

Teaching Memoryless Randomized Learners Without Feedback

Frank J. Balbach¹ and Thomas Zeugmann^{*2}

¹ Institut für Theoretische Informatik, Universität zu Lübeck
Ratzeburger Allee 160, 23538 Lübeck, Germany

`balbach@tcs.uni-luebeck.de`

² Division of Computer Science
Hokkaido University, N-14, W-9, Sapporo 060-0814, Japan
`thomas@ist.hokudai.ac.jp`

Abstract. The present paper mainly studies the expected teaching time of memoryless randomized learners without feedback.

First, a characterization of optimal randomized learners is provided and, based on it, optimal teaching times for certain classes are established. Second, the problem of determining the *optimal teaching time* is shown to be \mathcal{NP} -hard. Third, an algorithm for approximating the optimal teaching time is given. Finally, two heuristics for teaching are studied, i.e., cyclic teachers and greedy teachers.

1 Introduction

Teaching studies scenarios in which a teacher gives examples of a target concept c , chosen from a prespecified concept class \mathcal{C} , to a student or a set of students with the aim that the student or all students, respectively, eventually hypothesize c . Classically, the admissible students are deterministic learning algorithms and the teaching performance is measured with respect to the worst case student.

Several models have been proposed to formalize these ideas mathematically. In the inductive inference framework, Freivalds *et al.* [8] and Jain *et al.* [14] developed a model of learning from good examples. Jackson and Tomkins [13] as well as Goldman and Mathias [11, 18] defined models of teacher/learner pairs where teachers and learners are constructed explicitly. In all these models, some kind of adversary disturbing the teaching process is necessary to avoid collusion between the teacher and the learner. Angluin and Kriķis' [1, 2] model prevents collusion by giving incompatible hypothesis spaces to teacher and learner.

Further approaches differ from the ones mentioned above by not constructing the learner but assume a learner or a set of learners is given. In Shinohara and Miyano's [20] model the students are all consistent deterministic learning algorithms and the teacher provides a set of examples for the target concept c such that c is the only concept in the class that is consistent with these examples.

* This work has been supported by the MEXT Grand-in-Aid for Scientific Research on Priority Areas under Grant No. 18049001

Goldman *et al.* [12] and Goldman and Kearns [10] also consider a helpful teacher within the online learning model and investigate how many mistakes a consistent learner can make in the worst case. This number equals the size of the smallest sample in Shinohara and Miyano’s [20] model. This number is called the *teaching dimension* of the target. Then, the difficulty of teaching a class \mathcal{C} is the maximum of the teaching dimensions taken over all $c \in \mathcal{C}$. Because of this similarity we will from now on refer to both models as the *teaching dimension (TD-)model*. The teaching dimension has been studied as a measure for the difficulty to teach a concept class. However, this measure does not always coincide with our intuition, since it can be as large as the maximum value possible, i.e., equal to size of the set of all examples (see, e.g., [4] for an illustrative example).

So, instead of looking at the worst-case, one has also studied the *average teaching dimension* (cf., e.g., [3, 4, 15, 16]). Nevertheless, the resulting model still does not allow to study interesting aspects of teaching such as teaching learners with limited memory or to investigate the difference to teach learners providing and not providing feedback, respectively (cf. [5] for a more detailed discussion). Therefore, in [5] we have introduced a new model for teaching randomized learners. This model is based on the *TD*-model but the set of deterministic learners is replaced by a single randomized one. The teacher gives in each round an example of the target concept to the randomized learner that in turn builds hypotheses. Moreover, the memory of the randomized learner may range from memoryless (just the example received can be used) to unlimited (all examples received so far are available). Additionally, the learner may or may not give feedback by showing its actual guess to the teacher. The teacher’s goal is to make the learner to hypothesize the target and to maintain it as quickly as possible. Now, the teaching performance is measured by the *expected* teaching time (cf. Sect. 2).

In [5] we showed that feedback is provably useful and that varying the learner’s memory size sensibly influences the expected teaching time. Thus, in this paper we focus our attention to randomized learners without feedback and limited memory. If there is *no* feedback, then the teacher can only present an infinite sequence of examples. Teaching infinite sequences introduces difficulties not present in the variant with feedback. As there are uncountably many teachers, there is no way to represent them all finitely. Also their teaching time cannot, in general, be calculated exactly. Finding optimal teachers in the set of all teachers also seems hard; it is not even clear that always an optimal one exists.

So, for getting started, we analyze the model of *memoryless* learners without feedback and ask occasionally which results generalize to any fixed constant memory size. First, we derive a characterization of optimal learners thereby showing that there is always an optimal one (Sect. 3). This enables us to calculate optimal teaching times for certain classes.

We then look at the computational problem of determining the optimal teaching time. No algorithm is known to solve this problem. We show that it is \mathcal{NP} -hard, and there is an (inefficient) algorithm to approximate this value (Sect. 4). Since optimal teachers are hard to find, we study two heuristics for teaching. The greedy one is sometimes optimal (checkable via the characterization in Sect. 3),

but can be arbitrarily far off the optimum (Sect. 5.2). In contrast, teachers iterating over the same sequence of examples forever can come arbitrarily close to optimal, but it is hard to determine whether they in fact *are* optimal (Sect. 5.1).

2 Preliminaries

2.1 Notations

Let X be a finite *instance space* and $\mathcal{X} = X \times \{0, 1\}$ the corresponding set of *examples*. A *concept class* is a set $\mathcal{C} \subseteq 2^X$ of *concepts* $c \subseteq X$. An example (x, v) is *positive* if $v = 1$ and *negative* if $v = 0$. We denote the set of all examples for a concept c by $\mathcal{X}(c) = \{(x, v) \mid v = 1 \iff x \in c\}$ and the set of all concepts *consistent* with an example z by $\mathcal{C}(z) = \{c \in \mathcal{C} \mid z \in \mathcal{X}(c)\}$.

A set $Z \subseteq \mathcal{X}$ is a *teaching set* for a concept $c \in \mathcal{C}$ with respect to a class \mathcal{C} iff c is the only concept in \mathcal{C} consistent with all examples in Z , i.e., $\bigcap_{z \in Z} \mathcal{C}(z) = \{c\}$.

For any set S , we denote by S^* the set of all finite lists of elements from S and by S^ℓ the set of all lists with length ℓ . By $[a, b]$ we mean $\{a, a + 1, \dots, b\}$.

We denote by \mathcal{M}_n the concept class of *monomials* over $X = \{0, 1\}^n$, that is conjunctions of Boolean literals over n variables. We exclude the empty concept from \mathcal{M}_n and can thus identify each monomial with a string from $\{0, 1, *\}^n$ and vice versa. The concept classes \mathcal{S}_n over $X = [1, n]$ are defined as $\mathcal{S}_n = \{[1, n] \setminus \{x\} \mid x \in [1, n]\} \cup \{[1, n]\}$.

