Teaching Randomized Learners with Feedback

Frank J. Balbach[a,2], Thomas Zeugmann[*,b,1]

[a] *Pattern Science AG, Freigericht, Germany*
[b] *Division of Computer Science, Hokkaido University, Sapporo 060-0814, Japan*

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 the learner's memory size and of the presence or absence of feedback. Moreover, in the new model the order in which examples are presented may influence the teaching process.

The resulting models are related to Markov decision processes, and characterizations of optimal teachers for memoryless learners with feedback and for learners with infinite memory and feedback are shown.

Furthermore, in the new model it is possible to investigate new aspects of teaching like teaching from positive data only or teaching with inconsistent teachers. Characterization theorems for teachability from positive data for both ordinary teachers and inconsistent teachers with and without feedback are provided.

Key words: Algorithmic teaching, randomized algorithms, computational complexity

1. Introduction

When preparing a lecture, a good teacher carefully selects *informative examples*. Additionally, a good teacher takes into account that students do not memorize everything previously taught. And usually we make a couple of assumptions about the learners. For example, They should neither be ignorant nor should they be lazy. Thus, it is only natural to ask whether or not such human behavior is at least partially reflected in some algorithmic learning and/or teaching models studied so far in the literature.

Learning concepts from examples has attracted considerable attention in learning theory and machine learning. Typically, a learner does not know much about the source of these examples. Usually the learner is required to learn from all such sources, regardless of their quality. This is even true for the query learning model introduced by

[*] Corresponding author

Email addresses: frank-balbach@gmx.de (Frank J. Balbach), thomas@ist.hokudai.ac.jp (Thomas Zeugmann)

[1] Supported by MEXT Grant-in-Aid for Scientific Research on Priority Areas under Grant No. 21013001.

[2] This research was done at the Institute for Theoretical Computer Science, University Lübeck, Germany

Angluin [1, 2], since the teacher or oracle, though answering truthfully, is assumed to behave adversarially whenever possible. Therefore, it was only natural to ask whether or not one can also model scenarios in which a helpful teacher is honestly interested in the learner's success.

Perhaps the first approach was proposed by Freivalds, Kinber, and Wiehagen [3, 4]. They developed a learning model in the inductive inference paradigm of identifying recursive functions in which the learner is provided with *good* examples chosen by an implicitly given teacher. Jain, Lange, and Nessel [5] adopted this model to learn recursively enumerable languages from *good* examples in the inductive inference paradigm.

The next step was to consider teaching as the natural counterpart of learning. Teaching has been modeled and investigated in various ways within algorithmic learning theory. However, the more classical models studied so far all follow one of two basically different approaches.

In the first approach, the goal is to find a teacher *and* a learner such that a given learning task can be carried out by them. Jackson and Tomkins [6] as well as Goldman and Mathias [7] and Mathias [8] 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. That is, when modeling teaching, a major problem consists in avoiding coding tricks. Though there is no generally accepted definition of coding tricks, it will be clear from our exposition that *no* form of coding tricks is used and thus no collusion occurs.

Angluin and Kriķis' [9, 10] model prevents collusion by giving incompatible hypothesis spaces to the teacher and the learner. This makes simple encoding of the target impossible.

In the second approach, a teacher has to be found that teaches *all* deterministic consistent learners. Here a learner is said to be *consistent* if its hypothesis correctly and completely reflects all examples received. This prevents collusion, since teaching happens the same way for all learners and cannot be tailored to a specific one. Goldman, Rivest, and Shapire [11] as well as Goldman and Kearns [12] substitute the adversarial teacher in the online learning model by a helpful one selecting good examples. They investigate how many mistakes a consistent learner can make in the worst case. In Shinohara and Miyano's [13] model the teacher produces a set of examples for the target concept such that it is the only consistent one in the concept class. The size of this set is the same as the worst case number of mistakes in the online model. This number is termed the *teaching dimension* of the target. Because of this similarity, from now on, we shall refer to both models as the *teaching dimension model* (abbr. TD model).

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.

A further difficulty of teaching in the TD model results from the fact that the teacher does not know anything about the learners besides them being consistent. In reality a teacher can benefit a lot from knowing the learners' behavior or their current hypotheses. It is therefore natural to ask how teaching can be improved if the teacher may observe the learners' hypotheses after each example. We refer to this scenario as teaching with *feedback*.

After translating this question into the TD model, one sees that there is no gain in the sample size at all. The current hypothesis of a consistent learner reveals nothing about its following hypothesis. Even if the teacher knew the hypothesis and provided a special example in response, he can only be sure that the learner's next hypothesis will be consistent. But this was already known to the teacher. So, in the TD model, feedback is useless.

There are also several other deficiencies in the teaching models studied so far. These deficiencies include that the order in which the teacher presents examples does not matter, and that teaching infinite concepts or infinite concept classes is severely limited.

Therefore, our goal has been to devise a teaching model that remedies the above mentioned flaws. In particular, our aim has been to develop a teaching model such that the following aspects do matter.

(1) The order in which the teacher presents the information should have an influence on the performance of the learner.
(2) Teaching should get harder when the memory size of the learners decreases, but it should not become impossible for small memory.
(3) Teaching should get easier when the learners give feedback to the teacher.
(4) Concepts that are more complex should be harder to teach.

In the present paper we propose a new teaching model that achieves all these goals (1) through (4). Our approach is rather radical. It is based on the observation that the worst case analysis style makes it impossible to investigate the influence of memory limitations or learner's feedback. Instead of following the more common remedy to perform an average case analysis (cf., e.g., [14, 15, 16, 17]), we replace the set of learners by a single one that is conceptually 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. To ensure that the learner maintains this correct hypothesis, we additionally require the learner to be *conservative*, i.e., it can change its hypothesis 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 (cf. Section 3).

The paper is organized as follows. In Section 2 the formal definition of our randomized teaching framework is provided. We continue with some examples already showing that the new models have the properties we aimed for (cf. Section 3).

Then we turn our attention to the main subject of the present paper, i.e., randomized learners with feedback (cf. Section 4). After explaining how our models can be regarded as Markov decision processes in Subsection 4.1, we discuss the simplest case, that is, memoryless learners with feedback and derive a characterization for *optimal teachers*. This characterization is then shown to be useful by applying it to the concept class of monomials (cf. Subsection 4.2).

In Subsection 4.3 we then continue with learners with infinite memory and show again a characterization of *optimal teachers*. However, computing the optimal teaching sets and times, respectively, turns out to be a difficult problem. Therefore, we also study its approximability and show that the optimal teaching time for a concept class $\mathcal{C}$ is hard to approximate within a factor of $\frac{1}{2}(1-\epsilon)\ln(|\mathcal{C}|-1)$ for any $\epsilon > 0$ under a standard complexity theoretic assumption. The hardness result concerning the approximability of the optimal teaching time is then extended to learners with infinite memory and without feedback.

Furthermore, we study two variations of our model, namely teaching from positive data and inconsistent teachers. Theorems characterizing the teachability within these models are shown in Section 5. Finally, in Section 6 we discuss the results obtained.

2. Preliminaries

We start by introducing the necessary notions and definitions. Let $\mathbb{N} = \{0, 1, \ldots\}$ denote the set of all natural numbers, and let $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$. By $\mathbb{R}$ we denote the set of all real numbers. For any set S we write $|S|$ and $\wp(S)$ to denote its cardinality and power set, respectively. Set inclusion and proper set inclusion is denoted by "$\subseteq$" and "$\subset$," respectively. For a set $S \subseteq \mathbb{N}$, we write $\min S$ and $\max S$ to denote the minimum and maximum of S, respectively, where, by convention, $\min \emptyset = \infty$ and $\max \emptyset = 0$. For any function $f\colon S \to \mathbb{R}$ over an arbitrary set S and any set $A \subseteq S$ we define

$$\operatorname*{argmin}_{x\in A} f(x) \quad = \quad \{x \mid x \in A,\ f(x) = \min\{f(x') \mid x' \in A\}\} \ .$$

If $\{f(x') \mid x' \in A\}$ has no minimum, then we set $\operatorname{argmin}_{x\in A} f(x) = \emptyset$. The definition of $\operatorname{argmax}_{x\in A} f(x)$ is analogous to the above.

For any set S, we denote by S^* the set of all finite sequences of elements from S. Let $m \in \mathbb{N}^+$; then we write S^m and $S^{\le m}$ to denote the set of all sequences with length m and at most length m, respectively. We use bold lowercase letters as identifiers for sequences. Elements forming a sequence are enclosed in angle brackets. The *empty sequence* is denoted by $\langle\ \rangle$ and the *length* of a sequence $\boldsymbol{s} \in S^*$ is denoted by $|\boldsymbol{s}|$. We refer to sequences of length 1 as *singleton sequences*. For the i-th element of a sequence $\boldsymbol{s}$ ($i = 1, \ldots, |\boldsymbol{s}|$) we write $\boldsymbol{s}[i]$. We use the symbol $\circ$ for the *concatenation* of sequences and the symbol $\circ_\mu$, where $\mu \in \mathbb{N}^+$, for a length-restricted concatenation with a singleton sequence, i.e.,

$$\langle x_1, \ldots, x_\ell\rangle \circ_\mu \langle y\rangle \quad = \quad \begin{cases} \langle x_1, \ldots, x_\ell, y\rangle \ , & \text{if } \ell < \mu \ ; \\ \langle x_{\ell-\mu+2}, \ldots, x_\ell, y\rangle \ , & \text{if } \ell \ge \mu \ . \end{cases}$$

To describe our teaching models we use mostly standard notations from algorithmic learning theory. We always assume a finite *learning domain* X. We refer to the elements of X as *instances*. Any subset $c \subseteq X$ is said to be a *concept*. Whenever appropriate, we identify a concept c with its characteristic function $c \colon X \to \{0,1\}$. A *concept class* is a set $\mathcal{C} \subseteq \wp(X)$ of concepts. As X is finite, every concept class is finite, too. A pair $(x,b) \in X \times \{0,1\}$ of an instance and a Boolean label is said to be an *example*. An example (x,b) is *positive* if $b=1$ and *negative* if $b=0$. The set of all examples is denoted by $\mathcal{X} = X \times \{0,1\}$. A set $S \subseteq \mathcal{X}$ of examples is also called *sample*. We denote the set of all examples for a concept c by $\mathcal{X}(c) = \{(x,c(x)) \mid x \in X\}$. An example (x,b) is called *consistent* with c iff $(x,b) \in \mathcal{X}(c)$. Let S be a sample and let $\mathcal{C}$ be a concept class; then we set $\mathcal{C}(S) = \{c \in \mathcal{C} \mid \text{all } z \in S \text{ are consistent with } c\}$. Furthermore, for a sequence $\boldsymbol{s}$ of examples, we set $\mathcal{C}(\boldsymbol{s}) = \{c \in \mathcal{C} \mid \text{all } \boldsymbol{s}[i] \text{ are consistent with } c, \text{ where } i = 1,\ldots,|\boldsymbol{s}|\}$.

A *teaching set* [12, 11] for a concept $c \in \mathcal{C}$ with respect to $\mathcal{C}$ is any sample S such that $\mathcal{C}(S) = \{c\}$. Note that teaching sets are also known as key [13], specifying set [18], discriminant [19], and witness set [20]. The *teaching dimension* of c with respect to $\mathcal{C}$ is defined as the size of c's smallest teaching set, i.e.,

$$\mathrm{TD}(c,\mathcal{C}) \quad = \quad \min\{|S| \ \mid \mathcal{C}(S) = \{c\}\} \ .$$

We simply write $\mathrm{TD}(c)$ if the concept class is clear from the context. The teaching dimension of the whole concept class $\mathcal{C}$ is defined as the maximum teaching dimension over all concepts, that is,

$$\mathrm{TD}(\mathcal{C}) \quad = \quad \max\{\mathrm{TD}(c,\mathcal{C}) \ \mid c \in \mathcal{C}\} \ .$$

To have some concept classes that allow us to exemplify certain effects, for $n \in \mathbb{N}^+$, we define $\mathcal{A}_n$ to be the class of *all* concepts over $\{1,\ldots,n\}$, i.e., $\mathcal{A}_n = \wp(\{1,\ldots,n\})$. The *co-singleton* concepts c_i over $\{1,\ldots,n\}$ are defined as $c_i = \{1,\ldots,n\} \setminus \{i\}$, $i = 1,\ldots,n$. We set $\mathcal{S}_n = \{c_i \mid i = 1,\ldots,n\} \cup \{\{1,\ldots,n\}\}$. Moreover, for $n \in \mathbb{N}^+$ we denote by $\mathcal{M}_n$ the concept class of all 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.

2.1. 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 memorized examples. The target and the hypotheses are taken from a concept class known to both the teacher and the learner.

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 the teacher is to teach the learner the target. Thus, the learner must eventually hypothesize the target *and 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 could 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.

So, we model the learner as an automaton with randomized state transitions. Whenever the learner has more than one hypothesis to choose between, it is supposed to pick one alternative uniformly at random. The goal is that the learner hypothesizes the target as quickly as possible. But now we do not measure the worst case time until this happens, rather we measure the *expected teaching time* of the learner.

To make 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 to be inconsistent with *init*. Thus, *init* is left after the first example and cannot be reached again. Moreover, the initial memory is empty. In other words, the randomized learners are randomized automata. Every state consists of a sequence $\boldsymbol{s} \in \mathcal{X}^*$ of memorized examples and a hypothesis $h \in \mathcal{C} \cup \{init\}$. The state space is thus $\mathcal{X}^* \times (\mathcal{C} \cup \{init\})$. More formally, we consider the following learners.

