Mathematical Logic

Learning theory in the arithmetic hierarchy II

Achilles A. Beros[1] · Konstantinos A. Beros[2] · Daniel Flores[1] · Umar Gaffar[1] · David J. Webb[1] · Soowhan Yoon[1]

Abstract

The present work determines the arithmetic complexity of the index sets of u.c.e. families which are learnable according to various criteria of algorithmic learning. Specifically, we prove that the index set of codes for families that are TxtFex_b^a-learnable is Σ_4^0-complete and that the index set of TxtFex_*^*-learnable and the index set of $\mathrm{TxtFext}_*^*$-learnable families are both Σ_5^0-complete.

Keywords Algorithmic learning theory · Computability theory · Recursion theory

Mathematics Subject Classification 03D15

1 Introduction

At its heart, the goal of algorithmic learning theory is a formalization of the intuitive concept of learning. Specifically, algorithmic learning theory aims to formalize the process of extrapolation. As a rudimentary example of extrapolation, if one is given the finite sequence

$$2, 4, 6, 8$$

a natural conclusion is that the sequence represents an initial segment of the infinite sequence of all even numbers. The process of extrapolating a generative algorithm

✉ Achilles A. Beros
beros@math.hawaii.edu

Konstantinos A. Beros
berosk@miamioh.edu

[1] Department of Mathematics, University of Hawaii at Manoa, 2565 McCarthy Mall, Keller Hall 401A, Honolulu, HI 96822, USA

[2] Department of Mathematics, Miami University, 123 Bachelor Hall, 301 S. Patterson Ave., Oxford, OH 45056, USA

from given finite data is modeled in a number of different ways in algorithmic learning theory. The standard model of learning – introduced in Gold [8] – is *explanatory learning* or *EX-learning*. A computable function $M : 2^{<\omega} \to \omega$ *EX-learns* a computably enumerable (c.e.) set $A \subseteq \omega$ iff, for any enumeration $f : \omega \to \omega$ of A,

$$(\exists e)\big(W_e = A \land (\forall^\infty n)(M(f \restriction n) = e)\big)$$

Recall that W_e is the c.e. set with Turing code e. The topic of learning families of c.e. sets has been studied extensively (for example, see [1,10]. A computable machine M EX-learns a uniformly c.e. (u.c.e.) family $\mathcal{F}$ if it EX-learns every $A \in \mathcal{F}$. While a family consisting of a single c.e. set can always be learned by an appropriate M, there are many families which are not EX-learnable. For instance, the family consisting of the finite sets $\{0, \ldots, n\}$ together with ω cannot be EX-learned. In this case, any learning machine which identifies the finite sets in this family will fail to identify ω from suitably chosen enumerations.

With this in mind, it is natural to ask which u.c.e. subsets of $\mathcal{P}(\omega)$ are EX-learnable. In [3], author A. Beros took up this question from the point of view of definable complexity. Specifically, A. Beros showed that the set of Turing codes for u.r.e families which are EX-learnable is Σ_4^0-complete. The lower bound was achieved by a finite injury priority construction which reduced an arbitrary Σ_4^0 predicate the set EXL of Turing codes for EX-learnable u.c.e. families. The upper bound (i.e., the statement that EXL is Σ_4^0) relied upon a result of Blum–Blum [4] to the effect that a u.c.e. family is EX-learnable if it is EX-learnable from computable enumerations. Whereas the definition of an EX-learnable family is apparently Π_1^1, the theorem of Blum–Blum guarantees that the initial universal quantifier can be restricted to only those functions $f : \omega \to \omega$ which are computable.

There are a number of learning criteria besides EX-learning which differ according to what type of convergence is required from the learning machine M [5,6] as well as the types of errors M is permitted to make [7] and the use of oracles [11]. For instance, a machine M *BC-learns* a set A if, given any enumeration f of A,

$$(\forall^\infty n)(W_{M(f \restriction n)} = A).$$

In other words, although $M(f \restriction n)$ is allowed to change infinitely many times as n increases, $M(f \restriction n)$ must be a Turing code for A for cofinitely many n. In [3], A. Beros showed that the set of all codes for u.c.e. families which BCL-learnable is Σ_5^0-complete. In this case, the upper bound required a BC-learning analogue of Blum–Blum's theorem: if a family is BC-learnable from Δ_2^0 enumerations, it is BC-learnable from arbitrary enumerations.

Vacillatory learning models strike a compomise between the infinitely many distinct hypotheses of BC-learning and the discrete convergence of EX-learning. Specifically, the learning machine is permitted to change its output infinitely many times (or *vacillate*) between finitely many distinct codes for the set to be learned, i.e.,

$$\{M(f \restriction n) : n \in \omega\}$$

must be a finite set.