2.2 The Teaching Model

The teaching process is divided into rounds. In each round the teacher gives the learner an example of a target concept. The learner chooses a new hypothesis based on this example and on its current hypothesis.

The Learner. As a minimum requirement we demand that the learner’s hypothesis is consistent with the example received in the last round. But the hypothesis is chosen at random from all consistent ones.

We define the goal of teaching as making the learner hypothesize the target *and maintain it*. Consistency alone cannot ensure this, since there may be several consistent hypotheses at every time and the learner would oscillate between them rather than maintaining a single one. To avoid this, the learner has to maintain its hypothesis as long as it is consistent to the new examples (*conservativeness*).

The following algorithm describes the choice of the next hypothesis by the *memoryless randomized learner* in one round of the teaching process.

Input: Current Hypothesis $h \in \mathcal{C}$, example $z \in \mathcal{X}$.

Output: Next Hypothesis $h' \in \mathcal{C}$.

1. **if** $z \notin \mathcal{X}(h)$ **then** pick h' uniformly at random from $\mathcal{C}(z)$;
2. **else** $h' := h$;

In the following the term “learner” refers to the memoryless randomized learner.

In order to make our results depend on \mathcal{C} alone, rather than on an arbitrary initial hypothesis from \mathcal{C} , we stipulate a special initial hypothesis, denoted *init*.

We consider every example inconsistent with *init* and thus *init* is automatically left after the first example and never reached again.

The definition of the learner contains implicitly a function $p: (\mathcal{C} \cup \{\textit{init}\}) \times \mathcal{X} \times (\mathcal{C} \cup \{\textit{init}\}) \rightarrow [0, 1]$ with $p(h, z, h')$ specifying the probability of a transition from hypothesis h to h' when receiving example z .

The Teacher. A teacher is an algorithm that is given a target concept c^* in the beginning and then outputs an example for c^* in each round. A teacher for c^* can thus be regarded as a function $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$.

Definition 1. Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$. Let $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$ be a teacher and $(h_t)_{t \in \mathbb{N}}$ be the series of random variables for the learner's hypothesis at round t . The event "teaching success in round t ," denoted by G_t , is defined as

$$h_{t-1} \neq c^* \quad \wedge \quad \forall t' \geq t: h_{t'} = c^* .$$

The success probability of T is $\Pr \left[\bigcup_{t \geq 1} G_t \right]$. A teacher is successful iff the success probability is 1. For such a teacher we therefore define the teaching time as $E_T(c^*, \mathcal{C}) := \sum_{t \geq 1} t \cdot \Pr[G_t]$. Then the teaching time for the concept c^* is

$$E(c^*, \mathcal{C}) := \inf_T E_T(c^*, \mathcal{C}) .$$

Although the teacher cannot observe the hypotheses, it can at least calculate the probability distribution $\delta: \mathcal{C} \cup \{\textit{init}\} \rightarrow [0, 1]$ over all possible hypotheses. Such a δ contains all knowledge of the teacher about the situation. The probability of being in c^* , however, is irrelevant for the teacher's decision. Only the relations of the probabilities for non-target states are important. Normalizing these probabilities yields a probability distribution $\gamma: \mathcal{C} \cup \{\textit{init}\} \setminus \{c^*\} \rightarrow [0, 1]$ over $\hat{\mathcal{C}} := \mathcal{C} \cup \{\textit{init}\} \setminus \{c^*\}$. Following Patek [19] we call γ an *information state*. We denote by $\gamma^{(0)}$ the initial information state, that is $\gamma^{(0)}(\textit{init}) = 1$.

The definition of the learner defines implicitly a *state transition function* $f: \Gamma \times \mathcal{X} \rightarrow \Gamma$, that is $f(\gamma, z)$ is the follow-up information state after teaching example z to a learner in state γ .

It is possible to describe teachers as functions $\tilde{T}: \Gamma \rightarrow \mathcal{X}(c^*)$ where Γ is the set of all information states. Such a teacher \tilde{T} , when applied to the initial state $\gamma^{(0)}$ and subsequently to all emerging states, yields a teacher $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$.

Remark. If we assume that the learner's hypothesis is observable as feedback then teachers become functions $T: \mathcal{C} \cup \{\textit{init}\} \rightarrow \mathcal{X}(c^*)$. In this model variant *with* feedback, teachers are finite objects (see Balbach and Zeugmann [5]).

Our teaching model without (with) feedback is a special case of an *unobservable (observable) stochastic shortest path problem*, (U)SSPP. Stochastic shortest path problems are more general in that they allow arbitrary transition probabilities and arbitrary costs assigned to each transition. In our teaching models, the transition probabilities are restricted to p and each example has unit cost. For more details on SSPPs see e.g., Bertsekas [6].

3 Existence of Optimal Teachers

The most interesting property of a target concept in our model is its optimal teaching time $E(c^*, \mathcal{C})$. One way to calculate $E(c^*, \mathcal{C})$ is to calculate $E_T(c^*, \mathcal{C})$

for an optimal teacher T . However, as there are uncountably many teachers, it is not even clear whether an optimal teacher exists at all. In this section we derive a characterization of optimal teachers which shows that there is always one. Moreover, it allows us to check whether a given teacher is optimal.

Our result is based on a characterization of optimal “policies” in USSPPs presented by Patek [19] that is applicable to USSPPs satisfying certain assumptions. As our teaching model is a special case of USSPPs, where “policies” correspond to teachers, we have to show that these assumptions hold in our teaching model.

To state Patek’s [19] characterization and the assumptions under which it works, it is inevitable to introduce some further technical notation. Moreover, the optimality criterion is based on information state teachers $\tilde{T}: \Gamma \rightarrow \mathcal{X}(c^*)$ rather than sequential teachers $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$.

Following Patek [19], we consider series $(\tilde{T}_t)_{t \in \mathbb{N}}$ of teachers. Such a series is called *stationary* if all teachers in it are identical. For such a series $(\tilde{T}_t)_{t \in \mathbb{N}}$, we denote by $\Pr^m(\gamma, \tilde{T})$ the probability that a learner reaches c^* within m rounds when it is started in $\gamma \in \Gamma$ and is taught by teacher \tilde{T}_t in round $t = 0, 1, \dots$.

