JMLR: Workshop and Conference Proceedings vol 49:1–26, 2016

Preference-based Teaching

Ziyuan Gao GAO257@CS.UREGINA.CA
Department of Computer Science, University of Regina

Christoph Ries CHRISTOPH.RIES@RUB.DE
Department of Mathematics, Ruhr-University Bochum

Hans U. Simon HANS.SIMON@RUB.DE
Department of Mathematics, Ruhr-University Bochum

Sandra Zilles ZILLES@CS.UREGINA.CA
Department of Computer Science, University of Regina

Abstract

We introduce a new model of teaching named "preference-based teaching" and a corresponding complexity parameter—the preference-based teaching dimension (PBTD)—representing the worst-case number of examples needed to teach any concept in a given concept class. Although the PBTD coincides with the well-known recursive teaching dimension (RTD) on finite classes, it is radically different on infinite ones: the RTD becomes infinite already for trivial infinite classes (such as half-intervals) whereas the PBTD evaluates to reasonably small values for a wide collection of infinite classes including classes consisting of so-called closed sets w.r.t. a given closure operator, including various classes related to linear sets over $\mathbb{N}_0$ (whose RTD had been studied quite recently) and including the class of Euclidean half-spaces (and some other geometric classes). On top of presenting these concrete results, we provide the reader with a theoretical framework (of a combinatorial flavor) which helps to derive bounds on the PBTD.

Keywords: teaching dimension, preference relation, recursive teaching dimension

1. Introduction

The classical model of teaching (Shinohara and Miyano, 1991; Goldman and Kearns, 1995) formulates the following interaction protocol between a teacher and a student:

- Both of them agree on a "classification-rule system", formally given by a concept class $\mathcal{L}$.
- In order to teach a specific concept $L \in \mathcal{L}$, the teacher presents to the student a *teaching set*, i.e., a set T of labeled examples so that L is the only concept in $\mathcal{L}$ that is consistent with T.
- The student determines L as the unique concept in $\mathcal{L}$ that is consistent with T.

Goldman and Mathias (1996) pointed out that this model of teaching is not powerful enough, since the teacher is required to make *any* consistent learner successful. A challenge is to model powerful teacher/student interactions without enabling unfair "coding tricks". Goldman and Mathias hence defined a notion of "valid teacher/learner pair" that is intuitively free of coding tricks while allowing for a much broader class of interaction protocols than the original teaching model. In particular, teaching may thus become more efficient in terms of the number of examples in the teaching sets. Further definitions of how to avoid unfair coding tricks have been suggested (Zilles et al., 2011), but they were less stringent than the one proposed by Goldman and Mathias.

The model of recursive teaching (Zilles et al., 2011; Mazadi et al., 2014), which is free of coding tricks according to the Goldman-Mathias definition, has recently gained attention because its complexity parameter, the recursive teaching dimension (RTD), has shown relations to the VC-dimension and to sample compression (Doliwa et al., 2014; Moran et al., 2015; Simon and Zilles, 2015), when focusing on finite concept classes. Below though we will give examples of rather simple infinite concept classes with infinite RTD, suggesting that the RTD is inadequate for addressing the complexity of teaching infinite classes.

In this paper, we introduce a model called *preference-based teaching*, in which the teacher and the student do not only agree on a classification-rule system $\mathcal{L}$ but also on a preference relation (a strict partial order) imposed on $\mathcal{L}$. If the labeled examples presented by the teacher allow for several consistent explanations (= consistent concepts) in $\mathcal{L}$, the student will choose a concept $L \in \mathcal{L}$ that she prefers most. This gives more flexibility to the teacher than the classical model: the set of labeled examples need not distinguish a target concept L from any other concept in $\mathcal{L}$ but only from those concepts L' over which L is not preferred. At the same time, preference-based teaching yields valid teacher/learner pairs according to Goldman and Mathias's definition. We will show that the new model, despite avoiding coding tricks, is quite powerful. Moreover, as we will see in the course of the paper, it often allows for a very natural design of teaching sets.

Assume teacher and student choose a preference relation that minimizes the worst-case number M of examples required for teaching any concept in the class $\mathcal{L}$. This number M is then called the preference-based teaching dimension (PBTD) of $\mathcal{L}$. In particular, we will show the following:

(i) Recursive teaching is a special case of preference-based teaching where the preference relation satisfies a so-called "finite-depth condition". It is precisely this additional condition that renders recursive teaching useless for many natural and apparently simple infinite concept classes. Preference-based teaching successfully addresses these shortcomings of recursive teaching, see Section 3. For finite classes, PBTD and RTD are equal.

(ii) A wide collection of geometric and algebraic concept classes with infinite RTD can be taught very efficiently, i.e., with low PBTD. To establish such results, we show in Section 4 that spanning sets can be used as preference-based teaching sets with positive examples only — a result that is very simple to obtain but quite useful.

(iii) In the preference-based model, linear sets over $\mathbb{N}_0$ with origin 0 and at most k generators can be taught with k positive examples, while recursive teaching with a bounded number of positive examples was previously shown to be impossible and it is unknown whether recursive teaching with a bounded number of positive and negative examples is possible for $k \geq 4$. We also give some almost matching upper and lower bounds on the PBTD for other classes of linear sets, see Section 6.

(iv) The PBTD of halfspaces in $\mathbb{R}^d$ is upper-bounded by $2d + 2$ (see Sections 7 and 8), while its RTD is infinite. This result is based on the design of a lexicographic preference-relation that can be described by a hierarchical rule system.

Moreover, in Section 9, we compute (bounds on) the PBTD of some geometric concept classes when both positive and negative examples are used for teaching. Based on our results and the naturalness of the teaching sets and preference relations used in their proofs, we claim that preference-based teaching is far more suitable to the study of infinite concept classes than recursive teaching.

To keep the main body of the paper accessible, some formal proofs are given in the appendix.

2. Basic Definitions and Facts

$\mathbb{N}_0$ denotes the set of all non-negative integers and $\mathbb{N}$ denotes the set of all positive integers. A *concept class* $\mathcal{L}$ is a family of subsets over a universe $\mathcal{X}$, i.e., $\mathcal{L} \subseteq 2^{\mathcal{X}}$ where $2^{\mathcal{X}}$ denotes the powerset of $\mathcal{X}$. The elements of $\mathcal{L}$ are called *concepts*. A *labeled example* is an element of $\mathcal{X} \times \{+, -\}$. Elements of $\mathcal{X}$ are called *examples*. Suppose that T is a set of labeled examples. Let $T^+ = \{x \in \mathcal{X} : (x, +) \in T\}$ and $T^- = \{x \in \mathcal{X} : (x, -) \in T\}$. A set $L \subseteq \mathcal{X}$ is *consistent with* T if it includes all examples in T that are labeled "$+$" and excludes all examples in T that are labeled "$-$", i.e, if $T^+ \subseteq L$ and $T^- \cap L = \emptyset$. A set of labeled examples that is consistent with L but not with L' is said to *distinguish* L *from* L'. The classical model of teaching is then defined as follows.

Definition 1 (Shinohara and Miyano (1991); Goldman and Kearns (1995)) *A* teaching set for $L \in \mathcal{L}$ w.r.t. $\mathcal{L}$ *is a set T of labeled examples such that L is the only concept in $\mathcal{L}$ that is consistent with T, i.e., T distinguishes L from any other concept in $\mathcal{L}$. Define $TD(L, \mathcal{L}) = \inf\{|T| : T$ is a teaching set for L w.r.t. $\mathcal{L}\}$. i.e., $TD(L, \mathcal{L})$ is the smallest possible size of a teaching set for L w.r.t. $\mathcal{L}$. If L has no finite teaching set w.r.t. $\mathcal{L}$, then $TD(L, \mathcal{L}) = \infty$. The number $TD(\mathcal{L}) = \sup_{L \in \mathcal{L}} TD(L, \mathcal{L}) \in \mathbb{N}_0 \cup \{\infty\}$ is called the* teaching dimension of $\mathcal{L}$.

For technical reasons, we will occasionally deal with the number $\mathrm{TD}_{min}(\mathcal{L}) = \inf_{L \in \mathcal{L}} \mathrm{TD}(L, \mathcal{L})$, i.e., the number of examples needed to teach the concept from $\mathcal{L}$ that is easiest to teach.

In this paper, we will examine a teaching model in which the teacher and the student do not only agree on a classification-rule system $\mathcal{L}$ but also on a preference relation, denoted as $\prec$, imposed on $\mathcal{L}$. We assume that $\prec$ is a *strict partial order* on $\mathcal{L}$, i.e., $\prec$ is asymmetric and transitive. The partial order that makes every pair $L \neq L' \in \mathcal{L}$ incomparable is denoted by $\prec_\emptyset$. For every $L \in \mathcal{L}$, let

$$\mathcal{L}_{\prec L} = \{L' \in \mathcal{L} : L' \prec L\}$$

be the set of concepts over which L is strictly preferred. Note that $\mathcal{L}_{\prec_\emptyset L} = \emptyset$ for every $L \in \mathcal{L}$.

As already noted above, a teaching set T of L w.r.t. $\mathcal{L}$ distinguishes L from any other concept in $\mathcal{L}$. If a preference relation comes into play, then T will be exempted from the obligation to distinguish L from the concepts in $\mathcal{L}_{\prec L}$ because L is strictly preferred over them anyway.

Definition 2 *A* teaching set for $L \subseteq X$ w.r.t. $(\mathcal{L}, \prec)$ *is defined as a teaching set for L w.r.t. $\mathcal{L} \setminus \mathcal{L}_{\prec L}$. Furthermore define*

$$\mathrm{PBTD}(L, \mathcal{L}, \prec) = \inf\{|T| : T \text{ is a teaching set for } L \text{ w.r.t. } (\mathcal{L}, \prec)\} \in \mathbb{N}_0 \cup \{\infty\} \enspace .$$

The number $\mathrm{PBTD}(\mathcal{L}, \prec) = \sup_{L \in \mathcal{L}} \mathrm{PBTD}(L, \mathcal{L}, \prec) \in \mathbb{N}_0 \cup \{\infty\}$ *is called the* teaching dimension of $(\mathcal{L}, \prec)$.

Definition 2 implies that

$$\mathrm{PBTD}(L, \mathcal{L}, \prec) = \mathrm{TD}(L, \mathcal{L} \setminus \mathcal{L}_{\prec L}) \enspace . \tag{1}$$

Let $L \mapsto T(L)$ be a mapping that assigns a teaching set for L w.r.t. $(\mathcal{L}, \prec)$ to every $L \in \mathcal{L}$. It is obvious from Definition 2 that T must be injective, i.e., $T(L) \neq T(L')$ if L and L' are distinct concepts from $\mathcal{L}$. The classical model of teaching is obtained from the model described in Definition 2 when we plug in the empty preference relation $\prec_\emptyset$ for $\prec$. In particular, $\mathrm{PBTD}(\mathcal{L}, \prec_\emptyset) = \mathrm{TD}(\mathcal{L})$.

To avoid unfair coding tricks, Goldman and Mathias (1996) required that the student identify a concept $L \in \mathcal{L}$ even from any superset $S \supseteq T(L)$ of the set $T(L)$ presented by the teacher, as long as L is consistent with S. It is easy to see that preference-based teaching fulfills this requirement.

We are interested in finding the partial order that is optimal for the purpose of teaching and we aim at determining the corresponding teaching dimension. This motivates the following notion:

Definition 3 *The* preference-based teaching dimension of $\mathcal{L}$ *is given by*

$$\mathrm{PBTD}(\mathcal{L}) = \inf\{\mathrm{PBTD}(\mathcal{L}, \prec) : \prec \textit{ is a strict partial order on } \mathcal{L}\} \enspace .$$

A relation R' on $\mathcal{L}$ is said to be an *extension of a relation* R if $R \subseteq R'$. The *order-extension principle* states that any partial order has a linear extension (Jech, 1973). The following result (whose second assertion follows from the first one in combination with the order-extension principle) is pretty obvious:

Lemma 4

1. *Suppose that $\prec'$ extends $\prec$. If T is a teaching set for L w.r.t. $(\mathcal{L}, \prec)$, then T is a teaching set for L w.r.t. $(\mathcal{L}, \prec')$. Moreover* $\mathrm{PBTD}(\mathcal{L}, \prec') \le \mathrm{PBTD}(\mathcal{L}, \prec)$.
2. $\mathrm{PBTD}(\mathcal{L}) = \inf\{\mathrm{PBTD}(\mathcal{L}, \prec) : \prec$ *is a strict linear order on* $\mathcal{L}\}$.

