Learning Families of Closed Sets in Matroids*

Ziyuan Gao[1], Frank Stephan[2], Guohua Wu[3] and Akihiro Yamamoto[4]

[1] Department of Mathematics,
National University of Singapore, Singapore 117543, Republic of Singapore.
Email: ziyuan84@yahoo.com
[2] Department of Mathematics and Department of Computer Science,
National University of Singapore, Singapore 117543, Republic of Singapore.
Email: fstephan@comp.nus.edu.sg
[3] School of Physical and Mathematical Sciences
Nanyang Technological University, Singapore 637371
Email: guohua@ntu.edu.sg
[4] Graduate School of Informatics, Kyoto University
Yoshida Honmachi, Sakyo-ku, Kyoto 606-850, Japan
Email: akihiro@i.kyoto-u.ac.jp

Abstract. In this paper it is studied for which oracles A and which types of A-r.e. matroids the class of all A-r.e. closed sets in the matroid is learnable by an unrelativised learner. The learning criteria considered comprise in particular criteria more general than behaviourally correct learning, namely behaviourally correct learning from recursive texts, partial learning and reliably partial learning. For various natural classes of matroids and learning criteria, characterisations of learnability are obtained.

1 Introduction

This paper extends previous studies on the learnability of mathematically defined objects. While Stephan and Ventsov [28] mainly investigated the learnability of the full class of substructures (all ideals of a ring, all subspaces of a vector space, ...), Harizanov and Stephan [12] looked more at structures where one wants to learn only r.e. subobjects of a certain type and ignores all other subobjects. The subobjects considered by Harizanov and Stephan [12] were the r.e. linear subspaces of a recursive vector space or, more generally, the r.e. closed sets in an r.e. matroid. Indeed, matroids are a generalisation of two known concepts: classes of sets closed under equivalence relations and classes of sets closed under linear combinations inside vector spaces. They are described by a closure operation Φ and in general one is interested in the class of the r.e. closed sets, that is, in the class

$$C_{\Phi} = \{L : \exists H\,[L = \Phi(H)] \wedge \exists \text{ recursive } f\,[H = \text{range}(f)]\}.$$

Harizanov and Stephan [12] investigated the learnability properties of this class in various contexts, in particular for k-thin vector spaces [16]. It turns out that the class C_{Φ} is behaviourally

* Part of this paper was written during F. Stephan's sabbatical leave to the University of Kyoto in August / September 2010. F. Stephan is partially supported by grant R252-000-420-112.

correctly learnable if and only if the matroid is Noetherian, that is, every closed set L is generated by a finite set D, that is, satisfies $L = \Phi(D)$ for a finite set D. Starting from this result, further investigations in the present work go into two direcions.

(a) It is shown how much the qualities of the learner in the learnable case can be improved. Indeed, it is shown that whenever every closed set is r.e. in an r.e. matroid, then the class of all closed sets is behaviourally correctly learnable and this learner can be chosen to be consistent, conservative and confident; actually, the learner is confident in the very strict way as it makes only a constant amount of mind changes where this constant bound is the dimension of the matroid. This carries over the results for recursive Artinian rings [28] to matroids; de Brecht, Kobayashi, Tokunaga and Yamamoto [2] considered more general closure systems of recursive sets which then do not have the mind change bound. Kasprzik and Kötzing [17] considered a related model of string extension learning.

(b) The second direction (which is the main novel part of the present work) is to ask what can be said about the learnability of more general matroids. As those which are not Noetherian are also not behaviourally correctly learnable, one has to consider learning criteria which are more general than behaviourally correct learning. Natural candidates are obtained by either making the notion of text more restrictive or by relaxing the convergence requirement.

- Behaviourally correct learning from recursive texts: Here the restrictions on the text simplify the learning task as the learner deals only with friendly modes of data presentation.
- Partial learning: Here the learner does not converge to a hypothesis but singles out the valid hypothesis by outputting it infinitely often while all other hypotheses are output only finitely often.
- Reliable partial learning: This is a restriction of partial learning where the learner either partially learns a set or outputs on every text for the set each hypothesis only finitely often.

Osherson, Stob and Weinstein [23, Exercise 7.5A] showed that already the class of all r.e. sets is partially learnable and therefore the same holds for the class of the r.e. closed sets of a given r.e. matroid. Hence one would like to study the limits of this learning notion by looking at the more general question for which A-r.e. matroids $(\mathbb{N}, \Phi)$ the class C_Φ^A of all closed A-r.e. sets is learnable under a suitable criterion. Here (reliably) partially learning is considered where the learner itself can conjecture indices from a given acceptable numbering of all A-r.e. sets, but the learner should not have access to the oracle A itself. As the learner can put some of the computations of finding the right hypothesis into the index of an A-recursive enumeration procedure, it follows that there is still some indirect access to A in this way. While partial learners and reliable partial learners cannot exploit that well, the behavioural correct learners can exploit this indirect usage of the oracle and so the corresponding learnability results often generalise to all A-r.e. matroids.

A further learning criterion considered is that of confident partial learning. On one hand, this criterion is much more restrictive than partial learning and does not permit to learn all cofinite sets. On the other hand, certain classes which are not behaviourally correctly learnble are learnable under this criterion and it permits to go beyond Angluin's tell-tale condition [1]. For matroids, it is shown that one one hand the class of closed sets of Noetherian K-r.e. matroids is confidently partially learnable while on the other hand there are some quite simple r.e. matroids where the class of closed r.e. sets is not confidently partially learnable.

Metakides and Nerode [18, 19] gave the definition of an r.e. matroid using a hull operation as follows; the reader is referred to the textbooks and papers of Calude [4, 5], Odifreddi [21, 22], Rogers [26] and Soare [27] for background on recursion theory.

Definition 1 (Metakides and Nerode [18, 19]). An r.e. matroid $(\mathbb{N}, \Phi)$ is given by an enumeration operator Φ mapping subsets of $\mathbb{N}$ to subsets of $\mathbb{N}$ such that the following axioms hold for all sets $R, S \subseteq \mathbb{N}$ and all $a, b \in \mathbb{N}$:

- $S \subseteq \Phi(S)$;
- $\Phi(\Phi(S)) = \Phi(S)$;
- $R \subseteq S \Rightarrow \Phi(R) \subseteq \Phi(S)$;
- If $a \in \Phi(\Phi(S) \cup \{b\}) - \Phi(S)$ then $b \in \Phi(\Phi(S) \cup \{a\})$;
- $\Phi(S) = \bigcup\{\Phi(D) : D \text{ is finite and } D \subseteq S\}$.

Furthermore, sets of the form $\Phi(S)$ are called closed sets and $C_\Phi = \{\Phi(S) : S \text{ is an r.e. subset of } \mathbb{N}\}$. A natural generalisation is that of an A-r.e. matroid where the closure operation is r.e. relative to A and also the class C_Φ^A of interest is the class of A-r.e. closed sets in the matroid.

Remark 2. If $\approx$ is an r.e. equivalence relation then $\Phi(S) = \{a : \exists b \in S\,[a \approx b]\}$ defines an r.e. matroid.

A matroid is called Noetherian iff for every S there is a finite subset $D \subseteq S$ with $\Phi(S) = \Phi(D)$. Noetherian matroids have a constant c, called the dimension, such that one can chose the D above as having at most c elements.

The linear hull in finite-dimensional vector spaces (coded into $\mathbb{N}$) is an example of a Noetherian r.e. matroid.

Let A be a coinfinite subset of $\mathbb{N}$. Then by letting $\Phi(S) = A$ for $S \subseteq A$ and $\Phi(S) = \mathbb{N}$ for $S \not\subseteq A$ one gets an A-r.e. matroid with exactly two closed sets: A and $\mathbb{N}$.

This example can be generalised to matroids for which there is a finite set D such that every closed set S satisfies $S = \Phi(D \cap S)$. For each A-r.e. matroid of this kind it holds that the class C_Φ^A can be explantorily learnt with at most $|D|$ mind changes, as the learner has only to track which elements of D have been seen so far and output an A-r.e. index of the set generated by the corresponding subset of D.