We also need two so called *dynamic programming operators* D and $D_{\tilde{T}}$ mapping functions $G: \Gamma \rightarrow \mathbb{R}$ to functions of the same type:

$$[D_{\tilde{T}}G](\gamma) = 1 + G(f(\gamma, \tilde{T}(\gamma))) \cdot \sum_{c, d \in \hat{\mathcal{C}}} \gamma(c) \cdot p(c, \tilde{T}(\gamma), d) ,$$

$$[DG](\gamma) = \min_{z \in \mathcal{X}(c^*)} \left(1 + G(f(\gamma, z)) \cdot \sum_{c, d \in \hat{\mathcal{C}}} \gamma(c) \cdot p(c, z, d) \right).$$

The sum $\sum_{c, d \in \hat{\mathcal{C}}} \gamma(c) \cdot p(c, \tilde{T}(\gamma), d)$ yields the probability for *not* reaching c^* in the next round after being taught $\tilde{T}(\gamma)$ in state γ . To get an intuition about above formulas, it is helpful to think of a value $G(\gamma)$ as the expected number of rounds to reach the target when the learner starts in state γ . Then $[D_{\tilde{T}}G](\gamma)$ specifies for every initial state γ the expected number of rounds under teacher \tilde{T} , assuming that for all other states the expectations are given by G .

Given a teacher series $(\tilde{T}_t)_{t \in \mathbb{N}}$, the expected time to reach the target when starting in $\gamma \in \Gamma$ is denoted by $G_{\tilde{T}}(\gamma)$. This yields a function $G_{\tilde{T}}: \Gamma \rightarrow \mathbb{R}$.

The characterization, in terms of the randomized teaching model, now is:

Theorem 2 ([19]). *Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$ a target. Assume that*

- (a) *There is a stationary series $(\tilde{T}_t)_{t \in \mathbb{N}}$ with $\lim_{m \rightarrow \infty} \Pr^m(\gamma, \tilde{T}) = 1$ for all $\gamma \in \Gamma$.*
- (b) *For every series $(\tilde{T}_t)_{t \in \mathbb{N}}$ not satisfying (a), a subsequence of*

$$([D_{\tilde{T}_0} D_{\tilde{T}_1} \cdots D_{\tilde{T}_t} \mathbf{0}](\gamma))_{t=0}^{\infty} \quad \text{tends to infinity for some } \gamma \in \Gamma .$$

Then

1. *The operator D has a fixed point G , that is $DG = G$.*
2. *A teacher $\tilde{T}: \Gamma \rightarrow \mathcal{X}(c^*)$ is optimal (i.e., has minimal teaching time) iff there is a $G: \Gamma \rightarrow \mathbb{R}$ such that $DG = G$ and $D_{\tilde{T}}G = G$.*

Roughly speaking, Theorem 2 says: If (a) there is a teacher successful from every initial state and if (b) every non-successful teacher has an infinite teaching time from at least one initial state, then there is an optimal teacher and its teaching time G is just the fixed point of the operator D .

We now show that conditions (a) and (b) hold for all classes and targets in our model. For condition (a) we show that a greedy teacher is always successful.

Definition 3. A teacher \tilde{T} for $c^* \in \mathcal{C}$ is called greedy iff for all $\gamma \in \Gamma$

$$\tilde{T}(\gamma) \in \operatorname{argmax}_{z \in \mathcal{X}(c^*)} \sum_{c \in \hat{\mathcal{C}}} \gamma(c) \cdot p(c, z, c^*).$$

Note that replacing γ with δ and $\hat{\mathcal{C}}$ with $\mathcal{C} \cup \{\text{init}\}$ yields an equivalent definition.

Lemma 4. Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$. Let T be the sequential teacher for some greedy teacher \tilde{T} . Then T is successful for c^* .

Proof. We denote by $\delta_t: \mathcal{C} \cup \{\text{init}\} \rightarrow [0, 1]$ the probabilities of the hypotheses in round t under teacher T . In each round t , T picks an example z maximizing $\sum_{c \in \mathcal{C} \cup \{\text{init}\}} \delta_t(c) \cdot p(c, z, c^*)$. We lower bound this value.

There is a concept $c' \neq c^*$ with $\delta_t(c') \geq (1 - \delta_t(c^*)) / |\mathcal{C}|$. Let z' be an example inconsistent with c' . Then $p(c', z, c^*) \geq 1 / |\mathcal{C}|$ and therefore $\sum_c \delta_t(c) \cdot p(c, z', c^*) \geq (1 - \delta_t(c^*)) / |\mathcal{C}|^2$. As T maximizes this sum, we have also for $z = T(t)$ that $\sum_c \delta_t(c) \cdot p(c, z, c^*) \geq (1 - \delta_t(c^*)) / |\mathcal{C}|^2$. This sum also equals $\delta_{t+1}(c^*) - \delta_t(c^*)$ and therefore

$$1 - \delta_{t+1}(c^*) \leq (1 - 1/|\mathcal{C}|^2) \cdot (1 - \delta_t(c^*)).$$

Hence, $1 - \delta_t(c^*) \rightarrow 0$ as $t \rightarrow \infty$ and the probability $\delta_t(c^*)$ tends to one. \square

We only sketch the technical proof that condition (b) is satisfied too.

Lemma 5. Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$ a target. Then (b) holds.

Proof. (Sketch) Let $(\tilde{T}_t)_{t \in \mathbb{N}}$ be a series that does not satisfy condition (a). Then there is a γ with $\lim_{m \rightarrow \infty} \Pr^m(\gamma, \tilde{T}) < 1$. This means $\lim_{m \rightarrow \infty} \delta_m(c^*) < 1$, where $(\delta_t)_{t \in \mathbb{N}}$ is the series resulting from application of \tilde{T} to γ . The expected number of rounds to reach c^* is infinite in this case. Patek [19] shows that this expectation is also $\liminf_{t \rightarrow \infty} [D_{\tilde{T}_0} \cdots D_{\tilde{T}_t} \mathbf{0}](\gamma)$, where $\mathbf{0}: \Gamma \rightarrow \mathbb{R}$ is the zero function.

Hence the sequence $([D_{\tilde{T}_0} D_{\tilde{T}_1} \cdots D_{\tilde{T}_t} \mathbf{0}](\gamma))_{t=0}^\infty$ tends to infinity. \square

Theorem 2 requires us to find a $G: \Gamma \rightarrow \mathbb{R}$ and to define a teacher $\tilde{T}: \Gamma \rightarrow \mathcal{X}(c^*)$. However, most of the states in Γ cannot be reached from the initial state $\gamma^{(0)}$ and it seems unnecessary to specify \tilde{T} 's behavior for the unreachable states too. As a matter of fact, it suffices to define G and \tilde{T} for the *reachable states* in Γ denoted by $\Gamma_0 = \{\gamma \in \Gamma \mid \exists t \exists z_0, \dots, z_t: \gamma = f(\dots f(f(\gamma^{(0)}, z_0), z_1) \dots, z_t)\}$. We omit the proof thereof and state the final version of the characterization.

Corollary 6. Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$ a target. A teacher $\tilde{T}: \Gamma_0 \rightarrow \mathcal{X}(c^*)$ is optimal iff there is a $G: \Gamma_0 \rightarrow \mathbb{R}$ such that $DG = G$ and $D_{\tilde{T}}G = G$, where D and $D_{\tilde{T}}$ have to be restricted suitably to work on functions $G: \Gamma_0 \rightarrow \mathbb{R}$.