Preference-based teaching with positive examples only. Suppose that $\mathcal{L}$ contains two concepts L, L' such that $L \subset L'$. In the classical teaching model, any teaching set for L w.r.t. $\mathcal{L}$ has to employ a negative example in order to distinguish L from L'. Symmetrically, any teaching set for L' w.r.t. $\mathcal{L}$ has to employ a positive example. Thus classical teaching cannot be performed with one type of examples only unless $\mathcal{L}$ is an antichain w.r.t. inclusion. As for preference-based teaching, the restriction to one type of examples is much less severe, as our results below will show.

A teaching set T for $L \in \mathcal{L}$ w.r.t. $(\mathcal{L}, \prec)$ is said to be *positive* if it does not make use of negatively labeled examples, i.e., if $T^- = \emptyset$. In the sequel, we will occasionally identify a positive teaching set T with T^+. A positive teaching set for L w.r.t. $(\mathcal{L}, \prec)$ can clearly not distinguish L from a proper superset of L in $\mathcal{L}$. Thus, the following holds:

Lemma 5 *Suppose that $L \mapsto T^+(L)$ maps each $L \in \mathcal{L}$ to a positive teaching set for L w.r.t. $(\mathcal{L}, \prec)$. Then $\prec$ must be an extension of $\supset$ (so that proper subsets of a set L are strictly preferred over L) and, for every $L \in \mathcal{L}$, the set $T^+(L)$ must distinguish L from every proper subset of L in $\mathcal{L}$.*

Define

$$\mathrm{PBTD}^+(L, \mathcal{L}, \prec) = \inf\{|T| : T \text{ is a positive teaching set for } L \text{ w.r.t. } (\mathcal{L}, \prec)\} \enspace . \tag{2}$$

The number $\mathrm{PBTD}^+(\mathcal{L}, \prec) = \sup_{L \in \mathcal{L}} \mathrm{PBTD}^+(L, \mathcal{L}, \prec)$ (possibly ∞) is called the *positive teaching dimension of* $(\mathcal{L}, \prec)$. The *positive preference-based teaching dimension of* $\mathcal{L}$ is then given by

$$\mathrm{PBTD}^+(\mathcal{L}) = \inf\{\mathrm{PBTD}^+(\mathcal{L}, \prec) : \prec \text{ is a strict partial order on } \mathcal{L}\} \enspace . \tag{3}$$

Monotonicity. A complexity measure K that assigns a number $K(\mathcal{L}) \in \mathbb{N}_0$ to a concept class $\mathcal{L}$ is said to be *monotonic* if $\mathcal{L}' \subseteq \mathcal{L}$ implies that $K(\mathcal{L}') \leq K(\mathcal{L})$. It is well known (and trivial to see) that TD is monotonic. It is fairly obvious that PBTD is monotonic, too:

Lemma 6 PBTD *and* PBTD^+ *are monotonic.*

As an application of monotonicity, we show the following result:

Lemma 7 *For every finite subclass $\mathcal{L}'$ of $\mathcal{L}$, we have* $\mathrm{PBTD}(\mathcal{L}) \geq \mathrm{PBTD}(\mathcal{L}') \geq TD_{min}(\mathcal{L}')$.

Proof The first inequality holds because PBTD is monotonic. The second inequality follows from the fact that a finite partially ordered set must contain a minimal element. Thus, for any fixed choice of $\prec$, $\mathcal{L}'$ must contain a concept L' such that $\mathcal{L}'_{\prec L'} = \emptyset$. Hence,

$$\mathrm{PBTD}(\mathcal{L}', \prec) \geq \mathrm{PBTD}(L', \mathcal{L}', \prec) \overset{(1)}{=} \mathrm{TD}(L', \mathcal{L}' \setminus \mathcal{L}'_{\prec L'}) = \mathrm{TD}(L', \mathcal{L}') \geq \mathrm{TD}_{min}(\mathcal{L}') \ .$$

Since this holds for any choice of $\prec$, we get $\mathrm{PBTD}(\mathcal{L}') \geq \mathrm{TD}_{min}(\mathcal{L}')$, as desired. ■

3. Preference-based versus Recursive Teaching

The preference-based teaching dimension is a relative of the recursive teaching dimension. In fact, both notions coincide on finite classes, as we will see shortly. We first recall the definitions of the recursive teaching dimension and of some related notions (Zilles et al., 2011; Mazadi et al., 2014).

A *teaching sequence for* $\mathcal{L}$ is a sequence of the form $\mathcal{S} = (\mathcal{L}_i, d_i)_{i \geq 1}$ where $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \ldots$ form a partition of $\mathcal{L}$ into non-empty sub-classes and, for every $i \geq 1$, we have that

$$d_i = \sup_{L \in \mathcal{L}_i} \mathrm{TD}\left(L, \mathcal{L} \setminus \cup_{j=1}^{i-1} \mathcal{L}_j\right) \ . \tag{4}$$

If, for every $i \geq 1$, d_i is the supremum over all $L \in \mathcal{L}_i$ of the smallest size of a *positive teaching set* for L w.r.t. $\cup_{j \geq i} \mathcal{L}_j$ (and $d_i = \infty$ if some $L \in \mathcal{L}_i$ does not have a positive teaching set w.r.t. $\cup_{j \geq i} \mathcal{L}_j$), then $\mathcal{S}$ is said to be a *positive teaching sequence for* $\mathcal{L}$. The *order* of a teaching sequence or a positive teaching sequence $\mathcal{S}$ (possibly ∞) is defined as $\mathrm{ord}(\mathcal{S}) = \sup_{i \geq 1} d_i$. The *recursive teaching dimension of* $\mathcal{L}$ (possibly ∞) is defined as the order of the teaching sequence of lowest order for $\mathcal{L}$. More formally, $\mathrm{RTD}(\mathcal{L}) = \inf_{\mathcal{S}} \mathrm{ord}(\mathcal{S})$ where $\mathcal{S}$ ranges over all teaching sequences for $\mathcal{L}$. Similarly, $\mathrm{RTD}^+(\mathcal{L}) = \inf_{\mathcal{S}} \mathrm{ord}(\mathcal{S})$, where $\mathcal{S}$ ranges over all positive teaching sequences for $\mathcal{L}$. Note that the following holds for every $\mathcal{L}' \subseteq \mathcal{L}$ and for every teaching sequence $\mathcal{S}$ for $\mathcal{L}'$ such that $\mathrm{ord}(\mathcal{S}) = \mathrm{RTD}(\mathcal{L}')$:

$$\mathrm{RTD}(\mathcal{L}) \geq \mathrm{RTD}(\mathcal{L}') \geq \mathrm{ord}(\mathcal{S}) \geq d_1 = \sup_{L \in \mathcal{L}_1} \mathrm{TD}(L, \mathcal{L}') \geq \mathrm{TD}_{min}(\mathcal{L}') \ . \tag{5}$$

The *depth* of $L \in \mathcal{L}$ w.r.t. a strict partial order imposed on $\mathcal{L}$ is defined as the length of the longest chain in $(\mathcal{L}, \prec)$ that ends in L (resp. as ∞ if there is no bound on the length of these chains). The recursive teaching dimension is related to the preference-based teaching dimension as follows:

Lemma 8 $\mathrm{RTD}(\mathcal{L}) = \inf_{\prec} \mathrm{PBTD}(\mathcal{L}, \prec)$ *and* $\mathrm{RTD}^+(\mathcal{L}) = \inf_{\prec} \mathrm{PBTD}^+(\mathcal{L}, \prec)$ *where* $\prec$ *ranges over all strict partial orders on* $\mathcal{L}$ *that satisfy the following "finite-depth condition": every* $L \in \mathcal{L}$ *has a finite depth w.r.t.* $\prec$.

The following is an immediate consequence of Lemma 8 and the trivial observation that the finite-depth condition is always satisfied if $\mathcal{L}$ is finite:

Corollary 9 $\mathrm{PBTD}(\mathcal{L}) \le \mathrm{RTD}(\mathcal{L})$, *with equality if* $\mathcal{L}$ *is finite.*

As for infinite classes, the gap between PBTD and RTD can be arbitrarily large:

Lemma 10 *There exists an infinite class* $\mathcal{L}_\infty$ *of VC-dimension* 1 *such that* $\mathrm{PBTD}^+(\mathcal{L}_\infty) = 1$ *and* $\mathrm{RTD}(\mathcal{L}_\infty) = \infty$. *Moreover, for every* $k \ge 1$, *there exists an infinite class* $\mathcal{L}_k$ *such that* $\mathrm{PBTD}^+(\mathcal{L}_k) = 1$ *and* $\mathrm{RTD}(\mathcal{L}_k) = k$.

The complete proof of Lemma 10 is given in Appendix A. Here we only specify the classes $\mathcal{L}_\infty$ and $\mathcal{L}_k$ that are employed in this proof:

- Choose $\mathcal{L}_\infty$ as the class of half-intervals $[0, a]$, where $0 \le a < 1$, over the universe $[0, 1)$.[1]
- Let $\mathcal{X} = [0, 2)$. For each $a = \sum_{n \ge 1} \alpha_n 2^{-n} \in [0, 1)$ and for all $i = 1, \dots, k$, let $1 \le a_i < 2$ be given by $a_i = 1 + \sum_{n \ge 0} \alpha_{kn+i} 2^{-n}$. Finally, let $I_a = [0, a) \cup \{a_1, \dots, a_k\} \subseteq \mathcal{X}$ and let $\mathcal{L}_k = \{I_a : 0 \le a < 1\}$.

.

4. Teaching with Positive Examples Only

The main purpose of this section is to relate positive preference-based teaching to "spanning sets" and "closure operators", which are well-studied concepts in the computational learning theory literature. For any subsets S and L of the universe, we say that S is a *spanning set* of L w.r.t. $\mathcal{L}$ if $S \subseteq L$ and any set in $\mathcal{L}$ that contains S must contain L as well. [2] In other words, L is the unique smallest concept in $\mathcal{L}$ that contains S. We say that S is a *weak spanning set* of L w.r.t. $\mathcal{L}$ if S is not contained in any proper subset of L in $\mathcal{L}$. We denote by $I(\mathcal{L})$ (resp. $I'(\mathcal{L})$) the smallest number k such that every concept $L \in \mathcal{L}$ has a spanning set (resp. a weak spanning set) w.r.t. $\mathcal{L}$ of size at most k. Note that S is a spanning set of L w.r.t. $\mathcal{L}$ iff S distinguishes L from all concepts in $\mathcal{L}$ except for supersets of L, i.e., iff S is a positive teaching set for L w.r.t. $(\mathcal{L}, \supset)$. Similarly, S is a weak spanning set of L w.r.t. $\mathcal{L}$ iff S distinguishes L from all its proper subsets in $\mathcal{L}$ (which is necessarily the case when S is a positive teaching set). These observations can be summarized as follows:

$$I'(\mathcal{L}) \le \mathrm{PBTD}^+(\mathcal{L}) \le \mathrm{PBTD}^+(\mathcal{L}, \supset) \le I(\mathcal{L}) \ . \tag{6}$$

Suppose $\mathcal{L}$ is intersection-closed. Then $\cap_{L \in \mathcal{L}: S \subseteq L} L$ is the unique smallest concept in $\mathcal{L}$ containing S. If $S \subseteq L_0$ is a weak spanning set of $L_0 \in \mathcal{L}$, then $\cap_{L \in \mathcal{L}: S \subseteq L} L = L_0$ because, on the one hand, $\cap_{L \in \mathcal{L}: S \subseteq L} L \subseteq L_0$ and, on the other hand, no proper subset of L_0 in $\mathcal{L}$ contains S. Thus the distinction between spanning sets and weak spanning sets is blurred for intersection-closed classes:

1. $\mathrm{RTD}(\mathcal{L}_\infty) = \infty$ had been observed by Moran et al. (2015) already.
2. This generalizes the classical definition of a spanning set (Helmbold et al., 1990), which is given w.r.t. intersection-closed classes only.