Definition 1. Let $\mathcal{C}$ be a concept class over X and let $\mu \in \mathbb{N} \cup \{\infty\}$. The following randomized algorithm is called the *randomized μ-memory learner using the hypothesis space* $\mathcal{C}$ and is denoted by $\mathcal{L}_{\mu,\mathcal{C}}$ (or by $\mathcal{L}_\mu$ if $\mathcal{C}$ is clear or not important):

Current state: Memory $\boldsymbol{s} \in \mathcal{X}^*$, hypothesis $h \in \mathcal{C} \cup \{init\}$.
Input: Example $z \in \mathcal{X}$.
Follow-up state: Memory $\boldsymbol{s}' \in \mathcal{X}^*$, hypothesis $h' \in \mathcal{C}$.

$\boldsymbol{s}' := \boldsymbol{s} \circ_\mu \langle z \rangle$;
if $z \notin \mathcal{X}(h)$ **then** choose h' uniformly at random from $\mathcal{C}(\boldsymbol{s}')$;
else $h' := h$.

Definition 1 implicitly defines the probabilities $p((\boldsymbol{s}, h), z, (\boldsymbol{s}', h'))$ of a state change from $(\boldsymbol{s}, h)$ to $(\boldsymbol{s}', h')$ on input $z \in \mathcal{X}$:

$$p((\boldsymbol{s}, h), z, (\boldsymbol{s}', h')) \quad = \quad \begin{cases} 1\ , & \text{if } z \in \mathcal{X}(h) \ \wedge \ \boldsymbol{s}' = \boldsymbol{s} \circ_\mu \langle z \rangle \ \wedge \ h = h'\ ; \\ 1/|\mathcal{C}(\boldsymbol{s}')|\ , & \text{if } z \notin \mathcal{X}(h) \ \wedge \ \boldsymbol{s}' = \boldsymbol{s} \circ_\mu \langle z \rangle \ \wedge \ h' \in \mathcal{C}(\boldsymbol{s}')\ ; \\ 0\ , & \text{otherwise}\ . \end{cases} \quad (1)$$

The Teacher. A teacher is an algorithm taking initially a given target concept c^* as input. In the *presence of feedback*, in each round it receives the follow-up state of the learner and outputs an example for c^*. Thus, in this case the teacher is a function

$$T\colon \mathcal{X}^* \times (\mathcal{C} \cup \{init\}) \to \mathcal{X}(c^*)\ .$$

In the *absence of feedback*, in each round it receives *nothing* and it just outputs an example for c^*. Hence, now the teacher is a function

$$T\colon \mathbb{N} \to \mathcal{X}(c^*)\ .$$

A learner $\mathcal{L}_{\mu,\mathcal{C}}$ and a teacher determine a teaching process. The state of the process in a round $t \in \mathbb{N}$ is described by the probability distribution over the learner's state space that specifies for each state the probability of the learner being in this state in round t. We denote this probability distribution by

$$\delta_T^{(t)} \colon \mathcal{X}^* \times (\mathcal{C} \cup \{init\}) \to [0,1] \ .$$

The initial distribution is $\delta^{(init)}$ with $\delta^{(init)}(\langle\ \rangle, init) = 1$, since initially the learner hypothesizes *init* and has an empty memory.

First, we consider the teaching process involving a teacher T *without feedback* and the learner $\mathcal{L}_{\mu,\mathcal{C}}$. Then the probability distributions evolve as follows. Let $\delta_T^{(t)}$ be the distribution in round t and let $z = T(t)$ be the example given in round t. Then for every state $(\boldsymbol{s}, h)$ the definition of $\mathcal{L}_{\mu,\mathcal{C}}$ implies a distribution over the follow-up states. The distribution $\delta_T^{(t+1)}$ for round $t+1$ is then the weighted sum over all the distributions for the single states of the learner. Thus, more formally we arrive at $\delta_T^{(0)} = \delta^{(init)}$, and for all $t \geq 0$:

$$\delta_T^{(t+1)}(\boldsymbol{s'}, h') \quad = \sum_{\substack{(\boldsymbol{s},h) \\ \in \mathcal{X}^* \times (\mathcal{C} \cup \{init\})}} \delta_T^{(t)}(\boldsymbol{s}, h) \cdot p((\boldsymbol{s}, h), T(t), (\boldsymbol{s'}, h')) \ . \tag{2}$$

Next, we look at the teaching process involving a teacher T *with feedback* and the learner $\mathcal{L}_{\mu,\mathcal{C}}$. Now, we have $\delta_T^{(0)} = \delta^{(init)}$, and for all $t \geq 0$:

$$\delta_T^{(t+1)}(\boldsymbol{s'}, h') \quad = \sum_{\substack{(\boldsymbol{s},h) \\ \in \mathcal{X}^* \times (\mathcal{C} \cup \{init\})}} \delta_T^{(t)}(\boldsymbol{s}, h) \cdot p((\boldsymbol{s}, h), T(\boldsymbol{s}, h), (\boldsymbol{s'}, h')) \ , \tag{3}$$

which is similar to Equation (2) except that $T(t)$ is replaced by $T(\boldsymbol{s}, h)$.

Since we are mostly interested in the probability for certain hypotheses, as opposed to the memory, we define as shortcut:

$$\delta_T^{(t)}(c) = \sum_{\boldsymbol{s} \in \mathcal{X}^*} \delta_T^{(t)}(\boldsymbol{s}, c) \ . \tag{4}$$

We distinguish two teaching success variants: finite and in the limit. Finite teaching success means that after finitely many rounds the probability of having reached the target is 1. Teaching success in the limit means that the probability of reaching the target converges to 1.

Definition 2. Let $\mathcal{C}$ be a concept class, $c^* \in \mathcal{C}$ be a target concept, and $\mu \in \mathbb{N} \cup \{\infty\}$. Furthermore, let T be a teacher and let $\left(\delta_T^{(t)}\right)_{t \in \mathbb{N}}$ be the series of probability distributions over states of $\mathcal{L}_{\mu,\mathcal{C}}$. The *success probability* of T is then

$$\lim_{t \to \infty} \delta_T^{(t)}(c^*) \ .$$

A teacher is *successful* iff its success probability equals 1. A successful teacher is called *finitely successful* iff there is a t such that $\delta_T^{(t)}(c^*) = 1$, otherwise it is called *successful in the limit.* For a successful teacher we define the *expected teaching time* as

$$\mathbb{E}[T, \mathcal{L}_{\mu,\mathcal{C}}, c^*] \quad = \quad \sum_{t \geq 1} t \cdot \left(\delta_T^{(t)}(c^*) - \delta_T^{(t-1)}(c^*)\right) \ .$$

Note that the expected teaching time does not need to be finite, even if the teacher is successful. Moreover, the limit $\lim_{t\to\infty} \delta_T^{(t)}(c^*)$ exists for every teacher, since $\delta_T^{(t)}(c^*)$ is monotonically increasing due to the conservativeness of the learner. The teachability of a concept is then measured by the minimal expected teaching time over all teachers.

Definition 3. Let $\mathcal{C}$ be a concept class, $c^* \in \mathcal{C}$ and $\mu \in \mathbb{N} \cup \{\infty\}$. The *optimal teaching time for teaching c^* with feedback to $\mathcal{L}_{\mu,\mathcal{C}}$* is

$$E_\mu^+(c^*, \mathcal{C}) \quad = \quad \inf_T \mathbb{E}[T, \mathcal{L}_{\mu,\mathcal{C}}, c^*]$$

where T ranges over all teachers $T\colon \mathcal{X}^* \times (\mathcal{C} \cup \{init\}) \to \mathcal{X}(c^*)$. The *optimal teaching time for teaching c^* without feedback to $\mathcal{L}_{\mu,\mathcal{C}}$* is

$$E_\mu^-(c^*, \mathcal{C}) \quad = \quad \inf_T \mathbb{E}[T, \mathcal{L}_{\mu,\mathcal{C}}, c^*] \ ,$$

where T ranges over all teachers $T\colon \mathbb{N} \to \mathcal{X}(c^*)$. For a class $\mathcal{C}$ we set $E_\mu^-(\mathcal{C}) = \max\{E_\mu^-(c, \mathcal{C}) \ \mid c \in \mathcal{C}\}$ and $E_\mu^+(\mathcal{C}) = \max\{E_\mu^+(c, \mathcal{C}) \ \mid c \in \mathcal{C}\}$.

If the concept class is clear, we may write $E_\mu^+(c)$ instead of $E_\mu^+(c, \mathcal{C})$ and $E_\mu^-(c)$ for $E_\mu^-(c, \mathcal{C})$. Note that, since we consider only finite concept classes, the infimum in the definition of $E_\mu^+(c)$ can be replaced by the minimum, because in this case there are only finitely many teachers with feedback.

We finish this subsection with some simple facts about the notions of teachability and the teaching times just defined.

Fact 1. *Let $\mathcal{C}$ be a concept class and let $\mu \in \mathbb{N}^+$. A concept c^* is teachable to $\mathcal{L}_{\mu,\mathcal{C}}$ finitely (with or without feedback) if and only if* $\mathrm{TD}(c^*, \mathcal{C}) \le \mu$.

Fact 2. *For all $\mathcal{C}$ and $\mu \in \mathbb{N} \cup \{\infty\}$ all $c^* \in \mathcal{C}$ and $\alpha \in \{+, -\}$ we have:*

(1) $E_\mu^+(c^*, \mathcal{C}) \le E_\mu^-(c^*, \mathcal{C})$,
(2) $E_\infty^\alpha(c^*, \mathcal{C}) \le E_{\mu+1}^\alpha(c^*, \mathcal{C}) \le E_\mu^\alpha(c^*, \mathcal{C})$.

Proper inequality holds for $\mathcal{C} = \mathcal{A}_n$ with $n \ge \mu + 3$.

3. Varying Feedback, Memory Size and the Order of Examples

We start by calculating the teaching times for the concept $\{1, \ldots, n\} \in \mathcal{S}_n$ for varying memory size and feedback in order to show that feedback and memory size have a somewhat realistic influence on the duration of teaching. First, we consider a teacher *with* feedback and the learner $\mathcal{L}_1$ using $\mathcal{S}_n$. Our teacher now gives an example inconsistent with the current hypothesis in every round until $\mathcal{L}_1$ reaches the target $\{1, \ldots, n\}$. The probability of reaching the target is $1/n$ in each round. Therefore the expected number of rounds until the target is reached is n. This teacher is optimal, because it is basically the only one. Giving an example consistent with the current hypothesis would not change the learner's state and would therefore be useless. Thus $E_1^+(\{1, \ldots, n\}, \mathcal{S}_n) = n$.

Generalizing this teacher to $1 < \mu \le n$ is easy. A small problem for calculating the expected teaching time is that the learner's memory needs some rounds to fill. More precisely, in the first round the probability of reaching the target is $1/n$, in the second round $1/(n-1)$ and in round $i \leq \mu$ it is $1/(n-i)$. Beginning with round μ the probability remains constant $1/(n-\mu)$. Thus on entering the μ-th round, the expected number of remaining rounds is $n-\mu$. Putting it all together and simplifying the expression, we get the following formula for the expected number of rounds until $\mathcal{L}_\mu$ ($1 \leq \mu \leq n$) reaches $\{1,\dots,n\}$: $\frac{\mu(\mu-1)}{2n}+n-\mu+1$. For memory sizes greater than n the teaching time improves no further. So we obtain the teaching time for $\mu = \infty$ by setting $\mu = n$. Again, also for $\mu > 1$ this is essentially the only teacher and its teaching time is therefore optimal:

$$E^{+}_{\mu}(\{1,\dots,n\},\mathcal{S}_n) \quad = \quad \frac{\mu(\mu-1)}{2n} + n - \mu + 1 \; . \tag{5}$$

Teaching is more difficult without feedback. In this situation the teacher can merely guess examples hoping that they are inconsistent with the current hypothesis. If a consistent example is presented no hypothesis change is possible, since the learner is conservative. Rather than using a provably optimal teacher, we use a "reasonable" teacher whose optimality for the special case $\mu = 1$ can be shown (cf. [21]).

In particular, giving one example a second time within an interval of μ rounds will certainly not trigger a hypothesis change. Therefore, it seems to be a good strategy to put a maximum length interval between two occurrences of the same example. This is achieved by the "reasonable" teacher T^- that gives all n examples for $\{1,\dots,n\}$ in the canonical order in an infinite loop, that is, $T^-(i) = (1 + i \bmod n, 1)$ for all $i \in \mathbb{N}$. The analysis of T^- is a bit more complicated and we sum it up in the following theorem, where we use H_n to denote the nth Harmonic number, i.e., $\mathrm{H}_n = \sum_{i=1}^{n} \frac{1}{i}$.

Theorem 3. *Let T^- be a teacher for $c^* = \{1,\dots,n\} \in \mathcal{S}_n$ with $T^-(i) = (1+i \bmod n, 1)$ for all $i \in \mathbb{N}$. Then*

$$\mathbb{E}[T^-,\mathcal{L}_{1,\mathcal{S}_n},c^*] \quad = \quad 1 + \frac{n(n-1)}{2}$$

and for $\mu \in \{2,\dots,n\}$:

$$\mathbb{E}[T^-,\mathcal{L}_{\mu,\mathcal{S}_n},c^*] \quad = \quad \frac{(n-\mu+1)(n+\mu)}{2} - H_n + H_{n-\mu+1} \; .$$

Proof. First, the case $\mu = 1$ is considered. To simplify notation we use F to denote the sought expectation. We derive some properties of F allowing us to find a formula for it.

Suppose the learner conjectures hypothesis c_i ($1 \leq i \leq n$) and the teacher T^- presents example $(i,1)$ next. The resulting probability distribution is the same as after the regular first example: all hypotheses except c_i have a probability of $1/n$ and the teacher presents $(i,1)$ only after all other examples have been presented. Therefore, the expected number of rounds to reach c^* for this learner is also F.

Now suppose the learner conjectures hypothesis c_i and the teacher T^- presents example $(i-1,1)$ next. This will not change the learners hypothesis because c_i is consistent with $(i-1,1)$. Only in the next round, in which T^- gives example $(i,1)$, the hypothesis changes. Moreover this hypothesis change is the same as before; it only happens one round later. Hence, the expectation for this learner is $F+1$. In general, if it takes $\ell \in \{0,\dots,n-2\}$ rounds before the next inconsistent example arrives then the expectation is $F+\ell$. Of course, starting in the target concept c^* has an expectation of 0.

