Teaching Randomized Learners

Frank J. Balbach[1] 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, Sapporo 060-0814, Japan
thomas@ist.hokudai.ac.jp

Abstract. The present paper introduces a new model for teaching *randomized learners*. Our new model, though based on the classical teaching dimension model, allows to study the influence of various parameters such as the learner's memory size, its ability to provide or to not provide feedback, and the influence of the order in which examples are presented. Furthermore, within the new model it is possible to investigate new aspects of teaching like teaching from positive data only or teaching with inconsistent teachers.

Furthermore, we provide characterization theorems for teachability from positive data for both ordinary teachers and inconsistent teachers with and without feedback.

1 Introduction

A natural teaching model consists of a teacher giving examples to a set of students with the goal that all students eventually hypothesize a certain target concept. Typically the admissible students are deterministic learning algorithms and the teaching performance is measured with respect to the worst case student. In the present paper we modify this model by assuming a partly randomized student and by measuring teaching performance in an average case fashion.

Our model is based on the teaching model introduced independently, and in different forms, by Shinohara and Miyano [19], Goldman *et al.* [11], Goldman and Kearns [9] as well as Anthony *et al.* [5]. Here, a teacher has to give enough examples to uniquely identify the target concept among all concepts in a given class. Thus, the students are all deterministic consistent learning algorithms.

By varying the set of admissible learners, the influence of different properties of the learners on the teaching process can be studied. For example, learners with limited memory should be harder to teach, whereas learners that show their current hypothesis to the teacher should ease the teaching process.

Let us consider the concept class of all Boolean functions over $\{0,1\}^n$. To teach a concept to all consistent learning algorithms, the teacher must present all 2^n examples. Teaching a concept to all consistent learners that can memorize less than 2^n examples is impossible; there is always a learner with a consistent, but

wrong hypothesis. So teaching gets indeed harder, but in a rather abrupt way. Moreover, it does not matter whether or not the teacher knows the learner's hypothesis, since there are deterministic learners choosing their next hypothesis independently of their current one.

It seems that the worst case analysis style makes it impossible to investigate the influence of memory limitations or learner's feedback. A common remedy for this is to perform an average case analysis instead. In the present paper, we propose a rather radical approach, i.e., we replace the set of learners by a single one that is intended to represent an "average learner."

We achieve this goal by substituting the set of deterministic learners by a single randomized one. Basically, such a learner picks a hypothesis at random from all hypotheses consistent with the known examples. Teaching is successful as soon as the learner hypothesizes the target concept. For ensuring that the learner maintains this correct hypothesis, we additionally require the learner to be *conservative*, i.e., it can change its hypotheses only on examples that are inconsistent with its current hypothesis. The complexity of teaching is measured by the *expected* teaching time (cf. Section 2).

Next, we explain why this model should work. Intuitively, since at every round there is a chance to reach the target, the target will eventually be reached even if, for instance, the randomized learner can only memorize few examples. Moreover, the ability of the teacher to observe the learner's current hypothesis should be advantageous, since it enables the teacher to teach an inconsistent example in every round. Recall that only these examples can cause a hypothesis change. In Section 3, we show these intuitions to be valid.

Randomized learners show another phenomenon, too: The complexity of the teaching process now does not only depend on the examples, but also on the order in which they are given.

The randomized teaching model can be regarded as a *Markov Decision Process*. Such processes have been studied for several decades and we will make use of some results from this theory (cf. Subsection 2.3).

Sections 4 and 5 study teaching with and without feedback, respectively. Here, we focus on computing the optimal teaching times. In Sections 6 to 8 we study variations of our model: teaching from positive data, inconsistent teachers and another restriction on teachers. Theorems characterizing teachability within these model are shown.

Note that there are also other approaches to teaching. They differ from the one discussed here, since the learner is not given, but constructed to fit to the teacher. One such model is learning from good examples (cf. Freivalds *et al.* [8] and Jain *et al.* [13]). Jackson and Tomkins [12] as well as Goldman and Mathias [10] and Mathias [15] defined models of teacher/learner pairs. In their models, a kind of adversary disturbing the teaching process is necessary to avoid collusion between the teacher and the learner. Angluin and Kriķis' [3, 4] model prevents collusion by giving incompatible hypothesis spaces to teacher and learner.

2   Preliminaries

2.1   Notations

Set inclusion and proper set inclusion is denoted by "$\subseteq$" and "$\subset$," respectively. For numbers a, b with $a < b$ we write $[a, b]$ to denote the set $\{a, a+1, \ldots, b\}$ or $\{a, a+1, \ldots\}$ if $b = \infty$.

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 $\mathcal{X}(c) = \{(x, v) \,|\, v = 1 \iff x \in c\} \subset \mathcal{X}$. An example (x, v) is called *consistent* with c iff $(x, v) \in \mathcal{X}(c)$.

A *teaching set* for a concept $c \in \mathcal{C}$ with respect to $\mathcal{C}$ is a set S of examples such that c is the only concept in $\mathcal{C}$ consistent with S. The *teaching dimension $TD(c)$* is the size of the smallest teaching set. We set $TD(\mathcal{C}) := \max\{TD(c) \,|\, c \in \mathcal{C}\}$.

For any set S, we denote by S^* the set of all finite lists of elements from S, by S^m and $S^{\leq m}$ the set of all lists with length m and at most length m, respectively. The operator $\circ_\mu$ concatenates a list of length at most μ with a single element resulting in a list of length at most μ: $\langle x_1, \ldots, x_\ell \rangle \circ_\mu \langle y \rangle$ equals $\langle x_1, \ldots, x_\ell, y \rangle$ if $\ell < \mu$ and $\langle x_2, \ldots, x_\ell, y \rangle$ if $\ell = \mu$. We regard $\circ_\infty$ as the usual list concatenation. For a list $\boldsymbol{x}$ of examples, we set $\mathcal{C}(\boldsymbol{x}) = \{c \in \mathcal{C} \,|\, \boldsymbol{x} \text{ is consistent with } c\}$.