Lemma 11 *Suppose that $\mathcal{L}$ is intersection-closed. Then $I'(\mathcal{L}) = \mathrm{PBTD}^+(\mathcal{L}) = I(\mathcal{L})$.*

Example 1 *Let $\mathcal{R}_d$ denote the class of d-dimensional axis-parallel hyper-rectangles (= d-dimensional boxes). This class is intersection-closed and clearly $I(\mathcal{R}_d) = 2$. Thus $\mathrm{PBTD}^+(\mathcal{R}_d) = 2$.*

A mapping $\mathrm{cl} : 2^{\mathcal{X}} \rightarrow 2^{\mathcal{X}}$ is said to be a *closure operator* on the universe $\mathcal{X}$ if the following conditions hold for all sets $A, B \subseteq \mathcal{X}$:

$$A \subseteq B \Rightarrow \mathrm{cl}(A) \subseteq \mathrm{cl}(B) \text{ and } A \subseteq \mathrm{cl}(A) = \mathrm{cl}(\mathrm{cl}(A)) \ .$$

The following notions refer to an arbitrary but fixed closure operator. The set $\mathrm{cl}(A)$ is called the *closure* of A. A set C is said to be *closed* if $\mathrm{cl}(C) = C$. It follows that precisely the sets $\mathrm{cl}(A)$ with $A \subseteq \mathcal{X}$ are closed. Let $\mathcal{C}$ denote the set of all closed subsets of $\mathcal{X}$. Then $L = \mathrm{cl}(S)$ and $S \subseteq L'$ for $L' \in \mathcal{C}$ implies that $L = \mathrm{cl}(S) \subseteq \mathrm{cl}(L') = L'$. Thus we obtain the following result:

Lemma 12 *If $L = \mathrm{cl}(S)$, then S is a spanning set of L w.r.t. $\mathcal{C}$.*

For every closed set $L \in \mathcal{L}$, let $s_{cl}(L)$ denote the size (possibly ∞) of the smallest set $S \subseteq \mathcal{X}$ such that $\mathrm{cl}(S) = L$. With this notation, we get the following (trivial but useful) result:

Theorem 13 *Given a closure operator, let $\mathcal{C}[m]$ be the class of all closed subsets $C \subseteq \mathcal{X}$ with $s_{cl}(C) \leq m$. Then $\mathrm{PBTD}^+(\mathcal{C}[m]) \leq \mathrm{PBTD}^+(\mathcal{C}[m], \supset) \leq m$. Moreover, this holds with equality provided that $\mathcal{C}[m] \setminus \mathcal{C}[m-1] \neq \emptyset$.*

Proof The inequality $\mathrm{PBTD}^+(\mathcal{C}[m], \supset) \leq m$ follows directly from Equation 6 and Lemma 12. Pick a concept $C_0 \in \mathcal{C}[m]$ such that $s_{cl}(C_0) = m$. Then any subset S of C_0 of size less than m spans only a proper subset of C_0, i.e., $\mathrm{cl}(S) \subset C_0$. Thus S does not distinguish C_0 from $\mathrm{cl}(S)$. It follows that there is no positive teaching set of size less than m for C_0 w.r.t. $\mathcal{C}[m]$. ■

Many natural classes can be cast as classes of the form $\mathcal{C}[m]$ by choosing the universe and the closure operator appropriately.

Example 2 *Let*

$$\mathit{LINSET}_k = \{\langle G \rangle : (G \subset \mathbb{N}) \wedge (1 \leq |G| \leq k)\}$$

where $\langle G \rangle = \left\{\sum_{g \in G} a(g)g : a(g) \in \mathbb{N}_0\right\}$. Note that the mapping $G \mapsto \langle G \rangle$ is a closure operator over the universe $\mathbb{N}_0$. Since obviously $\mathit{LINSET}_k \setminus \mathit{LINSET}_{k-1} \neq \emptyset$, we obtain $\mathrm{PBTD}^+(\mathit{LINSET}_k) = k$.
Let $\mathcal{X} = \mathbb{R}^2$ and let $\mathrm{cl}(S)$ be the convex closure of S. Then $\mathcal{C}[k]$ is the class of convex polygons with at most k vertices. It follows that $\mathrm{PBTD}^+(\mathcal{C}[k]) = k$.

5. A Convenient Technique for Proving Upper Bounds

In this section, we shall give an alternative definition of the preference-based teaching dimension using the notion of an "admissible mapping". Given a concept class $\mathcal{L}$ over a universe $\mathcal{X}$, let T be a mapping $L \mapsto T(L) \subseteq \mathcal{X} \times \{+, -\}$ that assigns a set $T(L)$ of labeled examples to every set $L \in \mathcal{L}$. The *order* of T, denoted as $\mathrm{ord}(T)$, is defined as $\sup_{L \in \mathcal{L}} |T(L)| \in \mathbb{N} \cup \{\infty\}$. Define the mappings

T^+ and T^- by setting $T^+(L) = \{x : (x,+) \in T(L)\}$ and $T^-(L) = \{x : (x,-) \in T(L)\}$ for every $L \in \mathcal{L}$. We say that T is *positive* if $T^-(L) = \emptyset$ for every $L \in \mathcal{L}$. In the sequel, we will occasionally identify a positive mapping $L \mapsto T(L)$ with the mapping $L \mapsto T^+(L)$. The symbol "$+$" as an upper index of T will always indicate that the underlying mapping T is positive.
Consider the following relation R_T on $\mathcal{L}$:

$$R_T = \{(L', L) \in \mathcal{L} \times \mathcal{L} : \ (L' \neq L) \wedge (L' \text{ is consistent with } T(L))\} \ .$$

The transitive closure of R_T is denoted as R_T^+ in the sequel. The following notion will play an important role in this paper:

Definition 14 *A mapping $L \mapsto T(L)$ with L ranging over all concepts in $\mathcal{L}$ is said to be* admissible *for $\mathcal{L}$ if the following holds:*

1. *For every $L \in \mathcal{L}$, L is consistent with $T(L)$.*
2. *The relation R_T^+ is asymmetric (which clearly implies that R_T is asymmetric too).*

If T is admissible, then R_T^+ is transitive and asymmetric, i.e., R_T^+ is a strict partial order on $\mathcal{L}$. We will therefore use the notation $\prec_T$ instead of R_T^+ whenever T is known to be admissible.

Lemma 15 *Suppose that T^+ is a positive admissible mapping for $\mathcal{L}$. Then the relation $\prec_{T^+}$ on $\mathcal{L}$ extends the relation $\supset$ on $\mathcal{L}$. More precisely, the following holds for all $L, L' \in \mathcal{L}$:*

$$L' \subset L \Rightarrow (L, L') \in R_{T^+} \Rightarrow L \prec_{T^+} L' \ .$$

Proof If T^+ is admissible, then L' is consistent with $T^+(L')$. Thus $T^+(L') \subseteq L' \subset L$ so that L is consistent with $T^+(L')$ too. Therefore $(L, L') \in R_{T^+}$, i.e., $L \prec_{T^+} L'$. ■

The following result clarifies how admissible mappings are related to preference-based teaching:

Lemma 16 *For each concept class $\mathcal{L}$, the following holds:*

$$\mathrm{PBTD}(\mathcal{L}) = \inf_T \mathrm{ord}(T) \ \text{ and } \ \mathrm{PBTD}^+(\mathcal{L}) = \inf_{T^+} \mathrm{ord}(T^+)$$

where T ranges over all mappings that are admissible for $\mathcal{L}$ and T^+ ranges over all positive mappings that are admissible for $\mathcal{L}$.

Proof We restrict ourselves to the proof for $\mathrm{PBTD}(\mathcal{L}) = \inf_T \mathrm{ord}(T)$ because the equation $\mathrm{PBTD}^+(\mathcal{L}) = \inf_{T^+} \mathrm{ord}(T^+)$ can be obtained in a similar fashion. We first prove that $\mathrm{PBTD}(\mathcal{L}) \leq \inf_T \mathrm{ord}(T)$. Let T be an admissible mapping for $\mathcal{L}$. It suffices to show that, for every $L \in \mathcal{L}$, $T(L)$ is a teaching set for L w.r.t. $(\mathcal{L}, \prec_T)$. Suppose $L' \in \mathcal{L} \setminus \{L\}$ is consistent with $T(L)$. Then $(L', L) \in R_T$ and thus $L' \prec_T L$. It follows that $\prec_T$ prefers L over all concepts $L' \in \mathcal{L} \setminus \{L\}$ that are consistent with $T(L)$. Thus T is a teaching set for L w.r.t. $(\mathcal{L}, \prec_T)$, as desired.
We now prove that $\inf_T \mathrm{ord}(T) \leq \mathrm{PBTD}(\mathcal{L})$. Let $\prec$ be a strict partial order on $\mathcal{L}$ and let T be a mapping such that, for every $L \in \mathcal{L}$, $T(L)$ is a teaching set for L w.r.t. $(\mathcal{L}, \prec)$. It suffices to show that T is admissible for $\mathcal{L}$. Consider a pair $(L', L) \in R_T$. The definition of R_T implies that $L' \neq L$ and that L' is consistent with $T(L)$. Since $T(L)$ is a teaching set w.r.t. $(\mathcal{L}, \prec)$, it follows that $L' \prec L$. Thus, $\prec$ is an extension of R_T. Since $\prec$ is transitive, it is even an extension of R_T^+. Because $\prec$ is asymmetric, R_T^+ must be asymmetric, too. It follows that T is admissible. ■

6. Preference-based Teaching of Linear Sets

In this section, we consider several concept classes over the universe $\mathcal{X} = \mathbb{N}_0$. Let $G = \{g_1, \dots, g_k\}$ be a finite subset of $\mathbb{N}$. We denote by $\langle G \rangle$ resp. by $\langle G \rangle_+$ the following sets:

$$\langle G \rangle = \left\{ \sum_{i=1}^{k} a_i g_i : \ a_1, \dots, a_k \in \mathbb{N}_0 \right\} \ \text{ and } \ \langle G \rangle_+ = \left\{ \sum_{i=1}^{k} a_i g_i : \ a_1, \dots, a_k \in \mathbb{N} \right\} \ .$$

In this section, we will determine (at least approximately) the preference-based teaching dimension of the following concept classes over $\mathbb{N}_0$:

$$\begin{aligned}
\text{LINSET}_k &= \{\langle G \rangle : \ (G \subset \mathbb{N}) \wedge (1 \le |G| \le k)\} \ . \\
\text{CF-LINSET}_k &= \{\langle G \rangle : \ (G \subset \mathbb{N}) \wedge (1 \le |G| \le k) \wedge (\gcd(G) = 1)\} \ . \\
\text{NE-LINSET}_k &= \{\langle G \rangle_+ : \ (G \subset \mathbb{N}) \wedge (1 \le |G| \le k)\} \ . \\
\text{NE-CF-LINSET}_k &= \{\langle G \rangle_+ : \ (G \subset \mathbb{N}) \wedge (1 \le |G| \le k) \wedge (gcd(G) = 1)\} \ .
\end{aligned}$$

A subset of $\mathbb{N}_0$ whose complement in $\mathbb{N}_0$ is finite is said to be *co-finite*. The letters "CF" in CF-LINSET mean "co-finite". The concepts in LINSET_k have the algebraic structure of a monoid w.r.t. addition. The concepts in CF-LINSET_k are also known as "numerical semigroups" (Rosales and García-Sánchez, 2009). A zero coefficient $a_j = 0$ erases g_j within the linear combination $\sum_{i=1}^{k} a_i g_i$. Coefficients from $\mathbb{N}$ are non-erasing in this sense. The letters "NE" in "NE-LINSET" mean "non-erasing".
The *shift-extension* $\mathcal{L}'$ of a concept class $\mathcal{L}$ over the universe $\mathbb{N}_0$ is defined as follows:

$$\mathcal{L}' = \{c + L : \ (c \in \mathbb{N}_0) \wedge (L \in \mathcal{L})\} \ . \tag{7}$$

The following bounds on RTD and RTD^+ (for sufficiently large values of k)[3] are known from (Gao et al., 2015):

	RTD^+	RTD
LINSET_k	$= \infty$	?
CF-LINSET_k	$= k$	$\in \{k-1, k\}$
$\text{NE-LINSET}'_k$	$= k+1$	$\in \{k-1, k, k+1\}$

Here $\text{NE-LINSET}'_k$ denotes the "shift extension" of NE-LINSET_k .
The following result shows the corresponding bounds with PBTD in place of RTD:

Theorem 17 *The bounds in the following table are valid:*

	PBTD^+	PBTD
LINSET_k	$= k$	$\in \{k-1, k\}$
CF-LINSET_k	$= k$	$\in \{k-1, k\}$
NE-LINSET_k	$\in \left[\left\lfloor \frac{k-1}{2} \right\rfloor : k\right]$	$\in \left[\left\lfloor \frac{k-1}{2} \right\rfloor : k\right]$
NE-CF-LINSET_k	$\in \left[\left\lfloor \frac{k-1}{2} \right\rfloor : k\right]$	$\in \left[\left\lfloor \frac{k-1}{2} \right\rfloor : k\right]$

3. For instance, $\text{RTD}^+(\text{LINSET}_k) = \infty$ holds for all $k \ge 2$ and $\text{RTD}(\text{LINSET}_k) = ?$ (where "?" means "unknown") holds for all $k \ge 4$.

