Sato. Computational analysis of plasma waves and particles in the auroral region observed by scientific satellite. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 426–437. Springer, 2001.
- [3173] A. Kasprzik. A learning algorithm for multi-dimensional trees, or: Learning beyond context-freeness. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 111–124. Springer, 2008.

- [3174] A. Kasprzik. Generalizing over several learning settings. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 288–292. Springer, 2010.
- [3175] A. Kasprzik. Four one-shot learners for regular tree languages and their polynomial characterizability. *Theoret. Comput. Sci.*, 485:85–106, 2013.
- [3176] A. Kasprzik and R. Yoshinaka. Distributional learning of simple context-free tree grammars. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 398–412, Berlin, Heidelberg, New York, 2011. Springer.
- [3177] S. Katayama, H. Kimura, and S. Kobayashi. A universal generalization for temporal-difference learning using haar basis functions. In *Proc. 17th International Conf. on Machine Learning*, pages 447–454. Morgan Kaufmann, San Francisco, CA, 2000.
- [3178] H. Kato, S. Matsumoto, and T. Miyahara. Learning of elementary formal systems with two clauses using queries. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 211–225. Springer, Oct. 2005.
- [3179] S. Katrenko and P. Adriaans. Grammatical inference in practice: A case study in the biomedical domain. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 188–200, Berlin, 2006. Springer.
- [3180] S. Katrenko and M. van Zaanen. Rademacher complexity and grammar induction algorithms: What it may (not) tell us. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 293–296. Springer, 2010.
- [3181] D. Kauchak, J. Smarr, and C. Elkan. Sources of success for boosted wrapper induction. *Journal of Machine Learning Research*, 5:499–527, 2004.
- [3182] E. Kaufmann, T. Bonald, and M. Lelarge. A spectral algorithm with additive clustering for the recovery of overlapping communities in networks. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*. Springer, 2016.
- [3183] E. Kaufmann, O. Cappé, and A. Garivier. On the complexity of A/B testing. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 461–481, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3184] E. Kaufmann and S. Kalyan Krishnan. Information complexity in bandit subset selection. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 228–251. JMLR.org, 2013.

- [3185] E. Kaufmann, N. Korda, and R. Munos. Thompson sampling: An asymptotically optimal finite-time analysis. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 199–213, Berlin, Heidelberg, New York, 2012. Springer.
- [3186] S. Kaufmann and F. Stephan. Resource bounded next value and explanatory identification: Learning automata, patterns and polynomials on-line. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 263–274, New York, NY, 1997. ACM Press.
- [3187] S. Kaufmann and F. Stephan. Robust learning with infinite additional information. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 316–330. Springer, 1997.
- [3188] S. Kaufmann and F. Stephan. Robust learning with infinite additional information. *Theoret. Comput. Sci.*, 259(1-2):427–454, 2001.
- [3189] H. Kautz and B. Selman. Speeding inference by acquiring new concepts. In *Proc. of AAAI-92*, San Jose, July 1992.
- [3190] B. Kavšek, N. Lavrač, and A. Ferligoj. Consensus decision trees: Using consensus hierarchical clustering for data relabelling and reduction. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 251–262. Springer, 2001.
- [3191] A. Kawachi. Circuit lower bounds from learning-theoretic approaches. *Theoret. Comput. Sci.*, 733:83–98, 2018. Special Issue on Learning Theory and Complexity.
- [3192] M. Kawanabe and K.-R. Müller. Estimating functions for blind separation when sources have variance dependencies. *Journal of Machine Learning Research*, 6:453–482, 2005.
- [3193] H. Kawano and M. Kawahara. Extended association algorithm based on ROC analysis for visual information navigator. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 640–649. Springer, 2001.
- [3194] C. Kaynak and E. Alpaydin. MultiStage cascading of multiple classifiers: One man’s noise is another man’s data. In *Proc. 17th International Conf. on Machine Learning*, pages 455–462. Morgan Kaufmann, San Francisco, CA, 2000.
- [3195] D. Kazakov and S. Manandhar. Unsupervised learning of word segmentation rules with genetic algorithms and inductive logic programming. *Machine Learning*, 43(1/2):121–162, 2001.
- [3196] R. Kazman. Simulating the child’s acquisition of the lexicon and syntax – experiences with *babel*. *Machine Learning*, 16(1/2):87–120, 1994.
- [3197] M. Kearns. Thoughts on hypothesis boosting. (Unpublished), Dec. 1988.
- [3198] M. Kearns. *The Computational Complexity of Machine Learning*. PhD thesis, Harvard University Center for Research in Computing Technology, May 1989. Technical Report TR-13-89. Also published by MIT Press as an ACM Distinguished Dissertation.

- [3199] M. Kearns. Efficient noise-tolerant learning from statistical queries. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 392–401, New York, NY, 1993. ACM Press.
- [3200] M. Kearns and M. Li. Learning in the presence of malicious errors. *SIAM J. Comput.*, 22:807–837, 1993.
- [3201] M. Kearns, M. Li, L. Pitt, and L. Valiant. On the learnability of Boolean formulae. In *Proc. 19th Annu. ACM Sympos. Theory Comput.*, pages 285–294, New York, NY, 1987. ACM Press.
- [3202] M. Kearns, M. Li, L. Pitt, and L. Valiant. Recent results on Boolean concept learning. In *Proc. of the Fourth International Workshop on Machine Learning*, pages 337–352, Los Altos, California, June 1987. Morgan Kaufmann.
- [3203] M. Kearns, M. Li, and L. Valiant. Learning Boolean formulas. *J. ACM*, 41(6):1298–1328, 1994.
- [3204] M. Kearns and Y. Mansour. A fast, bottom-up decision tree pruning algorithm with near-optimal generalization. In *Proc. 15th International Conf. on Machine Learning*, pages 269–277. Morgan Kaufmann, San Francisco, CA, 1998.
- [3205] M. Kearns and Y. Mansour. On the boosting ability of top-down decision tree learning algorithms. *J. of Comput. Syst. Sci.*, 58(1):109–128, 1999.
- [3206] M. Kearns, Y. Mansour, and A. Y. Ng. A sparse sampling algorithm for near-optimal planning in large Markov decision processes. *Machine Learning*, 49(2-3):193–208, 2002.
- [3207] M. Kearns, Y. Mansour, A. Y. Ng, and D. Ron. An experimental and theoretical comparison of model selection methods. *Machine Learning*, 27(1):7–50, 1997.
- [3208] M. Kearns, Y. Mansour, D. Ron, R. Rubinfeld, R. Schapire, and L. Sellie. On the learnability of discrete distributions. In *Proc. of Twenty-sixth ACM Symposium on Theory of Computing*, pages 273–282, 1994.
- [3209] M. Kearns and L. Pitt. A polynomial-time algorithm for learning  $k$ -variable pattern languages from examples. In *Proceedings of the Second Annual Workshop on Computational Learning Theory, Santa Cruz, CA*, pages 57–71, San Mateo, CA, 1989. Morgan Kaufmann.
- [3210] M. Kearns and D. Ron. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 152–162, New York, NY, 1997. ACM Press.
- [3211] M. Kearns and D. Ron. Testing problems with sub-learning sample complexity. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 268–279, New York, NY, 1998. ACM Press.
- [3212] M. Kearns and D. Ron. Testing problems with sublearning sample complexity. *J. Comput. Syst. Sci.*, 61(3):428–456, Dec. 2000.
- [3213] M. Kearns and H. S. Seung. Learning from a population of hypotheses. *Machine Learning*, 18(2/3):255–276, 1995. Special issue for COLT '93.
- [3214] M. Kearns and S. Singh. Near-optimal reinforcement learning in polynomial time. In *Proc. 15th International Conf. on Machine Learning*, pages 260–268. Morgan Kaufmann, San Francisco, CA, 1998.

- [3215] M. Kearns and S. Singh. Bias-variance error bounds for temporal difference updates. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 142–147. Morgan Kaufmann, San Francisco, 2000.
- [3216] M. Kearns and S. Singh. Near-optimal reinforcement learning in polynomial time. *Machine Learning*, 49(2-3):209–232, 2002.
- [3217] M. Kearns and L. Valiant. Cryptographic limitations on learning Boolean formulae and finite automata. *J. ACM*, 41(1):67–95, 1994.
- [3218] M. Kearns and L. G. Valiant. Learning Boolean formulae or finite automata is as hard as factoring. Technical Report TR 14-88, Harvard University Aiken Computation Laboratory, 1988.
- [3219] M. Kearns and L. G. Valiant. Cryptographic limitations on learning Boolean formulae and finite automata. In *Proc. of the 21st Symposium on Theory of Computing*, pages 433–444, New York, NY, 1989. ACM Press.
- [3220] M. Kearns and J. Wortman. Learning from collective behavior. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 99–110. Omnipress, 2008.
- [3221] M. Kearns and Z. S. Wu. Predicting with distributions. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1214–1241, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3222] M. J. Kearns and R. E. Schapire. Efficient distribution-free learning of probabilistic concepts. *J. of Comput. Syst. Sci.*, 48(3):464–497, 1994.
- [3223] M. J. Kearns and R. E. Schapire. Efficient distribution-free learning of probabilistic concepts. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 10, pages 289–329. MIT Press, 1994. Earlier version appeared in FOCS90.
- [3224] M. J. Kearns, R. E. Schapire, and L. M. Sellie. Toward efficient agnostic learning. *Machine Learning*, 17(2/3):115–142, 1994. Special issue for COLT 1992.
- [3225] M. J. Kearns and U. V. Vazirani. *An Introduction to Computational Learning Theory*. The MIT Press, Cambridge, Massachusetts, 1994.
- [3226] J. Kececioglu, M. Li, and J. Tromp. Inferring a DNA sequence from erroneous copies. *Theoret. Comput. Sci.*, 185(1):3–13, 1997. Special issue for ALT '95.
- [3227] S. S. Keerthi and D. DeCoste. A modified finite newton method for fast solution of large scale linear SVMs. *Journal of Machine Learning Research*, 6:341–361, 2005.
- [3228] S. S. Keerthi and E. G. Gilbert. Convergence of a generalized SMO algorithm for SVM classifier design. *Machine Learning*, 46(1-3):351–360, 2002.
- [3229] B. Kégl. Robust regression by boosting the median. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 258–272. Springer, 2003.

- [3230] B. Kégl. Open problem: A (missing) boosting-type convergence result for ADABOOST.MH with factorized multi-class classifiers. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1268–1275, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3231] B. Kégl, T. Linder, and G. Lugosi. Data-dependent margin-based generalization bounds for classification. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 368–384. Springer, 2001.
- [3232] D. A. Keim, F. Mansmann, D. Oelke, and H. Ziegler. Visual analytics: Combining automated discovery with interactive visualizations. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, page 2, Berlin, Oct. 2008. Springer.
- [3233] K. Kelly and C. Glymour. On convergence to the truth and nothing but the truth. Technical report, Carnegie Mellon University, Laboratory for Computational Linguistics, CMU-LCL-87-4, 1987.
- [3234] K. T. Kelly. Induction from the general to the more general. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 334–348, San Mateo, CA, 1989. Morgan Kaufmann.
- [3235] K. T. Kelly and C. Glymour. Inductive inference from theory laden data. *Journal of Philosophical Logic*, 21(4):391–444, 1992.
- [3236] D. Kempe and F. McSherry. A decentralized algorithm for spectral analysis. *J. Comput. Syst. Sci.*, 74(1):70–83, 2008.
- [3237] J. O. Kephart and G. J. Tesauro. Pseudo-convergent Q-learning by competitive pricebots. In *Proc. 17th International Conf. on Machine Learning*, pages 463–470. Morgan Kaufmann, San Francisco, CA, 2000.
- [3238] F. N. Kepler, S. L. S. Mergen, and C. Z. Billa. Simple variable length N-grams for probabilistic automata learning. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 254–258, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [3239] M. Kerber. Some aspects of analogy in mathematical reasoning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 231–242. Springer-Verlag, 1989.
- [3240] K. L. Kermanidis. Extracting shallow paraphrasing schemata from modern greek text using statistical significance testing and supervised learning. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 297–300. Springer, 2010.
- [3241] C. Kermorvant and P. Dupont. Stochastic grammatical inference with multinomial tests. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 149–160. Springer, 2002.

- [3242] C. Kermorvant and C. de la Higuera. Learning languages with help. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 161–173. Springer, 2002.
- [3243] V. Kešelj. Grammar model and grammar induction in the system NL PAGE. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 57–66. Springer, 1998.
- [3244] D. Keysers, W. Macherey, J. Dahmen, and H. Ney. Learning of variability for invariant statistical pattern recognition. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 263–275. Springer, 2001.
- [3245] A. Khaleghi and D. Ryabko. Nonparametric multiple change point estimation in highly dependent time series. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 382–396. Berlin, Heidelberg, New York, 2013. Springer.
- [3246] A. Khaleghi and D. Ryabko. Nonparametric multiple change point estimation in highly dependent time series. *Theoret. Comput. Sci.*, 620:119–133, 2016. Special Issue: Algorithmic Learning Theory.
- [3247] A. Khalili and A. Tacchella. Learning nondeterministic mealy machines. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 109–123. Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [3248] R. Khardon. On using the Fourier transform to learn disjoint DNF. *Information Processing Letters*, 49(5):219–222, Mar. 1994.
- [3249] R. Khardon. Learning first order universal Horn expressions. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 154–165. New York, NY, 1998. ACM Press.
- [3250] R. Khardon. Learning function-free Horn expressions. *Machine Learning*, 37(3):241–275, 1999.
- [3251] R. Khardon. Learning range restricted Horn expressions. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 111–125. Springer, 1999.
- [3252] R. Khardon. Learning to take actions. *Machine Learning*, 35(1):57–90, 1999.
- [3253] R. Khardon. Learning Horn expressions with LOGAN-H. In *Proc. 17th International Conf. on Machine Learning*, pages 471–478. Morgan Kaufmann, San Francisco, CA, 2000.
- [3254] R. Khardon and M. Arias. The subsumption lattice and query learning. *J. Comput. Syst. Sci.*, 72(1):72–94, 2006.
- [3255] R. Khardon and D. Roth. Learning to reason. *J. ACM*, 44(5):697–725, 1997.

- [3256] R. Khairon and D. Roth. Learning to reason with a restricted view. *Machine Learning*, 35(2):95–116, 1999.
- [3257] R. Khairon, D. Roth, and R. Servedio. Efficiency versus convergence of Boolean kernels for online learning algorithms. In *Advances in Neural Information Processing Systems 14*, pages 423–430. MIT Press, 2001.
- [3258] R. Khairon and R. A. Servedio. Maximum margin algorithms with Boolean kernels. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 87–101. Springer, 2003.
- [3259] M. Kharitonov. Cryptographic lower bounds for learnability of Boolean functions on the uniform distribution. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 29–36, New York, NY, 1992. ACM Press.
- [3260] M. Kharitonov. Cryptographic hardness of distribution-specific learning. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 372–381, New York, NY, 1993. ACM Press.
- [3261] M. Kharitonov. Cryptographic lower bounds for learnability of Boolean functions on the uniform distribution. *J. of Comput. Syst. Sci.*, 50(3):600–610, 1995.
- [3262] S. Khot and A. K. Ponnuswami. Minimizing wide range regret with time selection functions. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 81–86. Omnipress, 2008.
- [3263] T. Kida, T. Matsumoto, Y. Shibata, M. Takeda, A. Shinohara, and S. Arikawa. Collage system: a unifying framework for compressed pattern matching. *Theoret. Comput. Sci.*, 298(1):253–272, 2003.
- [3264] J.-U. Kietz. Some computational lower bounds for the computational complexity of inductive logic programming. In *1993 European Conference of Machine Learning*, Vienna, Austria, 1993.
- [3265] J.-U. Kietz. Some lower bounds for the computational complexity of inductive logic programming. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 115–123. Springer-Verlag, 1993.
- [3266] J.-U. Kietz and M. Lübke. An efficient subsumption algorithm for inductive logic programming. In *Proc. 11th International Conference on Machine Learning*, pages 130–138. Morgan Kaufmann, 1994.
- [3267] J.-U. Kietz and K. Morik. A polynomial approach to the constructive induction of structural knowledge. *Machine Learning*, 14:193–217, 1994.
- [3268] D. Kifer, A. Smith, and A. Thakurta. Private convex empirical risk minimization and high-dimensional regression. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 25.1–25.40, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [3269] B. Kijirikul, M. Numao, and M. Shimura. Discrimination-based constructive induction of logic programs. In *Proc. Tenth National Conference on Artificial Intelligence*, pages 44–49. AAAI Press, 1992.



- [3270] B. Kijssirikul and S. Sinthupinyo. Approximate ILP rules by backpropagation neural network: A result on thai character recognition. In *Inductive Logic Programming, 9th International Workshop, ILP-99, Bled, Slovenia, June 24-27, 1999, Proceedings*, volume 1634 of *Lecture Notes in Artificial Intelligence*, pages 162–173. Springer, 1999.
- [3271] B. Kijssirikul, S. Sinthupinyo, and K. Chongkasemwongse. Approximate match of rules using backpropagation neural networks. *Machine Learning*, 44(3):273–299, 2001.
- [3272] F. Kilander and C. G. Jansson. COBBIT - A control procedure for COBWEB in the presence of concept drift. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 244–261. Springer-Verlag, 1993.
- [3273] J. Kilian, K. J. Lang, and B. A. Pearlmutter. Playing the matching-shoulders lob-pass game with logarithmic regret. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 159–164, New York, NY, 1994. ACM Press.
- [3274] J. Kilian and H. Siegelmann. On the power of sigmoid neural networks. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 137–143, New York, NY, 1993. ACM Press.
- [3275] P. Kilpeläinen, H. Mannila, and E. Ukkonen. MDL learning of unions of simple pattern languages from positive examples. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 252–260. Springer, 1995.
- [3276] D. Kim and J. Lee. Handling continuous-valued attributes in decision tree with neural network modelling. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 211–219. Springer, 2000.
- [3277] H. Kim, P. Howland, and H. Park. Dimension reduction in text classification with support vector machines. *Journal of Machine Learning Research*, 6:37–53, 2005.
- [3278] J. H. Kim. *CONVINCE: A Conversational Inference Consolidation Engine*. PhD thesis, University of California, Los Angeles, 1983.
- [3279] J. H. Kim and J. R. Roche. On the optimal capacity of binary neural networks: rigorous combinatorial approaches. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 240–249, New York, NY, 1995. ACM Press.
- [3280] J. H. Kim and J. R. Roche. Covering cubes by random half cubes, with applications to binary neural networks. *J. of Comput. Syst. Sci.*, 56(2):223–252, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [3281] W. M. Kim. Learning by smoothing: a morphological approach. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 43–57, San Mateo, CA, 1991. Morgan Kaufmann.
- [3282] Z. W. Kim and R. Nevatia. Learning Bayesian networks for diverse and varying numbers of evidence sets. In *Proc. 17th International Conf. on Machine Learning*, pages 479–486. Morgan Kaufmann, San Francisco, CA, 2000.
- [3283] D. Kimber and P. M. Long. The learning complexity of smooth functions of a single variable. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 153–159, New York, NY, 1992. ACM Press.

- [3284] H. Kimoto and T. Iwadera. Learning effect of a dynamic thesaurus in associated information retrieval. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 47–60. IOS Press, 1992.
- [3285] H. Kimura and S. Kobayashi. An analysis of actor/critic algorithms using eligibility traces: reinforcement learning with imperfect value functions. In *Proc. 15th International Conf. on Machine Learning*, pages 278–286. Morgan Kaufmann, San Francisco, CA, 1998.
- [3286] H. Kimura and S. Kobayashi. Efficient non-linear control by combining Q-learning with local linear controllers. In *Proc. 16th International Conf. on Machine Learning*, pages 210–219. Morgan Kaufmann, San Francisco, CA, 1999.
- [3287] H. Kimura, K. Miyazaki, and S. Kobayashi. Reinforcement learning in POMDPs with function approximation. In *Proc. 14th International Conference on Machine Learning*, pages 152–160. Morgan Kaufmann, 1997.
- [3288] H. Kimura, M. Yamamura, and S. Kobayashi. Reinforcement learning by stochastic hill climbing on discounted reward. In *Proc. 12th International Conference on Machine Learning*, pages 295–303. Morgan Kaufmann, 1995.
- [3289] S. Kimura, A. Togashi, and N. Shiratori. Synthesis algorithm for recursive processes by  $\mu$ -calculus. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 379–394. Springer-Verlag, 1994.
- [3290] S. Kimura, A. Togashi, and N. Shiratori. Inductive synthesis of recursive processes from logical properties. *Inform. Comput.*, 163(2):257–284, Dec. 2000.
- [3291] E. Kinber. Some models of inductive syntactical synthesis from sample computations. In *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 213–252. Springer, Berlin, 1991.
- [3292] E. Kinber. Learning a class of regular expressions via restricted subset queries. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 232–243. Springer-Verlag, 1992.
- [3293] E. Kinber. Monotonicity versus efficiency for learning languages from texts. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 395–406. Springer-Verlag, 1994.
- [3294] E. Kinber. On learning regular expressions and patterns via membership and correction queries. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 125–138. Springer, 2008.
- [3295] E. Kinber. Learning regular expressions from representative examples and membership queries. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 94–108. Springer, 2010.

- [3296] E. Kinber, C. Papazian, C. Smith, and R. Wiehagen. On the intrinsic complexity of learning recursive functions. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 257–266, New York, NY, 1999. ACM Press.
- [3297] E. Kinber, C. H. Smith, M. Velauthapillai, and R. Wiehagen. On learning multiple concepts in parallel. *J. of Comput. Syst. Sci.*, 50(1):41–52, 1995. Earlier version in 6th COLT, 1993.
- [3298] E. Kinber and F. Stephan. Language learning from texts: mind changes, limited memory and monotonicity. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 182–189, New York, NY, 1995. ACM Press.
- [3299] E. Kinber and F. Stephan. Language learning from texts: Mindchanges, limited memory, and monotonicity. *Inform. Comput.*, 123(2):224–241, 1995.
- [3300] E. Kinber and T. Zeugmann. Monte-carlo inference and its relations to reliable frequency identification. In *Fundamentals of Computation Theory, International Conference FCT '89, Szeged, Hungary, August 1989, Proceedings*, volume 380 of *Lecture Notes in Computer Science*, pages 257–266. Springer-Verlag, 1989.
- [3301] E. Kinber and T. Zeugmann. One-sided error probabilistic inductive inference and reliable frequency identification. *Inform. Comput.*, 92(2):253–284, 1991.
- [3302] E. B. Kinber. On the synthesis in the limit of almost minimal Gödel numbers (in russian). In *Theory of Algorithms and Programs*, volume 1, pages 221–223. Latvian State University, Riga, 1974.
- [3303] E. B. Kinber. On comparison of limit identification and limit standardization of general recursive functions. *Uch. zap. Latv. univ.*, 233:45–56, 1975.
- [3304] E. B. Kinber. On a theory of inductive inference. In *Proc. of the 1977 International FCT-Conference*, volume 56 of *Lecture Notes in Computer Science*, pages 435–440. Springer-Verlag, 1977.
- [3305] E. B. Kinber. On limit identification of minimal Gödel numbers for functions from enumerable classes. *Theory of Algorithms and Programs, Latvian State University, Riga*, 3:35–56, 1977.
- [3306] E. B. Kinber. A note on limit identification of c-minimal indices. *Elektronische Informationsverarbeitung und Kybernetik*, 19:459–463, 1983.
- [3307] E. B. Kinber. Inductive syntactical synthesis of programs from sample computations. In *Proc. 13th Internat. Symposium on Mathematical Foundations of Computer Science*, volume 324 of *Lecture Notes in Computer Science*, pages 73–84, 1988.
- [3308] E. B. Kinber. Some problems of learning with an oracle. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 178–186, San Mateo, CA, 1990. Morgan Kaufmann.
- [3309] E. B. Kinber, W. I. Gasarch, T. Zeugmann, M. G. Pleszkoch, and C. H. Smith. Learning via queries with teams and anomalies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 327–337, San Mateo, CA, 1990. Morgan Kaufmann.
- [3310] E. B. Kinber and T. Zeugmann. Inductive inference of almost everywhere correct programs by reliably working strategies. *Elektronische Informationsverarbeitung und Kybernetik*, 21(3):91–100, 1985.

- [3311] E. B. Kimber and T. Zeugmann. Refined query inference. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 148–160. Springer-Verlag, 1989.
- [3312] R. D. King. The robot scientist project. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 12–12. Springer, Oct. 2005.
- [3313] P. Kisilev, M. Zibulevsky, and Y. Y. Zeevi. A multiscale framework for blind separation of linearly mixed signals. *Journal of Machine Learning Research*, 4:1339–1363, 2003.
- [3314] G. Kitagawa. Signal extraction and knowledge discovery based on statistical modeling. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 3–14. Springer, 2003.
- [3315] G. Kitagawa. Signal extraction and knowledge discovery based on statistical modeling. *Theoret. Comput. Sci.*, 364(1):132–142, 2006. Special issue for ALT 2003.
- [3316] G. Kitagawa, T. Higuchi, and F. N. Kondo. Smoothness prior approach to explore mean structure in large-scale time series. *Theoret. Comput. Sci.*, 292(2):431–446, 2003.
- [3317] G. Kitagawa, T. Takanami, A. Kuwano, Y. Murai, and H. Shimamura. Extraction of signal from high dimensional time series: Analysis of ocean bottom seismograph data. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 449–458. Springer, 2001.
- [3318] M. Kitsuregawa. Challenge for info-plosion. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, page 12, Berlin, Oct. 2007. Springer.
- [3319] J. Kivinen. Reliable and useful learning. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 365–380, San Mateo, CA, 1989. Morgan Kaufmann.
- [3320] J. Kivinen. Guest editor’s introduction. *Machine Learning*, 47(2/3):131–132, 2002.
- [3321] J. Kivinen, H. Mannila, and E. Ukkonen. Learning hierarchical rule sets. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 37–44, New York, NY, 1992. ACM Press.
- [3322] J. Kivinen, H. Mannila, and E. Ukkonen. Learning rules with local exceptions. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 35–46, Oxford, 1994. Oxford University Press.
- [3323] J. Kivinen and R. H. Sloan, editors. *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, Berlin, 2002. Springer.
- [3324] J. Kivinen, A. J. Smola, and R. C. Williamson. Large margin classification for moving targets. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 113–127. Springer, 2002.

- [3325] J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors. *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, Berlin, Heidelberg, New York, Oct. 2011. Springer.
- [3326] J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann. Editors' introduction. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 1–13, Berlin, Heidelberg, New York, 2011. Springer.
- [3327] J. Kivinen, C. Szepesvári, and T. Zeugmann. Guest editors' introduction. *Theoret. Comput. Sci.*, 519:1–3, 2014. Special Issue: Algorithmic Learning Theory.
- [3328] J. Kivinen and M. Warmuth. Using experts for predicting continuous outcomes. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 109–120, Oxford, 1994. Oxford University Press.
- [3329] J. Kivinen and M. K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. Technical Report UCSC-CRL-94-16, University of California, Santa Cruz, Computer Research Laboratory, June 1994. Revised December 7, 1995. An extended abstract to appeared in the STOC 95, pp. 209-218.
- [3330] J. Kivinen and M. K. Warmuth. Additive versus exponentiated gradient updates for linear prediction. In *Proc. 27th Annual ACM Symposium on Theory of Computing*, pages 209–218, New York, NY, 1995. ACM Press.
- [3331] J. Kivinen and M. K. Warmuth. The perceptron algorithm vs. Winnow: linear vs. logarithmic mistake bounds when few input variables are relevant. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 289–296, New York, NY, 1995. ACM Press.
- [3332] J. Kivinen and M. K. Warmuth. Averaging expert predictions. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 153–167. Springer, 1999.
- [3333] J. Kivinen and M. K. Warmuth. Boosting as entropy projection. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 134–144, New York, NY, 1999. ACM Press.
- [3334] J. Kivinen and M. K. Warmuth. Relative loss bounds for multidimensional regression problems. *Machine Learning*, 45(3):301–329, 2001.
- [3335] D. Klahr. Children, adults, and machines as discovery systems. *Machine Learning*, 14(3):313–320, 1994.
- [3336] N. Klasner and H. U. Simon. From noise-free to noise-tolerant and from on-line to batch learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 250–257, New York, NY, 1995. ACM Press.
- [3337] A. Klautau, N. Jevtić, and A. Orlitsky. On nearest-neighbor error-correcting output codes with application to all-pairs multiclass support vector machines. *Journal of Machine Learning Research*, 4:1–15, 2003.

- [3338] D. Klein. Unsupervised learning for natural language processing. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 5–6. Omnipress, 2008.
- [3339] J. Kleinberg and M. Sandler. Using mixture models for collaborative filtering. *J. Comput. Syst. Sci.*, 74(1):49–69, 2008.
- [3340] R. D. Kleinberg, A. Niculescu-Mizil, and Y. Sharma. Regret bounds for sleeping experts and bandits. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 425–436. Omnipress, 2008.
- [3341] M. Kleindessner and U. Luxburg. Uniqueness of ordinal embedding. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 40–67, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3342] M. Klenin. A practical approach for evaluating generalization performance. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 16, pages 279–300. MIT Press, 1997.
- [3343] R. Klette. Unterklassen in der Familie NUM aller effektiv numerierbaren Mengen von einstelligen allgemein rekursiven Funktionen. *Kompliziertheit von Lern und Erkennungsprozessen*, 1:62–66, 1975.
- [3344] R. Klette. Indexmengen und Erkennung rekursiver Funktionen. *Zeitschr. f. math. Logik und Grundlagen d. Math. Bd.*, 22:231–238, 1976.
- [3345] R. Klette. Recognizing algorithms for recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 12:227–243, 1976.
- [3346] R. Klette. A few results on the complexity of classes of identifiable recursive function sets. *Sektion Mathematik report 4 (M) 77, Friedrich Schiller Universitat, Jena*, pages 1–16, 1977.
- [3347] R. Klette and B. Goetze. Limes-rekursive Funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 11:586–589, 1975.
- [3348] R. Klette and R. Wiehagen. Research in the theory of inductive inference by GDR mathematicians – A survey. *Inform. Sci.*, 22:149–169, 1980.
- [3349] V. Klingspor, K. J. Morik, and A. D. Rieger. Learning concepts from sensor data of a mobile robot. *Machine Learning*, 23:305–332, 1996.
- [3350] R. Klinkenberg and T. Joachims. Detecting concept drift with support vector machines. In *Proc. 17th International Conf. on Machine Learning*, pages 487–494. Morgan Kaufmann, San Francisco, CA, 2000.
- [3351] A. Klivans, R. O’Donnell, and R. A. Servedio. Learning intersections and thresholds of halfspaces. In *43rd Annual Symposium on Foundations of Computer Science, 6-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 177–186. IEEE Computer Society, 2002.
- [3352] A. Klivans and R. Servedio. Learning DNF in time  $2^{O(n^{1/3})}$ . *J. Comput. Syst. Sci.*, 68(2):303–318, 2004.
- [3353] A. Klivans and R. A. Servedio. Learning DNF in time  $2^{O(n^{1/3})}$ . In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 258–265. ACM Press, 2001.

- [3354] A. R. Klivans, R. O'Donnell, and R. A. Servedio. Learning intersections and thresholds of halfspaces. *J. Comput. Syst. Sci.*, 68(4):808–840, 2004.
- [3355] A. R. Klivans and R. A. Servedio. Boosting and hard-core set construction. *Machine Learning*, 51(3):217–238, 2003.
- [3356] A. R. Klivans and R. A. Servedio. Learning intersections of halfspaces with a margin. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 348–362. Springer, 2004.
- [3357] A. R. Klivans and R. A. Servedio. Perceptron-like performance for intersections of halfspaces. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 639–640. Springer, 2004.
- [3358] A. R. Klivans and R. A. Servedio. Toward attribute efficient learning of decision lists and parities. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 224–238. Springer, 2004.
- [3359] A. R. Klivans and R. A. Servedio. Learning intersections of halfspaces with a margin. *J. Comput. Syst. Sci.*, 74(1):35–48, 2008.
- [3360] A. R. Klivans and A. A. Sherstov. Improved lower bounds for learning intersections of halfspaces. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 335–349, Berlin, 2006. Springer.
- [3361] A. R. Klivans and A. A. Sherstov. A lower bound for agnostically learning disjunctions. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 409–423, Berlin, 2007. Springer.
- [3362] A. R. Klivans and A. A. Sherstov. Unconditional lower bounds for learning intersections of halfspaces. *Machine Learning*, 69(2-3):97–114, 2007.
- [3363] A. R. Klivans and A. Shpilka. Learning arithmetic circuits via partial derivatives. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 463–476. Springer, 2003.
- [3364] B. Knobe and K. Knobe. A method for inferring context-free grammars. *Inform. Control*, 31:129–146, 1976.
- [3365] T. Knuutila. How to invent characterizable inference methods for regular languages. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 209–222. Springer-Verlag, 1993.
- [3366] T. Knuutila. Inductive inference from positive data: from heuristic to characterizing methods. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 22–47. Springer, 1996.

- [3367] T. Knuutila and M. Steinby. The inference of tree languages from finite samples: an algebraic approach. *Theoret. Comput. Sci.*, 129(2):337–367, 1994.
- [3368] K. Ko. On the complexity of learning minimum time-bounded Turing machines. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 82–96, San Mateo, CA, 1990. Morgan Kaufmann.
- [3369] K.-I. Ko, A. Marron, and W.-G. Tzeng. Learning string patterns and tree patterns from examples. In *Machine Learning: Proceedings of the Seventh International Conference on Machine Learning*, pages 384–391. Morgan Kaufmann, 1990.
- [3370] I. Kobayashi, K. Furukawa, T. Ozaki, and M. Imai. A computational model for children’s language acquisition using inductive logic programming. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 140–155. Springer, 2001.
- [3371] S. Kobayashi. Learning strategies using decision lists. In *Algorithmic Learning Theory, 4th International Workshop, ALT ’93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 370–383. Springer-Verlag, 1993.
- [3372] S. Kobayashi. Iterated transductions and efficient learning from positive data: A unifying view. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 157–170. Springer, 2000.
- [3373] S. Kobayashi and T. Yokomori. Learning concatenations of locally testable languages from positive data. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 407–422. Springer-Verlag, 1994.
- [3374] S. Kobayashi and T. Yokomori. On approximately identifying concept classes in the limit. In *Algorithmic Learning Theory, 6th International Workshop, ALT ’95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 298–312. Springer, 1995.
- [3375] S. Kobayashi and T. Yokomori. Families of noncounting languages and their learnability from positive data. *Int. J. Found. Comput. Sci.*, 7(4):309–328, 1996.
- [3376] S. Kobayashi and T. Yokomori. Identifiability of subspaces and homomorphic images of zero-reversible languages. In *Algorithmic Learning Theory, 8th International Workshop, ALT ’97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 48–61. Springer, 1997.
- [3377] S. Kobayashi and T. Yokomori. Learning approximately regular languages with reversible languages. *Theoret. Comput. Sci.*, 174(1-2):251–257, 1997.
- [3378] J. Köbler and W. Lindner. Oracles in  $\Sigma_2^P$  are sufficient for exact learning. In *Algorithmic Learning Theory, 8th International Workshop, ALT ’97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 277–290. Springer, 1997.
- [3379] J. Köbler and W. Lindner. The complexity of learning concept classes with polynomial general dimension. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 149–163. Springer, 2002.



- [3380] J. Köbler and W. Lindner. A general dimension for approximately learning Boolean functions. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 139–148. Springer, 2002.
- [3381] J. Köbler and W. Lindner. The complexity of learning concept classes with polynomial general dimension. *Theoret. Comput. Sci.*, 350(1):49–62, 2006. Special issue for ALT 2002.
- [3382] S. Kocabas. Conflict resolution as discovery in particle physics. *Machine Learning*, 6:277–309, 1991.
- [3383] Y. Kodratoff. A class of functions synthesized from a finite number of examples and a LISP program scheme. *Internat. J. Comput. Info. Sci.*, 8:489–521, 1979.
- [3384] Y. Kodratoff. Technical and scientific issues of KDD (or: Is KDD a science?). In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 261–265. Springer, 1995.
- [3385] Y. Kodratoff and J. Fargues. A sane algorithm for the synthesis of LISP functions from example problems. In *Proc. of AISB/GI Conference on Artificial Intelligence*, pages 169–175. AISB and GI, 1978.
- [3386] Y. Kodratoff, M. Franova, and D. Partridge. Why and how program synthesis? In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 45–59. Springer-Verlag, 1989.
- [3387] Y. Kodratoff and R. Michalski, editors. *Machine Learning: An Artificial Intelligence Approach*, volume III. Morgan Kaufmann, Los Altos, California, 1990.
- [3388] S. Koenig and R. G. Simmons. The effect of representation and knowledge on goal-directed exploration with reinforcement-learning algorithms. *Machine Learning*, 22:227–250, 1996.
- [3389] S. Koenig and R. G. Simmons. Passive distance learning for robot navigation. In *Proc. 13th International Conference on Machine Learning*, pages 266–274. Morgan Kaufmann, 1996.
- [3390] S. Koenig and Y. Smirnov. Graph learning with a nearest neighbor approach. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 19–28, New York, NY, 1996. ACM Press.
- [3391] Y. Koga, E. Hirowatari, and S. Arikawa. Explanation-based reuse of Prolog programs. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 149–160. Springer-Verlag, 1994.
- [3392] R. Kohavi and S. Benson. Research note on decision lists. *Machine Learning*, 13:131–134, 1993.
- [3393] R. Kohavi and G. H. John. Automatic parameter selection by minimizing estimated error. In *Proc. 12th International Conference on Machine Learning*, pages 304–312. Morgan Kaufmann, 1995.
- [3394] R. Kohavi and C. Kunz. Option decision trees with majority votes. In *Proc. 14th International Conference on Machine Learning*, pages 161–169. Morgan Kaufmann, 1997.

- [3395] R. Kohavi and D. H. Wolpert. Bias plus variance decomposition for zero-one loss functions. In *Proc. 13th International Conference on Machine Learning*, pages 275–283. Morgan Kaufmann, 1996.
- [3396] P. Koiran. Efficient learning of continuous neural networks. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 348–355, New York, NY, 1994. ACM Press.
- [3397] P. Koiran and E. D. Sontag. Vapnik-Chervonenkis dimension of recurrent neural networks. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 223–237. Springer, 1997.
- [3398] M. Koivisto. Parent assignment is hard for the MDL, AIC, and NML costs. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 289–303, Berlin, 2006. Springer.
- [3399] M. Koivisto, T. Kivioja, H. Mannila, P. Rastas, and E. Ukkonen. Hidden markov modelling techniques for haplotype analysis. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 37–52. Springer, 2004.
- [3400] M. Koivisto and K. Sood. Exact Bayesian structure discovery in Bayesian networks. *Journal of Machine Learning Research*, 5:549–573, 2004.
- [3401] M. M. Kokar. Determining arguments of invariant functional descriptions. *Machine Learning*, 1(4):403–422, 1986.
- [3402] D. Koller and R. Fratkin. Using learning for approximation in stochastic processes. In *Proc. 15th International Conf. on Machine Learning*, pages 287–295. Morgan Kaufmann, San Francisco, CA, 1998.
- [3403] D. Koller and M. Sahami. Toward optimal feature selection. In *Proc. 13th International Conference on Machine Learning*, pages 284–292. Morgan Kaufmann, 1996.
- [3404] D. Koller and M. Sahami. Hierarchically classifying documents using very few words. In *Proc. 14th International Conference on Machine Learning*, pages 170–178. Morgan Kaufmann, 1997.
- [3405] D. Kollmar and D. H. Hellmann. Feature selection for a real-world learning task. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 157–172. Springer, 2001.
- [3406] V. Koltchinskii. Entropy bounds for restricted convex hulls. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 741–742. Springer, 2003.
- [3407] V. Koltchinskii and O. Beznosova. Exponential convergence rates in classification. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 295–307. Springer, 2005.

- [3408] V. Koltchinskii, D. Panchenko, and S. Andonova. Generalization bounds for voting classifiers based on sparsity and clustering. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 492–505. Springer, 2003.
- [3409] V. Koltchinskii, D. Panchenko, and F. Lozano. Further explanation of the effectiveness of voting methods: The game between margins and weights. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 241–255. Springer, 2001.
- [3410] V. Koltchinskii and M. Yuan. Sparse recovery in large ensembles of kernel machines online learning and bandits. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 229–238. Omnipress, 2008.
- [3411] P. Komarek and A. Moore. A dynamic adaptation of AD-trees for efficient machine learning on large data sets. In *Proc. 17th International Conf. on Machine Learning*, pages 495–502. Morgan Kaufmann, San Francisco, CA, 2000.
- [3412] J. Komiyama, J. Honda, H. Kashima, and H. Nakagawa. Regret lower bound and optimal algorithm in dueling bandit problem. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1141–1154, Paris, France, 03–06 July 2015. PMLR.
- [3413] J. Kómlós and R. Paturi. Convergence results in the Hopfield model. Technical Report Technical Report Number CS-104, Department of Computer Science and Engineering, University of California, San Diego, Sept. 1987.
- [3414] A. Konagaya. A stochastic approach to genetic information processing. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 25–36. Springer-Verlag, 1993.
- [3415] E. B. Kong and T. G. Dietterich. Error-correcting output coding corrects bias and variance. In *Proc. 12th International Conference on Machine Learning*, pages 313–321. Morgan Kaufmann, 1995.
- [3416] I. Kononenko and I. Bratko. Information-based evaluation criterion for classifier’s performance. *Machine Learning*, 6:67–80, 1991.
- [3417] L. Kontorovich, C. Cortes, and M. Mohri. Learning linearly separable languages. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 288–303. Springer, Oct. 2006.
- [3418] L. A. Kontorovich, C. Cortes, and M. Mohri. Kernel methods for learning languages. *Theoret. Comput. Sci.*, 405(3):223–236, 2008. Special issue for ALT 2006.
- [3419] W. M. Koolen and S. de Rooij. Combining expert advice efficiently. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 275–286. Omnipress, 2008.

- [3420] W. M. Koolen and S. de Rooij. Switching investments. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 239–254, Berlin, Heidelberg, New York, 2010. Springer.
- [3421] W. M. Koolen and T. V. Erven. Second-order quantile methods for experts and combinatorial games. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1155–1175, Paris, France, 03–06 July 2015. PMLR.
- [3422] W. M. Koolen, J. Nie, and M. Warmuth. Learning a set of directions. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 851–866. JMLR.org, 2013.
- [3423] W. M. Koolen and S. de Rooij. Switching investments. *Theoret. Comput. Sci.*, 473:61–76, 2013. Special Issue on Learning Theory.
- [3424] W. M. Koolen and V. Vovk. Buy low, sell high. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 335–349, Berlin, Heidelberg, New York, 2012. Springer.
- [3425] W. M. Koolen and V. Vovk. Buy low, sell high. *Theoret. Comput. Sci.*, 558:144–158, 2014. Special Issue: Algorithmic Learning Theory.
- [3426] W. M. Koolen, M. K. Warmuth, and D. Adamskiy. Open problem: Online sabotaged shortest path. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1764–1766, Paris, France, 03–06 July 2015. PMLR.
- [3427] M. Koppel. Learning to predict non-deterministically generated strings. *Machine Learning*, 7:85–99, 1991.
- [3428] M. Koppel, A. M. Segre, and R. Feldman. Getting the most from flawed theories. In *Proc. 11th International Conference on Machine Learning*, pages 139–147. Morgan Kaufmann, 1994.
- [3429] K. B. Korb, L. R. Hope, and M. J. Hughes. The evaluation of predictive learners: Some theoretical and empirical results. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 276–287. Springer, 2001.
- [3430] T. Koren. Open problem: Fast stochastic exp-concave optimization. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1073–1075. JMLR.org, 2013.
- [3431] T. Koren, R. Livni, and Y. Mansour. Bandits with movement costs and adaptive pricing. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1242–1268, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [3432] T. Koshiha. Typed pattern languages and their learnability. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 367–379. Springer, 1995.
- [3433] T. Koshiha, E. Mäkinen, and Y. Takada. Learning strongly deterministic even linear languages from positive examples. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 41–54. Springer, 1995.
- [3434] T. Koshiha, E. Mäkinen, and Y. Takada. Learning deterministic even linear languages from positive examples. *Theoret. Comput. Sci.*, 185(1):63–79, 1997. Special issue for ALT '95.
- [3435] W. Kotłowski. On minimaxity of follow the leader strategy in the stochastic setting. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 261–275. Springer, 2016.
- [3436] W. Kotłowski. Scale-invariant unconstrained online learning. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 412–433, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3437] W. Kotłowski and P. Grünwald. Maximum likelihood vs. sequential normalized maximum likelihood in on-line density estimation. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 457–476, Budapest, Hungary, 09–11 June 2011. PMLR.
- [3438] W. Kotłowski, W. M. Koolen, and A. Malek. Online isotonic regression. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1165–1189, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [3439] W. Kotłowski and M. K. Warmuth. Minimax algorithm for learning rotations. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 821–824, Budapest, Hungary, 09–11 June 2011. PMLR.
- [3440] T. Kötzing. Iterative learning from positive data and counters. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 40–54, Berlin, Heidelberg, New York, 2011. Springer.
- [3441] T. Kötzing. Iterative learning from positive data and counters. *Theoret. Comput. Sci.*, 519:155–169, 2014. Special Issue: Algorithmic Learning Theory.
- [3442] T. Kötzing. A solution to Wiehagen’s thesis. In E. W. Mayr and N. Portier, editors, *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5–8, 2014, Lyon, France*, volume 25 of *LIPICs*, pages 494–505. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014.
- [3443] T. Kötzing. A solution to Wiehagens thesis. *Theory of Computing Systems*, 60(3):498–520, 2017.

- [3444] T. Kötzing and R. Palenta. A map of update constraints in inductive inference. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 40–54, Berlin, Heidelberg, New York, 2014. Springer.
- [3445] T. Kötzing and R. Palenta. A map of update constraints in inductive inference. *Theoret. Comput. Sci.*, 650:4–24, 2016. Special Issue: Algorithmic Learning Theory.
- [3446] T. Kötzing, M. Schirneck, and K. Seidel. Normal forms in semantic language identification. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 493–516, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3447] A. Kowalczyk. Dense shattering and teaching dimensions for differentiable families. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 143–151, New York, NY, 1997. ACM Press.
- [3448] A. Kowalczyk. Continuity of performance metrics for thin feature maps. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 343–357, Berlin, Oct. 2007. Springer.
- [3449] A. Kowalczyk and O. Chapelle. An analysis of the anti-learning phenomenon for the class symmetric polyhedron. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 78–91. Springer, Oct. 2005.
- [3450] K. Koyama and T. Lai. An interactive knowledge transfer model and analysis of Mastermind game. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 196–206. IOS Press, 1992.
- [3451] J. R. Koza. Evolution of a subsumption architecture that performs a wall following task for an autonomous mobile robot. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 19, pages 321–346. MIT Press, 1994.
- [3452] D. Kozen, J. Palsberg, and M. I. Schwartzbach. Efficient inference of partial types. *J. of Comput. Syst. Sci.*, 49(2):306–324, 1994.
- [3453] S. Kpotufe and S. Dasgupta. A tree-based regressor that adapts to intrinsic dimension. *J. Comput. Syst. Sci.*, 78(5):1496–1515, 2012.
- [3454] S. Kpotufe, R. Urner, and S. Ben-David. Hierarchical label queries with data-dependent partitions. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1176–1189, Paris, France, 03–06 July 2015. PMLR.
- [3455] H. P. Kramer. A note on the self-consistency of definitions of generalization and inductive inference. *J. ACM*, 9(2):280–281, 1962.
- [3456] S. Kramer and L. De Raedt. Feature construction with version spaces for biochemical applications. In *Proc. 18th International Conf. on Machine Learning*, pages 258–265. Morgan Kaufmann, San Francisco, CA, 2001.

- [3457] M. Krause. On the computational power of Boolean decision lists. In *STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings*, volume 2285 of *Lecture Notes in Computer Science*, pages 372–383. Springer, 2002.
- [3458] K. Krawiec. On the use of pairwise comparison of hypotheses in evolutionary learning applied to learning from visual examples. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 307–321. Springer, 2001.
- [3459] K. Krawiec. Pairwise comparison of hypotheses in evolutionary learning. In *Proc. 18th International Conf. on Machine Learning*, pages 266–273. Morgan Kaufmann, San Francisco, CA, 2001.
- [3460] A. Krieger, C. Long, and A. Wyner. Boosting noisy data. In *Proc. 18th International Conf. on Machine Learning*, pages 274–281. Morgan Kaufmann, San Francisco, CA, 2001.
- [3461] H.-U. Krieger. A corpus-driven context-free approximation of head-driven phrase structure grammar. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 199–210. Springer, 2004.
- [3462] M. R. K. Krishna Rao. Incremental learning of logic programs. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 95–109. Springer, 1995.
- [3463] M. R. K. Krishna Rao. A framework for incremental learning of logic programs. *Theoret. Comput. Sci.*, 185(1):193–213, 1997. Special issue for ALT '95.
- [3464] M. R. K. Krishna Rao. Some classes of prolog programs inferable from positive data. *Theoret. Comput. Sci.*, 241(1-2):211–234, 2000. Special issue for ALT '96.
- [3465] M. R. K. Krishna Rao. A class of prolog programs with non-linear outputs inferable from positive data. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 312–326. Springer, Oct. 2005.
- [3466] M. R. K. Krishna Rao, D. Kapur, and R. K. Shyamasundar. Transformational methodology for proving termination of logic programs. *The Journal of Logic Programming*, 34(1):1–41, 1998.
- [3467] M. R. K. Krishna Rao and A. Sattar. Learning from entailment of logic programs with local variables. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 143–157. Springer, 1998.
- [3468] M. R. K. Krishna Rao and A. Sattar. Polynomial-time learnability of logic programs with local variables from entailment. *Theoret. Comput. Sci.*, 268(2):179–198, 2001. Special issue for ALT '98.
- [3469] P. Krishnan, P. M. Long, and J. S. Vitter. Learning to make rent-to-buy decisions with systems applications. In *Proc. 12th International Conference on Machine Learning*, pages 322–330. Morgan Kaufmann, 1995.

- [3470] P. Krishnan and R. Vijay. On the limits of proper learnability of subclasses of DNF formulas. *Machine Learning*, 25:237–263, 1996.
- [3471] A. Krogh, M. Brown, I. S. Mian, K. Sjölander, and D. Haussler. Hidden Markov models in computational biology: Applications to protein modeling. *Journal of Molecular Biology*, 235:1501–1531, Feb. 1994.
- [3472] A. Krogh, I. S. Mian, and D. Haussler. A hidden Markov model that finds genes in *e. coli* DNA. *Nucleic Acids Research*, 22:4768–4778, 1994.
- [3473] S. Krovvidy and W. G. Wee. Wastewater treatment systems from case-based reasoning. *Machine Learning*, 10:341–363, 1993.
- [3474] R. Kruse and C. Borgelt. Data mining with graphical models. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 22–22. Springer, 2002.
- [3475] C. P. Kruskal. The chromatic number of the plane: The bounded case. *J. Comput. Syst. Sci.*, 74(4):598–627, 2008.
- [3476] C. P. Kruskal. The orchard visibility problem and some variants. *J. Comput. Syst. Sci.*, 74(4):587–597, 2008.
- [3477] A. Krzyżak. Nonlinear function learning and classification using optimal radial basis function networks. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 217–225. Springer, 2001.
- [3478] M. Kubat. Second tier for decision trees. In *Proc. 13th International Conference on Machine Learning*, pages 293–301. Morgan Kaufmann, 1996.
- [3479] M. Kubat and M. Cooperson, Jr. Voting nearest-neighbor subclassifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 503–510. Morgan Kaufmann, San Francisco, CA, 2000.
- [3480] M. Kubat, D. Flotzinger, and G. Pfurtscheller. Discovering patterns in EEG-signals: Comparative study of a few methods. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 366–371. Springer-Verlag, 1993.
- [3481] M. Kubat, R. C. Holte, and S. Matwin. Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, 30:195–215, 1998.
- [3482] M. Kubat and S. Matwin. Addressing the curse of imbalanced training sets: one-sided selection. In *Proc. 14th International Conference on Machine Learning*, pages 179–186. Morgan Kaufmann, 1997.
- [3483] L. Kucera, A. Marchetti-Spaccamela, and M. Protasi. On the learnability of DNF formulae. In *Proc. of the 15th International Colloquium on Automata, Languages, and Programming*, volume 317 of *Lecture Notes in Computer Science*, pages 347–361. Springer-Verlag, 1988.
- [3484] L. Kucera, A. Marchettispaccamela, and M. Protasi. On learning monotone DNF formulae under uniform distributions. *Inform. Comput.*, 110(1):84–95, 1994.
- [3485] I. Kucevalovs, O. Krasts, R. Freivalds, and T. Zeugmann. On the influence of technology on learning processes. *Parallel Processing Letters*, 24(2):63–79, 2014.



- [3486] Y. Kudoh, M. Haraguchi, and Y. Okubo. Data abstractions for decision tree induction. *Theoret. Comput. Sci.*, 292(2):387–416, 2003.
- [3487] P. Kugel. Induction, pure and simple. *Inform. Control*, 35:276–336, 1977.
- [3488] A. Kuh, T. Petsche, and R. L. Rivest. Learning time-varying concepts. In *Advances in Neural Information Processing Systems 3*, pages 183–189. Morgan Kaufmann, 1991.
- [3489] T. Kuh, T. Petsche, and R. Rivest. Learning time varying concepts. In *NIPS 3*. Morgan Kaufmann, 1990.
- [3490] T. Kuh, T. Petsche, and R. Rivest. Mistake bounds of incremental learners when concepts drift with applications to feedforward networks. In *NIPS 4*. Morgan Kaufmann, 1991.
- [3491] T. P. A. Kuh and R. L. Rivest. Incrementally learning time-varying half-planes. In *Advances in Neural Information Processing Systems 4*, pages 920–927. Morgan Kaufmann, 1992.
- [3492] C. Kuhlmann. On teaching and learning intersection-closed concept classes. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 168–182. Springer, 1999.
- [3493] C. Kuhlmann. Hardness results for general two-layer neural networks. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 275–285. Morgan Kaufmann, San Francisco, 2000.
- [3494] J. Kujala and T. Elomaa. On following the perturbed leader in the bandit setting. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 371–385. Springer, Oct. 2005.
- [3495] J. Kujala and T. Elomaa. Following the perturbed leader to gamble at multi-armed bandits. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 166–180, Berlin, Oct. 2007. Springer.
- [3496] S. Kulkarni. On metric entripy, Vapnik-Chervonenkis dimension, and learnability for a class of distributions. Technical Report LIDS-P-1910, Center for Intelligent Control Systems, MIT, 1989.
- [3497] S. Kulkarni and O. Zeitouni. On probably correct classification of concepts. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 111–116, New York, NY, 1993. ACM Press.
- [3498] S. R. Kulkarni. *Problems of computational and information complexity in machine vision and learning*. PhD thesis, MIT EECS, 1991. (Also CICS report CICS-TH-298.).
- [3499] S. R. Kulkarni, S. K. Mitter, and J. N. Tsitsiklis. Active learning using arbitrary binary valued queries. *Machine Learning*, 11:23–35, 1993.
- [3500] S. R. Kulkarni, S. K. Mitter, J. N. Tsitsiklis, and O. Zeitouni. PAC learning with generalized samples and an application to stochastic geometry. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 172–179, New York, NY, 1992. ACM Press.

- [3501] S. J. Kumaar, P. J. Abisha, and D. G. Thomas. Learning subclasses of parallel communicating grammar systems. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 301–304. Springer, 2010.
- [3502] M. Kummer. The strength of noninclusions for teams of finite learners. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 268–277, New York, NY, 1994. ACM Press.
- [3503] M. Kummer. A learning theoretic characterization of classes of recursive functions. *Information Processing Letters*, 54:205–211, 1995.
- [3504] M. Kummer and M. Ott. Learning branches and learning to win closed games. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 280–291, New York, NY, 1996. ACM Press.
- [3505] M. Kummer and F. Stephan. Inclusion problems in parallel learning and games. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 287–298, New York, NY, 1994. ACM Press.
- [3506] M. Kummer and F. Stephan. Inclusion problems in parallel learning and games. *J. of Comput. Syst. Sci.*, 52(3):403–420, 1996.
- [3507] M. Kummer and F. Stephan. On the structure of the degrees of inferability. *J. of Comput. Syst. Sci.*, 52(2):214–238, 1996. Earlier version in Sixth Annual Conference on Computational Learning Theory, 1993.
- [3508] J. Kun and L. Reyzin. Open problem: Learning quantum circuits with queries. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1767–1769, Paris, France, 03–06 July 2015. PMLR.
- [3509] L. I. Kuncheva and C. J. Whitaker. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine Learning*, 51(2):181–207, 2003.
- [3510] R. Kurai, S. ichi Minato, and T. Zeugmann. N-gram analysis based on zero-suppressed BDDs. Technical Report TCS-TR-A-06-16, Division of Computer Science, Hokkaido University, June 2006.
- [3511] K. Kurihara and T. Sato. Variational Bayesian grammar induction for natural language. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 84–96, Berlin, 2006. Springer.
- [3512] T. Kurita. Iterative weighted least squares algorithms for neural networks classifiers. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 77–86. Springer-Verlag, 1993.
- [3513] H. Kuroda, T. Katagiri, and Y. Kanada. Knowledge discovery in auto-tuning parallel numerical library. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 628–639. Springer, 2001.
- [3514] S. A. Kurtz and J. S. Royer. Prudence in language learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 143–156. Morgan Kaufmann, Inc., 1988.

- [3515] S. A. Kurtz and C. H. Smith. On the role of search for learning. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 303–311, San Mateo, CA, 1989. Morgan Kaufmann.
- [3516] S. A. Kurtz and C. H. Smith. A refutation of Barzdins’ conjecture. In *Analogical and Inductive Inference, International Workshop AII ’89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 171–176. Springer-Verlag, 1989.
- [3517] S. A. Kurtz, C. H. Smith, and R. Wiehagen. On the role of search for learning from examples. *J. of Experimental and Theoret. Artif. Intell.*, 13(1):25–43, 2001.
- [3518] S. Kushagra and S. Ben-David. Information preserving dimensionality reduction. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 239–253, Berlin, Heidelberg, New York, 2015. Springer.
- [3519] S. Kushagra, S. Samadi, and S. Ben-David. Finding meaningful cluster structure amidst background noise. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 339–354. Springer, 2016.
- [3520] E. Kushilevitz. A simple algorithm for learning  $O(\log n)$ -term DNF. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 266–269, New York, NY, 1996. ACM Press.
- [3521] E. Kushilevitz and Y. Mansour. Learning decision trees using the fourier spectrum. *SIAM J. Comput.*, 22(6):1331–1348, 1993. Earlier version appeared in STOC 1991.
- [3522] E. Kushilevitz and D. Roth. On learning visual concepts and DNF formulae. *Machine Learning*, 24:65–85, 1996.
- [3523] P. Kuusela and D. Ocone. Learning with side information: PAC learning bounds. *J. Comput. Syst. Sci.*, 68(3):521–545, 2004.
- [3524] D. Kuzmin and M. K. Warmuth. Optimum follow the leader algorithm. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 684–686. Springer, 2005.
- [3525] D. Kuzmin and M. K. Warmuth. Unlabeled compression schemes for maximum classes. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 591–605. Springer, 2005.
- [3526] V. Kuznetsov and M. Mohri. Generalization bounds for time series prediction with non-stationary processes. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 260–274, Berlin, Heidelberg, New York, 2014. Springer.
- [3527] V. Kuznetsov and M. Mohri. Time series prediction and online learning. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1190–1213, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [3528] W. Kwedlo and M. Krękowski. An evolutionary algorithm for cost-sensitive decision rule learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 288–299. Springer, 2001.
- [3529] S. Kwek. Learning disjunctions of features. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 401–415. Springer, 1997.
- [3530] S. Kwek. On a simple depth-first search strategy for exploring unknown graphs. In *Algorithms and Data Structures, 5th International Workshop, WADS '97, Halifax, Nova Scotia, Canada, August 6–8, 1997, Proceedings.*, volume 1272 of *Lecture Notes in Computer Science*, pages 345–353. Springer, 1997.
- [3531] S. Kwek and L. Pitt. PAC learning intersections of halfspaces with membership queries. *Algorithmica*, 22(1/2):53–75, 1998.
- [3532] S. S. Kwek. An apprentice learning model. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th–9th, 1999, Santa Cruz, California*, pages 63–74, New York, NY, 1999. ACM Press.
- [3533] S. S. Kwek. Learning intermediate concepts. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 151–166. Springer, 2001.
- [3534] J. Kwon, V. Perchet, and C. Vernade. Sparse stochastic bandits. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1269–1270, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3535] K. H. Kwon, K. Kang, and J. H. Oh. Generalization in partially connected layered neural networks. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 356–361, New York, NY, 1994. ACM Press.
- [3536] R. Kyng, A. Rao, S. Sachdeva, and D. A. Spielman. Algorithms for lipschitz learning on graphs. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1190–1223, Paris, France, 03–06 July 2015. PMLR.
- [3537] J. Lafferty. Additive models, boosting, and inference for generalized divergences. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th–9th, 1999, Santa Cruz, California*, pages 125–133, New York, NY, 1999. ACM Press.
- [3538] J. Lafferty and G. Lebanon. Diffusion kernels on statistical manifolds. *Journal of Machine Learning Research*, 6:129–163, 2005.
- [3539] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proc. 18th International Conf. on Machine Learning*, pages 282–289. Morgan Kaufmann, San Francisco, CA, 2001.
- [3540] J. Lafond. Low rank matrix completion with exponential family noise. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1224–1243, Paris, France, 03–06 July 2015. PMLR.