We denote by $\mathcal{M}_n$ the concept class of monomials over $\{0, 1\}^n$. 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. $\mathcal{D}_n$ is the set of all 2^n concepts over $[1, n]$. The singleton classes are defined as $\mathcal{S}_n = \{\{x\} \,|\, x \in [1, n]\}$.

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 memorizes this example and computes a new hypothesis based on its last hypothesis and the known examples.

The Learner. In a sense, consistency is a minimum requirement for a learner. We thus require our learners to be consistent with all examples they know. However, the hypothesis is chosen at *random* from all consistent ones.

The memory of our learners may be limited to $\mu \geq 1$ examples. If the memory is full and a new example arrives, the oldest example is erased. In other words, the memory works like a queue. Setting $\mu = \infty$ models unlimited memory.

The goal of teaching is making the learner to hypothesize the target *and to maintain it*. Consistency alone cannot guarantee this behavior if the memory is too small. In this case, there is more than one consistent hypothesis at every round and the learner would oscillate between them rather than maintaining a single one. To avoid this, *conservativeness* is required, i.e., the learner can change its hypothesis only when taught an example inconsistent with its current one.

To study the influence of the learners' feedback to the teacher, we distinguish between private and public output of the learner. The private output is the result of the calculation during a round (i.e., new memory content and hypothesis), the

public output is that part of the private one observable by the teacher. So, if the learner gives feedback, the teacher can observe in every round the complete hypothesis computed by the learner. If the learner does not give feedback, the teacher can observe nothing.

The following algorithm describes the behavior of the *μ-memory learner with/without feedback* (short: L_μ^+ / L_μ^-) during one round of the teaching process.

> *Input*: memory $\boldsymbol{x} \in \mathcal{X}^{\leq \mu}$, hypothesis $h \in \mathcal{C}$, example $z \in \mathcal{X}$.
> *Private Output*: memory $\boldsymbol{x}'$, hypothesis h'.
> *Public Output*: hypothesis h' / nothing.
> 1 $\boldsymbol{x}' := \boldsymbol{x} \circ_\mu \langle z \rangle$;
> 2 **if** $z \notin \mathcal{X}(h)$ **then** pick h' uniformly at random from $\mathcal{C}(\boldsymbol{x}')$;
> 3   **else** $h' := h$;

For making our results dependent on $\mathcal{C}$ alone, rather than on an arbitrary initial state of the learner, we stipulate a special initial hypothesis, called *init*. We assume every example inconsistent with *init*. Thus, *init* is left after the first example and cannot be reached again. Moreover, the initial memory is empty.

The Teacher. A teacher is an algorithm taking initially a given target concept c^* as input. Then, in each round, it receives the public output of the learner (if any) and outputs an example for c^*.

Definition 1. *Let $\mathcal{C}$ be a concept class and $c^* \in \mathcal{C}$. Let L_μ^σ be a learner ($\sigma \in \{+, -\}$) and T be a teacher and $(h_i)_{i \in \mathbb{N}}$ be the series of random variables for the hypothesis at round i. 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 teaching process is* successful *iff the success probability equals 1. A successful teaching process is called* finite *iff there is a t' such that* $\Pr\left[\bigcup_{1 \leq t \leq t'} G_t\right] = 1$, *otherwise it is called* infinite. *For a successful teaching process we define the* expected teaching time *as* $\mathbb{E}[T, L_\mu^\sigma, c^*, \mathcal{C}] := \sum_{t \geq 1} t \cdot \Pr[G_t]$.

Definition 2. *Let $\mathcal{C}$ be a concept class, $c^* \in \mathcal{C}$ and L_μ^σ a learner. We call c^** teachable *to L_μ^σ iff there is a successful teacher T. The* optimal teaching time *for c^* is*

$$E_\mu^\sigma(c^*) := \inf_T \mathbb{E}[T, L_\mu^\sigma, c^*, \mathcal{C}]$$

and the optimal teaching time for $\mathcal{C}$ is denoted by $E_\mu^\sigma(\mathcal{C}) := \max_{c \in \mathcal{C}} E_\mu^\sigma(c)$.

2.3   Markov Decision Processes

For an extensive treatment of this topic see Puterman [17] and Bertsekas [6]. A Markov Decision Process (MDP) is a probabilistic system whose state transitions can be influenced during the process by actions which incur costs. Formally, an

MDP consists of a finite set S of states, an initial state $s_0 \in S$, a finite set A of actions, a function $cost \colon S \times A \to \mathbb{R}$, and a function $p \colon S \times A \times S \to [0, 1]$; $cost(s, a)$ is the cost incurred by action a in state s; $p(s, a, s')$ is the probability for the MDP to change from state s to s' under action a.

In the *total cost infinite horizon* setting, the goal is to choose actions such that the expected total cost, when the MDP runs forever, is minimal. This makes sense only if there is a costless absorbing state $s^* \in S$. In the *finite horizon* setting the MDP is only run for finitely many rounds.

The actions chosen at each point in time are described by a *policy*. This is a function depending on the observed history of the MDP and the current state. A basic result says that there is a minimum-cost policy that is *stationary*, i.e., that depends only on the current state. A stationary policy $\pi \colon S \to A$ defines a Markov chain over S and for all $s \in S$ an expected time $H(s)$ to reach s^* from s. Such a policy is optimal iff for all $s \in S$:

$$\pi(s) \in \operatorname*{argmin}_{a \in A} \left(cost(s, a) + \sum_{s' \in S} p(s, a, s') \cdot H(s') \right) .$$

Finding optimal policies can be phrased as a linear programming problem and can thus be done in polynomial time in the representation size of the MDP.

3   Varying Memory Size and Feedback

As a simple example, we calculate the optimal teaching times for $\mathcal{D}_n$. To the learner L_μ^+ ($1 \le \mu \le n$) the teacher can give an example inconsistent with the current hypothesis in each round. For all such examples, there are $2^{n-\mu}$ hypotheses consistent with the μ examples in the learner's memory and learner chooses one of them. Therefore the probability of choosing the target concept is $2^{-(n-\mu)}$. Thus, considering that in the first $\mu - 1$ rounds the memory contains less then μ examples, $E_\mu^+(\mathcal{D}_n)$ is, for constant μ, asymptotically equal to $2^{n-\mu}$.