Let A be a proper subset of $\mathbb{N}$. Then there is an A-r.e. matroid having the closed sets A, $A \cup \{a\}$ for every $a \in \mathbb{N} - A$ and $\mathbb{N}$. This is a Noetherian matroid of dimension 2: $A = \Phi(\emptyset)$, $A \cup \{a\} = \Phi(\{a\})$ and $\mathbb{N} = \Phi(\{a, b\})$ where a, b are distinct nonelements of A.

If $(\mathbb{N}, \Phi)$ is a Noetherian r.e. matroid then the relation $a, D \mapsto \Phi(D)(a)$ is recursive. The reason is that one can, for a fixed set I of size c, search until one of the following two conditions hold:

- $a \in \Phi(D)$, that is, $\Phi(D)(a) = 1$;
- There is E with $|E| + 1 = c$, $D \subseteq \Phi(E)$ and $I \subseteq \Phi(E \cup \{a\})$, that is, $\Phi(D)(a) = 0$.

The function $a, D \mapsto \Phi(D)(a)$ is called the dependence relation associated to Φ.

A set S is called *independent* iff $a \notin \Phi(S - \{a\})$ for any $a \in S$. The emptyset is independent and subsets of independent sets are independent. A matroid is Noetherian iff every independent set is finite iff there is a maximum c of the possible cardinalities of independent sets. This constant c is called the dimension of the matroid.

Example 3. An r.e. matroid is called *almost Noetherian* iff there are finitely many r.e. sets $E_0, E_1, \ldots, E_n$ such that every r.e. closed set L is of the form $\Phi(D \cup E_m)$ for some finite D and $m \in \{0, 1, \ldots, n\}$. Every k-thin vector space [16] is an example of an almost Noetherian r.e. matroid which is not Noetherian.

2 Learnability

Gold [11] laid the foundations of inductive inference and defined that a class C of r.e. sets is explanatory learnable iff there is a learner M which outputs on every text T for a language $L \in C$ an infinite sequence $e_0, e_1, \ldots$ of indices such that almost all e_n are actually the same index e for the set L with respect to a preassigned hypothesis space, usually just a fixed acceptable numbering of all r.e. sets. Here a text is an infinite sequence $a_0\, a_1\, \ldots$ of elements from $\mathbb{N}$ plus perhaps a pause symbol and T is *a text for* L if the range of T, that is, the set of all members of $\mathbb{N}$ occurring in T, is equal to L. Furthermore, for a finite prefix σ of T, that is, for a σ of the form $a_0\, a_1\, \ldots\, a_n$, one let $M(\sigma)$ denote the last hypothesis output by M while processing this initial part of T; without loss of generality M starts with a hypothesis of $\emptyset$ so that $M(\sigma)$ is always defined. It is required that M is recursive and thus induces a recursive function from $\mathbb{N}^*$ to $\mathbb{N}$ which maps σ to $M(\sigma)$. In this context it is noted that, whenever talking of sequences of data, it is understood that pause symbols can arise as well and so L^* actually means $(L \cup \{\#\})^*$, but for convenience reasons the first is written in place of the second. It is clear from the context which of the two is meant by "L^*".

Subsequently, the model of Gold had been expanded and thoroughly studied. So in addition to explantory convergence, additional criteria where introduced which were satisfied by some but not all learners.

- M is consistent iff $W_{M(\sigma)}$ contains range(σ) for all strings σ;
- M is conservative iff range$(\sigma\tau) \not\subseteq W_{M(\sigma)}$ whenever $M(\sigma\tau) \neq M(\sigma)$;
- M makes at most c mind changes iff there are no $c+2$ strings $\tau_0, \tau_1, \ldots, \tau_c, \tau_{c+1}$ such that $M(\tau_0\tau_1 \ldots \tau_b) \neq M(\tau_0\tau_1 \ldots \tau_{b+1})$ for $b = 0, 1, \ldots, c$.

Note that when learning the closed sets in matroids, one can assume that the first hypothesis is $\Phi(\emptyset)$ and therefore the definition of learning with bounded mind changes does not need to conjecture a special waiting symbol until the first hypothesis is ready.

The interested reader is referred to the two editions of the text book "Systems that learn" for more information on inductive inference [13, 23]. The next result shows that Noetherian matroids have learners which satisfy all these properties simultaneously.

Theorem 4. *The class C_Φ of the closed sets in a Notherian r.e. matroid of dimension c has a consistent and conservative learner which makes at most c mind changes.*

Proof. The learner produces from the text $a_0\, a_1\, \ldots$ all sets of the form $\{a_{d_0}, a_{d_1}, \ldots, a_{d_e}\}$ where $d_k = \min\{b : \forall \ell < k\, [d_\ell < b] \wedge a_b \notin \Phi(\{a_\ell : \ell < b\})\}$ for $k = 0, 1, \ldots, e$. Note that one can find a_{d_k} only for $k < c$ and the initial hypothesis is $\Phi(\emptyset)$ which will be replaced by sets of the form $\Phi(\{a_{d_0}, a_{d_1}, \ldots, a_{d_e}\})$ for $e = 0, 1, \ldots, c-1$; each mind change is caused by a non-element of the previous hypothesis and thus the learner is conservative. Furthermore, it is consistent as an update is done whenever a datum observed is not generated by the previous hypothesis. ∎

A learner M is behaviourally correctly learning a set L iff M outputs on every text T of L an infinite sequence $e_0\, e_1\, \ldots$ of hypotheses such that $W_{e_k} = L$ for almost all k. Note that a behaviourally correct learner only converges semantically and not syntactically, so it is okay if every hypothesis is different from all previous ones. A class C is behaviourally correctly learnable iff there is a learner M which behaviourally correctly learns every language in C. In the case of learning A-r.e. languages, the learner uses A-r.e. indices from a given acceptable numbering of all A-r.e. sets. Here acceptable is meant in the strict way such that for every further A-r.e. numbering $V_0, V_1, \ldots$ of A-r.e. sets there is a recursive function f such that $W^A_{f(e)} = V_e$ for all e; several learning algorithms would fail in the case that one takes the weaker notion of acceptable numbering where the above f is permitted to be A-recursive in place of recursive. Note that acceptable numberings of all A-r.e. sets in the strict sense exist for every oracle A. The following result is mainly a mirror image of the corresponding results in the world of recursive sets [2, 3], but some adjustments have to be done so that a recursive learner can deal with A-r.e. sets.

Theorem 5. *If $(\mathbb{N}, \Phi)$ is an A-r.e. matroid then the class C^A_Φ of the A-r.e. closed sets is behaviourally correctly learnable (using A-r.e. indices and a recursive learner) iff $(\mathbb{N}, \Phi)$ is Noetherian.*

Proof. If $(\mathbb{N}, \Phi)$ is an A-r.e. Noetherian matroid then one can at each stage conjecture $\Phi(D)$ where D is the data observed so far; note that there is a recursive function f which finds for D an A-r.e. index $f(D)$ of $\Phi(D)$. Namely there is an A-r.e. enumeration of sets $V_D = \Phi(D)$ whose indices are finite sets (given in some canonical form) and f is then the recursive translations from the indices of the sets V_D into indices of the form $W^A_{f(D)}$ with respect to suitable acceptable numbering of all A-r.e. sets. As every closed set L is generated by a finite subset E, that is, $L = \Phi(D)$ for all D with $E \subseteq D \subseteq L$, the indices $f(\mathrm{range}(\sigma))$ are correct for all sufficiently long prefixes σ of any fixed text for L. Hence the class is behaviourally correctly learnable.

For the converse direction, assume that M is a behaviourally correct learner for the matroid $(\mathbb{N}, \Phi)$ and L an A-r.e. closed set. Then there is a locking sequence σ for L, that is, $\sigma \in L^*$ and $M(\sigma\tau)$ is an index for L for all $\tau \in L^*$. As the range D of σ is finite, $\Phi(D)$ is an A-r.e. subset of L and M learns that set as well. As $M(\sigma\tau)$ is an index for L for all $\tau \in \Phi(D)^*$ it follows that these two sets must coincide and $L = \Phi(D)^*$. Hence every closed A-r.e. set is finitely generated.