Now let us go back to our "real" teaching process in which the learner starts in *init*. After the first example, the learner assumes hypothesis c_1 with probability 0 and all other hypotheses $c^*, c_2, \ldots, c_n$ with probability $1/n$. This means that with probability $1/n$ the learner is in a state, namely c_2, in which the next example triggers a hypothesis change. More generally, the learner is with probability $1/n$ in a state in which after $\ell = 0, \ldots, n-2$ examples a hypothesis change is triggered (the states are $c_2, c_3, \ldots, c_n$). The expected number F of rounds is thus composed of $n-1$ individual expectations, each of which is to be weighted by $1/n$. This yields

$$F \quad = \quad 1 + \sum_{\ell=0}^{n-2} \frac{1}{n} \cdot (F + \ell) \tag{6}$$

which is a linear equation with one variable whose solution is

$$F \quad = \quad 1 + \frac{n(n-1)}{2} \ .$$

Next we consider the case $\mu > 1$. Again, the situation is a bit more complicated, since it takes μ rounds until the memory is filled. We consider this initial phase later and focus first on the situation in which the memory already contains μ examples.

The arguments are similar to the above for the case $\mu = 1$. Suppose a learner conjectures hypothesis c_i and memorizes the last μ examples $\langle (i-\mu, 1), \ldots, (i-1, 1) \rangle$. This means example $(i, 1)$ comes next. We denote the expected number of rounds for this learner to reach the target c^* by F'. After the example $(i, 1)$ is given, the learner is with a probability of $1/(n-\mu+1)$ in each hypothesis consistent with the new memory $\langle (i-\mu+1, 1), \ldots, (i, 1) \rangle$.

More generally, suppose the learner assumes $c_{i+\ell}$ for some $\ell > 0$ and the memory is the same as before. Then it takes $1 + \ell$ rounds until a hypothesis change happens. The situation reached afterwards is essentially the same as in the special case $\ell = 0$ just discussed: the learner is with a probability of $1/(n-\mu+1)$ in each hypothesis consistent with the new memory. Of course, the consistent hypotheses are different for different ℓ. The point, however, is that there is always a hypothesis that is inconsistent with the next example, one that is consistent with the next but inconsistent with the example after the next and so on. In effect the expectation for a learner starting in $c_{i+\ell}$ and receiving example $(i, 1)$ next is $F' + \ell$.

Our observations above allow us to state a formula for F' similar to the case $\mu = 1$:

$$F' \quad = \quad 1 + \sum_{\ell=0}^{n-\mu-1} \frac{1}{n-\mu+1} (F' + \ell) \ . \tag{7}$$

Note that for $\mu = 1$ we get Equation (6). The solution of Equation (7) is

$$F' = 1 + \frac{(n-\mu)(n-\mu+1)}{2} \ . \tag{8}$$

Now, if n is large compared to μ the initial μ rounds can be neglected and Equation (8) gives a good approximation for the true expectation. We now derive the exact values.

The probability of reaching the target in the first round is $1/n$. Otherwise after $i \leq \mu$ rounds the learner knows the examples $(1, 1), \ldots, (i, 1)$ and hypothesizes $c^*, c_{i+1}, \ldots, c_n$

with equal probability, namely $p_i = 1/(n-i+1)$. It follows that the probability of reaching the target c^* in round $i > 1$ is $p_i - p_{i-1} = \frac{1}{(n-i+1)(n-i+2)}$. Calculating the expectation would therefore begin like this:

$$\frac{1}{n} \cdot 1 + \sum_{i=2}^{\mu} \frac{1}{(n-i+1)(n-i+2)} \cdot i \; . \tag{9}$$

Computing the "in-the-target-probabilities" p_i for $i \geq \mu$ is more difficult, but we can use the values for F' instead.

After μ examples have been given, the learner memorizes $(1,1), \ldots, (\mu, 1)$ and hypothesizes $c^*, c_{\mu+1}, \ldots, c_n$ with probability $1/(n-\mu+1)$ each. Then for all $\ell \in \{0, \ldots, n-\mu-1\}$ there is a hypothesis, namely $c_{\mu+\ell}$, that is inconsistent only with the example given ℓ rounds later. Thus, the expected number of rounds to reach the target from this probability distribution is

$$\sum_{\ell=0}^{n-\mu-1} \frac{1}{n-\mu+1} \cdot (F' + \ell) \; .$$

But this probability distribution is reached after μ rounds. Thus the expectations have to be considered higher by μ. Therefore we have to add to Equation (9) the expression:

$$\sum_{\ell=0}^{n-\mu-1} \frac{1}{n-\mu+1} \cdot (F' + \ell + \mu) \quad = \quad \frac{(n-\mu)(1 + (n-\mu)^2 + 2n)}{2(n-\mu+1)}$$

which yields for the sought expectation

$$\frac{1}{n} + \sum_{i=2}^{\mu} \frac{i}{(n-i+1)(n-i+2)} + \frac{(n-\mu)(1 + (n-\mu)^2 + 2n)}{2(n-\mu+1)} \; . \tag{10}$$

Next, we simplify the sum in (10) as follows.

$$\begin{aligned}
\frac{1}{n} + \sum_{i=2}^{\mu} \frac{i}{(n-i+1)(n-i+2)} &= \frac{1}{n} + \sum_{i=2}^{\mu} \left(\frac{1}{n-i+1} - \frac{1}{n-i+2} \right) \cdot i \\
= \quad \frac{1}{n} - \frac{2}{n} + \sum_{i=2}^{\mu-1} \frac{i+1}{n-i} - \sum_{i=2}^{\mu-2} \frac{i+2}{n-i} &= \frac{\mu}{n-\mu+1} - \sum_{i=2}^{\mu-2} \frac{1}{n-i} \\
= \quad \frac{\mu}{n-\mu+1} - \mathrm{H}_n + \mathrm{H}_{n-\mu+1} \; . &
\end{aligned}$$

All what is left is to add the right expression of (10) to the term just obtained. Then we factor out $n - \mu + 1$ and obtain

$$\begin{aligned}
\mathbb{E}[T^-, \mathcal{L}_{\mu, \mathcal{S}_n}, c^*] &= \frac{\mu}{n-\mu+1} + \frac{(n-\mu)(1 + (n-\mu)^2 + 2n)}{2(n-\mu+1)} - \mathrm{H}_n + \mathrm{H}_{n-\mu+1} \\
&= \frac{(n-\mu+1)(n+\mu)}{2} - \mathrm{H}_n + \mathrm{H}_{n-\mu+1}
\end{aligned}$$

as claimed. $\square$

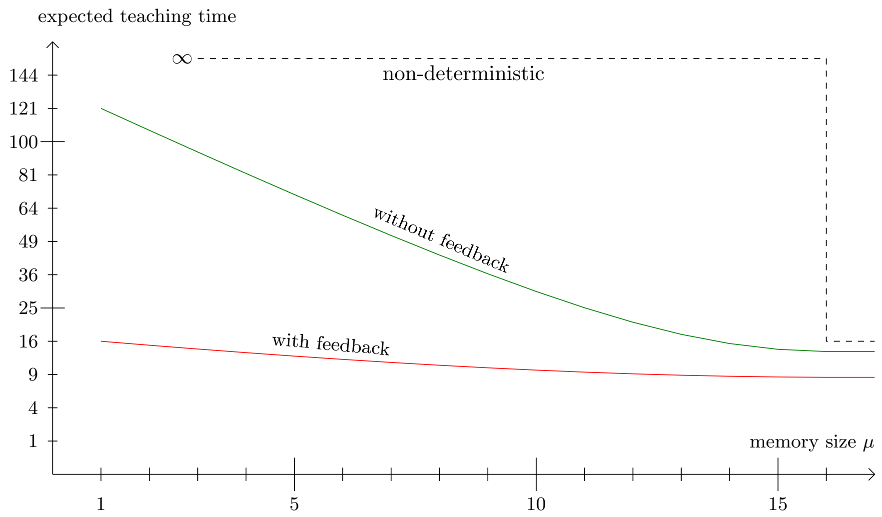

Figure 1: Influence of feedback and memory size on the expected teaching time. The concept $\{1, \ldots, 16\} \in \mathcal{S}_{16}$ is taught to the randomized learners $\mathcal{L}_\mu$ with and without feedback. The values for $\mu = 1$ and those for "with feedback" are the optimal teaching times. The values for $1 < \mu \leq 16$ without feedback are based on a reasonable, supposedly optimal, teacher. In contrast, teaching is impossible in the TD model unless the memory size is at least 16.

Comparing Equation (5) with Theorem 3, we see that for teaching with feedback the expected teaching time is $\Theta(n)$ while for teaching without feedback it is $\Theta(n^2)$.

As an illustration, all teaching times for $n = 16$ and $\mu = 1, \ldots, 16$ are shown in Figure 1. Clearly, teaching becomes faster with growing μ. Moreover the teaching speed increases continuously with μ and not abruptly as in the TD model. In particular, teaching is possible even with the smallest memory size ($\mu = 1$).

In general, the influence of the memory size varies between concept classes. At one end of the spectrum there is the concept class of all singleton concepts over $X = \{1, \ldots, n\}$. Here all concepts have a teaching dimension of one, and increasing the memory size beyond $\mu = 1$ will not improve the expected teaching time.

On the other end of the spectrum, there is the class of *all* concepts over $X = \{1, \ldots, n\}$, in which all concepts have a teaching dimension of 2^n. For a given target there is, up to symmetry, only one teacher with feedback giving an inconsistent example in every round. This teacher is therefore optimal. After $\mu < n$ rounds there are always $2^{n-\mu}$ hypotheses consistent with the memory, resulting in a probability of $2^{\mu-n}$ of reaching the target in any given round. Thus, for small μ, the exptected teaching time is approximately $2^{n-\mu}$. This shows that increasing the memory size by one roughly halves the teaching time.

In order to illustrate the influence of the order of examples in the randomized teaching model, we have calculated a numerical example. Figure 2 shows three teachers teaching the monomial $v_3 \wedge v_4$ without feedback to $\mathcal{L}_1$ using the hypothesis space $\mathcal{M}_4$. All teachers use the same four examples from a minimum teaching set. Every teacher, however, arranges these examples into a different sequence and teaches this sequence in an infinite loop.

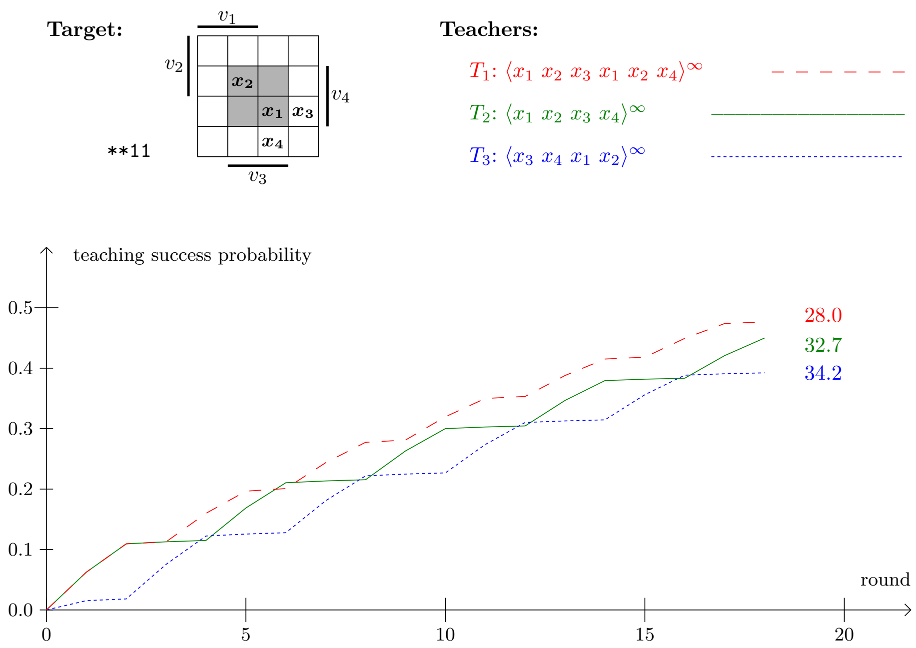

Figure 2: Influence of the order of examples on the teaching succes probability and expected time. The learner $\mathcal{L}_1$ using $\mathcal{M}_4$ is taught the monomial `**11` by three teachers T_1, T_2, T_3. All teachers use the same examples, but give them in different orders. This leads to different success probability curves and to different expected teaching times (the numbers on the right end of the curves).

We refrain from including all the numerical calculations of the teaching success probabilities in the curves of Figure 2. The expected teaching times have been proved by Balbach [22, Fact 8.18].

Already these simple examples show that the randomized model is sensitive to feedback, memory size, and the order of examples. This sensitivity is also qualitatively correct, that is, teaching becomes faster with growing memory or with feedback.

4. Teaching with Feedback as a Markov Decision Process

4.1. Markov Decision Processes

All our randomized teaching model variants can be regarded as special cases of so called *Markov decision processes* (MDP). These processes have been extensively studied, and we refer the reader to Puterman [23], and Bertsekas [24]. In this section we introduce some basic terminology.

An 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 *State* of states, an initial state $s_0 \in \mathit{State}$, a finite set *Action* of actions, a function $\mathit{cost}\colon \mathit{State} \times \mathit{Action} \to \mathbb{R}$, and a function $\mathit{prob}\colon \mathit{State} \times \mathit{Action} \times \mathit{State} \to [0,1]$. The value $\mathit{cost}(s,a)$ specifies the cost incurred if action a is performed in state s. The value $\mathit{prob}(s,a,s')$ specifies the probability for the MDP to change from state s to s' under action a. A *policy* $\pi\colon State \to Action$ assigns an action to every state and thus induces a Markov chain.