Clearly, teaching becomes faster with growing μ. Moreover the teaching speed increases continuously with μ and not abruptly as in the classical deterministic model. In particular, teaching is possible even with the smallest memory size ($\mu = 1$), although it takes very long (2^{n-1} rounds).

Teaching is more difficult when feedback is unavailable. In this situation the teacher can merely guess examples hoping that they are inconsistent with the current hypothesis. Roughly speaking, when teaching $\mathcal{D}_n$, the teacher needs two guesses on average to find such an example. Hence, the expected teaching time E_μ^- is about two times E_μ^+. Thus feedback doubles the teaching speed for $\mathcal{D}_n$.

Fact 3. *For all $\mathcal{C}$ and $\mu \in [1, \infty]$ all $c^* \in \mathcal{C}$ and $\sigma \in \{+, -\}$:*
    (1) $E_\mu^+(c^*) \le E_\mu^-(c^*)$,   (2) $E_\infty^\sigma(c^*) \le E_{\mu+1}^\sigma(c^*) \le E_\mu^\sigma(c^*)$.
Proper inequality holds for the concepts in $\mathcal{D}_n$.

Next, we relate the deterministic model (in terms of the teaching dimension) to the randomized model (in terms of the expected teaching time). Essentially, the teaching dimension can be used to lower bound the teaching time.

Lemma 4. *Let $\mathcal{C}$ be a class and let $c^* \in \mathcal{C}$ be a target. For all $\mu \in [1, TD(c^*)]$,*

$$E_\mu^-(c^*) \geq E_\mu^+(c^*) \geq \frac{\mu(\mu-1)}{2\,TD(c^*)} + TD(c^*) + 1 - \mu,$$

and for all $\mu > TD(c^*)$, $E_\mu^-(c^*) \geq E_\mu^+(c^*) \geq TD(c^*)/2$.

Proof. Let $k = TD(c^*)$ and $\mu \in [1, TD(c^*)]$. We show the statement for E_μ^+.

Claim: For i examples $z_0, \ldots, z_{i-1} \in \mathcal{X}(c^*)$: $|\mathcal{C}(z_0, \ldots, z_{i-1})| \geq k + 1 - i$.

Proof: Suppose $|\mathcal{C}(z_0, \ldots, z_{i-1})| \leq k - i$. Then c^* can be specified with $k-i-1$ examples with respect to $\mathcal{C}(z_0, \ldots, z_{i-1})$ (each example rules out at least one concept). Thus, c^* can be specified with $z_0, \ldots, z_{i-1}$ plus $k-i-1$ other examples, which amounts to $k - 1$ examples. This contradicts $TD(c^*) = k$.   □ Claim

Using the claim we upper bound the probabilities for reaching the target in round $i = 0, \ldots, \mu - 2$. After round i the learner knows $i + 1$ examples and therefore can choose between at least $k - i$ consistent hypotheses (see Claim). Thus, the probability for reaching c^* in round i is at most $p_i := \frac{1}{k-i}$. Beginning with round $\mu - 1$, the learner knows μ examples and has in each following round $i \geq \mu - 1$ a probability of at most $p_i = p_{\mu-1} = 1/(k+1-\mu)$ of reaching c^*.

No teaching process can be faster than one with the probabilities p_i described above. The expectation of such a process is

$$\sum_{i=0}^{\mu-2}(i+1) \cdot p_i \cdot \prod_{j=0}^{i-1}(1-p_j) \; + \; \sum_{i=\mu-1}^{\infty}(i+1) \cdot p_i \cdot \prod_{j=0}^{i-1}(1-p_j) \;. \qquad (1)$$

We start with the second sum in (1). Since $\prod_{j=0}^{\mu-2}(1-p_j) = \frac{k-\mu+1}{k}$ the product $\prod_{j=0}^{i-1}(1-p_j)$ in the this sum equals $\frac{k-\mu+1}{k} \cdot (1-p_{\mu-1})^{i-\mu+1}$. So, this sum is

$$\sum_{i=\mu-1}^{\infty}(i+1) \cdot p_{\mu-1} \cdot \tfrac{k-\mu+1}{k} \cdot (1-p_{\mu-1})^{i-\mu+1}$$

$$= \tfrac{k-\mu+1}{k} \cdot \sum_{i=0}^{\infty}(\mu+i) \cdot p_{\mu-1} \cdot (1-p_{\mu-1})^i$$

$$= \tfrac{k-\mu+1}{k} \cdot \left(\mu - 1 + \sum_{i=0}^{\infty}(i+1) \cdot p_{\mu-1} \cdot (1-p_{\mu-1})^i \right) \;.$$

The sum appearing in the last line is the expectation of the first success in a Bernoulli experiment with probability $p_{\mu-1}$ and thus equals $1/p_{\mu-1} = k - \mu + 1$. For the second sum in (1) we therefore get $\frac{k-\mu+1}{k} \cdot (\mu - 1 + k - \mu + 1) = k - \mu + 1$.

Calculating the first sum in (1) yields

$$\sum_{i=0}^{\mu-2}(i+1) \cdot \tfrac{1}{k-i} \cdot \prod_{j=0}^{i-1}\tfrac{k-j-1}{k-j} = \sum_{i=0}^{\mu-2}(i+1) \cdot \tfrac{1}{k-i} \cdot \tfrac{k-i}{k} = \frac{\mu(\mu-1)}{2k} \;.$$

Putting it together we obtain $\frac{\mu(\mu-1)}{2k} + k + 1 - \mu$ as the value of (1).