Now assume that there is a set $\Phi(S)$ such that $\Phi(S) \neq \Phi(D)$ for all finite D. As shown before, the superset $\mathbb{N}$ is finitely generated and $\mathbb{N} = \Phi(D)$ for some finite set D, let D be the minimal such set so that D is independent. As $\Phi(S)$ is not finitely generated, there is an independent subset E of $\Phi(S)$ with $|E| = |D| + 1$. It follows that there is some $a \in E$ such that $D \cup \{a\}$ is independent by the rules of independent sets in matroids. This gives then that $a \notin \Phi(D)$ and

$\Phi(D) \neq \mathbb{N}$, in contradiction to the assumption. Hence every closed set is finitely generated and A-r.e.; so $(\mathbb{N}, \Phi)$ is a Noetherian matroid. ∎

This characterisation shows that the study of behaviourally correct learning in the context of matroids is only interesting when the setting is a bit modified in order to make the learner more powerful. So it might be interesting to ask for which r.e. matroids the class of all closed r.e. sets is learnable from recursive texts. This is not true for the full matroid where every set is closed: the class of all r.e. languages is not behaviourally correctly learnable from recursive texts. Example 6 and Theorem 7 show that there are non-Noetherian matriods for which the class of closed r.e. sets is behaviourally correctly learnable from recursive texts.

Example 6. Let B be a maximal set, that is, a coinfinite r.e. set satisfying that every r.e. superset is either a finite variant of B or is cofinite. Let $\approx$ be defined such that

$$x \approx y \Leftrightarrow \forall z\,[\min\{x, y\} \leq z < \max\{x, y\} \Rightarrow z \in B]$$

and let Φ be the corresponding closure operator. Then the class C_Φ of all r.e. closed sets consists only of finite and cofinite sets. Furthermore, $\mathbb{N}$ is the ascending union of finite members of C_Φ and so C_Φ is not behaviourally correctly learnable. But, as the class of finite and cofinite sets is an indexed family which is a superclass of C_Φ, the class C_Φ is behaviourally correctly learnable from recursive texts [6, Theorem 28 and Corollary 33].

Theorem 7. *Let C_Φ be the class of all closed r.e. sets in an almost Noetherian r.e. matroid $(\mathbb{N}, \Phi)$. Then C_Φ is behaviourally correctly learnable from recursive texts.*

Proof. Let $E_0, E_1, \ldots, E_n$ be the n r.e. sets witnessing that $(\mathbb{N}, \Phi)$ is almost Noetherian and for each m let $E_{m,s}$ be the set of numbers enumerated into E_m within s steps. Let $\tilde{\varphi}_e$ be the e-th recursive text from a list of all partial-recursive texts; note that some of the $\tilde{\varphi}_e$ are partial. Now define the following behaviourally correct learner M for C_Φ.

$M(\sigma)$ outputs an r.e. index for $\Phi(X_\sigma)$ where X_σ is the union of range(σ) and all E_m for which there exists an e with
- $\sigma \preceq \tilde{\varphi}_e$;
- $E_{m,e} \subseteq \text{range}(\sigma)$;
- $E_{m,|\sigma|} \subseteq \text{range}(\tilde{\varphi}_e)$.

The first condition on σ makes sure that $\tilde{\varphi}_e$ is a text extending σ; the second condition wants that the index is permitted by L so that only finitely many e get permitted when $E_m \not\subseteq L$; the third condition wants to make sure that every index e which gets permitted infinitely often is a superset of E_m.

Now it is shown that when learning $L \in C_\Phi$ from a recursive text $\tilde{\varphi}_e$, then for almost all prefixes σ of this text, X_σ is the union of range(σ) and all those E_m which satisfy $E_m \subseteq L$. Hence X_σ contains in particular the "right E_m", so that $L = \Phi(D \cup E_m)$ for some D, as well as some E_m which are "not harmful"; furthermore, the elements of the just mentioned D are also in X_σ for sufficiently long prefixes σ of $\tilde{\varphi}_e$. Hence almost all conjectures of M are correct.

Now the verification of the learner is given in detail. Let σ be a prefix of $\tilde{\varphi}_e$ which is so long that the following conditions are satisfied:

- $e \leq |\sigma|$;
- $D \subseteq \text{range}(\sigma)$;
- if $E_m \not\subseteq L$ then the least $c \in E_m - L$ satisfies that there is no index $d \leq c$ such that $\sigma \preceq \tilde{\varphi}_d \wedge E_{m,|\sigma|} \subseteq \text{range}(\tilde{\varphi}_d)$.

It is clear that the first two conditions are true for sufficiently long prefixes σ of $\tilde{\varphi}_e$. The third condition is true for sufficiently long σ as either φ_d is a finite sequence and therefore cannot extend sufficiently long σ or φ_d is a text which is not for L and therefore does not extend sufficiently long prefixes of the text $\tilde{\varphi}_e$ for L or φ_d is a text for L which then does not enumerate c and therefore $E_{m,|\sigma|} \not\subseteq \text{range}(\tilde{\varphi}_d)$ for sufficiently long $\sigma \preceq \tilde{\varphi}_e$.

As there are only finitely many E_m and as the c in the third condition bounds the number of d over which is quantified in the third condition, it follows that one has the conjunction of all these conditions is satisfied for sufficiently long prefixes σ of $\tilde{\varphi}_e$. As a consequence, X_σ is, for sufficiently long σ, just the union of D and range(σ) and those E_m which are subsets of L; hence $\Phi(X_\sigma) = L$ for those σ and the learner M behaviourally correctly learns L from each recursive text of L. ∎

Remark 8. Every A-recursive behaviourally correct learner (from A-recursive texts) can be turned into a recursive behaviourally correct learner (from A-recursive texts). The reason is that the algorithm to compute the hypothesis can be incorporated into the hypothesis where the computation then, as it is an A-recursive enumeration, has access to the oracle A. Thus the class of the closed sets of any A-r.e. Noetherian matroid has a recursive behaviourally correct learner using A-r.e. indices.

Furthermore, one can generalise Example 6 and Theorem 7 to show that there are A-r.e. matroids where the class of A-r.e. closed sets has a recursive learner which learns these sets behaviourally correctly from A-recursive text outputting A-r.e. indices.

One can generalise behaviourally correct learning in the way that one only requires that almost every hypothesis is a finite variant of the target language; here the number of errors might be different from hypothesis to hypothesis [7, 8, 24]. Harrington [7] showed that the class of all recursive functions can be learnt by an algorithm which outputs on data for a function f a sequence $e_0, e_1, \ldots$ of hypotheses such that almost all φ_{e_k} are finite variants of f and furthermore total. Hence one can also, when reading an A-recursive text, produce a sequence $e_0, e_1, \ldots$ of indices such that $\tilde{\varphi}^A_{e_k}$ is for almost all k a text which is a finite variant of the input text; thus, when looking at the ranges $W^A_{e_k}$ in place of the texts itself, the corresponding indices are almost all finite variants of the set to be learnt. It follows that the class of all A-r.e. sets is behaviourally correctly learnable with at most finitely many errors at almost every hypothesis from A-recursive texts.

3 Partial Learning

Osherson, Stob and Weinstein [23, Exercise 7.5A] introduced the notion of a partial learner and showed that a partial learner can learn the class of all r.e. sets. Minicozzi [20] introduced the notion of reliable learning; although she introduced it for the learning criterion of explanatory learning, it can be brought over to other learning criteria as well. Osherson, Stob and Weinstein [23, Section 4.6.2] also introduced the notion of confident learning where a confident learner has to converge on every text, even a text for a non-r.e. language, to some hypothesis according with respect to the convergence criterion which applies.

Definition 9 (Minicozzi [20], Osherson, Stob and Weinstein [23]). A learner M partially learns a language L iff M, given any text T for L, outputs on this text exactly one index infinitely often and that is an index for L; M partially learns a class C iff M partially learns every $L \in C$.