Moreover

$$\mathrm{PBTD}^+(\mathcal{L}') = k+1 \ \wedge\ \mathrm{PBTD}(\mathcal{L}') \in \{k-1, k, k+1\} \tag{8}$$

holds for all $\mathcal{L} \in \{\textit{LINSET}_k, \textit{CF-LINSET}_k, \textit{NE-LINSET}_k, \textit{NE-CF-LINSET}_k\}$.

Note that the equation $\mathrm{PBTD}^+(\mathrm{LINSET}_k) = k$ was already stated in Example 2. All upper bounds in Theorem 17 are easy to derive from this equation. The lower bounds in Theorem 17 are much harder to obtain. A complete proof of Theorem 17 will be given in Appendix B.

7. Hierarchical Preference-based Teaching

Suppose that $\mathcal{L}$ is a parametrized concept class in the sense that any concept of $\mathcal{L}$ can be fixed by assigning real values $q = (q_1, \ldots, q_d)$ to "programmable parameters" $Q = (Q_1, \ldots, Q_d)$. The concept resulting from setting $Q_i = q_i$ for $i = 1, \ldots, d$ is denoted as L_q. Let $\mathcal{D} \subseteq \mathbb{R}^d$ be a set which makes the representation q of L_q unique, i.e., $L_q = L_{q'}$ with $q, q' \in \mathcal{D}$ implies that $q = q'$. We will then identify a preference relation over $\mathcal{L}$ with a preference relation over $\mathcal{D}$. For every $p \in \{\downarrow, \uparrow\}^d$, let $\prec_p$ be the following algorithmically defined (lexicographic) preference relation:

1. Given $q \neq q' \in \mathcal{D}$, find the smallest index $i \in [d]$ such that $q_i \neq q'_i$, say $q_i < q'_i$.
2. If $p_i =\downarrow$ (resp. $p_i =\uparrow$), then $q' \prec_p q$ (resp. $q \prec q'$).

Imagine a student with this preference relation who has seen a collection of labeled examples. The following hierarchical system of Rules $i = 1, \ldots, d$ clarifies which value q'_i she should assign to the unknown parameter Q_i:

Rule i: With i-highest priority do the following. Choose q'_i as small as possible if $p_i =\downarrow$ and as large as possible if $p_i =\uparrow$. Assign the value q'_i to the parameter Q_i.

It is important to understand that this rule system is hierarchical (as expressed by the distinct priorities of the rules): when Rule i becomes active, then the values of the parameters $Q_1, \ldots, Q_{i-1}$ (in accordance with the rules $1, \ldots, i-1$) have been chosen already.

Suppose that L_q with $q \in \mathcal{D}$ is the target concept. A teacher who designs a teaching set for L_q w.r.t. $(\mathcal{L}, \prec_p)$ can proceed in stages $i = 1, \ldots, d$ as follows:

Stage i: Suppose that $p_i =\downarrow$ (resp. $p_i =\uparrow$). Choose the next part T_i of the teaching set so that every hypothesis $L_{q'}$ with $q' \in \mathcal{D}$, and $q'_1 = q_1, \ldots, q'_{i-1} = q_{i-1}$ satisfies the following condition: if $L_{q'}$ is consistent with $T_1 \cup \ldots \cup T_{i-1} \cup T_i$, then $q'_i \geq q_i$ (resp. $q'_i \leq q_i$).

In other words, the teacher chooses T_1 so that the student with preference relation $\prec_p$ will assign the value q_1 to Q_1. Given that $Q_1 = q_1$, the teacher chooses T_2 so that the student will next assign the value q_2 to Q_2, and so on.

This basic technique can be made a little bit more flexible by allowing to handle more than one parameter in a single stage.

8. Application: Teaching Halfspaces

Suppose that $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$. The *halfspace induced by* w *and* b is then given by

$$H_{w,b} = \{x \in \mathbb{R}^d : \ w^\top x + b \ge 0\} \ .$$

Let $\mathcal{H}_d = \{H_{w,b} : \ (w \in \mathbb{R}^d) \wedge (b \in \mathbb{R})\}$ denote the class of d-dimensional Euclidean halfspaces. $\mathcal{H}_d$ decomposes into $\mathcal{H}_d[b < 0]$ and $\mathcal{H}_d[b \ge 0]$ (with the obvious meaning of the two sub-classes).

Lemma 18 *At the expense of just* 1 *example, it can be taught to which of the two sub-classes* $\mathcal{H}_d[b < 0]$ *and* $\mathcal{H}_d[b \ge 0]$ *a target concept taken from* $\mathcal{H}_d$ *belongs.*

Proof A target concept from $\mathcal{H}_d$ that is consistent with $(\mathbf{0}, -)$ (resp. with $(\mathbf{0}, +)$) must belong to $\mathcal{H}_d[b < 0]$ (resp. to $\mathcal{H}_d[b \ge 0]$). ■

Lemma 19 $\mathrm{PBTD}(\mathcal{H}_d[b < 0]) \le d$.

The full proof of Lemma 19 is given in Appendix C. Here we only specify the (lexicographic) preference relation that is used in the proof.
We normalize the representation (w, b) of a halfspace $H_{w,b}$ by setting $\|w\|_\infty = 1$. Initially, we impose the constraints $\|W\|_\infty = 1$ and $B < 0$ on the unknown weight vector $W \in \mathbb{R}^d$ and the unknown bias $B \in \mathbb{R}$. We choose the preference relation according to the following rules:

Rule 1: With highest priority, choose the bias B as small as possible.

Rule 2: With second highest priority, choose $(|W_1|, \dots, |W_d|)$ as small as possible w.r.t. the partial order $\le$ in $\mathbb{R}^d$.

Lemma 20 $\mathrm{PBTD}(\mathcal{H}_d[b \ge 0]) \le 2d + 1$.

The full proof of Lemma 20 is given in Appendix C. Here we only specify the (lexicographic) preference relation that is used in the proof.
We normalize the representation (w, b) of a halfspace $H_{w,b}$ by setting $\|w\|_1 = 1$. Initially, we impose the constraints $\|W\|_1 = 1$ and $B \ge 0$ on the unknown weight vector $W \in \mathbb{R}^d$ and the unknown bias $B \in \mathbb{R}$. We choose the preference relation according to the following rules:

Rule 1: With highest priority, set $B = 0$ unless this rules out consistency.

Rule 2: With second highest priority, assign the value 0 to as many components of W as possible.

Rule 3: With third highest priority, choose the bias B as small as possible.

Rule 4: Let $I \subseteq [d]$ be the set of indices $i \in [d]$ with $W_i \ne 0$ (in accordance with Rule 2). With fourth highest priority, choose $(|W_i|)_{i \in I}$ as large as possible w.r.t. the partial order $\le$ in $\mathbb{R}^{|I|}$.

Lemmas 18, 19 and 20 can be combined to obtain the following result:

Theorem 21 $\mathrm{PBTD}(\mathcal{H}_d) \le 2d + 2$.

We briefly note that the RTD of halfspaces is infinite simply because (1-dimensional) half-intervals can be seen as special (d-dimensional) halfspaces.

9. Some Geometric Classes and Their PBTD

Let $\mathcal{P}_k$ denote the class of convex polygons in the plane with at most k vertices. We know from Example 2 that $\mathrm{PBTD}^+(\mathcal{P}_k) = k$. We claim that $\mathrm{PBTD}(\mathcal{P}_6) \geq 4$. This can be seen as follows. Let P_6 denote a convex polygon with vertices $P1, \ldots, P6$ as shown in Fig. 1. Let $\mathcal{P}_6' \subset \mathcal{P}_6$ denote the family of polygons that result from P_6 by using a subset of the four possible shortcuts. For instance the polygon with vertices $P1, M0, P4, P5, M1$ is among the possible choices. Note the one-to-one correspondence between the four shortcuts and the four shaded triangles in Fig. 1. In order to distinguish between the polygons in $\mathcal{P}_6'$, the following observations can be made:

- The only informative points are found within the four triangles.
- Points within the same triangle give the same information to the student: the point is marked "-" iff the corresponding shortcut is used.

It follows that teaching $\mathcal{P}_6'$ is equivalent to teaching the powerset over a universe of size 4. Hence $\mathrm{PBTD}(\mathcal{P}_6) \geq \mathrm{PBTD}(\mathcal{P}_6') = 4$. A similar but more general reasoning leads to the following result:

Theorem 22 *For all $k \geq 3$, the following holds:*

$$\mathrm{PBTD}(\mathcal{P}_k) \geq \begin{cases} 2\lfloor k/3 \rfloor + 1 & \text{if } k \equiv 2 \pmod 3 \\ 2\lfloor k/3 \rfloor & \text{otherwise} \end{cases}$$

Proof The construction from Fig. 1 can be generalized in the obvious fashion to convex polygons with k vertices, using $2\lfloor k/3 \rfloor$ resp. $2\lfloor k/3 \rfloor + 1$ shortcuts. It is easy to fill in the details. ■

Let $\mathcal{L}$ be a concept class over a universe $\mathcal{X}$. We denote by $\mathcal{L}^k$ the class of concepts which can be written as a union of at most k concepts from $\mathcal{L}$. We denote by $\dot{\mathcal{L}}^k$ the set of concepts which can be written as a union of at most k concepts from $\mathcal{L}$ provided that these concepts have at least one element in common. We remind the reader that $\mathcal{R}_d$ denotes the class of d-dimensional boxes.

Theorem 23 $2^d(k - d + 1) \leq \mathrm{PBTD}^+(\dot{\mathcal{R}}_d^k) \leq 2^d k$.

The proof of the upper bound makes use of the following fact (which is illustrated in Fig. 2): if B is a d-dimensional box with vertices $P_1, \ldots, P_{2^d}$ and w is any point in $\mathbb{R}^d$, then the boxes spanned by $\{w, P_1\}, \ldots, \{w, P_{2^d}\}$, respectively, fully cover B. (In case that $w \in B$, they even form a partition of B.) This can be used to argue that the total set of vertices of (up to) k boxes that have a common point forms a spanning set of their union w.r.t. $\dot{\mathcal{R}}_d^k$. The full proof of Theorem 23 is given in Section D.

It seems that the class $\mathcal{R}_d^k$ poses a harder teaching problem in the preference-based setting than its subclass $\dot{\mathcal{R}}_d^k$. Presenting the set of all (up to $k2^d$ many) vertices does not seem to be sufficient because the student does not know how to decompose them into (up to) k clusters consisting of 2^d vertices each (and this problem occurs in general for classes $\mathcal{L}^k$ whose concepts are k-unions of concepts taken from a basic class $\mathcal{L}$). As a matter of fact, this "clustering problem" can be solved by adding only few negative examples when we deal with disjoint unions of topologically open boxes:

Theorem 24 *The PBTD of the class of disjoint unions of up to k topologically open d-dimensional boxes is upper-bounded by $k(2d + 1)$.*

Proof The proof is very simple. Choose the following hierarchical preference relation. With first priority choose a union of as few as possible boxes. With second priority, prefer a concept R (= union of boxes) over all concepts R' that are proper subsets of R. Fig. 3 shows how the teaching set for each single box of the target concept should be chosen. If the student gets (up to) k teaching sets of this kind, she will put the largest possible box around each positive example. The negative examples at the border of the box ensure that these boxes do not become too large. ■

10. Conclusions

Preference-based teaching uses the natural notion of preference relation to extend the classical teaching model. The resulting model is (i) more powerful than the classical one, (ii) resolves difficulties with the recursive teaching model in the case of infinite concept classes, and (iii) is at the same time free of coding tricks even according to the definition by Goldman and Mathias (1996). Our examples of algebraic and geometric concept classes demonstrate that preference-based teaching can be achieved very efficiently with naturally defined teaching sets and based on intuitive preference relations such as inclusion. We believe that further studies of the PBTD will provide insights into structural properties of concept classes that render them easy or hard to learn in a variety of formal learning models.