There are further learning criteria whereby the learning machine is permitted to output codes for sets which are merely equal to the set to be learned modulo finite sets. For instance, a computable function $M : 2^{<\omega} \to \omega$ TxtFex$^*_*$-*learns* a c.e. set A if, given any enumeration $f : \omega \to \omega$ of A, there is a finite set F such that

- $(\forall^\infty n)(M(f \restriction n) \in F)$ and
- $(\forall e \in F)(W_e$ and A have finite symmetric difference).

TxtFex$^*_*$-learning is therefore a criterion which permits both vacillation and errors. Placing fixed bounds on the number of vacillations and/or the number of errors permitted yields an infinite number of distinct learning criteria. If the symmetric difference $W_e \triangle A$ is permitted to have cardinality at most a and only b many vacillations are allowed, the resulting learning criterion is denoted TxtFex^a_b. If $a \in \omega$, a family is TxtFex$^a_*$-learnable if only a many errors, but arbitrarily many vacillations are permitted. Details of these definitions are given in the next section.

The results in the present work cover a variety of different vacillatory learning criteria and establish definable complexity results in each case.

Theorem 3.5

For $a \in \omega$ and $b \in \omega \cup \{*\}$, the set of Turing codes for u.c.e. families which are TxtFex^a_b-learnable is Σ^0_4-complete.

Theorem 4.2

The set of Turing codes for u.c.e. families which are TxtFex$^*_*$-learnable is Σ^0_5-complete.

This latter result uses machinery developed by author A. Beros in his doctoral thesis and published in [3].

2 Preliminaries and notation

2.1 Fundamental notions

For the most part, the terminology of this paper is standard. For reference, this section contains a summary of the relevant notation. We refer the reader to [12] for more on computability theory and to [9] for learning theory.

For $e \in \omega$, let φ_e be the Turing machine coded by e. The domain of φ_e is denoted by W_e. If M is any computable function, $M(\alpha) \downarrow$ indicates that M converges on input α. The notation $M(\alpha) \downarrow_t$ indicates that M has converged on input α within t computation stages. Similarly, $M(\alpha) \uparrow$ indicates that M never halts on input α and $M(\alpha) \uparrow_t$ indicates that M has not yet converged on input α after t stages.

Let $|A|$ denote the cardinality of A. For sets $A, B \subseteq \omega$, let $A \triangle B$ denote the symmetric difference of A and B. Let $A =^* B$ indicate that A and B are equal mod finite, i.e., $|A \triangle B| < \infty$. If $|A \triangle B| \leq a$ for some fixed $a \in \omega$, write $A =^a B$.

If $f : \omega \to \omega$ is any function, let $f \restriction n$ denote the restriction of f to the set $\{0, \ldots, n-1\}$.

For $A \subseteq \omega$, let $A^{<\omega}$ denote the set of finite strings of elements of A. If $\sigma, \tau \in A^{<\omega}$, write $\sigma \preceq \tau$ to indicate that σ is a prefix of τ. Likewise, if $f : \omega \to A$, let $\sigma \prec f$ indicate that σ is a prefix of f.

2.2 Vacillatory learning preliminaries

The present paper focuses on the following family of vacillatory learning criteria.

Definition 2.1 Let $a, b \in \omega \cup \{*\}$. A partial computable function $M : \omega^{<\omega} \to \omega$ TxtFex$_b^a$ -*identifies* a c.e. set $A \subseteq \omega$ iff, for any $f : \omega \to A$ which enumerates A, there is a set $F \subseteq \omega$ such that

- F is finite if $b = *$ and $|F| \leq b$ if $b \in \omega \setminus \{0\}$,
- $(\forall^\infty n)(M(f \restriction n) \downarrow \ \& \ M(f \restriction n) \in F)$ and
- $(\forall e \in F)(W_e =^a A)$.

When $a = *$, the only requirement in the final clause is that W_e equal A modulo finite sets.

The function M *TxtFex$_b^a$-learns* a u.c.e. family $\mathcal{F} \subseteq \mathcal{P}(\omega)$ iff it identifies each $A \in \mathcal{F}$. Finally, a u.c.e. family is *TxtFex$_b^a$-learnable* iff there exists a machine M which learns $\mathcal{F}$.

Theorem 4.2 below concerns another form of vacillatory learning.

Definition 2.2 A computable function on strings *TxtFext$_*^*$-identifies* a c.e. set A iff it TxtFex$_*^*$-identifies A and the finite set F in the definition above has the additional property that

$$(\forall e_1, e_2 \in F)(W_{e_1} = W_{e_2}).$$

The definitions of "TxtFext$_*^*$-learning" and "TxtFext$_*^*$-learnable" are analogous to the ones above in the TxtFex$_*^*$ case.