For $\mu > TD(c^*)$ the teaching process described above takes at most $TD(c^*)$ rounds. The lower bound is therefore the same as for $\mu = TD(c^*)$.   □

Input: Target $c^* \in \mathcal{M}_n$ represented by $\gamma_1 \ldots \gamma_n \in \{0, 1, *\}^n$;
        Hypothesis $h \in \mathcal{M}_n$ represented by $\eta_1 \ldots \eta_n \in \{0, 1, *\}^n$.
Output: Example $z \in \mathcal{X}(c^*)$.

1 **if** $h \supset c^*$ **output** $(\chi_1 \ldots \chi_n, 0)$ with $\chi_i = \begin{cases} \gamma_i & \text{if } \eta_i = \gamma_i \neq *, \\ 1 - \gamma_i & \text{if } i = \min\{j \mid \eta_j = * \neq \gamma_j\}, \\ 0 & \text{otherwise}; \end{cases}$

2 **else output** $(x, 1)$ with arbitrary $x \in c^*$.

Fig. 1. Optimal teacher for monomials and the learner L_1^+.

4 Learners with Feedback

4.1 Learners with 1-Memory

A teaching process involving L_1^+ can be modeled as an MDP with $S = \mathcal{C} \cup \{init\}$, $A = \mathcal{X}(c^*)$, $cost(h, z) = 1$ for $h \neq c^*$ and $cost(c^*, z) = 0$. Furthermore, for $h \neq c^*$, $p(h, z, h') = 1/|\mathcal{C}(z)|$ if $z \in \mathcal{X}(h') \setminus \mathcal{X}(h)$ and $p(h, z, h') = 0$ otherwise; finally $p(c^*, z, c^*) = 1$. The initial state is $init$ and the state c^* is costless and absorbing. The memory does not need to be part of the state, since the next hypothesis only depends on the newly given example which is modeled as an action.

An example $z \in \mathcal{X}(h)$ does not change the learner's state h and is therefore useless. An optimal teacher refrains from teaching such examples and thus we can derive the following criterion by using the results from Subsection 2.3.

Lemma 5. *Let $\mathcal{C}$ be a class over X and c^* be a target. A teacher $T : \mathcal{C} \cup \{init\} \to \mathcal{X}(c^*)$ with expectations $H : \mathcal{C} \cup \{init\} \to \mathbb{R}$ is optimal iff for all $h \in \mathcal{C} \cup \{init\}$:*

$$T(h) \in \operatorname*{argmin}_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{X}(h)}} \left(1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h') \right).$$

This criterion can be used to prove optimality for teaching algorithms.

Fact 6. *The teacher in Fig. 1 is an optimal teacher for $\mathcal{M}_n$ and the learner L_1^+.*

Proof. We define $H : \mathcal{C} \cup \{init\} \to \mathbb{R}$ as $H(h) = \frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1}}{3^n - 2^n + 2^{k-1}}$ for $h \supset c^*$ and $H(h) = \frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1} + 2^{n+1} - 3^n}{3^n - 2^n + 2^{k-1}}$ for all other h, including $init$.

It is possible (though tedious) to show that H describes the teaching times for T and that T and H satisfy the criterion of Lemma 5. $\square$

The teacher from Fig. 1 can be computed in linear time. It outputs a positive example whenever possible (i.e., when $h \not\supset c^*$). Since there are 2^n hypotheses consistent with a positive example and $3^n - 2^n$ consistent with a negative one, this means following a greedy strategy minimizing the number of consistent hypotheses for the learner to choose from, thus maximizing the probability for reaching c^* in the next step.

Such a greedy strategy seems sensible and is provably optimal in the case of $\mathcal{M}_n$. However, there are classes where no greedy teacher is optimal.

Definition 7. *Let $\mathcal{C}$ be a class over X and $c^* \in \mathcal{C}$. A stationary teacher $T\colon \mathcal{C} \cup \{init\} \to \mathcal{X}$ for c^* is called* greedy *iff for all $h \in \mathcal{C}\colon T(h) \in \operatorname{argmin}_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{C}(h)}} |\mathcal{C}(z)|$.*

Fact 8. *There is a class $\mathcal{C}$ and target c_0 such that no greedy teacher is optimal.*

Proof. Figure 2 displays such a concept class $\mathcal{C}$ and target c_0. T^* with teaching times H^* is an optimal teacher and T^g with H^g is the only greedy teacher. $\square$

h	x_1	x_2	x_3	x_4	x_5	$T^*(h)$	$H^*(h)$	$T^g(h)$	$H^g(h)$
init	–	–	–	–	–	x_1	$176/35 = 5.0285\ldots$	x_1	$2536/504 = 5.0317\ldots$
c_0	1	1	1	1	1	–	0	–	0
c_1	0	0	0	0	1	x_1	$176/35$	x_1	$2536/504$
c_2	0	0	0	1	1	x_1	$176/35$	x_1	$2536/504$
c_3	0	0	1	0	1	x_1	$176/35$	x_1	$2536/504$
c_4	0	0	1	1	1	x_1	$176/35$	x_1	$2536/504$
c_5	0	1	0	1	1	x_1	$176/35$	x_1	$2536/504$
c_6	0	1	1	0	1	x_1	$176/35$	x_1	$2536/504$
c_7	0	1	1	1	1	x_1	$176/35$	x_1	$2536/504$
c_8	1	0	0	1	0	x_2	$186/35$	x_2	$2680/504$
c_9	1	1	1	0	0	x_5	$189/35$	x_4	$2725/504$
c_{10}	1	1	1	1	0	x_5	$189/35$	x_5	$2723/504$

Fig. 2. Class with an optimal teacher T^* and a greedy teacher T^g that is not optimal. Both teachers teach c_0 to the learner L_1^+.

We now compare E_1^+ with other dimensions. The comparison of E_1^+ with the number MQ of membership queries (see Angluin [1]) is interesting because MQ and E_1^+ are both lower bounded by the teaching dimension.