A special case of Markov decision processes, which is still more general than our teaching scenario, are *stochastic shortest path problems* (SSPP). In an SSPP there is a set $State_* \subset State$ of target states. Once a target state has been reached it cannot be left and all actions in a target state incur no costs. In an SSPP the costs are then interpreted as lengths and a minimum expected cost policy corresponds to a tour with minimum expected length from the initial state to any of the target states.

The basic relation between SSPPs and our teaching model is as follows. The set *State* contains all states of the learner, the set *Action* contains all examples for the target, *cost* is set to 1, except for the target states, which incur no costs, and policies correspond to teachers. The function *prob* is identical to the function p defined in Equation (1). The teaching time of a teacher corresponds to the expected length of the path from the initial state to the target state under the policy corresponding to that teacher. The optimal teaching time corresponds to the minimal expected path length over all policies. A policy $\pi\colon State \to Action$ defines a Markov chain over *State* and for all $s \in State$ an expected time $H_\pi(s)$ to reach the target c^* from s. These expectations, called *hitting times*, satisfy the following linear equations for all $s \in State$:

$$H_\pi(s) = cost(s, \pi(s)) + \sum_{s' \in State} prob(s, \pi(s), s') \cdot H_\pi(s') \; . \tag{11}$$

For a given policy π it is therefore possible, by solving a system of linear equation of size $|State|$, to calculate the hitting times.

Under certain assumptions optimal policies and their expectations for SSPPs can be characterized (cf., e.g., Bertsekas and Tsitsiklis [25], or Puterman [23, Chapter 7]). These assumptions are as follows. First, all costs, except in the target state, have to be positive. This assumption is satisfied in our teaching model, since all costs are 1. Second, a so-called *proper* policy has to exist. A sufficient condition for properness is that in every state an action is chosen such that there is a positive probability of reaching the target state in the next round. A straightforward teacher that corresponds to a proper policy is a teacher that gives for every state an example inconsistent with the hypothesis. Such an example triggers a hypothesis change that leads to the target with positive probability.

Now we are ready to state the optimality condition in terms of SSPPs. Interpretations in terms of the teaching model are given in Subsections 4.2 and 4.3.

Lemma 4. *All hitting times $H(s)$ simultaneously assume their minimal values if and only if for all states $s \in State$:*

$$H(s) \quad = \quad \min_{a \in Action} \left(cost(s, a) + \sum_{s' \in State} prob(s, a, s') \cdot H(s') \right) .$$

A policy π has minimal hitting times for all states if and only if for all states $s \in State$:

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

The hitting time for a state $s \in \mathit{State}_*$ is $H(s) = 0$.

A policy $\pi\colon \mathit{State} \to \mathit{Action}$ corresponds to a teacher that receives feedback and that can thus choose an action depending on the current state of the learner. If the teacher receives no feedback the results about SSPPs, including Lemma 4, do not apply. The notion corresponding to this teaching scenario is that of an *unobservable stochastic shortest path problem* (USSPP). Only recently Patek [26, 27] has analyzed such problems and derived an optimality characterization analogous to Lemma 4 for them.

In the following subsection we shall show how the theory developed so far can be applied to teaching randomized learners with feedback. We start with the simplest case, i. e., with memoryless learners with feedback and then turn our attention to learners with infinite memory and feedback.

4.2. Memoryless Learners

Teaching memoryless learners with feedback presents the simplest situation. The teacher faces no uncertainty about the current state of the learner and there are only few states. In this subsection we aim to apply Lemma 4 to the special case of teaching the memoryless learner with feedback and thus derive a characterization of optimal teachers (cf. Lemma 5). We then use this criterion to develop an optimal teacher for the monomials (cf. Fact 6). As we shall see, this optimal teacher is greedy. We therefore continue by asking whether or not greedy teachers are always optimal and answer this question negatively. Finally we compare the teachability measure E_1^+ with other popular measures of teachability and learnability (Fact 8).

When $\mathcal{L}_1$ receives an example z, the new memory $\boldsymbol{s'}$ will contain only this example, $\boldsymbol{s'} = \langle z \rangle$, and the follow-up hypothesis is chosen from $\mathcal{C}(\langle z \rangle)$. Thus the behavior of $\mathcal{L}_1$ in a state $(\boldsymbol{s}, h)$ does not depend on $\boldsymbol{s}$ and in effect the memory is not part of the state. Therefore the state can be described by the hypothesis alone. More precisely, the learner $\mathcal{L}_1$ looks as follows (cf. Definition 1), where below we write $\mathcal{C}(z)$ for $\mathcal{C}(\langle z \rangle)$:

Current state: Hypothesis $h \in \mathcal{C} \cup \{\mathit{init}\}$.
Input: Example $z \in \mathcal{X}$.
Follow-up state: Hypothesis $h' \in \mathcal{C}$.

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

A teacher for teaching c^* to $\mathcal{L}_1$ with feedback is then a function $T\colon \mathcal{C} \cup \{\mathit{init}\} \to \mathcal{X}(c^*)$.

A teaching process with feedback involving $\mathcal{L}_1$ can be modeled as a stochastic shortest path problem with $\mathit{State} = \mathcal{C} \cup \{\mathit{init}\}$, $\mathit{State}_* = \{c^*\}$, $\mathit{Action} = \mathcal{X}(c^*)$, $\mathit{cost}(h, z) = 1$ for $h \neq c^*$ and $\mathit{cost}(c^*, z) = 0$ for all $z \in \mathcal{X}(c^*)$. Furthermore,

$$\mathit{prob}(h, z, h') \quad = \quad \begin{cases} 1/|\mathcal{C}(z)| \ , & \text{if } z \in \mathcal{X}(h') \setminus \mathcal{X}(h) \ ; \\ 0 \ , & \text{otherwise} \ ; \end{cases}$$

and $\mathit{prob}(c^*, z, c^*) = 1$ for all $z \in \mathcal{X}(c^*)$. The initial state is *init*.

Next, we derive a characterization of optimal teachers and of the minimum teaching time from the characterization of optimal policies (cf. Lemma 4). Note that if $\mathcal{L}_1$ is in state h, an example $z \in \mathcal{X}(h)$ does not change its state and is therefore useless. An optimal teacher refrains from teaching such examples.

Lemma 5. *Let $\mathcal{C}$ be a finite concept class and $c^* \in \mathcal{C}$ be a target. Let $H\colon \mathcal{C} \cup \{init\} \to \mathbb{R}$ be such that for all $h \in \mathcal{C} \cup \{init\} \setminus \{c^*\}$,*

$$H(h) \quad = \quad \min_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{X}(h)}} \left(1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h') \right) \tag{12}$$

and $H(c^*) = 0$. A teacher $T\colon \mathcal{C} \cup \{init\} \to \mathcal{X}(c^*)$ is optimal for teaching c^* to $\mathcal{L}_1$ with feedback if and only if for all $h \in \mathcal{C} \cup \{init\} \setminus \{c^*\}$,

$$T(h) \in \operatorname*{argmin}_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{X}(h)}} \left(1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h') \right) . \tag{13}$$

The minimum teaching time for teaching c^* to $\mathcal{L}_1$ with feedback is $H(init)$.

The characterization in Lemma 5 can be used to prove the optimality of teachers and the optimal teaching time for concepts. We show this for the class of monomials $\mathcal{M}_n$. Recall that $\mathcal{M}_n$ does not contain the concept $\emptyset$. Also note that there are 2^n monomials consistent with any positive example, and $3^n - 2^n$ monomials consistent with any negative example.

Fact 6. *Let $n \in \mathbb{N}$, $n \geq 2$ and let $\mathcal{M}_n$ be the concept class of monomials. Then the optimal teaching time for the concept $\mathtt{1}^k \mathtt{*}^{n-k}$ is*

$$E_1^+(\mathtt{1}^k \mathtt{*}^{n-k}, \mathcal{M}_n) \quad = \quad \frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1} + 2^{n+1} - 3^n}{3^n - 2^n + 2^{k-1}}$$

for all $k \in \{1, \ldots, n\}$. The optimal teaching time for the all-concept is

$$E_1^+(*^n, \mathcal{M}_n) \quad = \quad 2^n \ .$$

Input: Target $c^* \in \{\mathtt{0}, \mathtt{1}, \mathtt{*}\}^n$, hypothesis $h \in \{\mathtt{0}, \mathtt{1}, \mathtt{*}\}^n$.
Output: Example $(x, b) \in \mathcal{X}(c^*)$.

if $h \supset c^*$ **then output** $(x, 0)$ with

$$x[i] \quad = \quad \begin{cases} 1 - c^*[i] \ , & \text{if } i = \min\{j \ \mid h[j] = * \neq c^*[j]\} \ ; \\ c^*[i] \ , & \text{if } i \neq \min\{j \ \mid h[j] = * \neq c^*[j]\} \text{ and } c^*[i] \neq * \ ; \\ 0 \ , & \text{otherwise.} \end{cases}$$

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

Figure 3: Optimal teacher with feedback for the concept class $\mathcal{M}_n$ and the learner $\mathcal{L}_1$. When the hypothesis encompasses the target, the teacher gives a negative examples that maximizes the probability that the learner reaches, in the next round, a hypothesis that does *not* encompass the target.

Proof. Let T be the teacher defined in Figure 3. We begin with the simpler case $k = 0$, that is, $c^* = *^n$, and claim that H with $H(h) = 2^n$ for all $h \neq c^*$ is optimal. In this case every teacher that always gives an arbitrary positive example until the learner is in the target state is optimal. Every such example leads to one of 2^n hypotheses with equal probability of 2^{-n}. Therefore, for all $x \in \mathcal{X}(c^*)$ and for all $h \neq c^*$ we have:

$$1 + \frac{1}{|\mathcal{C}(x)|} \sum_{h' \in \mathcal{C}(x)} H(h') \quad = \quad 1 + 2^{-n} \cdot (2^n - 1)2^n = 2^n = H(h) \ .$$

The expectations H thus satisfy the first condition in Lemma 5. The teacher T satisfies the second condition in Lemma 5.

Now let the target concept c^* be represented by $\mathtt{1}^k *^{n-k}$ with $k \geq 1$. The behavior of the teacher is based on a partition of all hypotheses into two groups. Within a group, all hypotheses are assigned the example in the same way and have the same expected teaching time. The first group contains all hypotheses h with $h \supset c^*$. We refer to these hypotheses as $\supset$*-hypotheses*. The second group contains the remaining hypotheses (including *init*), called the $\not\supset$*-hypotheses*.

Now we define the expectations $H \colon \mathcal{C} \cup \{init\} \to \mathbb{R}$ by

$$H(h) \quad = \quad H_{\supset} := \frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1}}{3^n - 2^n + 2^{k-1}}$$

for all $\supset$-hypotheses h and

$$H(h) = H_{\not\supset} := \frac{(3^n - 2^n)(2^n + 2^k) - 2^{n+k-1} + 2^{n+1} - 3^n}{3^n - 2^n + 2^{k-1}}$$

for all $\not\supset$-hypotheses h. Note that for $n \geq 2$, $H_{\not\supset} < H_{\supset}$.

We have to prove that H and T satisfy Lemma 5. To achieve this goal, we show two claims.

Claim 1. *Let $(x, 0)$ be consistent with c^*, that is, $x \notin c^*$. Then*

(a) *there are $3^n - 2^n$ hypotheses consistent with $(x, 0)$;*
(b) *x is of the form $y\{0, 1\}^{n-k}$ with y containing $\ell \geq 1$ zeros;*
(c) *the number of $\supset$-hypotheses consistent with $(x, 0)$ is exactly $2^k - 2^{k-\ell} - 1$.*

Proof. Without loss of generality let $x \in 0^\ell 1^{k-\ell}\{0, 1\}^{n-k}$ for some $\ell \geq 0$.

(a) There are 2^n concepts containing x, hence there are $3^n - 2^n$ concepts that are consistent with $(x, 0)$.

(b) If $\ell = 0$, then x would be of the form $1^k\{0, 1\}^{n-k}$ and thus in c^*, a contradiction.

(c) A concept $d \in \mathcal{M}_n$ encompasses c^* if and only if d is of the form $\{\mathtt{1}, *\}^k *^{n-k}$. Furthermore, d is consistent with $(x, 0)$ (that is, $x \notin d$) if and only if d is of the form $y\{\mathtt{1}, *\}^{k-\ell} *^{n-k}$ with $y \in \{\mathtt{1}, *\}^\ell$ containing at least one "1". There are exactly $(2^\ell - 1) \cdot 2^{k-\ell} = 2^k - 2^{k-\ell}$ concepts satisfying the latter condition. Since c^* does not count as $\supset$-hypothesis, the sought number is one less, as claimed. □ (Claim 1)

Claim 2. *For every positive example there are 2^n consistent hypotheses of which $2^k - 1$ are $\supset$-hypotheses.*

Proof. For every instance $x \in 1^k\{0,1\}^{n-k}$ there are exactly 2^n concepts containing x. The concepts that result from substituting 1's by $*$ in c^* are the only concepts containing x and c^*. There are $2^k - 1$ such concepts (c^* itself is not a $\supset$-hypothesis). $\square$ (Claim 2)

We continue by proving the necessary property for H.

Case 1. $h \supset c^*$.

Then only negative examples $z = (x, 0)$ are inconsistent with h. Without loss of generality, let $x \in 0^\ell 1^{k-\ell}\{0,1\}^{n-k}$ with $1 \le \ell \le k$. Note that, by Claim 1, Assertion (b), the case $\ell = 0$ cannot occur. Then

$$1 + \frac{1}{|\mathcal{C}(x)|} \sum_{h' \in \mathcal{C}(x)} H(h') \quad = \quad 1 + \frac{2^k - 2^{k-\ell} - 1}{3^n - 2^n} \cdot H_{\supset} + \frac{3^n - 2^n - 2^k + 2^{k-\ell}}{3^n - 2^n} \cdot H_{\not\supset} \; .$$

The sum of the coefficients of $H_{\supset}$ and $H_{\not\supset}$ is $(3^n - 2^n - 1)/(3^n - 2^n)$ and thus independent of ℓ. The right hand side of the equation becomes minimal for $\ell = 1$, since the coefficient of $H_{\supset}$ becomes minimal for $\ell = 1$ and, moreover, $H_{\supset} > H_{\not\supset}$. After plugging in $H_{\supset}$ and $H_{\not\supset}$, a tedious calculation shows that this minimal value equals $H_{\supset}$. Consequently, in Case 1 the Condition (12) is satisfied.