The following is a useful observation.

Proposition 2.3 *If M TxtFex$_b^a$-identifies or TxtFext$_*^*$-identifies a c.e. set $A \subseteq \omega$, there is a string $\sigma \in A^{<\omega}$ such that, for every $\tau \in A^{<\omega}$,*

$$M(\sigma^\frown \tau) \in \{M(\alpha) : \alpha \preceq \sigma\}.$$

Proof Suppose otherwise. This implies that every $\sigma \in A^{<\omega}$ has an extension $\sigma' \in A^{<\omega}$ such that

$$M(\sigma') \notin \{M(\alpha) : \alpha \preceq \sigma\}.$$

In other words, the set

$$X = \{\sigma \in A^{<\omega} : (\exists \sigma' \succeq \sigma)(M(\sigma') \notin \{M(\alpha) : \alpha \preceq \sigma\})\}$$

is dense in $A^{<\omega}$. It is therefore possible to inductively define an enumeration f of A such that

$$\{M(f \upharpoonright n) : n \in \omega\}$$

is infinite. In particular, M does not identify A from f. This is a contradiction. Note that, since X is c.e., f can be made recursive.   □

What follows is a powerful result due to John Case, reproduced for the reader's convenience. This result is a cornerstone of the complexity upper bounds obtained in the present paper.

Theorem 2.4 (Case [6], Theorem 5.5) *Let $a, b \in \omega \cup \{*\}$. If M TxtFex$_b^a$-learns a family $\mathcal{F} \subseteq \mathcal{P}(\omega)$ from computable enumerations, then there exists a total computable $\widehat{M} : \omega^{<\omega} \to \omega$ which TxtFex$_b^a$-learns $\mathcal{F}$ from arbitrary enumerations.*

Remark 2.5 A routine modification of Case's proof establishes the same result for TxtFext$_*^*$-learning.

3 Vacilation with bounded error

Definition 3.1 For $a \in \omega \cup \{*\}$ and $b \in \omega \cup \{*\}$, let FEXL$_b^a$ denote the index set of natural numbers coding u.c.e. families which are TxtFex$_b^a$-learnable.

Theorem 3.2 *For $a \in \omega$ and $b \in \omega \cup \{*\}$, the index set FEXL$_b^a$ has a Σ_4^0 description.*

First, a lemma.

Lemma 3.3 $\{\langle y, z \rangle : m \leq |W_y \triangle W_z| \leq n\}$ *is Δ_3^0 for all $m, n \in \omega$.*

Proof Fix $m, n \in \omega$ and let $A = \{\langle y, z \rangle : m \leq |W_y \triangle W_z| \leq n\}$.

$$\langle y, z \rangle \in A \iff |W_y \triangle W_z| \leq n \wedge |W_y \triangle W_z| \geq m.$$

In other words,

$$(\exists x_0, \ldots, x_{m-1}, s)(\forall t)\Big[(\forall i, j \leq m)(i \neq j \implies x_i \neq x_j) \wedge$$
$$\bigwedge_{i \leq m-1} (x_i \in W_{y,s} \wedge x_i \notin W_{z,t}) \vee (x_i \in W_{z,s} \wedge x_i \notin W_{y,t})\Big]$$
$$\wedge (\forall x_0, \ldots, x_n)\Big[(\forall i, j \leq n)(i \neq j \implies x_i \neq x_j) \implies$$
$$\bigvee_{i \leq n} (x_i \in W_y \cap W_z) \vee (x_i \notin W_y \cup W_z)\Big]$$

which is the conjunction of a Π_2^0 statement with a Σ_2^0 statement.   □

Proof (Proof of Theorem 3.2) For each $e \in \omega$, let $\mathcal{L}_e$ denote the u.c.e. family coded by e. For each i, let

$$\mathcal{L}_e[i] = \{j : \langle i, j \rangle \in \mathcal{L}_e\}.$$

That is, $\mathcal{L}_e[i]$ is the ith set in the family $\mathcal{L}_e$. For each $p \in \omega$, let

$$M_p : 2^{<\omega} \to \omega$$

be the learning machine coded by p.

First of all, if a family $\mathcal{L}_e$ is learnable, say by M, there is a total learner $\widehat{M}$ which also learns $\mathcal{L}_e$. Indeed, let $\widehat{M}$ be defined by

$$\widehat{M}(\sigma) = \begin{cases} M(\sigma \restriction n) & \text{where } n \text{ is largest such that } M(\sigma \restriction n) \\ & \text{converges in } |\sigma| \text{ computation stages,} \\ 0 & \text{otherwise.} \end{cases}$$

The learner $\widehat{M}$ learns every subset of ω that M learns. To see this, suppose that M learns a set $A \subseteq \omega$. Let $f : \omega \to \omega$ be any enumeration of A. Let n_0 be large enough that $W_{M(f \restriction n)} =^a A$ for each $n \geq n_0$. Let $n_1 \geq n_0$ be large enough that $M(f \restriction n_0) \downarrow$ within n_1 computation stages. It follows that, if $\sigma \prec f$ and $|\sigma| \geq n_1$, then $\widehat{M}(\sigma) = M(f \restriction n)$ for some $n \geq n_0$. Hence, $W_{\widehat{M}(\sigma)} =^a A$. This shows that $\widehat{M}$ learns A.