A partial learner M is reliable iff M for every set L either M outputs on each text for L exactly one index e infinitely often which satisfies $L = W_e$ or M outputs on each text for L each index only finitely often.

A partial learner M is confident iff M for every set L and every text for L outputs on this text exactly one index infinitely often.

Remark 10. One might also ask whether there is a counterpart of partial learning when combined with semantic convergence which gives someting interesting when learning in relativised worlds. Here one could say that a learner M behaviourally correctly partially learns a class C of A-r.e. languages iff M outputs on every text for a language $L \in C$ an infinite sequence $e_0, e_1, \ldots$ of indices such that $W^A_{e_k} = L$ for infinitely many k while for every $H \neq L$ there are only finitely many k with $W^A_{e_k} = H$. Unfortunately this notion does not give anything interesting: Relativising the result of Osherson, Stob and Weinstein [23], there is an A-recursive partial learner which learns all A-r.e. languages; one can in addition have that the learner uses an underlying one-one numbering so that no incorrect language appears on a text by giving infinitely many different indices for that language. Hence one can make a new learner N which on input σ outputs an A-r.e. index e_σ enumerating $W^A_{M(\sigma)}$. That learner has then the required learning properties. Hence this recursive learner learns even the class of all A-r.e. sets using A-r.e. indices.

Remark 11. If M partially learns L then there is a sequence $\sigma \in L^*$ and an index e for L such that for every $\tau \in L^*$ there is an $\eta \in L^*$ such that $M(\sigma\tau\eta) = e$; such a sequence σ is called a locking sequence for L. If such σ, e would not exist then one could find sequences $\tau_0, \tau_1, \ldots \in L^*$ such that for all indices $e_0, e_1, \ldots$ of L the following holds:

- The sequence $\tau_0\tau_1 \ldots \tau_d$ contains all elements of L below d plus perhaps some other elements of L;
- $M(\tau_0\tau_1 \ldots \tau_d\eta) \neq e_d$ for all $\eta \in L^*$.

These two conditions together would then give that $\tau_0\tau_1 \ldots$ is a text for L and that M does not output any index e_d of L on this text infinitely often.

If M is a reliable partial learner learning L, $H \neq L$ and e an index of L then there is a sequence $\sigma \in H^*$ such that $M(\sigma\tau) \neq e$ for all $\tau \in H^*$. If such a σ would not exist, one could construct a text for H on which M outputs e infinitely often in contrast to M being a reliable partial learner.

In the case of a partial learner which is not reliable, the sequence $\sigma \in H^*$ is only guaranteed to exist when H is in the class of sets to be learnt.

Theorem 12. *The class of all cofinite sets is not confidently partially learnable.*

Proof. Assume that M is a partial confident learner. Furthermore, choose $a_0, a_1, a_2, \ldots$ such that for every n, $a_{n+1} > a_n$ and $a_{n+1} > \varphi_e^K(a_0, a_1, \ldots, a_n)$ for all $e \leq n$ where $\varphi_e^K(a_0, a_1, \ldots, a_n)$ is defined. Let $L = \mathbb{N} - \{a_0, a_1, \ldots\}$. By confidence and Remark 11, there is an index d and a $\sigma \in L^*$ such that for all $\tau \in L^*$ there is an $\eta \in L^*$ with $M(\sigma\eta\tau) = d$. With except perhaps one exception, d is not the index of the cofinite set $\mathbb{N} - \{a_0, a_1, \ldots, a_n\}$ and not output infinitely often on any text for this set. Hence, for all but at most one n, there is a sequence $\tau_n \in (\mathbb{N} - \{a_0, a_1, \ldots, a_n\})^*$ such that for all $\eta \in (\mathbb{N} - \{a_0, a_1, \ldots, a_n\})^*$ it holds that $M(\sigma\tau_n\eta) \neq d$. The τ_n and the maximum value occuring in τ_n can be found using a partial K-recursive function, let $\varphi_e^K(a_0, a_1, \ldots, a_n)$ be this value and let n be so large that $n > e$ and $\varphi_e^K(a_0, a_1, \ldots, a_n)$ is defined. By the choice of a_{n+1} it holds that $a_{n+1} > \varphi_e^K(a_0, a_1, \ldots, a_n)$ and that therefore $\tau_n \in L^*$ which contradicts the property according to which σ was chosen. Hence the partial confident learner M cannot exist. ∎

It follows immediately that also the class of all recursive sets is not confidently partially learnable.

Furthermore, if A is a maximal set and $\Phi(S) = A \cup S$ then every r.e. closed set in $(\mathbb{N}, \Phi)$ is either a cofinite superset of A or the union of A with a finite set. This class is also not confidently partially learnable, as one can take the sequence $a_0, a_1, \ldots$ in above proof outside A. Hence there is a quite easy matroid for which the closed r.e. sets are not confidently partially learnable.

The next theorem shows that there is nevertheless a large amount of confidently partially learnable classes. As a corollary one obtains that the class of closed sets of any K-r.e. Noetherian matroid is confidently partially learnable.

Theorem 13. *If a class C of A-r.e. sets is confidently explanatorily learnable relative to the oracle K from informant then C is confidently partially learnable from text without any oracle usage.*

Proof. Let M be the K-recursive learner using A-recursive indices from an acceptable numbering of all A-r.e. sets. Without loss of generality, M codes at every mind change into the index e the finite sets $D_{pos(e)}$ and $D_{neg(e)}$ of positive and negative data, respectively, on which the new hypothesis is based. Furthermore, let M_s be an approximation of M for s steps such that M_s has the same behaviour with respect to the indices output as M.

The new confident partial learner N outputs an index e at least n times iff there is a stage $s > n$ and sets D_{pos^*}, D_{neg^*} such that D_{pos^*} is contained in the first s data items of the text, the first n data items of the text are contained in D_{pos^*}, no data item in D_{neg^*} belongs to the first s data items of the text, $\{0, 1, \ldots, n\} \subseteq D_{pos^*} \cup D_{neg^*}$, M_s with inputs D_{pos^*} and D_{neg^*} outputs e, $D_{pos(e)} \subseteq D_{pos^*}$ and $D_{neg(e)} \subseteq D_{neg^*}$.

Now consider any text T and let e be the index to which M converges on T. The index e is also output infinitely often by N. To see this one shows that N outputs e at least n times by taking for n the parameters $D_{pos^*} = (\{0, 1, 2, \ldots, n\} \cap L) \cup D_{pos(e)}$, $D_{neg^*} = (\{0, 1, 2, \ldots, n\} - L) \cup D_{neg(e)}$ and s so large that all elements of D_{pos^*} have appeared among the first s elements of the text and that M_s on the given input has converged to the value of M on this input.

If a further index d is an old hypothesis which is superseded by e then there is an n such that

$D_{pos(e)} \cup D_{neg(e)} \subseteq \{0, 1, \ldots, n\}$, all elements of $D_{pos(e)}$ have been observed within the first n elements of the text and for all $s \geq n$ and all disjoint supersets $D_{pos^*} \supseteq D_{pos(e)}$ and $D_{neg^*} \supseteq D_{neg(e)}$, M_s with inputs D_{pos^*} and D_{neg^*} outputs e or outputs a hypothesis d' with $D_{pos(d')} \supseteq D_{pos(e)}$ and $D_{neg(d')} \supseteq D_{neg(e)}$. That hypothesis d' is different from d as $D_{pos(d')} \neq D_{pos(d)} \vee D_{neg(d')} \neq D_{neg(d)}$. It follows that d is output by N at most n times.

If a hypothesis d is output by M_s but not by M on inputs $D_{pos(d)}$ and $D_{neg(d)}$ then N outputs d at most where n is the point from which on M_s is equal to M on the inputs $D_{pos(d)}$ and $D_{neg(d)}$.

If a hypothesis d is output by some M_s on some data $D_{pos(d)}$ and $D_{neg(d)}$ where $D_{neg(d)}$ contains some element within the first n elements of T for some n then d is output at most n times.