- [3541] M. G. Lagoudakis and M. L. Littman. Algorithm selection using reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 511–518. Morgan Kaufmann, San Francisco, CA, 2000.
- [3542] M. G. Lagoudakis and R. Parr. Least-squares policy iteration. *Journal of Machine Learning Research*, 4:1107–1149, 2003.
- [3543] H. Lähdesmäki, I. Shmulevich, and O. Yli-Harja. On learning gene regulatory networks under the Boolean network model. *Machine Learning*, 52(1-2):147–167, 2003.
- [3544] J. Laird, P. Rosenbloom, and A. Newell. Towards chunking as a general learning mechanism. In *Proc. AAAI-84*, pages 188–192, Aug. 1984.
- [3545] J. E. Laird, A. Newell, and P. S. Rosenbloom. SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33(1):1–64, Sept. 1987.
- [3546] J. E. Laird, P. S. Rosenbloom, and A. Newell. Chunking in soar: The anatomy of a general learning mechanism. *Machine Learning*, 1(1):11–46, 1986.
- [3547] P. Laird. Inductive inference by refinement. In *Proc. of AAAI-86*, pages 472–476. AAAI, 1986.
- [3548] P. Laird. Learning by making models. Technical Report RIA-88-4-12-0, NASA-Ames Research Center, 1988.
- [3549] P. Laird. A survey of computational learning theory. In *Formal Techniques in Artificial Intelligence: A Sourcebook*, pages 173–215. Elsevier Science, 1990.
- [3550] P. Laird. Identifying and using patterns in sequential data. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 1–18. Springer-Verlag, 1993.
- [3551] P. Laird and E. Gamble. Learning a probability distribution efficiently and reliably. Technical report, NASA Ames Research Center, Oct. 1988.
- [3552] P. Laird and E. Gamble. A ‘PAC’ algorithm for making feature maps. *Machine Learning*, 6:145–160, 1991.
- [3553] P. Laird, S. R., and P. Dunning. A model of sequence extrapolation. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 84–93, New York, NY, 1993. ACM Press.
- [3554] P. Laird and R. Saul. Discrete sequence prediction and its applications. *Machine Learning*, 15(1):43–68, 1994.
- [3555] P. D. Laird. Efficient unsupervised learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 297–311, San Mateo, CA, 1988. Morgan Kaufmann.
- [3556] P. D. Laird. Learning from good and bad data. In *Kluwer international series in engineering and computer science*. Kluwer Academic Publishers, Boston, 1988.
- [3557] B. Lambeau, C. Damas, and P. Dupont. State-merging DFA induction algorithms with mandatory merge constraints. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 139–153. Springer, 2008.

- [3558] E. Lamma, F. Riguzzi, and L. M. Pereira. Strategies in combined learning via logic programs. *Machine Learning*, 38(1/2):63–87, 2000.
- [3559] G. R. G. Lanckriet, N. Cristianini, P. Bartlett, L. E. Ghaoui, and M. I. Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, 5:27–72, 2004.
- [3560] G. R. G. Lanckriet, L. E. Ghaoui, C. Bhattacharyya, and M. I. Jordan. A robust minimax approach to classification. *Journal of Machine Learning Research*, 3:555–582, 2002.
- [3561] T. Lane and C. E. Brodley. Data reduction techniques for instance-based learning from human/computer interface data. In *Proc. 17th International Conf. on Machine Learning*, pages 519–526. Morgan Kaufmann, San Francisco, CA, 2000.
- [3562] T. Lane and C. E. Brodley. An empirical study of two approaches to sequence learning for anomaly detection. *Machine Learning*, 51(1):73–107, 2003.
- [3563] K. Lang. NewsWeeder: learning to filter netnews. In *Proc. 12th International Conference on Machine Learning*, pages 331–339. Morgan Kaufmann, 1995.
- [3564] K. J. Lang. Random DFA’s can be approximately learned from sparse uniform examples. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 45–52, New York, NY, 1992. ACM Press.
- [3565] K. J. Lang. Hill climbing beats genetic search on a Boolean circuit problem of Koza’s. In *Proc. 12th International Conference on Machine Learning*, pages 340–343. Morgan Kaufmann, 1995.
- [3566] K. J. Lang, B. A. Pearlmutter, and R. A. Price. Results of the Abbadingo one DFA learning competition and a new evidence-driven state merging algorithm. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 1–12. Springer, 1998.
- [3567] S. Lange. A decidability problem of Church-Rosser specifications for program synthesis. In *Analogical and Inductive Inference, International Workshop AII ’86, Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 105–124. Springer-Verlag, 1986.
- [3568] S. Lange. A program synthesis algorithm exemplified. In *Mathematical Methods of Specification and Synthesis of Software Systems ’85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 185–193. Springer-Verlag, 1986.
- [3569] S. Lange. Towards a set of inference rules for solving divergence in Knuth-Bendix completion. In *Analogical and Inductive Inference, International Workshop AII ’89, Reinhardsbrenn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 304–316. Springer-Verlag, 1989.
- [3570] S. Lange. A note on polynomial-time inference of  $k$ -variable pattern language. In *Non-monotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany, December 1990, Proceedings*, volume 543 of *Lecture Notes in Artificial Intelligence*, pages 178–183. Springer-Verlag, 1991.
- [3571] S. Lange. The representation of recursive languages and its impact on the efficiency of learning. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 256–267, New York, NY, 1994. ACM Press.

- [3572] S. Lange. *Algorithmic Learning of Recursive Languages*. Mensch & Buch Verlag, Berlin, 2000.
- [3573] S. Lange and G. Grieser. On variants of iterative learning. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 72–83. Springer, 1998.
- [3574] S. Lange and G. Grieser. On the strength of incremental learning. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 118–131. Springer, 1999.
- [3575] S. Lange and G. Grieser. On the power of incremental learning. *Theoret. Comput. Sci.*, 288(2):277–307, 2002. Special issue for ALT '99.
- [3576] S. Lange and G. Grieser. Variants of iterative learning. *Theoret. Comput. Sci.*, 292(2):359–376, 2003.
- [3577] S. Lange, G. Grieser, and K. P. Jantke. Extending elementary formal systems. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 332–347. Springer, 2001.
- [3578] S. Lange, G. Grieser, and K. P. Jantke. Advanced elementary formal systems. *Theoret. Comput. Sci.*, 298(1):51–70, 2003.
- [3579] S. Lange, G. Grieser, and T. Zeugmann. Inductive inference of approximations for recursive concepts. *Theoret. Comput. Sci.*, 348(1):15–40, 2005. Special issue for ALT 2000.
- [3580] S. Lange, S. E. Moelius III, and S. Zilles. Learning with temporary memory. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 449–463, Berlin, Oct. 2008. Springer.
- [3581] S. Lange and J. Nessel. Decision lists over regular patterns. *Theoret. Comput. Sci.*, 298(1):71–87, 2003.
- [3582] S. Lange, J. Nessel, and R. Wiehagen. Language learning from good examples. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 423–437. Springer-Verlag, 1994.
- [3583] S. Lange, J. Nessel, and R. Wiehagen. Learning recursive languages from good examples. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):27–52, 1998. Special issue for AII '94 and ALT '94.
- [3584] S. Lange and P. Watson. Machine discovery in the presence of incomplete or ambiguous data. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 438–452. Springer-Verlag, 1994.
- [3585] S. Lange and R. Wiehagen. Polynomial-time inference of arbitrary pattern languages. *New Generation Computing*, 8(4):361–370, 1991.

- [3586] S. Lange, R. Wiehagen, and T. Zeugmann. Learning by erasing. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 228–241. Springer, 1996.
- [3587] S. Lange and T. Zeugmann. Learning recursive languages with bounded mind changes. Technical Report 16/92, GOSLER-Report, FB Mathematik und Informatik, TH Leipzig, 1992.
- [3588] S. Lange and T. Zeugmann. On the power of monotonic language learning. Technical Report 5/92, GOSLER-Report, FB Mathematik und Informatik, TH Leipzig, 1992.
- [3589] S. Lange and T. Zeugmann. Types of monotonic language learning and their characterization. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 377–390, New York, NY, 1992. ACM Press.
- [3590] S. Lange and T. Zeugmann. A unifying approach to monotonic language learning on informant. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 244–259. Springer-Verlag, 1992.
- [3591] S. Lange and T. Zeugmann. Language learning in dependence on the space of hypotheses. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 127–136, New York, NY, 1993. ACM Press.
- [3592] S. Lange and T. Zeugmann. Language learning with a bounded number of mind changes. In *STACS 93, 10th Annual Symposium on Theoretical Aspects of Computer Science, Würzburg, Germany, February 1993, Proceedings*, volume 665 of *Lecture Notes in Computer Science*, pages 682–691. Springer-Verlag, 1993.
- [3593] S. Lange and T. Zeugmann. The learnability of recursive languages in dependence on the hypothesis space. Technical Report 20/93, GOSLER-Report, FB Mathematik und Informatik, TH Leipzig, 1993.
- [3594] S. Lange and T. Zeugmann. Learning recursive languages with a bounded number of mind changes. *Internat. Journal of Foundations of Computer Science*, 4(2):157–178, 1993.
- [3595] S. Lange and T. Zeugmann. Monotonic versus non-monotonic language learning. In G. Brewka, K. P. Jantke, and P. H. Schmitt, editors, *Nonmonotonic and Inductive Logic, Second International Workshop, Reinhardsbrunn Castle, Germany, December 1991*, volume 659 of *Lecture Notes in Artificial Intelligence*, pages 254–269, Berlin, Heidelberg, 1993. Springer-Verlag.
- [3596] S. Lange and T. Zeugmann. Characterization of language learning from informant under various monotonicity constraints. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):73–94, 1994. Special issue for AII '92.
- [3597] S. Lange and T. Zeugmann. Modeling incremental learning from positive data. Technical Report RIFIS-TR-CS-117, RIFIS, Kyushu University 33, 1995.
- [3598] S. Lange and T. Zeugmann. Refined incremental learning. In *Proc. 8th Australian Joint Conference on Artificial Intelligence - AI'95*, pages 147–154. World Scientific Publ. Co., 1995.
- [3599] S. Lange and T. Zeugmann. Trading monotonicity demands versus efficiency. *Bulletin of Informatics and Cybernetics*, 27(1):53–83, 1995.



- [3600] S. Lange and T. Zeugmann. Trading monotonicity demands versus mind changes. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 125–139. Springer, 1995.
- [3601] S. Lange and T. Zeugmann. Incremental learning from positive data. *J. of Comput. Syst. Sci.*, 53(1):88–103, 1996.
- [3602] S. Lange and T. Zeugmann. Set-driven and rearrangement-independent learning of recursive languages. *Math. Syst. Theory*, 29(6):599–634, 1996. Earlier version in Fifth ALT conf, 1994, Lecture Notes in AI 872.
- [3603] S. Lange, T. Zeugmann, and S. Kapur. Monotonic and dual monotonic language learning. *Theoret. Comput. Sci.*, 155(2):365–410, 1996.
- [3604] S. Lange, T. Zeugmann, and S. Zilles. Learning indexed families of recursive languages from positive data. Technical Report TCS-TR-A-07-31, Division of Computer Science, Hokkaido University, Oct. 2007.
- [3605] S. Lange, T. Zeugmann, and S. Zilles. Learning indexed families of recursive languages from positive data: A survey. *Theoret. Comput. Sci.*, 397(1-3):194–232, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [3606] S. Lange and S. Zilles. On the learnability of erasing pattern languages in the query model. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 129–143. Springer, 2003.
- [3607] S. Lange and S. Zilles. Comparison of query learning and Gold-style learning in dependence of the hypothesis space. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 99–113. Springer, 2004.
- [3608] S. Lange and S. Zilles. Formal language identification: query learning vs. Gold-style learning. *Information Processing Letters*, 91(6):285–292, 2004.
- [3609] S. Lange and S. Zilles. Replacing limit learners with equally powerful one-shot query learners. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 155–169. Springer, 2004.
- [3610] S. Lange and S. Zilles. Relations between Gold-style learning and query learning. *Inform. Comput.*, 203(2):211–237, 2005.
- [3611] J. Langford. The cross validation problem. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 687–688. Springer, 2005.
- [3612] J. Langford. Tutorial on practical prediction theory for classification. *Journal of Machine Learning Research*, 6:273–306, 2005.
- [3613] J. Langford and A. Beygelzimer. Sensitive error correcting output codes. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 158–172. Springer, 2005.

- [3614] J. Langford and A. Blum. Microchoice bounds and self bounding learning algorithms. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 209–214, New York, NY, 1999. ACM Press.
- [3615] J. Langford and A. Blum. Microchoice bounds and self bounding learning algorithms. *Machine Learning*, 51(2):165–179, 2003.
- [3616] J. Langford and D. McAllester. Computable shell decomposition bounds. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 25–34. Morgan Kaufmann, San Francisco, 2000.
- [3617] J. Langford and D. McAllester. Computable shell decomposition bounds. *Journal of Machine Learning Research*, 5:529–547, 2004.
- [3618] J. Langford, M. Seeger, and N. Megiddo. An improved predictive accuracy bound for averaging classifiers. In *Proc. 18th International Conf. on Machine Learning*, pages 290–297. Morgan Kaufmann, San Francisco, CA, 2001.
- [3619] P. Langley. Editorial: Human and machine learning. *Machine Learning*, 1(3):243–248, 1986.
- [3620] P. Langley. Editorial: The terminology of machine learning. *Machine Learning*, 1(2):141–144, 1986.
- [3621] P. Langley. A general theory of discrimination learning. Technical Report 86-18, University of California, Irvine, 1986.
- [3622] P. Langley. On machine learning. *Machine Learning*, 1(1):5–10, 1986.
- [3623] P. Langley. Machine learning and concept formation. *Machine Learning*, 2(2):99–102, 1987.
- [3624] P. Langley. Machine learning and grammar induction. *Machine Learning*, 2(1):5–8, 1987.
- [3625] P. Langley. Research papers in machine learning. *Machine Learning*, 2(3):195–198, 1987.
- [3626] P. Langley. Machine learning as an experimental science. *Machine Learning*, 3(1):5–8, 1988.
- [3627] P. Langley. Toward a unified science of machine learning. *Machine Learning*, 3:253–259, 1989.
- [3628] P. Langley. Advice to machine learning authors. *Machine Learning*, 5:233–237, 1990.
- [3629] P. Langley. Induction of recursive Bayesian classifiers. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 153–164. Springer-Verlag, 1993.
- [3630] P. Langley. Crafting papers on machine learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1207–1212. Morgan Kaufmann, San Francisco, CA, 2000.
- [3631] P. Langley and R. S. Michalski. Editorial: Machine learning and discovery. *Machine Learning*, 1(4):363–366, 1986.
- [3632] P. Langley and B. Nordhausen. A framework for empirical discovery. Technical Report 86-19, University of California, Irvine, 1986.
- [3633] P. Langley and K. Pflieger. Case-based acquisition of place knowledge. In *Proc. 12th International Conference on Machine Learning*, pages 344–352. Morgan Kaufmann, 1995.

- [3634] P. Langley, G. M. Provan, and P. Smyth. Learning with probabilistic representations. *Machine Learning*, 29(2-3):91–101, 1997.
- [3635] P. Langley and S. Sage. Scaling to domains with irrelevant features. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 4, pages 51–63. MIT Press, 1997.
- [3636] P. Langley and S. Sage. Tractable average-case analysis of naive Bayesian classifiers. In *Proc. 16th International Conf. on Machine Learning*, pages 220–228. Morgan Kaufmann, San Francisco, CA, 1999.
- [3637] P. Langley and S. Stromsten. Learning context-free grammars with a simplicity bias. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 220–228. Springer, 2000.
- [3638] H. Langseth and T. D. Nielsen. Fusion of domain knowledge with data for structural learning in object oriented domains. *Journal of Machine Learning Research*, 4:339–368, 2003.
- [3639] C. Lanquillon. Partially supervised text classification: Combining labeled and unlabeled documents using an EM-like scheme. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 229–237. Springer, 2000.
- [3640] A. S. Lapedes, E. W. Steeg, and R. M. Farber. Use of adaptive networks to define highly predictable protein secondary-structure classes. *Machine Learning*, 21:103–124, 1995.
- [3641] P. Laskov. Feasible direction decomposition algorithms for training support vector machines. *Machine Learning*, 46(1-3):315–349, 2002.
- [3642] R. H. Lathrop. *Efficient Methods for Massively Parallel Symbolic Induction: Algorithms and Implementation*. PhD thesis, MIT Department of Electrical Engineering and Computer Science, June 1990.
- [3643] R. H. Lathrop. On the learnability of the uncomputable. In *Proc. 13th International Conference on Machine Learning*, pages 302–309. Morgan Kaufmann, 1996.
- [3644] P. Latinne, M. Saerens, and C. Decaestecker. Adjusting the outputs of a classifier to new a priori probabilities may significantly improve classification accuracy: Evidence from a multi-class problem in remote sensing. In *Proc. 18th International Conf. on Machine Learning*, pages 298–305. Morgan Kaufmann, San Francisco, CA, 2001.
- [3645] M. Latourrette. Toward an explanatory similarity measure for nearest-neighbor classification. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 238–245. Springer, 2000.
- [3646] T. Lattimore. Regret analysis of the finite-horizon gittins index strategy for multi-armed bandits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1214–1245, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [3647] T. Lattimore, A. György, and C. Szepesvári. On learning the optimal waiting time. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 200–214, Berlin, Heidelberg, New York, 2014. Springer.
- [3648] T. Lattimore and M. Hutter. Asymptotically optimal agents. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 368–382, Berlin, Heidelberg, New York, 2011. Springer.
- [3649] T. Lattimore and M. Hutter. Time consistent discounting. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 383–397, Berlin, Heidelberg, New York, 2011. Springer.
- [3650] T. Lattimore and M. Hutter. PAC bounds for discounted MDPs. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 320–334, Berlin, Heidelberg, New York, 2012. Springer.
- [3651] T. Lattimore and M. Hutter. Bayesian reinforcement learning with exploration. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 170–184, Berlin, Heidelberg, New York, 2014. Springer.
- [3652] T. Lattimore and M. Hutter. General time consistent discounting. *Theoret. Comput. Sci.*, 519:140–154, 2014. Special Issue: Algorithmic Learning Theory.
- [3653] T. Lattimore and M. Hutter. Near-optimal PAC bounds for discounted MDPs. *Theoret. Comput. Sci.*, 558:125–143, 2014. Special Issue: Algorithmic Learning Theory.
- [3654] T. Lattimore, M. Hutter, and V. Gavane. Universal prediction of selected bits. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 262–276, Berlin, Heidelberg, New York, 2011. Springer.
- [3655] T. Lattimore, M. Hutter, and P. Sunehag. Concentration and confidence for discrete Bayesian sequence predictors. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 324–338, Berlin, Heidelberg, New York, 2013. Springer.
- [3656] T. Lau, P. Domingos, and D. S. Weld. Version space algebra and its application to programming by demonstration. In *Proc. 17th International Conf. on Machine Learning*, pages 527–534. Morgan Kaufmann, San Francisco, CA, 2000.
- [3657] T. Lau, S. A. Wolfman, P. Domingos, and D. S. Weld. Programming by demonstration using version space algebra. *Machine Learning*, 53(1-2):111–156, 2003.

- [3658] J. Laub and K.-R. Müller. Feature discovery in non-metric pairwise data. *Journal of Machine Learning Research*, 5:801–818, 2004.
- [3659] M. Lauer. A mixture approach to novelty detection using training data with outliers. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 300–311. Springer, 2001.
- [3660] M. Lauer and M. Riedmiller. An algorithm for distributed reinforcement learning in cooperative multi-agent systems. In *Proc. 17th International Conf. on Machine Learning*, pages 535–542. Morgan Kaufmann, San Francisco, CA, 2000.
- [3661] N. Lavrač and S. Džeroski. Background knowledge and declarative bias in inductive concept learning. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 51–71. Springer-Verlag, 1992.
- [3662] N. Lavrač and S. Džeroski. Weakening the language bias in LINUS. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):95–120, 1994. Special issue for AII '92.
- [3663] N. Lavrač and S. Džeroski. A reply to Pazzani’s book review of *inductive logic programming: techniques and applications*. *Machine Learning*, 23(1):109–111, 1996.
- [3664] N. Lavrač, S. Džeroski, and M. Numao. Inductive logic programming for relational knowledge discovery. *New Generation Computing*, 17(1):3–23, 1999.
- [3665] N. Lavrač, D. Gamberger, and P. Turney. Cost-sensitive feature reduction applied to a hybrid genetic algorithm. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 127–134. Springer, 1996.
- [3666] N. Lavrač, B. Kavčec, P. Flach, and L. Todorovski. Subgroup discovery with CN2-SD. *Journal of Machine Learning Research*, 5:153–188, 2004.
- [3667] M. H. Law and J. T. Kwok. Applying the Bayesian evidence framework to  $\nu$ -support vector regression. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 312–323. Springer, 2001.
- [3668] N. D. Lawrence and B. Schölkopf. Estimating a kernel Fisher discriminant in the presence of label noise. In *Proc. 18th International Conf. on Machine Learning*, pages 306–313. Morgan Kaufmann, San Francisco, CA, 2001.
- [3669] J. A. Laxminarayana, J. M. Sempere, and G. Nagaraja. Learning distinguishable linear grammars from positive data. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 279–280. Springer, 2004.
- [3670] A. Lazaric and R. Munos. Learning with stochastic inputs and adversarial outputs. *J. Comput. Syst. Sci.*, 78(5):1516–1537, 2012.
- [3671] E. G. Learned-Miller and J. W. F. III. ICA using spacings estimates of entropy. *Journal of Machine Learning Research*, 4:1271–1295, 2003.
- [3672] M. Lebowitz. Experiments with incremental concept formation: UNIMEM. *Machine Learning*, 2(2):103–138, 1987.

- [3673] G. Lecué. Optimal oracle inequality for aggregation of classifiers under low noise condition. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 364–378, Berlin, 2006. Springer.
- [3674] G. Lecué. Suboptimality of penalized empirical risk minimization in classification. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 142–156, Berlin, 2007. Springer.
- [3675] H. Lee, R. Ge, T. Ma, A. Risteski, and S. Arora. On the ability of neural nets to express distributions. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1271–1296, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3676] H. Lee, R. A. Servedio, and A. Wan. DNF are teachable in the average case. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 214–228, Berlin, 2006. Springer.
- [3677] H. K. H. Lee. A noninformative prior for neural networks. *Machine Learning*, 50(1-2):197–212, 2003.
- [3678] H. K. H. Lee and M. A. Clyde. Lossless online Bayesian bagging. *Journal of Machine Learning Research*, 5:143–151, 2004.
- [3679] J. D. Lee, M. Simchowitz, M. I. Jordan, and B. Recht. Gradient descent only converges to minimizers. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1246–1257, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [3680] K. Lee and S. Mahajan. A pattern classification approach to evaluation function learning. *Artificial Intelligence*, 36(1):1–26, Aug. 1988.
- [3681] M. D. Lee. A simple method for generating additive clustering models with limited complexity. *Machine Learning*, 49(1):39–58, 2002.
- [3682] T.-W. Lee, J.-F. Cardoso, E. Oja, and S. ichi Amari. Introduction to special issue on independent components analysis. *Journal of Machine Learning Research*, 4:1175–1176, 2003.
- [3683] W. S. Lee. Collaborative learning for recommender systems. In *Proc. 18th International Conf. on Machine Learning*, pages 314–321. Morgan Kaufmann, San Francisco, CA, 2001.
- [3684] W. S. Lee, P. L. Bartlett, and R. C. Williamson. Lower bounds on the VC-dimension of smoothly parametrized function classes. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 362–367, New York, NY, 1994. ACM Press.
- [3685] W. S. Lee, P. L. Bartlett, and R. C. Williamson. On efficient agnostic learning of linear combinations of basis functions. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 369–376, New York, NY, 1995. ACM Press.
- [3686] W. S. Lee, P. L. Bartlett, and R. C. Williamson. The importance of convexity in learning with squared loss. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 140–146, New York, NY, 1996. ACM Press.

- [3687] W. S. Lee and P. M. Long. A theoretical analysis of query selection for collaborative filtering. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 517–528. Springer, 2001.
- [3688] Y. Lee, B. G. Buchanan, and J. M. Aronis. Knowledge-based learning in exploratory science: learning rules to predict rodent carcinogenicity. *Machine Learning*, 30:217–240, 1998.
- [3689] Y. C. Lee, S. Qian, R. D. Jones, C. W. Barnes, G. W. Flake, M. K. O’Rourke, K. Lee, H. H. Chen, G. Z. Sun, Y. Q. Zhang, D. Chen, and C. L. Giles. Adaptive stochastic cellular automata: Theory. *Physica D*, 666, 1990.
- [3690] R. A. Legenstein and W. Maass. Neural circuits for pattern recognition with small total wire length. *Theoret. Comput. Sci.*, 287(1):239–249, 2002.
- [3691] S. Legg. Is there an elegant universal theory of prediction? In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 274–287. Springer, Oct. 2006.
- [3692] S. Legg, J. Poland, and T. Zeugmann. On the limits of learning with computational models. Technical Report TCS-TR-A-08-34, Division of Computer Science, Hokkaido University, Jan. 2008.
- [3693] J. Leike and M. Hutter. Indefinitely oscillating martingales. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 321–335, Berlin, Heidelberg, New York, 2014. Springer.
- [3694] J. Leike and M. Hutter. Bad universal priors and notions of optimality. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1244–1259, Paris, France, 03–06 July 2015. PMLR.
- [3695] J. Leike and M. Hutter. On the computability of Solomonoff induction and knowledge-seeking. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 364–378, Berlin, Heidelberg, New York, 2015. Springer.
- [3696] J. Leike and M. Hutter. Solomonoff induction violates nicod’s criterion. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 349–363, Berlin, Heidelberg, New York, 2015. Springer.
- [3697] J. Leike and M. Hutter. On the computability of Solomonoff induction and AIXI. *Theoret. Comput. Sci.*, 716:28–49, 2018. Special Issue on ALT 2015.
- [3698] M. Lelarge and L. Miolane. Fundamental limits of symmetric low-rank matrix estimation. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1297–1301, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [3699] Á. D. Lelkes and L. Reyzin. Interactive clustering of linear classes and cryptographic lower bounds. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 165–176, Berlin, Heidelberg, New York, 2015. Springer.
- [3700] A. Lemay, J. Niehren, and R. Gilleron. Learning  $n$ -ary node selecting tree transducers from completely annotated examples. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 253–267, Berlin, 2006. Springer.
- [3701] M. Lemmon. Prototype selection using competitive learning. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 12, pages 183–203. MIT Press, 1994.
- [3702] D. B. Lenat. When will machines learn? *Machine Learning*, 4:255–257, 1989.
- [3703] M. van Lent and J. Laird. Learning hierarchical performance knowledge by observation. In *Proc. 16th International Conf. on Machine Learning*, pages 229–238. Morgan Kaufmann, San Francisco, CA, 1999.
- [3704] E. Leopold and J. Kindermann. Text categorization with support vector machines. How to represent texts in input space? *Machine Learning*, 46(1-3):423–444, 2002.
- [3705] B. Leskes. The value of agreement, a new boosting algorithm. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27–30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 95–110. Springer, 2005.
- [3706] B. Leskes and L. Torenvliet. The value of agreement a new boosting algorithm. *J. Comput. Syst. Sci.*, 74(4):557–586, 2008.
- [3707] C. Leslie and R. Kuang. Fast kernels for inexact string matching. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 114–128. Springer, 2003.
- [3708] F. Letouzey, F. Denis, and R. Gilleron. Learning from positive and unlabeled examples. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 71–85. Springer, 2000.
- [3709] H. C. Leung. *The Use of Artificial Neural Networks for Phonetic Recognition*. PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, May 1989.
- [3710] G. Lever, F. Laviolette, and J. Shawe-Taylor. Distribution-dependent PAC-Bayes priors. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6–8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 119–133, Berlin, Heidelberg, New York, 2010. Springer.
- [3711] G. Lever, F. Laviolette, and J. Shawe-Taylor. Tighter PAC-Bayes bounds through distribution-dependent priors. *Theoret. Comput. Sci.*, 473:4–28, 2013. Special Issue on Learning Theory.



- [3712] E. Levin, N. Tishby, and S. Solla. A statistical approach to learning and generalization in neural networks. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*. Morgan Kaufmann, 1989.
- [3713] E. Levin, N. Tishby, and S. A. Solla. A statistical approach to learning and generalization in layered neural networks. *Proc. of the IEEE – Special issue on Neural Networks*, 1990.
- [3714] B. Levine. Derivatives of tree sets with applications to grammatical inference. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-3:285–293, 1981.
- [3715] B. Levine. The use of tree derivatives and a sample support parameter for inferring tree systems. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-4:25–34, 1982.
- [3716] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell System Technical Journal*, 62(4):1035–1074, Apr. 1983.
- [3717] D. D. Lewis. Challenges in machine learning for text classification. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, page 1, New York, NY, 1996. ACM Press.
- [3718] D. D. Lewis and J. Catlett. Heterogeneous uncertainty sampling for supervised learning. In *Proc. 11th International Conference on Machine Learning*, pages 148–156. Morgan Kaufmann, 1994.
- [3719] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li. RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*, 5:361–397, 2004.
- [3720] C. Li and G. Biswas. A Bayesian approach to temporal data clustering using hidden Markov models. In *Proc. 17th International Conf. on Machine Learning*, pages 543–550. Morgan Kaufmann, San Francisco, CA, 2000.
- [3721] C. Li, W. Jiang, and M. Tanner. General oracle inequalities for Gibbs posterior with application to ranking. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 512–521. JMLR.org, 2013.
- [3722] H. Li and F. Stephan. Splitting of learnable classes. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 109–121. Springer, 2010.
- [3723] J. Li and K. Ramamohanarao. The space of jumping emerging patterns and its incremental maintenance algorithms. In *Proc. 17th International Conf. on Machine Learning*, pages 551–558. Morgan Kaufmann, San Francisco, CA, 2000.
- [3724] J. Li and L. Schmidt. Robust and proper learning for mixtures of gaussians via systems of polynomial inequalities. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1302–1382, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3725] L. Li and O. Chapelle. Open problem: Regret bounds for thompson sampling. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 43.1–43.3, Edinburgh, Scotland, 25–27 June 2012. PMLR.

- [3726] L. K. Li. On computing decision regions with neural nets. *J. of Comput. Syst. Sci.*, 43(3):509–512, 1991.
- [3727] L. K. Li. Learning fixed point patterns by recurrent networks. *J. of Comput. Syst. Sci.*, 48(2):203–213, 1994.
- [3728] M. Li. Towards a DNA sequencing theory (learning a string). In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 125–134. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [3729] M. Li. Foreword. *Theoret. Comput. Sci.*, 261(1):1, 2001. Special issue for ALT '97.
- [3730] M. Li. Information distance and its extensions. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 39–39, Berlin, Heidelberg, New York, 2011. Springer.
- [3731] M. Li and A. Maruoka, editors. *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, Berlin/Heidelberg, Oct. 1997. Springer-Verlag.
- [3732] M. Li and U. Vazirani. On the learnability of finite automata. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 359–370, San Mateo, CA, 1988. Morgan Kaufmann.
- [3733] M. Li and P. Vitányi. Inductive reasoning and kolmogorov complexity. *J. Comput. Syst. Sci.*, 44:343–384, 1992.
- [3734] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Text and Monographs in Computer Science. Springer-Verlag, 1993.
- [3735] M. Li and P. M. B. Vitányi. A theory of learning simple concepts under simple distributions and average case complexity for the universal distribution. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 34–39. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [3736] M. Li and P. M. B. Vitányi. Learning simple concepts under simple distributions. *SIAM J. Comput.*, 20:911–935, 1991.
- [3737] P. Li, T. J. Hastie, and K. W. Church. Improving random projections using marginal information. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 635–649, Berlin, 2006. Springer.
- [3738] P. Li, T. J. Hastie, and K. W. Church. Nonlinear estimators and tail bounds for dimension reduction in  $l_1$  using cauchy random projections. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 514–529, Berlin, 2007. Springer.
- [3739] P. Li and C.-H. Zhang. A new algorithm for compressed counting with applications in shannon entropy estimation in dynamic data. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 477–496, Budapest, Hungary, 09–11 June 2011. PMLR.

- [3740] P. Li, C.-H. Zhang, and T. Zhang. Compressed counting meets compressed sensing. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1058–1077, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3741] Y. Li. Selective voting for perceptron-like online learning. In *Proc. 17th International Conf. on Machine Learning*, pages 559–566. Morgan Kaufmann, San Francisco, CA, 2000.
- [3742] Y. Li and P. M. Long. The relaxed online maximum margin algorithm. *Machine Learning*, 46(1-3):361–387, 2002.
- [3743] Y. Li, P. M. Long, and A. Srinivasan. Improved bounds on the sample complexity of learning. *J. Comput. Syst. Sci.*, 62(3):516–527, 2001.
- [3744] T. Liang, A. Rakhlin, and K. Sridharan. Learning with square loss: Localization through offset rademacher complexity. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1260–1285, Paris, France, 03–06 July 2015. PMLR.
- [3745] S. Lievesley and E. Atwell. ‘NAIL’: Artificial intelligence software for learning natural language. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 306–308. Springer, 2002.
- [3746] S. H. Lim and P. Auer. Autonomous exploration for navigating in MDPs. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 40.1–40.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [3747] T.-S. Lim, W.-Y. Loh, and Y.-S. Shih. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning*, 40(3):203–228, 2000.
- [3748] D. Lin. An information-theoretic definition of similarity. In *Proc. 15th International Conf. on Machine Learning*, pages 296–304. Morgan Kaufmann, San Francisco, CA, 1998.
- [3749] H.-T. Lin and L. Li. Large-margin thresholded ensembles for ordinal regression: Theory and practice. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 319–333. Springer, Oct. 2006.
- [3750] J. Lin and J. S. Vitter. Complexity issues in learning by neural nets. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 118–133, San Mateo, CA, 1989. Morgan Kaufmann.
- [3751] J.-H. Lin and J. S. Vitter. Complexity results on learning by neural networks. *Machine Learning*, 6:211–230, 1991.
- [3752] J.-H. Lin and J. S. Vitter. A theory for memory-based learning. *Machine Learning*, 17(2/3):143–167, 1994.
- [3753] L. Lin. Self-improving reactive agents: case studies of reinforcement learning frameworks. Technical Report CMU-CS-90-109, Carnegie Mellon Computer Science Department, Aug. 1990.

- [3754] Y. Lin, Y. Lee, and G. Wahba. Support vector machines for classification in nonstandard situations. *Machine Learning*, 46(1-3):191–202, 2002.
- [3755] Y. Lin, G. Wahba, H. Zhang, and Y. Lee. Statistical properties and adaptive tuning of support vector machines. *Machine Learning*, 48(1/3):115–136, 2002.
- [3756] L. ji Lin. Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine Learning*, 8:293–321, 1992.
- [3757] N. Lincoln and J. Skrzypek. Synergy of clustering multiple backpropagation networks. In *NIPS 2*. Morgan Kaufmann, 1989.
- [3758] R. Linder and S. J. Pöpl. How to automate neural net based learning. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 206–216. Springer, 2001.
- [3759] T. Lindgren and H. Boström. Classification with intersecting rules. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 395–402. Springer, 2002.
- [3760] R. Lindner. *Algorithmische Erkennung*. Dissertation B, Friedrich-Schiller-Universität, Jena, 1972.
- [3761] W. Lindner. Learning DNF by statistical and proper distance queries under the uniform distribution. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 198–210. Springer, Oct. 2005.
- [3762] W. Lindner, R. Schuler, and O. Watanabe. Resource-bounded measure and learnability. *Theory of Computing Systems*, 33(2):151–170, 2000.
- [3763] C. X. Ling and R. Buchal. Learning to control dynamic systems with automatic quantization. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 372–377. Springer-Verlag, 1993.
- [3764] C. X. Ling and H. Zhang. The representational power of discrete Bayesian networks. *Journal of Machine Learning Research*, 3:709–721, 2002.
- [3765] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 574–579. IEEE Computer Society Press, 1989.
- [3766] N. Linial, Y. Mansour, and R. L. Rivest. Results on learnability and the Vapnik-Chervonenkis dimension. *Inform. Comput.*, 90(1):33–49, Jan. 1991.
- [3767] R. P. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Magazine*, pages 4–22, Apr. 1987.
- [3768] M. Liquiere and J. Sallantin. Structural machine learning with Galois lattice and graphs. In *Proc. 15th International Conf. on Machine Learning*, pages 305–313. Morgan Kaufmann, San Francisco, CA, 1998.

- [3769] N. List. Convergence of a generalized gradient selection approach for the decomposition method. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 338–349. Springer, 2004.
- [3770] N. List. Generalized SMO-style decomposition algorithms. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 365–377, Berlin, 2007. Springer.
- [3771] N. List, D. Hush, C. Scovel, and I. Steinwart. Gaps in support vector optimization. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 336–348, Berlin, 2007. Springer.
- [3772] N. List and H. U. Simon. A general convergence theorem for the decomposition method. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 363–377. Springer, 2004.
- [3773] N. List and H. U. Simon. General polynomial time decomposition algorithms. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 308–322. Springer, 2005.
- [3774] N. List and H. U. Simon. General polynomial time decomposition algorithms. *Journal of Machine Learning Research*, 8:303–321, 2007.
- [3775] N. Littlestone. Learning in a layered network with many fixed-function hidden nodes. In *Proc. of the 1st Annual IEEE International Conference on Neural Networks*, volume II, pages 471–478, Washington, D.C., June 1987. IEEE Computer Society Press.
- [3776] N. Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2(4):285–318, 1988.
- [3777] N. Littlestone. From on-line to batch learning. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 269–284, San Mateo, CA, 1989. Morgan Kaufmann.
- [3778] N. Littlestone. *Mistake Bounds and Logarithmic Linear-threshold Learning Algorithms*. PhD thesis, Technical Report UCSC-CRL-89-11, University of California Santa Cruz, 1989.
- [3779] N. Littlestone. Redundant noisy attributes, attribute errors, and linear threshold learning using Winnow. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 147–156, San Mateo, CA, 1991. Morgan Kaufmann.
- [3780] N. Littlestone. Learning simple Boolean concepts. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 6–14. IOS Press, 1992.
- [3781] N. Littlestone. Comparing several linear-threshold learning algorithms on tasks involving superfluous attributes. In *Proc. 12th International Conference on Machine Learning*, pages 353–361. Morgan Kaufmann, 1995.
- [3782] N. Littlestone and P. Long. On-line learning with linear loss constraints. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 412–421, New York, NY, 1993. ACM Press.

- [3783] N. Littlestone, P. M. Long, and M. K. Warmuth. On-line learning of linear functions. *Journal of Computational Complexity*, 5:1–23, 1995. Was in STOC91.
- [3784] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Inform. Comput.*, 108(2):212–261, 1994.
- [3785] M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proc. 11th International Conference on Machine Learning*, pages 157–163. Morgan Kaufmann, 1994.
- [3786] M. L. Littman. Friend-or-foe Q-learning in general-sum games. In *Proc. 18th International Conf. on Machine Learning*, pages 322–328. Morgan Kaufmann, San Francisco, CA, 2001.
- [3787] M. L. Littman. Tutorial: Learning topics in game-theoretic decision making. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 1–1. Springer, 2003.
- [3788] M. L. Littman. Inducing partially observable markov decision processes. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 145–148, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [3789] M. L. Littman, A. R. Cassandra, and L. P. Kaelbling. Learning policies for partially observable environments: scaling up. In *Proc. 12th International Conference on Machine Learning*, pages 362–370. Morgan Kaufmann, 1995.
- [3790] M. L. Littman, F. Jiang, and G. A. Keim. Learning a language-independent representation for terms from a partially aligned corpus. In *Proc. 15th International Conf. on Machine Learning*, pages 314–322. Morgan Kaufmann, San Francisco, CA, 1998.
- [3791] M. L. Littman and C. Szepesvári. A generalized reinforcement-learning model: Convergence and applications. In *Proc. 13th International Conference on Machine Learning*, pages 310–318. Morgan Kaufmann, 1996.
- [3792] C.-Y. Liu and S. Bubeck. Most correlated arms identification. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 623–637, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3793] C.-Y. Liu and L. Li. On the prior sensitivity of Thompson sampling. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 321–336. Springer, 2016.
- [3794] H. Liu and R. Setiono. A probabilistic approach to feature selection – a filter solution. In *Proc. 13th International Conference on Machine Learning*, pages 319–327. Morgan Kaufmann, 1996.
- [3795] W. Z. Liu and A. P. White. The importance of attribute selection measures in decision tree induction. *Machine Learning*, 15(1):25–41, 1994.
- [3796] Y. Liu. Abnormal data points in the data set: an algorithm for robust neural net regression. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 8, pages 121–136. MIT Press, 1997.

- [3797] R. Livni and P. Simon. Honest compressions and their application to compression schemes. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 77–92. JMLR.org, 2013.
- [3798] F. F. Liza and M. Grześ. A spectral method that worked well in the SPiCe’16 competition. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 143–148, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [3799] X. Llorà and J. M. Garrell. Inducing partially-defined instances with evolutionary algorithms. In *Proc. 18th International Conf. on Machine Learning*, pages 337–344. Morgan Kaufmann, San Francisco, CA, 2001.
- [3800] A. Locatelli, A. Carpentier, and S. Kpotufe. Adaptivity to noise parameters in nonparametric active learning. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1383–1416, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3801] A. Locatelli, A. Carpentier, and S. Kpotufe. An adaptive strategy for active learning with smooth decision boundary. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 547–571. PMLR, 07–09 Apr. 2018.
- [3802] J. Loch and S. Singh. Using eligibility traces to find the best memoryless policy in partially observable Markov decision processes. In *Proc. 15th International Conf. on Machine Learning*, pages 323–331. Morgan Kaufmann, San Francisco, CA, 1998.
- [3803] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins. Text classification using string kernels. *Journal of Machine Learning Research*, 2:419–444, 2002.
- [3804] P.-L. Loh and S. Nowozin. Faster Hoeffding racing: Bernstein races via jackknife estimates. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 203–217, Berlin, Heidelberg, New York, 2013. Springer.
- [3805] P. Long and F. Stephan. Preface. *Theoret. Comput. Sci.*, 405(3):207–208, 2008. Special issue for ALT 2006.
- [3806] P. M. Long. *Towards a more comprehensive theory of learning in computers*. PhD thesis, UC Santa Cruz, 1992.
- [3807] P. M. Long. On the sample complexity of PAC learning half-spaces against the uniform distribution. *IEEE Transactions on Neural Networks*, 6(6):1556–1559, 1995.
- [3808] P. M. Long. Improved bounds about on-line learning of smooth functions of a single variable. In *Algorithmic Learning Theory, 7th International Workshop, ALT ’96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 26–36. Springer, 1996.
- [3809] P. M. Long. Guest editor’s introduction. *Machine Learning*, 27(1):5, 1997.
- [3810] P. M. Long. On-line evaluation and prediction using linear functions. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 21–31, New York, NY, 1997. ACM Press.

- [3811] P. M. Long. On the sample complexity of learning functions with bounded variation. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 126–133, New York, NY, 1998. ACM Press.
- [3812] P. M. Long. The complexity of learning according to two models of a drifting environment. *Machine Learning*, 37(3):337–354, 1999.
- [3813] P. M. Long. Improved bounds about on-line learning of smooth-functions of a single variable,. *Theoret. Comput. Sci.*, 241(1-2):25–35, 2000. Special issue for ALT '96.
- [3814] P. M. Long. On agnostic learning with  $\{0, *, 1\}$ -valued and real-valued hypotheses. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 289–302. Springer, 2001.
- [3815] P. M. Long. Introduction to the special issue on computational learning theory. *Journal of Machine Learning Research*, 3:361–362, 2002.
- [3816] P. M. Long. An upper bound on the sample complexity of PAC-learning halfspaces with respect to the uniform distribution. *Information Processing Letters*, 87(5):229–234, 2003.
- [3817] P. M. Long. Efficient algorithms for learning functions with bounded variation. *Inform. Comput.*, 188(1):99–115, 2004.
- [3818] P. M. Long. New bounds on the price of bandit feedback for mistake-bounded online multi-class learning. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 3–10, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3819] P. M. Long and R. A. Servedio. Martingale boosting. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 79–94. Springer, 2005.
- [3820] P. M. Long and R. A. Servedio. Discriminative learning can succeed where generative learning fails. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 319–334, Berlin, 2006. Springer.
- [3821] P. M. Long and L. Tan. PAC learning axis-aligned rectangles with respect to product distributions from multiple-instance examples. *Machine Learning*, 30(1):7–21, 1998. Earlier version in COLT96.
- [3822] P. M. Long and V. B. Vega. Boosting and microarray data. *Machine Learning*, 52(1-2):31–44, 2003.
- [3823] P. M. Long and M. K. Warmuth. Composite geometric concepts and polynomial predictability. *Inform. Comput.*, 113(2):230–252, 1994.
- [3824] D. López, J. Calera-Rubio, and J. Gallego-Sánchez. Inference of  $k$ -testable directed acyclic graph languages. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 149–163, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.



- [3825] L. M. López, I. F. Ruiz, R. M. Bueno, and F. T. Ruiz. Dynamic discretization of continuous values from time series. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 280–291. Springer, 2000.
- [3826] J. Lorenz, M. Marcinişzyn, and A. Steger. Observational learning in random networks. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 574–588, Berlin, 2007. Springer.
- [3827] D. Lorenzo. Inductive logic programming for discrete event systems. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 250–261. Springer, 1996.
- [3828] L. Lovász. Some mathematics behind graph property testing. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, page 3, Berlin, Oct. 2008. Springer.
- [3829] S. Lovett and J. Zhang. Noisy population recovery from unknown noise. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1417–1431, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3830] V. Lozin, I. Razgon, V. Zamaraev, E. Zamaraeva, and N. Y. Zolotykh. Specifying a positive threshold function via extremal points. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 208–222, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3831] C.-J. Lu and W.-F. Lu. Making online decisions with bounded memory. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 249–261, Berlin, Heidelberg, New York, 2011. Springer.
- [3832] J. Lu and J. Arima. Inductive logic programming beyond logical implication. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 185–198. Springer, 1996.
- [3833] D. J. Lubinsky. Increasing the performance and consistency of classification trees by using the accuracy criterion at the leaves. In *Proc. 12th International Conference on Machine Learning*, pages 371–377. Morgan Kaufmann, 1995.
- [3834] S. M. Lucas. Grammatical inference and games: Extended abstract. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 1–4. Springer, 2010.
- [3835] M.-C. Ludl and G. Widmer. Relative unsupervised discretization for regression problems. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 246–253. Springer, 2000.

- [3836] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, MA, 1984.
- [3837] G. Lugosi. Concentration inequalities. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 7–8. Omnipress, 2008.
- [3838] G. Lugosi, S. Mannor, and G. Stoltz. Strategies for prediction under imperfect monitoring. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 248–262, Berlin, 2007. Springer.
- [3839] G. Lugosi and M. Pintér. A data dependent skeleton estimate for learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 51–56, New York, NY, 1996. ACM Press.
- [3840] G. Lugosi and H. U. Simon, editors. *Learning Theory, 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, Berlin, 2006. Springer.
- [3841] G. Lugosi and N. Vayatis. A consistent strategy for boosting algorithms. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 303–318. Springer, 2002.
- [3842] Y. Lui, R. Emery, D. Charabarti, W. Burgard, and S. Thrun. Using EM to learn 3D models of indoor environments with mobile robots. In *Proc. 18th International Conf. on Machine Learning*, pages 329–336. Morgan Kaufmann, San Francisco, CA, 2001.
- [3843] H. Luo and R. E. Schapire. Achieving all with no parameters: Adanormalhedge. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1286–1304, Paris, France, 03–06 July 2015. PMLR.
- [3844] T. Luo, K. Kramer, D. B. Goldgof, L. O. Hall, S. Samson, A. Remsen, and T. Hopkins. Active learning to recognize multiple types of plankton. *Journal of Machine Learning Research*, 6:589–613, 2005.
- [3845] W. Luo. Compute inclusion depth of a pattern. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 689–690. Springer, 2005.
- [3846] W. Luo and O. Schulte. Mind change efficient learning. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 398–412. Springer, 2005.
- [3847] W. Luo and O. Schulte. Mind change efficient learning. *Inform. Comput.*, 204(6):989–1011, 2006.
- [3848] F. M. Luque and G. Infante-Lopez. Bounding the maximal parsing performance of non-terminally separated grammars. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 135–147. Springer, 2010.

- [3849] F. M. Luque and G. Infante-Lopez. PAC-learning unambiguous  $k, l$ -NTS $^{\leq}$  languages. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 122–134. Springer, 2010.
- [3850] U. von Luxburg and O. Bousquet. Distance-based classification with Lipschitz functions. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 314–328. Springer, 2003.
- [3851] U. von Luxburg and O. Bousquet. Distance-based classification with lipschitz functions. *Journal of Machine Learning Research*, 5:669–695, 2004.
- [3852] U. von Luxburg, O. Bousquet, and M. Belkin. On the convergence of spectral clustering on random samples: The normalized case. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 457–471. Springer, 2004.
- [3853] U. von Luxburg, O. Bousquet, and B. Schölkopf. A compression approach to support vector model selection. *Journal of Machine Learning Research*, 5:293–323, 2004.
- [3854] J. Lv, J. Zhang, F. Wang, Z. Wang, and C. Zhang. Compressed learning with regular concept. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 163–178, Berlin, Heidelberg, New York, 2010. Springer.
- [3855] W. Maass. On-line learning with an oblivious environment and the power of randomization. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 167–175, San Mateo, CA, 1991. Morgan Kaufmann.
- [3856] W. Maass. Efficient agnostic PAC-learning with simple hypotheses. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 67–75, New York, NY, 1994. ACM Press.
- [3857] W. Maass. On the complexity of learning on neural nets. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 1–17, Oxford, 1994. Oxford University Press.
- [3858] W. Maass, editor. *Proceedings of the Eighth Annual Conference on Computational Learning Theory*, New York, N.Y., 1995. ACM Press.
- [3859] W. Maass. Bounds for the computational power and learning complexity of analog neural nets. *SIAM J. Comput.*, 26(3):708–732, 1997. first in STOC '93.
- [3860] W. Maass. On the relevance of time in neural computation and learning. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 364–384. Springer, 1997.
- [3861] W. Maass. On the relevance of time in neural computation and learnin. *Theoret. Comput. Sci.*, 261(1):157–178, 2001. Special issue for ALT '97.

- [3862] W. Maass and M. Schmitt. On the complexity of learning for a spiking neuron. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 54–61, New York, NY, 1997. ACM Press.
- [3863] W. Maass and M. Schmitt. On the complexity of learning for spiking neurons with temporal coding. *Inform. Comput.*, 153(1):26–46, 1999.
- [3864] W. Maass, G. Schnitger, and E. D. Sontag. On the computational power of sigmoid versus Boolean threshold circuits. In *Proc. of the 32nd Symposium on the Foundations of Comp. Sci.*, pages 767–776. IEEE Computer Society Press, Los Alamitos, CA, 1991.
- [3865] W. Maass and G. Turán. On the complexity of learning from counterexamples. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 262–267. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [3866] W. Maass and G. Turán. On the complexity of learning from counterexamples and membership queries. In *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, 22-24 October 1990*, pages 203–210, Los Alamitos, CA, 1990. IEEE Computer Society Press.
- [3867] W. Maass and G. Turán. Lower bound methods and separation results for on-line learning models. *Machine Learning*, 9:107–145, 1992.
- [3868] W. Maass and G. Turán. Algorithms and lower bounds for on-line learning of geometrical concepts. *Machine Learning*, 14:251–269, 1994.
- [3869] W. Maass and G. Turán. How fast can a threshold gate learn? In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 13, pages 381–414. MIT Press, 1994. Earlier versions appeared in FOCS89 and FOCS90.
- [3870] W. Maass and M. K. Warmuth. Efficient learning with virtual threshold gates. In *Proc. 12th International Conference on Machine Learning*, pages 378–386. Morgan Kaufmann, 1995.
- [3871] A. Macintyre and E. D. Sontag. Finiteness results for sigmoid. In *Proc. 25th Annu. ACM Sympos. Theory Comput.*, pages 325–334, New York, NY, 1993. ACM Press.
- [3872] D. MacKay. *Bayesian Methods for Adaptive Models*. PhD thesis, California Institute of Technology, 1992.
- [3873] D. J. C. MacKay. Choice of basis for Laplace approximation. *Machine Learning*, 33(1):77–86, 1998.
- [3874] R. Maclin and J. W. Shavlik. Using knowledge-based neural networks to improve algorithms: Refining the chou-fasman algorithm for protein folding. *Machine Learning*, 11:195–215, 1993.
- [3875] R. Maclin and J. W. Shavlik. Refining algorithms with knowledge-based neural networks: improving the Cho-Fasman algorithm for protein folding. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 9, pages 249–286. MIT Press, 1994.
- [3876] R. Maclin and J. W. Shavlik. Creating advice-taking reinforcement learners. *Machine Learning*, 22:251–281, 1996.

- [3877] O. Madani, S. Hanks, and A. Condon. On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence, July 18-22, 1999, Orlando, Florida, USA*, pages 541–548. AAAI Press / The MIT Press, 1999.
- [3878] O. Madani, D. J. Lizotte, and R. Greiner. The budgeted multi-armed bandit problem. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 643–645. Springer, 2004.
- [3879] P. Maes. How to do the right thing. *Connection Science Journal, Special Issue on Hybrid Systems*, 1, 1990.
- [3880] P. Maes and R. A. Brooks. Learning to coordinate behaviors. In *Proc. of AAAI-90*, volume 2, pages 796–802, Boston, MA, July 1990. AAAI Press/The MIT Press.
- [3881] M. Magdon-Ismail and J. Sill. Using a linear fit to determine monotonicity directions. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 477–491. Springer, 2003.
- [3882] D. M. Magerman. Learning grammatical structure using statistical decision-trees. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 1–21. Springer, 1996.
- [3883] L. Magnoni, M. Mirolli, F. Montagna, and G. Simi. PAC learning of probability distributions over a discrete domain. *Theoret. Comput. Sci.*, 299(1-3):37–63, 2003.
- [3884] S. Magureanu, R. Combes, and A. Proutiere. Lipschitz bandits: Regret lower bound and optimal algorithms. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 975–999, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3885] S. Mahadevan. To discount or not to discount in reinforcement learning: a case study comparing R learning and Q learning. In *Proc. 11th International Conference on Machine Learning*, pages 164–172. Morgan Kaufmann, 1994.
- [3886] S. Mahadevan. Average reward reinforcement learning: Foundations, algorithms, and empirical results. *Machine Learning*, 22:159–195, 1996.
- [3887] S. Mahadevan. Sensitive discount optimality: unifying discounted and average reward reinforcement learning. In *Proc. 13th International Conference on Machine Learning*, pages 328–336. Morgan Kaufmann, 1996.
- [3888] S. Mahadevan and J. Connell. Automatic programming of behavior-based robots using reinforcement learning. Technical report, IBM Research at Yorktown Heights, Dec. 1990.
- [3889] S. Mahadevan, N. Marchallick, T. K. Das, and A. Gosavi. Self-improving factory simulation using continuous-time average-reward reinforcement learning. In *Proc. 14th International Conference on Machine Learning*, pages 202–210. Morgan Kaufmann, 1997.
- [3890] S. Mahadevan and P. Tadepalli. Quantifying prior determination knowledge using the PAC learning model. *Machine Learning*, 17(1):69–105, 1994.

- [3891] S. Mahalanabis and D. Stefankovic. Density estimation in linear time. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 503–512. Omnipress, 2008.
- [3892] M. Mahdavi and R. Jin. Passive learning with target risk. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 252–269. JMLR.org, 2013.
- [3893] M. Mahdavi, L. Zhang, and R. Jin. Lower and upper bounds on the generalization of stochastic exponentially concave optimization. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1305–1320, Paris, France, 03–06 July 2015. PMLR.
- [3894] S. Mahloujifar, D. I. Diochnos, and M. Mahmoodi. Learning under  $p$ -tampering attacks. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 572–596. PMLR, 07–09 Apr. 2018.
- [3895] M. M. H. Mahmud. On universal transfer learning. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 135–149, Berlin, Oct. 2007. Springer.
- [3896] M. M. H. Mahmud. On universal transfer learning. *Theoret. Comput. Sci.*, 410(9):1826–1846, 2009. Special issue for ALT 2007.
- [3897] J. J. Mahoney and R. J. Mooney. Comparing methods for refining certainty-factor rule-bases. In *Proc. 11th International Conference on Machine Learning*, pages 173–180. Morgan Kaufmann, 1994.
- [3898] R. E. Mahony and R. C. Williamson. Prior knowledge and preferential structures in gradient descent learning algorithms. *Journal of Machine Learning Research*, 1:311–355, 2001.
- [3899] M. Maier, M. Hein, and U. von Luxburg. Cluster identification in nearest-neighbor graphs. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 196–210, Berlin, Oct. 2007. Springer.
- [3900] M. Maier, M. Hein, and U. von Luxburg. Optimal construction of  $k$ -nearest-neighbor graphs for identifying noisy clusters. *Theoret. Comput. Sci.*, 410(9):1749–1764, 2009. Special issue for ALT 2007.
- [3901] O.-A. Maillard. Robust risk-averse stochastic multi-armed bandits. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 218–233, Berlin, Heidelberg, New York, 2013. Springer.
- [3902] O.-A. Maillard. Boundary crossing for general exponential families. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 151–184, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [3903] O.-A. Maillard, R. Munos, and G. Stoltz. A finite-time analysis of multi-armed bandits problems with kullback-leibler divergences. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 497–514, Budapest, Hungary, 09–11 June 2011. PMLR.

- [3904] K. Makarychev, Y. Makarychev, and A. Vijayaraghavan. Correlation clustering with noisy partial information. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1321–1342, Paris, France, 03–06 July 2015. PMLR.
- [3905] K. Makarychev, Y. Makarychev, and A. Vijayaraghavan. Learning communities in the presence of errors. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1258–1291, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [3906] J. Makhoul, S. Roucos, and H. Gish. Vector quantization in speech coding. *Proceedings of the IEEE*, 73(11):1551–1588, Nov. 1985.
- [3907] E. Mäkinen. On inferring linear single-tree languages. *Information Processing Letters*, 73(1-2):1–3, 2000.
- [3908] K. Makino and T. Ibaraki. The maximum latency and identification of positive Boolean functions. *SIAM J. Comput.*, 26(5):1363–1383, 1997.
- [3909] O. Maler and A. Pnueli. On the learnability of infinitary regular sets. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 128–136, San Mateo, CA, 1991. Morgan Kaufmann.
- [3910] O. Maler and A. Pnueli. Learning  $\omega$ -regular languages from queries and counter-examples (A preliminary report). In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 161–170. Springer-Verlag, 1989.
- [3911] D. Malerba, F. Esposito, A. Lanza, and F. A. Lisi. First-order rule induction for the recognition of morphological patterns in topographic maps. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 88–101. Springer, 2001.
- [3912] M. A. Maloof. An initial study of an adaptive hierarchical vision system. In *Proc. 17th International Conf. on Machine Learning*, pages 567–573. Morgan Kaufmann, San Francisco, CA, 2000.
- [3913] M. A. Maloof, P. Langley, T. O. Binford, R. Nevatia, and S. Sage. Improved rooftop detection in aerial images with machine learning. *Machine Learning*, 53(1-2):157–191, 2003.
- [3914] M. A. Maloof and R. S. Michalski. Selecting examples for partial memory learning. *Machine Learning*, 41(1):27–52, 2000.
- [3915] D. Malzahn and M. Opper. An approximate analytical approach to resampling averages. *Journal of Machine Learning Research*, 4:1151–1173, 2003.
- [3916] H. Mamitsuka and N. Abe. Efficient mining from large database by query learning. In *Proc. 17th International Conf. on Machine Learning*, pages 575–582. Morgan Kaufmann, San Francisco, CA, 2000.
- [3917] H. Mamitsuka and N. Abe. Efficient data mining by active learning. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 258–267. Springer, 2001.

- [3918] H. Mamitsuka and K. Yamanishi. Protein secondary structure prediction based on stochastic-rule learning. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 240–251. Springer-Verlag, 1993.
- [3919] L. M. Manevitz and M. Yousef. One-class SVMs for document classification. *Journal of Machine Learning Research*, 2:139–154, 2001.
- [3920] L. Mangasarian and D. R. Musicant. Large scale kernel regression via linear programming. *Machine Learning*, 46(1-3):255–269, 2002.
- [3921] O. L. Mangasarian and D. R. Musicant. Lagrangian support vector machines. *Journal of Machine Learning Research*, 1:161–177, 2001.
- [3922] L. Mangu and E. Brill. Automatic rule acquisition for spelling correction. In *Proc. 14th International Conference on Machine Learning*, pages 187–194. Morgan Kaufmann, 1997.
- [3923] H. Mannila. Data mining and machine learning (abstract). In *Proc. 13th International Conference on Machine Learning*, page 555. Morgan Kaufmann, 1996.
- [3924] H. Mannila. Combining discrete algorithmic and probabilistic approaches in data mining. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, page 601. Springer, 2001.
- [3925] H. Mannila. Finding total and partial orders from data for seriation. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, page 4, Berlin, Oct. 2008. Springer.
- [3926] S. Mannor. Reinforcement learning for average reward zero-sum games. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 49–63. Springer, 2004.
- [3927] S. Mannor.  $k$ -armed bandit. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 561–563. Springer, New York, 2010.
- [3928] S. Mannor and R. Meir. Geometric bounds for generalization in boosting. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 461–472. Springer, 2001.
- [3929] S. Mannor and R. Meir. On the existence of linear weak learners and applications to boosting. *Machine Learning*, 48(1/3):219–251, 2002.
- [3930] S. Mannor, R. Meir, and T. Zhang. The consistency of greedy algorithms for classification. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 319–333. Springer, 2002.
- [3931] S. Mannor, R. Meir, and T. Zhang. Greedy algorithms for classification – consistency, convergence rates, and adaptivity. *Journal of Machine Learning Research*, 4:713–742, 2003.