Case 2. $h \not\supset c^*$.

If z is a positive example then by Claim 2,

$$1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h') \quad = \quad 1 + \frac{2^k - 1}{2^n} \cdot H_{\supset} + \frac{2^n - 2^k}{2^n} \cdot H_{\not\supset} = H_{\not\supset}$$

where the last equality is due to a tedious calculation. If z is a negative example then

$$1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h') \quad = \quad 1 + \frac{2^k - 2^{k-\ell} - 1}{3^n - 2^n} \cdot H_{\supset} + \frac{3^n - 2^n - 2^k + 2^{k-\ell}}{3^n - 2^n} \cdot H_{\not\supset}.$$

Again this expression is minimized for $\ell = 1$ and its minimal value is $H_{\supset}$. The value for a positive example is therefore smaller. Consequently, the minimal value of $1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h')$ is $H_{\not\supset}$, and the Condition (12) is also satisfied for $\not\supset$-hypotheses.

In Case 1 and 2 above we have identified examples minimizing the value of $1 + \frac{1}{|\mathcal{C}(z)|} \sum_{h' \in \mathcal{C}(z)} H(h')$. The teacher in Figure 3 always teaches such examples. Therefore, the teacher in Figure 3 satisfies Condition (13) in Lemma 5 and is thus optimal. $\square$

The teacher from Figure 3 can be computed in linear time. It outputs a positive example whenever possible (that is, 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 that T follows 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 round.

Definition 4. Let $\mathcal{C}$ be a concept class over X and $c^* \in \mathcal{C}$. A teacher $T \colon \mathcal{C} \cup \{init\} \to \mathcal{X}$ for c^* is called *greedy* iff for all $h \in \mathcal{C}$: $T(h) \in \operatorname*{argmin}_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{C}(h)}} |\mathcal{C}(z)|$.

The notion of greedy teacher cannot be generalized to arbitrary stochastic shortest path problems, since in general the target state cannot be reached from all states under all actions. The question of how good a greedy teacher can be is thus a teaching-specific question which cannot be answered directly by the MDP theory. Such a greedy strategy seems sensible in general and is provably optimal in the case of monomials. However, there are classes where no greedy teacher is optimal.

	x_1	x_2	x_3	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	T'	T^g
init	–	–	–	–	–	–	–	–	–	–	–	–	–	$(x_3, 1)$	$(x_3, 1)$
c^*	1	1	1	1	1	1	1	1	1	1	1	1	1	–	–
	0	0	1	0	1	1	1	1	1	1	1	1	1	$(x_1, 1)$	$(x_2, 1)$
	0	1	1	1	0	1	1	1	1	1	1	1	1	$(x_1, 1)$	$(x_1, 1)$
	0	1	1	1	1	0	1	1	1	1	1	1	1	$(x_1, 1)$	$(x_1, 1)$
	1	0	0	1	1	1	0	1	1	1	1	1	1	$(x_3, 1)$	$(x_3, 1)$
	1	0	0	1	1	1	1	0	1	1	1	1	1	$(x_3, 1)$	$(x_3, 1)$
	1	0	0	1	1	1	1	1	0	1	1	1	1	$(x_3, 1)$	$(x_3, 1)$
	1	1	0	1	1	1	1	1	1	0	1	1	1	$(x_3, 1)$	$(x_3, 1)$
	1	1	0	1	1	1	1	1	1	1	0	1	1	$(x_3, 1)$	$(x_3, 1)$
	1	1	0	1	1	1	1	1	1	1	1	0	1	$(x_3, 1)$	$(x_3, 1)$
	1	1	0	1	1	1	1	1	1	1	1	1	0	$(x_3, 1)$	$(x_3, 1)$

Figure 4: The concept class $\mathcal{C}_1$ from the proof of Fact 7. For a family of generalizations of this class, the greedy teacher T^g has an expected teaching time of $\Omega(|\mathcal{C}|)$ rounds greater than teacher T' has.

Fact 7. *There is a family of concept classes $\mathcal{C}_n$ with $|\mathcal{C}_n| = 10n + 1$ and target concepts such that the difference between the teaching time of the greedy teacher and the optimal teaching time is $\Omega(n)$.*

PROOF. Let $n \geq 1$ and $X_n = \{x_1, x_2, x_3\}$, and let the learning domain be $Y_n = X_n \cup \{y_1, \ldots, y_{10n}\}$. The class $\mathcal{C}_n$ consists of the target concept $c^* = Y_n$, n concepts c with $c \cap X_n = \{x_3\}$, $2n$ concepts c with $c \cap X_n = \{x_2, x_3\}$, $3n$ concepts c with $c \cap X_n = \{x_1\}$, and $4n$ concepts c with $c \cap X_n = \{x_1, x_2\}$. Each non-target concept contains all but one of the instances $y_1, \ldots, y_{10n}$, and each such instance is missing from exactly one non-target concept. This differentiates the concepts that are equal wrt. X_n (see $\mathcal{C}_1$ in Figure 4).

The unique greedy teacher is $T^g(c) = (x_3, 1)$ if $c(x_3) = 0$ or $c = init$, $T^g(c) = (x_2, 1)$ if $c \cap X_n = \{x_3\}$, and $T^g(c) = (x_1, 1)$ for all remaining concepts except the target.

Denoting the teaching times when starting in a state c with $T^g(c) = (x_j, 1)$ by H_j, $j = 1, 2, 3$, we have

$$\begin{aligned} H_1 &= 1 + 7n/(7n+1) \cdot H_3 \\ H_2 &= 1 + 2n/(6n+1) \cdot H_1 + 4n/(6n+1) \cdot H_3 \\ H_3 &= 1 + 2n/(3n+1) \cdot H_1 + n/(3n+1) \cdot H_2 \ . \end{aligned}$$

This yields $H_3 = \frac{266n^3+122n^2+19n+1}{63n^2+16n+1}$ for the expected teaching time of T^g.

We define the teacher T' like T^g except that $T'(c) = (x_1, 1)$ for all concepts c with $T^g(c) = (x_2, 1)$. Analogously to the above, one determines that the expected teaching time of the teacher T' is $H' = \frac{(6n+1)(7n+1)}{10n+1}$. Furthermore, an explicit calculation shows that $H_3 - H' = \Theta(n)$. Since H' is an upper bound for the optimal teaching time, it follows that $H_3 - E_1^+(c^*, \mathcal{C}_n) = \Omega(n)$. □

Next, we compare E_1^+ with other dimensions that occur in learning theory. In particular, 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 8. (1) *For all concept classes $\mathcal{C}$ and all concepts $c \in \mathcal{C}$, $E_1^+(c, \mathcal{C}) \geq \mathrm{TD}(c, \mathcal{C})$.*
(2) *There is no function of* TD *upper bounding E_1^+.*
(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. For every example $x \in \mathcal{X}(c)$ there are at least $\mathrm{TD}(c, \mathcal{C})$ consistent hypotheses. Consequently, in every round the probability of reaching the target is at most $1/\mathrm{TD}(c, \mathcal{C})$. The expected number of rounds is therefore at least $\mathrm{TD}(c, \mathcal{C})$. This proves Assertion (1).

Let $\mathcal{C}_n = \{c \subseteq \{1, \ldots, n\} \mid |c| = 2\}$. Then $\mathrm{TD}(\mathcal{C}_n) = 2$, but $E_1^+(\mathcal{C}_n) = n - 1$, since the optimal teacher gives positive examples all the time and there are $n - 1$ hypotheses consistent with such an example. Thus Assertion (2) is shown.

Let $\mathcal{C}_n = \{c \subseteq \{1, \ldots, n\} \mid |c| = 1\}$. Then $E_1^+(c, \mathcal{C}_n) = 1$ for all $c \in \mathcal{C}_n$, but $MQ(\mathcal{C}_n) = n - 1$ and Assertion (3) follows.

As an easy calculation shows $MQ(\mathcal{A}_n) = n$ and $E_1^+(\mathcal{A}_n) = 2^{n-1}$ and Assertion (4) is proved.

It is known (see, for example, 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 Assertion (5). □

Roughly speaking, teaching the learner $\mathcal{L}_1$ can take arbitrarily longer than teaching in the teaching dimension model, but is still incomparable with membership query learning.

4.3. Learners with Infinite Memory

Learners with infinite memory can have infinitely many states $(\boldsymbol{s}, h) \in \mathcal{X}^* \times (\mathcal{C} \cup \{init\})$. But the behavior of the learner $\mathcal{L}_\infty$ is not affected by memorizing the same example multiple times. This makes it pointless to teach the same example twice. Thus it suffices to consider the finitely many memories of length at most $|X|$. The number of states is therefore only finite. From the SSPP optimality criterion Lemma 4 we can then immediately derive a characterization of optimal teachers for $\mathcal{L}_\infty$. This characterization is more complicated than in the $\mathcal{L}_1$ case (Lemma 5) and difficult to write in a closed form. Our first task is thus to simplify this criterion (Lemma 11) by proving that an optimal teacher always gives examples that are inconsistent with the current hypothesis (see Proposition 10).

The criterion also yields an algorithm, called *backward induction*, for computing the optimal teaching time. This algorithm's runtime, however, is not polynomial in the representation of the concept class. A straightforward idea to improve the backward induction is to consider only the first $\mathrm{TD}(c^*)$ rounds of the teaching process, since there is always a teacher successful after that many rounds. But, as we show in Fact 13, this modified algorithm does not always yield the optimal teaching time. Indeed, that an efficient algorithm for computing E^+_∞ is unlikely to exist is then shown in Theorem 15. This result is based on a general lemma (cf. Lemma 12) that relates E^+_∞ to the teaching dimension.

It is not difficult to formally describe the SSPP corresponding to teaching $c^* \in \mathcal{C}$ to $\mathcal{L}_\infty$ with feedback: We stipulate that in every state $(\boldsymbol{s}, h)$ only examples $z \notin \boldsymbol{s}$ can be given. As already mentioned, an optimal teacher would not teach other examples. Thus, we only have to consider memories in which no example occurs twice. Therefore, the set of states is

$$\mathit{State} \quad = \quad \{(\boldsymbol{s}, h) \in \mathcal{X}(c^*)^{\leq |X|} \times (\mathcal{C} \cup \{\mathit{init}\}) \ \mid h \in \mathcal{C}(\boldsymbol{s}) \ \wedge \ (i \neq j \Rightarrow \boldsymbol{s}[i] \neq \boldsymbol{s}[j])\} \ .$$

The initial state is $(\langle\ \rangle, \mathit{init})$ and the set of target states is

$$\mathit{State}_* \quad = \quad \{(\boldsymbol{s}, h) \mid (\boldsymbol{s}, h) \in \mathit{State},\ h = c^*\} \ .$$

For a state $(\boldsymbol{s}, h)$ and an example $z \notin \boldsymbol{s}$ the transition probabilities are

$$\begin{aligned} &\mathit{prob}((\boldsymbol{s}, h), z, (\boldsymbol{s'}, h')) \\ &= p((\boldsymbol{s}, h), z, (\boldsymbol{s'}, h')) = \begin{cases} 1\ , & \text{if } z \in \mathcal{X}(h) \ \wedge \ \boldsymbol{s'} = \boldsymbol{s} \circ \langle z \rangle \ \wedge \ h' = h\ ; \\ 1/|\mathcal{C}(\boldsymbol{s'})|\ , & \text{if } z \notin \mathcal{X}(h) \ \wedge \ \boldsymbol{s'} = \boldsymbol{s} \circ \langle z \rangle \ \wedge \ h' \in \mathcal{C}(\boldsymbol{s'})\ ; \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

As usual, the costs are 1 for each example, except the examples given in the target state, which are costless, i.e.,

$$\mathit{cost}((\boldsymbol{s}, h), x) \quad = \quad \begin{cases} 1\ , & \text{if } h \neq c^*\ ; \\ 0\ , & \text{if } h = c^*\ . \end{cases}$$

Plugging the above into Lemma 4 yields an optimality characterization that is hard to write concisely. This is because we have to distinguish three cases for p. In comparison, Lemma 5 looks rather simple because we could confine the actions to examples that are inconsistent with the current hypothesis. This was possible, since a consistent example would not change the state of $\mathcal{L}_1$. Giving a consistent example to $\mathcal{L}_\infty$, however, does change the learner's state. Below we show that it nevertheless suffices to consider teachers that always give inconsistent examples. This not only yields a simpler characterization, it is also interesting in its own right.

Our first step towards this result is that the order of the examples in an infinite memory is not important. Therefore, we can regard the memory as a set rather than as a sequence. As a by-product, this reduces the number of states we have to consider.

Proposition 9. *Let H satisfy the condition in Lemma 4 for an SSPP corresponding to teaching $c^* \in \mathcal{C}$ to $\mathcal{L}_\infty$ with feedback. Let $(\boldsymbol{s}, h), (\tilde{\boldsymbol{s}}, h) \in \mathit{State}$ with $\boldsymbol{s}, \tilde{\boldsymbol{s}} \in \mathcal{X}^k$ and $\{\boldsymbol{s}[1], \ldots, \boldsymbol{s}[k]\} = \{\tilde{\boldsymbol{s}}[1], \ldots, \tilde{\boldsymbol{s}}[k]\}$. Then $H(\boldsymbol{s}, h) = H(\tilde{\boldsymbol{s}}, h)$.*

PROOF. The proof is by induction on the length k of $\boldsymbol{s}$ and $\tilde{\boldsymbol{s}}$. We start at the maximal length $k = |X|$. Let $(\boldsymbol{s}, h), (\tilde{\boldsymbol{z}}, h) \in \mathit{State}$ with $|\boldsymbol{s}| = |X|$. Then both $\boldsymbol{s}$ and $\tilde{\boldsymbol{s}}$ contain all examples in $\mathcal{X}(c^*)$ and thus $h = c^*$. Therefore $H(\boldsymbol{s}, h) = H(\tilde{\boldsymbol{s}}, h) = 0$.