From this case distinction follows that every d different from e is output only finitely often by N and hence N is a confident partial learner which converges (in the partial sense on the text) to the same indices as M (in the explanatory sense on an informant for range(T)). ∎

This result together with Theorem 4 directly gives the following corollary.

Corollary 14. *Assume that $A \leq_T K$ and $(\mathbb{N}, \Phi)$ is a Noetherian A-r.e. matroid. Then the class C_Φ^A of the closed sets in this matroid is confidently partially learnable.*

Remark 15. Gold's class [11] consisting of all finite sets and one infinite set can be confidently partially learned by outputting a fixed index of the infinite set when a new element is observed and a canonical index for the current range when no new element is observed. But this class is not confidently learnable from informant relative to any oracle, hence Theorem 13 is not a characterisation.

The next result gives a useful criterion for classes to be reliably partially learnable.

Theorem 16. *Every uniformly K-recursive class is reliably partially learnable using a padded version of the given indexing as hypothesis space.*

Proof. Let $V_0, V_1, \ldots$ be a uniformly K-recursive numbering containing the class C to be learnt and let f be a two-place recursive function with $W^K_{f(d,e)} = V_d$ for all d, e. The second parameter is just a padding parameter used to code from which point onwards the language V_d is different from all $V_{d'}$ with $d' < d$. For the reliable partial learner, it is enough to say how often an index has to be output at least; that information can then be used to generate the learner. So the learner for C will output only indices of the form $f(d, e)$ and it will output on a text $a_0\, a_1\, a_2\, \ldots$ at least n times iff there is a stage $s > n$ such that

- e is the least number such that for all $d' < d$ there is an $x < e$ with $V_{d',s}(x) \neq V_{d,s}(x)$ and
- for all $x < n$, $x \in V_{d,s}$ iff $x \in \{a_0, a_1, \ldots, a_s\}$.

If d is the least index of the set L to be learnt and if e is the least number such that for every $d' < d$ there is an $x < e$ with $V_{d'}(x) \neq V_d(x)$ then there is for every n an $s > n$ where $f(d, e)$ qualifies to be output. Hence $f(d, e)$ is output infinitely often.

If $V_d(x) \neq L(x)$ for some x then for all sufficiently large n and all $s > n$, $V_{d,s}(x) \neq L(x)$ and x has appeared within the first n elements of the text iff x is in L. One can then see that $f(d, e)$

is not output n times for these large n and hence output only finitely often.

If there is $d' < d$ with $V_{d'} = V_d$ then every bound e will qualify only finitely often as for all sufficiently large s there is no $x < e$ with $V_{d',s}(x) \neq V_{d,s}(x)$. Hence each $f(d, e)$ is output only finitely often.

If d is the minimal index of L but e is the incorrect bound then one can again see that $f(d, e)$ is output only finitely often.

This case distinction shows that on an input text for a language L, the index $f(d, e)$ is output infinitely often iff $V_d = L$, d is the least index with this property and e is the least bound witnessing that d is the least index. This shows that the learner is correct and also that the learner is reliable. ∎

Example 17. Not every reliably partially learnable class is uniformly K-recursive. For example, let $T_0, T_1, T_2, \ldots$ be a uniformly recursive sequence of trees over $\mathbb{N}^*$ which have at most one infinite branch and for which the tree is unique to the branch, that is, no two trees have the same infinite branch. Furthermore, let $V_e = \{(x, f_e(x)) : x \in \mathbb{N}\}$ where f_e is the unique infinite branch of T_e. Note that the class $\{V_0, V_1, \ldots\}$ needs not to be uniformly K-recursive as there are recursive trees with exactly one infinite branch f such that f is not even arithmetic, that is, much more complicated than K-recursive. Now let g be a recursive function which translates the V-index into that of an A-r.e. set with respect to some acceptable numbering of these sets where A is a suitably chosen oracle. The learner outputs $g(e)$ at least n times iff after some time for $x = 0, 1, \ldots, n$ exactly one pair (x, y_x) has been observed and $y_0\, y_1\, \ldots\, y_x$ is a node on the tree T_e. It is clear that the learner outputs an index $g(e)$ infinitely often iff the observed data is the set V_e.

4 Learnability Properties of Full and Noetherian Matroids

In this section the two extremes of A-r.e. matroids are investigated: The first case is the full matroid where every A-r.e. set is also closed in the matroid, as Φ is the identity operator. The second case are the Noetherian matroids; these satisfy that the closed sets are all uniformly A-recursive.

For the convenience of the reader, the following well-known facts and properties of Turing degrees are summarised: $A \leq_T K$ means that A is Turing reducible to the halting problem. By Shoenfield's Limit Lemma, this is equivalent to saying that A can be approximated in the limit by a sequence of uniformly recursive sets A_s; that is, if $x \in A$ then $x \in A_s$ for almost all s and if $x \notin A$ then $x \notin A_s$ for almost all s. Furthermore, there is a real $r \equiv_T A$ such that there is some recursive sequence $r_0, r_1, \ldots$ of rationals with $r = \lim_s r_s$ and $\{m : r_m < r\} \equiv_T A$. Here the number r itself can be chosen not to be a rational, so $r \neq r_m$ for all m. In the case that A is not recursive (what is the only interesting case), one can in addition chose r such that $\{q \in \mathbb{Q} : q > r\}$ is not an r.e. set.

The set A' denotes the Turing jump of A or the halting problem relative to A; A' is A-r.e. and for every A-r.e. set B there is a recursive function f with $x \in B \Leftrightarrow f(x) \in A'$. A set A has low Turing degree (or just "A is low") iff $A' \leq_T K$, that is, the halting problem relative to A

is not more complicated than the unrelativised halting problem. A is low$_2$ iff $A'' \leq_T K'$. Note that every low set is low$_2$, A is low implies $A \leq_T K$ and A is low$_2$ implies $A \leq_T K'$. Of special interest are those sets A which are low$_2$ and satisfy in addition that $A \leq_T K$. They coincide with those sets such that the class of all A-recursive sets has a uniformly K-recursive listing.

Note that the full matroid satisfies that every set is closed. Hence the class of its closed A-r.e. sets is exactly the class of all A-r.e. sets. Therefore it is interesting to ask when this class is partially learnable. The following theorem gives an answer.

Theorem 18. *Assume that $A \leq_T K'$. Let C be the class of all A-r.e. sets. Then the following statements are equivalent.*

1. *A is low, that is, $A' \leq_T K$;*
2. *C is reliably partially learnable;*
3. *C is partially learnable.*

Proof. The proof is done by a case distinction over several parts.

First, consider the case that A is low. In this case, the class of A-r.e. sets is uniformly K-recursive and hence by Theorem 16 reliably partially learnable; note there that the set $\{(e, x) : x \in W_e^A\}$ is already K-recursive, hence the reliable partial learner from Theorem 16 can actually be made to use the A-r.e. indices. It follows that in this case all three conditions are true.