- [3932] S. Mannor, V. Perchet, and G. Stoltz. Robust approachability and regret minimization in games with partial monitoring. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 515–536, Budapest, Hungary, 09–11 June 2011. PMLR.
- [3933] S. Mannor, V. Perchet, and G. Stoltz. Approachability in unknown games: Online learning meets multi-objective optimization. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 339–355, Barcelona, Spain, 13–15 June 2014. PMLR.
- [3934] S. Mannor, J. S. Shamma, and G. Arslan. Online calibrated forecasts: Memory efficiency versus universality for learning in games. *Machine Learning*, 67(1-2):77–115, 2007.
- [3935] S. Mannor and N. Shimkin. Adaptive strategies and regret minimization in arbitrarily varying Markov environments. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 128–142. Springer, 2001.
- [3936] S. Mannor and N. Shimkin. On-line learning with imperfect monitoring. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 552–566. Springer, 2003.
- [3937] S. Mannor and N. Shimkin. A geometric approach to multi-criterion reinforcement learning. *Journal of Machine Learning Research*, 5:325–360, 2004.
- [3938] S. Mannor and N. Shimkin. Online learning with variable stage duration. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 408–422, Berlin, 2006. Springer.
- [3939] S. Mannor, N. Srebro, and R. C. Williamson. Preface. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 1.1–1.2, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [3940] S. Mannor, N. Srebro, and R. C. Williamson, editors. *Proceedings of the Conference on Learning Theory, 25-27 June 2012, Edinburgh, Scotland*, volume 23 of *Proceedings of Machine Learning Research*. PMLR, 2012.
- [3941] S. Mannor and J. N. Tsitsiklis. Lower bounds on the sample complexity of exploration in the multi-armed bandit problem. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 418–432. Springer, 2003.
- [3942] S. Mannor and J. N. Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5:623–648, 2004.
- [3943] S. Mannor and J. N. Tsitsiklis. Online learning with constraints. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 529–543, Berlin, 2006. Springer.

- [3944] Y. Mansour. Learning via fourier transform. (see KushilevitzMa91), Apr. 1990.
- [3945] Y. Mansour. An  $O(n^{\log \log n})$  learning algorithm for DNF under the uniform distribution. *J. of Comput. Syst. Sci.*, 50(3):543–550, 1995.
- [3946] Y. Mansour. Pessimistic decision tree pruning based on tree size. In *Proc. 14th International Conference on Machine Learning*, pages 195–201. Morgan Kaufmann, 1997.
- [3947] Y. Mansour. Reinforcement learning and mistake bounded algorithms. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 183–192, New York, NY, 1999. ACM Press.
- [3948] Y. Mansour and D. McAllester. Boosting using branching programs. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 220–224. Morgan Kaufmann, San Francisco, 2000.
- [3949] Y. Mansour and D. McAllester. Generalization bounds for decision trees. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 69–80. Morgan Kaufmann, San Francisco, 2000.
- [3950] Y. Mansour and D. A. McAllester. Boosting using branching programs. *J. Comput. Syst. Sci.*, 64(1):103–112, 2002. Special issue for COLT 2000.
- [3951] Y. Mansour and S. Sahar. Implementation issues in the fourier transform algorithm. *Machine Learning*, 40(1):5–33, 2000.
- [3952] Y. Mansour and M. Schain. Learning with maximum-entropy distributions. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 201–210, New York, NY, 1997. ACM Press.
- [3953] Y. Mansour and M. Schain. Learning with maximum-entropy distributions. *Machine Learning*, 45(2):123–145, 2001.
- [3954] R. L. D. Mántaras. A distance-based attribute selection measure for decision tree induction. *Machine Learning*, 6:81–92, 1991.
- [3955] R. L. de Mántaras and E. Plaza, editors. *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, Berlin, 2000. Springer.
- [3956] P. Manurangsi and A. Rubinfeld. Inapproximability of VC dimension and littlestones dimension. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1432–1460, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [3957] C. Mao, J. Weed, and P. Rigollet. Minimax rates and efficient algorithms for noisy sorting. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 821–847. PMLR, 07–09 Apr. 2018.
- [3958] M. Marchand and J. Shawe-Taylor. Learning with the set covering machine. In *Proc. 18th International Conf. on Machine Learning*, pages 345–352. Morgan Kaufmann, San Francisco, CA, 2001.
- [3959] M. Marchand and J. Shawe-Taylor. The set covering machine. *Journal of Machine Learning Research*, 3:723–746, 2002.

- [3960] M. Marchand and M. Sokolova. Learning with decision lists of data-dependent features. *Journal of Machine Learning Research*, 6:427–451, 2005.
- [3961] A. Marchetti-Spaccamela and M. Protasi. Learning under uniform distribution. In *Fundamentals of Computation Theory, International Conference FCT '89, Szeged, Hungary, August 1989, Proceedings*, volume 380 of *Lecture Notes in Computer Science*, pages 327–338. Springer-Verlag, 1989.
- [3962] D. Margaritis and S. Thrun. Learning to locate an object in 3D space from a sequence of camera images. In *Proc. 15th International Conf. on Machine Learning*, pages 332–340. Morgan Kaufmann, San Francisco, CA, 1998.
- [3963] D. D. Margineantu and T. G. Dietterich. Pruning adaptive boosting. In *Proc. 14th International Conference on Machine Learning*, pages 211–218. Morgan Kaufmann, 1997.
- [3964] D. D. Margineantu and T. G. Dietterich. Bootstrap methods for the cost-sensitive evaluation of classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 583–590. Morgan Kaufmann, San Francisco, CA, 2000.
- [3965] C. E. Mariano and E. F. Morales. DQL: A new updating strategy for reinforcement learning based on Q-learning. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 324–335. Springer, 2001.
- [3966] Z. Markov and I. Marinchev. Metric-based inductive learning using semantic height functions. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 254–262. Springer, 2000.
- [3967] S. Markovitch and D. Rosenstein. Feature generation using general constructor functions. *Machine Learning*, 49(1):59–98, 2002.
- [3968] S. Markovitch and P. D. Scott. Information filtering: Selection mechanisms in learning systems. *Machine Learning*, 10:113–151, 1993.
- [3969] S. Markovitch and A. Shatil. Speedup learning for repair-based search by identifying redundant steps. *Journal of Machine Learning Research*, 4:649–682, 2003.
- [3970] O. Maron and A. L. Ratan. Multiple-instance learning for natural scene classification. In *Proc. 15th International Conf. on Machine Learning*, pages 341–349. Morgan Kaufmann, San Francisco, CA, 1998.
- [3971] L. Màrquez, L. Padró, and H. Rodríguez. A machine learning approach to POS tagging. *Machine Learning*, 39(1):59–91, 2000.
- [3972] A. Marron. Learning pattern languages from a single initial example and from queries. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 345–358, San Mateo, CA, 1988. Morgan Kaufmann.
- [3973] A. Marron and K. Ko. Identification of pattern languages from examples and queries. *Inform. Comput.*, 74(2):91–112, Aug. 1987.
- [3974] E. Martin and D. Osherson. A note on the use of probabilities by mechanical learners. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 261–271. Springer, 1995.

- [3975] E. Martin and D. N. Osherson. Scientific discovery based on belief revision. *The Journal of Symbolic Logic*, 62(4):1352–1370, Dec. 1997.
- [3976] E. Martin and D. N. Osherson. Belief revision in the service of scientific discovery. *Mathematical Social Sciences*, 36(1):57–68, 1998.
- [3977] E. Martin and D. N. Osherson. *Elements of Scientific Inquiry*. The MIT Press, Cambridge, Massachusetts, 1998.
- [3978] E. Martin and D. N. Osherson. Induction by enumeration. *Inform. Comput.*, 171(1):50–68, Nov. 2001.
- [3979] E. Martin and A. Sharma. On a syntactic characterization of classification with a mind change bound. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 413–428. Springer, 2005.
- [3980] E. Martin, A. Sharma, and F. Stephan. Learning, logic, and topology in a common framework. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 248–262. Springer, 2002.
- [3981] E. Martin, A. Sharma, and F. Stephan. Learning power and language expressiveness. *Theoret. Comput. Sci.*, 298(2):365–383, 2003.
- [3982] E. Martin, A. Sharma, and F. Stephan. On ordinal VC-dimension and some notions of complexity. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 54–68. Springer, 2003.
- [3983] E. Martin, A. Sharma, and F. Stephan. On the data consumption benefits of accepting increased uncertainty. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 83–98. Springer, 2004.
- [3984] E. Martin, A. Sharma, and F. Stephan. On ordinal VC-dimension and some notions of complexity. *Theoret. Comput. Sci.*, 364(1):62–76, 2006. Special issue for ALT 2003.
- [3985] É. Martin, A. Sharma, and F. Stephan. Unifying logic, topology and learning in parametric logic. *Theoret. Comput. Sci.*, 350(1):103–124, 2006. Special issue for ALT 2002.
- [3986] E. Martin, A. Sharma, and F. Stephan. On the data consumption benefits of accepting increased uncertainty. *Theoret. Comput. Sci.*, 382(3):170–182, 2007. Special issue for ALT 2004.
- [3987] J. D. Martin and D. O. Billman. Acquiring and combining overlapping concepts. *Machine Learning*, 16(1/2):121–155, 1994.
- [3988] J. K. Martin. An exact probability metric for decision tree splitting and stopping. *Machine Learning*, 28(2-3):257–291, 1997.
- [3989] L. Martin and F. Moal. A language-based similarity measure. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 336–347. Springer, 2001.

- [3990] L. Martin and C. Vrain. Induction of constraint logic programs. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 169–176. Springer, 1996.
- [3991] P. Martinek. An inverse limit of context-free grammars - A new approach to identifiability in the limit. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 171–185. Springer, 2000.
- [3992] J. F. Martínez-Trinidad and G. Sanches-D'íaz. LC: A conceptual clustering algorithm. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 117–127. Springer, 2001.
- [3993] A. Maruoka and E. Takimoto. Structured weight-based prediction algorithms. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 127–142. Springer, 1998.
- [3994] A. Maruoka and E. Takimoto. On-line algorithm to predict nearly as well as the best pruning of a decision tree. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 296–306. Springer, 2001.
- [3995] O. Maruyama and S. Miyano. Inferring a tree from walks. Technical Report RIFIS-TR-CS-51, Kyushu University, Dec. 1991.
- [3996] O. Maruyama and S. Miyano. Inferring a tree from walks. *Theoret. Comput. Sci.*, 161(1-2):289–300, 1996.
- [3997] O. Maruyama and S. Miyano. Design aspects of discovery systems. *IEICE Transactions on Information and Systems*, E83-D(1):61–70, 2000.
- [3998] Z. Marx, I. Dagan, and J. Buhmann. Coupled clustering: a method for detecting structural correspondence. In *Proc. 18th International Conf. on Machine Learning*, pages 353–360. Morgan Kaufmann, San Francisco, CA, 2001.
- [3999] Z. Marx, I. Dagan, J. M. Buhmann, and E. Shamir. Coupled clustering: A method for detecting structural correspondence. *Journal of Machine Learning Research*, 3:747–780, 2002.
- [4000] F. J. Maryanski and T. L. Booth. Inference of finite-state probabilistic grammars. *IEEE Trans. on Computers*, C-26:521–536, 1977.
- [4001] V. Maslov and V. V'yugin. Maximum entropy principle in non-ordered setting. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 221–233. Springer, 2004.
- [4002] L. Mason, P. L. Bartlett, and J. Baxter. Improved generalization through explicit optimization of margins. *Machine Learning*, 38(3):243–255, 2000.
- [4003] P. Massart and C. Meynet. Some rates of convergence for the selected lasso estimator. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012*,

- Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 17–33, Berlin, Heidelberg, New York, 2012. Springer.
- [4004] L. Massoulié and K. Xu. On the capacity of information processing systems. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1292–1297, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4005] S. Matar. Learning with minimal number of queries. Master’s thesis, University of Alberta, Canada, 1993.
- [4006] M. J. Mataric. Reward functions for accelerated learning. In *Proc. 11th International Conference on Machine Learning*, pages 181–189. Morgan Kaufmann, 1994.
- [4007] M. J. Mataric. Using communication to reduce locality in distributed multiagent learning. *J. of Experimental and Theoret. Artif. Intell.*, 10(3):357–369, 1998.
- [4008] D. H. Mathias. DNF - if you can’t learn ’em, teach ’em: an interactive model of teaching. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 222–229, New York, NY, 1995. ACM Press.
- [4009] H. D. Mathias. A model of interactive teaching. *J. of Comput. Syst. Sci.*, 54(3):487–501, 1997.
- [4010] I. Matsumoto, K. Hatano, and E. Takimoto. Online density estimation of bradley-terry models. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1343–1359, Paris, France, 03–06 July 2015. PMLR.
- [4011] S. Matsumoto, Y. Hayashi, and T. Shoudai. Polynomial time inductive inference of regular term tree languages from positive data. In *Algorithmic Learning Theory, 8th International Workshop, ALT ’97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 212–227. Springer, 1997.
- [4012] S. Matsumoto and A. Shinohara. Refutably probably approximately correct learning. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 469–483. Springer-Verlag, 1994.
- [4013] S. Matsumoto and A. Shinohara. Learning pattern languages using queries. In *Computational Learning Theory, Third European Conference, EuroCOLT ’97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 185–197. Springer, 1997.
- [4014] S. Matsumoto and T. Shoudai. Learning of ordered tree languages with height-bounded variables using queries. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 425–439. Springer, 2004.
- [4015] S. Matsumoto, Y. Suzuki, T. Shoudai, T. Miyahara, and T. Uchida. Learning of finite unions of tree patterns with repeated internal structured variables from queries. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 144–158. Springer, 2003.

- [4016] Y. Matsumoto. Parsing without grammar rules. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 1–6, Berlin, 2006. Springer.
- [4017] H. Matsunaga and K. Urahama. Partially supervised learning for nearest neighbor classifiers. *IEICE Transactions on Information and Systems*, E79-D(2):130–135, 1996.
- [4018] M. Matsuoka. An application of Bernstein polynomials in PAC model. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 220–228. Springer-Verlag, 1993.
- [4019] R. J. Matthews and W. Demopoulos. *Learnability and Linguistic Theory*. Kluwer, 1989.
- [4020] A. Maurer. Algorithmic stability and meta-learning. *Journal of Machine Learning Research*, 6:967–994, 2005.
- [4021] A. Maurer. The rademacher complexity of linear transformation classes. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 65–78, Berlin, 2006. Springer.
- [4022] A. Maurer. Unsupervised slow subspace-learning from stationary processes. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 363–377. Springer, Oct. 2006.
- [4023] A. Maurer. Unsupervised slow subspace-learning from stationary processes. *Theoret. Comput. Sci.*, 405(3):237–255, 2008. Special issue for ALT 2006.
- [4024] A. Maurer. A chain rule for the expected suprema of Gaussian processes. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 245–259, Berlin, Heidelberg, New York, 2014. Springer.
- [4025] A. Maurer. A chain rule for the expected suprema of Gaussian processes. *Theoret. Comput. Sci.*, 650:109–122, 2016. Special Issue: Algorithmic Learning Theory.
- [4026] A. Maurer. A vector-contraction inequality for rademacher complexities. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 3–17. Springer, 2016.
- [4027] A. Maurer. A second-order look at stability and generalization. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1461–1475, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4028] A. Maurer and M. Pontil. Generalization bounds for K-dimensional coding schemes in hilbert spaces. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 79–91, Berlin, Oct. 2008. Springer.

- [4029] A. Maurer and M. Pontil. A uniform lower error bound for half-space learning. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 70–78, Berlin, Oct. 2008. Springer.
- [4030] A. Maurer, M. Pontil, and B. Romera-Paredes. An inequality with applications to structured sparsity and multitask dictionary learning. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 440–460, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4031] E. Mayoraz and M. Moreira. On the decomposition of polychotomies into dichotomies. In *Proc. 14th International Conference on Machine Learning*, pages 219–226. Morgan Kaufmann, 1997.
- [4032] M. de la Maza. The generate, test, and explain discovery system architecture. In *Proc. 11th International Conference on Machine Learning*, pages 46–52. Morgan Kaufmann, 1994.
- [4033] D. McAllester. Simplified PAC-Bayesian margin bounds. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 203–215. Springer, 2003.
- [4034] D. McAllester, M. Collins, and F. Pereira. Case-factor diagrams for structured probabilistic modeling. *J. Comput. Syst. Sci.*, 74(1):84–96, 2008.
- [4035] D. McAllester and L. Ortiz. Concentration inequalities for the missing mass and for histogram rule error. *Journal of Machine Learning Research*, 4:895–911, 2003.
- [4036] D. McAllester and R. E. Schapire. On the convergence rate of good-Turing estimators. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 1–6. Morgan Kaufmann, San Francisco, 2000.
- [4037] D. A. McAllester. PAC-Bayesian model averaging. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 164–170, New York, NY, 1999. ACM Press.
- [4038] D. A. McAllester. Some PAC-Bayesian theorems. *Machine Learning*, 37(3):355–363, 1999.
- [4039] D. A. McAllester. PAC-Bayesian stochastic model selection. *Machine Learning*, 51(1):5–21, 2003.
- [4040] A. McCallum, D. Freitag, and F. Pereira. Maximum entropy Markov models for information extraction and segmentation. In *Proc. 17th International Conf. on Machine Learning*, pages 591–598. Morgan Kaufmann, San Francisco, CA, 2000.
- [4041] A. McCallum, R. Rosenfeld, T. Mitchell, and A. Y. Ng. Improving text classification by shrinkage in a hierarchy of classes. In *Proc. 15th International Conf. on Machine Learning*, pages 359–367. Morgan Kaufmann, San Francisco, CA, 1998.
- [4042] A. K. McCallum and K. Nigam. Employing EM and pool-based active learning for text classification. In *Proc. 15th International Conf. on Machine Learning*, pages 350–358. Morgan Kaufmann, San Francisco, CA, 1998.
- [4043] R. A. McCallum. Instance-based utile distinctions for reinforcement learning with hidden state. In *Proc. 12th International Conference on Machine Learning*, pages 387–395. Morgan Kaufmann, 1995.



- [4044] J. L. McClelland and D. E. Rumelhart, editors. *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. MIT Press, Cambridge, Massachusetts, 1988.
- [4045] T. L. McCluskey and M. M. West. A case study in the use of theory revision in requirements validation. In *Proc. 15th International Conf. on Machine Learning*, pages 368–376. Morgan Kaufmann, San Francisco, CA, 1998.
- [4046] E. McCreath. A probabilistic identification result. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 131–140. Springer, 2000.
- [4047] E. McCreath and A. Sharma. LIME: A system for learning relations. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 336–374. Springer, 1998.
- [4048] J. McDermott. The world would be a better place if non-programmers could program. *Machine Learning*, 4:337–338, 1989.
- [4049] A. McGovern and A. G. Barto. Automatic discovery of subgoals in reinforcement learning using diverse density. In *Proc. 18th International Conf. on Machine Learning*, pages 361–368. Morgan Kaufmann, San Francisco, CA, 2001.
- [4050] A. McGovern, E. Moss, and A. G. Barto. Building a basic block instruction scheduler with reinforcement learning and rollouts. *Machine Learning*, 49(2-3):141–160, 2002.
- [4051] D. McKoskey and D. Boley. Error analysis of automatic speech recognition using principal direction divisive partitioning. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 263–270. Springer, 2000.
- [4052] G. McLachlan. Mixtures of factor analyzers. In *Proc. 17th International Conf. on Machine Learning*, pages 599–606. Morgan Kaufmann, San Francisco, CA, 2000.
- [4053] H. B. McMahan and A. Blum. Online geometric optimization in the bandit setting against an adaptive adversary. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 109–123. Springer, 2004.
- [4054] H. B. McMahan and F. Orabona. Unconstrained online linear learning in hilbert spaces: Minimax algorithms and normal approximations. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1020–1039, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4055] H. B. McMahan and M. Streeter. Open problem: Better bounds for online logistic regression. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 44.1–44.3, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4056] V. Medina-Chico, A. Suárez, and J. F. Lutsko. Backpropagation in decision trees for regression. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 348–359. Springer, 2001.

- [4057] C. Meek, B. Thiesson, and D. Heckerman. The learning-curve sampling method applied to model-based clustering. *Journal of Machine Learning Research*, 2:397–418, 2002.
- [4058] B. Megyesi. Shallow parsing with PoS taggers and linguistic features. *Journal of Machine Learning Research*, 2:639–668, 2002.
- [4059] G. Mehlman, H. Kaindl, and W. Barth. Feature construction during tree learning. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 391–403. Springer, 1995.
- [4060] D. Mehta and V. Raghavan. Decision tree approximations of Boolean functions. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 16–24. Morgan Kaufmann, San Francisco, 2000.
- [4061] S. Mei, T. Misiakiewicz, A. Montanari, and R. I. Oliveira. Solving SDPs for synchronization and maxcut problems via the grothendieck inequality. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1476–1515, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4062] M. Meilă. An accelerated Chow and Liu algorithm: fitting tree distributions to high-dimensional sparse data. In *Proc. 16th International Conf. on Machine Learning*, pages 249–257. Morgan Kaufmann, San Francisco, CA, 1999.
- [4063] M. Meila. Comparing clusterings by the variation of information. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 173–187. Springer, 2003.
- [4064] M. Meila and D. Heckerman. An experimental comparison of model-based clustering methods. *Machine Learning*, 42(1/2):9–29, 2001.
- [4065] M. Meila and M. I. Jordan. Learning with mixtures of trees. *Journal of Machine Learning Research*, 1:1–48, 2000.
- [4066] K. Meinke. CGE: A sequential learning algorithm for mealy automata. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 148–162. Springer, 2010.
- [4067] R. Meir. Performance bounds for nonlinear time series prediction. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 122–129, New York, NY, 1997. ACM Press.
- [4068] R. Meir. Nonparametric time series prediction through adaptive model selection. *Machine Learning*, 39(1):5–34, 2000.
- [4069] R. Meir, R. El-Yaniv, and S. Ben-David. Localized boosting. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 190–199. Morgan Kaufmann, San Francisco, 2000.
- [4070] R. Meir and J. F. Fontanari. On learning noisy threshold functions with finite precision weights. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 280–286, New York, NY, 1992. ACM Press.
- [4071] R. Meir and N. Merhav. On the stochastic complexity of learning realizable and unrealizable rules. *Machine Learning*, 19(3):241–261, 1995.

- [4072] R. Meir and T. Zhang. Generalization error bounds for Bayesian mixture algorithms. *Journal of Machine Learning Research*, 4:839–860, 2003.
- [4073] G. Melideo and S. Varricchio. Learning unary output two-tape automata from multiplicity and equivalence queries. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 87–102. Springer, 1998.
- [4074] C. Melis and E. Melis. Some considerations about formalization of analogical reasoning. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 125–134. Springer-Verlag, 1986.
- [4075] T. Melluish, C. Saunders, I. Nouretdinov, and V. Vovk. Comparing the Bayes and typicalness frameworks. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 360–371. Springer, 2001.
- [4076] O. Melnik. Decision region connectivity analysis: A method for analyzing high-dimensional classifiers. *Machine Learning*, 48(1/3):321–351, 2002.
- [4077] F. S. Melo and M. I. Ribeiro.  $Q$ -learning with linear function approximation. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13–15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 308–322, Berlin, 2007. Springer.
- [4078] P. Ménard and A. Garivier. A minimax and asymptotically optimal algorithm for stochastic bandits. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 223–237, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4079] F. Menczer. ARACHNID: Adaptive retrieval agents choosing heuristic neighborhoods for information discovery. In *Proc. 14th International Conference on Machine Learning*, pages 227–235. Morgan Kaufmann, 1997.
- [4080] F. Menczer and R. K. Belew. Adaptive retrieval agents: Internalizing local context and scaling up to the web. *Machine Learning*, 39(2/3):203–242, 2000.
- [4081] S. Mendelson. Geometric methods in the analysis of Glivenko-Cantelli classes. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 256–272. Springer, 2001.
- [4082] S. Mendelson. Learning relatively small classes. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 273–288. Springer, 2001.
- [4083] S. Mendelson. On the size of convex hulls of small sets. *Journal of Machine Learning Research*, 2:1–18, 2001.
- [4084] S. Mendelson. Geometric parameters of kernel machines. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 29–43. Springer, 2002.

- [4085] S. Mendelson. On the performance of kernel classes. *Journal of Machine Learning Research*, 4:759–771, 2003.
- [4086] S. Mendelson. On the limitations of embedding methods. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 353–365. Springer, 2005.
- [4087] S. Mendelson. Learning without concentration. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 25–39, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4088] S. Mendelson and A. Pajor. Ellipsoid approximation using random vectors. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 429–443. Springer, 2005.
- [4089] S. Mendelson and P. Philips. Random subclass bounds. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 329–343. Springer, 2003.
- [4090] S. Mendelson and P. Philips. On the importance of small coordinate projections. *Journal of Machine Learning Research*, 5:219–238, 2004.
- [4091] S. Mendelson and N. Tishby. Statistical sufficiency for classes in empirical  $L_2$  spaces. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 81–89. Morgan Kaufmann, San Francisco, 2000.
- [4092] S. Mendelson and R. Vershynin. Entropy, combinatorial dimensions and random averages. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 14–28. Springer, 2002.
- [4093] S. Mendelson and R. C. Williamson. Agnostic learning nonconvex function classes. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 1–13. Springer, 2002.
- [4094] A. K. Menon and R. C. Williamson. Bayes-optimal scorers for bipartite ranking. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 68–106, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4095] D. Meretakís, H. Lu, and B. Wüthrich. A study on the performance of large Bayes classifier. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 271–279. Springer, 2000.
- [4096] N. Merhav and M. Feder. Universal sequential learning and decisions from individual data sequences. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 413–427, New York, NY, 1992. ACM Press.
- [4097] W. Merkle and F. Stephan. Trees and learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 270–279, New York, NY, 1996. ACM Press.

- [4098] W. Merkle and F. Stephan. Refuting learning revisited. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 299–314. Springer, 2001.
- [4099] W. Merkle and F. Stephan. Refuting learning revisited. *Theoret. Comput. Sci.*, 298(1):145–177, 2003.
- [4100] W. Merkle and F. Stephan. Trees and learning. *J. Comput. Syst. Sci.*, 68(1):134–156, 2004.
- [4101] C. J. Merz. Using correspondence analysis to combine classifiers. *Machine Learning*, 36(1/2):33–58, 1999.
- [4102] C. J. Merz and M. J. Pazzani. A principal components approach to combining regression estimates. *Machine Learning*, 36(1/2):9–32, 1999.
- [4103] C. Mesterharm. Tracking linear-threshold concepts with Winnow. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 138–152. Springer, 2002.
- [4104] C. Mesterharm. Tracking linear-threshold concepts with Winnow. *Journal of Machine Learning Research*, 4:819–838, 2003.
- [4105] C. Mesterharm. On-line learning with delayed label feedback. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 399–413. Springer, Oct. 2005.
- [4106] N. Meuleau and P. Bourgin. Exploration of multi-state environments: Local measures and back-propagation of uncertainty. *Machine Learning*, 35(2):117–154, 1999.
- [4107] A. Meyer. Program size in restricted programming languages. *Inform. Control*, 21:382–394, 1972.
- [4108] L. Meyer. Monotonic and dual-monotonic probabilistic language learning of indexed families with high probability. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 66–78. Springer, 1997.
- [4109] L. Meyer. Probabilistic language learning under monotonicity constraint. *Theoret. Comput. Sci.*, 185(1):81–128, 1997. Special issue for ALT '95.
- [4110] L. Meyer. Aspects of complexity of conservative probabilistic learning. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 72–78, New York, NY, 1998. ACM Press.
- [4111] L. Meyer. Comparing the power of probabilistic learning and oracle identification under monotonicity constraints. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 306–320. Springer, 1998.
- [4112] L. Meyer. Aspects of complexity of probabilistic learning under monotonicity constraints. *Theoret. Comput. Sci.*, 268(2):275–322, 2001. Special issue for ALT '98.

- [4113] C. A. Micchelli and M. Pontil. A function representation for learning in banach spaces. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 255–269. Springer, 2004.
- [4114] C. A. Micchelli and M. Pontil. Feature space perspectives for learning the kernel. *Machine Learning*, 66(2-3):297–319, 2007.
- [4115] L. Michael. Missing information impediments to learnability. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 825–828, Budapest, Hungary, 09–11 June 2011. PMLR.
- [4116] R. Michalski. Pattern recognition as rule guided inductive inference. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-2:349–361, 1980.
- [4117] R. Michalski and G. Tecuci, editors. *Machine Learning: A Multistrategy Approach*, volume IV. Morgan Kaufmann, Los Altos, California, 1993.
- [4118] R. M. Michalski. *Understanding the Nature of Learning: Issues and Research Directions*, pages 3–26. Morgan Kaufmann, Los Altos, CA, 1986.
- [4119] R. S. Michalski. A theory and methodology of inductive inference. In *Machine Learning Vol2*, pages 215–244. Morgan Kaufman, 1986.
- [4120] R. S. Michalski. Inferential theory of learning as a conceptual basis for multistrategy learning. *Machine Learning*, 11:111–151, 1993.
- [4121] R. S. Michalski. LEARNABLE EVOLUTION MODEL: Evolutionary processes guided by machine learning. *Machine Learning*, 38(1/2):9–40, 2000.
- [4122] R. S. Michalski and P. Brazdil. Introduction. *Machine Learning*, 50(3):219–222, 2003.
- [4123] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, editors. *Machine Learning: An Artificial Intelligence Approach*, volume I. Morgan Kaufmann, Los Altos, California, 1983.
- [4124] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, editors. *Machine Learning: An Artificial Intelligence Approach*, volume II. Morgan Kaufmann, Los Altos, California, 1986.
- [4125] R. S. Michalski and J. Wnek. Guest editors’ introduction. *Machine Learning*, 27(3):205–208, 1997.
- [4126] D. Michie. Inductive rule generation in the context of the fifth generation. In *Proc. of the First International Workshop on Machine Learning*, pages 66–70, Monticello, Illinois, June 1983.
- [4127] L. Miclet. Regular inference with a tail-clustering method. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-10:737–743, 1980.
- [4128] L. Miclet, N. Barbot, and B. Jeudy. A lattice of sets of alignments built on the common subwords in a finite language. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 164–176, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.

- [4129] S. Miesbach. High accuracy path tracking by neural linearization techniques. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 2, pages 21–42. MIT Press, 1995.
- [4130] O. Mihatsch and R. Neuneier. Risk-sensitive reinforcement learning. *Machine Learning*, 49(2-3):267–290, 2002.
- [4131] J. D. R. Millán and C. Torras. A reinforcement connectionist approach to robot path finding in non-maze-like environments. *Machine Learning*, 8:363–395, 1992.
- [4132] W. T. Miller, F. H. Glanz, and L. G. Kraft. Application of a general learning algorithm to the control of robotic manipulators. *International Journal of Robotics Research*, 6(2):84–98, 1987.
- [4133] A. Milosavljević. Discovering dependencies via algorithmic mutual information: A case study in DNA sequence comparisons. *Machine Learning*, 21:35–50, 1995.
- [4134] A. Milosavljević, D. Haussler, and J. Jurka. Informed parsimonious inference of prototypical genetic sequences. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 102–117, San Mateo, CA, 1989. Morgan Kaufmann.
- [4135] A. Milosavljević and J. Jurka. Discovery by minimal length encoding: A case study in molecular evolution. *Machine Learning*, 12:69–87, 1993.
- [4136] M. Miltrup and G. Schnitger. Neural networks and efficient associative memory. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 235–246, New York, NY, 1998. ACM Press.
- [4137] J. Mingers. An empirical comparison of pruning methods for decision tree induction. *Machine Learning*, 4:227–243, 1989.
- [4138] J. Mingers. An empirical comparison of selection measures for decision-tree induction. *Machine Learning*, 3:319–342, 1989.
- [4139] H. Q. Minh and T. Hofmann. Learning over compact metric spaces. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 239–254. Springer, 2004.
- [4140] H. Q. Minh, P. Niyogi, and Y. Yao. Mercer’s theorem, feature maps, and smoothing. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 154–168, Berlin, 2006. Springer.
- [4141] E. Minicozzi. Some natural properties of strong identification in inductive inference. *Theoret. Comput. Sci.*, 2(3):345–360, 1976.
- [4142] S. Minsker. Estimation of extreme values and associated level sets of a regression function via selective sampling. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 105–121. JMLR.org, 2013.
- [4143] M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge, MA, 1969.
- [4144] S. Minton. A reply to Zito-Wolf’s book review of learning search control knowledge: An explanation-based approach. *Machine Learning*, 6:205–206, 1991.

- [4145] B. Mirkin. Concept learning and feature selection based on square-error clustering. *Machine Learning*, 35(1):25–39, 1999.
- [4146] B. Mirkin. Reinterpreting the category utility function. *Machine Learning*, 45(2):219–228, 2001.
- [4147] N. Mishra and L. Pitt. Generating all maximal independent sets of bounded-degree hypergraphs. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 211–217, New York, NY, 1997. ACM Press.
- [4148] N. Mishra, D. Ron, and R. Swaminathan. On finding large conjunctive clusters. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24–27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 448–462. Springer, 2003.
- [4149] S. Mitaim and B. Kosko. Stochastic resonance with adaptive fuzzy systems. In *Proc. 15th International Conf. on Machine Learning*, pages 377–385. Morgan Kaufmann, San Francisco, CA, 1998.
- [4150] A. Mitchell. “Boosting” a positive-data-only learner. In *Proc. 17th International Conf. on Machine Learning*, pages 607–614. Morgan Kaufmann, San Francisco, CA, 2000.
- [4151] A. Mitchell, T. Scheffer, A. Sharma, and F. Stephan. The VC-dimension of subclasses of pattern languages. In *Algorithmic Learning Theory, 10th International Conference, ALT ’99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 93–105. Springer, 1999.
- [4152] A. R. Mitchell. Learnability of a subclass of extended pattern languages. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 64–71, New York, NY, 1998. ACM Press.
- [4153] T. M. Mitchell. Version spaces: A candidate elimination approach to rule learning. In *Proc. IJCAI-77*, pages 305–310, Cambridge, Mass., Aug. 1977. International Joint Committee for Artificial Intelligence.
- [4154] T. M. Mitchell. The need for biases in learning generalizations. Technical Report CBM-TR-117, Rutgers University, New Brunswick, NJ, 1980.
- [4155] T. M. Mitchell. Generalization as search. *Art. Int.*, 18:203–226, 1982.
- [4156] T. M. Mitchell. *Machine Learning*. WCB/McGraw-Hill, Boston, Massachusetts, 1997.
- [4157] T. M. Mitchell. Computational models of neural representations in the human brain. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 5–6, Berlin, Oct. 2008. Springer.
- [4158] T. M. Mitchell, R. Caruana, D. Freitag, J. P. McDermott, and D. Zabowski. Experience with a learning personal assistant. *Communications of the ACM*, 37(7):80–91, 1994.
- [4159] T. M. Mitchell, R. M. Keller, and S. T. Kedar-Cabelli. Explanation-based generalization: A unifying view. *Machine Learning*, 1(1):47–80, 1986.
- [4160] O. Mitina and N. Vereshchagin. How to use expert advice in the case when actual values of estimated events remain unknown. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 91–97, New York, NY, 1995. ACM Press.



- [4161] V. Mitrană and C. Tîrnăucă. New bounds for the query complexity of an algorithm that learns DFAs with correction and equivalence queries. *Acta Informatica*, 48(1):43–50, 2011.
- [4162] T. Miyahara. Analogical reasoning using elementary formal system with mismatch. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 224–230. IOS Press, 1992.
- [4163] T. Miyahara. Refutable inference of functions computed by loop programs. Technical Report RIFIS-TR-CS-112, RIFIS, Kyushu University 33, 1995.
- [4164] T. Miyahara, T. Uchida, T. S. T. Kuboyama, K. Takahashi, and H. Ueda. Discovering knowledge from graph structured data by using refutably inductive inference of formal graph systems. *IEICE Transactions on Information and Systems*, E84-D(1):48–56, 2001.
- [4165] S. Miyano. Learning theory toward genome informatics. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 19–36. Springer-Verlag, 1993.
- [4166] S. Miyano, A. Shinohara, and T. Shinohara. Which classes of elementary formal systems are polynomial-time learnable? In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 139–150. IOS Press, 1992.
- [4167] S. Miyano, A. Shinohara, and T. Shinohara. Polynomial-time learning of elementary formal systems. *New Generation Computing*, 18(3):217–242, 2000.
- [4168] A. Miyata, J. Tarui, and E. Tomita. Learning Boolean functions in  $AC^0$  on attribute and classification noise. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 142–155. Springer, 2004.
- [4169] H. Mizumoto, S. Todoroki, Diptarama, R. Yoshinaka, and A. Shinohara. An efficient query learning algorithm for zero-suppressed binary decision diagrams. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 360–371, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4170] F. Mizushima and T. Toyoshima. Suprasymbolic grammar induction by recurrent self-organizing maps. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 346–348, Berlin, 2006. Springer.
- [4171] D. Mladenić and M. Grobelnik. Feature selection for unbalanced class distribution and Naive Bayes. In *Proc. 16th International Conf. on Machine Learning*, pages 258–267. Morgan Kaufmann, San Francisco, CA, 1999.
- [4172] D. S. Modha and E. Masry. Prequential and cross-validated regression estimation. *Machine Learning*, 33(1):5–39, 1998.
- [4173] D. S. Modha and W. S. Spangler. Feature weighting in  $k$ -means clustering. *Machine Learning*, 52(3):217–237, 2003.
- [4174] A. Modi, N. Jiang, S. Singh, and A. Tewari. Markov decision processes with continuous side information. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 597–618. PMLR, 07–09 Apr. 2018.

- [4175] M. Modrzejewski. Feature selection using rough sets theory. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 213–226. Springer-Verlag, 1993.
- [4176] S. E. Moelius III and S. Zilles. Learning without coding. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 300–314, Berlin, Heidelberg, New York, 2010. Springer.
- [4177] S. E. Moelius III and S. Zilles. Erratum: Learning without coding. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 452–452, Berlin, Heidelberg, New York, 2011. Springer.
- [4178] C. R. Mofizur and M. Numao. Constructive induction for recursive programs. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 161–175. Springer-Verlag, 1994.
- [4179] M. H. Mohamed, T. Minamoto, and K. Nijjima. Convergence rate of minimization learning for neural networks. In *Proc. 10th European Conference on Machine Learning*, volume 1398 of *Lecture Notes in Artificial Intelligence*, pages 412–417. Springer, 1998.
- [4180] M. Mohri. Learning from uncertain data. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 656–670. Springer, 2003.
- [4181] M. Mohri and A. M. Medina. New analysis and algorithm for learning with drifting distributions. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 124–138, Berlin, Heidelberg, New York, 2012. Springer.
- [4182] M. Mohri and K. Sridharan. Algorithmic Learning Theory (ALT) 2018: Preface. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 1–2. PMLR, 07–09 Apr. 2018.
- [4183] M. Mohri and S. Yang. Structural online learning. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 223–237. Springer, 2016.
- [4184] A. Molina and F. Pla. Shallow parsing using specialized HMMs. *Journal of Machine Learning Research*, 2:595–613, 2002.
- [4185] R. Moll, T. J. Perkins, and A. G. Barto. Machine learning for subproblem selection. In *Proc. 17th International Conf. on Machine Learning*, pages 615–622. Morgan Kaufmann, San Francisco, CA, 2000.

- [4186] M. Momma and K. P. Bennett. Sparse kernel partial least squares regression. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 216–230. Springer, 2003.
- [4187] F. Montagna. Investigations on measure-one identification of classes of languages. *Inform. Comput.*, 143(1):74–107, 1998.
- [4188] F. Montagna and D. Osherson. Learning to coordinate; a recursion theoretic perspective. *Synthese*, 118(3):363–382, 1999.
- [4189] F. Montagna and G. Simi. Paradigms in measure theoretic learning and in informant learning. *Studia Logica*, 62(2):243–268, 1999.
- [4190] D. J. Montana. Empirical learning using rule threshold optimization for detection of events in synthetic images. *Machine Learning*, 5:427–450, 1990.
- [4191] C. Monteleoni. Efficient algorithms for generalized active learning. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 650–652, Berlin, 2006. Springer.
- [4192] S. Monti, P. Tamayo, J. Mesirov, and T. Golub. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning*, 52(1-2):91–118, 2003.
- [4193] J. Moody. Fast learning in multi-resolution hierarchies. In *Advances in Neural Information Processing Systems I*, pages 29–39, San Mateo, CA, 1989. Morgan Kaufmann.
- [4194] J. Moody and C. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1(2):281–294, 1989.
- [4195] R. J. Mooney. Induction over the unexplained: Using overly-general domain theories to aid concept learning. *Machine Learning*, 10:79–110, 1993.
- [4196] R. J. Mooney. Encouraging experimental results on learning CNF. *Machine Learning*, 19(1):79–92, 1995.
- [4197] R. J. Mooney. A preliminary PAC analysis of theory revision. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 3, pages 43–53. MIT Press, 1995.
- [4198] A. W. Moore. Reinforcement learning in factories: the auton project (abstract). In *Proc. 13th International Conference on Machine Learning*, page 556. Morgan Kaufmann, 1996.
- [4199] A. W. Moore and C. G. Atkeson. Prioritized sweeping: Reinforcement learning with less data and less time. *Machine Learning*, 13:103–130, 1993.
- [4200] A. W. Moore and C. G. Atkeson. The parti-game algorithm for variable resolution reinforcement learning in multidimensional state-spaces. *Machine Learning*, 21:199–233, 1995.
- [4201] A. W. Moore, D. J. Hill, and M. P. Johnson. An empirical investigation of brute force to choose features, smoothers and function approximators. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 19, pages 361–379. MIT Press, 1995.

- [4202] A. W. Moore and M. S. Lee. Efficient algorithms for minimizing cross validation error. In *Proc. 11th International Conference on Machine Learning*, pages 190–198. Morgan Kaufmann, 1994.
- [4203] A. W. Moore, J. Schneider, and K. Deng. Efficient locally weighted polynomial regression predictions. In *Proc. 14th International Conference on Machine Learning*, pages 236–244. Morgan Kaufmann, 1997.
- [4204] A. W. Moore, J. G. Schneider, J. A. Boyan, and M. S. Lee. Q2: memory-based active learning for optimizing noisy continuous functions. In *Proc. 15th International Conf. on Machine Learning*, pages 386–394. Morgan Kaufmann, San Francisco, CA, 1998.
- [4205] J. H. Moore, J. S. Parker, and L. W. Hahn. Symbolic discriminant analysis for mining gene expression patterns. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 372–381. Springer, 2001.
- [4206] E. F. Morales. PAL: A pattern-based first-order inductive system. *Machine Learning*, 26(2-3):227–252, 1997.
- [4207] S. Moran and M. K. Warmuth. Labeled compression schemes for extremal classes. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 34–49. Springer, 2016.
- [4208] E. Moreau. Partial learning using link grammars data. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 211–222. Springer, 2004.
- [4209] J. Morgenstern and T. Roughgarden. Learning simple auctions. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1298–1318, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4210] D. E. Moriarty and R. Miikkulainen. Efficient learning from delayed rewards through symbolic evolution. In *Proc. 12th International Conference on Machine Learning*, pages 396–404. Morgan Kaufmann, 1995.
- [4211] D. E. Moriarty and R. Miikkulainen. Efficient reinforcement learning through symbiotic evolution. *Machine Learning*, 22:11–32, 1996.
- [4212] J. ichi Moribe, K. Hatano, E. Takimoto, and M. Takeda. Smooth boosting for margin-based ranking. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 227–239, Berlin, Oct. 2008. Springer.
- [4213] K. Morik. Balanced cooperative modeling. *Machine Learning*, 11:217–235, 1993.
- [4214] K. Morik. Tailoring representations to different requirements. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 1–12. Springer, 1999.

- [4215] K. Morik. The representation race – preprocessing for handling time phenomena. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 4–19. Springer, 2000.
- [4216] K. Morik, F. Bergadano, and W. Buntine. Guest editorial. *Machine Learning*, 14:137–138, 1994. This issue was a special issue from the Workshop on Evaluating and Changing representation from 1991.
- [4217] K. Morik and P. Brockhausen. A multistrategy approach to relational knowledge discovery in databases. *Machine Learning*, 27(3):287–312, 1997.
- [4218] K. Morik, P. Brockhausen, and T. Joachims. Combining statistical learning with a knowledge-based approach - a case study in intensive care monitoring. In *Proc. 16th International Conf. on Machine Learning*, pages 268–277. Morgan Kaufmann, San Francisco, CA, 1999.
- [4219] J. Morimoto and K. Doya. Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 623–630. Morgan Kaufmann, San Francisco, CA, 2000.
- [4220] Y. Morimoto, H. Ishii, and S. Morishita. Efficient construction of regression trees with range and region splitting. *Machine Learning*, 45(3):235–259, 2001.
- [4221] E. Morin and E. Martienne. Using a symbolic machine learning tool to refine lexico-syntactic patterns. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 292–299. Springer, 2000.
- [4222] S. Morishita. Computing optimal hypotheses efficiently for boosting. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 471–481. Springer, 2001.
- [4223] N. Morita, M. Haraguchi, and Y. Okubo. A method of similarity-driven knowledge revision for type specification. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 194–205. Springer, 1999.
- [4224] T. Moriyama and M. Sato. Properties of language classes with finite elasticity. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 187–196. Springer-Verlag, 1993.
- [4225] T. Moriyama and M. Sato. Properties of language classes with finite elasticity. *IEICE TRANSACTIONS on Information and Systems*, E78-D(5):532–538, 1995.
- [4226] E. Moroshko and K. Crammer. Weighted last-step min-max algorithm with improved sub-logarithmic regret. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 245–259. Berlin, Heidelberg, New York, 2012. Springer.
- [4227] E. Moroshko and K. Crammer. Weighted last-step min-max algorithm with improved sub-logarithmic regret. *Theoret. Comput. Sci.*, 558:107–124, 2014. Special Issue: Algorithmic Learning Theory.

- [4228] D. Moshkovitz and M. Moshkovitz. Mixing implies lower bounds for space bounded learning. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1516–1566, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4229] E. Mossel, J. Neeman, and A. Sly. Belief propagation, robust reconstruction and optimal recovery of block models. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 356–370, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4230] E. Mossel, R. O’Donnell, and R. A. Servedio. Learning juntas. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 9-11, 2003*, pages 206–212. ACM Press, 2003.
- [4231] E. Mossel, R. O’Donnell, and R. A. Servedio. Learning functions of  $k$  relevant variables. *J. Comput. Syst. Sci.*, 69(3):421–434, 2004.
- [4232] E. Mossel and S. Roch. Learning nonsingular phylogenies and hidden Markov models. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 366–375. ACM Press, 2005.
- [4233] E. Mossel and S. Roch. Slow emergence of cooperation for win-stay lose-shift on trees. *Machine Learning*, 67(1-2):7–22, 2007.
- [4234] E. Mossel and J. Xu. Density evolution in the degree-correlated stochastic block model. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1319–1356, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4235] T. Motoki. Inductive inference from all positive and some negative data. Unpublished, 1992.
- [4236] T. Motoki, T. Shinohara, and K. Wright. The correct definition of finite elasticity: corrigendum to Identification of unions. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, page 375, San Mateo, CA, 1991. Morgan Kaufmann.
- [4237] V. Mottl, S. Dvoenko, O. Seredin, C. Kulikowski, and I. Muchnik. Featureless pattern recognition in an imaginary hilbert space and its applicaton to protein fold classification. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 322–336. Springer, 2001.
- [4238] J. Mourtada, O.-A. M. Mourtada, and O.-A. Maillard. Efficient tracking of a growing number of experts. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 517–539, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4239] J. R. Movellan and J. L. McClelland. Contrastive learning with graded random networks. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 14, pages 215–232. MIT Press, 1994.
- [4240] J. R. Movellan and P. Mineiro. Robust sensor fusion: Analysis and application to audio visual speech recognition. *Machine Learning*, 32(2):85–100, 1998.

- [4241] M. C. Mozer and J. Bachrach. Discovering the structure of a reactive environment by exploration. Technical report, University of Colorado Department of Computer Science, Nov. 1989. Report CU-CS-451-89.
- [4242] M. C. Mozer and J. Bachrach. SLUG: A connectionist architecture for inferring the structure of finite-state environments. *Machine Learning*, 7:139–160, 1991.
- [4243] F. Mráz, F. Otto, and M. Plátek. Learning analysis by reduction from positive data. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 125–136, Berlin, 2006. Springer.
- [4244] S. Muggleton. Inductive logic programming. *New Generation Computing*, 8(4):295–318, 1991.
- [4245] S. Muggleton. Inductive logic programming: Derivations, successes and shortcomings. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 21–37. Springer-Verlag, 1993.
- [4246] S. Muggleton. Optimal layered learning: a PAC approach to incremental sampling. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 37–44. Springer-Verlag, 1993.
- [4247] S. Muggleton. Bayesian inductive logic programming. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 3–11, New York, NY, 1994. ACM Press.
- [4248] S. Muggleton. Predicate invention and utilization. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):121–130, 1994. Special issue for AII '92.
- [4249] S. Muggleton. Recent advances in inductive logic programming. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 3–11, New York, NY, 1994. ACM Press.
- [4250] S. Muggleton, C. H. Bryant, and A. Srinivasan. Measuring performance when positives are rare: Relative advantage versus predictive accuracy - A biological case study. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 300–312. Springer, 2000.
- [4251] S. Muggleton and D. Page. Guest editors' introduction. *Machine Learning*, 26(2-3):97–98, 1997.
- [4252] S. H. Muggleton, C. H. Bryant, and A. Srinivasan. Learning chomsky-like grammars for biological sequence families. In *Proc. 17th International Conf. on Machine Learning*, pages 631–638. Morgan Kaufmann, San Francisco, CA, 2000.
- [4253] I. Mukherjee, C. Rudin, and R. E. Schapire. The rate of convergence of adaboost. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 537–558, Budapest, Hungary, 09–11 June 2011. PMLR.
- [4254] I. Mukherjee and R. E. Schapire. Learning with continuous experts using drifting games. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 240–255, Berlin, Oct. 2008. Springer.

- [4255] I. Mukherjee and R. E. Schapire. Learning with continuous experts using drifting games. *Theoret. Comput. Sci.*, 411(29-30):2670–2683, 2010. Special issue for ALT 2008.
- [4256] S. Mukherjee and T. Oates. Estimating graph parameters using graph grammars. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 292–294. Springer, 2008.
- [4257] Y. Mukouchi. Characterization of finite identification. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 260–267, 1992.
- [4258] Y. Mukouchi. Characterization of pattern languages. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 93–104. IOS Press, 1992.
- [4259] Y. Mukouchi. Inductive inference with bounded mind changes. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 125–134. Springer-Verlag, 1993.
- [4260] Y. Mukouchi. Induction inference of an approximate concept from positive data. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 484–499. Springer-Verlag, 1994.
- [4261] Y. Mukouchi. *Inductive inference of recursive concepts*. PhD thesis, Research Institute of Fundamental Information Science, Kyushu University 33, Fukuoka, Japan, 1994.
- [4262] Y. Mukouchi. Inferring a system from examples with time passage. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 197–211. Springer, 1997.
- [4263] Y. Mukouchi and S. Arikawa. Inductive inference machines that can refute hypothesis spaces. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 123–136. Springer-Verlag, 1993.
- [4264] Y. Mukouchi and S. Arikawa. Towards a mathematical theory of machine discovery from facts. *Theoret. Comput. Sci.*, 137(1):53–84, 1995. Special issue for ALT '93.
- [4265] Y. Mukouchi and M. Sato. Language learning with a neighbor system. In *Discovery Science, Third International Conference, DS 2000, Kyoto, Japan, December 2000, Proceedings*, volume 1967 of *Lecture Notes in Artificial Intelligence*, pages 183–196. Springer, 2000.
- [4266] Y. Mukouchi and M. Sato. Refutable language learning with a neighbor system. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 267–282. Springer, 2001.
- [4267] Y. Mukouchi and M. Sato. Refutable language learning with a neighbor system. *Theoret. Comput. Sci.*, 298(1):89–110, 2003.



- [4268] Y. Mukouchi and M. Sato. Learning languages generated by elementary formal systems and its application to SH languages. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 380–394. Springer, 2004.
- [4269] Y. Mukouchi, I. Yamue, and M. Sato. Inferring a rewriting system from examples. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 93–104. Springer, 1998.
- [4270] W. Mulder and P. Adriaans. Using grammar induction to model adaptive behavior of networks of collaborative agents. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 163–177. Springer, 2010.
- [4271] M. Mullin and R. Sukthankar. Complete cross-validation for nearest neighbor classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 639–646. Morgan Kaufmann, San Francisco, CA, 2000.
- [4272] R. Munos. A convergent reinforcement learning algorithm in the continuous case: the finite-element reinforcement learning. In *Proc. 13th International Conference on Machine Learning*, pages 337–345. Morgan Kaufmann, 1996.
- [4273] R. Munos. A study of reinforcement learning in the continuous case by the means of viscosity solutions. *Machine Learning*, 40(3):265–299, 2000.
- [4274] R. Munos and A. Moore. Variable resolution discretization in optimal control. *Machine Learning*, 49(2-3):291–323, 2002.
- [4275] R. Munos and A. W. Moore. Rates of convergence for variable resolution schemes in optimal control. In *Proc. 17th International Conf. on Machine Learning*, pages 647–654. Morgan Kaufmann, San Francisco, CA, 2000.
- [4276] P. M. Murphy and M. J. Pazzani. Revision of production system rule-bases. In *Proc. 11th International Conference on Machine Learning*, pages 199–207. Morgan Kaufmann, 1994.
- [4277] P. M. Murphy and M. J. Pazzani. Exploring the decision forest: an empirical investigation of Occam’s razor in decision tree induction. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 11, pages 171–187. MIT Press, 1997.
- [4278] J. F. Murray, G. F. Hughes, and K. Kreutz-Delgado. Machine learning methods for predicting failures in hard drives: A multiple-instance application. *Journal of Machine Learning Research*, 6:783–816, 2005.
- [4279] M. Muselli. Is the pocket algorithm optimal? In *Computational Learning Theory, Second European Conference, EuroCOLT ’95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 287–297. Springer, 1995.
- [4280] M. A. Musen. Automated support for building and extending expert models. *Machine Learning*, 4:347–375, 1989.
- [4281] J. Mycielski. A learning algorithm for linear operators. *Proc. of the American Mathematical Society*, 103(2):547–550, 1988.

- [4282] K. Myers, M. Kearns, S. Singh, and M. A. Walker. A boosting approach to topic spotting on subdialogues. In *Proc. 17th International Conf. on Machine Learning*, pages 655–662. Morgan Kaufmann, San Francisco, CA, 2000.
- [4283] C. Nadeau and Y. Bengio. Inference for the generalization error. *Machine Learning*, 52(3):239–281, 2003.
- [4284] H. Nagao, T. Higuchi, T. Iyemori, and T. Araki. Automatic detection of geomagnetic jerks by applying a statistical time series model to geomagnetic monthly means. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 360–371. Springer, 2001.
- [4285] P. B. Nair, A. Choudhury, and A. J. Keane. Some greed algorithms for sparse nonlinear regression. In *Proc. 18th International Conf. on Machine Learning*, pages 369–376. Morgan Kaufmann, San Francisco, CA, 2001.
- [4286] P. B. Nair, A. Choudhury, and A. J. Keane. Some greedy learning algorithms for sparse regression and classification with mercer kernels. *Journal of Machine Learning Research*, 3:781–801, 2002.
- [4287] A. Nakamura. An efficient exact learning algorithm for ordered binary decision diagrams. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 307–322. Springer, 1997.
- [4288] A. Nakamura. Learning specialist decision lists. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 215–225, New York, NY, 1999. ACM Press.
- [4289] A. Nakamura. Query learning of bounded-width OBDDs. *Theoret. Comput. Sci.*, 241(1-2):83–114, 2000. Special issue for ALT '96.
- [4290] A. Nakamura. An efficient query learning algorithm for ordered binary decision diagrams. *Inform. Comput.*, 201(2):178–198, 2005.
- [4291] A. Nakamura. Learning-related complexity of linear ranking functions. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 378–392. Springer, Oct. 2006.
- [4292] A. Nakamura and N. Abe. Exact learning of linear combinations of monotone terms from function value queries. *Theoret. Comput. Sci.*, 137(1):159–176, 1995. Special issue for ALT '93.
- [4293] A. Nakamura and N. Abe. On-line learning of binary and  $n$ -ary relations over multi-dimensional clusters. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 214–221, New York, NY, 1995. ACM Press.
- [4294] A. Nakamura and N. Abe. Collaborative filtering using weighted majority prediction algorithms. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 395–403. Morgan Kaufmann, San Francisco, CA, 1998.
- [4295] A. Nakamura and N. Abe. Online learning of binary and  $n$ -ary relations over clustered domains. *J. Comput. Syst. Sci.*, 65(2):224–256, 2002.

- [4296] A. Nakamura and S. Miura. Learning sparse linear combinations of basis functions over a finite domain. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 138–150. Springer, 1995.
- [4297] A. Nakamura, M. Schmitt, N. Schmitt, and H. U. Simon. Bayesian networks and inner product spaces. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 518–533. Springer, 2004.
- [4298] A. Nakamura, J. Takeuchi, and N. Abe. Efficient distribution-free population learning of simple concepts. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):53–82, 1998. Special issue for AII '94 and ALT '94.
- [4299] K. Nakamura. Extending incremental learning of context free grammars in synapse. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11–13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 281–282. Springer, 2004.
- [4300] K. Nakamura. Incremental learning of context free grammars by bridging rule generation and search for semi-optimum rule sets. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 72–83, Berlin, 2006. Springer.
- [4301] K. Nakamura and T. Ishiwata. Synthesizing context free grammars from sample strings based on inductive CYK algorithm. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 186–195. Springer, 2000.
- [4302] K. Nakamura and M. Matsumoto. Incremental learning of context free grammars. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23–25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 174–184. Springer, 2002.
- [4303] R. Nakano and K. Saito. Discovering polynomials to fit multivariate data having numeric and nominal variables. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 482–493. Springer, 2001.
- [4304] T. Nakazono, K. ichiro Moridomi, K. Hatano, and E. Takimoto. A combinatorial metrical task system problem under the uniform metric. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 276–287. Springer, 2016.
- [4305] D. Nam, S. Seo, and S. Kim. An efficient top-down search algorithm for learning Boolean networks of gene expression. *Machine Learning*, 65(1):229–245, 2006.
- [4306] A. Namatame. Knowledge-intensive learning in connectionist networks. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 35–46. IOS Press, 1992.
- [4307] S. C. Nandy, S. Das, and P. P. Goswami. An efficient  $k$  nearest neighbors searching algorithm for a query line. *Theoret. Comput. Sci.*, 299(1–3):273–288, 2003.

- [4308] K. S. Narendra and M. A. L. Thathachar. Learning automata – A survey. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-4(4):323–334, 1974.
- [4309] K. S. Narendra and M. A. L. Thathachar. *Learning Automata – An Introduction*. Prentice Hall, 1989.
- [4310] B. Natarajan. On learning functions from noise-free and noisy samples via occam’s razor. *SIAM J. Comput.*, 29(3):712–727, 1999.
- [4311] B. K. Natarajan. Learning functions from examples. Technical Report CMU-RI-TR-87-19, Robotics Institute, Carnegie-Mellon University, 1987.
- [4312] B. K. Natarajan. On learning Boolean functions. In *Proc. 19th Annu. ACM Sympos. Theory Comput.*, pages 296–304, New York, NY, 1987. ACM Press.
- [4313] B. K. Natarajan. On learning from exercises. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 72–87, San Mateo, CA, 1989. Morgan Kaufmann.
- [4314] B. K. Natarajan. On learning sets and functions. *Machine Learning*, 4:67–97, 1989.
- [4315] B. K. Natarajan. Probably-approximate learning over classes of distributions. Technical Report HPL-SAL-89-29, Hewlett Packard Labs, Palo Alto, CA, 1989.
- [4316] B. K. Natarajan. Some results on learning. Technical Report CMU-RI-TR-89-6, Carnegie Mellon, 1989.
- [4317] B. K. Natarajan. *Machine Learning: A Theoretical Approach*. Morgan Kaufmann, San Mateo, CA, 1991.
- [4318] B. K. Natarajan. Probably approximate learning of sets and functions. *SIAM J. Comput.*, 20(2):328–351, Apr. 1991.
- [4319] B. K. Natarajan. Occam’s razor for functions. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 370–376, New York, NY, 1993. ACM Press.
- [4320] B. K. Natarajan. A reply to Hellerstein’s book review of machine learning: A theoretical approach. *Machine Learning*, 13:151–152, 1993.
- [4321] B. K. Natarajan and P. Tadepalli. Two new frameworks for learning. In *Proc. of the 5th International Conference on Machine Learning*, pages 402–415, San Mateo, CA, June 1988. published by Morgan Kaufmann.
- [4322] T. Natschläger and W. Maass. Spiking neurons and the induction of finite state machines. *Theoret. Comput. Sci.*, 287(1):251–265, 2002.
- [4323] R. M. Neal. Learning stochastic feedforward networks. Technical report, Department of Computer Science, University of Toronto, Nov. 1990.
- [4324] R. M. Neal. Bayesian training of backpropagation networks by the hybrid monte carlo method. Technical Report CRG-TR-92-1, Department of Computer Science, U. of Toronto, Apr. 1992.
- [4325] R. M. Neal and G. E. Hinton. A new view of the EM algorithm that justifies incremental, sparse and other variants. In M. I. Jordan, editor, *Learning in Graphical Models*, pages 355–368. Kluwer Academic Publishers, 1998.