Fact 9. (1) *For all $\mathcal{C}$ and $c^* \in \mathcal{C}\colon E_1^+(c^*) \geq TD(c^*)$.*
(2) *There is no function of TD upper bounding $E_1^+(c)$.*
(3) *There is no function of E_1^+ upper bounding MQ.*
(4) *There is a concept class $\mathcal{C}$ with $E_1^+(\mathcal{C}) > MQ(\mathcal{C})$.*
(5) *For all concept classes $\mathcal{C}$, $E_1^+(\mathcal{C}) \leq 2^{MQ(\mathcal{C})}$.*

Proof. (1) This follows from Lemma 4. (2) Let $\mathcal{C}_n = \{c \subseteq [1, n] \mid |c| = 2\}$. Then $TD(\mathcal{C}_n) = 2$, but $E_1^+(\mathcal{C}_n) = n - 1$ because the optimal teacher gives positive examples all the time and there are $n - 1$ hypotheses consistent to such an example. (3) $E_1^+(c) = 1$ for all $c \in \mathcal{S}_n$, but $MQ(\mathcal{S}_n) = n-1$. (4) $MQ(\mathcal{D}_n) = n$ and $E_1^+(\mathcal{D}) = 2^{n-1}$. (5) It is known (see e.g., Angluin [2]) that $\log |\mathcal{C}| \leq MQ(\mathcal{C})$ for all classes $\mathcal{C}$. Also, $E_1^+(\mathcal{C}) \leq |\mathcal{C}|$ because in every step the learner cannot choose from more than $|\mathcal{C}|$ hypotheses. Combining both inequalities yields the fact. $\square$

Roughly speaking, teaching L_1^+ can take arbitrarily longer than teaching in the classical model, but is still incomparable with membership query learning.

4.2 Learners with ∞-Memory

A straightforward MDP for teaching c^* to L_∞^+ has states $S = (\mathcal{C} \cup \{init\}) \times \mathcal{X}(c^*)^{\leq |X|}$. The number of states can be reduced because two states (h, m) and (h, m') with $\mathcal{C}(m) = \mathcal{C}(m')$ are equivalent from a teacher's perspective, but in general the size of the resulting MDP will not be polynomial in the size of the matrix representation of $\mathcal{C}$. Therefore, optimal teachers cannot be computed efficiently by the known general MDP algorithms.

A similar criterion as Lemma 5 can be stated for the L_∞^+ learner, too, and used to prove optimality of algorithms. We mention, without the technical proof, that a slight modification of the algorithm in Fig. 1 is optimal for L_∞^+ and $\mathcal{M}_n$.

That computing E_∞^+ is already a hard problem can be seen as follows. First, there is always a teacher that needs at most $TD(c^*)$ rounds by giving a minimal teaching set, hence $E_\infty^+(c^*) \leq TD(c^*)$. Second, it follows from Lemma 4 that $E_\infty^+(c^*) \geq TD(c^*)/2$. This means that every algorithm computing $E_\infty^+(c^*)$ also computes a factor 2 approximation of the teaching dimension.

As it has often been noted [19, 5, 9], the problem of computing the teaching dimension is essentially equivalent to the SET-COVER (or HITTING-SET) problem which is a difficult approximation problem. Raz and Safra [18] have shown that there is no polynomial time constant-factor approximation (unless $\mathcal{P} = \mathcal{NP}$). Moreover, Feige [7] proved that SET-COVER cannot be approximated better than within a logarithmic factor (unless $\mathcal{NP} \subseteq DTime(n^{\log \log n})$).

Corollary 10. *Unless $\mathcal{NP} \subseteq DTime(n^{\log \log n})$, computing E_∞^+ is $\mathcal{NP}$-hard and cannot be approximated with a factor of $(1 - \epsilon) \log |\mathcal{C}|$ for any $\epsilon > 0$.*

Fact 11. *Let $\mathcal{C}$ be a concept class and $c^* \in \mathcal{C}$ a target. Then there is a successful teacher for the learner L_∞^+ halting after at most $|X|$ rounds that is also optimal.*

Proof. Every given example is memorized forever. Hence, an optimal teacher never presents the same example twice and after at most $|X|$ rounds there is only one consistent hypothesis for the learner to choose from, namely c^*. □

As there is always a successful teacher giving at most $TD(c^*)$ examples, one could conjecture that there is also an optimal teacher teaching finitely within at most $TD(c^*)$ rounds. But this not the case.

Fact 12. *There is a concept class $\mathcal{C}$ and a concept $c^* \in \mathcal{C}$ such that all teachers teaching c^* to the learner L_∞^+ finitely within $TD(c^*)$ rounds are suboptimal.*

Proof. (Sketch) The concept class $\mathcal{C}$ and the concept c^* are defined by Figure 3. The teaching dimension of c^* is three and the only smallest teaching set $S := \{(x_1, 1), (x_2, 1), (x_3, 1)\}$. The only teachers finite after 3 rounds are those always giving an inconsistent example from S. Their expected teaching time is 2.6.

A teacher starting with $(x_4, 1)$ and then giving examples from S is not finite after three rounds, but has an expected teaching time of only 2.5 rounds. □

$$
\begin{array}{c c c c c c c}
 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\
\mathit{init}\text{:} & - & - & - & - & - & - \\
c^*\text{:} & 1 & 1 & 1 & 1 & 1 & 1 \\
 & 1 & 1 & 0 & 1 & 1 & 1 \\
 & 1 & 0 & 1 & 1 & 1 & 1 \\
 & 0 & 1 & 1 & 1 & 1 & 1 \\
 & 1 & 1 & 0 & 0 & 1 & 0 \\
 & 1 & 0 & 1 & 0 & 1 & 0 \\
 & 0 & 1 & 1 & 0 & 1 & 0 \\
 & 1 & 1 & 0 & 0 & 0 & 1 \\
 & 1 & 0 & 1 & 0 & 0 & 1 \\
 & 0 & 1 & 1 & 0 & 0 & 1 \\
 & 1 & 1 & 0 & 0 & 0 & 0 \\
 & 1 & 0 & 1 & 0 & 0 & 0 \\
 & 0 & 1 & 1 & 0 & 0 & 0 \\
\end{array}
$$

Fig. 3. Concept class and target whose optimal L_∞^+-teacher is not finite after $TD(c^*) = 3$ rounds. The optimal teacher starts with x_4 and is finite after 4 rounds (see Fact 12).