We have shown that spanning sets lead to a general-purpose construction for preference-based teaching sets of only positive examples. While this result is fairly obvious, it provides further justification of the model of preference-based teaching, since the teaching sets it yields are often intuitively exactly those a teacher would choose in the classroom (for instance, one would represent convex polygons by their vertices). It should be noted, too, that it can sometimes be difficult to establish whether the upper bound on PBTD obtained this way is tight, or whether the use of negative examples or preference relations other than inclusion yield smaller teaching sets. A further challenge is posed by the study of unions of geometric objects such as axis-aligned boxes. There seems to be no obvious way of combining preference-based teaching sets for a number of objects to a preference-based teaching set for their union, and it is unclear how to choose preference relations in the best possible way. Generally, the choice of preference relation provides a degree of freedom that increases the power of the teacher but also increases the difficulty of establishing lower bounds on the number of examples required for teaching.

Acknowledgements. Sandra Zilles was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), in the Discovery Grant and Canada Research Chairs programs.

References

Thorsten Doliwa, Gaojian Fan, Hans Ulrich Simon, and Sandra Zilles. Recursive teaching dimension, VC-dimension and sample compression. *Journal of Machine Learning Research*, 15: 3107–3131, 2014.

Ziyuan Gao, Hans U. Simon, and Sandra Zilles. On the teaching complexity of linear sets. In *Proceedings of the 26th International Conference on Algorithmic Learning Theory*, pages 102–116, 2015.

Sally A. Goldman and Michael J. Kearns. On the complexity of teaching. *Journal of Computer and System Sciences*, 50(1):20–31, 1995.

Sally A. Goldman and H. David Mathias. Teaching a smarter learner. *Journal of Computer and System Sciences*, 52(2):255–267, 1996.

David Helmbold, Robert Sloan, and Manfred K. Warmuth. Learning nested differences of intersection-closed concept classes. *Machine Learning*, 5:165–196, 1990.

Thomas J. Jech. *The Axiom of Choice*. North-Holland Pub. Co., Amsterdam, 1973.

Zeinab Mazadi, Ziyuan Gao, and Sandra Zilles. Distinguishing pattern languages with membership examples. In *Proceedings of the 8th International Conference on Language and Automata Theory and Applications*, pages 528–540, 2014.

Shay Moran, Amir Shpilka, Avi Wigderson, and Amir Yehudayoff. Compressing and teaching for low VC-dimension. In *Proceedings of the 56'th Annual Symposium on the Foundations of Computer Science*, pages 40–51, 2015.

José C. Rosales and Pedro A. García-Sánchez. *Numerical Semigroups*. Springer, 2009.

Ayumi Shinohara and Satoru Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.

Hans U. Simon and Sandra Zilles. Open problem: Recursive teaching dimension versus VC dimension. In *Proceedings of the 28th Annual Conference on Learning Theory*, pages 1770–1772, 2015.

Sandra Zilles, Steffen Lange, Robert Holte, and Martin Zinkevich. Models of cooperative teaching and learning. *Journal of Machine Learning Research*, 12:349–384, 2011.

Appendix A. Proof of Lemma 10

We first show that there exists a class of VC-dimension 1, say L_∞, such that $\mathrm{PBTD}^+(\mathcal{L}_\infty) = 1$ while $\mathrm{RTD}(\mathcal{L}_\infty) = \infty$. To this end, let $\mathcal{L}_\infty$ be the family of half-intervals over $[0, 1)$, i.e., $\mathcal{L}_\infty = \{[0, a] : 0 \leq a < 1\}$. We first prove that $\mathrm{PBTD}^+(\mathcal{L}_\infty) = 1$. Consider the preference relation given by $[0, b] \prec [0, a]$ iff $a < b$. Then, for each $0 \leq a < 1$, we have

$$\mathrm{PBTD}([0, a], \mathcal{L}_\infty, \prec) \stackrel{(1)}{=} \mathrm{TD}([0, a], \{[0, b] : \ 0 \leq b \leq a\}) = 1$$

because the single example $(a, +)$ suffices for distinguishing $[0, a]$ from any interval $[0, b]$ with $b < a$.

It had been observed by Moran et al. (2015) already that $\mathrm{RTD}(\mathcal{L}_\infty) = \infty$ because every teaching set for some $[0, a]$ must contain an infinite sequence of rationals that converges from above to a. Thus $\mathrm{RTD}(\mathcal{L}_\infty) \geq \mathrm{TD}_{min}(\mathcal{L}_\infty) = \infty$.

Next we show that, for every $k \geq 1$, there exists a class, say $\mathcal{L}_k$, such that $\mathrm{PBTD}^+(\mathcal{L}_k) = 1$ while $\mathrm{RTD}(\mathcal{L}_k) = k$. To this end, let $\mathcal{X} = [0, 2)$. For each $a = \sum_{n \geq 1} \alpha_n 2^{-n} \in [0, 1)$ and for all $i = 1, \ldots, k$, let $1 \leq a_i < 2$ be given by $a_i = 1 + \sum_{n \geq 0} \alpha_{kn+i} 2^{-n}$. Finally, let $I_a =$

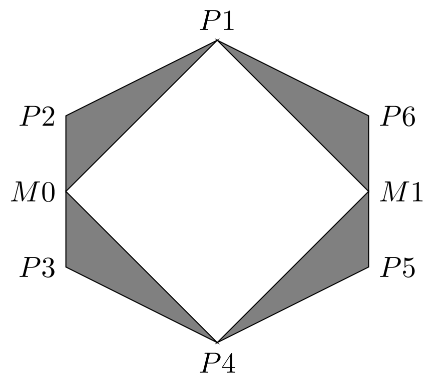

Figure 1: The polygon P_6 with vertices $P1, \ldots, P6$, its four shortcuts $(P1, M0), (M0, P4), (P4, M1), (M1, P1)$ and the corresponding (shaded) triangles.

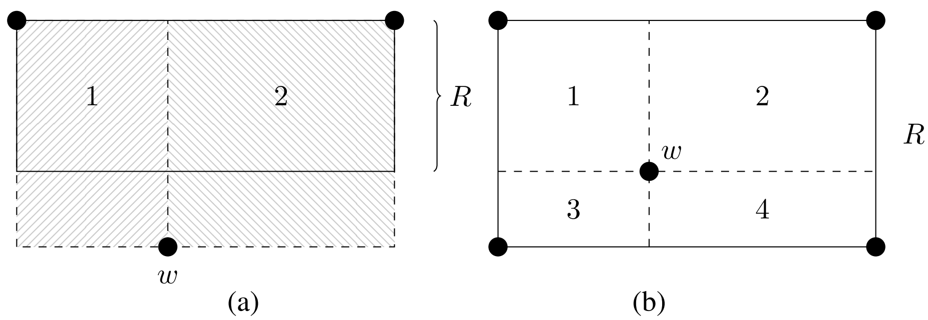

Figure 2: (a) a box R covered by the boxes induced by $w \notin R$ (b) a box R partitioned by the sub-boxes induced by $w \in R$

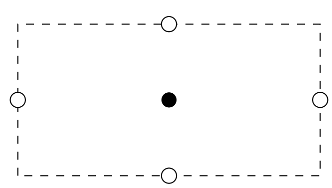

Figure 3: A teaching set for a topologically open box: hollow points indicate negative examples whereas the black point in the center serves a positive example.

$[0, a) \cup \{a_1, \ldots, a_k\} \subseteq \mathcal{X}$ and let $\mathcal{L}_k = \{I_a : 0 \leq a < 1\}$. Clearly $\mathrm{PBTD}^+(\mathcal{L}_k) = 1$ because, using the preference relation given by $I_b \prec I_a$ iff $a < b$, we can teach I_a w.r.t. $\mathcal{L}_k$ by presenting the single example $(a, +)$ (the same strategy as for half-intervals). Moreover, note that I_a is the only concept in $\mathcal{L}_k$ that contains $a_1, \ldots, a_k$, i.e., $\{a_1, \ldots, a_k\}$ is a positive teaching set for I_a w.r.t. $\mathcal{L}_k$. It follows that $\mathrm{RTD}(\mathcal{L}_k) \leq \mathrm{TD}(\mathcal{L}_k) \leq k$. It remains to show that $\mathrm{RTD}(\mathcal{L}_k) \geq k$. To this end, we consider the subclass $\mathcal{L}'_k$ consisting of all concepts I_a such that a has only finitely many 1's in its binary representation. Pick any concept $I_a \in \mathcal{L}'_k$. Let T be any set of at most $k-1$ examples labeled consistently according to I_a. At least one of the positive examples $a_1, \ldots, a_k$ must be missing, say a_i is missing. Let J be the set of indices given by $J = \{n \in \mathbb{N}_0 : \ \alpha_{kn+i} = 0\}$. The following observations show that there exists $a' \in \mathcal{X} \setminus \{a\}$ such that $I_{a'}$ is consistent with T.

- When we set some (at least one but only finitely many) of the bits α_{kn+i} with $n \in J$ from 0 to 1 (while keeping fixed the remaining bits of the binary representation of a), then we obtain a number $a' \neq a$ such that $I_{a'}$ is still consistent with all positive examples in T (including the example $(a, +)$ which might be in T).

- Note that J is an infinite set. It is therefore possible to choose the bits that are set from 0 to 1 in such a fashion that the finitely many bit patterns represented by the numbers in $T^- \cap [1, 2)$ are avoided.

- It is furthermore possible to choose the bits that are set from 0 to 1 in such a fashion that the resulting number a' is as close to a as we like so that $I_{a'}$ is also consistent with the negative examples from $T^- \cap [0.1)$.

It follows from this discussion that no set with less than k examples can possibly be a teaching set for I_a. Since this holds for an arbitrary choice of a, we may conclude that $\mathrm{RTD}(\mathcal{L}_k) \geq \mathrm{RTD}(\mathcal{L}'_k) \geq \mathrm{TD}_{min}(\mathcal{L}'_k) = k$.

Appendix B. Proof of Theorem 17

In Section B.1, we present a general result which helps to verify the upper bounds within Theorem 17. These upper bounds are then derived in Section B.2. Section B.3 is devoted to the derivation of the lower bounds.

B.1. The Shift Lemma

In this section, we assume that $\mathcal{L}$ is a concept class over a universe $\mathcal{X} \in \{\mathbb{N}_0, \mathbb{Q}_0^+, \mathbb{R}_0^+\}$. We furthermore assume that 0 is contained in every concept $L \in \mathcal{L}$. We can extend $\mathcal{L}$ to a larger class, namely the shift-extension $\mathcal{L}'$ of $\mathcal{L}$, by allowing each of its concepts to be shifted by some constant which is taken from $\mathcal{X}$:

$$\mathcal{L}' = \{c + L : \ (c \in \mathcal{X}) \wedge (L \in \mathcal{L})\} \ .$$

The next result states that this extension has little effect only on the complexity measures PBTD and PBTD^+:

Lemma 25 (Shift Lemma) *With the above notation and assumptions, the following holds:*

$$\mathrm{PBTD}(\mathcal{L}) \leq \mathrm{PBTD}(\mathcal{L}') \leq 1{+}\mathrm{PBTD}(\mathcal{L}) \ \textit{and} \ \mathrm{PBTD}^+(\mathcal{L}) \leq \mathrm{PBTD}^+(\mathcal{L}') \leq 1{+}\mathrm{PBTD}^+(\mathcal{L}) \ .$$

Proof It suffices to verify the inequalities $\mathrm{PBTD}(\mathcal{L}') \le 1 + \mathrm{PBTD}(\mathcal{L})$ and $\mathrm{PBTD}^+(\mathcal{L}') \le 1 + \mathrm{PBTD}^+(\mathcal{L})$ because the other inequalities hold by virtue of monotonicity. Let T be an admissible mapping for $\mathcal{L}$. It suffices to show that T can be transformed into an admissible mapping T' for $\mathcal{L}'$ such that $\mathrm{ord}(T') \le 1 + \mathrm{ord}(T)$ and such that T' is positive provided that T is positive. To this end, we define T' as follows:

$$T'(c + L) = \{(c, +)\} \cup \{(c + x, b) : (x, b) \in T(L)\} \ .$$

Obviously $\mathrm{ord}(T') \le 1 + \mathrm{ord}(T)$. Note that $c \in c + L$ because of our assumption that 0 is contained in every concept in $\mathcal{L}$. Moreover, since the admissibility of T implies that L is consistent with $T(L)$, the above definition of $T'(c + L)$ makes sure that $c + L$ is consistent with $T'(c + L)$. It suffices therefore to show that the relation $R^+_{T'}$ is asymmetric. Consider a pair $(c' + L', c + L) \in R_{T'}$. By the definition of $R_{T'}$, it follows that $c' + L'$ is consistent with $T'(c + L)$. Because of $(c, +) \in T'(c + L)$, we must have $c' \le c$. Suppose that $c' = c$. In this case, L' must be consistent with $T(L)$. Thus $L' \prec_T L$. This discussion shows that $(c' + L', c + L) \in R_{T'}$ can happen only if either $c' < c$ or $(c' = c) \wedge (L' \prec_T L)$. Since $\prec_T$ is asymmetric, we may now conclude that $R^+_{T'}$ is asymmetric, as desired. Finally note that, according to our definition above, the mapping T' is positive provided that T is positive. This concludes the proof. ■

B.2. The Upper Bounds in Theorem 17

We remind the reader that the equality $\mathrm{PBTD}^+(\text{LINSET}_k) = k$ was stated in Example 2. In combination with the Shift Lemma, this implies that $\mathrm{PBTD}^+(\text{LINSET}'_k) \le k + 1$. All remaining upper bounds in Theorem 17 follow now by virtue of monotonicity.