- [4326] F. Neri. Mining TCP/IP traffic for network intrusion detection by using a distributed genetic algorithm. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 313–322. Springer, 2000.
- [4327] F. Neri. Multi level knowledge in modeling qualitative physics learning. *Machine Learning*, 38(1/2):181–211, 2000.
- [4328] J. Nessel. Birds can fly... In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 56–63, New York, NY, 1998. ACM Press.
- [4329] J. Nessel. Learnability of enumerable classes of recursive functions from “typical” examples. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 264–275. Springer, 1999.
- [4330] J. Nessel and S. Lange. Learning erasing pattern languages with queries. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 86–100. Springer, 2000.
- [4331] J. Nessel and S. Lange. Learning erasing pattern languages with queries. *Theoret. Comput. Sci.*, 348(1):41–57, 2005. Special issue for ALT 2000.
- [4332] G. Neu. First-order regret bounds for combinatorial semi-bandits. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1360–1375, Paris, France, 03–06 July 2015. PMLR.
- [4333] G. Neu and G. Bartók. An efficient algorithm for learning with semi-bandit feedback. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 234–248, Berlin, Heidelberg, New York, 2013. Springer.
- [4334] G. Neu and V. Gómez. Fast rates for online learning in Linearly Solvable Markov Decision Processes. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1567–1588, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4335] F. Nevado, J.-A. Sánchez, and J. M. Benedí. Combination of estimation algorithms and grammatical inference techniques to learn stochastic context-free grammars. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 196–206. Springer, 2000.
- [4336] A. J. Nevins. A branch and bound incremental conceptual clusterer. *Machine Learning*, 18(1):5–22, 1995.
- [4337] Z. Nevo and R. El-Yaniv. On online learning of decision lists. *Journal of Machine Learning Research*, 3:271–301, 2002.
- [4338] B. Neyshabur, Y. Makarychev, and N. Srebro. Clustering, hamming embedding, generalized LSH and the max norm. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia,*

- October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 306–320, Berlin, Heidelberg, New York, 2014. Springer.
- [4339] B. Neyshabur, R. Tomioka, and N. Srebro. Norm-based capacity control in neural networks. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1376–1401, Paris, France, 03–06 July 2015. PMLR.
- [4340] A. Y. Ng. Preventing overfitting of cross-validation data. In *Proc. 14th International Conference on Machine Learning*, pages 245–253. Morgan Kaufmann, 1997.
- [4341] A. Y. Ng. On feature selection: learning with exponentially many irrelevant features as training examples. In *Proc. 15th International Conf. on Machine Learning*, pages 404–412. Morgan Kaufmann, San Francisco, CA, 1998.
- [4342] A. Y. Ng. Reinforcement learning and apprenticeship learning for robotic control. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 29–31. Springer, Oct. 2006.
- [4343] A. Y. Ng, D. Harada, and S. Russell. Policy invariance under reward transformations: theory and application to reward shaping. In *Proc. 16th International Conf. on Machine Learning*, pages 278–287. Morgan Kaufmann, San Francisco, CA, 1999.
- [4344] A. Y. Ng and M. I. Jordan. Convergence rates of the voting gibbs classifier, with application to Bayesian feature selection. In *Proc. 18th International Conf. on Machine Learning*, pages 377–384. Morgan Kaufmann, San Francisco, CA, 2001.
- [4345] A. Y. Ng and S. Russell. Algorithms for inverse reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 663–670. Morgan Kaufmann, San Francisco, CA, 2000.
- [4346] Y. K. Ng and T. Shinohara. Inferring unions of the pattern languages by the most fitting covers. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 269–282. Springer, Oct. 2005.
- [4347] Y. K. Ng and T. Shinohara. Characteristic sets for inferring the unions of the tree pattern languages by the most fitting hypotheses. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 307–319, Berlin, 2006. Springer.
- [4348] Y. K. Ng and T. Shinohara. Developments from enquiries into the learnability of the pattern languages from positive data. *Theoret. Comput. Sci.*, 397(1-3):150–165, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [4349] V. C. Nguyen, L. S. T. Ho, and V. Dinh. Generalization and robustness of batched weighted average algorithm with V-geometrically ergodic Markov data. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 264–278, Berlin, Heidelberg, New York, 2013. Springer.
- [4350] T. Niblett. Constructing decision trees in noisy domains. In *Progress in Machine Learning—Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 67–78, Bled, Yugoslavia, May 1987.

- [4351] T. Niblett. A note on refinement operators. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 329–335. Springer-Verlag, 1993.
- [4352] J. Nie, W. Kotłowski, and M. K. Warmuth. Online PCA with optimal regrets. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 98–112, Berlin, Heidelberg, New York, 2013. Springer.
- [4353] J. Nie, M. K. Warmuth, S. V. N. Vishwanathan, and X. Zhang. Open problem: Lower bounds for boosting with hadamard matrices. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1076–1079. JMLR.org, 2013.
- [4354] S. H. Nienhuys-Cheng and M. Polman. Complexity dimensions and learnability. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 348–353. Springer-Verlag, 1993.
- [4355] K. Nigam, A. K. McCallum, S. Thrun, and T. Mitchell. Text classification from labeled and unlabeled documents using EM. *Machine Learning*, 39(2/3):103–134, 2000.
- [4356] K. Niiijima. Domains of attraction in autoassociative memory networks for character pattern recognition. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 87–98. Springer-Verlag, 1993.
- [4357] K. Niiijima. Learning of associative memory networks based upon cone-like domains of attraction. *Neural Networks*, 10(9):1649–1658, 1997.
- [4358] K. Niiijima and S. Takano. Finding of signal and image by integer-type haar lifting wavelet transform. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 494–503. Springer, 2001.
- [4359] K. Niiijima, H. Uchida, E. Hirowatari, and S. Arikawa. Discovery of differential equations from numerical data. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 364–374. Springer, 1998.
- [4360] Z. Nikita. Optimal learning via local entropies and sample compression. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 2023–2065, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4361] D. Nikovski and I. Nourbakhsh. Learning probabilistic models for decision-theoretic navigation of mobile robots. In *Proc. 17th International Conf. on Machine Learning*, pages 671–678. Morgan Kaufmann, San Francisco, CA, 2000.
- [4362] N. J. Nilsson. *Learning Machines*. McGraw-Hill, New York, 1965.
- [4363] N. J. Nilsson. *The Mathematical Foundations of Learning Machines*. Morgan Kaufmann, 1990. (Republication of 1965 edition, with a new introduction by T.J. Sejnowski and H. White.).

- [4364] J. Nishii. A learning model for oscillatory networks. *Neural Networks*, 11(2):249–257, 1998.
- [4365] K. Nissim and U. Stemmer. Clustering algorithms for the centralized and local models. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 619–653. PMLR, 07–09 Apr. 2018.
- [4366] R. P. Nix. *Editing By Example (Ph.D Thesis)*. PhD thesis, Department of Computer Science, Yale University, 1983. Computer Science TR-280.
- [4367] R. P. Nix. Editing by example. In *Conference Record of the Eleventh Annual ACM Symposium on Principles of Programming Languages, Salt Lake City, Utah, January 1984*, pages 186–195. ACM Press, 1984.
- [4368] R. P. Nix. Editing by example. *ACM Transactions on Programming Languages and Systems*, 7(4):600–621, 1985.
- [4369] P. Niyogi. Free to choose: investigating the sample complexity of active learning of real valued functions. In *Proc. 12th International Conference on Machine Learning*, pages 405–412. Morgan Kaufmann, 1995.
- [4370] P. Niyogi and N. Karmarkar. An approach to data reduction and clustering with theoretical guarantees. In *Proc. 17th International Conf. on Machine Learning*, pages 679–686. Morgan Kaufmann, San Francisco, CA, 2000.
- [4371] P. Niyogi, S. Smale, and S. Weinberger. A topological view of unsupervised learning from noisy data. *SIAM J. Comput.*, 40(3):646–663, 2011.
- [4372] A. Nobel. Evaluating the performance of a simple inductive procedure in the presence of overfitting error. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 267–274, San Mateo, CA, 1991. Morgan Kaufmann.
- [4373] R. Nock. Complexity in the case against accuracy: When building one function-free horn clause is as hard as any. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 182–193. Springer, 1999.
- [4374] R. Nock and O. Gascuel. On learning decision committees. In *Proc. 12th International Conference on Machine Learning*, pages 413–420. Morgan Kaufmann, 1995.
- [4375] R. Nock and P. Jappy. On the power of decision lists. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 413–420. Morgan Kaufmann, San Francisco, CA, 1998.
- [4376] R. Nock and M. Sebban. Sharper bounds for the hardness of prototype and feature selection. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 224–237. Springer, 2000.
- [4377] K. Noé. The structure of scientific discovery: From a philosophical point of view. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 31–39. Springer, 2001.



- [4378] T. Nomoto and Y. Matsumoto. Comparing the minimum description length principle and boosting in the automatic analysis of discourse. In *Proc. 17th International Conf. on Machine Learning*, pages 687–694. Morgan Kaufmann, San Francisco, CA, 2000.
- [4379] B. Nordhausen and P. Langley. An integrated framework for empirical discovery. *Machine Learning*, 12:17–47, 1993.
- [4380] I. Nouretdinov, T. Melluish, and V. Vovk. Ridge regression confidence machine. In *Proc. 18th International Conf. on Machine Learning*, pages 385–392. Morgan Kaufmann, San Francisco, CA, 2001.
- [4381] I. Nouretdinov and V. Vovk. Criterion of calibration for transductive confidence machine with limited feedback. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 259–267. Springer, 2003.
- [4382] I. Nouretdinov and V. Vovk. Criterion of calibration for transductive confidence machine with limited feedback. *Theoret. Comput. Sci.*, 364(1):3–9, 2006. Special issue for ALT 2003.
- [4383] I. Nouretdinov, V. Vovk, M. Vyugin, and A. Gammerman. Pattern recognition and density estimation under the general i.i.d assumption. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 337–353. Springer, 2001.
- [4384] I. Nouretdinov, V. V’yugin, and A. Gammerman. Transductive confidence machine is universal. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 283–297. Springer, 2003.
- [4385] A. Nowé, J. Parent, and K. Verbeeck. Social agents playing a periodical policy. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 382–393. Springer, 2001.
- [4386] M. Núñez. The use of background knowledge in decision tree induction. *Machine Learning*, 6:231–250, 1991.
- [4387] M. Núñez. Learning patterns of behavior by observing system events. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 323–330. Springer, 2000.
- [4388] T. Oates, T. Armstrong, L. B. Bonache, and M. Atamas. Inferring grammars for mildly context sensitive languages in polynomial-time. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 137–147. Berlin, 2006. Springer.
- [4389] T. Oates and P. R. Cohen. Searching for structure in multiple streams of data. In *Proc. 13th International Conference on Machine Learning*, pages 346–354. Morgan Kaufmann, 1996.
- [4390] T. Oates and B. Heeringa. Estimating grammar parameters using bounded memory. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23–25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 185–198. Springer, 2002.

- [4391] T. Oates and D. Jensen. The effects of training set size on decision tree complexity. In *Proc. 14th International Conference on Machine Learning*, pages 254–262. Morgan Kaufmann, 1997.
- [4392] E. M. Oblow. Implementing valiant’s learnability theory using random sets. *Machine Learning*, 8:45–73, 1992.
- [4393] D. Obradovic. Recurrent neural networks with continuous topology adaptation, Kalman filter based training. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 2, pages 17–33. MIT Press, 1997.
- [4394] Z. Obradović and I. Parberry. Learning with discrete multivalued neurons. *J. of Comput. Syst. Sci.*, 49(2):375–390, 1994.
- [4395] W. of Occam. *Quodlibeta Septem* (in translation). Entities should not be multiplied unnecessarily. Comment regarding superfluous elaboration., Circa 1320.
- [4396] P. G. Odifreddi. *Classical Recursion Theory, Vol. II*. North Holland, Amsterdam, 1999.
- [4397] L. Ogiela, R. Tadeusiewicz, and M. R. Ogiela. Graph-based structural data mining in cognitive pattern interpretation. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 349–350, Berlin, 2006. Springer.
- [4398] A. T. Ogielski. Minimal samples of positive examples identifying  $k$ -CNF Boolean functions. *Inform. Comput.*, 113(2):220–229, 1994.
- [4399] M. I. Ohannessian and M. A. Dahleh. Rare probability estimation under regularly varying heavy tails. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 21.1–21.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4400] S. O’Hara. A model of the ‘redescription’ process in the context of geometric proportional analogy problems. In *Analogical and Inductive Inference, International Workshop AII ’92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 268–293. Springer-Verlag, 1992.
- [4401] Y. Ohsawa. Discovery of chances underlying real data. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 168–177. Springer, 2001.
- [4402] H. Ohwada and F. Mizoguchi. Integrating information visualization and retrieval for WWW information discovery. *Theoret. Comput. Sci.*, 292(2):547–571, 2003.
- [4403] E. Oja. Neural networks, principle components, and subspaces. *Internat. J. Neural Syst.*, 1(1):61–68, 1989.
- [4404] E. Oja. Unsupervised learning in neural computation. *Theoret. Comput. Sci.*, 287(1):187–207, 2002.
- [4405] J. Ok, S.-Y. Yun, A. Proutiere, and R. Mochaourab. Collaborative clustering: Sample complexity and efficient algorithms. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 288–329, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.

- [4406] M. Okada. Ideal concepts, intuitions, and mathematical knowledge acquisitions in Husserl and Hilbert. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 40–77. Springer, 2001.
- [4407] R. Okada, S. Matsumoto, T. Uchida, Y. Suzuki, and T. Shoudai. Exact learning of finite unions of graph patterns from queries. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 298–312, Berlin, Oct. 2007. Springer.
- [4408] T. Okadome. Simple flat languages: A learnable class in the limit from positive data. *International Journal of Foundations of Computer Science*, 10(4):483–502, 1999.
- [4409] S. Okamoto and N. Yugami. Theoretical analysis of the nearest neighbor classifier in noisy domains. In *Proc. 13th International Conference on Machine Learning*, pages 355–363. Morgan Kaufmann, 1996.
- [4410] S. Okamoto and N. Yugami. Generalized average-case analyses of the nearest neighbor algorithm. In *Proc. 17th International Conf. on Machine Learning*, pages 695–702. Morgan Kaufmann, San Francisco, CA, 2000.
- [4411] S. Okamoto and N. Yugami. Effects of domain characteristics on instance-based learning algorithms. *Theoret. Comput. Sci.*, 298(1):207–233, 2003.
- [4412] T. Okazaki, K. Sadakane, and H. Imai. Data compression method combining properties of PPM and CTW. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 268–283. Springer, 2001.
- [4413] Y. Okubo and M. Haraguchi. Planning with abstraction based on partial predicate mappings. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 183–194. Springer-Verlag, 1993.
- [4414] Y. Okubo and M. Haraguchi. Constructing predicate mappings for goal-dependent abstraction. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):169–197, 1998. Special issue for AII '94 and ALT '94.
- [4415] A. L. Oliveira, editor. *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11-13, 2000, Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, Berlin, 2000. Springer.
- [4416] A. L. Oliveira and S. Edwards. Limits of exact algorithms for inference of minimum size finite state machines. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 59–66. Springer, 1996.
- [4417] A. L. Oliveira and A. Sangiovanni-Vincentelli. Inferring reduced ordered decision graphs of minimum decision length. In *Proc. 12th International Conference on Machine Learning*, pages 421–429. Morgan Kaufmann, 1995.
- [4418] A. L. Oliveira and A. Sangiovanni-Vincentelli. Using the minimum description length principle to infer reduced ordered decision graphs. *Machine Learning*, 25:23–50, 1996.
- [4419] A. L. Oliveira and J. P. M. Silva. Efficient algorithms for the inference of minimum size DFAs. *Machine Learning*, 44(1/2):93–119, 2001.

- [4420] J. J. Oliver, R. A. Baxter, and C. S. Wallace. Unsupervised learning using MML. In *Proc. 13th International Conference on Machine Learning*, pages 364–372. Morgan Kaufmann, 1996.
- [4421] J. J. Oliver and D. J. Hand. On pruning and averaging decision trees. In *Proc. 12th International Conference on Machine Learning*, pages 430–437. Morgan Kaufmann, 1995.
- [4422] S. Omohundro. Efficient algorithms with neural networks behavior. *Complex Systems*, 1:273–347, 1987.
- [4423] J. Oncina. The data driven approach applied to the OSTIA algorithm. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 50–56. Springer, 1998.
- [4424] J. Oncina. Using multiplicity automata to identify transducer relations from membership and equivalence queries. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 154–162. Springer, 2008.
- [4425] J. Oncina and P. Garcia. Inferring regular languages in polynomial update time. In N. P. de la Blanca, A. Sanfeliu, and E. Vidal, editors, *Pattern Recognition and Image Analysis*, volume 1 of *Series in Machine Perception and Artificial Intelligence*, pages 49–61. World Scientific, Singapore, 1992.
- [4426] J. Oncina and M. A. Varó. Using domain information during the learning of a subsequential transducer. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 301–312. Springer, 1996.
- [4427] H. Ono, K. Makino, and T. Ibaraki. Logical analysis of data with decomposable structures. In *Computing and Combinatorics, 6th Annual International Conference, COCOON 2000, Sydney, Australia, July 26-28, 2000, Proceedings*, volume 1858 of *Lecture Notes in Computer Science*, pages 396–406. Springer, 2000.
- [4428] antiago Ontañón Enric Plaza. Learning when to collaborate among learning agents. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 394–405. Springer, 2001.
- [4429] D. W. Opitz. The effective size of a neural network: A principal component approach. In *Proc. 14th International Conference on Machine Learning*, pages 263–271. Morgan Kaufmann, 1997.
- [4430] D. W. Opitz and J. W. Shavlik. Using genetic search to refine knowledge-based neural networks. In *Proc. 11th International Conference on Machine Learning*, pages 208–216. Morgan Kaufmann, 1994.
- [4431] D. W. Opitz and J. W. Shavlik. Using heuristic search to expand knowledge-based neural networks. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 1, pages 3–19. MIT Press, 1995.
- [4432] M. Opper and D. Haussler. Calculation of the learning curve of Bayes optimal classification algorithm for learning a perceptron with noise. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 75–87, San Mateo, CA, 1991. Morgan Kaufmann.

- [4433] M. Opper and D. Haussler. Bounds for predictive errors in the statistical mechanics of in supervised learning. *Physical Review Letters*, 75(20):3772–3775, 1995.
- [4434] M. Opper and D. Haussler. General bounds for predictive errors in supervised learning. In *Proc. Workshop on the Theory of Neural Networks: The Statistical Mechanics Perspective*. World Scientific, 1995.
- [4435] M. Opper and O. Winther. Approximate inference in probabilistic models. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 494–504. Springer, 2004.
- [4436] F. Orabona and D. Pál. Scale-free algorithms for online linear optimization. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 287–301, Berlin, Heidelberg, New York, 2015. Springer.
- [4437] F. Orabona and D. Pál. Open problem: Parameter-free and scale-free online algorithms. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1659–1664, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4438] F. Orabona and D. Pál. Scale-free online learning. *Theoret. Comput. Sci.*, 716:50–69, 2018. Special Issue on ALT 2015.
- [4439] E. Ordentlich and T. Cover. On-line portfolio selection. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 310–313, New York, NY, 1996. ACM Press.
- [4440] R. Orihara. Analogical reasoning as a form of hypothetical reasoning and justification-based knowledge acquisition. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 243–254. IOS Press, 1992.
- [4441] D. Ormoneit and Š. Sen. Kernel-based reinforcement learning. *Machine Learning*, 49(2-3):161–178, 2002.
- [4442] P. O’Rorke. LT revisited: Explanation-based learning and the logic of principia mathematica. *Machine Learning*, 4:117–159, 1989.
- [4443] P. Orponen and R. Greiner. On the sample complexity of finding good search strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 352–358, San Mateo, CA, 1990. Morgan Kaufmann.
- [4444] L. Orseau. Optimality issues of universal greedy agents with static priors. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 345–359, Berlin, Heidelberg, New York, 2010. Springer.
- [4445] L. Orseau. Universal knowledge-seeking agents. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 353–367, Berlin, Heidelberg, New York, 2011. Springer.

- [4446] L. Orseau. Asymptotic non-learnability of universal agents with computable horizon functions. *Theoret. Comput. Sci.*, 473:149–156, 2013. Special Issue on Learning Theory.
- [4447] L. Orseau. Universal knowledge-seeking agents. *Theoret. Comput. Sci.*, 519:127–139, 2014. Special Issue: Algorithmic Learning Theory.
- [4448] L. Orseau, T. Lattimore, and M. Hutter. Universal knowledge-seeking agents for stochastic environments. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 158–172, Berlin, Heidelberg, New York, 2013. Springer.
- [4449] L. Orseau, T. Lattimore, and S. Legg. Soft-Bayes: Prod for mixtures of experts with log-loss. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory, volume 76 of Proceedings of Machine Learning Research*, pages 372–399, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4450] R. Ortner. Pseudometrics for state aggregation in average reward Markov decision processes. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 373–387, Berlin, Oct. 2007. Springer.
- [4451] R. Ortner. Online regret bounds for Markov decision processes with deterministic transitions. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 123–137, Berlin, Oct. 2008. Springer.
- [4452] R. Ortner. Online regret bounds for markov decision processes with deterministic transitions. *Theoret. Comput. Sci.*, 411(29-30):2684–2695, 2010. Special issue for ALT 2008.
- [4453] R. Ortner, O.-A. Maillard, and D. Ryabko. Selecting near-optimal approximate state representations in reinforcement learning. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 140–154, Berlin, Heidelberg, New York, 2014. Springer.
- [4454] R. Ortner, D. Ryabko, P. Auer, and R. Munos. Regret bounds for restless Markov bandits. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 214–228, Berlin, Heidelberg, New York, 2012. Springer.
- [4455] R. Ortner, D. Ryabko, P. Auer, and R. Munos. Regret bounds for restless markov bandits. *Theoret. Comput. Sci.*, 558:62–76, 2014. Special Issue: Algorithmic Learning Theory.
- [4456] R. Ortner, H. U. Simon, and S. Zilles, editors. *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*. Springer, 2016.
- [4457] M. Osborne. Shallow parsing using noisy and non-stationary training material. *Journal of Machine Learning Research*, 2:695–719, 2002.
- [4458] D. Osherson. Computer output. *Cognition*, 20:261–264, 1985.
- [4459] D. Osherson, M. Stob, and S. Weinstein. Ideal learning machines. *Cognitive Science*, 6:277–290, 1982.

- [4460] D. Osherson, M. Stob, and S. Weinstein. Note on a central lemma of learning theory. *J. of Math. Psych.*, 27:86–92, 1983.
- [4461] D. Osherson, M. Stob, and S. Weinstein. An analysis of a learning paradigm. In *Language Learning and Concept Acquisition*. Ablex Publ. Co., 1986.
- [4462] D. Osherson and S. Weinstein. A note on formal learning theory. *Cognition*, 11:77–88, 1982.
- [4463] D. Osherson and S. Weinstein. Formal learning theory. In *Handbook of Cognitive Neurology*. Plenum Press, New York, 1983.
- [4464] D. Osherson and S. Weinstein. Models of language acquisition. In *The Biology of Learning*. Springer-Verlag, Berlin, New York, 1984.
- [4465] D. Osherson and S. Weinstein. On the study of first language acquisition. Manuscript, 1992.
- [4466] D. N. Osherson, M. Stob, and S. Weinstein. Learning strategies. *Inform. Control*, 53(1/2):32–51, 1982.
- [4467] D. N. Osherson, M. Stob, and S. Weinstein. Learning theory and natural language. *Cognition*, 17:1–28, 1984.
- [4468] D. N. Osherson, M. Stob, and S. Weinstein. Aggregating inductive expertise. *Inform. Control*, 70(1):69–95, 1986.
- [4469] D. N. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, Cambridge, Massachusetts, 1986.
- [4470] D. N. Osherson, M. Stob, and S. Weinstein. Mathematical/mechanical? learners pay a price for Bayesianism. *J. Symbolic Logic*, 53(4):1245–1251, 1988.
- [4471] D. N. Osherson, M. Stob, and S. Weinstein. Synthesising inductive expertise. *Inform. Comput.*, 77(2):138–161, 1988.
- [4472] D. N. Osherson, M. Stob, and S. Weinstein. On approximate truth. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 88–101, San Mateo, CA, 1989. Morgan Kaufmann.
- [4473] D. N. Osherson, M. Stob, and S. Weinstein. A mechanical method of successful scientific inquiry. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 187–201, San Mateo, CA, 1990. Morgan Kaufmann.
- [4474] D. N. Osherson, M. Stob, and S. Weinstein. A universal inductive inference machine. *The Journal of Symbolic Logic*, 56(2):661–672, June 1991.
- [4475] D. N. Osherson, M. Stob, and S. Weinstein. A universal method of scientific inquiry. *Machine Learning*, 9:261–271, 1992.
- [4476] D. N. Osherson, M. Stob, and S. Weinstein. Logic and learning. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 1, pages 3–23. MIT Press, 1994.
- [4477] D. N. Osherson and S. Weinstein. Criteria of language learning. *Inform. Control*, 52(2):123–138, 1982.
- [4478] D. N. Osherson and S. Weinstein. Identification in the limit of first order structures. *Journal of Philosophical Logic*, 15:55–81, 1986.

- [4479] J. O’Sullivan, J. Langford, R. Caruana, and A. Blum. FeatureBoost: A meta-learning algorithm that improves model robustness. In *Proc. 17th International Conf. on Machine Learning*, pages 703–710. Morgan Kaufmann, San Francisco, CA, 2000.
- [4480] K. Otaki and A. Yamamoto. Estimation of generating processes of strings represented with patterns and refinements. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 177–182, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [4481] S. Otsuki. Discovery learning in intelligent tutoring systems. In *Algorithmic Learning Theory, Third Workshop, ALT ’92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 3–12. Springer-Verlag, 1993.
- [4482] M. Ott and F. Stephan. Avoiding coding tricks by hyperrobust learning. In *Computational Learning Theory, 4th European Conference, EuroCOLT ’99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 183–197. Springer, 1999.
- [4483] M. Ott and F. Stephan. Structural measures for games and process control in the branch learning model. *Theoret. Comput. Sci.*, 244(1-2):135–165, 2000. first in EuroColt ’97.
- [4484] M. Ott and F. Stephan. Avoiding coding tricks by hyperrobust learning. *Theoret. Comput. Sci.*, 284(1):161–180, 2002.
- [4485] C. Owens. Integrating feature extraction and memory search. *Machine Learning*, 10:311–339, 1993.
- [4486] A. Paccanaro and G. E. Hinton. Learning distributed representations by mapping concepts and relations into a linear space. In *Proc. 17th International Conf. on Machine Learning*, pages 711–718. Morgan Kaufmann, San Francisco, CA, 2000.
- [4487] G. Pagallo and D. Haussler. A greedy method for learning  $\mu$ -DNF functions under the uniform distribution. Technical report, University of California at Santa Cruz, UCSC-CRL-89-12, 1989.
- [4488] G. Pagallo and D. Haussler. Boolean feature discovery in empirical learning. *Machine Learning*, 5:71–99, 1990.
- [4489] D. Page and A. Srinivasan. ILP: A short look back and a longer look forward. *Journal of Machine Learning Research*, 4:415–430, 2003.
- [4490] M. Palhang and A. Sowmya. Learning discriminatory and descriptive rules by an inductive logic programming system. In *Proc. 16th International Conf. on Machine Learning*, pages 288–297. Morgan Kaufmann, San Francisco, CA, 1999.
- [4491] G. Paliouras, C. Papatheodorou, V. Karkaletsis, and C. D. Spyropoulos. Clustering the users of large web sites into communities. In *Proc. 17th International Conf. on Machine Learning*, pages 719–726. Morgan Kaufmann, San Francisco, CA, 2000.
- [4492] G. Paliouras and Y. Sakakibara. Guest editorial to the special issue on grammatical inference. *Machine Learning*, 66(1):3–5, 2007.
- [4493] N. Palmer and P. W. Goldberg. PAC-learnability of probabilistic deterministic finite state automata in terms of variation distance. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 157–170. Springer, Oct. 2005.



- [4494] N. Palmer and P. W. Goldberg. PAC-learnability of probabilistic deterministic finite state automata in terms of variation distance. *Theoret. Comput. Sci.*, 387(1):18–31, 2007. Special issue for ALT 2005.
- [4495] R. Pampapathi, B. Mirkin, and M. Levene. A suffix tree approach to anti-spam email filtering. *Machine Learning*, 65(1):309–338, 2006.
- [4496] C. Panagiotakopoulos and P. Tsampouka. The perceptron with dynamic margin. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 204–218, Berlin, Heidelberg, New York, 2011. Springer.
- [4497] T. W. Pao. A solution of the syntactical induction-inference problem for regular languages. *Comput. Lang.*, 3:53–64, 1978.
- [4498] T. W. Pao and J. W. I. Carr. A solution of the syntactical induction-inference problem for regular languages. *Comput. Lang.*, 3:53–64, 1978.
- [4499] G. Papa, P. Bianchi, and S. Cléménçon. Adaptive sampling for incremental optimization using stochastic gradient descent. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 317–331, Berlin, Heidelberg, New York, 2015. Springer.
- [4500] C. Papadimitriou. Learning the internet. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 396–396. Springer, 2002.
- [4501] C. Papadimitriou, S. Petti, and S. Vempala. Cortical computation via iterative constructions. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1357–1375, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4502] C. H. Papadimitriou and S. S. Vempala. Cortical learning via prediction. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1402–1422, Paris, France, 03–06 July 2015. PMLR.
- [4503] C. H. Papadimitriou and M. Yannakakis. On limited nondeterminism and the complexity of the V-C dimension. *J. Comput. Syst. Sci.*, 53(2):161–170, 1996.
- [4504] A. Papagelis and D. Kalles. Breeding decision trees using evolutionary techniques. In *Proc. 18th International Conf. on Machine Learning*, pages 393–400. Morgan Kaufmann, San Francisco, CA, 2001.
- [4505] R. Parekh and V. Honavar. An incremental interactive algorithm for grammar inference. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 238–249. Springer, 1996.
- [4506] R. Parekh and V. Honavar. Learning DFA from simple examples. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 116–131. Springer, 1997.

- [4507] R. Parekh and V. Honavar. Simple DFA are polynomially probably exactly learnable from simple examples. In *Proc. 16th International Conf. on Machine Learning*, pages 298–306. Morgan Kaufmann, San Francisco, CA, 1999.
- [4508] R. Parekh and V. Honavar. On the relationship between models for learning in helpful environments. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 207–220. Springer, 2000.
- [4509] R. Parekh and V. Honavar. Learning DFA from simple examples. *Machine Learning*, 44(1/2):9–35, 2001.
- [4510] R. Parekh, C. Nichitiu, and V. Honavar. A polynomial time incremental algorithm for learning DFA. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 37–49. Springer, 1998.
- [4511] L. Parra and P. Sajda. Blind source separation via generalized eigenvalue decomposition. *Journal of Machine Learning Research*, 4:1261–1269, 2003.
- [4512] R. J. Parsons, S. Forrest, and C. Burks. Genetic algorithms, operators, and DNA fragment assembly. *Machine Learning*, 21:11–33, 1995.
- [4513] G. Pask and H. V. Foerster. A predictive model for self organizing systems, part I. *Cybernetica*, III:258–300, 1960.
- [4514] G. Pask and H. V. Foerster. A predictive model for self organizing systems, part II. *Cybernetica*, IV:20–55, 1961.
- [4515] S. Pasteris, F. Vitale, C. Gentile, and M. Herbster. On similarity prediction and pairwise clustering. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 654–681. PMLR, 07–09 Apr. 2018.
- [4516] R. Paturi, S. Rajasekaran, and J. Reif. The light bulb problem. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 261–268, San Mateo, CA, 1989. Morgan Kaufmann.
- [4517] R. Paturi and M. E. Saks. On threshold circuits for parity. In *Proc. of the 31st Symposium on the Foundations of Comp. Sci.*, pages 397–404. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [4518] T. R. Payne and P. Edwards. Dimensionality reduction through sub-space mapping for nearest neighbor algorithms. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 331–343. Springer, 2000.
- [4519] M. Pazzani. Learning causal patterns: Making a transition from data-driven to theory-driven learning. *Machine Learning*, 11:173–194, 1993.
- [4520] M. Pazzani. A reply to Cohen’s book review of creating a memory of causal relationships. *Machine Learning*, 10:185–190, 1993.
- [4521] M. Pazzani. Book review: inductive logic programming: techniques and applications. *Machine Learning*, 23(1):103–108, 1996.

- [4522] M. Pazzani. Review of *inductive logic programming: techniques and applications* by nada lavrač, sašo Džeroski. *Machine Learning*, 23(1):103–108, 1996.
- [4523] M. Pazzani and D. Billsus. Learning and revising user profiles: The identification of interesting web sites. *Machine Learning*, 27(3):313–331, 1997.
- [4524] M. Pazzani and D. Kibler. The utility of knowledge in inductive learning. *Machine Learning*, 9:57–94, 1992.
- [4525] M. Pazzani, C. Merz, P. Murphy, K. Ali, T. Hume, and C. Brunk. Reducing misclassification costs. In *Proc. 11th International Conference on Machine Learning*, pages 217–225. Morgan Kaufmann, 1994.
- [4526] M. J. Pazzani. Guest editor’s introduction. *Machine Learning*, 16(1/2):7–9, 1994.
- [4527] M. J. Pazzani and W. Sarrett. A framework for average case analysis of conjunctive learning algorithms. *Machine Learning*, 9:349–372, 1992.
- [4528] J. Pearl. On the connection between the complexity and credibility of inferred models. *Journal of General Systems*, 4:255–264, 1978.
- [4529] J. Pearl. Capacity and error estimates for Boolean classifiers with limited complexity. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-1(4):350–355, Oct. 1979.
- [4530] J. Pearl. Jeffrey’s rule and the problem of autonomous inference agents. Class notes for 274A, Fall 1984, 1984.
- [4531] J. Pearl. Bayes decision methods. Technical Report CSD-850023, UCLA Computer Science Department, June 1985.
- [4532] J. Pearl. Bayesian networks: A model of self-activated memory for evidential reasoning. Technical Report CSD-850021, R-43, UCLA Computer Science Department, June 1985.
- [4533] J. Pearl. How to do with probabilities what people say you can’t. Technical Report CSD-850031, R-49, UCLA Computer Science Department, Sept. 1985.
- [4534] J. Pearl. On evidential reasoning in A hierarchy of hypothesis. Technical Report CSD-850032, UCLA Computer Science Department, Sept. 1985.
- [4535] J. Pearl and R. Dechter. Learning structure from data: a survey. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 230–244, San Mateo, CA, 1989. Morgan Kaufmann.
- [4536] J. Pearl and A. Paz. GRAPHOIDS: A graph-based logic for reasoning about relevance relations, or when would x tell you more about y if you already know z. Technical report, UCLA Computer Science Department, Dec. 1985.
- [4537] J. Pearl and A. Paz. On the logic of representing dependencies by graphs. In *Proc. 1986 Canadian AI Conference*, May 1986.
- [4538] J. Pearl and T. S. Verma. A formal theory of inductive causation. Technical Report R-1555 (I), UCLA Cognitive Systems Laboratory, Oct. 1990.
- [4539] R. A. Pearson and E. K. T. Smith. Reducing complexity of decision trees with two variable tests. In *Algorithmic Learning Theory, 7th International Workshop, ALT ’96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 91–99. Springer, 1996.

- [4540] E. Pełkalska, P. Paclík, and R. P. W. Duin. A generalized kernel approach to dissimilarity-based classification. *Journal of Machine Learning Research*, 2:175–211, 2001.
- [4541] D. Pelleg and A. Moore.  $X$ -means: Extending  $K$ -means with efficient estimation of the number of clusters. In *Proc. 17th International Conf. on Machine Learning*, pages 727–734. Morgan Kaufmann, San Francisco, CA, 2000.
- [4542] D. Pelleg and A. Moore. Mixtures of rectangles: Interpretable soft clustering. In *Proc. 18th International Conf. on Machine Learning*, pages 401–408. Morgan Kaufmann, San Francisco, CA, 2001.
- [4543] G. Pellegrino, C. Hammerschmidt, Q. Lin, and S. Verwer. Learning deterministic finite automata from infinite alphabets. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 120–131, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [4544] J. M. Peña, J. A. Lozano, and P. Larrañaga. Learning recursive Bayesian multinets for data clustering by means of constructive induction. *Machine Learning*, 47(1):63–89, 2002.
- [4545] M. D. Pendrith and M. J. McGarity. An analysis of direct reinforcement learning in non-Markovian domains. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 421–429. Morgan Kaufmann, San Francisco, CA, 1998.
- [4546] M. D. Pendrith and M. R. K. Ryan. Actual return reinforcement learning versus temporal differences: some theoretical and experimental results. In *Proc. 13th International Conference on Machine Learning*, pages 373–381. Morgan Kaufmann, 1996.
- [4547] J. Peng. Efficient memory-based dynamic programming. In *Proc. 12th International Conference on Machine Learning*, pages 438–446. Morgan Kaufmann, 1995.
- [4548] J. Peng and R. J. Williams. Incremental multi-step Q-learning. *Machine Learning*, 22:283–290, 1996.
- [4549] J. Peng and R. J. Williams. Technical note: incremental multi-step Q-learning. *Machine Learning*, 22(1/2/3):283–290, 1996.
- [4550] R. Peng, H. Sun, and L. Zanetti. Partitioning well-clustered graphs: Spectral clustering works! In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1423–1455, Paris, France, 03–06 July 2015. PMLR.
- [4551] D. M. Pennock, P. Maynard-Reid II, C. L. Giles, and E. Horvitz. A normative examination of ensemble learning algorithms. In *Proc. 17th International Conf. on Machine Learning*, pages 735–742. Morgan Kaufmann, San Francisco, CA, 2000.
- [4552] A. Pentina and S. Ben-David. Multi-task and lifelong learning of kernels. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 194–208, Berlin, Heidelberg, New York, 2015. Springer.
- [4553] A. Pentina and S. Ben-David. Multi-task Kernel Learning based on Probabilistic Lipschitzness. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 682–701. PMLR, 07–09 Apr. 2018.

- [4554] V. Perchet. Calibration and internal no-regret with random signals. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 68–82, Berlin/Heidelberg, 2009. Springer.
- [4555] V. Perchet and S. Mannor. Approachability, fast and slow. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 474–488. JMLR.org, 2013.
- [4556] V. Perchet, P. Rigollet, S. Chassang, and E. Snowberg. Batched bandit problems. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1456–1456, Paris, France, 03–06 July 2015. PMLR.
- [4557] F. C. N. Pereira. Frequencies vs. biases: machine learning problems in natural language processing – abstract. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, page 12, New York, NY, 1994. ACM Press.
- [4558] F. C. N. Pereira and Y. Singer. An efficient extension to mixture techniques for prediction and decision trees. *Machine Learning*, 36(3):183–199, 1999.
- [4559] C. N. Pereira, Fernando. Frequencies vs biases: machine learning problems in natural language processing - abstract. In *Proc. 11th International Conference on Machine Learning*, page 380. Morgan Kaufmann, 1994.
- [4560] E. Pérez and L. A. Rendell. Using multidimensional projection to find relations. In *Proc. 12th International Conference on Machine Learning*, pages 447–455. Morgan Kaufmann, 1995.
- [4561] E. Pérez and L. A. Rendell. Learning despite concept variation by finding structure in attribute-based data. In *Proc. 13th International Conference on Machine Learning*, pages 391–399. Morgan Kaufmann, 1996.
- [4562] M. A. Pérez. representing and learning quality-improving search control knowledge. In *Proc. 13th International Conference on Machine Learning*, pages 382–390. Morgan Kaufmann, 1996.
- [4563] P. Peris and D. López. Transducer inference by assembling specific languages. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 178–188. Springer, 2010.
- [4564] P. Peris, D. López, M. Campos, and J. M. Sempere. Protein motif prediction by grammatical inference. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 175–187, Berlin, 2006. Springer.
- [4565] S. Perkins, K. Lacker, and J. Theiler. Grafting: Fast, incremental feature selection by gradient descent in function space. *Journal of Machine Learning Research*, 3:1333–1356, 2003.
- [4566] T. J. Perkins and A. G. Barto. Lyapunov-constrained action sets for reinforcement learning. In *Proc. 18th International Conf. on Machine Learning*, pages 409–416. Morgan Kaufmann, San Francisco, CA, 2001.

- [4567] T. J. Perkins and A. G. Barto. Lyapunov design for safe reinforcement learning. *Journal of Machine Learning Research*, 3:803–832, 2002.
- [4568] C. Perlich, F. Provost, and J. S. Simonoff. Tree induction vs. logistic regression: A learning-curve analysis. *Journal of Machine Learning Research*, 4:211–255, 2003.
- [4569] P. Perner. Are case-based reasoning and dissimilarity-based classification two sides of the same coin? In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 35–51. Springer, 2001.
- [4570] P. Perner, editor. *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*. Springer, 2001.
- [4571] P. Perner and T. Belikova. A hybrid tool for data mining in picture archiving system. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 141–156. Springer, 2001.
- [4572] M. P. Perrone and B. S. Blais. Regression NSS: an alternative to cross validation. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 385–391, New York, NY, 1995. ACM Press.
- [4573] S. Persson. *Some Sequence Extrapolating Programs: a Study of Representation and Modeling in Inquiring Systems*. PhD thesis, Stanford University, 1966.
- [4574] N. Perugini. Neural network learning: Effects of network and training set size. Master’s thesis, MIT Department of Electrical Engineering and Computer Science, June 1989.
- [4575] L. Peshkin, N. Meuleau, and L. P. Kaelbling. Learning policies with external memory. In *Proc. 16th International Conf. on Machine Learning*, pages 307–314. Morgan Kaufmann, San Francisco, CA, 1999.
- [4576] L. Peshkin and S. Mukherjee. Bounds on sample size for policy evaluation in Markov environments. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 616–629. Springer, 2001.
- [4577] V. Pestov. PAC learnability of a concept class under non-atomic measures: A problem by Vidyasagar. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 134–147, Berlin, Heidelberg, New York, 2010. Springer.
- [4578] V. Pestov. PAC learnability under non-atomic measures: A problem by Vidyasagar. *Theoret. Comput. Sci.*, 473:29–45, 2013. Special Issue on Learning Theory.
- [4579] G. Petasis, G. Paliouras, C. D. Spyropoulos, and C. Halatsis. eg-GRIDS: Context-free grammatical inference from positive examples using genetic search. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 223–234. Springer, 2004.

- [4580] C. Peterson and J. R. Anderson. A mean field theory learning algorithm for neural networks. *Complex Systems*, 1:995–1019, 1987.
- [4581] V. Petridis and V. G. Kaburlasos. FINkNN: A fuzzy interval number  $k$ -nearest neighbor classifier for prediction of sugar production from populations of samples. *Journal of Machine Learning Research*, 4:17–37, 2003.
- [4582] B. Pfahringer. Compression-based discretization of continuous attributes. In *Proc. 12th International Conference on Machine Learning*, pages 456–463. Morgan Kaufmann, 1995.
- [4583] B. Pfahringer, H. Bensusan, and C. Giraud-Carrier. Meta-learning by landmarking various learning algorithms. In *Proc. 17th International Conf. on Machine Learning*, pages 743–750. Morgan Kaufmann, San Francisco, CA, 2000.
- [4584] T. V. Pham, M. Worring, and A. W. M. Smeulders. Face detection by aggregated Bayesian network classifiers. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 249–262. Springer, 2001.
- [4585] J. M. Phillips and Y. Zheng. Subsampling in smoothed range spaces. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 224–238, Berlin, Heidelberg, New York, 2015. Springer.
- [4586] J. H. Piater, P. R. Cohen, X. Zhang, and M. Atighetchi. A randomized ANOVA procedure for comparing performance curves. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 430–438. Morgan Kaufmann, San Francisco, CA, 1998.
- [4587] J. H. Piater and R. A. Grupen. Constructive feature learning and the development of visual expertise. In *Proc. 17th International Conf. on Machine Learning*, pages 751–758. Morgan Kaufmann, San Francisco, CA, 2000.
- [4588] A. Piccolboni and C. Schindelhauer. Discrete prediction games with arbitrary feedback and loss. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 208–223. Springer, 2001.
- [4589] D. Picó and F. Casacuberta. Some statistical-estimation methods for stochastic finite-state transducers. *Machine Learning*, 44(1/2):121–141, 2001.
- [4590] D. Picó and E. Vidal. Transducer-learning experiments on language understanding. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 138–149. Springer, 1998.
- [4591] K. Pillaipakkamnatt and V. Raghavan. On the limits of proper learnability of subclasses of DNF formulas. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 118–129, New York, NY, 1994. ACM Press.
- [4592] K. Pillaipakkamnatt and V. Raghavan. Read-twice DNF formulas are properly learnable. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 121–132, Oxford, 1994. Oxford University Press.

- [4593] S. Pinker. Formal models of language learning. *Cognition*, 7:217–283, 1979.
- [4594] S. Pinker. Language acquisition. In *Foundations of Cognitive Science*. The MIT Press, Cambridge, 1990.
- [4595] S. Pinker and A. Prince. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. Technical Report Occasional Paper 33, MIT Center for Cognitive Science, 1987.
- [4596] C. C. Pinter. Complexity of network training for classes of neural networks. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 215–227. Springer, 1995.
- [4597] M. Pintér. On the rate of convergence of error estimates for the partitioning classification rule. *Theoret. Comput. Sci.*, 284(1):181–196, 2002.
- [4598] E. Pippig. Structuring neural networks and PAC learning. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 417–434. Springer, 1995.
- [4599] S. Piramuthu. Effects of feature selection with ‘blurring’ on neurofuzzy systems. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 135–142. Springer, 1996.
- [4600] B. Á. Pires and C. Szepesvári. Policy error bounds for model-based reinforcement learning with factored linear models. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 121–151, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4601] L. Pitt. *Probabilistic Inductive Inference (Ph.D Thesis)*. PhD thesis, Yale University, 1985. Computer Science Dept. TR-400.
- [4602] L. Pitt. Inductive inference, DFAs, and computational complexity. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 18–44. Springer-Verlag, 1989.
- [4603] L. Pitt. Probabilistic inductive inference. *J. ACM*, 36(2):383–433, 1989.
- [4604] L. Pitt. Introduction: Special issue on computational learning theory. *Machine Learning*, 5:117–120, 1990.
- [4605] L. Pitt, editor. *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, New York, N.Y., 1993. ACM Press.
- [4606] L. Pitt. On exploiting knowledge and concept use in learning theory. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 62–84. Springer, 1997.
- [4607] L. Pitt and R. Board. Semi-supervised learning. Technical report, University of Illinois, UIUCDCS-R-87-1372, 1987.
- [4608] L. Pitt and R. Board. On the necessity of Occam algorithms. In *Proc. 23th Annu. ACM Sympos. Theory Comput.*, pages 54–63. ACM Press, 1990.



- [4609] L. Pitt and R. E. Reinke. Criteria for polynomial-time (conceptual) clustering. *Machine Learning*, 2(4):371–396, 1988.
- [4610] L. Pitt and C. H. Smith. Probability and plurality for aggregations of learning machines. *Inform. Comput.*, 77(1):77–92, 1988.
- [4611] L. Pitt and L. G. Valiant. Computational limitations on learning from examples. *J. ACM*, 35(4):965–984, 1988.
- [4612] L. Pitt and M. K. Warmuth. Prediction preserving reducibility. *J. of Comput. Syst. Sci.*, 41(3):430–467, Dec. 1990. Special issue of the for the *Third Annual Conference of Structure in Complexity Theory* (Washington, DC., June 88).
- [4613] L. Pitt and M. K. Warmuth. The minimum consistent DFA problem cannot be approximated within any polynomial. *J. ACM*, 40(1):95–142, 1993.
- [4614] M. Pivar and M. Finkelstein. Automation, using LISP, of inductive inference on sequences. In *The Programming Language LISP*, pages 125–136. MIT Press, 1964.
- [4615] M. Pivar and E. Gord. The LISP program for inductive inference on sequences. In *The Programming Language LISP*, pages 260–289. MIT Press, 1964.
- [4616] G. D. Plotkin. A note on inductive generalization. In *Machine Intelligence 5*, pages 153–163. American Elsevier, 1970.
- [4617] G. D. Plotkin. *Automatic Methods of Inductive Inference*. PhD thesis, Edinburgh University, 1971.
- [4618] G. D. Plotkin. A further note on inductive generalization. In *Machine Intelligence 6*, pages 101–124. American Elsevier, 1971.
- [4619] K. Podnieks. Comparing various concepts of function prediction, Part 1. *Theory of Algorithms and Programs, Latvian State University, Riga*, 210:68–81, 1974.
- [4620] K. Podnieks. Comparing various concepts of function prediction, Part 2. *Theory of Algorithms and Programs, Latvian State University, Riga*, 233, 1975.
- [4621] K. Podnieks. Probabilistic program synthesis. *Theory of Algorithms and Programs, Latvian State University, Riga*, 16:57–88, 1977.
- [4622] K. M. Podnieks. Probabilistic synthesis of enumerated classes of functions. *Soviet Math. Doklady*, 16:1042–1045, 1975.
- [4623] D. Poetschke. Analogical reasoning for second generation expert systems. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 264–276. Springer-Verlag, 1989.
- [4624] T. Poggio and F. Girosi. Networks and the best approximation property. Technical Report AIM1164, MIT Artificial Intelligence Laboratory, Oct. 1989.
- [4625] T. Poggio and F. Girosi. Representation propoerties of networks: Kolmogorov’s theorm is irrelevant. *Neural Computation*, 1:465–469, 1989.
- [4626] T. Poggio and F. Girosi. Extensions of a theory of networks for approximation and learning: dimensionality and reduction and clustering. Technical Report AIM1167, MIT Artificial Intelligence Laboratory, Apr. 1990.

- [4627] T. Poggio and F. Girosi. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 247:978–982, Feb. 23 1990.
- [4628] T. Poggio and F. Girosi. A theory of networks for approximation and learning. *Proceedings of the IEEE*, 78(9), Sept. 1990. An earlier version is MIT TR AIM-1140.
- [4629] J. Poland. FPL analysis for adaptive bandits. In *Stochastic Algorithms: Foundations and Applications, Third International Symposium, SAGA 2005, Moscow, Russia, October 2005, Proceedings*, volume 3777 of *Lecture Notes in Computer Science*, pages 58–69. Springer, 2005.
- [4630] J. Poland. The missing consistency theorem for Bayesian learning: Stochastic model selection. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 259–273. Springer, Oct. 2006.
- [4631] J. Poland. Potential functions for stochastic model selection. Technical Report TCS-TR-A-06-11, Division of Computer Science, Hokkaido University, Mar. 2006.
- [4632] J. Poland. Consistency of discrete Bayesian learning. *Theoret. Comput. Sci.*, 405(3):256–273, 2008. Special issue for ALT 2006.
- [4633] J. Poland. Nonstochastic bandits: Countable decision set, unbounded costs and reactive environments. *Theoret. Comput. Sci.*, 397(1-3):77–93, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [4634] J. Poland and M. Hutter. Convergence of discrete MDL for sequential prediction. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 300–314. Springer, 2004.
- [4635] J. Poland and M. Hutter. On the convergence speed of MDL predictions for Bernoulli sequences. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 294–308. Springer, 2004.
- [4636] J. Poland and M. Hutter. Defensive universal learning with experts. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 356–370. Springer, Oct. 2005.
- [4637] J. Poland and T. Zeugmann. Clustering pairwise distances with missing data: Maximum cuts versus normalized cuts. In *Discovery Science, 9th International Conference, DS 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4265 of *Lecture Notes in Artificial Intelligence*, pages 197–208. Springer, Oct. 2006.
- [4638] J. Poland and T. Zeugmann. Clustering the google distance with eigenvectors and semidefinite programming. In *Knowledge Media Technologies, First International Core-to-Core Workshop*, volume 21 of *Diskussionsbeiträge, Institut für Medien und Kommunikationswissenschaft*, pages 61–69. Technische Universität Ilmenau, 2006.
- [4639] J. B. Pollack. The induction of dynamical recognizers. *Machine Learning*, 7:227–252, 1991.
- [4640] J. B. Pollack and A. D. Blair. Co-evolution in the successful learning of backgammon strategy. *Machine Learning*, 32(3):225–240, 1998.
- [4641] D. Pollard. *Convergence of Stochastic Processes*. Springer-Verlag, 1984.

- [4642] S. Pölt. Improved sample size bounds for PAB-decisions. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 229–239, Oxford, 1994. Oxford University Press.
- [4643] Y. Polyanskiy, A. T. Suresh, and Y. Wu. Sample complexity of population recovery. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1589–1618, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4644] D. A. Pomerleau. Neural network-based vision for precise control of a walking robot. *Machine Learning*, 15(2):125–135, 1994.
- [4645] D. A. Pomerleau. A reply to Towell’s book review of *neural network perception for mobile robot guidance*. *Machine Learning*, 18(1):121–122, 1995.
- [4646] U. Pompe. Noise-tolerant recursive best-first induction. In *Proc. 16th International Conf. on Machine Learning*, pages 315–324. Morgan Kaufmann, San Francisco, CA, 1999.
- [4647] M. Pontil and A. Maurer. Excess risk bounds for multitask learning with trace norm regularization. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 55–76. JMLR.org, 2013.
- [4648] M. Pontil, S. Mukherjee, and F. Girosi. On the noise model of support vector machines regression. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 316–324. Springer, 2000.
- [4649] S. Porat. Stability and looping in connectionist models with asymmetric weights. Technical Report TR 210, University of Rochester Computer Science Department, Mar. 1987.
- [4650] S. Porat and J. A. Feldman. Learning automata from ordered examples. *Machine Learning*, 7:109–138, 1991.
- [4651] B. W. Porter and D. F. Kibler. Experimental goal regression: A method for learning problem-solving heuristics. *Machine Learning*, 1(3):249–285, 1986.
- [4652] S. E. Posner and S. R. Kulkarni. On-line learning of functions of bounded variation under various sampling schemes. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 439–445, New York, NY, 1993. ACM Press.
- [4653] C. Posthoff and M. Schlosser. Optimal strategies – learning from examples – Boolean equations. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 363–390. Springer, 1995.
- [4654] G. Potamias. Content-based similarity assessment in multi-segmented medical image data bases. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 347–361. Springer, 2001.
- [4655] A. Potechin and D. Steurer. Exact tensor completion with sum-of-squares. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1619–1673, Amsterdam, Netherlands, 07–10 July 2017. PMLR.

- [4656] D. Pötschke. Analogical reasoning using graph transformations. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 135–144. Springer-Verlag, 1986.
- [4657] D. Pötschke. Formalizing analogical reasoning. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 215–218. Springer-Verlag, 1986.
- [4658] M. J. D. Powell. Radial basis functions for multivariable interpolation: A review. In *Algorithms for Approximation*, pages 143–167, Oxford, 1987. Clarendon Press.
- [4659] R. Powers, Y. Shoham, and T. Vu. A general criterion and an algorithmic framework for learning in multi-agent systems. *Machine Learning*, 67(1-2):45–76, 2007.
- [4660] L. A. Prashanth. Policy gradients for CVaR-constrained MDPs. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 155–169, Berlin, Heidelberg, New York, 2014. Springer.
- [4661] L. Y. Pratt. Experiments on the transfer of knowledge between neural networks. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 19, pages 523–560. MIT Press, 1994.
- [4662] D. Precup and R. S. Sutton. Exponentiated gradient methods for reinforcement learning. In *Proc. 14th International Conference on Machine Learning*, pages 272–277. Morgan Kaufmann, 1997.
- [4663] D. Precup, R. S. Sutton, and S. Dasgupta. Off-policy temporal-difference learning with function approximation. In *Proc. 18th International Conf. on Machine Learning*, pages 417–424. Morgan Kaufmann, San Francisco, CA, 2001.
- [4664] D. Precup, R. S. Sutton, and S. Singh. Eligibility traces for off-policy policy evaluation. In *Proc. 17th International Conf. on Machine Learning*, pages 759–766. Morgan Kaufmann, San Francisco, CA, 2000.
- [4665] D. Precup and P. E. Utgoff. Classification using  $\Phi$ -machines and constructive function approximation. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 439–444. Morgan Kaufmann, San Francisco, CA, 1998.
- [4666] J. B. Predd, S. R. Kulkarni, and H. V. Poor. Consistency in models for communication constrained distributed learning. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 442–456. Springer, 2004.
- [4667] B. Price and C. Boutilier. Implicit imitation in multiagent reinforcement learning. In *Proc. 16th International Conf. on Machine Learning*, pages 325–334. Morgan Kaufmann, San Francisco, CA, 1999.
- [4668] A. E. Prieditis. Machine discovery of effective admissible heuristics. *Machine Learning*, 12:117–141, 1993.
- [4669] F. Provost and P. Domingos. Tree induction for probability-based ranking. *Machine Learning*, 52(3):199–215, 2003.

- [4670] F. Provost and T. Fawcett. Robust classification for imprecise environments. *Machine Learning*, 42(3):203–231, 2001.
- [4671] F. Provost, T. Fawcett, and R. Kohavi. The case against accuracy estimation for comparing induction algorithms. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 445–453. Morgan Kaufmann, San Francisco, CA, 1998.
- [4672] F. J. Provost and J. M. Aronis. Scaling up inductive learning with massive parallelism. *Machine Learning*, 23:33–46, 1996.
- [4673] F. J. Provost and B. G. Buchanan. Inductive strengthening: The effects of a simple heuristic for restricting hypothesis space search. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 294–304. Springer-Verlag, 1992.
- [4674] J. F. Provost and B. G. Buchanan. Inductive policy: The pragmatics of bias selection. *Machine Learning*, 20:35–61, 1995.
- [4675] D. Pucci de Farias and B. Van Roy. Fixed points of approximate value iteration and temporal-difference learning. In *Proc. 17th International Conf. on Machine Learning*, pages 207–214. Morgan Kaufmann, San Francisco, CA, 2000.
- [4676] P. Pudlak and F. N. Springsteel. Complexity of mechanized hypothesis formation. *Theoret. Comput. Sci.*, 8:203–225, 1979.
- [4677] J.-F. Puget. Explicit representation of concept negation. *Machine Learning*, 14:233–247, 1994.
- [4678] P. W. Purdom, D. V. Gucht, and D. P. Groth. Average-case performance of the apriori algorithm. *SIAM J. Comput.*, 33(5):1223–1260, 2004.
- [4679] H. Putnam. Probability and confirmation. In *Mathematics, Matter and Method*. Cambridge University Press, 1975.
- [4680] S. Qian, Y. C. Lee, R. D. Jones, C. W. Barnes, G. W. Flake, M. K. O'Rourke, K. Lee, H. H. Chen, G. Z. Sun, Y. Q. Zhang, D. Chen, and C. L. Giles. Adaptive stochastic cellular automata: Theory. *Physica D*, 666, 1990.
- [4681] H. Qiao, N. S. V. Rao, and V. Protopopescu. PAC learning using Nadaraya-Watson estimator based on orthonormal systems. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 146–160. Springer, 1997.
- [4682] J. R. Quinlan. Inferno: A cautious approach to uncertain inference. *The Computer Journal*, 26(3):255–269, 1983.
- [4683] J. R. Quinlan. The effect of noise on concept learning. In *Machine Learning, An Artificial Intelligence Approach (Volume II)*, chapter 6, pages 149–166. Morgan Kaufmann, 1986.
- [4684] J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, 1986.
- [4685] J. R. Quinlan. Simplifying decision trees. *International Journal of Man-Machine Studies*, 27(3):221–234, 1987.
- [4686] J. R. Quinlan. Learning logical definitions from relations. *Machine Learning*, 5:239–266, 1990.

- [4687] J. R. Quinlan. Determinate literals in inductive logic programming. In *Machine Learning: Proceedings of the Eighth International Workshop*, San Mateo, CA, 1991. Morgan Kaufmann.
- [4688] J. R. Quinlan. Improved estimates for the accuracy of small disjuncts. *Machine Learning*, 6:93–98, 1991.
- [4689] J. R. Quinlan. *C4.5: Programs for machine learning*. Morgan Kaufmann, 1993.
- [4690] J. R. Quinlan. Comparing connectionist and symbolic learning methods. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 15, pages 445–456. MIT Press, 1994.
- [4691] J. R. Quinlan. The minimum description length principle and categorical theories. In *Proc. 11th International Conference on Machine Learning*, pages 233–241. Morgan Kaufmann, 1994.
- [4692] J. R. Quinlan. MDL and categorical theories (continued). In *Proc. 12th International Conference on Machine Learning*, pages 464–470. Morgan Kaufmann, 1995.
- [4693] J. R. Quinlan. Boosting first-order learning. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 143–155. Springer, 1996.
- [4694] J. R. Quinlan. Some elements of machine learning. In *Proc. 16th International Conf. on Machine Learning*, pages 523–525. Morgan Kaufmann, San Francisco, CA, 1999.
- [4695] J. R. Quinlan and R. M. Cameron-Jones. FOIL: A midterm report. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 3–20. Springer-Verlag, 1993.
- [4696] J. R. Quinlan and R. L. Rivest. Inferring decision trees using the Minimum Description Length Principle. *Inform. Comput.*, 80(3):227–248, Mar. 1989. (An early version appeared as MIT LCS Technical report MIT/LCS/TM-339 (September 1987).).
- [4697] M. Quist and G. Yona. Distributional scaling: An algorithm for structure-preserving embedding of metric and nonmetric spaces. *Journal of Machine Learning Research*, 5:399–420, 2004.
- [4698] J. del R. Millán, D. Posenato, and E. Dedieu. Continuous-action Q-learning. *Machine Learning*, 49(2-3):247–265, 2002.
- [4699] L. R. Rabiner and B. H. Juang. An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1):4–16, Jan. 1986.
- [4700] L. R. Rabiner, S. E. Levinson, and M. M. Sondhi. On the application of vector quantization and hidden Markov models to speaker-independent, isolated word recognition. *Bell System Technical Journal*, 62(4):1075–1105, Apr. 1983.
- [4701] G. Rabusseau and F. Denis. Maximizing a tree series in the representation space. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 124–138, Kyoto, Japan, 17–19 Sept. 2014. PMLR.