Now assume the statement holds for all sequences of length $k > 0$. We show the statement for all memories of length $k-1$. Let $\boldsymbol{s}, \tilde{\boldsymbol{s}} \in \mathcal{X}(c^*)^{k-1}$ be memories with identical range and h a hypothesis such that $(\boldsymbol{s}, h), (\tilde{\boldsymbol{s}}, h) \in \mathit{State}$. If $h = c^*$ then both H-values are again 0. We thus assume $h \neq c^*$. Since $\boldsymbol{s}$ and $\tilde{\boldsymbol{s}}$ have the same range, we have for all h' and all z by the definition of p:

$$p((\boldsymbol{s}, h), z, (\boldsymbol{s} \circ \langle z \rangle, h')) \quad = \quad p((\tilde{\boldsymbol{s}}, h), z, (\tilde{\boldsymbol{s}} \circ \langle z \rangle, h')) \ .$$

From the latter equality and Lemma 4 we thus obtain:

$$\begin{aligned}
H(\boldsymbol{s}, h) &= \min_{z \in \mathcal{X}(c^*)} \left(1 + \sum_{(\boldsymbol{s'}, h') \in \mathit{State}} p((\boldsymbol{s}, h), z, (\boldsymbol{s'}, h')) \cdot H(\boldsymbol{s'}, h') \right) \\
&= \min_{z \in \mathcal{X}(c^*)} \left(1 + \sum_{(\boldsymbol{s} \circ \langle z \rangle, h') \in \mathit{State}} p((\boldsymbol{s}, h), z, (\boldsymbol{s} \circ \langle z \rangle, h')) \cdot H(\boldsymbol{s} \circ \langle z \rangle, h') \right) \\
&= \min_{z \in \mathcal{X}(c^*)} \left(1 + \sum_{(\tilde{\boldsymbol{s}} \circ \langle z \rangle, h') \in \mathit{State}} p((\tilde{\boldsymbol{s}}, h), z, (\tilde{\boldsymbol{s}} \circ \langle z \rangle, h')) \cdot H(\tilde{\boldsymbol{s}} \circ \langle z \rangle, h') \right) \\
&= H(\tilde{\boldsymbol{s}}, h).
\end{aligned}$$

Note that the third equality is valid, since $H(\boldsymbol{s} \circ \langle z \rangle, h') = H(\tilde{\boldsymbol{s}} \circ \langle z \rangle, h')$ by the induction hypothesis. □

Proposition 10. *Let $\mathcal{C}$ be a concept class and c^* be a target. Then there is an optimal teacher T' for teaching c^* to $\mathcal{L}_\infty$ with feedback that never gives an example consistent with the current hypothesis, that is,*

$$T'(\boldsymbol{s}, h) \notin \mathcal{X}(h)$$

for all $(\boldsymbol{s}, h) \in \mathit{State} \setminus \mathit{State}_*$.

PROOF. Let H satisfy the first condition in Lemma 4. Furthermore, let T be a teacher that gives a consistent example $z_1 = T(\boldsymbol{s}, h) \in \mathcal{X}(h)$ when the learner is in a state $(\boldsymbol{s}, h) \in \mathit{State} \setminus \mathit{State}_*$. We assume that $\boldsymbol{s}$ is of maximal length with this property. Thus in the follow-up state $(\boldsymbol{s} \circ \langle z_1 \rangle, h)$ the teacher T gives an example $z_2 = T(\boldsymbol{s} \circ \langle z_1 \rangle, h) \notin \mathcal{X}(h)$.

We show that this teacher does not satisfy the second condition in Lemma 4 for state $(\boldsymbol{s}, h)$, i.e., we prove that

$$z_1 \notin \operatorname*{argmin}_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \boldsymbol{s}}} \left(\mathit{cost}((\boldsymbol{s}, h), z) + \sum_{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z \rangle)} \mathit{prob}((\boldsymbol{s}, h), z, (\boldsymbol{s} \circ \langle z \rangle, h)) \cdot H(\boldsymbol{s} \circ \langle z \rangle, c) \right) .$$

Let us denote the expression in the large parentheses by Y_z. The value of Y_z is

$$Y_z \quad = \quad 1 + \frac{1}{|\mathcal{C}(\boldsymbol{s} \circ \langle z \rangle)|} \cdot \sum_{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z \rangle)} H(\boldsymbol{s} \circ \langle z \rangle, c)$$

for examples $z \notin \mathcal{X}(h)$ and

$$Y_z \quad = \quad 1 + H(\boldsymbol{s} \circ \langle z \rangle, h)$$

for examples $z \in \mathcal{X}(h)$. The value of Y_z, with z set to z_1, is

$$Y_{z_1} \quad = \quad 1 + H(\boldsymbol{s} \circ \langle z_1 \rangle, h) = 2 + q \cdot \sum_{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z_1, z_2 \rangle)} H(\boldsymbol{s} \circ \langle z_1, z_2 \rangle, c) \tag{14}$$

with $q = 1/|\mathcal{C}(\boldsymbol{s} \circ \langle z_1, z_2 \rangle)|$.

Next, we prove that (14) is not the minimal value of Y_z over all examples z by showing that $Y_{z_2} < Y_{z_1}$. At first we have

$$Y_{z_2} \quad = \quad 1 + q' \cdot \sum_{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)} H(\boldsymbol{s} \circ \langle z_2 \rangle, c) \tag{15}$$

with $q' = 1/|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)|$. The values $H(\boldsymbol{s} \circ \langle z_2 \rangle, c)$ in the summation are

$$\begin{aligned}
&H(\boldsymbol{s} \circ \langle z_2 \rangle, c) \\
&\quad = \min_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \boldsymbol{s} \circ \langle z_2 \rangle}} \left(1 + \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z \rangle)} \mathit{prob}((\boldsymbol{s} \circ \langle z_2 \rangle, c), z, (\boldsymbol{s} \circ \langle z_2, z \rangle, c')) \cdot H(\boldsymbol{s} \circ \langle z_2, z \rangle, c') \right) \\
&\quad < 1 + \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} \mathit{prob}((\boldsymbol{s} \circ \langle z_2 \rangle, c), z_1, (\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c')) \cdot H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') \ ,
\end{aligned}$$

where the upper bound results from setting z to z_1. When substituting the upper bounds just derived for the H-values in (15), we have to distinguish between hypotheses for which z_1 is consistent and those for which z_1 triggers a hypothesis change. We get as upper bound for (15):

$$\begin{aligned}
Y_{z_2} \ < \ 1 + q' \Bigg(&\sum_{\substack{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle) \\ c \notin \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \\ c \neq c^*}} \left(1 + q'' \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') \right) \\
&+ \sum_{\substack{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \\ c \neq c^*}} (1 + H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c)) \Bigg)
\end{aligned}$$

with $q'' = 1/|\mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)| = q$. Now all occurring H-values have $\boldsymbol{s} \circ \langle z_2, z_1 \rangle$ as first argument. Removing the first summation yields

$$1 + q' \left(|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle) \setminus \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \setminus \{c^*\}| \cdot \left(1 + q'' \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') \right) \right.$$
$$\left. + \sum_{\substack{c \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \\ c \neq c^*}} (1 + H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c)) \right) .$$

We set $r = |\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle) \setminus \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \setminus \{c^*\}|$ and $r' = |\mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \setminus \{c^*\}|$. After multiplying out we obtain

$$Y_{z_2} < 1 + q'r' + q'r'q'' \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') + q'r + q'r \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c')$$

and after sorting the terms,

$$Y_{z_2} \quad < \quad 1 + q'r' + q'r + (q'rq'' + q') \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') \; . \tag{16}$$

The first three terms can be upper bounded by $1 + q'r' + q'r = 1 + q'(r + r') = 2 - q' < 2$. The coefficient $q'rq'' + q'$ of the summation can be upper bounded as follows:

$$\begin{aligned} q'rq'' + q' \quad &= \quad q'' \cdot \frac{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle) \setminus \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle) \setminus \{c^*\}|}{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)|} + q' \\ &= \quad q'' \cdot \left(1 - \frac{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)| + 1}{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)|} \right) + q' \\ &< \quad q'' \cdot \left(1 - \frac{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)|}{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)|} \right) + q' \\ &= \quad q'' - \frac{1}{|\mathcal{C}(\boldsymbol{s} \circ \langle z_2 \rangle)|} + q' \\ &= \quad q'' \; . \end{aligned}$$

Applying these upper bounds to (16), it follows

$$\begin{aligned} Y_{z_2} \quad &< \quad 2 + q'' \cdot \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') \\ &= \quad 2 + q \cdot \sum_{c' \in \mathcal{C}(\boldsymbol{s} \circ \langle z_2, z_1 \rangle)} H(\boldsymbol{s} \circ \langle z_1, z_2 \rangle, c') \\ &= \quad Y_{z_1} \; , \end{aligned}$$

where the first equality holds because $q'' = q$ and $H(\boldsymbol{s} \circ \langle z_2, z_1 \rangle, c') = H(\boldsymbol{s} \circ \langle z_1, z_2 \rangle, c')$ by Proposition 9. Therefore, Y_{z_2} is strictly less than Y_{z_1}. This means that the example z_1 is not in the set $\operatorname{argmin}_{z \in \mathcal{X}(c^*), z \notin \boldsymbol{s}} Y_z$. This shows that the teacher T does not satisfy the second condition in Lemma 4. $\square$

Now, we are in the position to state the optimality characterization.

Lemma 11. *Let $\mathcal{C}$ be a finite concept class and $c^* \in \mathcal{C}$ be a target. Furthermore, let $H\colon \mathcal{X}^{\leq |X|} \times (\mathcal{C} \cup \{init\}) \to \mathbb{R}$ be such that for all $(\boldsymbol{s}, h) \in State \setminus State_*$,*

$$H(\boldsymbol{s}, h) \quad = \quad \min_{\substack{z \in \mathcal{X}(c^*) \\ z \notin \mathcal{X}(h)}} \left(1 + \frac{1}{|\mathcal{C}(\boldsymbol{s} \circ \langle z \rangle)|} \sum_{h' \in \mathcal{C}(\boldsymbol{s} \circ \langle z \rangle))} H(\boldsymbol{s} \circ \langle z \rangle, h') \right)$$

and for all $(\boldsymbol{s}, h) \in State_*$, $H(\boldsymbol{s}, h) = 0$. A teacher $T\colon \mathcal{X}^* \times (\mathcal{C} \cup \{init\}) \to \mathcal{X}(c^*)$ is optimal for teaching c^* to $\mathcal{L}_\infty$ with feedback if and only if for all $(\boldsymbol{s}, h) \in State \setminus State_*$,

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

The minimum teaching time for teaching c^* to $\mathcal{L}_\infty$ with feedback is $H(\langle\ \rangle, init)$.

Lemma 11 is virtually the same as Lemma 5, only with a larger set of states. For a given state $(\boldsymbol{s}, h)$ the sum in the minimization ranges over the hitting times of all possible follow-up states. The main difference to the condition in Lemma 5 is that there are no cyclic dependencies within the H-values. Intuitively, the reason for this is that a learner with infinite memory cannot reach the same state twice during a teaching process because in each round the memory grows.

The lack of cyclic dependencies yields a straightforward inductive algorithm for computing all optimal hitting times. A state $(\boldsymbol{s}, h)$ with $|\boldsymbol{s}| = |X|$ has a hitting time of $H(\boldsymbol{s}, h) = 0$. The optimal hitting times for states with smaller memories can be computed using the first formula in Lemma 11, until finally the states with empty memory are reached. This algorithm is called *backward induction* and runs in time polynomial in the size the tabular representation of the MDP, but not polynomial in the representation size of the teaching problem, that is, in the matrix representation of $\mathcal{C}$.

A tempting idea for improvement is based on the observation that a teacher that gives a minimum teaching set is always successful after $\mathrm{TD}(c^*)$ rounds. We shall call these teachers *MTS teachers*. Not every such teacher is optimal, but one could conjecture that there is at least one optimal teacher among them. This is, however, not always the case. To show this, we use the following lemma, which gives an upper and lower bound for $E^+_\infty(c^*, \mathcal{C})$ in terms of the teaching dimension.

Lemma 12. *Let $\mathcal{C}$ be a concept class and let $c^* \in \mathcal{C}$ be a target. Then, for all $\mu \in \{1, \ldots, \mathrm{TD}(c^*, \mathcal{C})\}$ we have,*

$$E^-_\mu(c^*, \mathcal{C}) \quad \geq \quad E^+_\mu(c^*, \mathcal{C}) \ \geq\ \frac{\mu(\mu - 1)}{2\mathrm{TD}(c^*, \mathcal{C})} + \mathrm{TD}(c^*) + 1 - \mu\ ,$$

and for all $\mu > \mathrm{TD}(c^*, \mathcal{C})$ and for $\mu = \infty$,

$$\mathrm{TD}(c^*, \mathcal{C}) \quad \geq \quad E^-_\mu(c^*, \mathcal{C}) \ \geq\ E^+_\mu(c^*, \mathcal{C}) \ \geq\ \frac{\mathrm{TD}(c^*, \mathcal{C})}{2}\ .$$

PROOF. Let $k = \mathrm{TD}(c^*)$ and $\mu \in \{1, \ldots, \mathrm{TD}(c^*)\}$. It suffices to show the statement for E_μ^+. The proof is based on a simple observation.