Second consider the case where $A \leq_T K$ and C is partially learnable. Now let $a_0, a_1, a_2, \ldots$ be a A-recursive sequence of rationals converging to a real r such that the set $\{m : a_m < r\}$ is A-r.e. and has the Turing degree of A'. Such a sequence is known to exist. As $A \leq_T K$ there is a uniformly recursive approximation to the sequence such that $a_m = \lim_s a_{m,s}$. Now let, for a real number q,

$$B_q = \{(m, s) : a_{m,s} < q \vee \exists t > s\, [a_{m,t} \neq a_{m,s}]\}.$$

Furthermore, there is a locking sequence $\sigma \in B_r^*$ and an index e of B_r such that for every $\tau \in B_r^*$ there is an $\eta \in B_r^*$ with $M(\sigma\tau\eta) = e$. Let $b_0 = \max\{a_0, a_1, a_2, \ldots\}$ and note that $b_0 > r$. Now define the following sequence of rationals using the oracle K: Given $b_k > r$, search for a sequence $\tau \in B_{b_k}^*$ such that there is no $\eta \in B_{b_k}^*$ with $M(\sigma\tau\eta) = e$; this τ must exist as M does not output e infinitely often on an text for the r.e. set B_{b_k} which is a proper superset of B_r; note that $B_{b_k} \in C$ as every r.e. set is an A-r.e. set and M outputs on each text for B_{b_k} some other index than e infinitely often. Having τ, τ cannot be a member of B_r^* by the choice of σ. So there is at least one $(m, s) \in B_{b_k} - B_r$ which occurs in τ. Now let $b_{k+1} = \max\{a_{m,s} : (m, s)$ occurs in τ and $a_{m,t} = a_{m,s}$ for all $t > s\}$. So the $a_{m,s}$ over which the maximum is taken satisfy all $a_{m,s} < b_k$ and furthermore one of them satisfies $r < a_{m,s}$ as otherwise $\tau \in B_r^*$. It follows that $r < b_{k+1} < b_k$. The sequence of the b_k is K-recursive and a subsequence of $a_0, a_1, \ldots$; hence the sequence $b_0, b_1, \ldots$ converges to r from above. Furthermore, by the choice of the sequence $a_0, a_1, \ldots$ one can approximate r from below with an A-recursive and thus with a K-recursive sequence. Hence $r \leq_T K$ and $\{m : a_m < r\} \leq_T K$. It follows that $A' \leq_T K$ and A is low. Again all three conditions are satisfied.

Third, consider the case where $A \leq_T K'$ and $A \not\leq_T K$. Furthermore, let $r = \sum_n 2^{-n} A(n)$

or $r = \sum_n 2^{-n}(1 - A(n))$, just take that version for which $\{q \in \mathbb{Q} : q > r\}$ is not K-r.e.; note that $\{q \in \mathbb{Q} : q < r\}$ is an A-r.e. set; it is even A-recursive. There is a K-recursive sequence of rationals such that $r = \lim a_n$ and $\{n : a_n < r\} \equiv_T A \oplus K$. Furthermore, $\{n : a_n > r\}$ is not r.e. relative to K as otherwise $\{q \in \mathbb{Q} : q > r\}$ would be r.e. relative to K. Having this, one makes a proof similar to the previous case. There is a uniformly recursive approximation $a_{m,s}$ to a_m; so for all m and almost all s it holds that $a_m = a_{m,s}$. Now let, for each real q,

$$B_q = \{(m, s) : a_{m,s} < q \vee \exists t > s\,[a_{m,t} \neq a_{m,s}]\}.$$

Note that for every rational q and also for $q = r$ it holds that B_q is A-r.e. and hence there is a learner M which partially learns all sets B_q with $q \in \mathbb{Q}$ and also B_r. Hence, there is a locking sequence $\sigma \in B_r^*$ and an index e of B_r such that for every $\tau \in B_r^*$ there is an $\eta \in B_r^*$ with $M(\sigma\tau\eta) = e$. Let $b_0 = \max\{a_0, a_1, a_2, \ldots\}$ and note that $b_0 > r$. Now define the following sequence of rationals using the oracle K: Given $b_k > r$, search for a sequence $\tau \in B_{b_k}^*$ such that there is no $\eta \in B_{b_k}^*$ with $M(\sigma\tau\eta) = e$; this τ must exist as M does not output e infinitely often on any text for the r.e. set B_{b_k} which is a proper superset of B_r. Having τ, τ cannot be a member of B_r^* by the choice of σ. Hence there is at least one $(m, s) \in B_{b_k} - B_r$ which occurs in τ. Now let $b_{k+1} = \max\{a_{m,s} : (m, s)$ occurs in τ and $a_{m,t} = a_{m,s}$ for all $t > s\}$. So the $a_{m,s}$ over which the maximum is taken satisfy all $a_{m,s} < b_k$ and furthermore one of them satisfies $r < a_{m,s}$ as otherwise $\tau \in B_r^*$. It follows that $r < b_{k+1} < b_k$. The sequence of the b_k is K-recursive and a subsequence of $a_0, a_1, \ldots$; hence the sequence $b_0, b_1, \ldots$ converges to r from above. It follows that $\{m : a_m > r\} = \{m : \exists n\,[b_n < a_m]\}$ is a K-r.e. set in contradiction to its choice. From this contradiction follows that the class of all A-r.e. sets is neither partially learnable nor reliably partially learnable. So all three conditions are not satisfied in this case and hence they are equivalent also in this case. ∎

Theorem 19. *Given a set A, it holds that $A \leq_T K$ iff C_Φ^A is reliably partially learnable for every Noetherian A-r.e. matroid* $(\mathbb{N}, \Phi)$.

Proof. One distinguishes three cases:

- first $A \leq_T K$,
- second $A \not\leq_T K$ and every A-recursive function is majorised by a K-recursive one,
- third there is some $f \leq_T A$ which is not majorised by any K-recursive function.

For the first case, consider $A \leq_T K$. Note that for a Noetherian matroid the class C_Φ^A is uniformly A-recursive and thus uniformly K-recursive. Then C_Φ^A is reliably partially learnable by Theorem 16, using the given A-r.e. indices for the learning process.

For the second part of the proof, one assumes for a contradiction that every Noetherian matroid is reliably partially learnable, that $A \not\leq_T K$ and that every A-recursive function is majorised by a K-recursive one. Now one works over binary strings instead of natural numbers and considers trees of binary strings. One defines a matroid with two A-recursive sets: The set $Br_A = \{\lambda, A(0), A(0)A(1), A(0)A(1)A(2), \ldots\}$ of all prefixes of A and the full binary set $\{0,1\}^*$; $\Phi(S) = Br_A$ for all $S \subseteq Br_A$ and $\Phi(S) = \{0,1\}^*$ for all other S. Now assume that M is a partial reliable learner of this class which outputs the index e infinitely often on the ascending text of Br_A; e is an index for Br_A and therefore not output infinitely often on the text of any other set. Let $f(n)$ be the number of elements seen of the ascending text of Br_A when M outputs e for the n-th time. The function f is A-recursive and hence majorised by a K-recursive function g, that is, $f(n) \leq g(n)$ for all n. Now define the K-recursive tree T to be the set of all strings σ such that every $n \leq |\sigma|$ with $g(n) \leq |\sigma|$ satisfies that M outputs e at least n times on the input sequence $\overline{\sigma}$ which contains all the prefixes $\lambda, \sigma(0), \sigma(0)\sigma(1), \sigma(0)\sigma(1)\sigma(2), \ldots, \sigma$ of σ in ascending order. Then A is an infinite branch of T as M outputs e at least n times on the $g(n)$ first elements of the ascending text of Br_A; on the other hand no set $B \neq A$ is an infinite branch of T as M outputs e only n times on the ascending text of Br_B for some n and then $B(0)B(1)\ldots B(g(n))$ is not on T by the definition of T. So A is the only infinite branch on T and as $T \leq_T K$, it also holds that $A \leq_T K$, in contradiction to the assumption. Hence there is an Noetherian A-recursive matroid whose closed sets cannot be reliably partially learnt.