- [4702] J. Rachlin, S. Kasif, S. Salzberg, and D. W. Aha. Towards a better understanding of memory-based reasoning systems. In *Proc. 11th International Conference on Machine Learning*, pages 242–250. Morgan Kaufmann, 1994.
- [4703] L. D. Raedt and K. Kersting. Probabilistic inductive logic programming. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 19–36. Springer, 2004.
- [4704] H. Ragavan and L. Rendell. Learning disjunctive concepts using domain knowledge. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 6, pages 143–161. MIT Press, 1994.
- [4705] T. Ragg. Building committees by clustering models based on pairwise similarity values. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 406–418. Springer, 2001.
- [4706] P. Raghavan. Learning in threshold networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 19–27, San Mateo, CA, 1988. Morgan Kaufmann.
- [4707] V. Raghavan. Bounded degree graph inference from walks. *J. of Comput. Syst. Sci.*, 49(1):108–132, 1994. Earlier version in 5th Colt, 1991.
- [4708] V. Raghavan and S. R. Schach. Learning switch configurations. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 38–51, San Mateo, CA, 1990. Morgan Kaufmann.
- [4709] V. Raghavan and D. Wilkins. Learning  $\mu$ -branching programs with queries. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 27–36, New York, NY, 1993. ACM Press.
- [4710] M. Raginsky, A. Rakhlin, and M. Telgarsky. Non-convex learning via stochastic gradient langevin dynamics: a nonasymptotic analysis. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1674–1703, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4711] H. Rahmanian, D. P. Helmbold, and S. V. N. Vishwanathan. Online learning of combinatorial objects via extended formulation. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 702–724. PMLR, 07–09 Apr. 2018.
- [4712] S. A. Rajamoney. The design of discrimination experiments. *Machine Learning*, 12:185–203, 1993.
- [4713] A. Rajkumar and S. Agarwal. When can we rank well from comparisons of  $O(n \log(n))$  non-actively chosen pairs? In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1376–1401, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4714] A. Rakhlin. Foreword. *J. Comput. Syst. Sci.*, 78(5):1403, 2012.

- [4715] A. Rakhlin and K. Sridharan. Online learning with predictable sequences. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 993–1019. JMLR.org, 2013.
- [4716] A. Rakhlin and K. Sridharan. Online non-parametric regression. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1232–1264, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4717] A. Rakhlin and K. Sridharan. Hierarchies of relaxations for online prediction problems with evolving constraints. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1457–1479, Paris, France, 03–06 July 2015. PMLR.
- [4718] A. Rakhlin and K. Sridharan. On equivalence of martingale tail bounds and deterministic regret inequalities. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1704–1722, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [4719] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Beyond regret. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 559–594, Budapest, Hungary, 09–11 June 2011. PMLR.
- [4720] A. Rakotomamonjy. Variable selection using SVM-based criteria. *Journal of Machine Learning Research*, 3:1357–1370, 2003.
- [4721] A. Ram. Indexing and elaboration and refinement: Incremental learning of explanatory cases. *Machine Learning*, 10:201–248, 1993.
- [4722] S. Ramachandran and R. J. Mooney. Theory refinement for Bayesian networks with hidden variables. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 454–462. Morgan Kaufmann, San Francisco, CA, 1998.
- [4723] U. Ramacher. Hamiltonian dynamics of neural networks. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 16, pages 251–277. MIT Press, 1994.
- [4724] K. Ramamohanarao. Contrast pattern mining and its application for building robust classifiers. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, page 33, Berlin, Heidelberg, New York, 2010. Springer.
- [4725] H. G. Ramaswamy, B. S. Babu, S. Agarwal, and R. C. Williamson. On the consistency of output code based learning algorithms for multiclass learning problems. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 885–902, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4726] A. Ramdas and A. Singh. Algorithmic connections between active learning and stochastic convex optimization. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9,*



- 2013, *Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 339–353, Berlin, Heidelberg, New York, 2013. Springer.
- [4727] J. Ramon and C. Comendant. Open problem: Learning dynamic network models from a static snapshot. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 45.1–45.3, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4728] M. Ramoni and P. Sebastiani. Robust learning with missing data. *Machine Learning*, 45(2):147–170, 2001.
- [4729] M. Ramoni, P. Sebastiani, and P. Cohen. Bayesian clustering by dynamics. *Machine Learning*, 47(1):91–121, 2002.
- [4730] J. Randoøv. Shaping in reinforcement learning by changing the physics of the problem. In *Proc. 17th International Conf. on Machine Learning*, pages 767–774. Morgan Kaufmann, San Francisco, CA, 2000.
- [4731] J. Randoøv and P. Alstrøm. Learning to drive a bicycle using reinforcement learning and shaping. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 463–471. Morgan Kaufmann, San Francisco, CA, 1998.
- [4732] J. Randoøv, A. G. Barto, and M. T. Rosenstein. Combining reinforcement learning with a local control algorithm. In *Proc. 17th International Conf. on Machine Learning*, pages 775–782. Morgan Kaufmann, San Francisco, CA, 2000.
- [4733] M. R. K. K. Rao. Inductive inference of term rewriting systems from positive data. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 69–82. Springer, 2004.
- [4734] M. R. K. K. Rao. Some classes of term rewriting systems inferable from positive data. *Theoret. Comput. Sci.*, 397(1-3):129–149, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [4735] N. S. V. Rao and E. M. Oblow. N-learners problem: system of PAC learners. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 12, pages 189–210. MIT Press, 1997.
- [4736] N. S. V. Rao, E. M. Oblow, C. W. Glover, and G. E. Liepins. N-learners problem: Fusion of concepts. Technical Report ORNL/TM-11904, CESAR-91/23, Oak Ridge National Laboratory, Sept. 1991.
- [4737] R. B. Rao, D. Gordon, and W. Spears. For every generalization action is there really an equal and opposite reaction? analysis of the conservation law for generalization performance. In *Proc. 12th International Conference on Machine Learning*, pages 471–479. Morgan Kaufmann, 1995.
- [4738] C. Raphael. Aligning music audio with symbolic scores using a hybrid graphical model. *Machine Learning*, 65(2-3):389–409, 2006.
- [4739] B. Raskutti, H. Ferrá, and A. Kowalczyk. Second order features for maximising text classification performance. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 419–430. Springer, 2001.

- [4740] A. Ratnaparkhi. Learning to parse natural language with maximum entropy models. *Machine Learning*, 34:151–175, 1999.
- [4741] J. Ratsaby. On learning multicategory classification with sample queries. *Inform. Comput.*, 185(2):298–327, 2003.
- [4742] J. Ratsaby. A stochastic gradient descent algorithm for structural risk minimisation. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 205–220. Springer, 2003.
- [4743] J. Ratsaby and V. Maiorov. Generalization of the PAC-model for learning with partial information. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 51–65. Springer, 1997.
- [4744] J. Ratsaby and V. Maiorov. On the learnability of rich function classes. *J. of Comput. Syst. Sci.*, 58(1):183–192, 1999.
- [4745] J. Ratsaby, R. Meir, and V. Maiorov. Towards robust model selection using estimation and approximation error bounds. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 57–67, New York, NY, 1996. ACM Press.
- [4746] J. Ratsaby and S. S. Venkatesh. Learning from a mixture of labeled and unlabeled examples with parametric side information. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 412–417, New York, NY, 1995. ACM Press.
- [4747] G. Rätsch. Solving semi-infinite linear programs using boosting-like methods. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 10–11. Springer, Oct. 2006.
- [4748] G. Rätsch, A. Demiriz, and K. P. Bennett. Sparse regression ensembles in infinite and finite hypothesis spaces. *Machine Learning*, 48(1/3):189–218, 2002.
- [4749] G. Rätsch, T. Onoda, and K.-R. Müller. Soft margins for AdaBoost. *Machine Learning*, 42(3):287–320, 2001.
- [4750] G. Rätsch, M. Warmuth, S. Mika, T. Onoda, S. Lemm, and K. R. Müller. Barrier boosting. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 170–179. Morgan Kaufmann, San Francisco, 2000.
- [4751] G. Rätsch and M. K. Warmuth. Maximizing the margin with boosting. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 334–350. Springer, 2002.
- [4752] C. Ravi's'e and M. Sebag. An advanced evolution should not repeat its past errors. In *Proc. 13th International Conference on Machine Learning*, pages 400–408. Morgan Kaufmann, 1996.
- [4753] S. Ray and D. Page. Multiple instance regression. In *Proc. 18th International Conf. on Machine Learning*, pages 425–432. Morgan Kaufmann, San Francisco, CA, 2001.

- [4754] S. Raychaudhuri, H. Schütze, and R. B. Altman. Inclusion of textual documentation in the analysis of multidimensional data sets: Application to gene expression data. *Machine Learning*, 52(1-2):119–145, 2003.
- [4755] P. Rebeschini and A. Karbasi. Fast mixing for discrete point processes. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1480–1500, Paris, France, 03–06 July 2015. PMLR.
- [4756] B. Recht and C. Re. Toward a noncommutative arithmetic-geometric mean inequality: Conjectures, case-studies, and consequences. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 11.1–11.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4757] C. Reddy and P. Tadepalli. Learning goal-decomposition rules using exercises. In *Proc. 14th International Conference on Machine Learning*, pages 278–286. Morgan Kaufmann, 1997.
- [4758] C. Reddy and P. Tadepalli. Learning first-order acyclic Horn programs from entailment. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 472–480. Morgan Kaufmann, San Francisco, CA, 1998.
- [4759] C. Reddy and P. Tadepalli. Learning Horn definitions: Theory and an application to planning. *New Generation Computing*, 17(1):77–98, 1999.
- [4760] C. Reddy, P. Tadepalli, and S. Roncagliolo. Theory-guided empirical speedup learning of goal decomposition rules. In *Proc. 13th International Conference on Machine Learning*, pages 409–417. Morgan Kaufmann, 1996.
- [4761] R. A. Redner and H. F. Walker. Mixture densities, maximum likelihood, and the EM algorithm. *Siam Review*, 26:195–239, 1984.
- [4762] G. N. Reeke, Jr, O. Sporns, and G. M. Edelman. Synthetic neural modelling: Comparisons of population and connectionist approaches. In *Connectionism in Perspective*. Elsevier Science B. V. (North-Holland), 1989.
- [4763] H. Reeve, J. Mellor, and G. Brown. The K-nearest neighbour UCB algorithm for multi-armed bandits with covariates. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 725–752. PMLR, 07–09 Apr. 2018.
- [4764] H. W. J. Reeve and G. Brown. Minimax rates for cost-sensitive learning on manifolds with approximate nearest neighbours. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 11–56, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4765] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93. ACM Press, 2005.
- [4766] Y. Reich. Exemplar-based knowledge acquisition. *Machine Learning*, 6:99–103, 1991.

- [4767] M. D. Reid, R. M. Frongillo, R. C. Williamson, and N. Mehta. Generalized mixability via entropic duality. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1501–1522, Paris, France, 03–06 July 2015. PMLR.
- [4768] D. Reidenbach. A negative result on inductive inference of extended pattern languages. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 308–320. Springer, 2002.
- [4769] D. Reidenbach. A discontinuity in pattern inference. In *STACS 2004, 21st Annual Symposium on Theoretical Aspects of Computer Science, Montpellier, France, March 25-27, 2004, Proceedings*, volume 2996 of *Lecture Notes in Computer Science*, pages 129–140. Springer, 2004.
- [4770] D. Reidenbach. On the learnability of E-pattern languages over small alphabets. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 140–154. Springer, 2004.
- [4771] D. Reidenbach. *The Ambiguity of Morphisms in Free Monoids and its Impact on Algorithmic Properties of Pattern Languages*. PhD thesis, Department of Computer Science, University of Kaiserslautern, Germany, 2006.
- [4772] D. Reidenbach. A non-learnable class of E-pattern languages. *Theoret. Comput. Sci.*, 350(1):91–102, 2006. Special issue for ALT 2002.
- [4773] D. Reidenbach. Discontinuities in pattern inference. *Theoret. Comput. Sci.*, 397(1-3):166–193, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [4774] D. Reidenbach and M. L. Schmid. Patterns with bounded treewidth. *Inform. Comput.*, 239:87–99, 2014.
- [4775] R. Reischuk and T. Zeugmann. Analyzing the average-case behavior of conjunctive learning algorithms. Technical Report DOI-TR-153, Department of Informatics, Kyushu University, Aug. 1998.
- [4776] R. Reischuk and T. Zeugmann. Learning one-variable pattern languages in linear average time. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 198–208, New York, NY, 1998. ACM Press.
- [4777] R. Reischuk and T. Zeugmann. A complete and tight average-case analysis of learning monomials. In *STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science, Trier, Germany, March 1999, Proceedings*, volume 1563 of *Lecture Notes in Computer Science*, pages 414–423. Springer, 1999.
- [4778] R. Reischuk and T. Zeugmann. An average-case optimal one-variable pattern language learner. *J. Comput. Syst. Sci.*, 60(2):302–335, 2000. Special issue for COLT '98.
- [4779] L. Rendell. A general framework for induction and a study of selective induction. *Machine Learning*, 1(2):177–226, 1986.
- [4780] L. Rendell and H. Cho. Empirical learning as a function of concept character. *Machine Learning*, 5:267–298, 1990.

- [4781] L. Rendell and R. Seshu. Learning hard concepts through constructive induction: framework and rationale. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 5, pages 83–141. MIT Press, 1994.
- [4782] J. Rennie and A. K. McCallum. Using reinforcement learning to spider the web efficiently. In *Proc. 16th International Conf. on Machine Learning*, pages 335–343. Morgan Kaufmann, San Francisco, CA, 1999.
- [4783] J. Reunanen. Overfitting in making comparisons between variable selection methods. *Journal of Machine Learning Research*, 3:1371–1382, 2003.
- [4784] S. I. Reynolds. Adaptive resolution model-free reinforcement learning: Decision boundary partitioning. In *Proc. 17th International Conf. on Machine Learning*, pages 783–790. Morgan Kaufmann, San Francisco, CA, 2000.
- [4785] L. Reyzin. Data stability in clustering: A closer look. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 184–198, Berlin, Heidelberg, New York, 2012. Springer.
- [4786] L. Reyzin and N. Srivastava. Learning and verifying graphs using queries with a focus on edge counting. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 285–297, Berlin, Oct. 2007. Springer.
- [4787] B. L. Richards and R. J. Mooney. Automated refinement of first-order horn-clause domain theories. *Machine Learning*, 19(2):95–131, 1995.
- [4788] C. Richter and J. Stachowiak. Knowledge propagation in model-based reinforcement learning tasks. In *Proc. 17th International Conf. on Machine Learning*, pages 791–798. Morgan Kaufmann, San Francisco, CA, 2000.
- [4789] M. M. Richter. A unified approach to inductive logic and case-based reasoning. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 210–210. Springer-Verlag, 1994.
- [4790] M. M. Richter, C. H. Smith, R. Wiehagen, and T. Zeugmann, editors. *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Oct. 1998.
- [4791] M. M. Richter, C. H. Smith, R. Wiehagen, and T. Zeugmann. Editors' introduction. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 1–10. Springer, 1998.
- [4792] J. R. Rico-Juan, J. Calera-Rubio, and R. C. Carrasco. Probabilistic  $k$ -testable tree languages. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 221–228. Springer, 2000.

- [4793] J. R. Rico-Juan, J. Calera-Rubio, and R. C. Carrasco. Stochastic  $k$ -testable tree languages and applications. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 199–212. Springer, 2002.
- [4794] R. Rifkin and A. Klautau. In defense of one-vs-all classification. *Journal of Machine Learning Research*, 5:101–141, 2004.
- [4795] R. Rifkin, M. Pontil, and A. Verri. A note on support vector machine degeneracy. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 252–263. Springer, 1999.
- [4796] P. Rigollet and X. Tong. Neyman-pearson classification under a strict constraint. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 595–614, Budapest, Hungary, 09–11 June 2011. PMLR.
- [4797] M. B. Ring. CHILD: a first step towards continual learning. *Machine Learning*, 28(1):77–104, 1997.
- [4798] S. Risau-Gusman and M. B. Gordon. Hierarchical learning in polynomial support vector machines. *Machine Learning*, 46(1-3):53–70, 2002.
- [4799] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.
- [4800] J. Rissanen. A universal prior for integers and estimation by minimum description length. *Ann. of Statist.*, 11(2):416–431, 1983.
- [4801] J. Rissanen. Minimum description length principle. In S. Kotz and N. L. Johnson, editors, *Encyclopedia of Statistical Sciences*, 5, pages 523–527, New York, 1985. Wiley.
- [4802] J. Rissanen. Stochastic complexity and modeling. *Ann. of Statist.*, 14(3):1080–1100, 1986.
- [4803] J. Rissanen. Stochastic complexity and sufficient statistics. Technical report, IBM Research Laboratory (San Jose), 1986.
- [4804] J. Rissanen. *Stochastic Complexity in Statistical Inquiry*, volume 15 of *Series in Computer Science*. World Scientific, 1989.
- [4805] J. Rissanen. Stochastic complexity in learning. *J. Comput. Syst. Sci.*, 55(1):89–95, 1997. Special issue for EuroCOLT '95.
- [4806] J. Rissanen. Optimal estimation. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 37–37, Berlin, Heidelberg, New York, 2011. Springer.
- [4807] E. S. Ristad and P. N. Yianilos. Learning string edit distance. In *Proc. 14th International Conference on Machine Learning*, pages 287–295. Morgan Kaufmann, 1997.
- [4808] A. Risteski. How to calculate partition functions using convex programming hierarchies: provable bounds for variational methods. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1402–1416, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [4809] G. Ritschard. Exploring sequential data. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 35–35, Berlin, Heidelberg, New York, 2012. Springer.
- [4810] I. Rivals and L. Personnaz. MLPs (mono-layer polynomials and multi-layer perceptrons) for nonlinear modeling. *Journal of Machine Learning Research*, 3:1383–1398, 2003.
- [4811] R. Rivest, D. Haussler, and M. K. Warmuth, editors. *Proceedings of the Second Annual Workshop on Computational Learning Theory*, San Mateo, CA, 1989. Morgan Kaufmann, Inc.
- [4812] R. Rivest and R. Schapire. Inference of visible simple assignment automata with planned experiments. Preprint, MIT Laboratory for Computer Science., 1987.
- [4813] R. L. Rivest. Learning decision lists. *Machine Learning*, 2(3):229–246, 1987.
- [4814] R. L. Rivest. Theory of learning: What’s hard and what’s easy to learn. In *Research Directions in Computer Science: An MIT Perspective*, chapter 11, pages 217–227. MIT Press, 1991.
- [4815] R. L. Rivest and W. Remmele. Machine learning: the human connection. *Siemens Review*, 55(2/88):33–38, Mar./Apr. 1988.
- [4816] R. L. Rivest and W. Remmele. Machine learning. In *Applied Computer Science and Software: Turing Theory into Practice*, pages 186–200. Springer-Verlag, 1991.
- [4817] R. L. Rivest and R. E. Schapire. Diversity-based inference of finite automata. In *Proc. 28th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 78–87. IEEE Computer Society Press, Los Alamitos, CA, 1987.
- [4818] R. L. Rivest and R. E. Schapire. A new approach to unsupervised learning in deterministic environments. In *Proc. of the 4th International Workshop on Machine Learning*, pages 364–375, San Mateo, CA, June 1987. Morgan Kaufmann.
- [4819] R. L. Rivest and R. E. Schapire. A new approach to unsupervised learning in deterministic environments (reprint). In *Machine Learning, An Artificial Intelligence Approach*, volume III, pages 670–684. Morgan Kaufmann, 1990.
- [4820] R. L. Rivest and R. E. Schapire. Inference of finite automata using homing sequences. *Inform. Comput.*, 103(2):299–347, Apr. 1993.
- [4821] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public key cryptosystems. *Commun. ACM*, 21(2):120–126, 1978.
- [4822] R. L. Rivest and R. Sloan. Learning complicated concepts reliably and usefully. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 69–79, San Mateo, CA, 1988. Morgan Kaufmann.
- [4823] R. L. Rivest and R. Sloan. A new model for inductive inference. In *Proc. of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 13–27. Morgan Kaufmann, Mar. 1988.
- [4824] R. L. Rivest and R. Sloan. A formal model of hierarchical concept-learning. *Inform. Comput.*, 114(1):88–114, 1994.

- [4825] R. L. Rivest and R. H. Sloan. On choosing between experimenting and thinking when learning. *Inform. Comput.*, 106(1):1–25, 1993.
- [4826] R. L. Rivest and P. Winston. Machine learning research at MIT. *Siemens Review*, pages 12–14, 1990.
- [4827] R. L. Rivest and Y. Yin. Simulation results for a new two-armed bandit heuristic. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 17, pages 477–486. MIT Press, 1994. Earlier version in 1990 Conference on Computation Learning and Natural Learning at Princeton.
- [4828] R. L. Rivest and Y. L. Yin. Being taught can be faster than asking questions. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 144–151, New York, NY, 1995. ACM Press.
- [4829] G. G. Robertson and R. L. Riolo. A tale of two classifier systems. *Machine Learning*, 3:139–159, 1988.
- [4830] M. Robnik-Šikonja and I. Kononenko. An adaptation of Relief for attribute estimation in regression. In *Proc. 14th International Conference on Machine Learning*, pages 296–304. Morgan Kaufmann, 1997.
- [4831] M. Robnik-Šikonja and I. Kononenko. Attribute dependencies, understandability and split selection in tree based models. In *Proc. 16th International Conf. on Machine Learning*, pages 344–353. Morgan Kaufmann, San Francisco, CA, 1999.
- [4832] M. Robnik-Šikonja and I. Kononenko. Comprehensible interpretation of relief’s estimates. In *Proc. 18th International Conf. on Machine Learning*, pages 433–440. Morgan Kaufmann, San Francisco, CA, 2001.
- [4833] M. Robnik-Sikonja and I. Kononenko. Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, 53(1-2):23–69, 2003.
- [4834] K. Romanik. Testing as A dual to learning. Technical Report UMIACS-TR-91.93, CS-TR-2704, Dept. of Computer Science, University of Maryland, June 1991.
- [4835] K. Romanik. Approximate testing and learnability. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 327–332, New York, NY, 1992. ACM Press.
- [4836] K. Romanik. Approximate testing and its relationship to learning. *Theoret. Comput. Sci.*, 188(1-2):175–194, 1997.
- [4837] K. Romanik and J. S. Vitter. Using Vapnik-Chervonenkis dimension to analyze the testing complexity of program segments. *Inform. Comput.*, 128(2):87–108, Aug. 1996.
- [4838] D. Ron. Property testing: A learning theory perspective. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 1–2, Berlin, 2007. Springer.
- [4839] D. Ron and R. Rubinfeld. Learning fallible deterministic finite automata. *Machine Learning*, 18(2/3):149–185, 1995. Special issue for COLT ’93.
- [4840] D. Ron and R. Rubinfeld. Exactly learning automata of small cover time. *Machine Learning*, 27(1):69–96, 1997.



- [4841] D. Ron, Y. Singer, and N. Tishby. Learning probabilistic automata with variable memory length. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 35–46, New York, NY, 1994. ACM Press.
- [4842] D. Ron, Y. Singer, and N. Tishby. On the learnability and usage of acyclic probabilistic finite automata. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 31–40, New York, NY, 1995. ACM Press.
- [4843] D. Ron, Y. Singer, and N. Tishby. The power of amnesia: Learning probabilistic automata with variable memory length. *Machine Learning*, 25(2-3):117–149, 1996.
- [4844] D. Ron, Y. Singer, and N. Tishby. On the learnability and usage of acyclic probabilistic finite automata. *J. of Comput. Syst. Sci.*, 56(2):133–152, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [4845] J. P. Rosca and D. H. Ballard. Hierarchical self-organization in genetic programming. In *Proc. 11th International Conference on Machine Learning*, pages 251–258. Morgan Kaufmann, 1994.
- [4846] M. Röscheisen, R. Hoffman, and V. Tresp. Incorporating prior knowledge into networks of locally-tuned units. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 4, pages 53–64. MIT Press, 1994.
- [4847] D. Rose and P. Langley. Chemical discovery as belief revision. *Machine Learning*, 1(4):423–452, 1986.
- [4848] C. Rosenberg. Image color constancy using EM and cached statistics. In *Proc. 17th International Conf. on Machine Learning*, pages 799–806. Morgan Kaufmann, San Francisco, CA, 2000.
- [4849] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psych. Rev.*, 65:386–407, 1958. (Reprinted in *Neurocomputing* (MIT Press, 1988).).
- [4850] C. D. Rosin. Sample complexity of model-based search. *J. Comput. Syst. Sci.*, 60(2):278–301, 2000. Special issue for COLT '98.
- [4851] C. D. Rosin and R. K. Belew. A competitive approach to game learning. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 292–302, New York, NY, 1996. ACM Press.
- [4852] R. Rosipal and L. J. Trejo. Kernel partial least squares regression in reproducing kernel hilbert space. *Journal of Machine Learning Research*, 2:97–123, 2001.
- [4853] S. Rosset, G. Swirszcz, N. Srebro, and J. Zhu.  $\ell_1$  regularization in infinite dimensional feature spaces. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 544–558, Berlin, 2007. Springer.
- [4854] F. Rossi and A. Sperduti. Learning solution preferences in constraint problems. *J. of Experimental and Theoret. Artif. Intell.*, 10(1):103–116, 1998.
- [4855] P. Rossmanith. Learning from random text. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 132–144. Springer, 1999.

- [4856] P. Rossmanith and T. Zeugmann. Learning  $k$ -variable pattern languages efficiently stochastically finite on average from positive data. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 13–24. Springer, 1998.
- [4857] P. Rossmanith and T. Zeugmann. Stochastic finite learning of the pattern languages. *Machine Learning*, 44(1/2):67–91, 2001.
- [4858] C. Rouveirol. Flattening and saturation: Two representation changes for generalization. *Machine Learning*, 14:219–232, 1994.
- [4859] D. Roverso. Analogy access by mapping spreading and abstraction in large, multifunctional knowledge bases. In *Proc. 13th International Conference on Machine Learning*, pages 418–426. Morgan Kaufmann, 1996.
- [4860] N. Roy and A. McCallum. Toward optimal active learning through sampling estimation of error reduction. In *Proc. 18th International Conf. on Machine Learning*, pages 441–448. Morgan Kaufmann, San Francisco, CA, 2001.
- [4861] V. Roychowdhury, K. Siu, A. Orlitsky, and T. Kailath. A geometric approach to threshold circuit complexity. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 97–111, San Mateo, CA, 1991. Morgan Kaufmann.
- [4862] H. Royden. *Real Analysis*. Macmillan, New York, NY, 1963.
- [4863] J. S. Royer. Inductive inference of approximations. *Inform. Control*, 70(2/3):156–178, 1986.
- [4864] A. Rozsypal and M. Kubat. Using the genetic algorithm to reduce the size of a nearest-neighbor classifier and to select relevant attributes. In *Proc. 18th International Conf. on Machine Learning*, pages 449–456. Morgan Kaufmann, San Francisco, CA, 2001.
- [4865] J. H. Rubinstein and B. I. P. Rubinstein. Geometric & topological representations of maximum classes with applications to sample compression. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 299–310. Omnipress, 2008.
- [4866] M. Rudelson and S. Zhou. Reconstruction from anisotropic random measurements. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 10.1–10.24, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4867] S. Rudich. Inferring the structure of a Markov chain from its output. In *Proc. 27th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 321–326, Oct. 1985.
- [4868] C. Rudin. Ranking with a P-norm push. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 589–604, Berlin, 2006. Springer.
- [4869] C. Rudin, C. Cortes, M. Mohri, and R. E. Schapire. Margin-based ranking meets boosting in the middle. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 63–78. Springer, 2005.
- [4870] C. Rudin, B. Letham, A. Sallab-Aouissi, E. Kogan, and D. Madigan. Sequential event prediction with association rules. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 615–634, Budapest, Hungary, 09–11 June 2011. PMLR.

- [4871] C. Rudin, R. E. Schapire, and I. Daubechies. Boosting based on a smooth margin. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 502–517. Springer, 2004.
- [4872] C. Rudin, R. E. Schapire, and I. Daubechies. Open problem: Does adaboost always cycle? In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 46.1–46.4, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [4873] S. M. Ruger. A class of asymptotically stable algorithms for learning-rate adaptation. *Algorithmica*, 22(1/2):198–210, 1998.
- [4874] J. Ruiz, S. Espa˜na, and P. Garca. Locally threshold testable languages in strict sense: Application to the inference problem. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 150–161. Springer, 1998.
- [4875] J. Ruiz and P. Garcia. Learning  $k$ -piecewise testable languages from positive data. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 203–210. Springer, 1996.
- [4876] D. E. Rumelhart, G. E. Hinton, and J. L. McClelland. A general framework for parallel distributed processing. In *Parallel Distributed Processing (Volume I: Foundations)*, chapter 2, pages 45–76. MIT Press, 1986.
- [4877] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing – Explorations in the Microstructure of Cognition*, chapter 8, pages 318–362. MIT Press, 1986.
- [4878] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.
- [4879] D. E. Rumelhart and J. L. McClelland, editors. *Parallel Distributed Processing (Volume I: Foundations)*. MIT Press, 1986.
- [4880] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group. *Parallel Distributed Processing – Explorations in the Microstructure of Cognition*, volume 2. MIT Press, Cambridge, Massachusetts, 1986.
- [4881] D. E. Rumelhart and D. Zipser. Feature discovery by competitive learning. *Cognitive Science*, 9:75–112, 1985.
- [4882] D. Rusakov and D. Geiger. Asymptotic model selection for naive Bayesian networks. *Journal of Machine Learning Research*, 6:1–35, 2005.
- [4883] S. Russell. Learning agents for uncertain environments. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 101–103, New York, NY, 1998. ACM Press.
- [4884] D. Russo. Simple Bayesian algorithms for best arm identification. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1417–1418, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [4885] D. Ryabko. Application of classical nonparametric predictors to learning conditionally I.I.D. data. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 171–180. Springer, 2004.
- [4886] D. Ryabko. On computability of pattern recognition problems. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 148–156. Springer, Oct. 2005.
- [4887] D. Ryabko. Some sufficient conditions on an arbitrary class of stochastic processes for the existence of a predictor. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 169–182, Berlin, Oct. 2008. Springer.
- [4888] D. Ryabko. Unsupervised model-free representation learning. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 354–366, Berlin, Heidelberg, New York, 2013. Springer.
- [4889] D. Ryabko. Things Bayes can’t do. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 253–260. Springer, 2016.
- [4890] D. Ryabko. Hypotheses testing on infinite random graphs. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 400–411, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4891] D. Ryabko. Universality of Bayesian mixture predictors. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 57–71, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [4892] D. Ryabko and M. Hutter. Asymptotic learnability of reinforcement problems with arbitrary dependence. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 334–347. Springer, Oct. 2006.
- [4893] D. Ryabko and M. Hutter. On the possibility of learning in reactive environments with arbitrary dependence. *Theoret. Comput. Sci.*, 405(3):274–284, 2008. Special issue for ALT 2006.
- [4894] M. Ryan and M. Reid. Learning to fly: An application of hierarchical reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 807–814. Morgan Kaufmann, San Francisco, CA, 2000.
- [4895] M. R. K. Ryan and M. D. Pendrith. RL-TOPs: an architecture for modularity and re-use in reinforcement learning. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML ’98)*, pages 481–487. Morgan Kaufmann, San Francisco, CA, 1998.
- [4896] M. Rychetsky, J. Shawe-Taylor, and M. Glesner. Direct Bayes point machines. In *Proc. 17th International Conf. on Machine Learning*, pages 815–822. Morgan Kaufmann, San Francisco, CA, 2000.

- [4897] S. Sabato. Submodular learning and covering with response-dependent costs. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 130–144. Springer, 2016.
- [4898] S. Sabato and T. Hess. Interactive algorithms: from pool to stream. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1419–1439, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [4899] S. Sabato and S. Shalev-Shwartz. Prediction by categorical features: Generalization properties and application to feature ranking. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 559–573, Berlin, 2007. Springer.
- [4900] G. Sablon, L. De Raedt, and M. Bruynooghe. Generalizing multiple examples in explanation based learning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 177–183. Springer-Verlag, 1989.
- [4901] U. Sack and M. Kamprad. Evaluation of clinical relevance of clinical laboratory investigations by data mining. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 12–22. Springer, 2001.
- [4902] K. Sadohara. Learning of Boolean functions using support vector machines. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 106–118. Springer, 2001.
- [4903] K. Sadohara and M. Haraguchi. Analogical logic program synthesis algorithm that can refute inappropriate similarities. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 266–281. Springer, 1995.
- [4904] M. Saerens. Non mean square error criteria for the training of learning machines. In *Proc. 13th International Conference on Machine Learning*, pages 427–434. Morgan Kaufmann, 1996.
- [4905] M. Sahami, M. Hearst, and E. Saund. Applying the multiple cause mixture model to text categorization. In *Proc. 13th International Conference on Machine Learning*, pages 435–443. Morgan Kaufmann, 1996.
- [4906] A. S. Saidi and S. Tayeb-bey. Grammatical inference in document recognition. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 175–186. Springer, 1998.
- [4907] L. Saitta and F. Bergadano. Pattern recognition and valiant’s learning framework. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(2):145–155, Feb. 1993.
- [4908] L. Saitta, M. Botta, and F. Neri. Multistrategy learning and theory revision. *Machine Learning*, 11:153–172, 1993.
- [4909] L. Saitta and F. Neri. Learning in the ‘real world’. *Machine Learning*, 30:133–163, 1998.

- [4910] Y. Sakai and A. Maruoka. Learning  $k$ -term monotone Boolean formulae. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 197–207. Springer-Verlag, 1993.
- [4911] Y. Sakai and A. Maruoka. Learning monotone log-term DNF formulas. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 165–172, New York, NY, 1994. ACM Press.
- [4912] Y. Sakai and A. Maruoka. Learning monotone Log-Term DNF formulas under the uniform distribution. *Theory of Computing Systems*, 33(1):17–33, 2000.
- [4913] Y. Sakai, E. Takimoto, and A. Maruoka. Proper learning algorithm for functions of  $k$  terms under smooth distributions. *Inform. Comput.*, 152(2):188–204, 1999. First appeared in COLT'95.
- [4914] Y. Sakaki. The logic of molecular geneticists for the understanding of genetic information. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 3–5. IOS Press, 1992.
- [4915] Y. Sakakibara. Inductive inference of logic programs based on algebraic semantics. Technical report, IIAS-SIS, Fujitsu Limited, Research Report No. 79, 1987.
- [4916] Y. Sakakibara. Inferring parsers of context free languages from structural examples. Technical Report 81, Fujitsu Ltd., 1987.
- [4917] Y. Sakakibara. Inferring parsers of context-free languages from structural examples. Technical report, IIAS-SIS, Fujitsu Limited, Research Report No. 81, 1987.
- [4918] Y. Sakakibara. Learnability in the presence of classification noise. In *Fujitsu IIAS-SIS Workshop on Computational Learning Theory*, pages 1–27, Nov. 1989.
- [4919] Y. Sakakibara. An efficient robust algorithm for learning decision lists. Technical Report No. 105, International Institute for Advanced Study of Social Information Science, Fujitsu Laboratories Ltd, Aug. 1990.
- [4920] Y. Sakakibara. Learning context-free grammars from structural data in polynomial time. *Theoret. Comput. Sci.*, 76(2-3):223–242, 1990.
- [4921] Y. Sakakibara. *Algorithmic Learning of Formal Languages and Decision Trees*. PhD thesis, Tokyo Institute of Technology, Oct. 1991. (International Institute for Advanced Study of Social Information Science, Fujitsu Laboratories Ltd, Research Report IIAS-RR-91-22E).
- [4922] Y. Sakakibara. On learning from queries and counterexamples in the presence of noise. *Inform. Proc. Lett.*, 37(5):279–284, Mar. 1991.
- [4923] Y. Sakakibara. Efficient learning of context-free grammars from positive structural examples. *Inform. Comput.*, 97:23–60, 1992.
- [4924] Y. Sakakibara. Noise-tolerant occam algorithms and their applications to learning decision trees. *Machine Learning*, 11:37–62, 1993.
- [4925] Y. Sakakibara. Recent advances of grammatical inference. *Theoret. Comput. Sci.*, 185(1):15–45, 1997. Special issue for ALT '95.
- [4926] Y. Sakakibara, M. Brown, R. Hughey, I. S. Mian, K. Sjölander, R. C. Underwood, and D. Haussler. Recent methods for RNA modeling using stochastic context-free grammars. In *Proc. of the Asilomar Conference on Combinatorial Pattern Matching*, New York, NY, 1994. Springer-Verlag.

- [4927] Y. Sakakibara, M. Brown, R. Hughey, I. S. Mian, K. Sjölander, R. C. Underwood, and D. Haussler. Stochastic context-free grammars for tRNA modeling. *Nucleic Acids Research*, 22:5112–5120, 1994.
- [4928] Y. Sakakibara, K. P. Jantke, and S. Lange. Learning languages by collecting cases and tuning parameters. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 532–546. Springer-Verlag, 1994.
- [4929] Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, and E. Tomita, editors. *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, Berlin, 2006. Springer.
- [4930] Y. Sakakibara and M. Kondo. GA-based learning of context-free grammars using tabular representations. In *Proc. 16th International Conf. on Machine Learning*, pages 354–360. Morgan Kaufmann, San Francisco, CA, 1999.
- [4931] Y. Sakakibara and H. Muramatsu. Learning context-free grammars from partially structured examples. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 229–240. Springer, 2000.
- [4932] Y. Sakakibara and R. Siromoney. A noise model on learning sets of strings. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 295–302, New York, NY, 1992. ACM Press.
- [4933] C. Sakama. Towards the integration of inductive and nonmonotonic logic programming. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 178–188. Springer, 2001.
- [4934] H. Sakamoto. Language learning from membership queries and characteristic examples. In *Algorithmic Learning Theory, 6th International Workshop, ALT '95, Fukuoka, Japan, October 18–20, 1995, Proceedings*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 55–65. Springer, 1995.
- [4935] H. Sakamoto. Learning simple deterministic finite-memory automata. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 416–431. Springer, 1997.
- [4936] H. Sakamoto. Finding a one-variable pattern from incomplete data. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 234–246. Springer, 1998.
- [4937] H. Sakamoto. Grammar compression: Grammatical inference by compression and its application to real data. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 3–20, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [4938] H. Sakamoto, H. Arimura, and S. Arikawa. Identification of tree translation rules from examples. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 241–255. Springer, 2000.

- [4939] H. Sakamoto, H. Arimura, and S. Arikawa. Knowledge discovery from semistructured texts. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 586–599. Springer, 2001.
- [4940] H. Sakamoto, K. Hirata, and H. Arimura. Learning elementary formal systems with queries. *Theoret. Comput. Sci.*, 298(1):21–50, 2003.
- [4941] M. F. Sakr, S. P. Levitan, D. M. Chiarulli, B. G. Horne, and C. L. Giles. Predicting multiprocessor memory access patterns with learning models. In *Proc. 14th International Conference on Machine Learning*, pages 305–312. Morgan Kaufmann, 1997.
- [4942] A. Sakurai. Inductive inference of formal languages from positive data enumerated primitive-recursively. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 73–83. IOS Press, 1992.
- [4943] A. Sakurai. On the VC-dimension of depth four threshold circuits and the complexity of Boolean-valued functions. *Theoret. Comput. Sci.*, 137(1):109–127, 1995. Special issue for ALT '93.
- [4944] S. Sakurai and M. Haraguchi. Towards learning by abstraction. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 288–298. IOS Press, 1992.
- [4945] M. Salganicoff and L. H. Ungar. Active exploration and learning in real-valued spaces using multi-armed bandit allocation indices. In *Proc. 12th International Conference on Machine Learning*, pages 480–487. Morgan Kaufmann, 1995.
- [4946] M. Salganicoff, L. H. Ungar, and R. Bajcsy. Active learning for vision-based robot grasping. *Machine Learning*, 23:251–278, 1996.
- [4947] J. Salmon and A. Dalalyan. Optimal aggregation of affine estimators. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 635–660, Budapest, Hungary, 09–11 June 2011. PMLR.
- [4948] A. Salomon and J.-Y. Audibert. Deviations of stochastic bandit regret. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 159–173, Berlin, Heidelberg, New York, 2011. Springer.
- [4949] A. Salomon and J.-Y. Audibert. Robustness of stochastic bandit policies. *Theoret. Comput. Sci.*, 519:46–67, 2014. Special Issue: Algorithmic Learning Theory.
- [4950] R. Sałustowicz and J. Schmidhuber. Evolving structured programs with hierarchical instructions and skip nodes. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 488–496. Morgan Kaufmann, San Francisco, CA, 1998.
- [4951] S. Salzberg. Exemplar-based learning: theory and implementation. Center for Research in Computing Technology TR-10-88, Harvard University, Cambridge, Mass., Oct. 1988.
- [4952] S. Salzberg. Nested hyper-rectangles for exemplar-based learning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 184–201. Springer-Verlag, 1989.



- [4953] S. Salzberg. A nearest hyperrectangle learning method. *Machine Learning*, 6:251–276, 1991.
- [4954] S. L. Salzberg. Book review: *C4.5: Programs for Machine Learning* by J. ross quinlan. morgan kaufmann publishers, inc., 1993. *Machine Learning*, 16(3):235–240, 1994.
- [4955] R. Samei, P. Semukhin, B. Yang, and S. Zilles. Sauer’s bound for a notion of teaching complexity. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 96–110, Berlin, Heidelberg, New York, 2012. Springer.
- [4956] R. Samei, P. Semukhin, B. Yang, and S. Zilles. Algebraic methods proving sauer’s bound for teaching complexity. *Theoret. Comput. Sci.*, 558:35–50, 2014. Special Issue: Algorithmic Learning Theory.
- [4957] R. Samei, P. Semukhin, B. Yang, and S. Zilles. Sample compression for multi-label concept classes. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 371–393, Barcelona, Spain, 13–15 June 2014. PMLR.
- [4958] R. Samei, B. Yang, and S. Zilles. Generalizing labeled and unlabeled sample compression to multi-label concept classes. In P. Auer, A. Clark, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 25th International Conference, ALT 2014, Bled, Slovenia, October 8–10, 2014, Proceedings*, volume 8776 of *Lecture Notes in Artificial Intelligence*, pages 275–290, Berlin, Heidelberg, New York, 2014. Springer.
- [4959] C. Sammut. Using background knowledge to build multistrategy learners. *Machine Learning*, 27(3):241–257, 1997.
- [4960] C. Sammut. Concept learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 205–208. Springer, New York, 2010.
- [4961] C. Sammut and R. Banerji. Learning concepts by asking questions. In *Machine Learning : An AI Approach, Vol2*. Morgan Kaufman, 1986.
- [4962] C. Sammut and G. I. Webb, editors. *Encyclopedia of Machine Learning*. Springer, New York, 2010.
- [4963] A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM J. of Res. Develop.*, 3:211–229, July 1959. (Reprinted in *Computers and Thought*, (eds. E. A. Feigenbaum and J. Feldman), McGraw-Hill, 1963, pages 39–70).
- [4964] A. L. Samuel. *Some studies in machine learning using the game of checkers*, pages 39–70. McGraw Hill, 1963.
- [4965] K. Samuel, S. Carberry, and K. Vijay-Shanker. An investigation of transformation-based learning in discourse. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML ’98)*, pages 497–505. Morgan Kaufmann, San Francisco, CA, 1998.
- [4966] C. Samuelsson, P. Tapanainen, and A. Voutilainen. Inducing constraint grammars. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 146–155. Springer, 1996.

- [4967] P. Sand and A. W. Moore. Repairing faulty mixture models using density estimation. In *Proc. 18th International Conf. on Machine Learning*, pages 457–464. Morgan Kaufmann, San Francisco, CA, 2001.
- [4968] T. D. Sanger. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473, 1989.
- [4969] S. Sanner, J. R. Anderson, C. Lebiere, and M. Lovett. Achieving efficient and cognitively plausible learning in Backgammon. In *Proc. 17th International Conf. on Machine Learning*, pages 823–830. Morgan Kaufmann, San Francisco, CA, 2000.
- [4970] P. Santi, V. Conitzer, and T. Sandholm. Towards a characterization of polynomial preference elicitation with value queries in combinatorial auctions: (extended abstract). In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 1–16. Springer, 2004.
- [4971] J. Sanz-González and D. Andina. Importance sampling techniques in neural detector training. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 431–441. Springer, 2001.
- [4972] A. Saoudi. Learning local and recognizable  $\omega$ -languages and monadic logic programs. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 157–169, Oxford, 1994. Oxford University Press.
- [4973] J. Särelä and H. Valpola. Denoising source separation. *Journal of Machine Learning Research*, 6:233–272, 2005.
- [4974] J. Särelä and R. Vigário. Overlearning in marginal distribution-based ICA: Analysis and solutions. *Journal of Machine Learning Research*, 4:1447–1469, 2003.
- [4975] U. Sarkans and J. Bārzdīņš. Using attribute grammars for description of inductive inference search space. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 400–408. Springer, 1998.
- [4976] M. Sarkar and T.-Y. Leong. Application of fuzzy similarity-based fractal dimensions to characterize medical time series. In *Proc. 18th International Conf. on Machine Learning*, pages 465–472. Morgan Kaufmann, San Francisco, CA, 2001.
- [4977] K. Sasahara, Y. Kakishita, T. Nishino, M. Takahasi, and K. Okanoya. Constructing song syntax by automata induction. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 351–353, Berlin, 2006. Springer.
- [4978] Y. Sasaki. Induction of logic programs based on  $\psi$ -terms. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 169–181. Springer, 1999.
- [4979] I. Sato, K. Chubachi, and Diptarama. Evaluation of machine learning methods on SPiCe. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 149–153, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.

- [4980] M. Sato. Inductive inference of formal languages. *Bulletin of Informatics and Cybernetics*, 27(1):85–106, 1995.
- [4981] M. Sato. Classical Brouwer-Heyting-Kolmogorov interpretation. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 176–196. Springer, 1997.
- [4982] M. Sato. Theory of judgments and derivations. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 78–122. Springer, 2001.
- [4983] M. Sato and S. Kobayashi. Average-reward reinforcement learning for variance penalized Markov decision problems. In *Proc. 18th International Conf. on Machine Learning*, pages 473–480. Morgan Kaufmann, San Francisco, CA, 2001.
- [4984] M. Sato, Y. Mukouchi, and M. Terada. Refutable/inductive learning from neighbor examples and its application to decision trees over patterns. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 201–213. Springer, 2001.
- [4985] M. Sato, Y. Mukouchi, and D. Zheng. Characteristic sets for unions of regular pattern languages and compactness. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 220–233. Springer, 1998.
- [4986] M. Sato and K. Umayahara. Inductive inferability for formal languages from positive data. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 84–92. IOS Press, 1992.
- [4987] T. Sato. EM learning for symbolic-statistical models in statistical abduction. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 189–200. Springer, 2001.
- [4988] T. Sato and S. Akiba. Inductive resolution. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 101–110. Springer-Verlag, 1993.
- [4989] K. Satoh. Analysis of case-based representability of Boolean functions by monotone theory. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 179–190. Springer, 1998.
- [4990] K. Satoh. Learning taxonomic relation by case-based reasoning. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 179–193. Springer, 2000.
- [4991] K. Satoh. Constructing a critical casebase to represent a lattice-based relation. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 214–223. Springer, 2001.
- [4992] K. Satoh. Learning taxonomic relation by case-based reasoning. *Theoret. Comput. Sci.*, 348(1):58–69, 2005. Special issue for ALT 2000.

- [4993] L. K. Saul. Automatic segmentation of continuous trajectories with invariance to nonlinear warpings of time. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 506–514. Morgan Kaufmann, San Francisco, CA, 1998.
- [4994] L. K. Saul and M. I. Jordan. Mixed memory Markov models: Decomposing complex stochastic processes as mixtures of simpler ones. *Machine Learning*, 37(1):75–87, 1999.
- [4995] L. K. Saul and M. G. Rahim. Markov processes on curves. *Machine Learning*, 41(3):345–363, 2000.
- [4996] L. K. Saul and S. T. Roweis. Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research*, 4:119–155, 2003.
- [4997] L. K. Saul and S. P. Singh. Markov decision processes in large state spaces. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 281–288, New York, NY, 1995. ACM Press.
- [4998] L. K. Saul and S. P. Singh. Learning curve bounds for a Markov decision process with undiscounted rewards. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 147–156, New York, NY, 1996. ACM Press.
- [4999] C. Saunders, A. Gammerman, and V. Vovk. Computationally efficient transductive machines. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 325–333. Springer, 2000.
- [5000] G. Saunders, A. Gammerman, and V. Vovk. Ridge regression learning algorithm in dual variables. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 515–521. Morgan Kaufmann, San Francisco, CA, 1998.
- [5001] J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy Gaussian process bandit optimization. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1723–1742, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5002] M. Schaefer. Deciding the Vapnik-Cervonenkis dimension is  $\Sigma_3^P$ -complete. *J. of Comput. Syst. Sci.*, 58(1):177–182, 1999.
- [5003] G. Schäfer. Some results in the theory of effective program synthesis: Learning by defective information. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 219–225. Springer-Verlag, 1986.
- [5004] G. Schäfer-Richter. *Über Eingabeabhängigkeit und Komplexität von Inferenzstrategien*. PhD thesis, RWTH Aachen, 1984.
- [5005] C. Schaffer. Bivariate scientific function finding in a sampled, real-data testbed. *Machine Learning*, 12:167–183, 1993.
- [5006] C. Schaffer. Overfitting avoidance as bias. *Machine Learning*, 10:153–178, 1993.
- [5007] C. Schaffer. Selecting a classification method by cross-validation. *Machine Learning*, 13:135–143, 1993.
- [5008] R. E. Schapire. Diversity-based inference of finite automata. Master’s thesis, MIT Lab. for Computer Science, May 1988. Technical Report MIT/LCS/TR-413.

- [5009] R. E. Schapire. Pattern languages are not learnable. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 122–129, San Mateo, CA, 1990. Morgan Kaufmann.
- [5010] R. E. Schapire. The strength of weak learnability. *Machine Learning*, 5(2):197–227, 1990.
- [5011] R. E. Schapire. *The Design and Analysis of Efficient Learning Algorithms*. MIT Press, Cambridge, MA, 1992.
- [5012] R. E. Schapire. Learning probabilistic read-once formulas on product distributions. *Machine Learning*, 14(1):47–81, 1994.
- [5013] R. E. Schapire. Using output codes to boost multiclass learning problems. In *Proc. 14th International Conference on Machine Learning*, pages 313–321. Morgan Kaufmann, 1997.
- [5014] R. E. Schapire. Guest editor’s foreword. *J. of Comput. Syst. Sci.*, 56(2):131, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [5015] R. E. Schapire. Drifting games. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 114–124, New York, NY, 1999. ACM Press.
- [5016] R. E. Schapire. Theoretical views of boosting. In *Computational Learning Theory, 4th European Conference, EuroCOLT ’99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 1–10. Springer, 1999.
- [5017] R. E. Schapire. Theoretical views of boosting and applications. In *Algorithmic Learning Theory, 10th International Conference, ALT ’99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 13–25. Springer, 1999.
- [5018] R. E. Schapire. Drifting games. *Machine Learning*, 43(3):265–291, 2001.
- [5019] R. E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee. Boosting the margin: a new explanation for the effectiveness of voting methods. In *Proc. 14th International Conference on Machine Learning*, pages 322–330. Morgan Kaufmann, 1997.
- [5020] R. E. Schapire and L. M. Sellie. Learning sparse multivariate polynomials over a field with queries and counterexamples. *J. of Comput. Syst. Sci.*, 52(2):201–213, 1996. Earlier version in 6th COLT, 1993.
- [5021] R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.
- [5022] R. E. Schapire and Y. Singer. BoosTexter: A boosting-based system for text categorization. *Machine Learning*, 39(2/3):135–168, 2000.
- [5023] R. E. Schapire and M. K. Warmuth. On the worst-case analysis of temporal-difference learning algorithms. *Machine Learning*, 22(1/2/3):95–121, 1996.
- [5024] T. Scheffer. Average-case analysis of classification algorithms for Boolean functions and decision trees. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 194–208. Springer, 2000.

- [5025] T. Scheffer. Nonparametric regularization of decision trees. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 344–356. Springer, 2000.
- [5026] T. Scheffer. Predicting the generalization performance of cross validatory model selection criteria. In *Proc. 17th International Conf. on Machine Learning*, pages 831–838. Morgan Kaufmann, San Francisco, CA, 2000.
- [5027] T. Scheffer, R. Greiner, and C. Darken. Why experimentation can be better than perfect guidance. In *Proc. 14th International Conference on Machine Learning*, pages 331–339. Morgan Kaufmann, 1997.
- [5028] T. Scheffer and T. Joachims. Expected error analysis for model selection. In *Proc. 16th International Conf. on Machine Learning*, pages 361–370. Morgan Kaufmann, San Francisco, CA, 1999.
- [5029] T. Scheffer and S. Wrobel. Incremental maximization of non-instance-averaging utility functions with applications to knowledge discovery problems. In *Proc. 18th International Conf. on Machine Learning*, pages 481–488. Morgan Kaufmann, San Francisco, CA, 2001.
- [5030] T. Scheffer and S. Wrobel. Finding the most interesting patterns in a database quickly by using sequential sampling. *Journal of Machine Learning Research*, 3:833–862, 2002.
- [5031] M. F. Schlang, K. Abraham-Fuchs, R. Neuneier, and J. Uebler. A comparison of RBF and MLP networks for classification of biomagnetic fields. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 20, pages 353–362. MIT Press, 1997.
- [5032] J. C. Schlimmer and R. H. Granger, JR. Incremental learning from noisy data. *Machine Learning*, 1(3):317–354, 1986.
- [5033] U. Schmid, M. Siebers, D. Seuß, M. Kunz, and S. Lautenbacher. Applying grammar inference to identify generalized patterns of facial expressions of pain. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 183–188, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5034] J. Schmidhuber. Discovering solutions with low Kolmogorov complexity and high generalization capability. In *Proc. 12th International Conference on Machine Learning*, pages 488–496. Morgan Kaufmann, 1995.
- [5035] J. Schmidhuber. The speed prior: A new simplicity measure yielding near-optimal computable predictions. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 216–228. Springer, 2002.
- [5036] J. Schmidhuber. Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity and creativity. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 32–33, Berlin, Oct. 2007. Springer.
- [5037] J. Schmidhuber, J. Zhao, and M. Wiering. Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement. *Machine Learning*, 28(1):105–130, 1997.

- [5038] R. Schmidt and L. Gierl. Temporal abstractions and case-based reasoning for medical course data: Two prognostic applications. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 23–34. Springer, 2001.
- [5039] W. C. Schmidt and C. X. Ling. A decision-tree model of balance scale development. *Machine Learning*, 24:203–229, 1996.
- [5040] M. Schmitt. On the sample complexity for neural trees. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 375–384. Springer, 1998.
- [5041] M. Schmitt. On the sample complexity for nonoverlapping neural networks. *Machine Learning*, 37(2):131–141, 1999.
- [5042] M. Schmitt. Radial basis function neural networks have superlinear VC dimension. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 14–30. Springer, 2001.
- [5043] M. Schmitt. RBF neural networks and descartes' rule of signs. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 321–335. Springer, 2002.
- [5044] M. Schmitt. An improved VC dimension bound for sparse polynomials. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 393–407. Springer, 2004.
- [5045] J. Schneider, W.-K. Wong, A. Moore, and M. Riedmiller. Distributed value functions. In *Proc. 16th International Conf. on Machine Learning*, pages 371–378. Morgan Kaufmann, San Francisco, CA, 1999.
- [5046] J. G. Schneider, J. A. Boyan, and A. W. Moore. Value function based production scheduling. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 522–530. Morgan Kaufmann, San Francisco, CA, 1998.
- [5047] G. Schohn and D. Cohn. Less is more: Active learning with support vector machines. In *Proc. 17th International Conf. on Machine Learning*, pages 839–846. Morgan Kaufmann, San Francisco, CA, 2000.
- [5048] B. Schölkopf, R. Herbrich, and A. J. Smola. A generalized representer theorem. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 416–426. Springer, 2001.
- [5049] B. Schölkopf and M. K. Warmuth, editors. *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, Berlin, 2003. Springer.

- [5050] T. Schramm and D. Steurer. Fast and robust tensor decomposition with applications to dictionary learning. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1760–1793, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5051] N. N. Schraudolph and R. K. Belew. Dynamic parameter encoding for genetic algorithms. *Machine Learning*, 9:9–21, 1992.
- [5052] R. Schuler. Universal distributions and time-bounded kolmogorov complexity. In *Proc. 16th Annual Symposium on Theoretical Aspects of Computer Science - STACS '99*, volume 1563 of *Lecture Notes in Computer Science*, pages 434–443. Springer, 1999.
- [5053] O. Schulte, W. Luo, and R. Greiner. Mind change optimal learning of Bayes net structure. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 187–202, Berlin, 2007. Springer.
- [5054] O. Schulte, W. Luo, and R. Greiner. Mind change optimal learning of Bayes net structure from dependency and independency data. *Inform. Comput.*, 208(1):63–82, 2010.
- [5055] T. R. Schultz, D. Mareschal, and W. C. Schmidt. Modeling cognitive development on balance scale phenomena. *Machine Learning*, 16(1/2):57–86, 1994.
- [5056] S. Schulze-Kreme. Genetic algorithms for protein tertiary structure prediction. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 262–279. Springer-Verlag, 1993.
- [5057] D. Schuurmans. Characterizing rational versus exponential learning curves. *J. Comput. Syst. Sci.*, 55(1):140–160, 1997. Special issue for EuroCOLT '95.
- [5058] D. Schuurmans and R. Greiner. Learning default concepts. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Proc. of the Tenth Canadian Conference on Artificial Intelligence*, pages 519–523, 1994.
- [5059] D. Schuurmans and R. Greiner. Practical PAC learning. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, IJCAI 95, Montréal, Québec, Canada, August 20-25 1995*, pages 1169–1177, 1995.
- [5060] D. Schuurmans and R. Greiner. Sequential PAC learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 377–384, New York, NY, 1995. ACM Press.
- [5061] D. Schuurmans and R. Greiner. Fast distribution-specific learning. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 10. MIT Press, 1997.
- [5062] D. Schuurmans and R. Greiner. Learning to classify incomplete examples. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 6, pages 87–105. MIT Press, 1997.
- [5063] D. Schuurmans and F. Southey. An adaptive regularization criterion for supervised learning. In *Proc. 17th International Conf. on Machine Learning*, pages 847–854. Morgan Kaufmann, San Francisco, CA, 2000.



- [5064] D. Schuurmans and F. Southey. Metric-based methods for adaptive model selection and regularization. *Machine Learning*, 48(1/3):51–84, 2002.
- [5065] D. Schuurmans, L. H. Ungar, and D. P. Foster. Characterizing the generalization performance of model selection strategies. In *Proc. 14th International Conference on Machine Learning*, pages 340–348. Morgan Kaufmann, 1997.
- [5066] M. Schwabacher and P. Langley. Discovering communicable scientific knowledge from spatio-temporal data. In *Proc. 18th International Conf. on Machine Learning*, pages 489–496. Morgan Kaufmann, San Francisco, CA, 2001.
- [5067] R. W. Schwanke and J. S. Hanson. Using neural networks to modularize software. *Machine Learning*, 15(2):137–168, 1994.
- [5068] D. B. Schwartz, R. E. Howard, and W. E. Hubbard. Adaptive neural networks using MOS charge storage. In *Proc. 1988 Neural Information Processing Conference*, pages 761–768, 1989.
- [5069] J. Scicluna and C. de la Higuera. Grammatical inference of some probabilistic context-free grammars from positive data using minimum satisfiability. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 139–152, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [5070] C. Scott, G. Blanchard, and G. Handy. Classification with asymmetric label noise: Consistency and maximal denoising. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 489–511. JMLR.org, 2013.
- [5071] P. D. Scott and S. Markovitch. Experience selection and problem choice in an exploratory learning system. *Machine Learning*, 12:49–67, 1993.
- [5072] S. Scott and S. Matwin. Feature engineering for text classification. In *Proc. 16th International Conf. on Machine Learning*, pages 379–388. Morgan Kaufmann, San Francisco, CA, 1999.
- [5073] D. B. Searls. Molecules, languages and automata. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 5–10. Springer, 2010.
- [5074] M. Sebag. A constraint-based induction algorithm in FOL. In *Proc. 11th International Conference on Machine Learning*, pages 275–283. Morgan Kaufmann, 1994.
- [5075] M. Sebag. Delaying the choice of bias: a disjunctive version space approach. In *Proc. 13th International Conference on Machine Learning*, pages 444–452. Morgan Kaufmann, 1996.
- [5076] M. Sebag and C. Rouveirol. Resource-bounded relational reasoning: Induction and deduction through stochastic matching. *Machine Learning*, 38(1/2):41–62, 2000.
- [5077] P. Sebastiani, I. S. Kohane, and M. F. Ramoni. Editorial: Methods in functional genomics. *Machine Learning*, 52(1-2):5–9, 2003.
- [5078] P. Sebastiani and M. Ramoni. Clustering continuous time series. In *Proc. 18th International Conf. on Machine Learning*, pages 497–504. Morgan Kaufmann, San Francisco, CA, 2001.