B.3. The Lower Bounds in Theorem 17

The lower bounds in Theorem 17 are an immediate consequence of the following result:

Lemma 26 *The following lower bounds are valid:*

$$\mathrm{PBTD}^+(\textit{NE-CF-LINSET}'_k) \ \ge \ k + 1 \ . \tag{9}$$
$$\mathrm{PBTD}(\textit{NE-CF-LINSET}'_k) \ \ge \ k - 1 \ . \tag{10}$$
$$\mathrm{PBTD}(\textit{NE-CF-LINSET}_k) \ \ge \ \frac{k-1}{2} \ . \tag{11}$$
$$\mathrm{PBTD}(\textit{CF-LINSET}_k) \ \ge \ k - 1 \ . \tag{12}$$

This lemma can be seen as an extension and a strengthening of a similar result in (Gao et al., 2015) where the following lower bounds were shown:

$$\begin{aligned} \mathrm{RTD}^+(\text{NE-LINSET}'_k) &\ge k + 1 \\ \mathrm{RTD}(\text{NE-LINSET}'_k) &\ge k - 1 \\ \mathrm{RTD}(\text{CF-LINSET}_k) &\ge k - 1 \end{aligned}$$

The proof of Lemma 26 builds on some ideas that are found in (Gao et al., 2015) already, but it requires some elaboration to get the stronger results.

We now briefly explain why the lower bounds in Theorem 17 directly follow from Lemma 26. Note that the lower bound $k-1$ in (8) is immediate from (10) and a monotonicity argument. Note furthermore that $\mathrm{PBTD}^+(\text{CF-LINSET}'_k) \geq k+1$ because of (9) and a monotonicity argument. Then the Shift Lemma implies that $\mathrm{PBTD}(\text{CF-LINSET}_k) \geq k$. All remaining lower bounds in Theorem 17 are obtained from these observations by virtue of monotonicity.

The proof of Theorem 17 can therefore be accomplished by proving Lemma 26. It turns out that the proof of this lemma is quite involved. We will present in Section B.3.1 some theoretical prerequisites. Sections B.3.2 and B.3.3 are devoted to the actual proof of the lemma.

B.3.1. Some Basic Concepts in the Theory of Numerical Semigroups

Recall from Section 6 that $\langle G\rangle = \left\{\sum_{g\in G} a(g)g : a(g) \in \mathbb{N}_0\right\}$. The elements of G are called *generators* of $\langle G\rangle$. A set $P \subset \mathbb{N}$ is said to be *independent* if none of the elements in P can be written as a linear combination (with coefficients from $\mathbb{N}_0$) of the remaining elements (so that $\langle P'\rangle$ is a proper subset of $\langle P\rangle$ for every proper subset P' of P). It is well known (Rosales and García-Sánchez, 2009) that independence makes generating systems unique, i.e., if P, P' are independent, then $\langle P\rangle = \langle P'\rangle$ implies that $P = P'$. Moreover, for every independent set P, the following implication is valid:

$$(S \subseteq \langle P\rangle \wedge P \not\subseteq S) \;\Rightarrow\; (\langle S\rangle \subset \langle P\rangle) \ . \tag{13}$$

Let $P = \{a_1, \ldots, a_k\}$ be independent with $a_1 = \min P$. It is well known[4] and easy to see that the residues of $a_1, a_2, \ldots, a_k$ modulo a_1 must be pairwise distinct (because, otherwise, we would obtain a dependence). If a_1 is a prime and $|P| \geq 2$, then the independence of P implies that $\gcd(P) = 1$. Thus the following holds:

Lemma 27 *If $P \subset \mathbb{N}$ is an independent set of cardinality at least 2 and $\min P$ is a prime, then* $\gcd(P) = 1$.

In the remainder of the paper, the symbols P and P' are reserved for denoting independent sets of generators.

It is well known that $\langle G\rangle$ is co-finite iff $\gcd(G) = 1$ (Rosales and García-Sánchez, 2009). Let P be a finite (independent) subset of $\mathbb{N}$ such that $\gcd(P) = 1$. The largest number in $\mathbb{N} \setminus \langle P\rangle$ is called the *Frobenius number of P* and is denoted as $F(P)$. It is well known (Rosales and García-Sánchez, 2009) that

$$F(\{p, q\}) = pq - p - q \tag{14}$$

provided that $p, q \geq 2$ satisfy $\gcd(p, q) = 1$.

B.3.2. Proof of (9)

The shift-extension of NE-CF-LINSET$_k$ is (by way of definition) the following class:

$$\text{NE-CF-LINSET}'_k = \{c + \langle P\rangle_+ :\ (c \in \mathbb{N}_0) \wedge (P \subset \mathbb{N}) \wedge (|P| \leq k) \wedge (\gcd(P) = 1)\} \ . \tag{15}$$

4. E.g., see (Rosales and García-Sánchez, 2009)

It is easy to see that this can be written alternatively in the form

$$\text{NE-CF-LINSET}'_k = \left\{ N + \langle P \rangle : \ N \in \mathbb{N}_0 \wedge P \subset \mathbb{N} \wedge |P| \le k \wedge \gcd(P) = 1 \wedge \sum_{p \in P} p \le N \right\} \tag{16}$$

where N in (16) corresponds to $c + \sum_{p \in P} p$ in (15).

For technical reasons, we define the following subfamilies of NE-CF-LINSET$'_k$. For each $N \ge 0$, let

$$\text{NE-CF-LINSET}'_k[N] = \{N + L : L \in \text{LINSET}_k[N]\}$$

where

$$\text{LINSET}_k[N] = \left\{ \langle P \rangle \in \text{LINSET}_k : (\gcd(P) = 1) \wedge \left(\sum_{p \in P} p \le N \right) \right\} \ .$$

In other words, NE-CF-LINSET$'_k[N]$ is the subclass consisting of all concepts in NE-CF-LINSET$'_k$ (written in the form (16)) whose constant is N.
A central notion for proving (9) is the following one:

Definition 28 *Let $k, N \ge 2$ be integers. We say that a set $L \in$ NE-CF-LINSET$'$ is (k, N)-*special *if it is of the form $L = N + \langle P \rangle$ such that the following holds:*

1. *P is an independent set of cardinality k and $\min P$ is a prime (so that $\gcd(P) = 1$ according to Lemma 27, which furthermore implies that $\langle P \rangle$ is co-finite).*

2. *Let $q(P)$ denote the smallest prime that is greater than $F(P)$ and greater than $\max P$. For $a = \min P$ and $r = 0, \ldots, a-1$, let*

$$t_r(P) = \min\{s \in \langle P \rangle : s \equiv r \pmod a\} \ \textit{and} \ t_{max}(P) = \max_{0 \le r \le a-1} t_r(P) \ .$$

Then

$$N \ge k(a + t_{max}(P)) \ \textit{and} \ N \ge q(P) + \sum_{p \in P \setminus \{a\}} p \ . \tag{17}$$

We need at least k positive examples in order to distinguish a (k, N)-special set from all its proper subsets in NE-CF-LINSET$'_k[N]$, as the following result shows:

Lemma 29 *For all $k \ge 2$, the following holds. If $L \in$ NE-CF-LINSET$'$ is (k, N)-special, then $L \in$ NE-CF-LINSET$'[N]$ and $I'(L,$ NE-CF-LINSET$_k[N]) \ge k$.*

Proof Suppose that $L = N + \langle P \rangle$ is of the form as described in Definition 28. Let $P = \{a, a_2 \ldots, a_k\}$ with $a = \min P$. For the sake of simplicity, we will write t_r instead of $t_r(P)$ and t_{max} instead of $t_{max}(P)$. The independence of P implies that $t_{a_i \bmod a} = a_i$ for $i = 2, \ldots, k$. It follows that $t_{max} \ge \max P$. Since, by assumption, $N \ge k \cdot t_{max}$, it becomes obvious that $L \in$ NE-CF-LINSET$'[N]$.
Assume by way of contradiction that the following holds:

(A) There is a weak spanning set S of size $k - 1$ for L w.r.t. NE-CF-LINSET$'_k[N]$.

Since N is contained in any concept from NE-CF-LINSET$'_k[N]$, we may assume that $N \notin S$ so that S is of the form $S = \{N + x_1, \ldots, N + x_{k-1}\}$ for integers $x_i \geq 1$. For $i = 1, \ldots, k-1$, let $r_i = x_i \bmod a \in \{0, 1, \ldots, a-1\}$. It follows that each x_i is of the form $x_i = q_i a + t_{r_i}$ for some integer $q_i \geq 0$. Let $X = \{x_1, \ldots, x_{k-1}\}$. We proceed by case analysis:

Case 1: $X \subseteq \{a_2, \ldots, a_k\}$ (so that, in view of $|X| = k-1$, we even have $X = \{a_2, \ldots, a_k\}$).
Let $L' = N + \langle X \rangle$. Then $S \subseteq L'$. Note that $X \subseteq P$ but $P \not\subseteq X$. We may conclude from (13) that $\langle X \rangle \subset \langle P \rangle$ and, therefore, $L' \subset L$. Thus L' is a proper subset of L which contains S. Note that (17) implies that $N \geq \sum_{i=2}^{k} a_i = \sum_{i=1}^{k-1} x_i$. If $\gcd(X) = 1$, then $L' \in$ NE-CF-LINSET$[N]$ and we have an immediate contradiction to the above assumption (A). Otherwise, if $\gcd(X) \geq 2$, then we define $L'' = N + \langle X \cup \{q(P)\} \rangle$. Note that $S \subseteq L' \subseteq L''$. Since $q(P) > F(P)$, we have $X \cup \{q(P)\} \subseteq \langle P \rangle$ and, since $q(P) > \max P$, we have $P \not\subseteq X \cup \{q(P)\}$. We may conclude from (13) that $\langle X \cup \{q(P)\} \rangle \subset \langle P \rangle$ and, therefore, $L'' \subset L$. Thus, L'' is a proper subset of L which contains S. Because $X = \{a_2, \ldots, a_k\}$ and $q(P)$ is a prime that is greater than $\max P$, it follows that $\gcd(X \cup \{q(P)\}) = 1$. In combination with (17), it easily follows now that $L'' \in$ NE-CF-LINSET$[N]$. Putting everything together, we arrive at a contradiction to the assumption (A).

Case 2: $X \not\subseteq \{a_2, \ldots, a_k\}$.
If $r_i = 0$ for $i = 1, \ldots, k-1$, then each x_i is a multiple of a. In this case, $N + \langle a, q(P) \rangle$ is a proper subset of $L = N + \langle P \rangle$ that is consistent with S, and we arrive at a contradiction. We may therefore assume that there exists $i' \in \{1, \ldots, k-1\}$ such that $r_{i'} \neq 0$. From the case assumption, $X \not\subseteq \{a_2, \ldots, a_k\}$, it follows that there must exist an index $i'' \in \{1, \ldots, k-1\}$ such that $q_{i''} \geq 1$ or $t_{r_{i''}} \notin \{a_2, \ldots, a_k\}$. For $i = 1, \ldots, k-1$, let $q'_i = \min\{q_i, 1\}$ and $x'_i = q'_i a + t_{r_i}$. Note that $q'_{i''} = 1$ iff $q_{i''} \geq 1$. Define $L'' = N + \langle X' \rangle$ for $X' = \{a, x'_1, \ldots, x'_{k-1}\}$ and observe the following. First, the set L'' clearly contains S. Second, the choice of $x'_1, \ldots, x'_{k-1}$ implies that $X' \subseteq \langle P \rangle$. Third, it easily follows from $q'_{i''} = 1$ or $t_{r_{i''}} \notin \{a_2, \ldots, a_k\}$ that $P \not\subseteq \{a, x'_1, \ldots, x'_{k-1}\}$. We may conclude from (13) that $\langle X' \rangle \subset \langle P \rangle$ and, therefore, $L'' \subset L$. Thus, L'' is a proper subset of L which contains S. Since $r_{i'} \neq 0$ and a is a prime, it follows that $\gcd(a, x'_{i'}) = 1$ and, therefore, $\gcd(X') = 1$. In combination with (17), it easily follows now that $L'' \in$ NE-CF-LINSET$[N]$. Putting everything together, we arrive again at a contradiction to the assumption (A).