One advantage of using Γ_0 instead of Γ is that we have to consider only one state with $\gamma(\text{init}) > 0$, namely the initial state $\gamma^{(0)}$. For illustration, we apply Corollary 6 to the class \mathcal{S}_n in order to find an optimal teacher for $[1, n]$.

Fact 7. *Let $c^* = [1, n] \in \mathcal{S}_n$. Then the teacher $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$ with $T(i) = (1 + (i \bmod n), 1)$ is an optimal teacher for c^* with teaching time $n(n-1)/2 + 1$.*

Proof. The proof proceeds in several steps. First we define a teacher $\tilde{T}: \Gamma_0 \rightarrow \mathcal{X}(c^*)$ and a function $G: \Gamma_0 \rightarrow \mathbb{R}$. Then we show that $DG = G$ and $D_{\tilde{T}} = D$ from which we conclude that \tilde{T} is optimal. Finally we show that \tilde{T} , when applied to $\gamma^{(0)}$, generates the same example sequence as T .

For a $\gamma \in \Gamma$ and $i \in [1, n]$ we set as shortcut $\gamma_i := \gamma(c)$ for $c = [1, n] \setminus \{i\}$. A positive example $(x, 1)$ is inconsistent only with the concept $[1, n] \setminus \{x\}$. Teaching $(x, 1)$ in a state $\gamma \neq \gamma^{(0)}$ results in a state $f(\gamma, (x, 1)) = \hat{\gamma}$ with $\hat{\gamma}_i = \frac{\gamma_i + \gamma_x/n}{1 - \gamma_x/n}$ for $i \neq x$, and $\hat{\gamma}_x = 0$. For $\gamma = \gamma^{(0)}$ we have $\hat{\gamma}_i = 1/(n-1)$ for all $i \neq x$ and $\hat{\gamma}_x = 0$.

We define \tilde{T} to be a greedy teacher. If there are several equally “greedy” examples, \tilde{T} picks the one with smallest instance. As every example is inconsistent with exactly one concept, \tilde{T} greedily picks an example that is inconsistent with a most probable hypothesis.

For defining G , let $\gamma \in \Gamma_0 \setminus \{\gamma^{(0)}\}$ and assume without loss of generality $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_n$. Let $F = \frac{(n-1)n}{2}$. Then we define

$$G(\gamma) := F + \sum_{i=1}^n \gamma_i \cdot i \quad \text{and} \quad G(\gamma^{(0)}) := F + 1.$$

Next we show $DG = G$. Let $\gamma \in \Gamma_0 \setminus \{\gamma^{(0)}\}$ and recall that $\gamma_1 \geq \dots \geq \gamma_n$. We have to show that $[DG](\gamma) = G(\gamma)$, in other words that $1 + \min_{(x,1) \in \mathcal{X}} G(f(\gamma, (x, 1))) \cdot \sum_{c,d \in \hat{\mathcal{C}}} p(c, (x, 1), d) = G(\gamma)$. Since $\sum_{c,d} p(c, (x, 1), d) = 1 - \gamma_x/n$ this means

$$1 + \min_{(x,1) \in \mathcal{X}} G(f(\gamma, (x, 1))) \cdot (1 - \gamma_x/n) = G(\gamma). \quad (1)$$

Let $z = (x, 1) \in \mathcal{X}$ and $\hat{\gamma} = f(\gamma, z)$. Then $\hat{\gamma}_1 \geq \dots \geq \hat{\gamma}_{z-1} \geq \hat{\gamma}_{z+1} \geq \dots \geq \hat{\gamma}_n \geq \hat{\gamma}_z = 0$. The expression to be minimized is

$$\begin{aligned} \left(1 - \frac{\gamma_x}{n}\right) \cdot G(\hat{\gamma}) &= \left(1 - \frac{\gamma_x}{n}\right) \cdot \left(F + \sum_{i \leq x-1} i \hat{\gamma}_i + \sum_{i \geq x+1} (i-1) \hat{\gamma}_i\right) \\ &= \left(1 - \frac{\gamma_x}{n}\right) \cdot \left(F + \sum_{i \leq x-1} i \cdot \frac{\gamma_i + \gamma_x/n}{1 - \gamma_x/n} + \sum_{i \geq x+1} (i-1) \frac{\gamma_i + \gamma_x/n}{1 - \gamma_x/n}\right) \\ &= F + \sum_{i=1}^n i \gamma_i - \left(x \cdot \gamma_x + \sum_{i \geq x+1} \gamma_i\right). \end{aligned} \quad (*)$$

From $\gamma_1 \geq \dots \geq \gamma_n$, it follows $1\gamma_1 + \sum_{i \geq 2} \gamma_i \geq 2\gamma_2 + \sum_{i \geq 3} \gamma_i \geq \dots \geq n\gamma_n$. This means that the expression (*) is minimal for $x = 1$, or $\gamma_x = \gamma_1$. Setting $x = 1$ yields $\min_{(x,1)} G(f(\gamma, (x, 1))) \cdot (1 - \gamma_x/n) = F - 1 + \sum_{i=1}^n i \gamma_i = G(\gamma) - 1$ and thus Equation (1) is satisfied.

It remains to show $[DG](\gamma^{(0)}) = G(\gamma^{(0)})$. For all examples $(x, 1) \in \mathcal{X}$ we have $[DG](\gamma^{(0)}) = 1 + (1 - \frac{1}{n})G(f(\gamma^{(0)}, (x, 1))) = 1 + (1 - \frac{1}{n}) \cdot \left(F + \sum_{i=1}^{n-1} i \frac{1/n}{1-1/n} \right) = 1 + \frac{n-1}{n} \cdot \left(F + \frac{1}{n-1} \cdot \frac{n(n-1)}{2} \right) = 1 + \frac{n(n-1)}{2} = F + 1 = G(\gamma^{(0)})$.

It follows that $[DG](\gamma) = G(\gamma)$ for all $\gamma \in \Gamma_0$. Moreover, teacher \tilde{T} always picks the example $(x, 1)$ minimizing the term in Equation (1), thus $D_{\tilde{T}}G = G$ and \tilde{T} is optimal according to Corollary 6.

The teacher \tilde{T} , when started in $\gamma^{(0)}$, generates the same sequence of examples as T . By the definition of \tilde{T} , $\tilde{T}(\gamma^{(0)}) = (1, 1)$ and for $\gamma \neq \gamma^{(0)}$ with $\gamma_1 \geq \dots \geq \gamma_n$ (w.l.o.g.) \tilde{T} chooses example $(1, 1)$ and the next information state is $\hat{\gamma}$ with $\hat{\gamma}_2 \geq \dots \geq \hat{\gamma}_n \geq \hat{\gamma}_1 = 0$. Therefore, \tilde{T} chooses $(2, 1)$ next and so on. \square

With *feedback* $[1, n] \in \mathcal{S}_n$ can be taught in expected n rounds: A teacher T observing the learner's hypothesis can always choose an inconsistent example. Under T , the learner has in each round a probability of $1/n$ of reaching the target. But teaching $[1, n] \in \mathcal{S}_n$ *without* feedback requires $\Omega(n^2)$ rounds (cf. Fact 7).