- [5079] M. Sebban and R. Nock. Instance pruning as an information preserving problem. In *Proc. 17th International Conf. on Machine Learning*, pages 855–862. Morgan Kaufmann, San Francisco, CA, 2000.
- [5080] M. Sebban, R. Nock, and S. Lallich. Boosting neighborhood-based classifiers. In *Proc. 18th International Conf. on Machine Learning*, pages 505–512. Morgan Kaufmann, San Francisco, CA, 2001.
- [5081] M. Sebban, R. Nock, and S. Lallich. Stopping criterion for boosting-based data reduction techniques: from binary to multiclass problem. *Journal of Machine Learning Research*, 3:863–885, 2002.
- [5082] M. Seeger. PAC-Bayesian generalisation error bounds for gaussian process classification. *Journal of Machine Learning Research*, 3:233–269, 2002.
- [5083] E. Segal, D. Pe’er, A. Regev, D. Koller, and N. Friedman. Learning module networks. *Journal of Machine Learning Research*, 6:557–588, 2005.
- [5084] R. B. Segal and J. O. Kephart. Incremental learning in swiftFile. In *Proc. 17th International Conf. on Machine Learning*, pages 863–870. Morgan Kaufmann, San Francisco, CA, 2000.
- [5085] Y. Seginer. Fast learning from strings of 2-letter rigid grammars. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 213–224. Springer, 2002.
- [5086] A. Segre, C. Elkan, and A. Russell. A critical look at experimental evaluations of EBL. *Machine Learning*, 6:183–195, 1991.
- [5087] G. Seidelmann. Using heuristics to speed up induction on continuous-valued attributes. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 390–395. Springer-Verlag, 1993.
- [5088] C. M. Seifert, K. J. Hammond, H. M. Johnson, T. M. Converse, T. F. Mcdoughal, and S. W. Vanderstoep. Case-based learning: Predictive features in indexing. *Machine Learning*, 16(1/2):37–56, 1994.
- [5089] T. J. Sejnowski, P. Dayan, and P. R. Montague. Predictive Hebbian learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 15–18, New York, NY, 1995. ACM Press.
- [5090] T. J. Sejnowski and C. R. Rosenberg. Parallel networks that learn to pronounce english text. *Journal of Complex Systems*, 1(1):145–168, Feb. 1987.
- [5091] S. Seki and S. Kobayashi. Efficient learning of  $k$ -reversible context-free grammars from positive structural examples. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 285–287. Springer, 2004.
- [5092] J. Selbig. Knowledge acquisition by inductive learning from examples. In *Analogical and Inductive Inference, International Workshop AII ’86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 145–163. Springer-Verlag, 1986.

- [5093] Y. Seldin, G. Bejerano, and N. Tishby. Unsupervised sequence segmentation by a mixture of switching variable memory Markov sources. In *Proc. 18th International Conf. on Machine Learning*, pages 513–520. Morgan Kaufmann, San Francisco, CA, 2001.
- [5094] Y. Seldin, K. Crammer, and P. Bartlett. Open problem: Adversarial multiarmed bandits with limited advice. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1067–1072. JMLR.org, 2013.
- [5095] Y. Seldin and G. Lugosi. An improved parametrization and analysis of the EXP3++ algorithm for stochastic and adversarial bandits. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1743–1759, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5096] L. Sellie. Learning random monotone DNF under the uniform distribution. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 181–192. Omnipress, 2008.
- [5097] B. Selman and H. Kautz. Knowledge compilation using Horn approximations. In *Proc. of AAAI-91, Anaheim, Aug. 1991*.
- [5098] J. Sempere and P. García. Learning locally testable even linear languages from positive data. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 225–236. Springer, 2002.
- [5099] J. M. Sempere. Learning reversible languages with terminal distinguishability. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 354–355, Berlin, 2006. Springer.
- [5100] J. M. Sempere. Learning context-sensitive languages from linear structural information. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 175–186. Springer, 2008.
- [5101] J. M. Sempere and A. Fos. Learning linear grammars from structural information. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 126–133. Springer, 1996.
- [5102] J. M. Sempere and G. Nagaraja. Learning a subclass of linear languages from positive structural information. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 162–174. Springer, 1998.
- [5103] S. Sen and M. Sekaran. Individual learning of coordination knowledge. *J. of Experimental and Theoret. Artif. Intell.*, 10(3):333–356, 1998.
- [5104] D. Servan-Schreiber, A. Cleeremans, and J. L. McClelland. Graded state machines: The representation of temporal contingencies in simple recurrent networks. *Machine Learning*, 7:161–193, 1991.

- [5105] R. Servedio. PAC analogues of perceptron and Winnow via boosting the margin. *Machine Learning*, 47(2/3):133–151, 2002.
- [5106] R. Servedio, L.-Y. Tan, and J. Thaler. Attribute-efficient learning and weight-degree tradeoffs for polynomial threshold functions. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 14.1–14.19, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5107] R. A. Servedio. On PAC learning using Winnow, perceptron, and a perceptron-like algorithm. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 296–307, New York, NY, 1999. ACM Press.
- [5108] R. A. Servedio. Computational sample complexity and attribute-efficient learning. *J. of Comput. Syst. Sci.*, 60(1):161–178, 2000.
- [5109] R. A. Servedio. PAC analogues of perceptron and winnow via boosting the margin. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 148–157. Morgan Kaufmann, San Francisco, 2000.
- [5110] R. A. Servedio. *Efficient Algorithms in Computational Learning Theory*. PhD thesis, The Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, 2001.
- [5111] R. A. Servedio. On learning monotone DNF under product distributions. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 558–573. Springer, 2001.
- [5112] R. A. Servedio. On the limits of efficient teachability. *Information Processing Letters*, 79(6):267–272, 2001.
- [5113] R. A. Servedio. Separating quantum and classical learning. In *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 1065–1080. Springer, 2001.
- [5114] R. A. Servedio. Smooth boosting and learning with malicious noise. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 473–489. Springer, 2001.
- [5115] R. A. Servedio. On learning embedded midbit functions. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 69–82. Springer, 2002.
- [5116] R. A. Servedio. Smooth boosting and learning with malicious noise. *Journal of Machine Learning Research*, 4:633–648, 2003.
- [5117] R. A. Servedio. On learning monotone DNF under product distributions. *Inform. Comput.*, 193(1):57–74, 2004.
- [5118] R. A. Servedio. On learning embedded midbit functions. *Theoret. Comput. Sci.*, 350(1):13–23, 2006. Special issue for ALT 2002.

- [5119] R. A. Servedio. On PAC learning algorithms for rich Boolean function classes. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 442–451, Berlin, 2006. Springer.
- [5120] R. A. Servedio and S. J. Gortler. Equivalences and separations between quantum and classical learnability. *SIAM J. Comput.*, 33(5):1067–1092, 2004.
- [5121] R. Seshu. Binary decision trees and an ‘average-case’ model for concept learning: implications for feature construction and the study of bias. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 8, pages 213–247. MIT Press, 1994.
- [5122] H. S. Seung, M. Opper, and H. Sompolinsky. Query by committee. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 287–294, New York, NY, 1992. ACM Press.
- [5123] F. Sha, L. K. Saul, and D. D. Lee. Multiplicative updates for large margin classifiers. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 188–202. Springer, 2003.
- [5124] G. Shackelford and D. Volper. Learning  $k$ -DNF with noise in the attributes. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 97–103, San Mateo, CA, 1988. Morgan Kaufmann.
- [5125] C. Shaffer. A conservation law for generalization performance. In *Proc. 11th International Conference on Machine Learning*, pages 259–265. Morgan Kaufmann, 1994.
- [5126] U. Shaked and Y. Theodor.  $H_\infty$ -optimal estimation: A tutorial. In *Proc. IEEE Conference on Decision and Control*, pages 2278–2286. IEEE, Dec. 1992.
- [5127] G. Shakhnarovich, R. El-Yaniv, and Y. Baram. Smoothed bootstrap and statistical data cloning for classifier evaluation. In *Proc. 18th International Conf. on Machine Learning*, pages 521–528. Morgan Kaufmann, San Francisco, CA, 2001.
- [5128] S. Shalev-Shwartz. Learnability beyond uniform convergence. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 13–16, Berlin, Heidelberg, New York, 2012. Springer.
- [5129] S. Shalev-Shwartz, O. Shamir, and K. Sridharan. Learning kernel-based halfspaces with the 0-1 loss. *SIAM J. Comput.*, 40(6):1623–1646, 2011.
- [5130] S. Shalev-Shwartz and Y. Singer. A new perspective on an old perceptron algorithm. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 264–278. Springer, 2005.
- [5131] S. Shalev-Shwartz and Y. Singer. Online learning meets optimization in the dual. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 423–437, Berlin, 2006. Springer.

- [5132] S. Shalev-Shwartz and Y. Singer. A primal-dual perspective of online learning algorithms. *Machine Learning*, 69(2-3):115–142, 2007.
- [5133] S. Shalev-Shwartz and Y. Singer. On the equivalence of weak learnability and linear separability: New relaxations and efficient boosting algorithms. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 311–322. Omnipress, 2008.
- [5134] S. Shalev-Shwartz and I. Steinwart, editors. *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12-14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*. JMLR.org, 2013.
- [5135] S. Shalev-Shwartz and I. Steinwart. Preface. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 1–2. JMLR.org, 2013.
- [5136] E. Shamir. A remark on discovery algorithms for grammars. *Inform. Control*, 5:246–251, 1962.
- [5137] E. Shamir and C. Schwartzman. Learning by extended statistical queries and its relation to PAC learning. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 357–366. Springer, 1995.
- [5138] O. Shamir. Open problem: Is averaging needed for strongly convex stochastic gradient descent? In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 47.1–47.3, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5139] O. Shamir. On the complexity of bandit and derivative-free stochastic convex optimization. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 3–24. JMLR.org, 2013.
- [5140] O. Shamir. On the complexity of bandit linear optimization. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1523–1551, Paris, France, 03–06 July 2015. PMLR.
- [5141] O. Shamir, S. Sabato, and N. Tishby. Learning and generalization with the information bottleneck. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 92–107, Berlin, Oct. 2008. Springer.
- [5142] O. Shamir, S. Sabato, and N. Tishby. Learning and generalization with the information bottleneck. *Theoret. Comput. Sci.*, 411(29-30):2696–2711, 2010. Special issue for ALT 2008.
- [5143] O. Shamir and S. Shalev-Shwartz. Collaborative filtering with the trace norm: Learning, bounding, and transducing. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 661–678, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5144] O. Shamir and N. Tishby. Model selection and stability in  $k$ -means clustering. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 367–378. Omnipress, 2008.

- [5145] E. Shapiro. A general incremental algorithm that infers theory from facts. In *Internat. Joint Conf. on Artif. Intell.*, 1981.
- [5146] E. Y. Shapiro. Inductive inference of theories from facts. Technical Report Research Report 192, Yale University Department of Computer Science, Feb. 1981.
- [5147] E. Y. Shapiro. *Algorithmic Program DeBugging*. MIT Press, Cambridge, MA, 1983.
- [5148] N. Shapiro. Review of “Limiting recursion” by E.M. Gold and “Trial and error predicates and the solution to a problem of Mostowski” by H. Putnam. *J. Symbolic Logic*, 36:342, 1971.
- [5149] A. Sharma. *Program Size and Teams for Computational Learning*. PhD thesis, SUNY at Buffalo, 1990.
- [5150] A. Sharma. A note on batch and incremental learnability. *J. of Comput. Syst. Sci.*, 56(3):272–276, 1998.
- [5151] A. Sharma. Editorial. *Theoret. Comput. Sci.*, 241(1-2):1–2, 2000. Special issue for ALT ’96.
- [5152] A. Sharma, F. Stephan, and Y. Ventsov. Generalized notions of mind change complexity. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee*, pages 96–108, New York, NY, 1997. ACM Press.
- [5153] A. Sharma, F. Stephan, and Y. Ventsov. Generalized notions of mind change complexity. *Inform. Comput.*, 189(2):235–262, 2004.
- [5154] H. Shatkay and L. P. Kaelbling. Heading in the right direction. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML ’98)*, pages 531–539. Morgan Kaufmann, San Francisco, CA, 1998.
- [5155] J. Shavlik. Combining symbolic and neural learning, extended abstract. *Machine Learning*, 14(3):321–331, 1994.
- [5156] J. Shavlik, L. Hunter, and D. Searls. Introduction. *Machine Learning*, 21:5–9, 1995.
- [5157] J. W. Shavlik. Acquiring recursive and iterative concepts with explanation-based learning. *Machine Learning*, 5:39–40, 1990.
- [5158] J. W. Shavlik. Combining symbolic and neural learning. *Machine Learning*, 14:321–331, 1994.
- [5159] J. W. Shavlik, R. J. Mooney, and G. G. Towell. Symbolic and neural learning algorithms: An experimental comparison. *Machine Learning*, 6:111–143, 1991.
- [5160] S. W. Shaw, D and C. Green. Inferring LISP programs from example problems. In *Proc. 4th IJCAI*, pages 260–267. IJCAI, 1975.
- [5161] J. Shawe-Taylor. Sample sizes for sigmoidal neural networks. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 258–264, New York, NY, 1995. ACM Press.
- [5162] J. Shawe-Taylor. Confidence estimates of classification accuracy on new examples. In *Computational Learning Theory, Third European Conference, EuroCOLT ’97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 260–271. Springer, 1997.
- [5163] J. Shawe-Taylor. Classification accuracy based on observed margin. *Algorithmica*, 22(1/2):157–172, 1998.

- [5164] J. Shawe-Taylor. Introducing the special issue of machine learning selected from papers presented at the 1997 conference on computational learning theory, COLT'97. *Machine Learning*, 35(3):191–192, 1999.
- [5165] J. Shawe-Taylor, M. Anthony, and R. L. Biggs. Bounding sample size with the vapnik-chervonenkis dimension. Technical Report CSD-TR-618, University of London, Royal Holloway and New Bedford College, 1989.
- [5166] J. Shawe-Taylor and P. L. Bartlett. Structural risk minimization over data-dependent hierarchies. *IEEE Trans. on Information Theory*, 44(5):1926–1940, 1998.
- [5167] J. Shawe-Taylor, P. L. Bartlett, R. C. Williamson, and M. Anthony. A framework for structural risk minimization. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 68–76, New York, NY, 1996. ACM Press.
- [5168] J. Shawe-Taylor and N. Cristianini. Margin distribution bounds on generalization. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 263–273. Springer, 1999.
- [5169] J. Shawe-Taylor and N. Cristianini. Further results on the margin distribution. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 278–285, New York, NY, 1999. ACM Press.
- [5170] J. Shawe-Taylor and Y. Singer, editors. *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, Berlin, 2004. Springer.
- [5171] J. Shawe-Taylor, C. Williams, N. Cristianini, and J. Kandola. On the eigenspectrum of the gram matrix and its relationship to the operator eigenspectrum. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 23–40. Springer, 2002.
- [5172] J. Shawe-Taylor and R. C. Williamson. A PAC analysis of a Bayesian estimator. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th-9th, 1997, Nashville, Tennessee*, pages 2–9, New York, NY, 1997. ACM Press.
- [5173] W.-M. Shen. Discovery as autonomous learning from the environment. *Machine Learning*, 12:143–165, 1993.
- [5174] W.-M. Shen. Learning finite automata using local distinguishing experiments. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 12, pages 225–242. MIT Press, 1995.
- [5175] J. W. Sheppard. Colearning in differential games. *Machine Learning*, 33(2/3):201–233, 1998.
- [5176] V. N. Shevchenko and N. Y. Zolotykh. Lower bounds for the complexity of learning half-spaces with membership queries. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 61–71. Springer, 1998.
- [5177] M. Shi, T. Wakabayashi, W. Ohyama, and F. Kimura. Mirror image learning for handwritten numeral recognition. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 239–248. Springer, 2001.



- [5178] C. Shibata. Inferring  $(k,l)$ -context-sensitive probabilistic context-free grammars using hierarchical pitman-yor processes. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 153–166, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [5179] C. Shibata and J. Heinz. Predicting sequential data with LSTMs augmented with strictly 2-piecewise input vectors. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 137–142, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [5180] C. Shibata and R. Yoshinaka. Marginalizing out transition probabilities for several subclasses of PFAs. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 259–263, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5181] C. Shibata and R. Yoshinaka. PAC learning of some subclasses of context-free grammars with basic distributional properties from positive data. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 143–157, Berlin, Heidelberg, New York, 2013. Springer.
- [5182] C. Shibata and R. Yoshinaka. Probabilistic learnability of context-free grammars with basic distributional properties from positive examples. *Theoret. Comput. Sci.*, 620:46–72, 2016. Special Issue: Algorithmic Learning Theory.
- [5183] T. Shibata, R. Yoshinaka, and T. Chikayama. Probabilistic generalization of simple grammars and its application to reinforcement learning. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 348–362. Springer, Oct. 2006.
- [5184] E. Shibayama, M. Toyoda, J. Yabe, and S. Takahashi. Scalable and comprehensible visualization for discovery of knowledge from the internet. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 661–671. Springer, 2001.
- [5185] S. Shimozone, A. Shinohara, T. Shinohara, S. Miyano, S. Kuhara, and S. Arikawa. Knowledge acquisition from amino acid sequences by machine learning system BONSAI. *Trans. Information Processing Society of Japan*, 35:2009–2018, 1994.
- [5186] K. Shin and T. Kuboyama. Polynomial summaries of positive semidefinite kernels. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 313–327, Berlin, Oct. 2007. Springer.
- [5187] K. Shin and T. Kuboyama. Polynomial summaries of positive semidefinite kernels. *Theoret. Comput. Sci.*, 410(9):1847–1862, 2009. Special issue for ALT 2007.
- [5188] A. Shinohara. Complexity of computing Vapnik-Chervonenkis dimension and some generalized dimensions. *Theoret. Comput. Sci.*, 137(1):129–144, 1995. Special issue for ALT '93.

- [5189] A. Shinohara. String pattern discovery. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 1–13. Springer, 2004.
- [5190] A. Shinohara and S. Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.
- [5191] A. Shinohara, M. Takeda, S. Arikawa, M. Hirao, H. Hoshino, and S. Inenaga. Finding best patterns practically. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 307–317. Springer, 2001.
- [5192] T. Shinohara. Polynomial time inference of pattern languages and its applications. Proc. 7th IBM Symposium on Mathematical Foundations of Computer Science, 1982.
- [5193] T. Shinohara. Inferring unions of two pattern languages. *Bulletin of Informatics and Cybernetics*, 8:83–88, 1983.
- [5194] T. Shinohara. Polynomial time inference of extended regular pattern languages. In *RIMS Symposium on Software Science and Engineering, Kyoto, 1982, Proceedings*, volume 147 of *Lecture Notes in Computer Science*, pages 115–127. Springer-Verlag, 1983.
- [5195] T. Shinohara. Some problems on inductive inference from positive data. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 41–58. Springer-Verlag, 1986.
- [5196] T. Shinohara. *Studies on Inductive Inference from Positive Data*. PhD thesis, Kyushu University, Kyushu, Japan, 1986.
- [5197] T. Shinohara. Inductive inference from positive data is powerful. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 97–110. Morgan Kaufmann, 1990.
- [5198] T. Shinohara. Inductive inference of monotonic formal systems from positive data. *New Generation Computing*, 8(4):371–384, 1991.
- [5199] T. Shinohara. Rich classes inferable from positive data: Length-bounded elementary formal systems. *Inform. Comput.*, 108(2):175–186, 1994.
- [5200] T. Shinohara and S. Arikawa. Pattern inference. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 259–291. Springer, 1995.
- [5201] T. Shinohara and H. Arimura. Inductive inference of unbounded unions of pattern languages from positive data. *Theoret. Comput. Sci.*, 241(1-2):191–209, 2000. Special issue for ALT '96.
- [5202] T. Shinohara and H. Ishizaka. On dimension reduction mappings for approximate retrieval of multi-dimensional data. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 224–231. Springer, 2001.
- [5203] T. Shinohara, C. H. Smith, and T. Zeugmann. Foreword. *Theoret. Comput. Sci.*, 298(1):1–4, 2003.

- [5204] Y. Shkel, M. Raginsky, and S. Verdú. Sequential prediction with coded side information under logarithmic loss. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 753–769. PMLR, 07–09 Apr. 2018.
- [5205] L. Shklar and H. Hirsh. Imposing bounds on the number of categories for incremental concept formation. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 3, pages 35–49. MIT Press, 1997.
- [5206] D. B. Shmoys. Approximation algorithms for clustering problems. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 100–101, New York, NY, 1999. ACM Press.
- [5207] Ilya Shmulevich, A. D. Korshunov, and J. Astola. Almost all monotone Boolean functions are polynomially learnable using membership queries. *Information Processing Letters*, 79(5):211–213, 2001.
- [5208] B. Shneiderman. Inventing discovery tools: Combining information visualization with data mining. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, page 58. Springer, 2001.
- [5209] J. Shrager. Theory change via view application in instructionless learning. *Machine Learning*, 2(3):247–276, 1987.
- [5210] Y. M. Shtarkov. Universal sequential coding of single messages. *Prob. Pered. Inf.*, 23:175–186, 1987.
- [5211] T. R. Shultz, D. Buckingham, and Y. Oshima-Takane. A connectionist model of the learning of personal pronouns in English. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 20, pages 347–362. MIT Press, 1994.
- [5212] T. R. Shultz and F. Rivest. Using knowledge to speed learning: A comparison knowledge-based cascade-correlation and multi-task learning. In *Proc. 17th International Conf. on Machine Learning*, pages 871–878. Morgan Kaufmann, San Francisco, CA, 2000.
- [5213] H. Shvaytser. Linear manifolds are learnable from positive examples. Unpublished manuscript, 1988.
- [5214] H. Shvaytser. Non-learnable classes of Boolean formulae that are closed under variable permutation. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 155–166, San Mateo, CA, 1988. Morgan Kaufmann.
- [5215] H. Shvaytser. A necessary condition for learning from positive examples. *Machine Learning*, 5:101–113, 1990.
- [5216] L.-Z. L. nad Si-Wei Luo, M. Tian, and L.-W. Zhao. Learning overcomplete representations with a generalized Gaussian prior. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 432–441, Berlin, 2006. Springer.
- [5217] H. T. Siegelmann and E. D. Sontag. On the computational power of neural nets. *J. of Comput. Syst. Sci.*, 50(1):132–150, 1995. Earlier version in 5th COLT, 1992.

- [5218] L. Siklossy and D. Sykes. Automatic program synthesis for example problems. In *Proc. 4th IJCAI*, pages 268–273. IJCAI, 1975.
- [5219] R. B. de Andrade e Silva and T. B. Ludermir. Obtaining simplified rule bases by hybrid learning. In *Proc. 17th International Conf. on Machine Learning*, pages 879–886. Morgan Kaufmann, San Francisco, CA, 2000.
- [5220] J. Šíma. Minimizing the quadratic training error of a sigmoid neuron is hard. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 92–105. Springer, 2001.
- [5221] J. Šíma. On the complexity of training a single perceptron with programmable synaptic delays. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 221–233. Springer, 2003.
- [5222] M. Simchowitz, K. Jamieson, and B. Recht. Best-of-K-bandits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1440–1489, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [5223] M. Simchowitz, K. Jamieson, and B. Recht. The simulator: Understanding adaptive sampling in the moderate-confidence regime. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1794–1834, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5224] H. A. Simon. Why should machines learn? In *Machine Learning, An Artificial Intelligence Approach*. Tioga, Palo Alto, California, 1983.
- [5225] H. A. Simon and K. Kotovsky. Human acquisition of concepts for sequential patterns. *Psych. Rev.*, 70:534–546, 1963.
- [5226] H. U. Simon. On the number of examples and stages needed for learning decision trees. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 303–313. Morgan Kaufmann, 1990.
- [5227] H. U. Simon. Learning decision lists and trees with equivalence-queries. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, number 904 in *Lecture Notes in Artificial Intelligence*, pages 322–336. Springer, 1995.
- [5228] H. U. Simon. General bounds on the number of examples needed for learning probabilistic concepts. *J. of Comput. Syst. Sci.*, 52(2):239–254, 1996. Earlier version in 6th COLT, 1993.
- [5229] H. U. Simon. Bounds on the number of examples needed for learning functions. *SIAM J. Comput.*, 26(3):751–763, 1997. first in EuroColt '93.
- [5230] H.-U. Simon. How many queries are needed to learn one bit of information? In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 1–13. Springer, 2001.

- [5231] H. U. Simon. How many missing answers can be tolerated by query learners? In *STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings*, volume 2285 of *Lecture Notes in Computer Science*, pages 384–395. Springer, 2002.
- [5232] H. U. Simon. On the complexity of working set selection. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 324–337. Springer, 2004.
- [5233] H. U. Simon. Spectral norm in learning theory: Some selected topics. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 13–27. Springer, Oct. 2006.
- [5234] H. U. Simon. On the complexity of working set selection. *Theoret. Comput. Sci.*, 382(3):262–279, 2007. Special issue for ALT 2004.
- [5235] H. U. Simon. Smart PAC-learners. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 353–367, Berlin/Heidelberg, 2009. Springer.
- [5236] H. U. Simon. An almost optimal PAC algorithm. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1552–1563, Paris, France, 03–06 July 2015. PMLR.
- [5237] H. U. Simon and E. Tomita. Guest editors’ foreword. *Theoret. Comput. Sci.*, 387(1):1–3, 2007. Special issue for ALT 2005.
- [5238] H. U. Simon and S. Zilles. Open problem: Recursive teaching dimension versus VC dimension. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1770–1772, Paris, France, 03–06 July 2015. PMLR.
- [5239] B. Singer and M. Veloso. Learning to predict performance from formula modeling and training data. In *Proc. 17th International Conf. on Machine Learning*, pages 887–894. Morgan Kaufmann, San Francisco, CA, 2000.
- [5240] B. Singer and M. Veloso. Learning to generate fast signal processing implementations. In *Proc. 18th International Conf. on Machine Learning*, pages 529–536. Morgan Kaufmann, San Francisco, CA, 2001.
- [5241] B. Singer and M. Veloso. Learning to construct fast signal processing implementations. *Journal of Machine Learning Research*, 3:887–919, 2002.
- [5242] Y. Singer. Guest editors’ introduction. *Machine Learning*, 43(3):171–172, 2001.
- [5243] A. Singh, R. Nowak, and C. Scott. Adaptive hausdorff estimation of density level sets. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 491–502. Omnipress, 2008.
- [5244] M. Singh and G. M. Provan. A comparison of inductive algorithms for selective and non-selective Bayesian classifiers. In *Proc. 12th International Conference on Machine Learning*, pages 497–505. Morgan Kaufmann, 1995.

- [5245] M. Singh and G. M. Provan. Efficient learning of selective Bayesian network classifiers. In *Proc. 13th International Conference on Machine Learning*, pages 453–461. Morgan Kaufmann, 1996.
- [5246] S. Singh. Introduction. *Machine Learning*, 49(2-3):107–109, 2002.
- [5247] S. Singh, T. Jaakkola, M. L. Littman, and C. Szepesvári. Convergence results for single-step on-policy reinforcement-learning algorithms. *Machine Learning*, 38(3):287–308, 2000.
- [5248] S. P. Singh. Transfer of learning by composing solutions of elemental sequential tasks. *Machine Learning*, 8:323–339, 1992.
- [5249] S. P. Singh, T. Jaakkola, and M. I. Jordan. Learning without state-estimation in partially observable Markovian decision processes. In *Proc. 11th International Conference on Machine Learning*, pages 284–292. Morgan Kaufmann, 1994.
- [5250] S. P. Singh and R. S. Sutton. Reinforcement learning with replacing eligibility traces. *Machine Learning*, 22:123–158, 1996.
- [5251] S. P. Singh and R. C. Yee. An upper bound on the loss from approximate optimal-value functions. *Machine Learning*, 16(3):227–233, 1994.
- [5252] R. Siromoney, D. G. Thomas, K. G. Subramanian, and V. R. Dare. Thue systems and DNA – A learning algorithm for a subclass. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 314–327. Springer-Verlag, 1993.
- [5253] R. C. Sison, M. Numao, and M. Shimura. Multistrategy discovery and detection of novice programmer errors. *Machine Learning*, 38(1/2):157–180, 2000.
- [5254] M. Sitharam and T. Straney. Derandomized learning of Boolean functions. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 100–115. Springer, 1997.
- [5255] K.-Y. Siu, A. Dembo, and T. Kailath. On the perceptron learning algorithm on data with high precision. *J. of Comput. Syst. Sci.*, 48(2):347–356, 1994.
- [5256] P. Siyari and M. Gallé. The generalized smallest grammar problem. In S. Verwer, M. van Zanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 79–92, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [5257] D. B. Skalak. Prototype and feature selection by sampling and random mutation hill climbing algorithms. In *Proc. 11th International Conference on Machine Learning*, pages 293–301. Morgan Kaufmann, 1994.
- [5258] T. A. Slaman and R. M. Solovay. When oracles do not help. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 379–383. Morgan Kaufmann, 1991.
- [5259] S. Slattery and T. Mitchell. Discovering test set regularities in relational domains. In *Proc. 17th International Conf. on Machine Learning*, pages 895–902. Morgan Kaufmann, San Francisco, CA, 2000.

- [5260] D. Sleeman. Introduction to the abstracts of the invited talks presented at ML92 conference in aberdeen, 1-3 july 1992. *Machine Learning*, 14:303–303, 1994.
- [5261] D. Sleeman, H. Hirsh, I. Ellery, and In-Yung Kim. Extending domain theories: Two case studies in student modeling. *Machine Learning*, 5:11–37, 1990.
- [5262] A. Slivkins. Contextual bandits with similarity information. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 679–702, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5263] A. Slivkins. Monotone multi-armed bandit allocations. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 829–834, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5264] A. Slivkins and E. Upfal. Adapting to a changing environment: the brownian restless bandits. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 343–354. Omnipress, 2008.
- [5265] R. Sloan. Types of noise in data for concept learning. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 91–96, San Mateo, CA, 1988. Morgan Kaufmann.
- [5266] R. Sloan. Special issue on computational learning theory. *J. Comput. Syst. Sci.*, 70(4):433–434, 2005. Special issue on COLT 2002.
- [5267] R. H. Sloan. Some notes on Chernoff bounds. (Unpublished), 1987.
- [5268] R. H. Sloan. *Computational Learning Theory: New Models and Algorithms*. PhD thesis, MIT, 1989. Issued as MIT/LCS/TR-448.
- [5269] R. H. Sloan. Corrigendum to types of noise in data for concept learning. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, page 450, New York, NY, 1992. ACM Press.
- [5270] R. H. Sloan. Four types of noise in data for PAC learning. *Information Processing Letters*, 54:157–162, 1995.
- [5271] R. H. Sloan and B. Szörényi. Revising projective DNF in the presence of noise. In *Proc. Kalmár Workshop on Logic and Computer Science*, pages 143–152, Szeged, Hungary, 2003. Dept. of Informatics, University of Szeged.
- [5272] R. H. Sloan, B. Szörényi, and G. Turán. Projective DNF formulae and their revision. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 625–639. Springer, 2003.
- [5273] R. H. Sloan, B. Szörényi, and G. Turán. Revising threshold functions. *Theoret. Comput. Sci.*, 382(3):198–208, 2007. Special issue for ALT 2004.
- [5274] R. H. Sloan and G. Turán. Learning with queries but incomplete information. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 237–245, New York, NY, 1994. ACM Press.

- [5275] R. H. Sloan and G. Turán. Learning from incomplete boundary queries using split graphs and hypergraphs. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, number 1208 in Lecture Notes in Artificial Intelligence, pages 38–50. Springer, 1997.
- [5276] R. H. Sloan and G. Turán. On theory revision with queries. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 41–52, New York, NY, 1999. ACM Press.
- [5277] N. R. Smalheiser. The Arrowsmith project: 2005 status report,. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 11–11. Springer, Oct. 2005.
- [5278] W. D. Smart and L. P. Kaelbling. Practical reinforcement learning in continuous spaces. In *Proc. 17th International Conf. on Machine Learning*, pages 903–910. Morgan Kaufmann, San Francisco, CA, 2000.
- [5279] R. Smetsers, M. Volpato, F. Vaandrager, and S. Verwer. Bigger is not always better: on the quality of hypotheses in active automata learning. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 167–181, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [5280] A. Smith and A. Thakurta. Differentially private feature selection via stability arguments, and the robustness of the lasso. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 819–850. JMLR.org, 2013.
- [5281] A. M. Smith and M. K. Warmuth. Learning rotations. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, page 517. Omnipress, 2008.
- [5282] C. H. Smith. *Hierarchies of Identification Criteria for Mechanized Inductive Inference*. PhD thesis, SUNY Buffalo, 1979.
- [5283] C. H. Smith. An inductive inference bibliography. Technical report, Purdue University Computer Science Dept., CSD TR 323, 1979.
- [5284] C. H. Smith. The power of pluralism for automatic program synthesis. *J. ACM*, 29(4):1144–1165, 1982.
- [5285] C. H. Smith. *A Recursive Introduction to the Theory of Computation*. Springer-Verlag, New York, Berlin, Heidelberg, 1994.
- [5286] C. H. Smith. Three decades of team learning. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 211–228. Springer-Verlag, 1994.
- [5287] C. H. Smith and W. I. Gasarch. Recursion theoretic models of learning: some results and intuitions. *Annals of Mathematics and Artificial Intelligence*, 15(2):151–166, 1995.



- [5288] C. H. Smith and M. Velauthapillai. On the inference of programs approximately computing the desired function. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 164–176. Springer-Verlag, 1986.
- [5289] C. H. Smith and M. Velauthapillai. On the inference of approximate programs. *Theoret. Comput. Sci.*, 77(3):249–266, 1990.
- [5290] C. H. Smith, R. Wiehagen, and T. Zeugmann. Classifying predicates and languages. *International Journal of Foundations of Computer Science*, 8(1):15–41, 1997.
- [5291] D. R. Smith. A class of synthesizable LISP programs. Technical report, Duke University Computer Science Dept., CS-1977-4, 1977.
- [5292] D. R. Smith. A survey of the synthesis of LISP programs from examples. In *Automatic Program Construction Techniques*. MacMillan Co., 1982.
- [5293] A. Smola, A. Gretton, L. Song, and B. Schölkopf. A hilbert space embedding for distributions. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 13–31, Berlin, Oct. 2007. Springer.
- [5294] A. J. Smola and R. Kondor. Kernels and regularization on graphs. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 144–158. Springer, 2003.
- [5295] A. J. Smola, S. Mika, B. Schölkopf, and R. C. Williamson. Regularized principal manifolds. *Journal of Machine Learning Research*, 1:179–209, 2001.
- [5296] A. J. Smola and B. Schölkopf. On a kernel-based method for pattern recognition, regression, approximation, and operator inversion. *Algorithmica*, 22(1/2):211–231, 1998.
- [5297] A. J. Smola and B. Schölkopf. Sparse greedy matrix approximation for machine learning. In *Proc. 17th International Conf. on Machine Learning*, pages 911–918. Morgan Kaufmann, San Francisco, CA, 2000.
- [5298] A. J. Smola, R. C. Williamson, S. Mika, and B. Schölkopf. Regularized principal manifolds. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 214–229. Springer, 1999.
- [5299] J. Smotrovs. Closedness properties in team learning of recursive functions. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 79–93. Springer, 1997.
- [5300] B. Smyth and E. McKenna. An efficient and effective procedure for updating a competence model for case-based reasoner. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 357–368. Springer, 2000.

- [5301] P. Smyth. Probability density estimation and local basis function neural networks. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 15, pages 233–248. MIT Press, 1994.
- [5302] P. Smyth. Learning with probabilistic supervision. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 9, pages 163–182. MIT Press, 1995.
- [5303] P. Smyth. A guided tour of finite mixture models: from pearson to the web. In *Proc. 18th International Conf. on Machine Learning*, pages 642–642. Morgan Kaufmann, San Francisco, CA, 2001.
- [5304] P. Smyth. Data-driven discovery using probabilistic hidden variable models. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, page 28. Springer, Oct. 2006.
- [5305] P. Smyth, A. Gray, and U. M. Fayyad. Retrofitting decision tree classifiers using kernel density estimation. In *Proc. 12th International Conference on Machine Learning*, pages 506–514. Morgan Kaufmann, 1995.
- [5306] P. Smyth and D. Wolpert. Linearly combining density estimators via stacking. *Machine Learning*, 36(1/2):59–83, 1999.
- [5307] S. Soderland. Learning information extraction rules for semi-structured and free text. *Machine Learning*, 34(1-3):233–272, 1999.
- [5308] D. D. Sofronie, I. Tellier, and M. Tommasi. A tool for language learning based on categorial grammars and semantic information. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 303–305. Springer, 2002.
- [5309] L.-K. Soh and C. Tsatsoulis. Using learning by discovery to segment remotely sensed images. In *Proc. 17th International Conf. on Machine Learning*, pages 919–926. Morgan Kaufmann, San Francisco, CA, 2000.
- [5310] S. A. Solla, E. Levin, and M. Fleisher. Accelerated learning in layered neural networks. *Complex Systems*, 2:625–639, 1988.
- [5311] P. Sollich. Bayesian methods for support vector machines: Evidence and predictive class probabilities. *Machine Learning*, 46(1-3):21–52, 2002.
- [5312] R. J. Solomonoff. A formal theory of inductive inference: Part 1. *Inform. Control*, 7(1):1–22, 1964.
- [5313] R. J. Solomonoff. A formal theory of inductive inference: Part 2. *Inform. Control*, 7(2):224–254, 1964.
- [5314] R. J. Solomonoff. Inductive inference theory - a unified approach to problems in pattern recognition and artificial intelligence. In *Proc. 4th IJCAI*, pages 274–280. IJCAI, 1975.
- [5315] R. J. Solomonoff. Complexity-based induction systems: comparisons and convergence theorems. *IEEE Trans. Inform. Theory*, IT-24:422–432, 1978.

- [5316] R. J. Solomonoff. The discovery of algorithmic probability. *J. Comput. Syst. Sci.*, 55(1):73–88, 1997. Special issue for EuroCOLT '95.
- [5317] M. van Someren E. Plaza, A. Aamodt, A. Ram, W. van de Velde. Integrated learning architectures. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 429–441. Springer-Verlag, 1993.
- [5318] H. Sompolinsky, H. S. Seung, and N. Tishby. Learning curves in large neural networks. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 112–127, San Mateo, CA, 1991. Morgan Kaufmann.
- [5319] H. Sompolinsky, N. Tishby, and H. S. Seung. Learning from examples in large neural networks. *Phys. Rev. Lett.*, 65:1683–1686, 1990.
- [5320] E. D. Sontag. Critical points for least-squares problems involving certain analytic functions, with applications to sigmoidal nets. Technical Report Unpublished, Rutgers University, New Brunswick, Department of Mathematics, 1995.
- [5321] E. D. Sontag and H. J. Sussmann. Backpropagation can give rise to spurious local minima even for networks without hidden layers. *Complex Systems*, 3(1):91–106, Feb. 1989.
- [5322] E. D. Sontag and H. J. Sussmann. Back propagation separates where perceptrons do. *Neural Networks*, 4(3), 1991.
- [5323] N. Soonthornphisaj and B. Kijirikul. The effects of different feature sets on the web page categorization problem using the iterative cross-training algorithm. In *ICEIS 2001, Proc. of the 3rd International Conference on Enterprise Information Systems, Setubal, Portugal, July 7–10, 2001*, volume 1, pages 404–410, 2001.
- [5324] W. M. Spears, K. A. D. Jong, T. Bäck, D. B. Fogel, and H. de Garis. An overview of evolutionary computation. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 442–459. Springer-Verlag, 1993.
- [5325] D. A. Spielman, H. Wang, and J. Wright. Exact recovery of sparsely-used dictionaries. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 37.1–37.18, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5326] L. Spirkovska and M. B. Reid. Higher-order neural networks applied to 2D and 3D object recognition. *Machine Learning*, 15(2):169–199, 1994.
- [5327] V. I. Spitzkovsky, H. Alshawi, and D. Jurafsky. Bootstrapping dependency grammar inducers from incomplete sentence fragments via austere models. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 189–194, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5328] B. Squires and C. Sammut. Automatic speaker recognition: an application of machine learning. In *Proc. 12th International Conference on Machine Learning*, pages 515–521. Morgan Kaufmann, 1995.

- [5329] N. Srebro. Are there local maxima in the infinite-sample likelihood of gaussian mixture estimation? In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 628–629, Berlin, 2007. Springer.
- [5330] N. Srebro. How good is a kernel when used as a similarity measure? In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 323–335, Berlin, 2007. Springer.
- [5331] N. Srebro and S. Ben-David. Learning bounds for support vector machines with learned kernels. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 169–183, Berlin, 2006. Springer.
- [5332] N. Srebro and A. Shraibman. Rank, trace-norm and max-norm. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 545–560. Springer, 2005.
- [5333] K. Sridharan and S. M. Kakade. An information theoretic framework for multi-view learning. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 403–414. Omnipress, 2008.
- [5334] M. Sridharan and G. Tesauro. Multi-agent Q-learning and regression trees for automated pricing decisions. In *Proc. 17th International Conf. on Machine Learning*, pages 927–934. Morgan Kaufmann, San Francisco, CA, 2000.
- [5335] A. Srinivasan. Extracting context-sensitive models in inductive logic programming. *Machine Learning*, 44(3):301–324, 2001.
- [5336] A. Srinivasan, R. D. King, and M. E. Bain. An empirical study of the use of relevance information in inductive logic programming. *Journal of Machine Learning Research*, 4:369–383, 2003.
- [5337] B. K. Sriperumbudur, A. Gretton, K. Fukumizu, G. R. G. Lanckriet, and B. Schölkopf. Injective hilbert space embeddings of probability measures. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 111–122. Omnipress, 2008.
- [5338] E. Stabler. Towards a rationalist theory of language acquisition. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 21–32, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [5339] I. Stahl. Predicate invention in ILP - an overview. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 313–322. Springer-Verlag, 1993.
- [5340] I. Stahl. The appropriateness of predicate invention as bias shift operation in ILP. *Machine Learning*, 20(1/2):95–117, 1995.
- [5341] I. Stahl, B. Tausend, and R. Wirth. Two methods for improving inductive logic programming systems. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 41–55. Springer-Verlag, 1993.

- [5342] I. Stainvas and D. Lowe. A generative model for separating illumination and reflectance from images. *Journal of Machine Learning Research*, 4:1499–1519, 2003.
- [5343] B. Starkie. Inferring attribute grammars with structured data for natural language processing. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 237–248. Springer, 2002.
- [5344] B. Starkie, F. Coste, and M. van Zaanen. The omphalos context-free grammar learning competition. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 16–27. Springer, 2004.
- [5345] B. Starkie and H. Fernau. The boisdale algorithm – an induction method for a subclass of unification grammar from positive data. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 235–247. Springer, 2004.
- [5346] B. Starkie, G. Findlow, K. Ho, A. Hui, L. Law, L. Lightwood, S. Michnowicz, and C. Walder. Lyrebird<sup>tm</sup>: Developing spoken dialog systems using examples. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 309–311. Springer, 2002.
- [5347] B. Starkie, M. van Zaanen, and D. Estival. The tenjinno machine translation competition. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 214–226, Berlin, 2006. Springer.
- [5348] B. Steffen, F. Howar, and M. Isberner. Active automata learning: From DFAs to interface programs and beyond. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 195–209, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5349] H. Stehouwer and M. van Zaanen. Enhanced suffix arrays as language models: Virtual  $k$ -testable languages. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 305–308. Springer, 2010.
- [5350] W. Stein. Consistent polynomial identification in the limit. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 424–438. Springer, 1998.
- [5351] J. Steinhardt and J. Duchi. Minimax rates for memory-bounded sparse linear regression. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1564–1587, Paris, France, 03–06 July 2015. PMLR.
- [5352] J. Steinhardt, G. Valiant, and S. Wager. Memory, communication, and statistical queries. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1490–1516, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.

- [5353] T. Steinke and J. Ullman. Interactive fingerprinting codes and the hardness of preventing false discovery. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1588–1628, Paris, France, 03–06 July 2015. PMLR.
- [5354] I. Steinwart. On the influence of the kernel on the consistency of support vector machines. *Journal of Machine Learning Research*, 2:67–93, 2001.
- [5355] I. Steinwart. Sparseness of support vector machines. *Journal of Machine Learning Research*, 4:1071–1105, 2003.
- [5356] I. Steinwart. Adaptive density level set clustering. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 703–738, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5357] I. Steinwart, D. Hush, and C. Scovel. A classification framework for anomaly detection. *Journal of Machine Learning Research*, 6:211–232, 2005.
- [5358] I. Steinwart, D. Hush, and C. Scovel. Function classes that approximate the Bayes risk. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 79–93, Berlin, 2006. Springer.
- [5359] I. Steinwart, C. Pasin, R. Williamson, and S. Zhang. Elicitation and identification of properties. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 482–526, Barcelona, Spain, 13–15 June 2014. PMLR.
- [5360] I. Steinwart and C. Scovel. Fast rates for support vector machines. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 279–294. Springer, 2005.
- [5361] G. Stempfel and L. Ralaivola. Learning kernel perceptrons on noisy data using random projections. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 328–342, Berlin, Oct. 2007. Springer.
- [5362] F. Stephan. Learning via queries and oracles. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 162–169, New York, NY, 1995. ACM Press.
- [5363] F. Stephan. Noisy inference and oracles. *Theoret. Comput. Sci.*, 185(1):129–157, 1997. Special issue for ALT '95.
- [5364] F. Stephan and S. A. Terwijn. The complexity of universal text-learners. *Inform. Comput.*, 154(2):149–166, 1999.
- [5365] F. Stephan and S. A. Terwijn. Counting extensional differences in BC-learning. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 256–269. Springer, 2000.

- [5366] F. Stephan and Y. Ventsov. Learning algebraic structures from text using semantical knowledge. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 321–335. Springer, 1998.
- [5367] F. Stephan and Y. Ventsov. Learning algebraic structures from text. *Theoret. Comput. Sci.*, 268(2):221–273, 2001. Special issue for ALT '98.
- [5368] F. Stephan, R. Yoshinaka, and T. Zeugmann. On the parameterised complexity of learning patterns. In E. Gelenbe, R. Lent, and G. Sakellari, editors, *Computer and Information Sciences II, 26th International Symposium on Computer and Information Sciences*, pages 277–281. Springer, 2011.
- [5369] F. Stephan and T. Zeugmann. Learning classes of approximations to non-recursive functions. Technical Report DOI-TR-166, Department of Informatics, Kyushu University, July 1999.
- [5370] F. Stephan and T. Zeugmann. On the uniform learnability of approximations to non-recursive functions. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 276–290. Springer, 1999.
- [5371] F. Stephan and T. Zeugmann. Average-case complexity of learning polynomials. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 59–68. Morgan Kaufmann, San Francisco, 2000.
- [5372] F. Stephan and T. Zeugmann. Learning classes of approximations to non-recursive functions. *Theoret. Comput. Sci.*, 288(2):309–341, 2002. Special issue for ALT '99.
- [5373] V. Sterzing and B. Schürmann. Recurrent neural networks with time-dependent inputs and outputs. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 7, pages 115–130. MIT Press, 1995.
- [5374] A. Stolcke and S. Omohundro. Inducing probabilistic grammars by Bayesian model merging. In *Proc. 2nd International Colloquium on Grammatical Inference - ICGI '94*, volume 862 of *Lecture Notes in Artificial Intelligence*, pages 106–118. Springer-Verlag, 1994.
- [5375] S. J. Stolfo. *Automatic Discovery Heuristics for Nondeterministic Programs From Sample Execution Traces*. PhD thesis, New York University, 1979.
- [5376] S. J. Stolfo and M. C. Harrison. Automatic discovery of heuristics for nondeterministic programs. Technical report, New York University Computer Science Dept., No. 7, 1979.
- [5377] G. Stoltz and G. Lugosi. Internal regret in on-line portfolio selection. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 403–417. Springer, 2003.
- [5378] P. Stone. TPOT-RL applied to network routing. In *Proc. 17th International Conf. on Machine Learning*, pages 935–942. Morgan Kaufmann, San Francisco, CA, 2000.
- [5379] P. Stone and R. S. Sutton. Scaling reinforcement learning toward RoboCup soccer. In *Proc. 18th International Conf. on Machine Learning*, pages 537–544. Morgan Kaufmann, San Francisco, CA, 2001.

- [5380] P. Stone and M. M. Veloso. Layered learning. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 369–381. Springer, 2000.
- [5381] H. Stoppiglia, G. Dreyfus, R. Dubois, and Y. Oussar. Ranking a random feature for variable and feature selection. *Journal of Machine Learning Research*, 3:1399–1414, 2003.
- [5382] D. G. Stork. Toward a computational theory of data acquisition and truthing. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 194–207. Springer, 2001.
- [5383] G. D. Stormo and D. Haussler. Optimally parsing a sequence into different classes based on multiple types of information. In *Proc. of Second Int. Conf. on Intelligent Systems for Molecular Biology*, pages 369–375, Menlo Park, CA, Aug. 1994. AAAI/MIT Press.
- [5384] W. N. Street. A neural network model for prognostic prediction. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 540–546. Morgan Kaufmann, San Francisco, CA, 1998.
- [5385] W. N. Street, O. L. Mangasarian, and W. H. Wolberg. An inductive learning approach to prognostic prediction. In *Proc. 12th International Conference on Machine Learning*, pages 522–530. Morgan Kaufmann, 1995.
- [5386] A. Strehl and J. Ghosh. Cluster ensembles — A knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research*, 3:583–617, 2002.
- [5387] M. Strens. A Bayesian framework for reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 943–950. Morgan Kaufmann, San Francisco, CA, 2000.
- [5388] M. J. A. Strens and A. W. Moore. Direct policy search using paired statistical tests. In *Proc. 18th International Conf. on Machine Learning*, pages 545–552. Morgan Kaufmann, San Francisco, CA, 2001.
- [5389] M. J. A. Strens and A. W. Moore. Policy search using paired comparisons. *Journal of Machine Learning Research*, 3:921–950, 2002.
- [5390] K. Strother-Garcia, J. Heinz, and H. J. Hwangbo. Using model theory for grammatical inference: a case study from phonology. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 66–78, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [5391] J. M. Stuart and E. Bradley. Learning the grammar of dance. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 547–555. Morgan Kaufmann, San Francisco, CA, 1998.
- [5392] D. Subramanian. Shifting vocabulary bias in speedup learning. *Machine Learning*, 20(1/2):155–191, 1995.
- [5393] D. Suc and I. Bratko. Problem decomposition for behavioural cloning. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 382–391. Springer, 2000.



- [5394] D. Šuc and I. Bratko. Induction of qualitative trees. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 442–453. Springer, 2001.
- [5395] D. Suehiro, K. Hatano, S. Kijima, E. Takimoto, and K. Nagano. Online prediction under submodular constraints. In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 260–274. Berlin, Heidelberg, New York, 2012. Springer.
- [5396] D. Suehiro, K. Hatano, and E. Takimoto. Approximate reduction from AUC maximization to 1-norm soft margin optimization. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 324–337. Berlin, Heidelberg, New York, 2011. Springer.
- [5397] N. Suematsu, A. Hayashi, and S. Li. A Bayesian approach to model learning in non-Markovian environments. In *Proc. 14th International Conference on Machine Learning*, pages 349–357. Morgan Kaufmann, 1997.
- [5398] T. Sugawara and V. Lesser. Learning to improve coordinated actions in cooperative distributed problem-solving. *Machine Learning*, 33(2/3):129–153, 1998.
- [5399] N. Sugimoto. Learnability of translations from positive examples. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 169–178. Springer, 1998.
- [5400] N. Sugimoto, K. Hirata, and H. Ishizaka. Constructive learning of translations based on dictionaries. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 177–184. Springer, 1996.
- [5401] N. Sugimoto, T. Toyoshima, S. Shimozone, and K. Hirata. Constructive learning of context-free languages with a subpansive tree. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 270–283. Springer, 2000.
- [5402] M. Sugiyama, E. Hirowatari, H. Tsuiki, and A. Yamamoto. Learning figures with the Hausdorff metric by fractals. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 315–329. Berlin, Heidelberg, New York, 2010. Springer.
- [5403] M. Sugiyama and K.-R. Müller;. The subspace information criterion for infinite dimensional hypothesis spaces. *Journal of Machine Learning Research*, 3:323–359, 2002.
- [5404] M. Sugiyama and H. Ogawa. Theoretical and experimental evaluation of the subspace information criterion. *Machine Learning*, 48(1/3):25–50, 2002.
- [5405] P. D. Summers. *Program Construction From Examples*. PhD thesis, Yale University, 1976.

- [5406] P. D. Summers. A methodology for LISP program construction from examples. *J. ACM*, 24(1):161–175, Jan. 1977.
- [5407] X. Sun and A. Nobel. Significance and recovery of block structures in binary matrices with noise. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 109–122, Berlin, 2006. Springer.
- [5408] P. Sunehag and M. Hutter. Consistency of feature markov processes. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 360–374, Berlin, Heidelberg, New York, 2010. Springer.
- [5409] P. Sunehag and M. Hutter. Axioms for rational reinforcement learning. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 338–352, Berlin, Heidelberg, New York, 2011. Springer.
- [5410] P. Suppes, M. Böttner, and L. Liang. Comprehension grammars generated from machine learning of natural languages. *Machine Learning*, 19(2):133–152, 1995.
- [5411] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, 1988.
- [5412] R. S. Sutton. Reinforcement learning architectures for animats. In *First International Conference on Simulation of Adaptive Behavior*, 1991.
- [5413] R. S. Sutton. TD models: modeling the world at a mixture of time scales. In *Proc. 12th International Conference on Machine Learning*, pages 531–539. Morgan Kaufmann, 1995.
- [5414] R. S. Sutton. Open theoretical questions in reinforcement learning. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 11–17. Springer, 1999.
- [5415] R. S. Sutton and A. G. Barto. Toward a modern theory of adaptive networks: Expectation and prediction. *Psych. Rev.*, 88(2):135–170, 1981.
- [5416] R. S. Sutton and A. G. Barto. An adaptive network that constructs and uses an internal model of its world. , 1985.
- [5417] R. S. Sutton, D. Precup, and S. Singh. Intra-option learning about temporally abstract actions. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 556–564. Morgan Kaufmann, San Francisco, CA, 1998.
- [5418] M. Suwa and H. Motoda. The use of abstract primitives in representing the meaning of verbs for understanding metaphors. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 231–242. IOS Press, 1992.
- [5419] M. Suwa and H. Motoda. A perceptual criterion for visually controlling learning. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 356–369. Springer-Verlag, 1993.

- [5420] E. Suzuki. In pursuit of interesting patterns with undirected discovery of exception rules. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 504–517. Springer, 2001.
- [5421] J. Suzuki. Learning Bayesian belief networks based on the minimum description length principle: an efficient algorithm using the B & B technique. In *Proc. 13th International Conference on Machine Learning*, pages 462–470. Morgan Kaufmann, 1996.
- [5422] J. Suzuki. Learning Bayesian belief networks based on the MDL principle: An efficient algorithm using the branch and bound technique. *IEICE Transactions on Information and Systems*, E82-D(2):356–367, 1999.
- [5423] J. Suzuki. Learning Bayesian belief networks based on the minimum description length principle: Basic properties. *IEICE Transactions on Fundamentals*, E82-A(10):2237–2245, 1999.
- [5424] T. Suzuki. PAC-Bayesian bound for gaussian process regression and multiple kernel additive model. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 8.1–8.20, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5425] Y. Suzuki, R. Akanuma, T. Shoudai, T. Miyahara, and T. Uchida. Polynomial time inductive inference of ordered tree patterns with internal structured variables from positive data. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 169–184. Springer, 2002.
- [5426] Y. Suzuki, T. Shoudai, S. Matsumoto, T. Uchida, and T. Miyahara. Efficient learning of ordered and unordered tree patterns with contractible variables. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 114–128. Springer, 2003.
- [5427] Y. Suzuki, T. Shoudai, T. Uchida, and T. Miyahara. Ordered term tree languages which are polynomial time inductively inferable from positive data. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 188–202. Springer, 2002.
- [5428] Y. Suzuki, T. Shoudai, T. Uchida, and T. Miyahara. Ordered term tree languages which are polynomial time inductively inferable from positive data. *Theoret. Comput. Sci.*, 350(1):63–90, 2006. Special issue for ALT 2002.
- [5429] B. K. Sy and A. K. Gupta. Data mining approach based on information-statistical analysis: Application to temporal-spatial data. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 128–140. Springer, 2001.
- [5430] Z. Szabó. Stratified inductive hypothesis generation. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 177–195. Springer-Verlag, 1986.
- [5431] I. Szita and C. Szepesvári. Agnostic KWIK learning and efficient approximate reinforcement learning. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual*

- Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 739–772, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5432] I. Szita, B. Takács, and A. Lőrincz.  $\epsilon$ -MDPs: Learning in varying environments. *Journal of Machine Learning Research*, 3:145–174, 2002.
- [5433] B. Szörényi. Characterizing statistical query learning: Simplified notions and proofs. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 186–200, Berlin/Heidelberg, 2009. Springer.
- [5434] T. Tabe and T. Zeugmann. Two variations of inductive inference of languages from positive data. Technical Report RIFIS-TR-CS-105, RIFIS, Kyushu University 33, 1995.
- [5435] P. Tadepalli and T. G. Dietterich. Hierarchical explanation-based reinforcement learning. In *Proc. 14th International Conference on Machine Learning*, pages 358–366. Morgan Kaufmann, 1997.
- [5436] P. Tadepalli and D. Ok. Scaling up average reward reinforcement learning by approximating the domain models and the value function. In *Proc. 13th International Conference on Machine Learning*, pages 471–479. Morgan Kaufmann, 1996.
- [5437] P. Tadepalli and S. Russell. Learning from examples and membership queries with structured determinations. *Machine Learning*, 32(3):245–295, 1998.
- [5438] V. Tadić. Convergence analysis of temporal-difference learning algorithms with linear function approximation. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 193–202, New York, NY, 1999. ACM Press.
- [5439] V. Tadic. On the asymptotic behaviour of a constant stepsize temporal-difference learning algorithm. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 126–137. Springer, 1999.
- [5440] V. Tadic. On the convergence of temporal-difference learning with linear function approximation. *Machine Learning*, 42(3):241–267, 2001.
- [5441] K. S. Tae and D. J. Cook. Experimental knowledge acquisition for planning. In *Proc. 13th International Conference on Machine Learning*, pages 480–488. Morgan Kaufmann, 1996.
- [5442] H. Taira and M. Haruno. Text categorization using transductive boosting. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 454–465. Springer, 2001.
- [5443] I. Tajika, E. Takimoto, and A. Maruoka. An on-line prediction algorithm combining several prediction strategies in the shared bet model. *IEICE Transactions on Information and Systems*, E82-D(2):348–355, 1999.
- [5444] Y. Tajima and G. Kikui. An example distribution for probabilistic query learning of simple deterministic languages. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 182–192, Kyoto, Japan, 17–19 Sept. 2014. PMLR.

- [5445] Y. Tajima and Y. Kotani. Polynomial time probabilistic learning of a subclass of linear languages with queries. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 187–199. Springer, 2008.
- [5446] Y. Tajima, Y. Kotani, and M. Terada. An analysis of examples and a search space for PAC learning of simple deterministic languages with membership queries. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 288–289. Springer, 2004.
- [5447] Y. Tajima and M. Terada. A PAC learnability of simple deterministic languages. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 249–260. Springer, 2002.
- [5448] Y. Tajima and E. Tomita. A polynomial time learning algorithm simple deterministic languages via membership queries and a representative sample. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 284–297. Springer, 2000.
- [5449] Y. Tajima, E. Tomita, and M. Wakatsuki. Polynomial time learnability of simple deterministic languages from MAT and a representative sample. *IEICE Transactions on Information and Systems*, E83-D(4):757–765, 2000.
- [5450] Y. Tajima, E. Tomita, M. Wakatsuki, and M. Terada. Polynomial time learning of simple deterministic languages via queries and a representative sample. *Theoret. Comput. Sci.*, 329(1-3):203–221, 2004.
- [5451] Y. Takada. Grammatical inference for even linear languages based on control sets. *Inform. Process. Lett.*, 28:193–199, 1988.
- [5452] Y. Takada. Learning semilinear sets from examples and via queries. *Theoret. Comput. Sci.*, 104(2):207–233, 1992.
- [5453] Y. Takada. A hierarchy of language families learnable by regular language learners. In *Proc. 2nd International Colloquium on Grammatical Inference - ICGI '94*, volume 862 of *Lecture Notes in Artificial Intelligence*, pages 16–24. Springer-Verlag, 1994.
- [5454] Y. Takada. Learning formal languages based on control sets. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 316–339. Springer, 1995.
- [5455] Y. Takada and T. Y. Nishida. A note on grammatical inference of slender context-free languages. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 117–125. Springer, 1996.
- [5456] Y. Takada, Y. Sakakibara, and T. Ohtani. ACE: a syntax-directed editor customizable from examples and queries. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 126–138. IOS Press, 1992.

- [5457] Y. Takada and R. Siromoney. On identifying DNA splicing systems from examples. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 305–319. Springer-Verlag, 1992.
- [5458] K. Takahashi and M. Hagiya. Searching for mutual exclusion algorithms using BDDs. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 1–18. Springer, 2001.
- [5459] I. Takamatsu, M. Kobayashi, H. Tokunaga, and A. Yamamoto. Computing characteristic sets of bounded unions of polynomial ideals. In *JSAI 2007 Conference and Workshops, Miyazaki, Japan, June 18-22, 2007, Revised Selected Papers*, volume 4914 of *Lecture Notes in Artificial Intelligence*, pages 318–329. Springer, Berlin, 2008.
- [5460] A. Takano. Association computation for information access. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, page 15. Springer, 2003.
- [5461] A. Takasu. Classification of object sequences using syntactical structure. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 318–326. Springer, 2001.
- [5462] M. Takeda, T. Fukuda, and I. Nanri. Mining from literary texts: Pattern discovery and similarity computation. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 518–531. Springer, 2001.
- [5463] M. Takeda, T. Fukuda, I. Nanri, M. Yamasaki, and K. Tamari. Discovering instances of poetic allusion from anthologies of classical japanese poems. *Theoret. Comput. Sci.*, 292(2):497–524, 2003.
- [5464] M. Takeda, T. Matsumoto, T. Fukuda, and I. Nanri. Discovering characteristic expressions in literary works. *Theoret. Comput. Sci.*, 292(2):525–546, 2003.
- [5465] J. Takeuchi. Some improved sample complexity bounds in the probabilistic PAC learning model. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 208–219. Springer-Verlag, 1993.
- [5466] J. Takeuchi, N. Abe, and S. ichi Amari. The lob-pass problem. *J. Comput. Syst. Sci.*, 61(3):523–557, Dec. 2000.
- [5467] E. Takimoto and K. Hatano. Efficient algorithms for combinatorial online prediction. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 22–32, Berlin, Heidelberg, New York, 2013. Springer.
- [5468] E. Takimoto, K. Hirai, and A. Maruoka. A simple algorithm for predicting nearly as well as the best pruning labeled with the best prediction values of a decision tree. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 385–400. Springer, 1997.