■

For the sake of brevity, let $\mathcal{L} =$ NE-CF-LINSET$'$. Assume by way of contradiction that there exists a positive mapping T of order k that is admissible for $\mathcal{L}_k$. We will pursue the following strategy:

1. We define a set $L \in \mathcal{L}_k$ of the form $L = N + p + \langle 1 \rangle$.

2. We define a second set $L' = N + \langle G \rangle \in \mathcal{L}$ that is (k, N)-special and consistent with $T^+(L)$. Moreover, $L' \setminus L = \{N\}$.

If this can be achieved, then the proof will be accomplished as follows:

- According to Lemma 29, $T^+(L')$ must contain at least k examples (all of which are different from N) for distinguishing L' from all its proper subsets in $\mathcal{L}_k[N]$.

- Since L' is consistent with $T^+(L)$, the set $T^+(L')$ must contain an example which distinguishes L' from L. But the only example which fits this purpose is $(N, +)$.
- The discussion shows that $T^+(L')$ must contain k examples in order to distinguish L' from all its proper subsets in $\mathcal{L}_k$ plus one additional example, N, needed to distinguish L' from L.
- We arrived at a contradiction to our initial assumption that T^+ is of order k.

We still have to describe how our proof strategy can actually be implemented. We start with the definition of L. Pick the smallest prime $p \geq k+1$. Then $\{p, p+1, \ldots, p+k\}$ is independent. Let $M = F(\{p, p+1\}) \overset{(14)}{=} p(p+1) - p - (p+1)$. An easy calculation shows that $k \geq 2$ and $p \geq k+1$ imply that $M \geq p+k$. Let $I = \{p, p+1, \ldots, M\}$. Choose N large enough so that all concepts of the form

$$N + \langle P \rangle \text{ where } |P| = k,\ p = \min P \text{ and } P \subseteq I$$

are (k, N)-special. With these choices of p and N, let $L = N+p+\langle 1 \rangle$. Note that $N+p, N+p+1 \in T^+(L)$ because, otherwise, one of the concepts $N+p+1+\langle 1 \rangle, N+p+\langle 2, 3 \rangle \subset L$ would be consistent with $T^+(L)$ whereas $T^+(L)$ must distinguish L from all its proper subsets in $\mathcal{L}_k$. Setting $A = \{x : N + x \in T^+(L)\}$, it follows that $|A| = |T^+(L)| \leq k$ and $p, p+1 \in A$. The set A is not necessarily independent but it contains an independent subset B such that $p, p+1 \in B$ and $\langle A \rangle = \langle B \rangle$. Since $M = F(\{p, p+1\})$, it follows that any integer greater than M is contained in $\langle p, p+1 \rangle$. Since B is an independent extension of $\{p, p+1\}$, it cannot contain any integer greater than M. It follows that $B \subseteq I$. Clearly, $|B| \leq k$ and $\gcd(B) = 1$. We would like to transform B into another generating system $G \subseteq I$ such that

$$\langle B \rangle \subseteq \langle G \rangle, \gcd(G) = 1 \text{ and } |G| = k \ \ .$$

If $|B| = k$, we can simply set $G = B$. If $|B| < k$, then we make use of the elements in the independent set $\{p, p+1, \ldots, p+k\} \subseteq I$ and add them, one after the other, to B (thereby removing other elements from B whenever their removal leaves $\langle B \rangle$ invariant) until the resulting set G contains k elements. We now define the set L' by setting $L' = N + \langle G \rangle$. Since $G \subseteq I = \{p, p+1, \ldots, M\}$, and $p, p+1 \in G$, it follows that $p = \min G$, $\gcd(G) = 1$ and $\min(L' \setminus \{N\})$ is $N + p$. Thus, $L' \setminus L = \{N\}$, as desired. Moreover, since N had been chosen large enough, the set L' is (k, N)-special. Thus L and L' have all properties that are required by our proof strategy and the proof of (9) is complete.

B.3.3. PROOF OF (10), (11) AND (12)

We make use of some well known (and trivial) lower bounds on TD_{min}:

Example 3 *For every $k \in \mathbb{N}$, let $[k] = \{1, 2, \ldots, k\}$, let $2^{[k]}$ denote the powerset of $[k]$ and, for all $\ell = 0, 1, \ldots, k$, let*

$$\binom{[k]}{\ell} = \{S \subseteq [k] :\ |S| = \ell\}$$

denote the class of those subsets of $[k]$ that have exactly ℓ elements. It is trivial to verify that

$$TD_{min}\left(2^{[k]}\right) = k \ \text{ and } \ TD_{min}\left(\binom{[k]}{\ell}\right) = \min\{\ell, k-\ell\} \ \ .$$

In view of $\mathrm{PBTD}^+(\text{LINSET}_k) = k$, the next results show that negative examples are of limited help only as far as preference-based teaching of concepts from LINSET_k is concerned:

Lemma 30 *For every $k \geq 1$ and for all $\ell = 0, \ldots, k-1$, let*

$$\begin{aligned} \mathcal{L}_k &= \{\langle k, p_1, \ldots, p_{k-1}\rangle : \ p_i \in \{k+i, 2k+i\}\} \ , \\ \mathcal{L}_{k,\ell} &= \{\{\langle k, p_1, \ldots, p_{k-1}\rangle \in \mathcal{L}_k : \ |\{i : p_i = k+i\}| = \ell\} \ . \end{aligned}$$

With this notation, the following holds:

$$\textit{TD}_{min}(\mathcal{L}_k) \geq k-1 \ \textit{ and } \ \textit{TD}_{min}(\mathcal{L}_{k,\ell}) \geq \min\{\ell, k-1-\ell\} \ .$$

Proof For $k = 1$, the assertion in the lemma is vacuous. Suppose therefore that $k \geq 2$. An inspection of the generators $k, p_1, \ldots, p_{k-1}$ with $p_i \in \{k+i, 2k+i\}$ shows that

$$\begin{aligned} \mathcal{L}_k &= \{L_{k,S} : \ S \subseteq \{k+1, k+2, \ldots, 2k-1\}\} \\ \mathcal{L}_{k,\ell} &= \{L_{k,S} : \ (S \subseteq \{k+1, k+2, \ldots, 2k-1\}) \wedge (|S| = \ell)\} \end{aligned}$$

where

$$L_{k,S} = \{0, k\} \cup \{2k, 2k+1, \ldots\} \cup S \ .$$

Note that the examples in $\{0, 1, \ldots, k\} \cup \{2k, 2k+1, \ldots, \}$ are redundant because they do not distinguish between distinct concepts from $\mathcal{L}_k$. The only useful examples are therefore contained in the interval $\{k+1, k+2, \ldots, 2k-1\}$. From this discussion, it follows that teaching the concepts of $\mathcal{L}_k$ (resp. of $\mathcal{L}_{k,\ell}$) is not essentially different from teaching the concepts of $2^{[k-1]}$ $\left(\text{resp. of } \binom{[k-1]}{\ell}\right)$. This completes the proof of the lemma because we know from Example 3 that $\mathrm{TD}_{min}(2^{[k-1]}) = k-1$ and $\mathrm{TD}_{min}\left(\binom{[k-1]}{\ell}\right) = \min\{\ell, k-1-\ell\}$. ∎

We claim now that the inequalities (10), (11) and (12) are valid, i.e., we claim that the following holds:.

1. $\mathrm{PBTD}(\text{CF-LINSET}_k) \geq k-1$.
2. $\mathrm{PBTD}(\text{NE-CF-LINSET}_k) \geq \lfloor (k-1)/2 \rfloor$.
3. $\mathrm{PBTD}(\text{NE-CF-LINSET}'_k) \geq k-1$.

Proof For $k = 1$, the inequalities are obviously valid. Suppose therefore that $k \geq 2$.

1. Since $\gcd(k, k+1) = \gcd(k, 2k+1) = 1$, it follows that $\mathcal{L}_k$ is a finite subclass of CF-LINSET_k. Thus $\mathrm{PBTD}(\text{CF-LINSET}_k) \geq \mathrm{PBTD}(\mathcal{L}_k) \geq \mathrm{TD}_{min}(\mathcal{L}_k) \geq k-1$.
2. Define $\mathcal{L}_k[N] = \{N + L : \ L \in \mathcal{L}_k\}$ and $\mathcal{L}_{k,\ell}[N] = \{N + L : \ L \in \mathcal{L}_{k,\ell}\}$. Clearly $\mathrm{TD}_{min}(\mathcal{L}_k[N]) = \mathrm{TD}_{min}(\mathcal{L}_k)$ and $\mathrm{TD}_{min}(\mathcal{L}_{k,\ell}[N]) = \mathrm{TD}_{min}(\mathcal{L}_{k,\ell})$ holds for every $N \geq 0$. It follows that the lower bounds in Lemma 30 are also valid for the classes $\mathcal{L}_k[N]$ and $\mathcal{L}_{k,\ell}[N]$ in place of $\mathcal{L}_k$ and $\mathcal{L}_{k,\ell}$, respectively. Let

$$N(k) = k^2 + (k-1-\lfloor (k-1)/2 \rfloor)k + \sum_{i=1}^{k-1} i = k^2 + (k-1-\lfloor (k-1)/2 \rfloor)k + \frac{1}{2}(k-1)k \ . \quad (18)$$

It suffices to show that $N(k) + \mathcal{L}_{k,\lfloor (k-1)/2 \rfloor}$ is a finite subclass of NE-CF-LINSET$_k$. To this end, first note that

$$\langle k, p_1, \ldots, p_{k-1} \rangle_+ = k + \sum_{i=1}^{k-1} p_i + \langle k, p_1, \ldots, p_{k-1} \rangle \ .$$

Call p_i "light" if $p_i = k + i$ and call it "heavy" if $p_i = 2k + i$. Note that a concept L from $N(k) + \mathcal{L}_{k,\ell}$ is of the general form

$$L = N(k) + \langle k, p_1, \ldots, p_{k-1} \rangle \tag{19}$$

with exactly ℓ light parameters among $p_1, \ldots, p_{k-1}$. A straightforward calculation shows that, for $\ell = \lfloor (k-1)/2 \rfloor$, the sum $k + \sum_{i=1}^{k-1} p_i$ equals the number $N(k)$ as defined in (18). Thus, the concept L from (19) with exactly $\lfloor (k-1)/2 \rfloor$ light parameters among $\{p_1, \ldots, p_{k-1}\}$ can be rewritten as follows:

$$L = N(k) + \langle k, p_1, \ldots, p_{k-1} \rangle = \langle k, p_1, \ldots, p_{k-1} \rangle_+ \ .$$

This shows that $L \in$ NE-CF-LINSET$_k$. As L is a concept from $N(k) + \mathcal{L}_{k,\lfloor (k-1)/2 \rfloor}$ in general form, we may conclude that $N(k) + \mathcal{L}_{k,\lfloor (k-1)/2 \rfloor}$ is a finite subclass of NE-CF-LINSET$_k$, as desired.

3. The proof of third inequality is similar to the above proof of the second one. It suffices to show that, for every $k \geq 2$, there exists $N \in \mathbb{N}$ such that $N + \mathcal{L}_k$ is a subclass of NE-CF-LINSET$'_k$. To this end, we set $N = 3k^2$. A concept L from $3k^2 + \mathcal{L}_k$ is of the general form

$$L = 3k^2 + \langle k, p_1, \ldots, p_{k-1} \rangle$$

with $p_i \in \{k + i, 2k + i\}$ (but without control over the number of light parameters). It is easy to see that the constant $3k^2$ is large enough so that L can be rewritten as

$$L = 3k^2 - \left(k + \sum_{i=1}^{k-1} p_i \right) + \langle k, p_1, \ldots, p_{k-1} \rangle_+$$

where $3k^2 - \left(k + \sum_{i=1}^{k-1} p_i \right) \geq 0$. This shows that $L \in$ NE-CF-LINSET$'_k$. As L is a concept from $3k^2 + \mathcal{L}_k$ in general form, we may conclude that $3k^2 + \mathcal{L}_k$ is a finite subclass of NE-CF-LINSET$'_k$, as desired.