As the previous fact also shows, to prove the optimality of a sequential teacher, we have to take a detour via information state teachers. Thus finding the “right” information state teacher is the crucial step in applying Corollary 6.

4 Finding Optimal Teachers

Now that we know that there is always an optimal teacher, we ask how to find one effectively. But as these teachers are infinite sequences of examples, it is unclear how an “optimal teacher finding”-algorithm should output one. Alternatively, we could seek a *generic optimal teacher*, that is an algorithm receiving a class, a target c^* , and a finite example sequence, and outputting an example such that its repeated application yields an optimal teacher for c^* .

A closely related task is to determine the teaching time of an optimal teacher, that is $E(c^*, \mathcal{C})$.

Definition 8. *We call the following problem OPT-TEACHING-TIME.*

Input: *Concept class \mathcal{C} , concept $c^* \in \mathcal{C}$, rational number F .*

Question: *Is $E(c^*, \mathcal{C}) \leq F$?*

In the more general setting of USSPPs the analog problem is undecidable (see Madani *et al.* [17] and Blondel and Canterini [7]). This can be seen as evidence for the undecidability of OPT-TEACHING-TIME. On the other hand, USSPPs differ from our model in some complexity aspects. For example, deciding whether there is a teacher with at least a given success probability is easy (because there is always one), whereas the analog problem for USSPPs is undecidable [17, 7].

Although the decidability of OPT-TEACHING-TIME is open, we can at least show it is \mathcal{NP} -hard. So even if there is an algorithm, it is presumably inefficient.

The proof is by reduction from the EXACT-3-COVER (X3C) problem [9]. The following algorithm computes OPT-TEACHING-TIME instances from X3C instances in polynomial time (see Fig.1 for an example).

	x_1	x_2	x_3	y_1	y_2	y_3	y_4	y_5	y_6
$B = \{1, 2, 3, 4, 5, 6\},$	c^*	1	1	1	1	1	1	1	1
$A_1 = \{2, 4, 5\},$	c_1	1	0	0	0	1	1	1	1
$A_2 = \{1, 3, 5\},$	c_2	0	1	1	1	0	1	1	1
$A_3 = \{1, 3, 6\}$	c_3	1	0	0	1	1	0	1	1
	c_4	0	1	1	1	1	0	1	1
	c_5	0	0	1	1	1	1	0	1
	c_6	1	1	0	1	1	1	1	0

Fig. 1. Illustration of the reduction from X3C to OPT-TEACHING-TIME. Every example on the left is inconsistent with exactly three concepts; y_1, \dots, y_6 are “dummy” instances making all concepts unique. The examples $(x_1, 1), (x_3, 1)$ have the X3C property.

Input: Set $B = [1, 3n]$ ($n \in \mathbb{N}$), sets $A_1, \dots, A_m \subseteq B$ with $|A_i| = 3$.

1. $X := \{x_1, \dots, x_m\} \cup \{y_1, \dots, y_{3n}\}$
2. $c_j := \{x_i \mid j \notin A_i\} \cup \{y_i \mid i \neq j\}$ for $j = 1, \dots, 3n$
3. $c^* := X$
4. $\mathcal{C} := \{c^*, c_1, \dots, c_{3n}\}$
5. **Output** $\langle \mathcal{C}, c^*, 1 + \frac{3}{2}n(n-1) \rangle$

We call a concept class \mathcal{C} output by this algorithm a *positive* or *negative X3C class* depending on whether the input was a positive or negative X3C instance.

An X3C class is positive iff there are examples $z_1, \dots, z_n \in \mathcal{X}(c^*)$ such that the sets $\mathcal{C} \setminus \mathcal{C}(z_j)$ are pairwise disjoint for $j = 1, \dots, n$ and $\bigcup_j (\mathcal{C} \setminus \mathcal{C}(z_j)) = \mathcal{C} \setminus \{c^*\}$. Examples z_1, \dots, z_n satisfying the property just stated have the *X3C property*.

If A_1, \dots, A_m consists of all $m = \binom{3n}{3}$ subsets of B we call the class a *full X3C class*. Every full X3C class is a positive X3C class.

Of all X3C classes, the full X3C classes are easiest to analyze because of their intrinsic symmetries. Moreover, the optimal teachers are just the greedy teachers, which simplifies the application of our optimality criterion. Note that for arbitrary X3C classes a greedy teacher need not to be optimal.

Lemma 9. *Let $n \geq 2$, let \mathcal{C} be a full X3C class for n and c^* be the concept containing all instances. Then a teacher $\tilde{T}: \Gamma_0 \rightarrow \mathcal{X}(c^*)$ is optimal if and only if \tilde{T} is greedy. The teaching time, when starting in $\gamma^{(0)}$, is $1 + \frac{3}{2}n(n-1)$.*

Proof. (sketch) The class \mathcal{C} is similar to the class \mathcal{S} only with three zeros per column instead of one. Consequently the proof that all greedy teachers are optimal is similar to that of Fact 7. That the “dummy” examples are never chosen by an optimal teacher and that all optimal teachers are greedy can be proved by straightforward but technically involved application of Corollary 6. \square

The next lemma describes the optimal teachers as example sequences rather than in terms of the information states.

Lemma 10. *Let $n \geq 2$, let \mathcal{C} be a full X3C class for n and c^* be the concept containing all instances. A teacher $T: \mathbb{N} \rightarrow \mathcal{X}(c^*)$ is optimal if and only if $T(t) = z_{t \bmod n}$ for all t with the examples z_0, \dots, z_{n-1} having the X3C property.*

Proof. This proof is similar to the last paragraph of the proof of Fact 7. We omit the technical details. \square

So far, we have characterized the optimal teachers for *full* X3C classes.

Lemma 11. *Let \mathcal{C} be an X3C class. Then $E(c^*, \mathcal{C}) = 1 + \frac{3}{2}n(n-1)$ if and only if \mathcal{C} is a positive X3C class.*

Proof. For the if-direction, let $z_1, \dots, z_n \in \mathcal{X}(c^*)$ have the X3C property.

The teacher T defined by $T(t) = z_{t \bmod n}$ has a teaching time of $1 + \frac{3}{2}n(n-1)$. This follows similar to Lemma 10. If there was a better teacher, this teacher would also have a smaller teaching time when applied to the full X3C class, thus contradicting Lemma 10.