5 Learners without Feedback

The problem of finding the optimal cost in an MDP whose states cannot be observed is much harder than in an observable MDP. In general, it is not even decidable whether the optimal cost is below a given threshold (see Madani, Hanks, and Condon [14]). We know of no obvious algorithm to decide this problem in the special case of teaching 1-memory learners.

Teaching ∞-memory learners can be seen as a finite horizon unobservable MDP since any reasonable teacher presents a different example in every round and thus can stop after at most $|X|$ rounds. The decision problem for finite horizon unobservable MDPs is $\mathcal{NP}$-complete (Mundhenk *et. al.* [16]) and the inapproximability result of Corollary 10 holds for the feedbackless case as well, since $TD(c^*)/2 \le E_\infty^-(c^*) \le TD(c^*)$.

6 Teaching Positive Examples Only

The learnability of classes from positive data is a typical question in learning theory. Similar restrictions on the data can be posed in teaching models, too. In contrast to teaching with positive *and* negative data, where all classes are teachable, we now get classes that are not teachable. More precisely we have the following characterization for teachability with positive data.

Theorem 13. *Let $\mathcal{C}$ be a concept class and $c^* \in \mathcal{C}$ a target concept. Then for all learners L_μ^σ with $\mu \in [1, \infty]$, $\sigma \in \{+, -\}$: The concept c^* is teachable from positive data iff there is no $c \in \mathcal{C}$ with $c \supset c^*$.*

Proof. For the if part, assume there is no proper superset of c^* in the class. Then the set S^+ of all positive examples for c^* is a teaching set for c^*. Learners with ∞-memory can be taught by presenting S^+, since they remember all examples and

are always consistent. Learners with smaller memory can be taught by infinitely repeating S^+ in any order.

For the only-if part, assume there is a $c \in \mathcal{C}$ with $c \supset c^*$. Let $z = (x, 1) \in \mathcal{X}(c^*)$ be the first example taught. Then $c \in \mathcal{C}(z)$ and therefore there is a positive probability that the randomized learner picks c as first hypothesis. In this case, it is impossible to trigger any further mind changes by giving positive examples. Thus, with positive probability the number of examples is infinite, leading to an infinite expected number of examples. □

Theorem 13 also characterizes teaching with positive data in the classical teaching dimension model. If there is no $c \supset c^*$, the set of all positive examples of c^* is a teaching set, but if there is a $c \supset c^*$, then every set of positive examples for c^* is also consistent with c.

We have seen that teachability with positive data has a simple characterization. Things become a little more complicated when combined with inconsistent teachers discussed in the next section.

7   Inconsistent Teachers

Until now, teachers were required to always tell the truth, i.e., to provide examples $z \in \mathcal{X}(c^*)$. In reality it might sometimes be worthwhile to teach something which is, strictly speaking, not fully correct, but nevertheless helpful for the students. For example, human teachers sometimes oversimplify to give a clearer, yet slightly incorrect, view on the subject matter.

To model this we allow the teacher to present any example from $\mathcal{X} \times \{0, 1\}$, even inconsistent ones. One can see this as an analog to inconsistent learners in learning theory, as these learners also contradict something they actually know.

Clearly, teaching learners with ∞-memory becomes difficult after giving an inconsistent example because the target is not consistent with the memory contents any more. Even worse, there might be no consistent hypothesis available. However, the model can be adapted to this, e.g., by stipulating that a memorized example (x, v) can be "erased" by the example $(x, 1 - v)$, but here we will not pursue this further. We restrict ourselves to consider only the 1-memory learner.

We first look at inconsistent teachers in combination with teaching from positive data. In this case, for a target concept c, the only inconsistent examples allowed are of the form $(x, 1)$, where $x \notin c$. The class $\mathcal{C}_1$ in Figure 4 shows that, when only positive data are allowed, inconsistent teachers can teach concepts to L_1^+ that consistent teachers cannot. First, the teacher gives $(x_1, 1)$. If the learner guesses c^*, we are done. Otherwise, the learner must return c_1 and the teacher gives $(x_3, 1)$ which is *inconsistent* with c^*. Now, the learner has to guess c_2. Next, $(x_1, 1)$ is again given and the process is iterated until the learner returns c^*.

However, consistent teachers with both positive and negative data are more powerful as we show next.

Fact 14. *There is a class that cannot be taught to L_1^+ by an inconsistent teacher from positive data.*

	x_1 x_2 x_3	T
init:	$-$ $-$ $-$	x_1
c^*:	1 0 0	$-$
c_1:	1 1 0	x_3
c_2:	0 0 1	x_1

$\mathcal{C}_1$:

	x_1 x_2 x_3
init:	$-$ $-$ $-$
c^*:	0 1 0
c_1:	1 1 0
c_2:	0 1 1

$\mathcal{C}_2$:

Fig. 4. The class $\mathcal{C}_1$ can be taught to L_1^+ by the inconsistent positive-data teacher T, but cannot be taught by a consistent positive-data teacher (Theorem 13). The class $\mathcal{C}_2$ cannot be taught by an inconsistent positive-data teacher (Fact 14).

Proof. We show that $\mathcal{C}_2$ from Figure 4 is such a class. Let T be a teacher for L_1^+ mapping $\mathcal{C}_2 \cup \{init\}$ to $\{x_1, x_2, x_3\} \times \{1\}$. No matter what $T(init)$ is, the probability that the learner switches to c_1 or c_2 is positive. If the learner guesses c_1 (the c_2 case is analog), the teacher must teach $(x_3, 1)$, since all other examples are consistent with the current hypothesis c_1. But the only hypothesis consistent with $(x_3, 1)$ is c_2. Analogously, T must give $(x_1, 1)$ when the learner is in c_2, leading again to c_1. So, the probability that L_1^+ never reaches c^* is positive. □

Classes teachable by inconsistent teachers from positive data can be characterized. We associate a directed graph with the class $\mathcal{C}$. Define the graph $G(\mathcal{C}) = (V, A)$ by $V = \mathcal{C}$ and $A = \{(c, d) \,|\, d \setminus c \neq \emptyset\}$, i.e., there is an arc from c to d iff there is a positive example inconsistent with c but consistent with d.