■

Appendix C. Proof of Lemmas 19 and 20

Proof [Lemma 19] We have to show that $\mathrm{PBTD}(\mathcal{H}_d[b < 0]) \leq d$. To this end, we normalize the representation (w, b) of a halfspace $H_{w,b}$ by setting $\|w\|_\infty = 1$. Initially, we impose the constraints $\|W\|_\infty = 1$ and $B < 0$ on the unknown weight vector $W \in \mathbb{R}^d$ and the unknown bias $B \in \mathbb{R}$. We choose the preference relation according to the following rules:

Rule 1: With highest priority, choose the bias B as small as possible.

Rule 2: With second highest priority, choose $(|W_1|, \dots, |W_d|)$ as small as possible w.r.t. the partial order $\leq$ in $\mathbb{R}^d$.

Let $H_{w,b}$ be the target concept. Let $I = \{i \in [d] : w_i \neq 0\}$. For every $i \in I$, choose the example $((-b/w_i)\mathbf{1}_i, +)$ where $\mathbf{1}_i$ denotes the vector with 1 in position i and zeroes elsewhere. Since the label is positive, the pair (W, B) can be consistent with these examples only if $-bW_i/w_i + B \geq 0$. Since $b, B < 0$, it follows that $\mathrm{sgn}(W_i) = \mathrm{sgn}(w_i)$ so that (in combination with $|W_i| \leq 1$) the constraint $-bW_i/w_i + B \geq 0$ can be written in the form

$$|b|/|w_i| - |B| \geq |b| \cdot |W_i|/|w_i| - |B| \geq 0 \ .$$

Making use of $\|w\|_\infty = 1$, we obtain $|B| \leq \min_{i \in I}(|b|/|w_i|) = |b|$. From $b, B < 0$, we may conclude that $0 > B \geq b$. According to Rule 1, the bias B will be chosen as $B = b$. Now the constraint $|b| \cdot |W_i|/|w_i| - |B| \geq 0$ can be written as $|b| \cdot |W_i|/|w_i| - |b| \geq 0$ which, after cancellation of $|b|$, becomes $|W_i| \geq |w_i|$. According to Rule 2, the parameter W_i with $i \in I$ will be chosen such that $|W_i| = |w_i|$ (implying that $W_i = w_i$ because $\mathrm{sgn}(W_i) = \mathrm{sgn}(w_i)$). Moreover, the parameter W_j with $j \notin I$ will be chosen such that $W_j = |W_j| = 0$. Thus $W = w$. It follows from this discussion that the above $|I| \leq d$ examples form a preference-based teaching set for $H_{w,b}$. This concludes the proof of the lemma. ■

Proof [Lemma 20] We have to show that $\mathrm{PBTD}(\mathcal{H}_d[b \geq 0]) \leq 2d+1$. To this end, we normalize the representation (w, b) of a halfspace $H_{w,b}$ by setting $\|w\|_1 = 1$. Initially, we impose the constraints $\|W\|_1 = 1$ and $B \geq 0$ on the unknown weight vector $W \in \mathbb{R}^d$ and the unknown bias $B \in \mathbb{R}$. We choose the preference relation according to the following rules:

Rule 1: With highest priority, set $B = 0$ unless this rules out consistency.

Rule 2: With second highest priority, assign the value 0 to as many components of W as possible.

Rule 3: With third highest priority, choose the bias B as small as possible.

Rule 4: Let $I \subseteq [d]$ be the set of indices $i \in [d]$ with $W_i \neq 0$ (in accordance with Rule 2). With fourth highest priority, choose $(|W_i|)_{i \in I}$ as large as possible w.r.t. the partial order $\leq$ in $\mathbb{R}^{|I|}$.

We first discuss the case that the target concept is represented by $(w, 0)$ with $\|w\|_1 = 1$. Let $u_1, \dots, u_{d-1} \in \mathbb{R}^d$ be a basis of the $(d-1)$-dimensional space $\langle w \rangle^\perp$. Choose a teaching set T of size $2d-1$ that consists of $(u_i, +), (-u_i, +)$ for $i = 1, \dots, d-1$ and of $(w, +)$. According to Rule 1, the bias B will be chosen as $B = 0$. Given that $B = 0$, setting $W = w$ is the only chance for achieving consistency with T and for satisfying the normalization condition $\|W\|_1 = 1$.
Suppose now that the target concept is represented by (w, b) with $\|w\|_1 = 1$ and $b > 0$. Let $I = \{i \in [d] \ w_i \neq 0\}$. Note that $I \neq \emptyset$ because $w \neq \mathbf{0}$. Let $\sigma_i = \mathrm{sgn}(w_i)$ and $\sigma = (\sigma_1, \dots, \sigma_d)^\top$. Choose a teaching set T that consists of the following examples:

- the example $(u, +) = (-b\sigma, +)$,
- the example $(v_i, +) = ((-b/w_i)\mathbf{1}_i, +)$ for every $i \in I$,

- the example $(2v_i, -)$ for every $i \in I$.

It is easy to check that $H_{w,b}$ is consistent with T:

$$\begin{aligned} w^\top u + b &= -b\sum_{i=1}^{d} \sigma_i w_i + b = -b\|w\|_1 + b = 0 \ , \\ w^\top v_i + b &= -b + b = 0 \ , \\ w^\top (2v_i) + b &= -2b + b = -b < 0 \ . \end{aligned}$$

For every $i \in I$, the hypothesis $H_{W,B}$ is consistent with $(v_i, +)$ and $(2v_i, -)$ iff

$$W^\top v_i + B = -b\frac{W_i}{w_i} + B \geq 0 \ \text{ and } \ W^\top (2v_i) + B = -2b\frac{W_i}{w_i} + B < 0 \ . \tag{20}$$

An inspection of (20) shows that, for every $i \in I$, the hypothesis $H_{W,B}$ can be consistent with $(2v_i, -)$ only if $\mathrm{sgn}(W_i) = \sigma_i$. It follows that (W, B) with $B \geq 0$ can be consistent with $(v_i, +)$ and $(2v_i, -)$ only if $B > 0$. Since $B = 0$ is not possible, Rule 2 applies so that W_j is set to $W_j = 0$ for every $j \in [d] \setminus I$. In combination with the normalization condition $\|W\|_1 = 1$, we obtain $\sum_{i \in I} |W_i| = 1$. Since the components W_j with $j \notin I$ have been set to 0 already, we even have $\mathrm{sgn}(W_i) = \sigma_i$ for all $i = 1, \ldots, d$. Now Rule 3 becomes active. The pair (W, B) is consistent with $(u, +)$ iff

$$W^\top u + B = -b\sum_{i=1}^{d} \sigma_i W_i + B = -b\|W\|_1 + B = -b + B \geq 0 \ ,$$

i.e., iff $B \geq b$. According to Rule 3, the parameter B is now set to $B = b$, and Rule 4 becomes active. Suppose that $i \in I$. Given that $B = b$ and $\mathrm{sgn}(W_i) = \mathrm{sgn}(w_i) = \sigma_i$, condition (20) implies that $-b|W_i|/|w_i| + b = -bW_i/w_i + b \geq 0$. Canceling the factor b from this inequality and resolving for $|W_i|$, one obtains $|W_i| \leq |w_i|$. According to Rule 4, the parameter W_i is chosen so that $|W_i| = |w_i|$. Since the signs of W_i and w_i coincide, it follows that $W = w$. The discussion shows that T is a preference-based teaching set for $H_{w,b}$ of size $2|I| + 1 \leq 2d + 1$. This concludes the proof of the lemma. ∎

Appendix D. Proof of Theorem 23

The proof makes use of the following fairly obvious result:

Lemma 31 *Let R be a d-dimensional box and $w \in \mathbb{R}^d$. Let $P_1, \ldots, P_{2^d}$ denote the vertices of R. Then the boxes spanned by $(w, P_1), \ldots, (w, P_{2^d})$, respectively, fully cover R. Moreover, if $w \in R$, then these boxes form a partition of R.*

See Fig. 2 for an illustration.

We first show that $\mathrm{PBTD}(\dot{\mathcal{R}}_d^k) \leq 2^k d$. Let $R = R_1 \cup \ldots \cup R_m$ be a union of d-dimensional boxes such that $1 \leq m \leq k$ and $R_1 \cap \ldots \cap R_m \neq \emptyset$. Suppose that R is to be taught. For $i = 1, \ldots, m$, let V_i be the set of the (2^d many) vertices of R_i. It suffices to show that $V = V_1 \cup \ldots \cup V_m$ is a spanning set for R. Consider any concept from $\dot{\mathcal{R}}_d^k$ that contains V, say the concept $S = S_1 \cup \ldots \cup S_n$

such that $1 \leq n \leq k$, $S_1 \cap \ldots \cap S_n \neq \emptyset$ and $V \subseteq S$. We have to show that $R \subseteq S$. Let w be a common point of the boxes $S_1, \ldots, S_n$. Clearly S must contain all boxes spanned by w and one of the vertices in V. According to Lemma 31, these boxes cover the box R_i for every $i \in [m]$. Thus $R \subseteq S$, as desired.

We now show that $\mathrm{PBTD}(\dot{\mathcal{R}}_d^k) \geq (k-d+1)2^d$. We may clearly assume that $k \geq d$. Let $R \in \dot{\mathcal{R}}_d^k$ be the union of $m \leq k-d+1$ boxes, say $R = R_1 \cup \ldots \cup R_m$ for $1 \leq m \leq k-d+1$, so that these m boxes have the additional property that $V_i \cap R_j = \emptyset$ for every $1 \leq i \neq j \leq m$. In other words, the vertices of the i-th box R_i are not contained in any other box R_j. See Figure 4 for an illustration in $\mathbb{R}^2$. Let $V = V_1 \cup \ldots \cup V_m$ be the total set of vertices in R. It suffices to show that any set V' that does not contain V is not even a weak spanning set for R w.r.t. $\dot{\mathcal{R}}_d^k$. The crucial observation which confirms this claim is the following one:

Let B be any d-dimensional box and, let P be a vertex of B and let $Q \subset B$ be a finite subset of $B \setminus \{P\}$. Then there exists a union $R(B, P, Q)$ of at most d boxes such that $Q \subset R(B, Q, P) \subset B \setminus \{P\}$.

Why is this observation true? Simply move every hyperface H incident to vertex P slightly towards the interior of B so that P becomes excluded but all vertices in $Q \setminus H$ are still included. See Figure 5 for an illustration.

Why does this observation show that V' is not a weak spanning set for R w.r.t. $\dot{\mathcal{R}}_d^k$? The answer is simple. Suppose that a vertex P of R_i does not belong to V'. Let $Q = R_i \cap V'$. Then we may replace the box R_i in the union R by $R(R_i, Q, P)$. Let $R' = R(R_i, Q, P) \cup \cup_{j \neq i} R_j$. Note that R' is still a union of at most d boxes. Moreover, by construction, we have $V' \subseteq R' \subset R$. Thus V' is not a weak spanning set for R w.r.t. $\dot{\mathcal{R}}_d^k$, as desired.

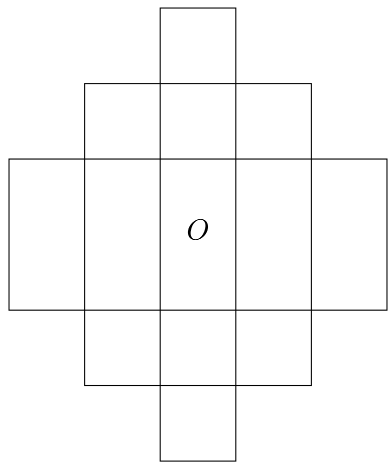

Figure 4: A collection of boxes such no vertex of one box belongs to another one.

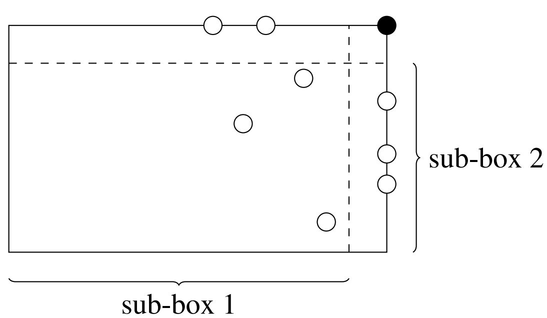

Figure 5: Excluding a vertex of a box without excluding another finite set of points in the box: a union of $d = 2$ sub-boxes suffices.