For the only-if direction, assume $E(c^*, \mathcal{C}) = 1 + \frac{3}{2}n(n-1)$ and suppose that \mathcal{C} is a negative X3C class. Then there is a teacher \bar{T} for c^* with teaching time $1 + \frac{3}{2}n(n-1)$, but not iterating over a sequence of examples $z_1, \dots, z_n \in \mathcal{X}(c^*)$ with the X3C property (because negative X3C classes have no such examples). The teacher T would then have the same teaching time with respect to a full X3C class, too. Hence, T would be an optimal teacher for the full X3C class, a contradiction to Lemma 10. \square

Using Lemma 11 we can show our main result.

Theorem 12. *The problem OPT-TEACHING-TIME is \mathcal{NP} -hard.*

Proof. Let $\langle B, A_1, \dots, A_m \rangle$ with $B = [1, 3n]$ be an instance of X3C and let $\langle \mathcal{C}, c^*, 1 + \frac{3}{2}n(n-1) \rangle$ be the instance of OPT-TEACHING-TIME resulting from the polynomial time reduction on Page 9.

By definition $\langle B, A_1, \dots, A_m \rangle$ is a positive instance of X3C iff \mathcal{C} is a positive X3C class. The latter holds iff $E(c^*, \mathcal{C}) = 1 + \frac{3}{2}n(n-1)$ (by Lemma 11). This in turn holds iff $\langle \mathcal{C}, c^*, 1 + \frac{3}{2}n(n-1) \rangle$ is a positive OPT-TEACHING-TIME instance. \square

The last theorem implies that no polynomial time generic optimal teacher exists (unless $\mathcal{P} = \mathcal{NP}$).

In our teaching model it is at least possible to effectively approximate E with arbitrary precision.

Fact 13. *There is an algorithm with:*

Input: *Concept class \mathcal{C} , concept $c \in \mathcal{C}$, precision $\varepsilon > 0$.*

Output: *$F \in \mathbb{R}$ with $|F - E(c, \mathcal{C})| < \varepsilon$.*

Proof. Roughly speaking, the probability for not being in the target state tends to zero as the sequence of examples given by the teacher grows. The idea of the algorithm in Fig. 2 is to approximate the expectations for growing finite sequences of examples until the probability of not being in the target state becomes negligibly small.

The values $\Pr[h_i = c \wedge h_{i-1} \neq c]$ can be calculated according to the state transition function f . Its values are always rational numbers which can be calculated and stored exactly. The value D is an upper bound for the expected

Input: Concept class \mathcal{C} , concept $c \in \mathcal{C}$, rational number $\varepsilon > 0$.

1. $D := |X| \cdot |\mathcal{C}|$
2. **for** $\ell = 1, 2, \dots$:
3. **for all** $\alpha \in \mathcal{X}(c)^\ell$:
 - // denote with h_i ($i = 1, \dots, \ell$) the random variable for the
 - // hypothesis the teacher after round i when taught α .
4. $b(\alpha) := \sum_{i=1}^{\ell} i \cdot \Pr[h_i = c \wedge h_{i-1} \neq c] + (\ell + 1) \cdot \Pr[h_\ell \neq c]$
5. $B(\alpha) := \sum_{i=1}^{\ell} i \cdot \Pr[h_i = c \wedge h_{i-1} \neq c] + (\ell + D) \cdot \Pr[h_\ell \neq c]$
6. $b_\ell := \min\{b(\alpha) \mid \alpha \in \mathcal{X}(c)^\ell\}$
7. **if** $\exists \alpha \in \mathcal{X}(c)^\ell : B(\alpha) - b_\ell < \varepsilon$:
8. **Output** $B(\alpha)$.

Fig. 2. Algorithm computing an approximation of $E(c^*, \mathcal{C})$.

number of rounds to reach c regardless of the initial state of the learner. That means that in every state of the learner teaching can be continued such that the target is reached in expected at most D rounds.

The values $b(\alpha)$ and $B(\alpha)$ are a lower and an upper bound for the teaching time of a teacher starting with example sequence α . To verify this note that $\sum_{i=1}^{\ell} i \cdot \Pr[h_i = c \wedge h_{i-1} \neq c]$ is the expectation considering the first ℓ rounds only. The remaining probability mass $\Pr[h_\ell \neq c]$ needs at least 1 and at most D additional rounds, which yields $b(\alpha)$ and $B(\alpha)$, respectively.

It follows that $B(\alpha) \geq E(c, \mathcal{C})$ for all $\alpha \in \mathcal{X}(c)^*$. Moreover, since every teacher starts with some example series $\alpha \in \mathcal{X}(c)^\ell$, the values b_ℓ are all lower bounds for $E(c, \mathcal{C})$, that is $b_\ell \leq E(c, \mathcal{C})$ for all $\ell \geq 1$. Therefore the output $B(\alpha)$ with $B(\alpha) - b_\ell < \varepsilon$ is an ε -approximation for $E(c, \mathcal{C})$.

It remains to show the termination of the algorithm. To this end we show:

Claim. $\lim_{\ell \rightarrow \infty} b_\ell = E(c, \mathcal{C})$.

Proof. Let $\delta > 0$ and set $\ell_0 := (D \cdot E(c, \mathcal{C}))/\delta$. We show that for all $\ell \geq \ell_0$, $|E(c, \mathcal{C}) - b_\ell| < \delta$. Let $\ell \geq \ell_0$. Then $\ell \geq (D \cdot E(c, \mathcal{C}))/\delta$.

Let $\alpha \in \mathcal{X}(c)^\ell$ such that $b(\alpha) = b_\ell$. Then $b(\alpha) \leq E(c, \mathcal{C})$ and therefore $(\ell + 1) \cdot \Pr[h_\ell \neq c] \leq E(c, \mathcal{C})$. It follows $\Pr[h_\ell \neq c] \leq E(c, \mathcal{C})/(\ell + 1)$.

For $B(\alpha)$ we have

$$B(\alpha) = b(\alpha) + \Pr[h_\ell \neq c] \cdot (D - 1) \leq b(\alpha) + \frac{E(c, \mathcal{C})}{\ell + 1} \cdot (D - 1) < b(\alpha) + \frac{E(c, \mathcal{C})}{\ell} \cdot D.$$

Using $1/\ell \leq \delta/(D \cdot E(c, \mathcal{C}))$, we get $B(\alpha) < b(\alpha) + \delta$.

On the other hand, $E(c, \mathcal{C}) \leq B(\alpha)$ and therefore $E(c, \mathcal{C}) < b(\alpha) + \delta$, hence $E(c, \mathcal{C}) - b_\ell = E(c, \mathcal{C}) - b(\alpha) < \delta$. □ Claim

To prove the termination of the algorithm we have to show that there is an α such that $B(\alpha) - b_\ell < \varepsilon$. Let $T: \mathbb{N} \rightarrow \mathcal{X}(c)$ be an optimal teacher and denote $\langle T(0), \dots, T(\ell - 1) \rangle \in \mathcal{X}(c)^\ell$ by $T_{0:\ell}$. Then $\lim_{\ell \rightarrow \infty} B(T_{0:\ell}) = E(c, \mathcal{C})$. Together with $\lim_{\ell \rightarrow \infty} b_\ell = E(c, \mathcal{C})$ it follows $\lim_{\ell \rightarrow \infty} (B(T_{0:\ell}) - b_\ell) = 0$. That means for sufficiently long $\alpha = T_{0:\ell}$, the condition $B(T_{0:\ell}) - b_\ell < \varepsilon$ is satisfied. □

5 Heuristics for Teaching

As it seems difficult to find an optimal teacher, next we study teaching heuristics.