Theorem 15. *Let $\mathcal{C}$ be a concept class and $G(\mathcal{C}) = (V, A)$ its associated graph. For the learner L_1^+ a concept $c^* \in \mathcal{C}$ is teachable by an inconsistent teacher from positive data iff for all $c \in V$ there is a path to c^* in $G(\mathcal{C})$.*

Proof. For the if part we have to describe a teacher. For each c let c' be a neighbor of c on a shortest path to c^*. Let T be such that for all c, $T(c)$ is consistent with c', but not with c. There is always such an example due to the definition of $G(\mathcal{C})$ and the reachability assumption.

Denote by $n = |\mathcal{C}|$ and by $p = 1/n$ the minimum probability for reaching c' when the learner receives $T(c)$ in state c. If the learner is in any state c, there is a probability of at least $p^n > 0$ for reaching c^* within the next n rounds by traversing the shortest path from c to c^*. Therefore, no matter in which state the learner is, the expected number of n-round blocks until reaching the target is at most $1/p^n$. Thus, the expected time to reach the target from any state, in particular from *init*, is at most $n/p^n < \infty$.

For the only-if part, let T be a teacher for $c^* \in \mathcal{C}$. Suppose there is a state c with no path to c^*. Then $c \supset c^*$ (otherwise $c^* \setminus c \neq \emptyset$ and $(c, c^*) \in A$). At some time, T must teach an example consistent with c^*, which is then also consistent with c. Hence, the probability for reaching c during the teaching process is positive. The graph $G(\mathcal{C})$ contains all transitions that are possible between the hypotheses by positive examples. Since c^* is not reachable from c in $G(\mathcal{C})$ there is no sequence of positive examples that can trigger hypothesis changes from c to c^*. Thus, the expected teaching time from c is infinite and hence the expected teaching time altogether. A contradiction to c^* being teachable by T. □

The criterion in Theorem 15 requires to check the reachability of a certain node from all other nodes in a directed graph. This problem is related to the REACHABILITY problem and also complete for the complexity class $\mathcal{NL}$.

While inconsistent teachers can teach classes to 1-memory learners with feedback from positive data that consistent teachers cannot teach to L_1^+ (cf. Figure 4), the situation changes if no feedback is available. That is, 1-memory learners without feedback can be taught the same classes by inconsistent teachers as by consistent teachers (cf. Theorem 13 and Theorem 16 below).

Theorem 16. *For the learner L_1^- a concept $c^* \in \mathcal{C}$ is teachable by an inconsistent teacher from positive data iff there is no $c \in \mathcal{C}$ with $c \supset c^*$.*

Proof. The if-direction follows from Theorem 13.

For the only-if part suppose that c^* is teachable by a teacher T and there is a c with $c \supset c^*$. Let $(z_i)_{i \in \mathbb{N}}$ be the series of examples taught.

Claim: T teaches inconsistent examples only finitely often.

Proof: Suppose T teaches an example $(x, 1) \notin \mathcal{X}(c^*)$ infinitely often. Without loss of generality we assume that there is a concept containing x (otherwise $(x, 1)$ would be useless and a teacher T' never giving this example would be successful, too). Whenever $(x, 1)$ is taught, the learner will not be in state c^* afterwards, i.e., there are infinitely many t such that $\Pr[h_t \neq c^*] = 1$. It follows that $\Pr[G_t] = \Pr[h_{t-1} \neq c^* \wedge \forall t' \geq t : h_t = c^*] = 0$ for all $t \geq 1$. This means that the success probability is zero, a contradiction. This proves the claim.   □ Claim

From the claim it follows that there is a t' such that $z_t \in \mathcal{X}(c^*)$ for all $t \geq t'$. We now show that $\Pr[h_{t'} = c^*] < 1$, i.e., it is uncertain whether the learner is in the target state. Suppose that $\Pr[h_{t'} = c^*] = 1$. Let $t \leq t'$ be minimal with $\Pr[h_t = c^*] = 1$. If $z_t \in \mathcal{X}(c^*)$ then z_t is consistent with c, too, and thus $\Pr[h_t = c^*] \leq 1/2$. If $z_t \notin \mathcal{X}(c^*)$, then $\Pr[h_t = c^*] = 0$, a contradiction.

Hence, the probability that the learner is not in the target state at time t' is positive. After t' only consistent examples are given. So there is a probability of at least $1/|\mathcal{C}|$ that the learner switches to c on the next example. As $c \supset c^*$ the target cannot be reached by positive examples any more. Thus, the success probability is less than one, a contradiction. This proves the only-if part.   □

8   Mind Change Forcing Teachers

In this section we deal again with consistent teachers. When teaching L_1^+ it is useless to provide an example consistent with the current hypothesis, since it does not change the state of L_1^+. In this situation the optimal teacher is necessarily "mind change forcing." But if we look at L_μ^+ ($\mu > 1$), it is not obvious that an optimal teacher has to force the learner to change its mind in every round until successful learning. While we could prove that for L_∞^+ an optimal teacher can be made "mind change forcing," it remains open whether a similar statement is true for L_μ^+ with $1 < \mu < \infty$.

Theorem 17. *Let $\mathcal{C}$ be a class and c^* be a target. Then there is an optimal teacher for L_∞^+ never giving an example consistent with the current hypothesis.*

Proof. Let T be a successful teacher that teaches a consistent example z_1 in a state (h, m) that is reached with positive probability and is not a target state. Then $z_1 \in \mathcal{X}(h) \cap \mathcal{X}(c^*)$ and $h \neq c^*$. We show that T is not optimal for this state and hence for the initial state. We do this by showing that there is another teacher T' giving an inconsistent example and being not worse than T.