Third, assume that $f \leq_T A$ is not majorised by any K-recursive function. In other words, the range of f is a set which is hyperimmune relative to K. Now consider the matroid $(\mathbb{N}, \Phi)$ with $\Phi(B) = \mathbb{N} - \text{range}(f)$ if B is disjoint to the range of f and $\Phi(B) = \mathbb{N}$ if B meets the range of f. Assume that there is a reliable partial learner M, an index e and a locking-sequence $\sigma \in (\mathbb{N} - \text{range}(f))^*$ such that for all $\tau \in (\mathbb{N} - \text{range}(f))^*$ there is an $\eta \in (\mathbb{N} - \text{range}(f))^*$ with $M(\sigma\tau\eta) = e$. There is a K-recursive function g such that for every n and every finite set $D \subseteq \{0, 1, \ldots, n\} - \text{range}(\sigma)$ there is a $\tau_D \in (\{0, 1, \ldots, g(n)\} - D)^*$ such that $M(\sigma\tau_D\eta) \neq e$ for all $\eta \in (\mathbb{N} - D)^*$. This function exists as M does not output e infinitely often on any text of a confinite set (by reliability) and therefore one can use K to find the τ_D for each finite set $D \subseteq \{0, 1, \ldots, n\} - \text{range}(\sigma)$ and then take the $g(n)$ to be the maximum of the ranges of the finitely many strings $\sigma\tau_D$ for these D. Now it must be that for $D = \{0, 1, \ldots, n\} - \text{range}(f)$ there is an element of range(f) between n and $g(n)$ as otherwise $\tau_D \in (\mathbb{N} - \text{range}(f))^*$ and $M(\sigma\tau_D\eta) \neq e$ for all $\eta \in (\mathbb{N} - \text{range}(f))^*$ in contradiction to the choice of σ. So it follows that for every $n \in \mathbb{N}$ the intersection $\{n+1, n+2, \ldots, g(n)\} \cap \text{range}(f)$ is not empty and that contradicts the fact that range(f) is hyperimmune relative to K. Thus this case does not occur and the reliable partial learner M cannot exist. So also in this case there exists a Noetherian A-recursive matroid whose closed sets cannot be reliably partially learnt. ∎

The two preceding results give together the following corollary.

Corollary 20. *The class of all A-r.e. sets is reliably partially learnable iff A is low.*

5 The Class of all A-Recursive Sets

The following results deal with the question when the class of all A-recursive sets is partially learnable.

Theorem 21. *The class C of all A-recursive sets is reliably partially learnable iff A is low$_2$ and $A \leq_T K$.*

Proof. If $A \leq_T K$ and A is GL_2 then A is low$_2$ then the class of all A-recursive sets is uniformly K-recursive [14]: One could fix an uniformly A-recursive numbering of all partial A-recursive functions and consider a K-recursive function g which dominates the convergence time of each total function in this list; such a g exists as $A \leq_T K$ and A is low$_2$. Now one defines

$$V_d(x) = \begin{cases} 0 & \text{if } \varphi^A_{g(d+x)}(x)\downarrow = 0 \text{ and } \forall y \leq x\,[\varphi^A_{g(d+x)}(x)\downarrow \in \{0,1\}]; \\ 1 & \text{if } \varphi^A_{g(d+x)}(x)\downarrow = 1 \text{ and } \forall y \leq x\,[\varphi^A_{g(d+x)}(x)\downarrow \in \{0,1\}]; \\ 2 & \text{otherwise.} \end{cases}$$

Note that those indices d where $V_d(x) = 2$ for some x are invalid indices for which the padded versions $f(d,e)$ in Theorem 16 will each be output only finitely often. Hence the learner from Theorem 16 will converge to one index of the form $f(d,e)$ where $V_d(x) \in \{0,1\}$ for all x. This is then a characteristic A-recursive index of the set to be learnt.

It follows from Theorem 19 that whenever $A \not\leq_T K$ then the class of all A-recursive sets is partially reliably learnable.

So consider the last case that $A \leq_T K$ and A is not low$_2$. For the way of a contradiction, assume that M is a reliable partial learner of the class of all A-recursive sets. Let ASC be the set of all strinctly ascending sequences. One can now show the following: $(*)$ For every $\sigma \in ASC$ there is τ_σ with $\sigma\tau_\sigma \in ASC$ such that for all η with $\sigma\tau\eta \in ASC$ it holds that $M(\sigma\tau\eta) > \max(\text{range}(\sigma))$.

If this would not be true, then one could find a $\sigma \in ASC$ such that for every τ with $\sigma\tau \in ASC$ there is an η such that $\sigma\tau\eta \in ASC$ and $M(\sigma\tau\eta) < \max(\text{range}(\sigma))$. This would permit to build more than $\max(\text{range}(\sigma)) + 1$ many recursive ascending texts for pairwise different sets on which M outputs an index below $\max(\text{range}(\sigma))$ infinitely often, in particular one such index would be output infinitely often on at least two different ascending texts of two different sets, thus the condition $(*)$ is true.

There is now a K-recursive function which finds one possible τ_σ for σ. Furthermore, let $\tau_{\sigma,s}$ be an approximation to τ_σ such that $\tau_{\sigma,s} \in ASC$ for all $\sigma \in ASC$ and all s.

Let $g(n)$ be the first s such that for all $t \geq s$ and all $\sigma \in ASC$ with $\max(\text{range}(\sigma)) \leq n$ it holds that $\max(\text{range}(\sigma\tau_{\sigma,s})) < s$ and $\tau_{\sigma,t} = \tau_\sigma$. Note that g in monotonically increasing. As g is a K-recursive function and A is not low$_2$ there is an A-recursive function f such that $f(n) > g(g(n))$ for infinitely many n. Let $T = \lim \sigma_m$ be the ascending text which is obtained by starting with $\sigma_0 = 0$ and taking $\sigma_{m+1} = \sigma_m \tau_{\sigma_m, f(\max(\text{range}(\sigma_m)))}$. Now consider any n with $f(n) > g(g(n))$. There is a first m such that $\max(\text{range}(\sigma_m)) \geq n$. It follows from the definition of g that $\max(\text{range}(\sigma_m)) \leq g(n)$, $g(\max(\text{range}(\sigma_m))) \leq f(n) \leq f(\max(\text{range}(\sigma_m)))$ and $\sigma_{m+1} = \sigma_m\tau_{\sigma_m}$. So, beyond σ_{m+1}, M does not output again any index below n while processing T. As there are infinitely such n it follows that M outputs on T every index only finitely often. Now T is an A-recursive ascending text, hence $B = \text{range}(T)$ is an A-recursive set which M does not partially learn. It follows that there is no reliable partial learner for the class of all A-recursive sets in this case. ∎

Corollary 22. *The class of all A-r.e. sets is reliably partially learnable from A-recursive texts iff $A \leq_T K$ and A is low$_2$.*

Proof. The class $\tilde{C}$ in Theorem 21 is reliably partially learnable only if $A \leq_T K$ and A is low$_2$. The texts considered there were ascending and, as the sets in $\tilde{C}$ are A-recursive, these texts are also A-recursive. It follows that the class of all A-r.e. sets is reliably partially learnable from A-recursive texts only if $A \leq_T K$ and A is low$_2$.

For the converse direction, note that the class of all A-recursive texts of some A-r.e. sets is uniformly K-recursive and that one can therefore reliably partially learn the underlying class of A-recursive functions. Identifying the index of the e-th text $\tilde{\varphi}_e^A$ and its range W_e^A permits then to convert the function learner into a language learner. ▮

One might ask whether Theorem 21 can be improved to use a characterisation based on partial learning in place of reliable partial learning. The answer is that one cannot do this. The next result exhibits that there are uncountably many oracles A such that the class of all A-recursive sets is partially learnable.

Theorem 23. *There are uncountably many oracles A such that the class of all A-recursive sets is partially learnable.*

Proof. For this result, one needs a co-r.e. tree T without deadends and with a branching node above every internal node such that the following holds for each each two different infinite branches A, B and all sets E:

– $A \not\leq_T B$;
– if A has hyperimmune-free Turing degree and $E \leq_T A$ and E is not recursive then $A \leq_T E$.

Groszek and Slaman [10] give a construction with the second property, but they do not write whether the first is satisfied as well. Furthermore, the construction of the minimal rK-degree by Raichev and Stephan [25] provides a co-r.e. tree T with the second property where the first one could easily be added in, one just has to add a fourth condition that whenever for d, e with $d < e$ and $\sigma, \tau \in \{0,1\}^e$ there are marker positions for the markers m_σ, m_τ from that construction there is an x and there are marker positions η above the current position of m_σ and η' above the current position of m_τ inside the current version of the co-r.e. tree considered with $\varphi_e^\eta \downarrow \neq \eta'(x) \downarrow$ then one moves m_σ to η and m_τ to η' and cuts the laid off branches accordingly. The verification is left to the reader.