5.1 Cyclic Teachers

Probably the simplest teachers are those that give the same sequence of examples over and over again. Such a *cyclic teacher* is identified with the sequence $(z_0, \dots, z_{m-1}) \in \mathcal{X}^m$ of examples it teaches.

Fact 14. *Let \mathcal{C} be a concept class and $c^* \in \mathcal{C}$ a target concept. A cyclic teacher (z_0, \dots, z_{m-1}) is successful iff $\{z_0, \dots, z_{m-1}\}$ is a teaching set for c^* wrt. \mathcal{C} .*

Not only is success of a cyclic teacher easy to decide, the teaching time is also efficiently computable.

Lemma 15. *The teaching time of a cyclic teacher can be computed from the sequence of examples that the teacher repeats.*

Proof. Let \mathcal{C} be a concept class and let $c^* \in \mathcal{C}$. Let T be a cyclic teacher repeating z_0, \dots, z_{m-1} . We assume that the examples constitute a teaching set.

Teaching will be successful no matter at which of the examples z_i the loop starts. We denote by F_i ($0 \leq i < m$) the teaching time for the teacher $T_i: T(t) = z_{(i+t) \bmod m}$ starting with example z_i . For $h \in \mathcal{C}$ we denote by $F_i(h)$ the teaching time for teacher T_i when the learner's initial hypothesis is h . For convenience throughout this proof all subscripts of T, z , and F are to be taken modulo m .

We can now state a linear equation for F_i involving all F_j with $j \neq i$. Consider the teacher T_i and the learner's state δ after the first example, z_i , has been given. The learner assumes all hypotheses $h \in \mathcal{C}(z_i)$ with equal probability $\delta(h) = 1/|\mathcal{C}(z_i)|$ and all other hypotheses with probability $\delta(h) = 0$.

The expectation F_i is one plus the expectation for teacher T_{i+1} when the learner starts in state δ . This expectation equals the weighted sum of the expectations of teacher T_{i+1} starting in state h , that is

$$F_i = 1 + \sum_{h \in \mathcal{C} \setminus \{c^*\}} \delta(h) \cdot F_{i+1}(h) .$$

We now determine $F_{i+1}(h)$. Consider a learner in state $h \neq c^*$ and a teacher giving z_{i+1}, z_{i+2}, \dots . The learner will change its state only when the first example inconsistent with h arrives (such an example exists since the z_i 's form a teaching set for c^*). Let z_{i+k} be this example. Beginning with z_{i+k} , teaching proceeds as if teacher T_{i+k} had started from the *init* state. Therefore $F_{i+1}(h) = (k-1) + F_{i+k}$.

If we denote for $i = 0, \dots, m-1$ and for $k = 1, \dots, m$,

$$\mathcal{C}_{i,k} = \{h \in \mathcal{C} \setminus \{c^*\} \mid h \in \mathcal{C}(z_i), h \notin \mathcal{C}(z_{i+1}), \dots, h \notin \mathcal{C}(z_{i+k-1}), h \in \mathcal{C}(z_{i+k})\},$$

then we get the following linear equation for F_i :

$$F_i = 1 + \sum_{k=1}^m \frac{|\mathcal{C}_{i,k}|}{|\mathcal{C}(z_i)|} \cdot ((k-1) + F_{i+k}) .$$

In this manner we get m linear equations in the variables F_0, \dots, F_{m-1} . Denoting the solution vector by \mathbf{F} we get a linear equation system of the form $(\mathbf{1} - C) \cdot \mathbf{F} = U$, where $\mathbf{1}$ is the $m \times m$ unit matrix, C is a substochastic matrix composed of entries of the form $|C_{i,k}|/|C(z_i)|$ and zeros. Thus $\mathbf{1} - C$ is invertible and the values F_0, \dots, F_{m-1} are uniquely determined by the equation system. F_0 is the sought teaching time. \square

The algorithm in Fig. 2 computes an $\alpha \in \mathcal{X}^*$ such that every extension of α yields a teacher ε -close to optimal. In particular, this holds for the cyclic teacher α . Hence, cyclic teachers can be arbitrarily close to the optimal teacher.

A drawback of cyclic teachers T is that they do not directly yield an information state teacher \tilde{T} . Thus the optimality criterion cannot immediately be applied. But cyclic teachers can be used to calculate upper bounds on $E(c^*, \mathcal{C})$.

Fact 16. *Let $k \geq 3$, $n \geq k$. The monomial $1^k \ast^{n-k}$ has a teaching time of at most $\frac{-2+2^{k+1}+7 \cdot 2^n - 2^{n+k+2} + 2^{2+k} \cdot 3^n + 2^{n+2} \cdot 3^n - 2 \cdot 3^{n+1} - 4^{n+1} - 2(2^k - 2^{n+1} + 2 \cdot 3^n)k}{2 \cdot 3^n - 2^{n+1} + 2^k}$.*

Proof. This teaching time is achieved by the following cyclic teacher T . The teacher provides alternately positive and negative examples. The positive examples alternate between the two complementary examples $1^k 0^{n-k}$ and $1^k 1^{n-k}$.

The first k characters of the negative examples iterate through 01^{k-1} , 101^{k-2} , $\dots, 1^{k-1}0$, the last $n - k$ characters equal the last $n - k$ characters of the immediately preceding positive example. For example, let $k = 3$ and $n = 5$. Then the example sequence is $(11100, 1)$, $(01100, 0)$, $(11111, 1)$, $(10111, 0)$, $(11100, 1)$, $(11000, 0)$, $(11111, 1)$, $(01111, 0)$, $(11100, 1)$, $(10100, 0)$, $(11111, 1)$, $(11011, 0)$.

Applying the method of the proof of Lemma 15, the expected teaching time of this teacher can be calculated. We omit the details. \square

For comparison, the optimal teaching time for monomials in the scenario *with* feedback is $\frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1} + 2^{n+1} - 3^n}{3^n - 2^{n+2} + 2^{k-1}}$ (see [5]). A tedious analysis would show that the value given in Fact 16 is at most twice as high. Thus, teaching monomials without feedback takes at most twice as long as with feedback.

Corollary 17. *The following problem OPT-CYCLIC-TEACHING-TIME is \mathcal{NP} -hard.*

Input: Concept class \mathcal{C} , concept $c^* \in \mathcal{C}$, rational number F .