- [5469] E. Takimoto and A. M. S. Koya. Boosting based on divide and merge. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 127–141. Springer, 2004.
- [5470] E. Takimoto and A. Maruoka. Relationships between PAC-learning algorithms and weak Occam algorithms. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 151–162. IOS Press, 1992.
- [5471] E. Takimoto and A. Maruoka. Conservativeness and monotonicity for learning algorithms. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 377–383, New York, NY, 1993. ACM Press.
- [5472] E. Takimoto and A. Maruoka. On the sample complexity of consistent learning with one-sided error. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 265–278. Springer-Verlag, 1993.
- [5473] E. Takimoto and A. Maruoka. On the boosting algorithm for multiclass functions based on information-theoretic criterion for approximation. In *Proc. 1st International Conference on Discovery Science*, volume 1532 of *Lecture Notes in Artificial Intelligence*, pages 256–267. Springer, 1998.
- [5474] E. Takimoto and A. Maruoka. Top-down decision tree boosting and its applications. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 327–337. Springer, 2001.
- [5475] E. Takimoto and A. Maruoka. Top-down decision tree learning as information based boosting. *Theoret. Comput. Sci.*, 292(2):447–464, 2003.
- [5476] E. Takimoto, A. Maruoka, and V. Vovk. Predicting nearly as well as the best pruning of a decision tree through dynamic programming scheme. *Theoret. Comput. Sci.*, 261(1):179–209, 2001. Special issue for ALT '97.
- [5477] E. Takimoto, A. Miyashiro, A. Maruoka, and Y. Sakai. Learning orthogonal  $F$ -Horn formulas. *Theoret. Comput. Sci.*, 185(1):177–190, 1997. Special issue for ALT '95.
- [5478] E. Takimoto, Y. Sakai, and A. Maruoka. The learnability of exclusive-or expansions based on monotone DNF formulas. *Theoret. Comput. Sci.*, 241(1-2):37–50, 2000. Special issue for ALT '96.
- [5479] E. Takimoto, I. Tajika, and A. Maruoka. Mutual information gaining algorithm and its relation to PAC-learning algorithm. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 547–559. Springer-Verlag, 1994.
- [5480] E. Takimoto and M. Warmuth. The minimax strategy for Gaussian density estimation. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 100–106. Morgan Kaufmann, San Francisco, 2000.

- [5481] E. Takimoto and M. K. Warmuth. Predicting nearly as well as the best pruning of a planar decision graph. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 335–346. Springer, 1999.
- [5482] E. Takimoto and M. K. Warmuth. The last-step minimax algorithm. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 279–290. Springer, 2000.
- [5483] E. Takimoto and M. K. Warmuth. Path kernels and multiplicative updates. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 74–89. Springer, 2002.
- [5484] E. Takimoto and M. K. Warmuth. Predicting nearly as well as the best pruning graph. *Theoret. Comput. Sci.*, 288(2):217–235, 2002. Special issue for ALT '99.
- [5485] E. Takimoto and M. K. Warmuth. Path kernels and multiplicative updates. *Journal of Machine Learning Research*, 4:773–818, 2003.
- [5486] L. Talavera. Feature selection as a preprocessing step for hierarchical clustering. In *Proc. 16th International Conf. on Machine Learning*, pages 389–397. Morgan Kaufmann, San Francisco, CA, 1999.
- [5487] L. Talavera. Dynamic feature selection in incremental hierarchical clustering. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 392–403. Springer, 2000.
- [5488] L. Talavera. Feature selection and incremental learning of probabilistic concept hierarchies. In *Proc. 17th International Conf. on Machine Learning*, pages 951–958. Morgan Kaufmann, San Francisco, CA, 2000.
- [5489] D. A. Talbert and D. H. Fisher. OPT-KD: an algorithm for optimizing kd-trees. In *Proc. 16th International Conf. on Machine Learning*, pages 398–405. Morgan Kaufmann, San Francisco, CA, 1999.
- [5490] M. S. Talebi and O.-A. Maillard. Variance-aware regret bounds for undiscounted reinforcement learning in MDPs. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 770–805. PMLR, 07–09 Apr. 2018.
- [5491] J. L. Talmon, H. Fonteijn, and P. J. Braspenning. An analysis of the WITT algorithm. *Machine Learning*, 11:91–104, 1993.
- [5492] Y. Tamada, H. Bannai, O. Maruyama, and S. Miyano. Foundations of designing computational knowledge discovery processes. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 459–470. Springer, 2001.
- [5493] M. Tambe, A. Newell, and P. S. Rosenbloom. The problem of expensive chunks and its solution by restricting expressiveness. *Machine Learning*, 5:299–348, 1990.



- [5494] C. Tamon and J. Xiang. On the boosting pruning problem. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 404–412. Springer, 2000.
- [5495] M. Tan. Cost-sensitive learning of classification knowledge and its applications in robotics. *Machine Learning*, 13:7–33, 1993.
- [5496] T. Tanaka. Second difference method reinforced by grouping: A new tool for assistance in assignment of complex molecular spectra. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 532–542. Springer, 2001.
- [5497] T. Tanaka. Statistical learning in digital wireless communications. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 464–478. Springer, 2004.
- [5498] Y. Tanaka. Meme media for re-editing and redistributing intellectual assets and their application to interactive virtual information materialization. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 672–681. Springer, 2001.
- [5499] E. K. Tang, P. N. Suganthan, and X. Yao. An analysis of diversity measures. *Machine Learning*, 65(1):247–271, 2006.
- [5500] L. R. Tang and R. J. Mooney. Using multiple clause constructors in inductive logic programming for semantic parsing. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 466–477. Springer, 2001.
- [5501] S. Tangkitvanich and M. Shimura. Recovery from multiple faults in relational theory. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 277–287. IOS Press, 1992.
- [5502] S. Tani. On the role of equivalence queries in MAT learning. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 115–125. IOS Press, 1992.
- [5503] N. Tanida. Polynomial-time inference of paralleled even monogenic pure context-free languages. *IEICE Transactions on Information and Systems*, E81-D(3):261–270, 1998.
- [5504] N. Tanida and T. Yokomori. Polynomial-time identification of very regular languages in the limit. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 61–72. IOS Press, 1992.
- [5505] N. Tanida and T. Yokomori. Inductive inference of monogenic pure context-free languages. *IEICE Transactions on Information and Systems*, E79-D(11):1503–1510, 1996.
- [5506] F. Tantini, C. de la Higuera, and J.-C. Janodet. Identification in the limit of systematic-noisy languages. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20–22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 19–31, Berlin, 2006. Springer.
- [5507] F. Tantini, A. Terlutte, and F. Torre. Sequences classification by least general generalisations. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13–16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 189–202. Springer, 2010.

- [5508] N. Tao, J. Baxter, and L. Weaver. A multi-agent, policy-gradient approach to network routing. In *Proc. 18th International Conf. on Machine Learning*, pages 553–560. Morgan Kaufmann, San Francisco, CA, 2001.
- [5509] E. Tateishi, O. Maruyama, and S. Miyano. Extracting best consensus motifs from positive and negative examples. In *Proc. 13th Annual Symposium on Theoretical Aspects of Computer Science: STACS '96*, volume 1046 of *Lecture Notes in Computer Science*, pages 219–230. Springer, 1996.
- [5510] D. M. J. Tax and R. P. W. Duin. Uniform object generation for optimizing one-class classifiers. *Journal of Machine Learning Research*, 2:155–173, 2001.
- [5511] I. Tchoumatchenko and J.-G. Ganascia. A Bayesian framework to integrate symbolic and neural learning. In *Proc. 11th International Conference on Machine Learning*, pages 302–308. Morgan Kaufmann, 1994.
- [5512] G. Tecuci. Plausible justification trees: A framework for deep and dynamic integration of learning strategies. *Machine Learning*, 11:237–261, 1993.
- [5513] G. Tecuci and H. Keeling. Teaching an agent to test students. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML '98)*, pages 565–573. Morgan Kaufmann, San Francisco, CA, 1998.
- [5514] Y. W. Teh, M. Welling, S. Osindero, and G. E. Hinton. Energy-based models for sparse overcomplete representations. *Journal of Machine Learning Research*, 4:1235–1260, 2003.
- [5515] M. Telgarsky. Boosting with the logistic loss is consistent. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 911–965. JMLR.org, 2013.
- [5516] M. Telgarsky. benefits of depth in neural networks. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1517–1539, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [5517] M. Telgarsky, M. Dudík, and R. Schapire. Convex risk minimization and conditional probability estimation. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1629–1682, Paris, France, 03–06 July 2015. PMLR.
- [5518] A. Teller and M. Veloso. Efficient learning through evolution: Neural programming and internal reinforcement. In *Proc. 17th International Conf. on Machine Learning*, pages 959–966. Morgan Kaufmann, San Francisco, CA, 2000.
- [5519] I. Tellier. Meaning helps learning syntax. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 25–36. Springer, 1998.
- [5520] I. Tellier. How to split recursive automata. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 200–212. Springer, 2008.
- [5521] C. M. Teng. Correcting noisy data. In *Proc. 16th International Conf. on Machine Learning*, pages 239–248. Morgan Kaufmann, San Francisco, CA, 1999.

- [5522] L.-N. Teow and K.-F. Loe. Selection of support vector kernel parameters for improved generalization. In *Proc. 17th International Conf. on Machine Learning*, pages 967–974. Morgan Kaufmann, San Francisco, CA, 2000.
- [5523] S. A. Terwijn. Extensional set learning. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 243–248, New York, NY, 1999. ACM Press.
- [5524] S. A. Terwijn. On the learnability of hidden markov models. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 261–268. Springer, 2002.
- [5525] G. Tesauro. Scaling relationships in back-propagation learning: dependence on training set size. *Complex Systems*, 1:367–372, 1987.
- [5526] G. Tesauro. Neurogammon: A neural-network backgammon program. Technical Report RC 15619 (69436), IBM T.J. Watson Research Center, Mar. 1990.
- [5527] G. Tesauro. Practical issues in temporal difference learning. *Machine Learning*, 8:257–277, 1992.
- [5528] G. Tesauro. Comments on “Co-Evolution in the Successful Learning of Backgammon Strategy”. *Machine Learning*, 32(3):241–243, 1998.
- [5529] G. Tesauro and D. Cohn. Can neural networks do better than the Vapnik-Chervonenkis bounds? In *Advances in Neural Information Processing, Vol. 3*, pages 911–917. Morgan Kaufmann, 1991.
- [5530] G. Tesauro and B. Janssens. Scaling relationships in back-propagation learning. *Complex Systems*, 2:39–44, 1988.
- [5531] G. Tesauro and T. J. Sejnowski. A ‘neural’ network that learns to play backgammon. *Artificial Intelligence*, 39(3):357–390, July 1989.
- [5532] G. Tesauro and T. J. Sejnowski. A parallel network that learns to play backgammon. *Artificial Intelligence*, in press, 1989.
- [5533] A. Tewari and P. L. Bartlett. On the consistency of multiclass classification methods. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 143–157. Springer, 2005.
- [5534] A. Tewari and P. L. Bartlett. Bounded parameter markov decision processes with average reward criterion. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 263–277, Berlin, 2007. Springer.
- [5535] A. Tewari and P. L. Bartlett. On the consistency of multiclass classification methods. *Journal of Machine Learning Research*, 8:1007–1025, 2007.
- [5536] C. K. Tham and R. W. Prager. A modular Q-learning architecture for manipulator task decomposition. In *Proc. 11th International Conference on Machine Learning*, pages 309–317. Morgan Kaufmann, 1994.

- [5537] H. Theron and I. Cloete. BEXA: A covering algorithm for learning propositional concept descriptions. *Machine Learning*, 24:5–40, 1996.
- [5538] H. Thiele. Lernverfahren zur Erkennung formaler Sprachen. *Kybernetik-Forschung*, 3:11–93, 1973.
- [5539] H. Thiele. Zur Charakterisierung von Erkennungssystemen mit einbettendem Konvergenzbegriff. *Kompliziertheit von Lern- und Erkennungsprozessen*, 2:188–207, 1975. Jena.
- [5540] H. Thiele. A model theoretic oriented approach to analogy. In *Analogical and Inductive Inference, International Workshop AII '86, Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 196–208. Springer-Verlag, 1986.
- [5541] N. Thiemann, C. Igel, O. Wintenberger, and Y. Seldin. A strongly quasiconvex PAC-Bayesian bound. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 466–492, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [5542] B. Thiesson, C. Meek, and D. Heckerman. Accelerating EM for large databases. *Machine Learning*, 45(3):279–299, 2001.
- [5543] F. Thollard. Improving probabilistic grammatical inference core algorithms with post-processing techniques. In *Proc. 18th International Conf. on Machine Learning*, pages 561–568. Morgan Kaufmann, San Francisco, CA, 2001.
- [5544] F. Thollard and A. Clark. Shallow parsing using probabilistic grammatical inference. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 269–282. Springer, 2002.
- [5545] F. Thollard and A. Clark. Learning stochastic deterministic regular languages. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 248–259. Springer, 2004.
- [5546] F. Thollard, P. Dupont, and C. de la Higuera. Probabilistic DFA inference using Kullback-Leibler divergence and minimality. In *Proc. 17th International Conf. on Machine Learning*, pages 975–982. Morgan Kaufmann, San Francisco, CA, 2000.
- [5547] D. G. Thomas, M. H. Begam, K. G. Subramanian, and S. Gnanasekaran. Learning of regular bi- $\omega$  languages. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 283–292. Springer, 2002.
- [5548] M. Thomas and K. P. Jantke. Inductive inference for solving divergence in Knuth-Bendix completion. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 288–303. Springer-Verlag, 1989.
- [5549] C. A. Thompson, M. E. Califf, and R. J. Mooney. Active learning for natural language parsing and information extraction. In *Proc. 16th International Conf. on Machine Learning*, pages 406–414. Morgan Kaufmann, San Francisco, CA, 1999.

- [5550] R. Thonangi and V. Pudi. ACME: An associative classifier based on maximum entropy principle. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 122–134. Springer, Oct. 2005.
- [5551] C. J. Thornton. Analogical inference as generalised inductive inference. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 254–263. Springer-Verlag, 1989.
- [5552] C. Thrampoulidis, S. Oymak, and B. Hassibi. Regularized linear regression: A precise analysis of the estimation error. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1683–1709, Paris, France, 03–06 July 2015. PMLR.
- [5553] S. Thrun. Bayesian landmark learning for mobile robot localization. *Machine Learning*, 33(1):41–76, 1998.
- [5554] S. Thrun, J. C. Langford, and D. Fox. Monte Carlo hidden Markov models: Learning non-parametric models of partially observable stochastic processes. In *Proc. 16th International Conf. on Machine Learning*, pages 415–424. Morgan Kaufmann, San Francisco, CA, 1999.
- [5555] S. Thrun and J. O’Sullivan. Discovering structure in multiple learning tasks: the TC algorithm. In *Proc. 13th International Conference on Machine Learning*, pages 489–497. Morgan Kaufmann, 1996.
- [5556] Y. Tikochinsky, N. Z. Tishby, and R. D. Levine. Consistent inference of probabilities for reproducible experiments. *Phys. Rev. Lett.*, 52(16):1357–1360, 1984.
- [5557] C. Tillmann and H. Ney. Selection criteria for word trigger pairs in language modelling. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 95–106. Springer, 1996.
- [5558] K. M. Ting. The characterisation of predictive accuracy and decision combination. In *Proc. 13th International Conference on Machine Learning*, pages 498–506. Morgan Kaufmann, 1996.
- [5559] K. M. Ting. A comparative study of cost-sensitive boosting algorithms. In *Proc. 17th International Conf. on Machine Learning*, pages 983–990. Morgan Kaufmann, San Francisco, CA, 2000.
- [5560] K. M. Ting. An empirical study of metaCost using boosting algorithms. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 413–425. Springer, 2000.
- [5561] K. M. Ting and I. H. Witten. Stacking bagged and dagged models. In *Proc. 14th International Conference on Machine Learning*, pages 367–375. Morgan Kaufmann, 1997.
- [5562] P. Tino and G. Dorffner. Predicting the future of discrete sequences from fractal representations of the past. *Machine Learning*, 45(2):187–217, 2001.
- [5563] M. E. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.

- [5564] C. Tîrnăucă. A note on the relationship between different types of correction queries. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 213–223. Springer, 2008.
- [5565] C. Tîrnăucă and T. Knuutila. Polynomial time algorithms for learning  $k$ -reversible languages and pattern languages with correction queries. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 272–284, Berlin, Oct. 2007. Springer.
- [5566] H. Tirri. Learning with instance-based encodings. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 13, pages 205–213. MIT Press, 1994.
- [5567] H. Tirri, P. Kontkanen, and P. Myllymäki. Probabilistic instance-based learning. In *Proc. 13th International Conference on Machine Learning*, pages 507–515. Morgan Kaufmann, 1996.
- [5568] N. Tishby. Efficient data representations that preserve information. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, page 16. Springer, 2003.
- [5569] N. Tishby, E. Levin, and S. Solla. Consistent inference of probabilities in layered networks: predictions and generalizations. In *IJCNN International Joint Conference on Neural Networks*, volume II, pages 403–409. IEEE, 1989.
- [5570] E. F. Tjong Kim Sang. Memory-based shallow parsing. *Journal of Machine Learning Research*, 2:559–594, 2002.
- [5571] T. E. Tobely, Y. Yoshiki, R. Tsuda, N. Tsuruta, and M. Amamiya. Dynamic hand gesture recognition based on randomized self-organizing map algorithm. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 252–263. Springer, 2000.
- [5572] L. Todorovski and S. Džeroski. Declarative bias in equation discovery. In *Proc. 14th International Conference on Machine Learning*, pages 376–384. Morgan Kaufmann, 1997.
- [5573] L. Todorovski and S. Džeroski. Using domain knowledge on population dynamics modeling for equation discovery. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 478–490. Springer, 2001.
- [5574] L. Todorovski and S. Džeroski. Combining classifiers with meta decision trees. *Machine Learning*, 50(3):223–249, 2003.
- [5575] L. Todorovski, S. Džeroski, A. Srinivasan, J. Whiteley, and D. Gavaghan. Discovering the structure of partial differential equations from example behavior. In *Proc. 17th International Conf. on Machine Learning*, pages 991–998. Morgan Kaufmann, San Francisco, CA, 2000.
- [5576] K.-A. Toh. Training a reciprocal-sigmoid classifier by feature scaling-space. *Machine Learning*, 65(1):273–308, 2006.

- [5577] J. Toivonen, A. Visa, T. Vesanen, B. Back, and H. Vanharanta. Validation of text clustering based on document contents. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 184–195. Springer, 2001.
- [5578] I. Tolstikhin, G. Blanchard, and M. Kloft. Localized complexities for transductive learning. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 857–884, Barcelona, Spain, 13–15 June 2014. PMLR.
- [5579] I. Tolstikhin, N. Zhivotovskiy, and G. Blanchard. Permutational rademacher complexity. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 209–223, Berlin, Heidelberg, New York, 2015. Springer.
- [5580] H. Tong and T. X. Brown. Reinforcement learning for call admission control and routing under quality of service constraints in multimedia networks. *Machine Learning*, 49(2-3):111–139, 2002.
- [5581] S. Tong and D. Koller. Support vector machine active learning with applications to text classification. In *Proc. 17th International Conf. on Machine Learning*, pages 999–1006. Morgan Kaufmann, San Francisco, CA, 2000.
- [5582] S. Tong and D. Koller. Support vector machine active learning with applications to text classification. *Journal of Machine Learning Research*, 2:45–66, 2001.
- [5583] L. Torgo. Controlled redundancy in incremental rule learning. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 185–195. Springer-Verlag, 1993.
- [5584] L. Torgo. Rule combination in inductive learning. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 384–389. Springer-Verlag, 1993.
- [5585] L. Torgo. Functional models for regression tree leaves. In *Proc. 14th International Conference on Machine Learning*, pages 385–393. Morgan Kaufmann, 1997.
- [5586] L. Torgo. Partial linear trees. In *Proc. 17th International Conf. on Machine Learning*, pages 1007–1014. Morgan Kaufmann, San Francisco, CA, 2000.
- [5587] L. Torgo and J. P. da Costa. Clustered partial linear regression. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 426–436. Springer, 2000.
- [5588] L. Torgo and J. Pinto da Costa. Clustered partial linear regression. *Machine Learning*, 50(3):303–319, 2003.
- [5589] K. Torkkola. Feature extraction by non-parametric mutual information maximization. *Journal of Machine Learning Research*, 3:1415–1438, 2003.
- [5590] K. Torkkola and W. M. Campbell. Mutual information in learning feature transformations. In *Proc. 17th International Conf. on Machine Learning*, pages 1015–1022. Morgan Kaufmann, San Francisco, CA, 2000.

- [5591] D. S. Touretzky. Introduction. *Machine Learning*, 7:105–107, 1991.
- [5592] G. Towell. Book review: *Neural Network Perception for Mobile Robot Guidance* by dean A. pomerleau. kluwer academic publishers, 1993. *Machine Learning*, 18(1):115–119, 1995.
- [5593] G. Towell. Supervised learning using labeled and unlabeled examples. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 7, pages 107–119. MIT Press, 1997.
- [5594] G. Towell and R. Lehrer. A knowledge-based model of geometry learning. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 4, pages 55–74. MIT Press, 1995.
- [5595] G. Towell, E. M. Voorhees, N. K. Gupta, and B. Johnson-Laird. Learning collection fusion strategies for information retrieval. In *Proc. 12th International Conference on Machine Learning*, pages 540–548. Morgan Kaufmann, 1995.
- [5596] G. G. Towell. Local expert autoassociators for anomaly detection. In *Proc. 17th International Conf. on Machine Learning*, pages 1023–1030. Morgan Kaufmann, San Francisco, CA, 2000.
- [5597] G. G. Towell, T. Petsche, and M. R. Miller. Learning priorities from noisy examples. In *Proc. 17th International Conf. on Machine Learning*, pages 1031–1038. Morgan Kaufmann, San Francisco, CA, 2000.
- [5598] G. G. Towell and J. W. Shavlik. Extracting refined rules from knowledge-based neural networks. *Machine Learning*, 13:71–101, 1993.
- [5599] G. G. Towell and J. W. Shavlik. Using knowledge-based neural networks to refine roughly-correct information. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 5, pages 65–80. MIT Press, 1994.
- [5600] T. B. Trafalis and A. M. Malysheff. An analytic center machine. *Machine Learning*, 46(1-3):203–223, 2002.
- [5601] B. A. Trakhtenbrot and Y. M. Barzdin. *Finite Automata, Behavior and Synthesis*. North Holland, Amsterdam, 1973.
- [5602] V. Tresp, J. Hollatz, and S. Ahmad. Representing probabilistic rules with networks of gaussian basis functions. *Machine Learning*, 27(2):173–200, 1997.
- [5603] V. Tresp, J. Moody, and W.-R. DeLong. Neural network modeling of physiological processes. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 21, pages 363–378. MIT Press, 1994.
- [5604] I. Tsamardinos, L. E. Brown, and C. F. Aliferis. The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, 65(1):31–78, 2006.
- [5605] I. W. Tsang, J. T. Kwok, and P.-M. Cheung. Core vector machines: Fast SVM training on very large data sets. *Journal of Machine Learning Research*, 6:363–392, 2005.
- [5606] I. Tsapara and G. Turán. Learning atomic formulas with prescribed properties. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 166–174, New York, NY, 1998. ACM Press.



- [5607] J. N. Tsitsiklis. A lemma on the multiarmed bandit problem. *IEEE Transactions on Automatic Control*, AC-31(6), June 1986.
- [5608] J. N. Tsitsiklis. Asynchronous stochastic approximation and Q-learning. *Machine Learning*, 16(3):185–202, 1994.
- [5609] J. N. Tsitsiklis. On the convergence of optimistic policy iteration. *Journal of Machine Learning Research*, 3:59–72, 2002.
- [5610] J. N. Tsitsiklis and B. V. Roy. On average versus discounted reward temporal-difference learning. *Machine Learning*, 49(2-3):179–191, 2002.
- [5611] J. N. Tsitsiklis and B. van Roy. Feature-based methods for large scale dynamic programming. *Machine Learning*, 22(1/2/3):59–94, 1996.
- [5612] M. Tsuchiya, S. Kurohashi, and S. Sato. Discovery of definition patterns by compressing dictionary sentences. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 284–295. Springer, 2001.
- [5613] K. Tsuda, S. Akaho, and K. Asai. The *em* algorithm for kernel matrix completion with auxiliary data. *Journal of Machine Learning Research*, 4:67–81, 2003.
- [5614] Y. Tsukada and K. Koyama. A knowledge transfer model: from zero-knowledge toward full knowledge. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 187–195. IOS Press, 1992.
- [5615] H. Tsukimoto, M. Kakimoto, C. Morita, and Y. Kikuchi. Rule discovery from fMRI brain images by logical regression analysis. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 232–245. Springer, 2001.
- [5616] S. Tsumoto. Discovery of positive and negative knowledge in medical databases using rough sets. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 543–552. Springer, 2001.
- [5617] S. Tsumoto and H. Tanaka. Induction of probabilistic rules based on rough set theory. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 410–423. Springer-Verlag, 1993.
- [5618] A. B. Tsybakov. Optimal rates of aggregation. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 303–313. Springer, 2003.
- [5619] H.-C. Tu and C. H. Smith. Training digraphs. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII '94, 5th International Workshop on Algorithmic Learning Theory, ALT '94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 176–186. Springer-Verlag, 1994.

- [5620] K. Tu and V. Honavar. Unsupervised learning of probabilistic context-free grammar using iterative biclustering. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 224–237. Springer, 2008.
- [5621] G. Turán. Lower bounds for PAC learning with queries. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 384–391, New York, NY, 1993. ACM Press.
- [5622] G. Turan. On learning and logic. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 1–3, Berlin, 2006. Springer.
- [5623] M. Turcotte, S. H. Muggleton, and M. J. E. Sternberg. The effect of relational background knowledge on learning of protein three-dimensional fold signatures. *Machine Learning*, 43(1/2):81–95, 2001.
- [5624] P. Turney. Technical note: Bias and the quantification of stability. *Machine Learning*, 20(1/2):23–33, 1995.
- [5625] P. D. Turney. Exploiting context when learning to classify. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 402–407. Springer-Verlag, 1993.
- [5626] P. D. Turney. Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 491–502. Springer, 2001.
- [5627] W. Tzeng. The equivalence and learning of probabilistic automata. In *Proc. 30th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 268–273. IEEE Computer Society Press, Los Alamitos, CA, 1989.
- [5628] W.-G. Tzeng. Learning probabilistic automata and Markov chains via queries. *Machine Learning*, 8:151–166, 1992.
- [5629] T. Uchida, Y. Itokawa, T. Shoudai, T. Miyahara, and Y. Nakamura. A new framework for discovering knowledge from two-dimensional structured data using layout formal graph system. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 141–155. Springer, 2000.
- [5630] T. Uchiya, A. Nakamura, and M. Kudo. Algorithms for adversarial bandit problems with multiple plays. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6–8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 375–389, Berlin, Heidelberg, New York, 2010. Springer.
- [5631] K. Ueda and S. Nagano. Implementation of heuristic problem solving process including analogical reasoning. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 171–182. Springer-Verlag, 1993.

- [5632] N. Ueda and R. Nakano. EM algorithm with split and merge operations for mixture models. *IEICE Transactions on Information and Systems*, E83-D(12):2047–2055, 2000.
- [5633] K. Uehara and M. Shimada. Extraction of primitive motion and discovery of association rules from human motion data. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 338–348. Springer, 2001.
- [5634] R. Uehara, K. Tsuchida, and I. Wegener. Optimal attribute-efficient learning of disjunction, parity, and threshold functions. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 171–184. Springer, 1997.
- [5635] J. Uemura and M. Sato. Compactness and learning of classes of unions of erasing regular pattern languages. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 293–307. Springer, 2002.
- [5636] J. Uemura and M. Sato. Learning of erasing primitive formal systems from positive examples. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 69–83. Springer, 2003.
- [5637] J. Uemura and M. Sato. Learning of erasing primitive formal systems from positive examples. *Theoret. Comput. Sci.*, 364(1):98–114, 2006. Special issue for ALT 2003.
- [5638] G. Ueno, N. Nakamura, T. Higuchi, T. Tsuchiya, S. Machida, and T. Araki. Application of multivariate Maxwellian mixture model to plasma velocity distribution. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 372–383. Springer, 2001.
- [5639] J. Uitto and R. Wattenhofer. On competitive recommendations. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 83–97. Berlin, Heidelberg, New York, 2013. Springer.
- [5640] J. Uitto and R. Wattenhofer. On competitive recommendations. *Theoret. Comput. Sci.*, 620:4–14, 2016. Special Issue: Algorithmic Learning Theory.
- [5641] O. Unold. Inducing regular languages using grammar-based classifier system. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22–24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 298–300. Springer, 2008.
- [5642] O. Unold. Learning fuzzy context-free grammar – A preliminary report. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13–16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 309–312. Springer, 2010.
- [5643] O. Unold. Fuzzy grammar-based prediction of amyloidogenic regions. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 210–219, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.

- [5644] A. Unwin. Statistification or mystification? The need for statistical thought in visual data mining. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, page 602. Springer, 2001.
- [5645] R. Urner, S. Wulff, and S. Ben-David. PLAL: Cluster-based active learning. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 376–397. JMLR.org, 2013.
- [5646] P. Utgoff. Shift of bias for inductive concept learning. In *Machine Learning Vol2*. Morgan Kaufman, 1987.
- [5647] P. E. Utgoff. *Machine Learning of Inductive Bias*. Kluwer Academic, 1986.
- [5648] P. E. Utgoff. Incremental induction of decision trees. *Machine Learning*, 4:161–186, 1989.
- [5649] P. E. Utgoff. An improved algorithm for incremental induction of decision trees. In *Proc. 11th International Conference on Machine Learning*, pages 318–325. Morgan Kaufmann, 1994.
- [5650] P. E. Utgoff, N. C. Berkman, and J. A. Clouse. Decision tree induction based on efficient tree restructuring. *Machine Learning*, 29(1):5–44, 1997.
- [5651] P. E. Utgoff and D. J. Stracuzzi. Approximation via value unification. In *Proc. 16th International Conf. on Machine Learning*, pages 425–432. Morgan Kaufmann, San Francisco, CA, 1999.
- [5652] S. Vadhan. On learning vs. refutation. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1835–1848, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5653] L. M. Vaina, V. Sundareswaran, and J. G. Harris. Fast perceptual learning of motion in humans and neural networks. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 21, pages 363–378. MIT Press, 1997.
- [5654] D. Vainsencher, S. Mannor, and A. M. Bruckstein. The sample complexity of dictionary learning. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 773–788, Budapest, Hungary, 09–11 June 2011. PMLR.
- [5655] D. Vainsencher, S. Mannor, and H. Xu. Ignoring is a bliss: Learning with large noise through reweighting-minimization. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1849–1881, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5656] S. Vaithyanathan and B. Dom. Model selection in unsupervised learning with applications to document clustering. In *Proc. 16th International Conf. on Machine Learning*, pages 433–443. Morgan Kaufmann, San Francisco, CA, 1999.
- [5657] S. Vaithyanathan and B. Dom. Hierarchical unsupervised learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1039–1046. Morgan Kaufmann, San Francisco, CA, 2000.

- [5658] N. Vaits and K. Crammer. Re-adapting the regularization of weights for non-stationary regression. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 114–128, Berlin, Heidelberg, New York, 2011. Springer.
- [5659] R. E. Valdés-Pérez. Algebraic reasoning about reactions: Discovery of conserved properties in particle physics. *Machine Learning*, 17(1):47–67, 1994.
- [5660] R. E. Valdés-Pérez and A. Pérez. A powerful heuristic for the discovery of complex patterned behavior. In *Proc. 11th International Conference on Machine Learning*, pages 326–334. Morgan Kaufmann, 1994.
- [5661] G. Valentini and T. G. Dietterich. Bias-variance analysis of support vector machines for the development of SVM-based ensemble methods. *Journal of Machine Learning Research*, 5:725–775, 2004.
- [5662] L. Valiant. Deductive learning. *Phil. Trans. Roy. Soc. Lond. A*, 312:441–446, 1984.
- [5663] L. Valiant. Functionality in neural nets (at AAAI). In *AAAI*, pages 629–634, 1988.
- [5664] L. Valiant. A view of computational learning theory. In C. W. Gear, editor, *Computation & Cognition: Proceedings of the First NEC Research Symposium*, pages 32–51. SIAM, 1991.
- [5665] L. Valiant. A neuroidal model for cognitive functions. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, page 2, Oxford, 1994. Oxford University Press.
- [5666] L. G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, Nov. 1984.
- [5667] L. G. Valiant. Learning disjunctions of conjunctions. In *Proc. of the 9th International Joint Conference on Artificial Intelligence, vol. 1*, pages 560–566, Los Angeles, California, 1985. International Joint Committee for Artificial Intelligence.
- [5668] L. G. Valiant. Functionality in neural networks. In *Proc. 1st Annu. Workshop on Comput. Learning Theory*, pages 28–39, San Mateo, CA, 1988. Morgan Kaufmann.
- [5669] L. G. Valiant. Cognitive computation (extended abstract). In *Proc. of the 36th Annual Symposium on Foundations of Computer Science*, pages 2–3. IEEE Computer Society Press, Los Alamitos, CA, 1995.
- [5670] L. G. Valiant. Rationality. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 3–14, New York, NY, 1995. ACM Press.
- [5671] L. G. Valiant. Managing complexity in neuroidal circuits. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 1–11. Springer, 1996.
- [5672] L. G. Valiant. Projection learning. *Machine Learning*, 37(2):115–130, 1999.
- [5673] L. G. Valiant. A neuroidal architecture for cognitive computation. *J. ACM*, 47:854–882, 2000.
- [5674] L. G. Valiant. Robust logics. *Artificial Intelligence*, 117:231–253, 2000.
- [5675] L. G. Valiant and M. K. Warmuth, editors. *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, San Mateo, CA, 1991. Morgan Kaufmann, Inc.

- [5676] A. Van Den Bosch. Careful abstraction from instance families in memory-based language learning. *J. of Experimental and Theoret. Artif. Intell.*, 11(3):339–368, 1999.
- [5677] A. Van der Mude and A. Walker. On the inference of stochastic regular grammars. *Inform. Control*, 38:310–329, 1978.
- [5678] E. Van Nimwegen and J. P. Crutchfield. Optimizing epochal evolutionary search: Population-size dependent theory. *Machine Learning*, 45(1):77–114, 2001.
- [5679] M. van Zaanen and J. van de Loo. Learning interpretations using sequence classification. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 220–223, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5680] R. Vandermeulen and C. Scott. Consistency of robust kernel density estimators. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 568–591. JMLR.org, 2013.
- [5681] K. VanLehn. Learning one subprocedure per lesson. *Artificial Intelligence*, 31(1):1–40, Jan. 1987.
- [5682] K. Vanlehn. Efficient specialization of relational concepts. *Machine Learning*, 4:99–106, 1989.
- [5683] K. Vanlehn and W. Ball. A version space approach to learning context-free grammars. *Machine Learning*, 2(1):39–74, 1987.
- [5684] V. Vapnik. Statistical theory of generalization (abstract). In *Proc. 13th International Conference on Machine Learning*, page 557. Morgan Kaufmann, 1996.
- [5685] V. N. Vapnik. *Theorie der Zeichenerkennung*. Akademie-Verlag, 1979.
- [5686] V. N. Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer-Verlag, New York, 1982.
- [5687] V. N. Vapnik. Inductive principles of the search for empirical dependences (methods based on weak convergence of probability measures). In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 3–21, San Mateo, CA, Aug. 1989. Morgan Kaufmann.
- [5688] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, Berlin, 2nd edition, 2000.
- [5689] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probab. and its Applications*, 16(2):264–280, 1971.
- [5690] N. Vayatis and R. Azencott. Distribution-dependent Vapnik-Chervonenkis bounds. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29–31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 230–240. Springer, 1999.
- [5691] O.-A. M. N. Vayatis. Complexity versus agreement for many views. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 232–246, Berlin/Heidelberg, 2009. Springer.

- [5692] L. P. J. Veelenturf. Inference of sequential machines from sample computations. *IEEE Trans. on Computers*, C-27:167–170, 1978.
- [5693] S. Veerer. An evolutionary approach to evidence-based learning of deterministic finite automata. In *Proc. 17th International Conf. on Machine Learning*, pages 1071–1078. Morgan Kaufmann, San Francisco, CA, 2000.
- [5694] M. Velauthapillai. *On the Inductive Inference of Programs with Anomalies*. PhD thesis, University of Maryland, 1986.
- [5695] M. Velauthapillai. Inductive inference with bounded number of mind changes. In *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 200–213, San Mateo, CA, 1989. Morgan Kaufmann.
- [5696] M. M. Veloso and J. G. Carbonell. Derivational analogy in prodigy: Automating case acquisition, storage, and utilization. *Machine Learning*, 10:249–278, 1993.
- [5697] S. Vempala. A random sampling based algorithm for learning the intersection of half-spaces (extended abstract). In *Proc. 38th Annu. IEEE Symposium on Foundations of Computer Science*, pages 508–513. IEEE Computer Society Press, Los Alamitos, CA, 1997.
- [5698] S. Vempala and G. Wang. A spectral algorithm for learning mixtures of distributions. In *43rd Annual Symposium on Foundations of Computer Science, 6-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 113–122. IEEE Computer Society, 2002.
- [5699] S. Vempala and G. Wang. A spectral algorithm for learning mixture models. *J. Comput. Syst. Sci.*, 68(4):841–860, 2004.
- [5700] S. S. Vempala. Spectral algorithms for learning and clustering. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 3–4, Berlin, 2007. Springer.
- [5701] S. S. Vempala and Y. Xiao. Max vs min: Tensor decomposition and ICA with nearly linear sample complexity. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1710–1723, Paris, France, 03–06 July 2015. PMLR.
- [5702] A. Venkataraman. A procedure for unsupervised lexicon learning. In *Proc. 18th International Conf. on Machine Learning*, pages 569–576. Morgan Kaufmann, San Francisco, CA, 2001.
- [5703] S. S. Venkatesh. On learning binary weights for majority functions. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 257–266. Morgan Kaufmann, 1991.
- [5704] S. S. Venkatesh. Directed drift: A new linear threshold algorithm for learning binary weights on-line. *J. of Comput. Syst. Sci.*, 46(2):198–217, 1993.
- [5705] S. S. Venkatesh, R. R. Snapp, and D. Psaltis. BELLMAN STRIKES AGAIN! The growth rate of sample complexity with dimension for the nearest neighbor classifier. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 93–102, New York, NY, 1992. ACM Press.
- [5706] G. Venturini. SIA: A supervised inductive algorithm with genetic search for learning attributes based concepts. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 280–296. Springer-Verlag, 1993.

- [5707] K. Verbeurgt. Learning DNF under the uniform distribution in quasi-polynomial time. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 314–326, San Mateo, CA, 1990. Morgan Kaufmann.
- [5708] K. A. Verbeurgt. Learning sub-classes of monotone DNF on the uniform distribution. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 385–399. Springer, 1998.
- [5709] S. A. Vere. Multilevel counterfactuals for generalizations of relational concepts and productions. *Art. Int.*, 14:139–164, 1980.
- [5710] S. A. Vere. Constrained N-to-1 generalization. preprint, Jet Prop. Lab., 1981.
- [5711] N. Verma. Distance preserving embeddings for general  $n$ -dimensional manifolds. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 32.1–32.28, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5712] J.-P. Vert. Classification of biological sequences with kernel methods. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 7–18, Berlin, 2006. Springer.
- [5713] S. Verwer, M. de Weerdt, and C. Witteveen. A likelihood-ratio test for identifying probabilistic deterministic real-time automata from positive data. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 203–216. Springer, 2010.
- [5714] S. Verwer, R. Eyraud, and C. de la Higuera. Results of the PAutomaC probabilistic automaton learning competition. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 243–248, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [5715] S. Verwer, M. de Weerdt, and C. Witteveen. Polynomial distinguishability of timed automata. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 238–251. Springer, 2008.
- [5716] S. Verwer, M. de Weerdt, and C. Witteveen. One-clock deterministic timed automata are efficiently identifiable in the limit. In A. H. Dediu, A. M. Ionescu, and C. Martín-Vide, editors, *Language and Automata Theory and Applications, Third International Conference, LATA 2009, Tarragona, Spain, April 2009, Proceedings*, volume 5457 of *Lecture Notes in Computer Science*, pages 740–751. Springer, 2009.
- [5717] S. Verwer, M. de Weerdt, and C. Witteveen. The efficiency of identifying timed automata and the power of clocks. *Inform. Comput.*, 209(3):513–527, 2011. Special Issue: 3rd International Conference on Language and Automata Theory and Applications (LATA 2009).
- [5718] S. Verwer, M. van Zaanen, and R. Smetsers, editors. *International Conference on Grammatical Inference, 5-7 October 2016, Delft, The Netherlands*, volume 57 of *Proceedings of Machine Learning Research*. PMLR, 2016.



- [5719] S. Verwer, M. van Zaanen, and R. Smetsers. International conference on grammatical inference 2016: Preface. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 1–2, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [5720] R. Vicente, O. Kinouchi, and N. Caticha. Statistical mechanics of online learning of drifting concepts: A variational approach. *Machine Learning*, 32(2):179–201, 1998.
- [5721] E. Vidal and F. Casacuberta. Learning finite-state models for machine translation. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 3–15. Springer, 2004.
- [5722] E. Vidal and D. Llorens. Using knowledge to improve N-gram language modelling through the MGGI methodology. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 179–190. Springer, 1996.
- [5723] P. Viechnicki. A performance evaluation of automatic survey classifiers. In *Grammatical Inference, 4th International Colloquium, ICGI-98, Ames, Iowa, USA, July 1998, Proceedings*, volume 1433 of *Lecture Notes in Artificial Intelligence*, pages 244–256. Springer, 1998.
- [5724] S. Vijayakumar and S. Schaal. Locally weighted projection regression: An  $O(n)$  algorithm for incremental real time learning in high dimensional space. In *Proc. 17th International Conf. on Machine Learning*, pages 1079–1086. Morgan Kaufmann, San Francisco, CA, 2000.
- [5725] J. Viksna. Learning with growing quality. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 164–172. Springer-Verlag, 1993.
- [5726] J. Viksna. Probabilistic limit identification up to “small” sets. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 312–324. Springer, 1996.
- [5727] R. Vilalta, M. Brodie, D. Oblinger, and I. Rish. A unified framework for evaluation metrics in classification using decision trees. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 503–514. Springer, 2001.
- [5728] R. Vilalta and D. Oblinger. A quantification of distance-bias between evaluation metrics in classification. In *Proc. 17th International Conf. on Machine Learning*, pages 1087–1094. Morgan Kaufmann, San Francisco, CA, 2000.
- [5729] R. Vilalta and L. Rendell. Integrating feature construction with multiple classifiers in decision tree induction. In *Proc. 14th International Conference on Machine Learning*, pages 394–402. Morgan Kaufmann, 1997.
- [5730] J. M. Vilar. Query learning of subsequential transducer. In *Grammatical Inference: Learning Syntax from Sentences, 3rd International Colloquium, ICGI-96, Montpellier, France, September 25-27, 1996, Proceedings*, volume 1147 of *Lecture Notes in Artificial Intelligence*, pages 72–83. Springer, 1996.
- [5731] J. M. Vilar. Improve the learning of subsequential transducers by using alignments and dictionaries. In *Grammatical Inference: Algorithms and Applications, 5th International Colloquium, ICGI 2000, Lisbon, Portugal, September 11 - 13, 2000; Proceedings*, volume 1891 of *Lecture Notes in Artificial Intelligence*, pages 298–311. Springer, 2000.

- [5732] P. Vincent and Y. Bengio. Kernel matching pursuit. *Machine Learning*, 48(1/3):165–187, 2002.
- [5733] A. Visa. Technology of text mining. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 1–11. Springer, 2001.
- [5734] P. Vitányi. Genetic fitness optimization using rapidly mixing Markov chains. In *Algorithmic Learning Theory, 7th International Workshop, ALT '96, Sydney, Australia, October 1996, Proceedings*, volume 1160 of *Lecture Notes in Artificial Intelligence*, pages 67–82. Springer, 1996.
- [5735] P. Vitányi. A discipline of evolutionary programming. *Theoret. Comput. Sci.*, 241(1-2):3–23, 2000. Special issue for ALT '96.
- [5736] P. M. Vitányi and M. Li. Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Trans. on Information Theory*, 46(2):446–464, 2000.
- [5737] E. D. Vito, L. Rosasco, A. Caponnetto, U. D. Giovannini, and F. Odone. Learning from examples as an inverse problem. *Journal of Machine Learning Research*, 6:883–904, 2005.
- [5738] J. Vitter and J. Lin. Learning in parallel. *Inform. Comput.*, pages 179–202, 1992.
- [5739] J. Vivaldi, L. Màrquez, and H. Rodríguez. Improving term extraction by system combination using boosting. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 515–526. Springer, 2001.
- [5740] I. Volkovich. A guide to learning arithmetic circuits. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1540–1561, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [5741] E. M. Voorhees, C. Leacock, and G. Towell. Learning context to disambiguate word senses. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 15, pages 279–305. MIT Press, 1995.
- [5742] Y. Vorobeychik, M. P. Wellman, and S. Singh. Learning payoff functions in infinite games. *Machine Learning*, 67(1-2):145–168, 2007.
- [5743] V. Vovk. Universal forecasting algorithms. *Inform. Comput.*, 96(2):245–277, 1992.
- [5744] V. Vovk. Probability theory for the Brier game. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 323–338. Springer, 1997.
- [5745] V. Vovk. A game of prediction with expert advice. *J. of Comput. Syst. Sci.*, 56(2):153–173, 1998. Special issue: Eighth Annual Conference on Computational Learning Theory.
- [5746] V. Vovk. Probability theory for the Brier game. *Theoret. Comput. Sci.*, 261(1):57–79, 2001. Special issue for ALT '97.
- [5747] V. Vovk. Asymptotic optimality of transductive confidence machine. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 336–350. Springer, 2002.

- [5748] V. Vovk. On-line confidence machines are well-calibrated. In *43rd Annual Symposium on Foundations of Computer Science, 6-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 187–196. IEEE Computer Society, 2002.
- [5749] V. Vovk. Universal well-calibrated algorithm for on-line classification. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 358–372. Springer, 2003.
- [5750] V. Vovk. Well-calibrated predictions from online compression models. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 268–282. Springer, 2003.
- [5751] V. Vovk. A universal well-calibrated algorithm for on-line classification. *Journal of Machine Learning Research*, 5:575–604, 2004.
- [5752] V. Vovk. Defensive prediction with expert advice. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 444–458. Springer, Oct. 2005.
- [5753] V. Vovk. Non-asymptotic calibration and resolution. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 429–443. Springer, Oct. 2005.
- [5754] V. Vovk. Competing with wild prediction rules. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 559–573, Berlin, 2006. Springer.
- [5755] V. Vovk. Leading strategies in competitive on-line prediction. In *Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings*, volume 4264 of *Lecture Notes in Artificial Intelligence*, pages 214–228. Springer, Oct. 2006.
- [5756] V. Vovk. On-line regression competitive with reproducing kernel Hilbert spaces (extended abstract). In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 452–463, Berlin, 2006. Springer.
- [5757] V. Vovk. Predictions as statements and decisions. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 4–4, Berlin, 2006. Springer.
- [5758] V. Vovk. Well-calibrated predictions from on-line compression models. *Theoret. Comput. Sci.*, 364(1):10–26, 2006. Special issue for ALT 2003.
- [5759] V. Vovk. Competing with stationary prediction strategies. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 439–453, Berlin, 2007. Springer.
- [5760] V. Vovk. Competing with wild prediction rules. *Machine Learning*, 69(2-3):193–212, 2007.
- [5761] V. Vovk. Non-asymptotic calibration and resolution. *Theoret. Comput. Sci.*, 387(1):77–89, 2007. Special issue for ALT 2005.

- [5762] V. Vovk. Leading strategies in competitive on-line prediction. *Theoret. Comput. Sci.*, 405(3):285–296, 2008. Special issue for ALT 2006.
- [5763] V. Vovk, A. Gammerman, and C. Saunders. Machine-learning applications of algorithmic randomness. In *Proc. 16th International Conf. on Machine Learning*, pages 444–453. Morgan Kaufmann, San Francisco, CA, 1999.
- [5764] V. Vovk, I. Nourtdinov, A. Takemura, and G. Shafer. Defensive forecasting for linear protocols. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 459–473. Springer, Oct. 2005.
- [5765] V. Vovk and A. Shen. Prequential randomness. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 154–168, Berlin, Oct. 2008. Springer.
- [5766] V. Vovk and A. Shen. Prequential randomness and probability. *Theoret. Comput. Sci.*, 411(29-30):2632–2646, 2010. Special issue for ALT 2008.
- [5767] V. Vovk and C. Watkins. Universal portfolio selection. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 12–23, New York, NY, 1998. ACM Press.
- [5768] V. G. Vovk. Aggregating strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 371–383. Morgan Kaufmann, 1990.
- [5769] V. G. Vovk. An optimal-control application of two paradigms of on-line learning. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 98–109, New York, NY, 1994. ACM Press.
- [5770] V. G. Vovk. A game of prediction with expert advice. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 51–60, New York, NY, 1995. ACM Press.
- [5771] V. G. Vovk. Minimum description length estimators under the optimal coding scheme. In *Computational Learning Theory, Second European Conference, EuroCOLT '95, Barcelona, Spain, March 1995, Proceedings*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 237–251. Springer, 1995.
- [5772] V. G. Vovk. Learning an optimal decision strategy in an influence diagram with latent variables. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 110–121, New York, NY, 1996. ACM Press.
- [5773] V. G. Vovk. Learning about the parameter of the bernoulli model. *J. Comput. Syst. Sci.*, 55(1):96–104, 1997. Special issue for EuroCOLT '95.
- [5774] V. G. Vovk. Derandomizing stochastic prediction strategies. *Machine Learning*, 35(3):247–282, 1999.
- [5775] S. Vucetic and Z. Obradovic. Discovering homogeneous regions in spatial data through competition. In *Proc. 17th International Conf. on Machine Learning*, pages 1095–1102. Morgan Kaufmann, San Francisco, CA, 2000.
- [5776] S. Vucetic and Z. Obradovic. Classification on data with biased class distribution. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 527–538. Springer, 2001.

- [5777] M. V. Vyugin and V. V. V'yugin. Non-linear inequalities between predictive and Kolmogorov complexities. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 190–204. Springer, 2001.
- [5778] M. V. Vyugin and V. V. V'yugin. On complexity of easy predictable sequences. *Inform. Comput.*, 178(1):241–252, 2002.
- [5779] M. V. Vyugin and V. V. V'yugin. Predictive complexity and information. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 90–104. Springer, 2002.
- [5780] M. V. Vyugin and V. V. V'yugin. Predictive complexity and information. *J. Comput. Syst. Sci.*, 70(4):539–554, 2005. Special issue on COLT 2002.
- [5781] V. V. V'yugin. Most sequences are stochastic. *Inform. Comput.*, 169(2):252–263, Sept. 2001.
- [5782] V. V. V'yugin. On calibration error of randomized forecasting algorithms. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 388–402, Berlin, Oct. 2007. Springer.
- [5783] V. V. V'yugin. The follow perturbed leader algorithm protected from unbounded one-step losses. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 38–52, Berlin/Heidelberg, 2009. Springer.
- [5784] V. V. V'yugin. Learning volatility of discrete time series using prediction with expert advice. In O. Watanabe and T. Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo, Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, pages 16–30, Berlin/Heidelberg, 2009. Springer.
- [5785] V. V. V'yugin. On calibration error of randomized forecasting algorithms. *Theoret. Comput. Sci.*, 410(9):1781–1795, 2009. Special issue for ALT 2007.
- [5786] V. V. V'yugin. Universal algorithm for trading in stock market based on the method of calibration. In S. Jain, R. Munos, F. Stephan, and T. Zeugmann, editors, *Algorithmic Learning Theory, 24th International Conference, ALT 2013, Singapore, October 6–9, 2013, Proceedings*, volume 8139 of *Lecture Notes in Artificial Intelligence*, pages 53–67, Berlin, Heidelberg, New York, 2013. Springer.
- [5787] M. P. Wachowiak, R. Smolíková, M. G. Milanova, and A. S. Elmaghraby. Statistical and neural approaches for estimating parameters of a Speckle model based on the nakagami distribution. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 196–205. Springer, 2001.
- [5788] K. Wagstaff and C. Cardie. Clustering with instance-level constraints. In *Proc. 17th International Conf. on Machine Learning*, pages 1103–1110. Morgan Kaufmann, San Francisco, CA, 2000.

- [5789] K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl. Constrained K-means clustering with background knowledge. In *Proc. 18th International Conf. on Machine Learning*, pages 577–584. Morgan Kaufmann, San Francisco, CA, 2001.
- [5790] G. Wahba, C. Gu, Y. Wang, and R. Chappell. Soft classification, a.k.a. risk estimation, via penalized log likelihood and smoothing spline analysis of variance. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 8, pages 133–162. MIT Press, 1995.
- [5791] K. Waheed and F. M. Salem. Blind source recovery: A framework in the state space. *Journal of Machine Learning Research*, 4:1411–1446, 2003.
- [5792] A. Waibel. Introduction structured connectionist systems. *Machine Learning*, 15(2):121–123, 1994.
- [5793] A. Waibel, T. Hanazawa, G. Ginton, and K. Shikano. Phoneme recognition using time-delay neural networks. Technical report, ATR Interpreting Telephony Laboratories, 1987.
- [5794] K. Waiyama and L. Lakhal. Knowledge discovery from very large databases using frequent concept lattices. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 437–445. Springer, 2000.
- [5795] M. Wakatsuki, K. Teraguchi, and E. Tomita. Polynomial time identification of strict deterministic restricted one-counter automata in some class from positive data. In *Grammatical Inference: Algorithms and Applications, 7th International Colloquium, ICGI 2004, Athens, Greece, October 11-13, 2004, Proceedings*, volume 3264 of *Lecture Notes in Artificial Intelligence*, pages 260–272. Springer, 2004.
- [5796] M. Wakatsuki and E. Tomita. Polynomial time identification of strict prefix deterministic finite state transducers. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 313–316. Springer, 2010.
- [5797] M. Wakatsuki, E. Tomita, and G. Yamada. A unified algorithm for extending classes of languages identifiable in the limit from positive data. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 161–174. Berlin, 2006. Springer.
- [5798] M. Walker, J. Wright, and I. Langkilde. Using natural language processing and discourse features to identify understanding errors in a spoken dialogue system. In *Proc. 17th International Conf. on Machine Learning*, pages 1111–1118. Morgan Kaufmann, San Francisco, CA, 2000.
- [5799] N. Walkinshaw, K. Bogdanov, and K. Johnson. Evaluation and comparison of inferred regular grammars. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 252–265. Springer, 2008.
- [5800] C. Wallace, K. B. Korb, and H. Dai. Causal discovery via MML. In *Proc. 13th International Conference on Machine Learning*, pages 516–524. Morgan Kaufmann, 1996.

- [5801] C. S. Wallace and D. M. Boulton. An information measure for classification. *The Computer Journal*, 11(2):185–194, 1968.
- [5802] C. S. Wallace and J. D. Patrick. Coding decision trees. *Machine Learning*, 11:7–22, 1993.
- [5803] R. S. Wallace. *Finding Natural Clusters Through Entropy Minimization*. PhD thesis, Carnegie Mellon Computer Science Department, June 1989. Tech. Report CMU-CS-89-183.
- [5804] W. G. Walter. A machine that learns. *Scientific American*, pages 60–63, Aug. 1951.
- [5805] C. Wang and S. S. Venkatesh. Criteria for specifying machine complexity in learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 273–280, New York, NY, 1995. ACM Press.
- [5806] G. Wang and S. Mahadevan. Hierarchical optimization of policy-coupled semi-Markov decision processes. In *Proc. 16th International Conf. on Machine Learning*, pages 464–473. Morgan Kaufmann, San Francisco, CA, 1999.
- [5807] J. Wang, J. Lee, and C. Zhang. Kernel trick embedded gaussian mixture model. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 159–174. Springer, 2003.
- [5808] J. Wang, W. Wang, D. Garber, and N. Srebro. Efficient coordinate-wise leading eigenvector computation. In F. Janoos, M. Mohri, and K. Sridharan, editors, *Algorithmic Learning Theory, 7-9 April 2018*, volume 83 of *Proceedings of Machine Learning Research*, pages 806–820. PMLR, 07–09 Apr. 2018.
- [5809] J. Wang, W. Wang, and N. Srebro. Memory and communication efficient distributed stochastic optimization with minibatch prox. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1882–1919, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5810] J. Wang and J.-D. Zucker. Solving the multiple-instance problem: A lazy learning approach. In *Proc. 17th International Conf. on Machine Learning*, pages 1119–1125. Morgan Kaufmann, San Francisco, CA, 2000.
- [5811] L. Wang. On the margin explanation of boosting algorithms. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 479–490. Omnipress, 2008.
- [5812] L. Wang and J. Feng. Rademacher margin complexity. In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 620–621, Berlin, 2007. Springer.
- [5813] S. Wang and D. Schuurmans. Learning continuous latent variable models with bregman divergences. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 190–204. Springer, 2003.
- [5814] S. Wang, S. Wang, L. Cheng, R. Greiner, and D. Schuurmans. Stochastic analysis of lexical and semantic enhanced structural language model. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 97–111, Berlin, 2006. Springer.

- [5815] X. Wang. Learning by observation and practice: an incremental approach for planning operator acquisition. In *Proc. 12th International Conference on Machine Learning*, pages 549–557. Morgan Kaufmann, 1995.
- [5816] Y. Wang, Z.-C. Guo, and J. Ramon. Learning from networked examples. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 641–666, Kyoto University, Kyoto, Japan, 15–17 Oct. 2017. PMLR.
- [5817] Y. Wang, R. Khardon, D. Pechyony, and R. Jones. Generalization bounds for online learning algorithms with pairwise loss functions. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 13.1–13.22, Edinburgh, Scotland, 25–27 June 2012. PMLR.
- [5818] Y. Wang, L. Wang, Y. Li, D. He, and T.-Y. Liu. A theoretical analysis of NDCG type ranking measures. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 25–54. JMLR.org, 2013.
- [5819] M. Warmuth, editor. *Proceedings of the Seventh Annual ACM Conference on Computational Learning Theory*, New York, N.Y., 1994. ACM Press.
- [5820] M. Warmuth. Sample compression, learnability, and the Vapnik-Chervonenkis dimension. In *Computational Learning Theory, Third European Conference, EuroCOLT '97, Jerusalem, Israel, March 1997, Proceedings*, volume 1208 of *Lecture Notes in Artificial Intelligence*, pages 1–2. Springer, 1997.
- [5821] M. K. Warmuth. Towards representation independence in PAC-learning. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence 397*, pages 78–103. Springer-Verlag, 1989.
- [5822] M. K. Warmuth. Compressing to VC dimension many points. In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 743–744. Springer, 2003.
- [5823] M. K. Warmuth. The optimal PAC algorithm. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 641–642. Springer, 2004.
- [5824] M. K. Warmuth. Can entropic regularization be replaced by squared euclidean distance plus additional linear constraints. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 653–654, Berlin, 2006. Springer.
- [5825] M. K. Warmuth. When is there a free matrix lunch? In *Learning Theory, 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007, Proceedings*, volume 4539 of *Lecture Notes in Artificial Intelligence*, pages 630–632, Berlin, 2007. Springer.
- [5826] M. K. Warmuth. The blessing and the curse of the multiplicative updates. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th*



- International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, page 31, Berlin, Heidelberg, New York, 2010. Springer.
- [5827] M. K. Warmuth, K. A. Glocer, and S. V. N. Vishwanathan. Entropy regularized LPBoost. In *Algorithmic Learning Theory, 19th International Conference, ALT 2008, Budapest, Hungary, October 2008, Proceedings*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 256–271, Berlin, Oct. 2008. Springer.
- [5828] M. K. Warmuth and W. M. Koolen. Open problem: Shifting experts on easy data. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1295–1298, Barcelona, Spain, 13–15 June 2014. PMLR.
- [5829] M. K. Warmuth, W. M. Koolen, and D. P. Helmbold. Combining initial segments of lists. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 219–233, Berlin, Heidelberg, New York, 2011. Springer.
- [5830] M. K. Warmuth, W. M. Koolen, and D. P. Helmbold. Combining initial segments of lists. *Theoret. Comput. Sci.*, 519:29–45, 2014. Special Issue: Algorithmic Learning Theory.
- [5831] M. K. Warmuth, W. Kotłowski, and S. Zhou. Kernelization of matrix updates, when and how? In N. H. Bshouty, G. Stoltz, N. Vayatis, and T. Zeugmann, editors, *Algorithmic Learning Theory, 23rd International Conference, ALT 2012, Lyon, France, October 29–31, 2012, Proceedings*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 350–364, Berlin, Heidelberg, New York, 2012. Springer.
- [5832] M. K. Warmuth, W. Kotłowski, and S. Zhou. Kernelization of matrix updates, when and how? *Theoret. Comput. Sci.*, 558:159–178, 2014. Special Issue: Algorithmic Learning Theory.
- [5833] M. K. Warmuth and D. Kuzmin. Online variance minimization. In *Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings*, volume 4005 of *Lecture Notes in Artificial Intelligence*, pages 514–528, Berlin, 2006. Springer.
- [5834] M. K. Warmuth and S. V. N. Vishwanathan. Leaving the span. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 366–381. Springer, 2005.
- [5835] T. Washio and H. Motoda. Toward the discovery of first principle based scientific law equations. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 553–564. Springer, 2001.
- [5836] T. Washio, H. Motoda, and Y. Niwa. Enhancing the plausibility of law equation discovery. In *Proc. 17th International Conf. on Machine Learning*, pages 1127–1134. Morgan Kaufmann, San Francisco, CA, 2000.
- [5837] T. Washio, H. Motoda, and Y. Niwa. Discovering admissible simultaneous equation models from observed data. In *Machine Learning: ECML 2001, 12th European Conference on*

- Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 539–551. Springer, 2001.
- [5838] K. Watanabe and S. Watanabe. Estimation of the data region using extreme-value distributions. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 206–220. Springer, 2004.
- [5839] K. Watanabe and S. Watanabe. Stochastic complexity for mixture of exponential families in variational Bayes. In *Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 107–121. Springer, Oct. 2005.
- [5840] K. Watanabe and S. Watanabe. Stochastic complexity for mixture of exponential families in generalized variational Bayes. *Theoret. Comput. Sci.*, 387(1):4–17, 2007. Special issue for ALT 2005.
- [5841] O. Watanabe. A formal study of learning via queries. In *Automata, Languages and Programming, 17th International Colloquium Warwick University, England, July 16-20, 1990 Proceedings*, volume 443 of *Lecture Notes in Computer Science*, pages 139–152. Springer-Verlag, 1990.
- [5842] O. Watanabe. From computational learning theory to discovery science. In *Proc. 26th International Colloquium on Automata Languages and Programming, ICALP '99*, volume 1644 of *Lecture Notes in Computer Science*, pages 134–148. Springer, 1999.
- [5843] O. Watanabe. Sequential sampling techniques for algorithmic learning theory. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 27–40. Springer, 2000.
- [5844] O. Watanabe. Algorithmic aspects of boosting. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 349–359. Springer, 2001.
- [5845] O. Watanabe. Sequential sampling techniques for algorithmic learning theory. *Theoret. Comput. Sci.*, 348(1):3–14, 2005. Special issue for ALT 2000.
- [5846] O. Watanabe and A. Sharma. Preface. *Theoret. Comput. Sci.*, 288(2):195–196, 2002. Special issue for ALT '99.
- [5847] O. Watanabe and T. Yokomori, editors. *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Dec. 1999.
- [5848] O. Watanabe and T. Zeugmann, editors. *Stochastic Algorithms: Foundations and Applications, 5th International Symposium, SAGA 2009, Sapporo, Japan, October 2009, Proceedings*, volume 5792 of *Lecture Notes in Computer Science*, Berlin/Heidelberg, Oct. 2009. Springer.
- [5849] S. Watanabe. Information-theoretical aspects of inductive and deductive inference. *IBM J. of Res. Develop.*, 4:208–231, 1960.
- [5850] S. Watanabe. Algebraic analysis for singular statistical estimation. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 39–50. Springer, 1999.