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

Proof. Assume $|\mathcal{C}(\{z_0, \ldots, z_{i-1}\})| \leq k - i$. 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 $\mathrm{TD}(c^*) = k$. □ Claim 1

Using Claim 1, 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 1). Thus, the probability for reaching c^* in round i is at most $p_i = 1/(k-i)$. 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

$$\sum_{i=0}^{\mu-2}(i+1) \cdot p_i \cdot \prod_{j=0}^{i-1}(1-p_j) \;+\; \sum_{i=\mu-1}^{\infty}(i+1) \cdot p_i \cdot \prod_{j=0}^{i-1}(1-p_j)\ . \tag{17}$$

We first calculate the second sum in (17). 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 right sum equals $\frac{k-\mu+1}{k} \cdot (1 - p_{\mu-1})^{i-\mu+1}$ and the whole sum can be written as

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

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 (17) we therefore get

$$\tfrac{k-\mu+1}{k} \cdot (\mu - 1 + k - \mu + 1) \quad = \quad k - \mu + 1\ .$$

Calculating the first sum in (17) yields

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

Putting it together, we obtain $\frac{\mu(\mu-1)}{2k} + k + 1 - \mu$ as the value of (17).

	x_1	x_2	x_3	z	y_1	y_2	$\cdots$	y_{11}	y_{12}
init	−	−	−	−	−	−	$\cdots$	−	−
c^*	1	1	1	1	1	1	$\cdots$	1	1
	1	1	0	1	0	1	$\cdots$	1	1
	1	0	1	1	1	0		1	1
	0	1	1	1	1	1		1	1
	1	1	0	0	1	1		1	1
	1	0	1	0	1	1		1	1
	0	1	1	0	1	1		1	1
	1	1	0	0	1	1		1	1
	1	0	1	0	1	1		1	1
	0	1	1	0	1	1		1	1
	1	1	0	0	1	1		1	1
	1	0	1	0	1	1		0	1
	0	1	1	0	1	1	$\cdots$	1	0

Figure 5: Concept class and target for which the optimal $\mathcal{L}_\infty$-teacher with feedback has not finished after $\mathrm{TD}(c^*) = 3$ rounds. This class is $\mathcal{C}_{3,4}$ in the proof of Fact 13.

For $\mu > \mathrm{TD}(c^*)$ the teaching process described above takes at most $\mathrm{TD}(c^*)$ rounds. The lower bound is therefore the same as for $\mu = \mathrm{TD}(c^*)$. Moreover, a teacher giving the examples of a minimum teaching set is successful after $\mathrm{TD}(c^*)$ rounds, from which follows $\mathrm{TD}(c^*) \geq E^-_\mu(c^*) \geq E^+_\mu(c^*)$. □

The expected teaching time of an MTS teacher is at most the teaching dimension of the target. From Lemma 12 we know that an MTS teacher can have at most twice the teaching time of an optimal teacher. Next, we show this bound to be asymptotically tight.

Fact 13. *For every $k > 1$ there is a family of concept classes and target concepts with teaching dimension k such that the ratio of any MTS teacher's teaching time and the optimal teaching time tends to $2 - O(\frac{1}{k})$.*

Proof. For $k, n > 1$ we define the concept class $\mathcal{C}_{k,n}$ as follows. Let $X_k = \{x_1, \ldots, x_k\}$ and $Y_{k,n} = \{y_1, \ldots, y_{kn}\}$ be sets of instances, and let $Z_k = X_k \cup Y_{k,n} \cup \{z\}$ be the learning domain. The target concept is $c^* = Z_k$, and there are, for every $i = 1, \ldots, k$, exactly n concepts c with $X_k \cap c = X_k \setminus \{x_i\}$, out of which exactly one satisfies $c(z) = 1$. The instances in $Y_{k,n}$ serve to make all concepts distinct (see $\mathcal{C}_{3,4}$ in Figure 5).

There is only one minimum teaching set for c^*, namely $S = \{(x_1, 1), \ldots, (x_k, 1)\}$. For symmetry reasons, all MTS teachers have the same teaching time,

$$H_{k,n} \quad = \quad \frac{1}{(k-1)n+1} + \sum_{i=2}^{k} \prod_{j=1}^{i-1} \left(1 - \frac{1}{(k-j)n+1}\right) \cdot \frac{1}{(k-i)n+1} \cdot i \; .$$

This value is bounded from above by $|S| = k$, and the last summand tends to k as n tends to infinity; thus $\lim_{n\to\infty} H_{k,n} = k$.

Now we consider the teacher that first teaches $(z, 1)$ and then always the example from S that is inconsistent with the observed hypothesis. Its expected teaching time is

$$H'_{k,n} \;=\; \frac{1}{k+1} + \sum_{i=2}^{k+1} \prod_{j=1}^{i-1} \left(1 - \frac{1}{k-j+2}\right) \cdot \frac{1}{k-i+2} \cdot i \;=\; 1 + \frac{k}{2} \,.$$

Since $H'_{k,n} \geq E^{+}_{\infty}(c^*, \mathcal{C}_{k,n}) \geq k/2$, we have

$$2 \;\geq\; \lim_{n\to\infty} H_{k,n}/E^{+}_{\infty}(c^*, \mathcal{C}_{k,n}) \;\geq\; \lim_{n\to\infty} H_{k,n}/H'_{k,n} \;=\; 2 - \frac{4}{k+2} \,.$$

□

We are going to show that there is most likely no algorithm approximating $E^{+}_{\infty}(c^*, \mathcal{C})$ up to a constant factor and running in time polynomial in the size of the matrix representation of $\mathcal{C}$. Setting $\mu = \infty$, we conclude from Lemma 12 that

$$\mathrm{TD}(c^*, \mathcal{C}) \quad\geq\quad E^{+}_{\infty}(c^*, \mathcal{C}) \;\geq\; \frac{\mathrm{TD}(c^*, \mathcal{C})}{2}, \tag{18}$$

which means that every algorithm computing $E^{+}_{\infty}(c^*, \mathcal{C})$ also computes a factor 2 approximation of the teaching dimension. Therefore, we continue with a closer look at the complexity of deciding and approximating the teaching dimension. Deciding whether a given concept in a given class has a teaching dimension of less then a given value is $\mathcal{NP}$-complete [13, 12, 18]. This can be shown by a reduction from the `SET-COVER` problem.

As an optimization problem, `SET-COVER` has been studied intensively, and it is relatively easy to translate these results to the problem of computing optimal teaching sets. For the formal definition of the `MIN-TEACHING-SET` problem we assume that $X = \{1, \ldots, k\}$ for some $k \in \mathbb{N}^+$ and that $\mathcal{C}$ is represented as a binary $|\mathcal{C}| \times k$ matrix. A concept is represented as a binary string of length k. In the following definitions we use the terminology from Ausiello *et al.* [28].

Definition 5. The problem `MIN-TEACHING-SET` is the optimization problem with

- instances of the form $(\mathcal{C}, c^*)$,
- feasible solutions $sol(\mathcal{C}, c^*) = \{\boldsymbol{s} \in X^* \mid \mathcal{C}(\boldsymbol{s}) = \{c^*\}\}$,
- a measure mes with $mes((\mathcal{C}, c^*), \boldsymbol{s}) = |\boldsymbol{s}|$ for all $\boldsymbol{s} \in sol(\mathcal{C}, c^*)$.

Definition 6. The problem `SET-COVER` is the optimization problem with

- instances of the form $(U, V_1, \ldots, V_k)$ with a finite set U and sets $V_1, \ldots, V_k \subseteq U$,
- feasible solutions $sol(U, V_1, \ldots, V_k) = \{\boldsymbol{s} \in \{1, \ldots, k\}^* \mid \bigcup_{i=1}^{|\boldsymbol{s}|} V_{\boldsymbol{s}[i]} = U\}$,
- a measure mes with $mes((U, V_1, \ldots, V_k), \boldsymbol{s}) = |\boldsymbol{s}|$ for all $\boldsymbol{s} \in sol(U, V_1, \ldots, V_k)$.

The usual reductions from SET-COVER to MIN-TEACHING-SET considered in the literature (cf., e.g., [13, 12, 18]) map an instance $(U, V_1, \ldots, V_k)$ of SET-COVER to the MIN-TEACHING-SET instance $(\{c_1, \ldots, c_{|U|}, c^*\}, c^*)$ in which all concepts are subsets of $\{1, \ldots, k\}$: $c_j = \{i \in \{1, \ldots, k\} \mid j \in V_i\}$ and $c^* = \emptyset$. Then the sets $V_{x_1}, \ldots, V_{x_\ell}$ cover U if and only if the examples $(x_1, 0), \ldots, (x_\ell, 0)$ constitute a teaching set for c^*.

On the other hand, every MIN-TEACHING-SET instance $(\{c_1, \ldots, c_n, c^*\}, c^*)$ over a domain $X = \{1, \ldots, k\}$ can be mapped to the SET-COVER instance $(U, V_1, \ldots, V_{|X|})$, where $U = \{1, \ldots, n-1\}$ and $V_i = \{j \mid c_i(j) \neq c^*(j)\}$ for $i = 1, \ldots, |(|X)$. The examples $(x_1, c^*(x_1)), \ldots, (x_\ell, c^*(x_\ell))$ then constitute a teaching set for c^* if and only if the sets $V_{x_1}, \ldots, V_{x_\ell}$ cover U.

Thus, instances of both problems can be transferred into one another, with the representation size of the instances changing only slightly. In particular, for two corresponding instances we have $|\mathcal{C}| - 1 = |U|$. This means that the result by Feige [29], who shows non-approximability for SET-COVER within a factor of $(1-\epsilon)\ln|U|$, can be formulated as follows.

Theorem 14. *The problem of computing a minimal teaching set is $\mathcal{NP}$-hard. It cannot be approximated in polynomial time within a factor of $(1-\epsilon)\ln(|\mathcal{C}|-1)$ for all $\epsilon > 0$, unless $\mathcal{NP} \subseteq DTime(n^{\log\log n})$.*

In a similar way, other non-approximability results for SET-COVER (cf., e.g., Raz and Safra [30]) can be rephrased for MIN-TEACHING-SET.

The problem corresponding to MIN-TEACHING-SET in the TD model would be the problem of finding an optimal teacher for $\mathcal{L}_\infty$ with feedback. Such a teacher has a representation size of $O(|X|^{|X|!\cdot|\mathcal{C}|})$ and is thus not polynomial in the representation size $|\mathcal{C}| \cdot |X|$ of the teaching problem. Finding such an optimal teacher is therefore not an $\mathcal{NPO}$ optimization problem (see Ausiello *et al.* [28]), and results about the hardness of SET-COVER cannot be immediately transferred. But a closer look at the proofs of these hardness results shows that also approximating this "set cover number" is hard in some sense. We demonstrate such a reasoning in the next theorem, which shows that E^+_∞ is hard to approximate within a factor of $\frac{1}{2}(1-\epsilon)\ln(|\mathcal{C}|-1)$ for any $\epsilon > 0$.

Theorem 15. *If there is a polynomial time algorithm computing for all finite classes $\mathcal{C}$ and concepts $c \in \mathcal{C}$ a rational number $A(c, \mathcal{C})$ such that*

$$\frac{E^+_\infty(c,\mathcal{C})}{\frac{1}{2}\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)}} \quad \leq \quad A(c,\mathcal{C}) \quad \leq \quad \tfrac{1}{2}\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)} \cdot E^+_\infty(c,\mathcal{C})$$

for some $\epsilon > 0$, then $\mathcal{NP} \subseteq DTime(n^{O(\log\log n)})$.

Proof. Suppose that there is such a polynomial time algorithm. Applying (18) we get

$$\frac{\mathrm{TD}(c,\mathcal{C})}{\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)}} \quad \leq \quad A(c,\mathcal{C}) \quad \leq \quad \tfrac{1}{2}\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)} \cdot \mathrm{TD}(c,\mathcal{C}).$$

Using the correspondence between SET-COVER and MIN-TEACHING-SET instances (see above), we conclude that there is also a polynomial time algorithm A' computing for every SET-COVER instance $(U, V_1, \ldots, V_k)$ a value $A'(U, V_1, \ldots, V_k)$ with

$$\frac{SC(U, V_1, \ldots, V_k)}{\sqrt{(1-\epsilon)\ln|U|}} \quad \leq \quad A'(U, V_1, \ldots, V_k) \quad \leq \quad \tfrac{1}{2}\sqrt{(1-\epsilon)\ln|U|} \cdot SC(U, V_1, \ldots, V_k)$$

where $SC(U, V_1, \ldots, V_k)$ is the minimal number of sets in $V_1, \ldots, V_k$ needed to cover U.

Now let Π be an $\mathcal{NP}$ decision problem and let Π^+ and Π^- be the set of positive and negative instances, respectively. Feige [29, Theorem 4.4] shows that it is possible to map every instance $\pi \in \Pi$ in time $n^{O(\log\log n)}$ to a SET-COVER instance $(U, V_1, \ldots, V_k)$ such that for an easily computable value Q:

$$\begin{aligned} \pi \in \Pi^+ &\iff SC(U, V_1, \ldots, V_k) \leq Q, \\ \pi \in \Pi^- &\iff SC(U, V_1, \ldots, V_k) > (1-\epsilon)\ln|U| \cdot Q. \end{aligned} \tag{19}$$

By checking the condition "$A'(U, V_1, \ldots, V_k) < \sqrt{(1-\epsilon)\ln|U|} \cdot Q$" one can decide whether π is a positive or negative instance for Π: Assume the condition holds; then $SC(U, V_1, \ldots, V_k) < (1-\epsilon)\ln|U| \cdot Q$ and by (19) we have $\pi \in \Pi^+$. Now let $\pi \in \Pi^+$; then by (19) $SC(U, V_1, \ldots, V_k) \leq Q$ and $A'(U, V_1, \ldots, V_k) \leq \frac{1}{2}\sqrt{(1-\epsilon)\ln|U|} \cdot SC(U, V_1, \ldots, V_k) \leq \frac{1}{2}\sqrt{(1-\epsilon)\ln|U|} \cdot Q < \sqrt{(1-\epsilon)\ln|U|} \cdot Q$.