Question: Is there a cyclic teacher with teaching time at most F ?

5.2 Greedy Teachers

We know from Lemma 4 that a greedy teacher is always successful. Moreover, in contrast to cyclic teachers, they allow a direct application of the optimality criterion. Thus we were able to prove their optimality in Fact 7. However, they can be arbitrarily far off the optimal teacher.

Fact 18. *For every $d > 1$ there is a class \mathcal{C} and a target c^* such that for all greedy teachers T , $E_T(c^*, \mathcal{C}) > d \cdot E(c^*, \mathcal{C})$.*

	x_1	x_2	x_3	y_1	\dots	\dots	\dots	y_n	
c^*	1	1	1	1	\dots	\dots	\dots	1	
c_1	0	1	1	1	\dots	\dots	\dots	1	
c_2	1	0	1	1	\dots	\dots	\dots	1	
c_3	1	1	0	1	\dots	\dots	\dots	1	
c_4	0	0	1	0	1	\dots	1	1	}
	\vdots	\vdots	\vdots	1	0	1	\dots	1	
	\vdots	\vdots	\vdots	\vdots	1	\ddots	\vdots	\vdots	
	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\ddots	1	
	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\ddots	1	
c_{3+n}	0	0	1	1	1	\dots	1	0	

Fig. 3. For growing n , the greedy teacher for c^* becomes arbitrarily worse than the optimal teacher. See Fact 18. The examples on the right are “dummy” examples.

Proof. Figure 3 shows a family of classes parameterized by n . We sketch the main steps of the proof.

(1) The cyclic teacher $((x_1, 1), (x_2, 1), (x_3, 1))$ has a teaching time of $(16 + 5n)/(4 + n) < 5$. Therefore $E(c^*, \mathcal{C}) < 5$. This can be shown using Lemma 15.

(2) Setting $n := \frac{3}{2}(3^{2\ell-1} - 3) + 1$ makes the greedy teacher be a cyclic teacher of the form $((x_1, 1), (x_2, 1)^\ell, (x_3, 1))$ or $((x_2, 1), (x_1, 1)^\ell, (x_3, 1))$ giving ℓ times x_1, x_2 before giving x_3 . The value ℓ becomes larger with growing n because, intuitively, the example x_3 becomes less attractive for the greedy teacher.

(3) For given ℓ the cyclic teacher has $\frac{-n+9^\ell(16+5n)+4\ell(1+3^{2\ell+1}+9^\ell n)}{2^{(-1+9^\ell)(4+n)}}$ as teaching time. The proof is again an application of Lemma 15.

(4) For n as in (2) the greedy learner is cyclic with $\ell = \log(7 + 2n)/\log(9)$ and according to (3) with a teaching time of

$$\frac{(56 + n(33 + 5n)) \log(3) + (22 + n(13 + 2n)) \log(7 + 2n)}{(3 + n)(4 + n) \log(9)}$$

which is not bounded from above and can be larger than 5 by any factor d . \square

In general there can be more than one greedy teacher for a given class and concept. It is \mathcal{NP} -hard to compute the teaching time of the optimal one.

Corollary 19. *The following problem OPT-GREEDY-TEACHING-TIME is \mathcal{NP} -hard.*

Input: Concept class \mathcal{C} , concept $c^* \in \mathcal{C}$, rational number F .

Question: Is there a greedy teacher with teaching time at most F ?

Conclusion. We presented a model for teaching randomized memoryless learners without feedback, characterized optimal learners, and analyzed the expected teaching time of certain classes. We showed the problem of determining the optimal teaching time to be \mathcal{NP} -hard and studied useful heuristics for teaching.

Acknowledgment. The authors are very grateful to the ALT 2006 PC members for their many valuable comments.

References

- [1] D. Angluin and M. Krikis. Teachers, learners and black boxes. In *Proc. 10th Ann. Conf. on Comput. Learning Theory*, pp. 285–297. ACM Press, New York, 1997.
- [2] D. Angluin and M. Krikis. Learning from different teachers. *Machine Learning*, 51(2):137–163, 2003.
- [3] M. Anthony, G. Brightwell, D. Cohen, and J. Shawe-Taylor. On exact specification by examples. In *Proc. 5th Ann. ACM Works. on Comput. Learning Theory*, pp. 311–318. ACM Press, New York, NY, 1992.
- [4] F. J. Balbach. Teaching Classes with High Teaching Dimension Using Few Examples. In *Learning Theory, 18th Ann. Conf. on Learning Theory, COLT 2005, Bertinoro, Italy, June 2005, Proc., LNAI 3559*, pp. 668–683, Springer, 2005.
- [5] F. J. Balbach and T. Zeugmann. Teaching randomized learners. In *Learning Theory, 19th Ann. Conf. on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proc., LNAI 4005*, pp. 229–243, Springer, 2006.
- [6] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Sci., 2005.
- [7] V. D. Blondel and V. Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory of Computing Systems*, 36(3):231–245, 2003.
- [8] 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.
- [9] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, San Francisco, 1979.
- [10] S. A. Goldman and M. J. Kearns. On the complexity of teaching. *J. of Comput. Syst. Sci.*, 50(1):20–31, 1995.
- [11] S. A. Goldman and H. D. Mathias. Teaching a smarter learner. *J. of Comput. Syst. Sci.*, 52(2):255–267, 1996.
- [12] S. A. Goldman, R. L. Rivest, and R. E. Schapire. Learning binary relations and total orders. *SIAM J. Comput.*, 22(5):1006–1034, 1993.
- [13] J. Jackson and A. Tomkins. A computational model of teaching. In *Proc. 5th Ann. ACM Works. on Comput. Learning Theory*, pp. 319–326. ACM Press, 1992.
- [14] 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.
- [15] C. Kuhlmann. On Teaching and Learning Intersection-Closed Concept Classes. In *Computat. Learning Theory, 4th European Conf., EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proc., LNAI 1572*, pp. 168–182, Springer, 1999.
- [16] H. Lee, R.A. Servedio, and A. Wan. DNF Are Teachable in the Average Case. In *Learning Theory, 19th Ann. Conf. on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proc., LNAI 4005*, pp. 214–228, Springer, 2006.
- [17] O. Madani, S. Hanks, and A. Condon. On the undecidability of probabilistic planning and infinite-horizon partially observable markov decision problems. In *Proc. 16th Nat. Conf. on Artificial Intelligence & 11th Conf. on Innovative Applications of Artificial Intelligence*, pp. 541–548, AAAI Press/MIT Press, 1999.
- [18] H. D. Mathias. A model of interactive teaching. *J. of Comput. Syst. Sci.*, 54(3):487–501, 1997.
- [19] S. D. Patek. On partially observed stochastic shortest path problems. In *Proc. of the 40-th IEEE Conf. on Decision and Control*, pp. 5050–5055, 2001.
- [20] A. Shinohara and S. Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.