After receiving z_1, the learner reaches $(h, m \cup \{z_1\})$ due to the conservativeness property. Then T teaches $z_1 = T(h, m \cup \{z_1\})$ leading to $(h, m \cup \{z_1, z_2\})$ and so on. Since $h \neq c^*$, T must eventually teach an example $z_k \notin \mathcal{C}(h)$. After teaching $z_1, \ldots, z_k$ the learner has either reached the target or assumes one of the hypotheses in $\mathcal{C}(m \cup \{z_1, \ldots, z_k\}) \setminus \{c^*\}$ with equal probability $p := 1/|\mathcal{C}(m \cup \{z_1, \ldots, z_k\})| - 1$. For the expected teaching time we have

$$H(h, m) = k + \sum_{h' \in \mathcal{C}(m \cup \{z_1, \ldots, z_k\})} p \cdot H(h', m \cup \{z_1, \ldots, z_k\})$$

The teacher T' teaches the same examples $z_1, \ldots, z_k$, but in different order, namely $z_k, z_1, \ldots, z_{k-1}$, that is with the inconsistent example first. Formally: $T'(h, m) = z_k$ and furthermore for all $i = 0, \ldots, k - 1$ and for all $h' \in \mathcal{C}(m \cup \{z_k, z_1, \ldots, z_i\})$: $T'(h', m \cup \{z_k, z_1, \ldots, z_i\}) = z_{i+1}$.

Beginning in (h, m) and being taught by T' for k rounds, the learner has either arrived at the target or assumes one of the hypotheses in $\mathcal{C}(m \cup \{z_1, \ldots, z_k\}) \setminus \{c^*\}$. Furthermore all these hypotheses are equally likely. This follows inductively from the fact that whenever a hypothesis change is triggered, say after $z_k, z_1 \ldots, z_i$, all hypotheses from $\mathcal{C}(m \cup \{z_k, z_1, \ldots, z_i\})$, and in particular all hypotheses from the subset $\mathcal{C}(m \cup \{z_1, \ldots, z_k\})$, are equally likely; no hypothesis is preferred. The probability p' for each of these hypotheses is at most $1/(|\mathcal{C}(m \cup \{z_1, \ldots, z_k\})| - 1) = p$. The expected teaching time under T' is

$$H'(h, m) \leq k + \sum_{h' \in \mathcal{C}(m \cup \{z_1, \ldots, z_k\})} p' \cdot H'(h', m \cup \{z_1, \ldots, z_k\})$$

$$= k + \sum_{h' \in \mathcal{C}(m \cup \{z_1, \ldots, z_k\})} p' \cdot H(h', m \cup \{z_1, \ldots, z_k\}) \leq H(h, m)$$

where the equality in the second line holds because T and T' are identical in the states $(h', m \cup \{z_1, \ldots, z_k\})$.

We have shown that T' is not worse than T and gives an inconsistent example in (h, m). By repeating the above argument the states in which T gives consistent examples can be moved to the "end" where they finally disappear.   □

9   Conclusions and Future Work

We have presented a model for teaching randomized learners based on the classical teaching dimension model. In our model, teachability depends, in a qualitatively plausible way, on the learner's memory size, on its ability to give feedback, and on the order of the examples taught. The model also allows to study learning theory like questions such as teaching from positive data only or teaching

by inconsistent teachers. Randomization also gives more flexibility in defining the learner's behavior by using certain *a priori* probability distributions over the hypotheses. So, one can define and study learners preferring simple hypotheses.

References

[1] D. Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.

[2] D. Angluin. Queries revisited. *Theoret. Comput. Sci.*, 313(2):175–194, 2004.

[3] D. Angluin and M. Kriķis. Teachers, learners and black boxes. In *Proc. of the 10th Annual Conference on Computational Learning Theory*, pages 285–297. ACM Press, New York, NY, 1997.

[4] D. Angluin and M. Kriķis. Learning from different teachers. *Machine Learning*, 51(2):137–163, 2003.

[5] M. Anthony, G. Brightwell, D. Cohen, and J. Shawe-Taylor. On exact specification by examples. In *Proc. of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 311–318. ACM Press, New York, NY, 1992.

[6] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific.

[7] U. Feige. A threshold of ln n for approximating set cover. *J. of the ACM*, 45(4):634–652, 1998.

[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] S. A. Goldman and M. J. Kearns. On the complexity of teaching. *J. Comput. Syst. Sci.*, 50(1):20–31, 1995.

[10] S. A. Goldman and H. D. Mathias. Teaching a smarter learner. *J. Comput. Syst. Sci.*, 52(2):255–267, 1996.

[11] S. A. Goldman, R. L. Rivest, and R. E. Schapire. Learning binary relations and total orders. *SIAM J. on Computing*, 22(5):1006–1034, Oct. 1993.

[12] J. Jackson and A. Tomkins. A computational model of teaching. In *Proc. of the 5th Annual ACM Workshop on Computational Learning Theory* pages 319–326. ACM Press, New York, NY, 1992.

[13] 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*, Lecture Notes in Artificial Intelligence Vol. 1316, pages 32–47, Springer, 1997.

[14] O. Madani, S. Hanks, and A. Condon. On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems. In *Proc. of the 16th National Conference on Artificial Intelligence and 11th Conference on Innovative Applications of Artificial Intelligence*, pages 541–548, AAAI Press/The MIT Press, 1999.

[15] H. D. Mathias. A model of interactive teaching. *J. Comput. Syst. Sci.*, 54(3):487–501, 1997.

[16] M. Mundhenk, J. Goldsmith, C. Lusena, and E. Allender. Complexity of finite-horizon Markov decision process problems. *J. of the ACM*, 47(4):681–720, 2000.

[17] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[18] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proc. of the 29th ACM Symposium on Theory of Computing*, pages 475–484, 1997.

[19] A. Shinohara and S. Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.