It follows that it can be decided in time $n^{O(\log\log n)}$ whether any given π is a positive instance for the arbitrarily chosen $\mathcal{NP}$ problem Π. This means that $\mathcal{NP} \subseteq DTime(n^{O(\log\log n)})$. □

Although Lemma 12 is responsible for a negative result about the approximability of E^+_∞, we can also draw a positive conclusion from it: if the TD-value is known, there is often no need to compute the E^+_∞-value. For example, the teachabilities of concepts or classes for infinite memory learners with feedback can be compared by comparing the teaching dimensions of these concepts or classes.

Another consequence of Lemma 12 is that the optimal teaching times E^+_∞ and E^-_∞ of teaching with and without feedback differ by a factor of at most two. This means that feedback is not that much of a help when teaching randomized learners with infinite memory.

Finally, we briefly remark on the the hardness of approximating E^-_∞. From Lemma 12 we know that $\mathrm{TD}(c^*, \mathcal{C}) \geq E^-_\infty(c^*, \mathcal{C}) \geq \mathrm{TD}(c^*, \mathcal{C})/2$. We can thus prove an analog to Theorem 15.

Theorem 16. *If there is a polynomial time algorithm computing for all $c^*, \mathcal{C}$ a rational number $A(c^*, \mathcal{C})$ such that*

$$\frac{E^-_\infty(c, \mathcal{C})}{\frac{1}{2}\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)}} \quad \leq \quad A(c^*, \mathcal{C}) \quad \leq \quad \tfrac{1}{2}\sqrt{(1-\epsilon)\ln(|\mathcal{C}|-1)} \cdot E^-_\infty(c^*, \mathcal{C})$$

for some $\epsilon > 0$, then $\mathcal{NP} \subseteq DTime(n^{O(\log\log n)})$.

5. Teaching Positive Examples Only or Inconsistent Ones

5.1. Teaching Positive Examples Only

The learnability of classes from positive data is a typical question in learning theory (cf., e.g., Gold [31] as well as Osherson *et al.* [32]). 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 17. *Let $\mathcal{C}$ be a concept class and $c^* \in \mathcal{C}$ a target concept. Then for all learners $\mathcal{L}_\mu$, where $\mu \in \mathbb{N}^+ \cup \{\infty\}$, with or without feedback we have: The concept c^* is teachable from positive data if and only if 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 infinite 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. Consequently, with positive probability the number of examples is infinite; thus leading to an infinite expected number of examples. □

Theorem 17 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.

5.2. 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 if it is inconsistent with the target. 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 infinite 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.

But even in this scenario a few modifications of the definitions are necessary. First, we need to explain what the learner is supposed to do when given an example that is not consistent with any concept in $\mathcal{C}$, a situation that cannot arise with consistent teachers, since the target concept will always be consistent to any example given. We stipulate that in such a case, the learner simply does not change the hypothesis. This implies that such an example does not change the learner's state and is thus useless from a teacher's perspective.

The second minor change concerns the definition of success probability (cf. Definition 2). The limit mentioned in this definition does not need to exist if the teacher is inconsistent, because whenever it presents an inconsistent example the probability for the learner to assume the target hypothesis is zero. This is easily remedied by considering a teacher for which the limit does not exist as unsuccessful.

We consider inconsistent teachers only 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 6 shows that, when only positive data are allowed, inconsistent teachers can teach concepts to $\mathcal{L}_1$ with feedback that consistent teachers cannot. First, the teacher gives $(x_1, 1)$. If the learner guesses c^*, we are done. Otherwise, the learner must hypothesize 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 hypothesizes c^*. On the other hand, it follows from Theorem 17 that c^* cannot be taught by a consistent teacher from positive data, because $c_1 \supset c^*$.

$\mathcal{C}_1$:

	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

$\mathcal{C}_2$:

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

Figure 6: The class $\mathcal{C}_1$ can be taught to $\mathcal{L}_1$ with feedback by the inconsistent positive-data teacher T, but cannot be taught by a consistent positive-data teacher (Theorem 17). The class $\mathcal{C}_2$ cannot be taught by an inconsistent positive-data teacher (Fact 18).

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

Fact 18. *There is a class that cannot be taught to $\mathcal{L}_1$ with feedback by an inconsistent teacher from positive data.*

PROOF. We show that $\mathcal{C}_2$ from Figure 6 is such a class. Let $T\colon \mathcal{C}_2 \cup \{init\} \to \{x_1, x_2, x_3\} \times \{1\}$ be a teacher for $\mathcal{L}_1$ with feedback. 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 analogous), 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. Therefore the probability that $\mathcal{L}_1$ 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 $A = \{(c, d) \mid d \setminus c \neq \emptyset\}$, i.e., there is an arc from c to d iff there is a positive example inconsistent with c but consistent with d.

Theorem 19. *Let $\mathcal{C}$ be a concept class and $G(\mathcal{C}) = (V, A)$ its associated graph. For the learner $\mathcal{L}_1$ with feedback 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 that 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 19 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 $\mathcal{L}_1$ with feedback (cf. Figure 6), 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 17 and Theorem 20 below).

Theorem 20. *For the learner $\mathcal{L}_1$ without feedback 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 17.

For the only-if part suppose that c^* is teachable by a teacher T and there is a c with $c \supset c^*$.

Claim: T gives examples inconsistent with c^* only finitely often.

Proof. Suppose T gives an example $(x, 1) \notin \mathcal{X}(c^*)$ infinitely often. Without loss of generality we assume that there is a concept in $\mathcal{C}$ that contains x (otherwise $(x, 1)$ would be useless and a teacher never giving this example would be successful, too). Whenever $(x, 1)$ is presented, the learner will not be in state c^* afterwards, i.e., there are infinitely many t such that $\delta_T^{(t)}(c^*) = 0$. It follows that the limit $\lim_{t\to\infty} \delta_T^{(t)}(c^*)$ either does not exist or equals zero, which means that the teacher T is not successful; a contradiction that proves the claim. □ (Claim)

Let t' be the latest time point at which T presents an example inconsistent with c^* (the existence of t' has been proved by the claim). Therefore the learner is not in state c^* at round $t' + 1$. Moreover $T(t' + 1)$ will be consistent with c^* (by definition of t') and thus consistent with $c \supset c^*$ as well. For this reason, the learner will assume the hypothesis c in round $t' + 2$ with positive probability. Since all following examples are consistent with c^*, and thus with c, there is a positive probability that the learner will stay in c and never reach c^*. In other words, the success probability is less than one, a contradiction. This proves the only-if part. $\square$

6. 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. Intuitively, this holds, since the more examples the learner memorizes the higher the probability of reaching the target. Also, teachability depends on the learner's ability to give feedback, since this enables the teacher to trigger in every round a hypothesis change. Furthermore, as we have seen, teachability is influenced by 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.

As already mentioned, if the teacher receives no feedback the results about SSPPs, including Lemma 4, do not apply. The notion corresponding to this teaching scenario is that of an *unobservable stochastic shortest path problem* (USSPP). In this setting, it much more complicated to derive an optimality criterion and thus, we shall present results along this line of research in a subsequent paper and refer the reader to Balbach and Zeugmann [21] for a first insight. 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. But the resulting models seem to be much harder to analyze.

Furthermore, one could aim at modifying our model as follows. Instead of allowing the learner to pick a hypothesis uniformly at random from all consistent hypotheses, it may seem more "natural" to allow the learner to pick a hypothesis uniformly at random from all consistent hypotheses that are within a certain neighborhood of the learner's current hypothesis. Such an approach would require to define an appropriate neighborhood relation. As a matter of fact, we have already considered such a modification of the traditional (non-randomized) TD model (cf. [33, 22]). Again, in such models feedback can be very helpful. The order of examples is also crucial, and the learner's memory size influences the teaching time and also whether or not a concept is teachable at all. All these effects can be achieved by defining appropriate neighborhood relations. On the other hand, it will require further research to arrive at a notion of "natural" neighborhood relations.

Instead of considering the memory as queue as done here, one could and should also study learners that have a selective memory, i.e., they still memorize only μ examples but decide themselves which examples to keep in the memory. Models allowing such a selective memory have been studied in algorithmic learning theory and shown to be quite useful (cf., e.g., [34, 35]). Studies in this direction may also reveal more insight into the problem of how informative particular examples are, or more generally, into the question of how to measure the information content of a sample. In this regard, it should be mentioned that teaching sets should be considered as samples having a high information content. We did not focus on this connection in the present paper, since it seems that much more research is necessary to explore it.

Acknowledgment

We would like to thank the referees for many valuable comments and suggestions which have resulted in several improvements of the presentation of the paper.

References

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

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

[3] R. Freivalds, E. B. Kinber, R. Wiehagen, Inductive inference from good examples, in: Analogical and Inductive Inference, International Workshop AII '89, Reinhardsbrunn Castle, GDR, October 1989, Proceedings, Vol. 397 of Lecture Notes in Artificial Intelligence, Springer-Verlag, 1989, pp. 1–17.

[4] R. Freivalds, E. B. Kinber, R. Wiehagen, On the power of inductive inference from good examples, Theoret. Comput. Sci. 110 (1) (1993) 131–144.

[5] S. Jain, S. Lange, J. Nessel, On the learnability of recursively enumerable languages from good examples, Theoret. Comput. Sci. 261 (1) (2001) 3–29.

[6] J. Jackson, A. Tomkins, A computational model of teaching, in: Proc. 5th Annual ACM Workshop on Computational Learning Theory, ACM Press, New York, NY, 1992, pp. 319–326.

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

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

[9] D. Angluin, M. Kriķis, Teachers, learners and black boxes, in: Proceedings of the Tenth Annual Conference on Computational Learning Theory, July 6th–9th, 1997, Nashville, Tennessee, ACM Press, New York, NY, 1997, pp. 285–297.

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

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

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

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

[14] M. Anthony, G. Brightwell, J. Shawe-Taylor, On specifying Boolean functions by labelled examples, Discrete Applied Mathematics 61 (1) (1995) 1–25.

[15] F. J. Balbach, Teaching classes with high teaching dimension using few examples, in: Learning Theory: 18th Annual Conference on Learning Theory, COLT 2005, Bertinoro, Italy, June 27-30, 2005, Proceedings, Vol. 3559 of Lecture Notes in Artificial Intelligence, Springer, 2005, pp. 668–683.

[16] C. Kuhlmann, On teaching and learning intersection-closed concept classes, in: Computational Learning Theory, 4th European Conference, EuroCOLT '99, Nordkirchen, Germany, March 29-31, 1999, Proceedings, Vol. 1572 of Lecture Notes in Artificial Intelligence, Springer, 1999, pp. 168–182.

[17] H. Lee, R. A. Servedio, A. Wan, DNF are teachable in the average case, in: Learning Theory: 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 2006, Proceedings, Vol. 4005 of Lecture Notes in Artificial Intelligence, Springer, Berlin, 2006, pp. 214–228.

[18] M. Anthony, G. Brightwell, D. Cohen, J. Shawe-Taylor, On exact specification by examples, in: Proc. 5th Annual ACM Workshop on Computational Learning Theory, ACM Press, New York, NY, 1992, pp. 311–318.

[19] B. K. Natarajan, Machine Learning: A Theoretical Approach, Morgan Kaufmann, San Mateo, CA, 1991.

[20] E. Kushilevitz, N. Linial, Y. Rabinovich, M. Saks, Witness sets for families of binary vectors, Journal of Combinatorial Theory Series A 73 (2) (1996) 376–380.

[21] F. J. Balbach, T. Zeugmann, Teaching memoryless randomized learners without feedback, in: Algorithmic Learning Theory, 17th International Conference, ALT 2006, Barcelona, Spain, October 2006, Proceedings, Vol. 4264 of Lecture Notes in Artificial Intelligence, Springer, 2006, pp. 93–108.
[22] F. Balbach, Models for Algorithmic Teaching, Dissertation, Universität zu Lübeck, Lübeck, 2007.
[23] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, John Wiley & Sons, 1994.
[24] D. P. Bertsekas, Dynamic Programming and Optimal Control, Athena Scientific, 2001.
[25] D. P. Bertsekas, J. N. Tsitsiklis, An analysis of stochastic shortest path problems, Mathematics of Operations Research 16 (3) (1991) 580–595.
[26] S. D. Patek, On partially observed stochastic shortest path problems, in: Proceedings of the 40 th IEEE Conference on Decision and Control (CDC 2001). IEEE Part, 1999, pp. 5050–5055.
[27] S. D. Patek, Partially observed stochastic shortest path problems with approximate solution by neurodynamic programming, IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans 37 (5) (2007) 710–720.
[28] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, M. Protasi, Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties, Springer, New York, 1999.
[29] U. Feige, A threshold of ln n for approximating set cover, J. of the ACM 45 (4) (1998) 634–652.
[30] R. Raz, 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, ACM Press, 1997, pp. 475–484.
[31] E. M. Gold, Language identification in the limit, Inform. Control 10 (5) (1967) 447–474.
[32] D. N. Osherson, M. Stob, S. Weinstein, Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists, MIT Press, Cambridge, Massachusetts, 1986.
[33] F. J. Balbach, T. Zeugmann, Teaching learners with restricted mind changes, in: Algorithmic Learning Theory, 16th International Conference, ALT 2005, Singapore, October 2005, Proceedings, Vol. 3734 of Lecture Notes in Artificial Intelligence, Springer, 2005, pp. 474–489.
[34] J. Case, S. Jain, S. Lange, T. Zeugmann, Incremental concept learning for bounded data mining, Inform. Comput. 152 (1) (1999) 74–110.
[35] S. Lange, T. Zeugmann, Incremental learning from positive data, J. of Comput. Syst. Sci. 53 (1) (1996) 88–103.