- [5851] D. A. Waterman, W. S. Faugh, P. Klahr, S. J. Rosenschein, and R. Wesson. Design issues for exemplary programming. Technical report, Rand Corporation, N-1484-RC, 1980.
- [5852] T. L. H. Watkin, A. Rau, and M. Biehl. The stastical mechanics of learning a rule. *Rev. Mod. Phys.*, 65:499–556, 1993.
- [5853] C. J. C. H. Watkins. Combining cross-validation and search. In *Progress in Machine Learning—Proceedings of EWSL 87: 2nd European Working Session on Learning*, pages 79–90, Bled, Yugoslavia, May 1987.
- [5854] C. J. C. H. Watkins. *Learning from Delayed Rewards*. PhD thesis, King’s College, Oxford, May 1989. (To be reprinted by MIT Press.).
- [5855] C. J. C. H. Watkins and P. Dayan. Technical note Q-learning. *Machine Learning*, 8:279–292, 1992.
- [5856] R. L. Watrous. A comparison between squared error and relative entropy metrics using several optimization algorithms. *Complex Systems*, 6:495–505, 1992.
- [5857] P. Watson. Inductive inference of recurrence-term languages from positive data. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 292–315. Springer, 1995.
- [5858] P. Watson. Inductive learning with corroboration. In *Algorithmic Learning Theory, 10th International Conference, ALT ’99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 145–156. Springer, 1999.
- [5859] G. I. Webb. MultiBoosting: A technique for combining Boosting and Wagging. *Machine Learning*, 40(2):159–196, 2000.
- [5860] G. I. Webb, J. Wells, and Z. Zheng. An experimental evaluation of integrating machine learning with knowledge. *Machine Learning*, 35(1):5–23, 1999.
- [5861] J. Weed, V. Perchet, and P. Rigollet. Online learning in repeated auctions. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1562–1583, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [5862] A. Weigend, B. Huberman, and D. Rumelhart. Predicting the future: A connectionist approach. *Internat. J. Neural Syst.*, 1:193–209, 1990.
- [5863] A. S. Weigend and D. E. Rumelhart. Weight elimination and effective network size. In S. J. Hanson, G. A. Drastal, and R. L. Rivest, editors, *Computational Learning Theory and Natural Learning Systems*, volume I: Constraints and Prospects, chapter 16, pages 457–476. MIT Press, 1994.
- [5864] M. J. Weinberger, N. Merhav, and M. Feder. Optimal sequential probability assignment for individual sequences. *IEEE Transactions on Information Theory*, 40(2):384–396, Mar. 1994.
- [5865] G. M. Weiss. Learning with rare cases and small disjuncts. In *Proc. 12th International Conference on Machine Learning*, pages 558–565. Morgan Kaufmann, 1995.
- [5866] G. M. Weiss and H. Hirsh. The problem with noise and small disjuncts. In *Machine Learning, Proceedings of the Fifteenth International Conference (ICML ’98)*, pages 574–578. Morgan Kaufmann, San Francisco, CA, 1998.

- [5867] S. Weiss and C. Kulikowski. *Computer Systems that Learn*. Morgan Kaufmann, San Mateo, CA, 1991.
- [5868] S. M. Weiss and N. Indurkha. Small sample decision tree pruning. In *Proc. 11th International Conference on Machine Learning*, pages 335–342. Morgan Kaufmann, 1994.
- [5869] S. M. Weiss and N. Indurkha. Lightweight rule induction. In *Proc. 17th International Conf. on Machine Learning*, pages 1135–1142. Morgan Kaufmann, San Francisco, CA, 2000.
- [5870] T. Weissmann and N. Merhav. On prediction of individual sequences relative to a set of experts in the presence of noise. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 19–28, New York, NY, 1999. ACM Press.
- [5871] M. P. Wellman and J. Hu. Conjectural equilibrium in multiagent learning. *Machine Learning*, 33(2/3):179–200, 1998.
- [5872] R. S. Wenocur and R. M. Dudley. Some special Vapnik-Chervonenkis classes. *Discrete Mathematics*, 33:313–318, 1981.
- [5873] M. Westerdijk and W. Wiegerinck. Classification with multiple latent variable models using maximum entropy discrimination. In *Proc. 17th International Conf. on Machine Learning*, pages 1143–1150. Morgan Kaufmann, San Francisco, CA, 2000.
- [5874] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping. Use of the zero-norm with linear models and kernel methods. *Journal of Machine Learning Research*, 3:1439–1461, 2003.
- [5875] D. Wettschereck and T. G. Dietterich. An experimental comparison of the nearest-neighbor and nearest-hyperrectangle algorithms. *Machine Learning*, 19(1):5–27, 1995.
- [5876] K. Wexler. On extensional learnability. *Cognition*, 11:89–95, 1982.
- [5877] K. Wexler. The subset principle is an intensional principle. In E. Reuland and W. Abraham, editors, *Knowledge and Language*, volume 1, pages 217–239. Kluwer Academic Publishers, 1993.
- [5878] K. Wexler and P. W. Culicover. *Formal Principles of Language Acquisition*. MIT Press, Cambridge, Mass., 1980.
- [5879] K. Wexler and H. Hamburger. On the sufficiency of surface data for the learning of transformational languages. In *Approaches to Natural Language*. Dordrecht, D. reidel, 1973.
- [5880] R. M. Wharton. Approximate language identification. *Inform. Control*, 26:236–255, 1974.
- [5881] R. M. Wharton. Grammar enumeration and inference. *Inform. Control*, 33:253–272, 1977.
- [5882] A. P. White and W. Z. Liu. Bias in information-based measures in decision tree induction. *Machine Learning*, 15(3):321–329, 1994.
- [5883] H. White. Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, 3:535–549, 1990.
- [5884] H. White. Learning in artificial neural networks: a statistical perspective. *Neural Computation*, 1(4):425–464, 1990.
- [5885] S. D. Whitehead and D. H. Ballard. Learning to perceive and act by trial and error. *Machine Learning*, 7:45–83, 1991.

- [5886] D. Whitley, S. Dominic, R. Das, and C. W. Anderson. Genetic reinforcement learning for neurocontrol problems. *Machine Learning*, 13:259–284, 1993.
- [5887] G. Widmer. Recognition and exploitation of contextual clues via incremental meta-learning. In *Proc. 13th International Conference on Machine Learning*, pages 525–533. Morgan Kaufmann, 1996.
- [5888] G. Widmer. Tracking context changes through meta-learning. *Machine Learning*, 27(3):259–286, 1997.
- [5889] G. Widmer. Discovering strong principles of expressive music performance with the PLCG rule learning strategy. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 552–563. Springer, 2001.
- [5890] G. Widmer. The musical expression project: A challenge for machine learning and knowledge discovery. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 603–614. Springer, 2001.
- [5891] G. Widmer. In search of the horowitz factor: Interim report on a musical discovery project. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 41–41. Springer, 2002.
- [5892] G. Widmer. Guest editorial: Machine learning in and for music. *Machine Learning*, 65(2-3):343–346, 2006.
- [5893] G. Widmer and M. Kubat. Effective learning in dynamic environments by explicit context tracking. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 227–243. Springer-Verlag, 1993.
- [5894] G. Widmer and M. Kubat. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, 23(1):69–101, 1996.
- [5895] B. Widrow and M. E. Hoff. Adaptive switching circuits. *1960 IRE WESCON Convention Record*, pages 96–104, 1960. Reprinted in *Neurocomputing* (MIT Press, 1988).
- [5896] B. Widrow and S. Stearns. *Adaptive Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1985.
- [5897] W. Wiecek. A local search algorithm for grammatical inference. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13–16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 217–229. Springer, 2010.
- [5898] W. Wiecek. Induction of non-deterministic finite automata on supercomputers. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 237–242, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.

- [5899] W. Wieczorek and O. Unold. Induction of directed acyclic word graph in a bioinformatics task. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 207–217, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [5900] J. Wiedermann. The computational limits to the cognitive power of the neuroidal tabula rasa. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 63–76. Springer, 1999.
- [5901] R. Wiehagen. Inductive inference of recursive functions. In *Lecture Notes in Computer Science 32, Mathematical Foundations of Computer Science 75*, pages 462–464. Springer-Verlag, 1975.
- [5902] R. Wiehagen. Limes-Erkennung rekursiver Funktionen durch spezielle Strategien. *Elektronische Informationsverarbeitung und Kybernetik*, 12(1/2):93–99, 1976.
- [5903] R. Wiehagen. Identification of formal languages. In *Mathematical Foundations of Computer Science 1977, Proceedings, 6th Symposium, Tatranská Lomnica, September 5-9, 1977*, volume 53 of *Lecture Notes in Computer Science*, pages 571–579. Springer-Verlag, 1977.
- [5904] R. Wiehagen. Characterization problems in the theory of inductive inference. In *Automata, Languages and Programming, Fifth Colloquium, Udine, Italy, July 17-21, 1978*, volume 62 of *Lecture Notes in Computer Science*, pages 494–508. Springer-Verlag, 1978.
- [5905] R. Wiehagen. *Zur Theorie der Algorithmischen Erkennung*. Dissertation B, Humboldt-Universität zu Berlin, 1978.
- [5906] R. Wiehagen. How fast is program synthesis from examples. In *Mathematical Methods of Specification and Synthesis of Software Systems '85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 231–239. Springer-Verlag, 1986.
- [5907] R. Wiehagen. On the complexity of effective program synthesis. In *Analogical and Inductive Inference, International Workshop AII '86. Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 209–219. Springer-Verlag, 1986.
- [5908] R. Wiehagen. On the complexity of program synthesis from examples. *Elektronische Informationsverarbeitung und Kybernetik*, 22:305–323, 1986.
- [5909] R. Wiehagen. A thesis in inductive inference. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany, December 1990, Proceedings*, volume 543 of *Lecture Notes in Artificial Intelligence*, pages 184–207. Springer-Verlag, 1991.
- [5910] R. Wiehagen. From inductive inference to algorithmic learning theory. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 13–24. Springer-Verlag, 1993.
- [5911] R. Wiehagen, R. Freivalds, and E. B. Kinber. On the power of probabilistic strategies in inductive inference. *Theoret. Comput. Sci.*, 28(1-2):111–133, 1984.
- [5912] R. Wiehagen and H. Jung. Rekursionstheoretische Charakterisierung von erkennbaren Klassen rekursiver Funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 13:385–397, 1977.

- [5913] R. Wiehagen and W. Liepe. Characteristic properties of recognizable classes of recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 12:421–438, 1976.
- [5914] R. Wiehagen and W. Liepe. Charakteristische Eigenschaften von erkennbaren Klassen rekursiver Funktionen. *Elektronische Informationsverarbeitung und Kybernetik*, 12(8/9):421–438, 1976.
- [5915] R. Wiehagen and C. H. Smith. Generalization versus classification. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 224–230, New York, NY, 1992. ACM Press.
- [5916] R. Wiehagen, C. H. Smith, and T. Zeugmann. Classification of predicates and languages. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 171–181, Oxford, 1994. Oxford University Press.
- [5917] R. Wiehagen, C. H. Smith, and T. Zeugmann. Classifying recursive predicates and languages. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 174–189. Springer, 1995.
- [5918] R. Wiehagen and T. Zeugmann. Too much information can be too much for efficient learning. In *Analogical and Inductive Inference, International Workshop AII '92, Dagstuhl Castle, Germany, October 1992, Proceedings*, volume 642 of *Lecture Notes in Artificial Intelligence*, pages 72–86. Springer-Verlag, 1992.
- [5919] R. Wiehagen and T. Zeugmann. Ignoring data may be the only way to learn efficiently. *J. of Experimental and Theoret. Artif. Intell.*, 6(1):131–144, 1994. Special issue for AII '92.
- [5920] R. Wiehagen and T. Zeugmann. Learning and consistency. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 1–24. Springer, 1995.
- [5921] R. Wiehagen and T. Zeugmann. Foreword. *Theoret. Comput. Sci.*, 268(2):175–177, 2001. Special issue for ALT '98.
- [5922] M. Wiering. Multi-agent reinforcement learning for traffic light control. In *Proc. 17th International Conf. on Machine Learning*, pages 1151–1158. Morgan Kaufmann, San Francisco, CA, 2000.
- [5923] M. Wiering and J. Schmidhuber. Solving POMDPs with Levin search and EIRA. In *Proc. 13th International Conference on Machine Learning*, pages 534–542. Morgan Kaufmann, 1996.
- [5924] M. Wiering and J. Schmidhuber. Fast online  $Q(\lambda)$ . *Machine Learning*, 33(1):105–115, 1998.
- [5925] M. A. Wiering. Reinforcement learning in dynamic environments using instantiated information. In *Proc. 18th International Conf. on Machine Learning*, pages 585–592. Morgan Kaufmann, San Francisco, CA, 2001.
- [5926] H. Wiklicky. The neural network loading problem is undecidable. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 183–192, Oxford, 1994. Oxford University Press.
- [5927] C. K. I. Williams. On a connection between kernel PCA and metric multidimensional scaling. *Machine Learning*, 46(1-3):11–19, 2002.

- [5928] C. K. I. Williams and M. Seeger. The effect of the input density distribution on kernel-based classifiers. In *Proc. 17th International Conf. on Machine Learning*, pages 1159–1166. Morgan Kaufmann, San Francisco, CA, 2000.
- [5929] C. K. I. Williams and F. Vivarelli. Upper and lower bounds on the learning curve for gaussian processes. *Machine Learning*, 40(1):77–102, 2000.
- [5930] R. J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8:229–256, 1992.
- [5931] R. C. Williamson. The geometry of losses. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 1078–1108, Barcelona, Spain, 13–15 June 2014. PMLR.
- [5932] R. C. Williamson, A. J. Smola, and B. Schölkopf. Entropy numbers, operators and support vector kernels. In *Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings*, volume 1572 of *Lecture Notes in Artificial Intelligence*, pages 285–299. Springer, 1999.
- [5933] R. C. Williamson, A. J. Smola, and B. Schölkopf. Entropy numbers of linear function classes. In *Proc. 13th Annu. Conference on Comput. Learning Theory*, pages 309–319. Morgan Kaufmann, San Francisco, 2000.
- [5934] D. R. Wilson and T. R. Martinez. Instance pruning techniques. In *Proc. 14th International Conference on Machine Learning*, pages 403–411. Morgan Kaufmann, 1997.
- [5935] D. R. Wilson and T. R. Martinez. Reduction techniques for instance-based learning algorithms. *Machine Learning*, 38(3):257–286, 2000.
- [5936] S. W. Wilson. Knowledge growth in an artificial animal. In *Proc. of an International Conference on Genetic Algorithms and their Applications*, pages 16–23, July 1985.
- [5937] S. W. Wilson. Classifier systems and the animat problem. *Machine Learning*, 2(3):199–228, 1987.
- [5938] D. Wingate and K. D. Seppi. Prioritization methods for accelerating MDP solvers. *Journal of Machine Learning Research*, 6:851–881, 2005.
- [5939] J. Winn and C. M. Bishop. Variational message passing. *Journal of Machine Learning Research*, 6:661–694, 2005.
- [5940] P. H. Winston. Learning structural descriptions from examples. In *The Psychology of Computer Vision*. McGraw-Hill, New York, 1975.
- [5941] P. H. Winston. Learning new principles from precedents and exercises: The details. Technical Report AIM632, MIT Artificial Intelligence Laboratory, May 1981.
- [5942] I. H. Witten. An adaptive optimal controller for discrete-time Markov environments. *Inform. Control*, 34:286–295, 1977.
- [5943] I. H. Witten. Learning structure from sequences, with applications in a digital library. In *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany November 2002, Proceedings*, volume 2533 of *Lecture Notes in Artificial Intelligence*, pages 42–56. Springer, 2002.



- [5944] B. S. Wittner and J. S. Denker. Strategies for teaching layered networks classification tasks. In D. Z. Anderson, editor, *Neural Information Processing Systems*, pages 850–859, New York, 1988. American Insitute of Physics.
- [5945] J. Wnek and R. S. Michalski. Hypothesis-driven constructive induction in AQ17-HCI: A method and experiments. *Machine Learning*, 14(2):139–168, 1994.
- [5946] L. Wolf and A. Shashua. Learning over sets using kernel principal angles. *Journal of Machine Learning Research*, 4:913–931, 2003.
- [5947] D. Wolpert and W. G. Macready. An efficient method to estimate bagging’s generalization error. *Machine Learning*, 35(1):41–55, 1999.
- [5948] D. H. Wolpert. Filter likelihoods and exhaustive learning. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 3, pages 29–50. MIT Press, 1994.
- [5949] D. H. Wolpert. Horizontal generalization. In *Proc. 12th International Conference on Machine Learning*, pages 566–574. Morgan Kaufmann, 1995.
- [5950] D. H. Wolpert. On the Bayesian ‘Occam factors’ argument for Occam’s razor. In T. Petsche, editor, *Computational Learning Theory and Natural Learning Systems*, volume III: Selecting Good Models, chapter 11, pages 203–224. MIT Press, 1995.
- [5951] Y. fai Wong. A comparative study of the Kohonen self-organizing map and the elastic net. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 23, pages 401–413. MIT Press, 1994.
- [5952] C. Wongngamnit and D. Angluin. Robot localization in a grid. *Information Processing Letters*, 77(5-6):261–267, 2001.
- [5953] P. C. Woodland and S. G. Smyth. An experimental comparison of connectionist and conventional classification systems on natural data. *Speech Communication*, 9:73–82, Oct. 1989.
- [5954] D. Woodruff and Q. Zhang. Subspace embeddings and  $\ell_p$ -regression using exponential random variables. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 546–567. JMLR.org, 2013.
- [5955] B. Woodworth, S. Gunasekar, M. I. Ohannessian, and N. Srebro. Learning non-discriminatory predictors. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1920–1953, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [5956] K. Wright. Identification of unions of languages drawn from an identifiable class. In *Proceedings of the Second Annual Workshop on Computational Learning Theory, Santa Cruz, CA*, pages 328–333. Morgan Kaufmann, 1989. (See also the correction by Motoki, Shinohara and Wright in the Proceedings of the Fourth Annual Workshop on Computational Learning Theory, page 375, 1991).
- [5957] K. Wright. Inductive identification of pattern languages with restricted substitutions. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 111–121, San Mateo, CA, 1990. Morgan Kaufmann.

- [5958] S. Wrobel. On the proper definition of minimality in specialization and theory revision. In *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5–7, 1993, Proceedings*, volume 667 of *Lecture Notes in Artificial Intelligence*, pages 65–82. Springer-Verlag, 1993.
- [5959] S. Wrobel. Concept formation during interactive theory revision. *Machine Learning*, 14(2):169–191, 1994.
- [5960] S. Wrobel. Scalability issues in inductive logic programming. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 11–30. Springer, 1998.
- [5961] S. Wrobel. Scalability, search, and sampling: From smart algorithms to active discovery. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, page 615. Springer, 2001.
- [5962] C. Wu, M. Berry, and S. Shivakumar. Neural networks for full-scale protein sequence classification: Sequence encoding with singular value decomposition. *Machine Learning*, 21:177–193, 1995.
- [5963] C. Wu, M. Berry, S. Shivakumar, and J. McLarty. Neural networks for full-scale protein sequence classification: sequence encoding with singular value decomposition. *Machine Learning*, 21(1/2):177–193, 1995.
- [5964] D. Wu, K. P. Bennett, N. Cristianini, and J. Shawe-Taylor. Large margin trees for induction and transduction. In *Proc. 16th International Conf. on Machine Learning*, pages 474–483. Morgan Kaufmann, San Francisco, CA, 1999.
- [5965] Y. Wu and T. S. Huang. Towards self-exploring discriminating features. In *Machine Learning and Data Mining in Pattern Recognition, Second International Workshop, MLDM 2001, Leipzig, Germany, July 2001, Proceedings*, volume 2123 of *Lecture Notes in Artificial Intelligence*, pages 263–277. Springer, 2001.
- [5966] J. L. Wyatt. Exploration control in reinforcement learning using optimistic model selection. In *Proc. 18th International Conf. on Machine Learning*, pages 593–600. Morgan Kaufmann, San Francisco, CA, 2001.
- [5967] D. Xi and D. Zhuang. Model selection of sequence prediction algorithms by compression. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 160–163, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [5968] Y. Xiang and J. Lee. Learning decomposable markov networks in pseudo-independent domains with local evaluation. *Machine Learning*, 65(1):199–227, 2006.
- [5969] Y. Xiang, S. K. M. Wong, and N. Cercone. A ‘microscopic’ study of minimum entropy search in learning decomposable Markov networks. *Machine Learning*, 26(1):65–92, 1997.
- [5970] E. P. Xing, M. I. Jordan, and R. M. Karp. Feature selection for high-dimensional genomic microarray data. In *Proc. 18th International Conf. on Machine Learning*, pages 601–608. Morgan Kaufmann, San Francisco, CA, 2001.

- [5971] J. Xu, L. Massoulié, and M. Lelarge. Edge label inference in generalized stochastic block models: from spectral theory to impossibility results. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 903–920, Barcelona, Spain, 13–15 June 2014. PMLR.
- [5972] H. Yamakawa, K. Maruhashi, and Y. Nakao. Predicting type of protein-protein interactions using a multiple-instance learning model. In *New Frontiers in Artificial Intelligence, JSAI 2006 Conference and Workshops, Tokyo, Japan, June 2006, Revised Selected Papers*, volume 4384 of *Lecture Notes in Artificial Intelligence*, pages 42–53, Berlin, 2007. Springer.
- [5973] A. Yamamoto. Generalized unification as background knowledge in learning logic programs. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 111–122. Springer-Verlag, 1993.
- [5974] A. Yamamoto. Logical aspects of several bottom-up fittings. In *Algorithmic Learning Theory, 9th International Conference, ALT '98, Otzenhausen, Germany, October 1998, Proceedings*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 158–168. Springer, 1998.
- [5975] A. Yamamoto. An inference method for the complete inverse of relative subsumption. *New Generation Computing*, 17(1):99–117, 1999.
- [5976] A. Yamamoto. Revising the logical foundations of inductive logic programming systems with ground reduced programs. *New Generation Computing*, 17(1):119–127, 1999.
- [5977] A. Yamamoto. Hypothesis finding based on upward refinement of residue hypotheses. *Theoret. Comput. Sci.*, 298(1):5–19, 2003.
- [5978] A. Yamamoto and B. Fronhöfer. Hypotheses finding via residue hypotheses with the resolution principle. In *Algorithmic Learning Theory, 11th International Conference, ALT 2000, Sydney, Australia, December 2000, Proceedings*, volume 1968 of *Lecture Notes in Artificial Intelligence*, pages 156–165. Springer, 2000.
- [5979] A. Yamamoto and B. Fronhöfer. A theory of hypothesis finding in clausal logic. In S. Arikawa and A. Shinohara, editors, *Progress in Discovery Science, Final Report of the Japanese Discovery Science Project*, volume 2281 of *Lecture Notes in Artificial Intelligence*, pages 246–257. Springer, 2001.
- [5980] K. Yamanishi. A loss bound model for on-line stochastic prediction strategies. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 290–302, San Mateo, CA, 1991. Morgan Kaufmann.
- [5981] K. Yamanishi. A learning criterion for stochastic rules. *Machine Learning*, 9:165–203, 1992.
- [5982] K. Yamanishi. Learning non-parametric densities by finite-dimensional parametric hypotheses. In *Proc. 2nd Int. Workshop on Algorithmic Learning Theory*, pages 175–186. IOS Press, 1992.
- [5983] K. Yamanishi. Probably almost discriminative learning. In *Proc. 5th Annual ACM Workshop on Comput. Learning Theory*, pages 164–171, New York, NY, 1992. ACM Press.
- [5984] K. Yamanishi. On polynomial-time probably almost discriminative learnability. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 94–100, New York, NY, 1993. ACM Press.

- [5985] K. Yamanishi. Generalized stochastic complexity and its applications to learning. In *Proc. 1994 Conf. Information Science and Systems*, volume 2, pages 763–768, 1994.
- [5986] K. Yamanishi. Learning non-parametric smooth rules by stochastic rules with finite partitioning. In *Computational Learning Theory: EuroColt '93*, volume New Series Number 53 of *The Institute of Mathematics and its Applications Conference Series*, pages 217–227, Oxford, 1994. Oxford University Press.
- [5987] K. Yamanishi. The minimum  $l$ -complexity algorithm and its applications to learning non-parametric rules. In *Proc. 7th Annu. ACM Conf. on Comput. Learning Theory*, pages 173–182, New York, NY, 1994. ACM Press.
- [5988] K. Yamanishi. A loss bound model for on-line stochastic prediction algorithms. *Inform. Comput.*, 119(1):39–54, 1995.
- [5989] K. Yamanishi. Probably almost discriminative learning. *Machine Learning*, 18(1):23–50, 1995.
- [5990] K. Yamanishi. Randomized approximate aggregating strategies and their applications to prediction and discrimination. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 83–90, New York, NY, 1995. ACM Press.
- [5991] K. Yamanishi. A randomized approximation of the MDL for stochastic models with hidden variables. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 99–109, New York, NY, 1996. ACM Press.
- [5992] K. Yamanishi. On-line maximum likelihood prediction with respect to general loss functions. *J. Comput. Syst. Sci.*, 55(1):105–118, 1997. Special issue for EuroCOLT '95.
- [5993] K. Yamanishi. A decision-theoretic extension of stochastic complexity and its applications to learning. *IEEE Transactions on Information Theory*, 44(4):1424–1439, 1998.
- [5994] K. Yamanishi. Minimax relative loss analysis for sequential prediction algorithms using parametric hypotheses. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*, pages 32–43, New York, NY, 1998. ACM Press.
- [5995] K. Yamanishi. Distributed cooperative Bayesian learning strategies. *Inform. Comput.*, 150(1):25–56, 1999. First appeared in COLT'97.
- [5996] K. Yamanishi. Extended stochastic complexity and minimax relative loss analysis. In *Algorithmic Learning Theory, 10th International Conference, ALT '99, Tokyo, Japan, December 1999, Proceedings*, volume 1720 of *Lecture Notes in Artificial Intelligence*, pages 26–38. Springer, 1999.
- [5997] K. Yamazaki and S. Watanabe. Newton diagram and stochastic complexity in mixture of binomial distributions. In *Algorithmic Learning Theory, 15th International Conference, ALT 2004, Padova, Italy, October 2004, Proceedings*, volume 3244 of *Lecture Notes in Artificial Intelligence*, pages 350–364. Springer, 2004.
- [5998] T. Yamazaki, M. J. Pazzani, and C. Merz. Learning hierarchies from ambiguous natural language data. In *Proc. 12th International Conference on Machine Learning*, pages 575–583. Morgan Kaufmann, 1995.
- [5999] K. Yang. On learning correlated Boolean functions using statistical queries (extended abstract). In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 59–76. Springer, 2001.

- [6000] K. Yang. New lower bounds for statistical query learning. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 229–243. Springer, 2002.
- [6001] K. Yang. New lower bounds for statistical query learning. *J. Comput. Syst. Sci.*, 70(4):485–509, 2005. Special issue on COLT 2002.
- [6002] L. Yang, S. Hanneke, and J. Carbonell. Bayesian active learning using arbitrary binary valued queries. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 50–58, Berlin, Heidelberg, New York, 2010. Springer.
- [6003] L. Yang, S. Hanneke, and J. Carbonell. Identifiability of priors from bounded sample sizes with applications to transfer learning. In S. M. Kakade and U. von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 789–806, Budapest, Hungary, 09–11 June 2011. PMLR.
- [6004] L. Yang, S. Hanneke, and J. Carbonell. Bounds on the minimax rate for estimating a prior over a VC class from independent learning tasks. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 270–284, Berlin, Heidelberg, New York, 2015. Springer.
- [6005] L. Yang, S. Hanneke, and J. Carbonell. Bounds on the minimax rate for estimating a prior over a VC class from independent learning tasks. *Theoret. Comput. Sci.*, 716:124–140, 2018. Special Issue on ALT 2015.
- [6006] Y. Yang, T. Ault, and T. Pierce. Combining multiple learning strategies for effective cross validation. In *Proc. 17th International Conf. on Machine Learning*, pages 1167–1174. Morgan Kaufmann, San Francisco, CA, 2000.
- [6007] Y. Yang and J. O. Pedersen. A comparative study on feature selection in text categorization. In *Proc. 14th International Conference on Machine Learning*, pages 412–420. Morgan Kaufmann, 1997.
- [6008] Y. Yang and G. I. Webb. Proportional  $k$ -interval discretization for naive-Bayes classifiers. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 564–575. Springer, 2001.
- [6009] M. Yannakakis and D. Lee. Testing finite state machines. In *Proc. of the 23rd Symposium on Theory of Computing*, pages 476–485, New York, NY, 1991. ACM Press.
- [6010] J. Ye. Characterization of a family of algorithms for generalized discriminant analysis on undersampled problems. *Journal of Machine Learning Research*, 6:483–502, 2005.
- [6011] O. T. Yildiz and E. Alpaydin. Linear discriminant trees. In *Proc. 17th International Conf. on Machine Learning*, pages 1175–1182. Morgan Kaufmann, San Francisco, CA, 2000.
- [6012] Y. L. Yin. Reducing the number of queries in self-directed learning. In *Proc. 8th Annu. Conf. on Comput. Learning Theory*, pages 128–135, New York, NY, 1995. ACM Press.
- [6013] Y. Ying and C. Campbell. Learning coordinate gradients with multi-task kernels. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 217–228. Omnipress, 2008.

- [6014] T. Yokimori. Learning non-deterministic finite automata from queries and counterexamples. In K. Furukawa, D. Michie, and S. Muggleton, editors, *Machine Intelligence*, volume 13, pages 169–189. Oxford Univ. Press, 1994.
- [6015] T. Yokomori. Learning context-free languages efficiently – A report on recent results in japan. In *Analogical and Inductive Inference, International Workshop AII '89, Reinhardtsbrunn Castle, GDR, October 1989, Proceedings*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 104–123. Springer-Verlag, 1989.
- [6016] T. Yokomori. Polynomial-time learning of very simple grammars from positive data. In *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, pages 213–227, San Mateo, CA, 1991. Morgan Kaufmann.
- [6017] T. Yokomori. Learning two-tape automata from queries and counterexamples. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 228–235, New York, NY, 1993. ACM Press.
- [6018] T. Yokomori. On learning systolic languages. In *Algorithmic Learning Theory, Third Workshop, ALT '92, Tokyo, Japan, October 1992, Proceedings*, volume 743 of *Lecture Notes in Artificial Intelligence*, pages 41–52. Springer-Verlag, 1993.
- [6019] T. Yokomori. On polynomial-time learnability in the limit of strictly deterministic automata. *Machine Learning*, 19(2):153–179, 1995.
- [6020] T. Yokomori. Learning two-tape automata from queries and counterexamples. *Mathematical Systems Theory*, 29(3):259–270, 1996.
- [6021] T. Yokomori. Corrigendum learning two-type automata from queries and counterexamples. *Theory of Computing Systems*, 35(4):465–466, 2002.
- [6022] T. Yokomori. Polynomial-time identification of very simple grammars from positive data. *Theoret. Comput. Sci.*, 298(1):179–206, 2003.
- [6023] K. Yoshida, H. Motoda, and N. Indurkha. Unifying learning methods by colored digraphs. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 342–355. Springer-Verlag, 1993.
- [6024] R. Yoshinaka. Polynomial-time identification of an extension of very simple grammars from positive data. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 45–58, Berlin, 2006. Springer.
- [6025] R. Yoshinaka. Learning efficiency of very simple grammars from positive data. In *Algorithmic Learning Theory, 18th International Conference, ALT 2007, Sendai, Japan, October 2007, Proceedings*, volume 4754 of *Lecture Notes in Artificial Intelligence*, pages 227–241, Berlin, Oct. 2007. Springer.
- [6026] R. Yoshinaka. Identification in the limit of  $k, l$ -substitutable context-free languages. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 266–279. Springer, 2008.
- [6027] R. Yoshinaka. Learning efficiency of very simple grammars from positive data. *Theoret. Comput. Sci.*, 410(9):1807–1825, 2009. Special issue for ALT 2007.

- [6028] R. Yoshinaka. Learning mildly context-sensitive languages with multidimensional substitutability from positive data. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Algorithmic Learning Theory, 20th International Conference, ALT 2009, Porto, Portugal, October 2009, Proceedings*, volume 5809 of *Lecture Notes in Artificial Intelligence*, pages 278–292, Berlin/Heidelberg, 2009. Springer.
- [6029] R. Yoshinaka. Polynomial-time identification of multiple context-free languages from positive data and membership queries. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 230–244. Springer, 2010.
- [6030] R. Yoshinaka. Efficient learning of multiple context-free languages with multidimensional substitutability from positive data. *Theoret. Comput. Sci.*, 412(19):1821–1831, 2011.
- [6031] M. Young-Lai and F. W. Tompa. Stochastic grammatical inference of text database structure. *Machine Learning*, 40(2):111–137, 2000.
- [6032] H. Yu. On convergence of emphatic temporal-difference learning. In P. Grünwald, E. Hazan, and S. Kale, editors, *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 1724–1751, Paris, France, 03–06 July 2015. PMLR.
- [6033] Y. Yue, J. Broder, R. Kleinberg, and T. Joachims. The  $K$ -armed dueling bandits problem. *J. Comput. Syst. Sci.*, 78(5):1538–1556, 2012.
- [6034] S.-Y. Yun and A. Proutiere. Community detection via random and adaptive sampling. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 138–175, Barcelona, Spain, 13–15 June 2014. PMLR.
- [6035] M. V. V. Yuri Kalnishkan, Volodya Vovk. How many strings are easy to predict? In *Learning Theory and Kernel Machines, 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003, Proceedings*, volume 2777 of *Lecture Notes in Artificial Intelligence*, pages 522–536. Springer, 2003.
- [6036] M. Zaanen and N. Noord. Model merging versus model splitting context-free grammar induction. In J. Heinz, C. de la Higuera, and T. Oates, editors, *Proceedings of the Eleventh International Conference on Grammatical Inference*, volume 21 of *Proceedings of Machine Learning Research*, pages 224–236, University of Maryland, College Park, MD, USA, 05–08 Sept. 2012. PMLR.
- [6037] M. Zaanen and N. Noord. Evaluation of selection in context-free grammar learning systems. In A. Clark, M. Kanazawa, and R. Yoshinaka, editors, *The 12th International Conference on Grammatical Inference*, volume 34 of *Proceedings of Machine Learning Research*, pages 193–206, Kyoto, Japan, 17–19 Sept. 2014. PMLR.
- [6038] M. van Zaanen. Bootstrapping syntax and recursion using alignment-based learning. In *Proc. 17th International Conf. on Machine Learning*, pages 1063–1070. Morgan Kaufmann, San Francisco, CA, 2000.
- [6039] M. van Zaanen. Implementing alignment-based learning. In *Grammatical Inference: Algorithms and Applications, 6th International Colloquium, ICGI 2002, Amsterdam, The Netherlands, September 23-25, 2002, Proceedings*, volume 2484 of *Lecture Notes in Artificial Intelligence*, pages 312–314. Springer, 2002.

- [6040] M. van Zaanen and T. Gaustad. Grammatical inference as class discrimination. In J. M. Sempere and P. García, editors, *Grammatical Inference: Theoretical Results and Applications, 10th International Colloquium, ICGI 2010, Valencia, Spain, September 13-16, 2010 Proceedings*, volume 6339 of *Lecture Notes in Artificial Intelligence*, pages 245–257. Springer, 2010.
- [6041] M. van Zaanen and J. Geertzen. Grammatical inference for syntax-based statistical machine translation. In *Grammatical Inference: Algorithms and Applications, 8th International Colloquium, ICGI 2006, Tokyo, Japan, September 20-22, 2006, Proceedings*, volume 4201 of *Lecture Notes in Artificial Intelligence*, pages 356–357, Berlin, 2006. Springer.
- [6042] M. van Zaanen and J. Geertzen. Problems with evaluation of unsupervised empirical grammatical inference systems. In A. Clark, F. Coste, and L. Miclet, editors, *Grammatical Inference: Algorithms and Applications, 9th International Colloquium, ICGI 2008, Saint-Malo, France, September 22-24, 2008, Proceedings*, volume 5278 of *Lecture Notes in Artificial Intelligence*, pages 301–303. Springer, 2008.
- [6043] A. M. Zador and B. A. Pearlmutter. VC dimension of an integrate-and-fire neuron model. In *Proc. 9th Annu. Conf. on Comput. Learning Theory*, pages 10–18, New York, NY, 1996. ACM Press.
- [6044] A. M. Zador and B. A. Pearlmutter. The discriminative power of a dynamical model neuron. In R. Greiner, T. Petsche, and S. J. Hanson, editors, *Computational Learning Theory and Natural Learning Systems*, volume IV: Making Learning Systems Practical, chapter 13, pages 211–225. MIT Press, 1997.
- [6045] B. Zadrozny and C. Elkan. Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In *Proc. 18th International Conf. on Machine Learning*, pages 609–616. Morgan Kaufmann, San Francisco, CA, 2001.
- [6046] A. Zakai and Y. Ritov. How local should a learning method be? In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 205–216. Omnipress, 2008.
- [6047] M. J. Zaki. SPADE: An efficient algorithm for mining frequent sequences. *Machine Learning*, 42(1/2):31–60, 2001.
- [6048] D. Zelenko, C. Aone, and A. Richardella. Kernel methods for relation extraction. *Journal of Machine Learning Research*, 3:1083–1106, 2003.
- [6049] S. Zelikovitz and H. Hirsh. Improving short-text classification using unlabeled background knowledge to assess document similarity. In *Proc. 17th International Conf. on Machine Learning*, pages 1183–1190. Morgan Kaufmann, San Francisco, CA, 2000.
- [6050] J. M. Zelle, R. J. Mooney, and J. B. Konvisser. Combining top-down and bottom-up techniques in inductive logic programming. In *Proc. 11th International Conference on Machine Learning*, pages 343–351. Morgan Kaufmann, 1994.
- [6051] X. Zeng and T. R. Martinez. Distribution-balanced stratified cross-validation for accuracy estimation. *J. of Experimental and Theoret. Artif. Intell.*, 12(1):1–12, 2000.
- [6052] G. Zenobi and P. Cunningham. Using diversity in preparing ensembles of classifiers based on different feature subsets to minimize generalization error. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 576–587. Springer, 2001.



- [6053] T. Zeugmann. A-posteriori characterizations in inductive inference of recursive functions. *Elektronische Informationsverarbeitung und Kybernetik*, 19(10/11):559–594, 1983.
- [6054] T. Zeugmann. On the finite identification of fastest programs. In H. Rasiowa and H. Thiele, editors, *Symposium on Mathematical Foundations of Computer Science, December 6-11, 1982, Dierichshagen*, volume 52, pages 151–159. Seminarberichte der Sektion Mathematik der Humboldt-Universität zu Berlin, Berlin, Germany, 1983.
- [6055] T. Zeugmann. On the synthesis of fastest programs in inductive inference. *Elektronische Informationsverarbeitung und Kybernetik*, 19(12):625–642, 1983.
- [6056] T. Zeugmann. *Zur algorithmischen Synthese von schnellen Programmen*. PhD thesis, Humboldt-Universität zu Berlin, Sektion Mathematik, 1983.
- [6057] T. Zeugmann. On the nonboundability of total effective operators. *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 30(9–11):169–172, 1984.
- [6058] T. Zeugmann. On Barzdin’s conjecture. In *Analogical and Inductive Inference, International Workshop AII ’86, Wendisch-Rietz, GDR, October 1986, Proceedings*, volume 265 of *Lecture Notes in Computer Science*, pages 220–227. Springer-Verlag, 1986.
- [6059] T. Zeugmann. On recursive optimizers. In *Mathematical Methods of Specification and Synthesis of Software Systems ’85, Proceedings of the International Spring School Wendisch-Rietz, GDR, April 22-26, 1985*, volume 215 of *Lecture Notes in Computer Science*, pages 240–245. Springer-Verlag, 1986.
- [6060] T. Zeugmann. On the power of recursive optimizers. *Theoret. Comput. Sci.*, 62(3):289–310, 1988.
- [6061] T. Zeugmann. Inductive inference of optimal programs: A survey and open problems. In *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany, December 1990, Proceedings*, volume 543 of *Lecture Notes in Artificial Intelligence*, pages 208–222. Springer-Verlag, 1991.
- [6062] T. Zeugmann. *Algorithmisches Lernen von Funktionen und Sprachen*. Habilitationsschrift, Technische Hochschule Darmstadt, Fachbereich Informatik, 1993.
- [6063] T. Zeugmann. Average-case analysis of pattern language learning algorithms. In *Algorithmic Learning Theory, 4th International Workshop on Analogical and Inductive Inference, AII ’94, 5th International Workshop on Algorithmic Learning Theory, ALT ’94, Reinhardtsbrunn Castle, Germany, October 1994, Proceedings*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 8–9. Springer-Verlag, 1994.
- [6064] T. Zeugmann. Foreword. *Theoret. Comput. Sci.*, 185(1):1–1, 1997. Special issue for ALT ’95.
- [6065] T. Zeugmann. Lange and Wiehagen’s pattern language learning algorithm: An average-case analysis with respect to its total learning time. *Annals of Mathematics and Artificial Intelligence*, 23(1/2):117–145, 1998. Special issue for AII ’94 and ALT ’94, first appeared as RIFIS-TR-CS-111, 1995.
- [6066] T. Zeugmann. Stochastic finite learning. In *Stochastic Algorithms: Foundations and Applications, International Symposium, SAGA 2001, Berlin, Germany, December 13-14, 2001, Proceedings*, volume 2264 of *Lecture Notes in Computer Science*, pages 155–171. Springer, 2001.

- [6067] T. Zeugmann. Can learning in the limit be done efficiently? In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 17–38. Springer, 2003.
- [6068] T. Zeugmann. From learning in the limit to stochastic finite learning. Technical Report TCS-TR-A-05-8, Division of Computer Science, Hokkaido University, Aug. 2005.
- [6069] T. Zeugmann. From learning in the limit to stochastic finite learning. *Theoret. Comput. Sci.*, 364(1):77–97, 2006. Special issue for ALT 2003.
- [6070] T. Zeugmann. Inductive inference and language learning. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 464–473, Berlin, 2006. Springer.
- [6071] T. Zeugmann. Epsilon covers. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, page 326. Springer, New York, 2010.
- [6072] T. Zeugmann. Epsilon nets. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 326–327. Springer, New York, 2010.
- [6073] T. Zeugmann. PAC learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 745–753. Springer, New York, 2010.
- [6074] T. Zeugmann. Stochastic finite learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 925–928. Springer, New York, 2010.
- [6075] T. Zeugmann. VC dimension. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 1021–1024. Springer, New York, 2010.
- [6076] T. Zeugmann. Epsilon cover. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 408–409. Springer US, New York, first edition, 2017. Invited contribution.
- [6077] T. Zeugmann. Epsilon nets. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 409–410. Springer US, New York, first edition, 2017. Invited contribution.
- [6078] T. Zeugmann. PAC learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 949–959. Springer US, New York, first edition, 2017. Invited contribution.
- [6079] T. Zeugmann. Stochastic finite learning. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 1187–1191. Springer US, New York, first edition, 2017. Invited contribution.
- [6080] T. Zeugmann. VC dimension. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, pages 1323–1327. Springer US, New York, first edition, 2017. Invited contribution.
- [6081] T. Zeugmann. Guest editor’s foreword. *Theoret. Comput. Sci.*, 733:1–3, 2018. Special Issue on Learning Theory and Complexity.
- [6082] T. Zeugmann and S. Lange. A guided tour across the boundaries of learning recursive languages. In *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 190–258. Springer, 1995.

- [6083] T. Zeugmann, S. Lange, and S. Kapur. Characterizations of monotonic and dual monotonic language learning. *Inform. Comput.*, 120(2):155–173, 1995.
- [6084] T. Zeugmann and S. Zilles. Learning recursive functions. Technical Report TCS-TR-A-07-32, Division of Computer Science, Hokkaido University, Nov. 2007.
- [6085] T. Zeugmann and S. Zilles. Learning recursive functions: A survey. *Theoret. Comput. Sci.*, 397(1-3):4–56, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.
- [6086] B.-T. Zhang, J. Yang, and S. W. Chi. Self-organizing latent lattice models for temporal gene expression profiling. *Machine Learning*, 52(1-2):67–89, 2003.
- [6087] C. Zhang and K. Chaudhuri. The extended littlestone’s dimension for learning with mistakes and abstentions. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1584–1616, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [6088] H. Zhang and C. X. Ling. Geometric properties of naive Bayes in nominal domains. In *Machine Learning: ECML 2001, 12th European Conference on Machine Learning, Freiburg, Germany, September 5–7, 2001, Proceedings*, volume 2167 of *Lecture Notes in Artificial Intelligence*, pages 588–599. Springer, 2001.
- [6089] H. Zhang and C. X. Ling. Learnability of augmented naive Bayes in nominal domains. In *Proc. 18th International Conf. on Machine Learning*, pages 617–623. Morgan Kaufmann, San Francisco, CA, 2001.
- [6090] H. Zhang and S. Sra. First-order methods for geodesically convex optimization. In V. Feldman, A. Rakhlin, and O. Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1617–1638, Columbia University, New York, New York, USA, 23–26 June 2016. PMLR.
- [6091] J. Zhang and R. S. Michalski. An integration of rule induction and exemplar-based learning for graded concepts. *Machine Learning*, 21(3):235–267, 1995.
- [6092] L. Zhang, M. Mahdavi, R. Jin, T. Yang, and S. Zhu. Recovering the optimal solution by dual random projection. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 135–157. JMLR.org, 2013.
- [6093] L. Zhang, T. Yang, and R. Jin. Empirical risk minimization for stochastic convex optimization:  $O(1/n)$ - and  $O(1/n^2)$ -type of risk bounds. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1954–1979, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [6094] L. Zhang, T. Yang, R. Jin, and Z.-H. Zhou. Sparse learning for large-scale and high-dimensional data: A randomized convex-concave optimization approach. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19–21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 83–97. Springer, 2016.
- [6095] N. L. Zhang. Hierarchical latent class models for cluster analysis. *Journal of Machine Learning Research*, 5:697–723, 2004.

- [6096] T. Zhang. Theoretical analysis of a class of randomized regularization methods. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory, July 6th-9th, 1999, Santa Cruz, California*, pages 156–163, New York, NY, 1999. ACM Press.
- [6097] T. Zhang. A leave-one-out cross validation bound for kernel methods with applications in learning. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 427–443. Springer, 2001.
- [6098] T. Zhang. A sequential approximation bound for some sample-dependent convex optimization problems with applications in learning. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 65–81. Springer, 2001.
- [6099] T. Zhang. Some sparse approximation bounds for regression problems. In *Proc. 18th International Conf. on Machine Learning*, pages 624–631. Morgan Kaufmann, San Francisco, CA, 2001.
- [6100] T. Zhang. Covering number bounds of certain regularized linear function classes. *Journal of Machine Learning Research*, 2:527–550, 2002.
- [6101] T. Zhang. On the dual formulation of regularized linear systems with convex risk. *Machine Learning*, 46(1-3):91–129, 2002.
- [6102] T. Zhang. On the convergence of MDL density estimation. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 315–330. Springer, 2004.
- [6103] T. Zhang. Data dependent concentration bounds for sequential prediction algorithms. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 173–187. Springer, 2005.
- [6104] T. Zhang. Localized upper and lower bounds for some estimation problems. In *Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 516–530. Springer, 2005.
- [6105] T. Zhang, F. Damerau, and D. Johnson. Text chunking based on a generalization of Winnow. *Journal of Machine Learning Research*, 2:615–637, 2002.
- [6106] T. Zhang and V. S. Iyengar. Recommender systems using linear classifiers. *Journal of Machine Learning Research*, 2:313–334, 2002.
- [6107] T. Zhang and F. J. Oles. A probability analysis on the value of unlabeled data for classification problems. In *Proc. 17th International Conf. on Machine Learning*, pages 1191–1198. Morgan Kaufmann, San Francisco, CA, 2000.
- [6108] W. Zhang. A region-based learning approach to discovering temporal structures in data. In *Proc. 16th International Conf. on Machine Learning*, pages 484–492. Morgan Kaufmann, San Francisco, CA, 1999.

- [6109] W. Zhang. Some improvements on event-sequence temporal region methods. In *Machine Learning: ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings*, volume 1810 of *Lecture Notes in Artificial Intelligence*, pages 446–458. Springer, 2000.
- [6110] X. Zhang. Kernel methods. In C. Sammut and G. I. Webb, editors, *Encyclopedia of Machine Learning*, pages 566–570. Springer, New York, 2010.
- [6111] X. Zhang and M. Numao. Learning and revising theories in noisy domains. In *Algorithmic Learning Theory, 8th International Workshop, ALT '97, Sendai, Japan, October 1997, Proceedings*, volume 1316 of *Lecture Notes in Artificial Intelligence*, pages 339–351. Springer, 1997.
- [6112] X. Zhang, A. Saha, and S. V. N. Vishwanathan. Accelerated training of max-margin Markov networks with kernels. In J. Kivinen, C. Szepesvári, E. Ukkonen, and T. Zeugmann, editors, *Algorithmic Learning Theory, 22nd International Conference, ALT 2011, Aalto University, Espoo, Finland October 5–7, 2011, Proceedings*, volume 6925 of *Lecture Notes in Artificial Intelligence*, pages 292–307, Berlin, Heidelberg, New York, 2011. Springer.
- [6113] X. Zhang, A. Saha, and S. V. N. Vishwanathan. Accelerated training of max-margin Markov networks with kernels. *Theoret. Comput. Sci.*, 519:88–102, 2014. Special Issue: Algorithmic Learning Theory.
- [6114] Y. Zhang, J. Duchi, and M. Wainwright. Divide and conquer kernel ridge regression. In S. Shalev-Shwartz and I. Steinwart, editors, *COLT 2013 - The 26th Annual Conference on Learning Theory, June 12–14, 2013, Princeton University, NJ, USA*, volume 30 of *JMLR Proceedings*, pages 592–617. JMLR.org, 2013.
- [6115] Y. Zhang, P. Liang, and M. Charikar. A hitting time analysis of stochastic gradient langevin dynamics. In S. Kale and O. Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1980–2022, Amsterdam, Netherlands, 07–10 July 2017. PMLR.
- [6116] Y. Zhang, M. J. Wainwright, and M. I. Jordan. Lower bounds on the performance of polynomial-time algorithms for sparse linear regression. In M. F. Balcan, V. Feldman, and C. Szepesvári, editors, *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 921–948, Barcelona, Spain, 13–15 June 2014. PMLR.
- [6117] Y. Zhao and C. G. Atkeson. Projection pursuit learning: some theoretical issues. In S. J. Hanson, T. Petsche, R. L. Rivest, and M. Kearns, editors, *Computational Learning Theory and Natural Learning Systems*, volume II: Intersections Between Theory and Experiment, chapter 22, pages 379–400. MIT Press, 1994.
- [6118] Y. Zhao, S. Chu, Y. Zhou, and K. Tu. Sequence prediction using neural network classifiers. In S. Verwer, M. van Zaanen, and R. Smetsers, editors, *Proceedings of the 13th International Conference on Grammatical Inference*, volume 57 of *Proceedings of Machine Learning Research*, pages 164–169, Delft, The Netherlands, 05–07 Oct. 2017. PMLR.
- [6119] F. Zhdanov and Y. Kalnishkan. An identity for kernel ridge regression. In M. Hutter, F. Stephan, V. Vovk, and T. Zeugmann, editors, *Algorithmic Learning Theory, 21th International Conference, ALT 2010, Canberra, Australia, October 6-8, 2010, Proceedings*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 405–419, Berlin, Heidelberg, New York, 2010. Springer.

- [6120] F. Zhdanov and Y. Kalnishkan. An identity for kernel ridge regression. *Theoret. Comput. Sci.*, 473:157–178, 2013. Special Issue on Learning Theory.
- [6121] D. Zheng, J. Wang, and Y. Zhao. Time series predictions using multi-scale support vector regressions. In *Theory and Applications of Models of Computation, Third International Conference, TAMC 2006, Beijing, China, May 2006, Proceedings*, volume 3959 of *Lecture Notes in Computer Science*, pages 474–481, Berlin, 2006. Springer.
- [6122] Z. Zheng. Constructing X-of-N attributes for decision tree learning. *Machine Learning*, 40(1):35–75, 2000.
- [6123] Z. Zheng and G. I. Webb. Lazy learning of Bayesian rules. *Machine Learning*, 41(1):53–84, 2000.
- [6124] Z. Zheng, G. I. Webb, and K. M. Ting. Lazy Bayesian rules: a lazy semi-naive Bayesian learning technique competitive to boosting decision trees. In *Proc. 16th International Conf. on Machine Learning*, pages 493–502. Morgan Kaufmann, San Francisco, CA, 1999.
- [6125] N. Zhivotovskiy and S. Hanneke. Localization of VC classes: Beyond local rademacher complexities. In R. Ortner, H. U. Simon, and S. Zilles, editors, *Algorithmic Learning Theory, 27th International Conference, ALT 2016, Bari, Italy, October 19-21, 2016, Proceedings*, volume 9925 of *Lecture Notes in Artificial Intelligence*, pages 18–33. Springer, 2016.
- [6126] K. Zhong, P. Jain, and I. S. Dhillon. Efficient matrix sensing using rank-1 gaussian measurements. In K. Chaudhuri, C. Gentile, and S. Zilles, editors, *Algorithmic Learning Theory, 26th International Conference, ALT 2015, Banff, AB, Canada, October 4–6, 2015, Proceedings*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 3–18, Berlin, Heidelberg, New York, 2015. Springer.
- [6127] N. Zhong and S. Ohsuga. A decomposition based induction model for discovering concept clusters from databases. In *Algorithmic Learning Theory, 4th International Workshop, ALT '93, Tokyo, Japan, November 1993, Proceedings*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 384–397. Springer-Verlag, 1993.
- [6128] S. Zhong. Semi-supervised model-based document clustering: A comparative study. *Machine Learning*, 65(1):3–29, 2006.
- [6129] S. Zhong and J. Ghosh. A unified framework for model-based clustering. *Journal of Machine Learning Research*, 4:1001–1037, 2003.
- [6130] H. H. Zhou. CSM: A computational model of cumulative learning. *Machine Learning*, 5:383–406, 1990.
- [6131] S. Zhou, J. Lafferty, and L. Wasserman. Time varying undirected graphs. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 455–466. Omnipress, 2008.
- [6132] Y. Zhou and C. Brodley. A hybrid lazy-eager approach to reducing the computation and memory requirements of local parametric learning algorithms. In *Proc. 16th International Conf. on Machine Learning*, pages 503–512. Morgan Kaufmann, San Francisco, CA, 1999.
- [6133] Y. Zhuravlev and Y. Kogan. Realization of boolean functions with a small number of zeros by disjunctive normal forms, and related problems. *Soviet Math. Doklady*, 32:771–775, 1985.

- [6134] A. Ziehe, M. Kawanabe, S. Harmeling, and K.-R. Müller. Blind separation of post-nonlinear mixtures using linearizing transformations and temporal decorrelation. *Journal of Machine Learning Research*, 4:1319–1338, 2003.
- [6135] A. Ziehe, P. Laskov, G. Nolte, and K.-R. Müller. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *Journal of Machine Learning Research*, 5:777–800, 2004.
- [6136] S. Zilles. On the comparison of inductive inference criteria for uniform learning of finite classes. In *Algorithmic Learning Theory, 12th International Conference, ALT 2001, Washington, DC, USA, November 25–28, 2001, Proceedings*, volume 2225 of *Lecture Notes in Artificial Intelligence*, pages 251–266. Springer, 2001.
- [6137] S. Zilles. On the synthesis of strategies identifying recursive functions. In *14th Annual Conference on Computational Learning Theory, COLT 2001 and 5th European Conference on Computational Learning Theory, EuroCOLT 2001, Amsterdam, The Netherlands, July 2001, Proceedings*, volume 2111 of *Lecture Notes in Artificial Intelligence*, pages 160–176. Springer, 2001.
- [6138] S. Zilles. Merging uniform inductive learners. In *15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 2002, Proceedings*, volume 2375 of *Lecture Notes in Artificial Intelligence*, pages 201–215. Springer, 2002.
- [6139] S. Zilles. Intrinsic complexity of uniform learning. In *Algorithmic Learning Theory, 14th International Conference, ALT 2003, Sapporo, Japan, October 2003, Proceedings*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 39–53. Springer, 2003.
- [6140] S. Zilles. *Uniform Learning of Recursive Functions*, volume 278 of *Dissertationen zur Künstlichen Intelligenz*. Akademische Verlagsgesellschaft, Aka GmbH, Berlin, 2003.
- [6141] S. Zilles. Separation of uniform learning classes. *Theoret. Comput. Sci.*, 313(2):229–265, 2004. Special issue for ALT 2001.
- [6142] S. Zilles. Increasing the power of uniform inductive learners. *J. Comput. Syst. Sci.*, 70(4):510–538, 2005. Special issue on COLT 2002.
- [6143] S. Zilles. An approach to intrinsic complexity of uniform learning. *Theoret. Comput. Sci.*, 364(1):42–61, 2006. Special issue for ALT 2003.
- [6144] S. Zilles. Inductive inference (Case, Smith 1983). In M.-Y. Kao, editor, *Encyclopedia of Algorithms*, pages 411–413. Springer, Berlin, 2008.
- [6145] S. Zilles, S. Lange, R. Holte, and M. Zinkevich. Teaching dimensions based on cooperative learning. In *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 135–146. Omnipress, 2008.
- [6146] M. Zinkevich and T. Balch. Symmetry in Markov decision processes and its implications for single agent and multiagent learning. In *Proc. 18th International Conf. on Machine Learning*, pages 632–639. Morgan Kaufmann, San Francisco, CA, 2001.
- [6147] R. J. Zito-Wolf. Book review. *Machine Learning*, 6:197–204, 1991.
- [6148] R. J. Zito-Wolf. Learning search control knowledge: An explanation-based approach. *Machine Learning*, 6:197–204, 1991.
- [6149] M. Zloof. Query-by-example: a data base language. *IBM Syst. J.*, 16:324–343, 1977.

- [6150] J.-D. Zucker and J.-G. Ganascia. Selective reformulation of examples in concept learning. In *Proc. 11th International Conference on Machine Learning*, pages 352–360. Morgan Kaufmann, 1994.
- [6151] J.-D. Zucker and J.-G. Ganascia. Representation changes for efficient learning in structural domains. In *Proc. 13th International Conference on Machine Learning*, pages 543–551. Morgan Kaufmann, 1996.
- [6152] B. Zupan, M. Bohanec, I. Bratko, and J. Demšar. Machine learning by function decomposition. In *Proc. 14th International Conference on Machine Learning*, pages 421–429. Morgan Kaufmann, 1997.
- [6153] B. Zupan, I. Bratko, M. Bohanec, and J. Demšar. Induction of concept hierarchies from noisy data. In *Proc. 17th International Conf. on Machine Learning*, pages 1199–1206. Morgan Kaufmann, San Francisco, CA, 2000.
- [6154] L. Zwald, O. Bousquet, and G. Blanchard. Statistical properties of kernel principal component analysis. In *Learning Theory: 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1-4, 2004. Proceedings*, volume 3120 of *Lecture Notes in Artificial Intelligence*, pages 594–608. Springer, 2004.
- [6155] J. M. Żytkow. Introduction: Cognitive autonomy in machine discovery. *Machine Learning*, 12:7–16, 1993.
- [6156] J. M. Zytkow and H. A. Simon. A theory of historical discovery: The construction of componential models. *Machine Learning*, 1(1):107–137, 1986.