Now, given such a co-r.e. tree T, it follows from the Hyperimmune-free Basis Theorem of Jockusch and Soare [15] that T has uncountably many infinite branches A of hyperimmune-free Turing degree; fix such an A. Note that there is no other infinite branch B of T such that $B \leq_T A$. Now let f be a recursive one-one function with three inputs such that the following holds:

– if $j = 0$ then $W_{f(i,j,n)}^A$ is the recursive set $\{x : \forall y < x\, [\varphi_i(y) \downarrow \in \{0,1\}] \wedge \varphi_i(x) \downarrow = 1\}$;
– if $j > 0$ then $W_{f(i,j,n)}^A$ is the A-recursive set $\{x : \forall y < x\, [\varphi_i^A(y) \downarrow \in \{0,1\}] \wedge \varphi_i^A(x) \downarrow = 1\}$.

The learner will on a text for an A-recursive set R output exactly one triple $f(i,j,n)$ infinitely often with the following properties:

- $W^A_{f(i,j,n)} = R$;
- If R is recursive then $j = 0$;
- If R is nonrecursive then $j = k+1$ and $A = \varphi_k^R$ and φ_i^B, φ_k^B are total for all oracles B;
- The number n is the number of times some triple $\langle i', j', n' \rangle$ with $\langle i', j' \rangle < \langle i, j \rangle$ had been output.

In order to make this algorithm work one has to specify when a pair $\langle i, j \rangle$ qualifies such that it happens m times that, with current values of n, the index $f(i, j, n)$ is output; the algorithm is then to output $f(i, j, n)$ at stage s iff $\langle i, j \rangle$ qualifies at this step for another time and the pairs $\langle i', j' \rangle < \langle i, j \rangle$ have altogether qualfied n times up to stage s. For this, one defines that $\langle i, j \rangle$ qualifies at least m times iff the following conditions are true.

- $j = 0$ and $\varphi_i(x)$ is defined for $x = 0, 1, \ldots, m$ and there is a stage $s > m$ such that the set D of the first s elements in the text satisfies $D(x) = \varphi_i(x)$ for all $x \leq m$.
- $j = k+1$ and there are bounds s, t such that the set D of the first t elements from the input text satisfies the following conditions:
  - $\varphi_k^E(x)$ is defined for all oracles E and all $x \leq s$ within t steps of computation;
  - $\varphi_i^E(x)$ is defined for all oracles E and all $x \leq m$ within s steps of computation and $\varphi_i^E(x)$ does not query beyond bound s;
  - $\sigma = \varphi_k^D(0)\varphi_k^D(1)\ldots\varphi_k^D(s)$ is not enumerated into the complement of T within t steps;
  - $D(x) = \varphi_i^\sigma(x)$ for all $x \leq m$.

If R is recursive then the first but not the second condition can qualify infinitely often; the second would require that R computes an infinite branch of T which is impossible. If R is not recursive then the first condition cannot qualify infinitely often while the second qualifies infinitely often iff $j > 0$, φ_i and φ_k with $k = j-1$ are truth-table reductions, that is, Turing reductions defined for all oracles and $A = \varphi_k^R$ and $R = \varphi_i^A$. There is a least pair $\langle i, j \rangle$ which qualifies infinitely often and for this the parameter n converges to a value n which is the number of times some pair $\langle i', j' \rangle$ qualifies. Then $f(i, j, n)$ is output infinitely often while all $f(i', j', n')$ with $\langle i', j' \rangle > \langle i, j \rangle$ are output only finitely often as the parameter n for those triples does not converge but goes to infinity. Hence exactly one index $f(i, j, n)$ is output infinitely often and that index satisfies that $R = W^A_{f(i,j,n)}$. ∎

If A is one of the oracles for which the class of all A-recursive sets is partially learnable then one has of course that this is also true for the closed sets of every Noetherian A-r.e. matroid. Hence one gets the following corollary which shows that reliably partially learnable cannot be replaced by partially learnable in Theorem 19.

Corollary 24. *There are uncountably many sets A such that C_Φ^A is partially learnable for every Noetherian matroid* $(\mathbb{N}, \Phi)$.

Acknowledgments. The authors would like to thank Matthew de Brecht, Noam Greenberg, Sanjay Jain and Dan Turetsky for discussions and comments.

References

1. Dana Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.
2. Matthew de Brecht, Masanori Kobayashi, Hiroo Tokunaga and Akihiro Yamamoto. Inferability of closed set systems from positive data. *New Frontiers in Artificial Intelligence*, JSAI 2006. Conference and Workshops, Tokyo, Japan, June 5-9 2006, Revised Selected Papers. Springer LNCS, 4384:265–275, 2007.
3. Matthew de Brecht and Akihiro Yamamoto. Topological properties of concept spaces. *Information annd Computation*, 208:327-340, 2010.
4. Cristian S. Calude. Relativized topological size of sets of partial recursive functions. *Theoretical Computer Science*, 87:347-352, 1991.
5. Cristian S. Calude. *Information and Randomness - An Algorithmic Perspective.* Second Edition, Springer, Heidelberg, 2002.
6. John Case and Sanjay Jain. Synthesizing learners tolerating computable noisy data. *Journal of Computer and System Sciences*, 62:413–441, 2001.
7. John Case and Chris Lynes. Machine inductive inference and language identification. *Proceedings of the nineth International Colloquium on Automata, Languages and Programming*, Springer LNCS 140:107–115, 1982.
8. John Case and Carl H. Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
9. Mark Fulk. Prudence and other conditions on formal language learning. *Information and Computation*, 85:1–11, 1990.
10. Marcia J. Groszek and Theodore A. Slaman. Π^0_1 classes and minimal degrees. *Annals of Pure and Applied Logic*, 87:117–144, 1997.
11. E. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
12. Valentina S.Harizanov and Frank Stephan. On the learnability of vector spaces. *Journal of Computer and System Sciences*, 73:109-122, 2007.
13. Sanjay Jain, Daniel Osherson, James S. Royer and Arun Sharma. *Systems That Learn: An Introduction to Learning Theory.* Second Edition. MIT-Press, Boston, MA., 1999.
14. Carl Jockusch. Degrees in which recursive sets are uniformly recursive. *Canadian Journal of Mathematics*, 24:1092–1099, 1972.
15. Carl Jockusch and Robert Soare. Π^0_1 classes and degrees of theories. *Transactions of the American Mathematical Society*, 173:33–56, 1972.
16. Iraj Kalantari and Allen Retzlaff. Maximal vector spaces under automorphisms of the lattice of recursively enumerable vector spaces. *The Journal of Symbolic Logic*, 42:481–491, 1977.
17. Anna Kasprzik and Timo Kötzing. String extension learning using lattices. *Language and Automata Theory and Applications*, fourth international conference, LATA 2010, Trier, Germany, May 2010, Proceedings. *Springer LNCS*, 6031:380–391, 2010.
18. George Metakides and Anil Nerode. Recursively enumerable vector spaces, *Annals of Mathematical Logic*, 11:147–171, 1977.

19. George Metakides and Anil Nerode. Recursion theory on fields and abstract dependence, *Journal of Algebra*, 65:36–59, 1980.
20. Eliana Minicozzi. Some natural properties of strong-identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.
21. Piergiorgio Odifreddi. *Classical Recursion Theory*, *Studies in Logic and the Foundations of Mathematics*, volume 125. North-Holland, Amsterdam, 1989.
22. Piergiorgio Odifreddi. *Classical Recursion Theory II*, *Studies in Logic and the Foundations of Mathematics*, volume 143. Elsevier, Amsterdam, 1999.
23. Daniel N. Osherson, Michael Stob and Scott Weinstein. *Systems That Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists*, MIT Press, 1986.
24. Daniel Osherson and Scott Weinstein. Criteria of language learning. *Information and Control*, 52:123–138, 1982.
25. Alexander Raichev and Frank Stephan. A minimal rK-degree. *Computational Prospects of Infinity*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, 15:261–269, 2008.
26. Hartley Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, New York, 1967.
27. Robert I. Soare. *Recursively enumerable sets and degrees.* Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987.
28. Frank Stephan and Yuri Ventsov. Learning algebraic structures from text. *Theoretical Computer Science*, 268:221–273, 2001.