Next, by Case's Theorem (Theorem 2.4), it follows that a family is learnable iff it is learnable from computable enumerations. Therefore, a family $\mathcal{L}_e$ is TxtFex$_b^a$ learnable iff there exists $p \in \omega$ such that, for all $x, i \in \omega$,

(1) M_p is total and,
(2) if φ_x is total and enumerates $\mathcal{L}_e[i]$, then
(a) there exists a finite set $D \subseteq \omega$ such that
(i) $|D| \leq b$ (if $b \neq *$) and
(ii) $(\forall^\infty n)(\forall t)(M_p(\varphi_x \restriction n) \downarrow_t \implies M_p(\varphi_x \restriction n) \in D)$
(b) and $(\forall j)(|W_j \triangle \mathcal{L}_e[i]| > a \implies (\forall^\infty n)(\forall t)(M_p(\varphi_x \restriction n) \downarrow_t \implies M_p(\varphi_x \restriction n) \neq j))$.

Note that, in the case of $b = *$, condition (i) above is removed. Condition (b) says that if j is an incorrect code for $\mathcal{L}_e[i]$, there are cofinitely many n such that $M_p(\varphi_x \restriction n) \neq j$.

To see that the predicate "$\mathcal{L}_e$ is learnable" is Σ_4^0, observe first that condition (1) above is Σ_2^0. Next, notice that condition (i) is computable and condition (ii) is Σ_2^0. Finally, condition (b) is Π_3^0 by Lemma 3.3. It follows that condition (2) is Π_3^0 and therefore, "$\mathcal{L}_e$ is learnable" is Σ_4^0. □

The next objective is to show that TxtFex$_b^a$ is Σ_4^0-hard. This is established by Theorem 3.5 below. The following example serves the basic building block in the proof of this theorem.

Example 3.4 For each $e \in \omega$, define a family $\mathcal{L}_e$ of c.e. sets as follows:

- if $e \in$ INF, i.e., W_e is infinite, let $\mathcal{L}_e = \{[e, \infty)\}$
- if $e \in$ FIN, i.e., W_e is finite, let $\mathcal{L}_e = \{[e, e + i + |W_e|] : i \in \omega\}$

Now let $\mathcal{F} = \bigcup_e \mathcal{L}_e$.

In the first place, $\mathcal{F}$ is u.c.e. To see this, observe that $\mathcal{F}$ is the collection of all sets of the form

$$L_{e,i} = \{x : (\exists s)(e \leq x \leq e + i + |W_{e,s}|)\}.$$

Note that, if $e \in$ INF, then $L_{e,i} = [e, \infty)$ for each $i \in \omega$.

The key attribute of $\mathcal{F}$ is that it is not TxtFex_b^a-learnable for any $a \in \omega$ and $b \in \omega \cup \{*\}$. To prove this, it will suffice to show that $\mathcal{F}$ is not TxtFex_*^a-learnable for any $a \in \omega$. To this end, suppose that the contrary is true. Say M TxtFex_*^a-learns $\mathcal{F}$. By Case's Theorem (Theorem 2.4 above), it is no loss to assume that M is total. Consider the Σ_2^0 predicate Q given by

$$Q(e) \iff (\exists \sigma \in [e, \infty)^{<\omega})(\forall \tau \in [e, \infty)^{<\omega})(M(\sigma^\frown \tau) \in \{M(\alpha) : \alpha \preceq \sigma\}).$$

Since INF is Π_2^0-complete, a contradiction will be obtained by showing that INF $= Q$.

Suppose $e \in$ INF. As M identifies $L_{e,i} = [e, \infty)$, it follows from Proposition 2.3 that there is a string $\sigma \in [e, \infty)$ with

$$M(\sigma^\frown \tau) \in \{M(\alpha) : \alpha \preceq \sigma\}$$

for all $\tau \in [e, \infty)^{<\omega}$. In other words, $Q(e)$ holds.

For $e \in$ FIN, the machine M identifies each

$$L_{e,i} = [e, e + i + |W_e|].$$

Given $\sigma \in [e, \infty)^{<\omega}$, let $i > \max(\mathrm{content}(\sigma))$ be large enough that, for all $\alpha \preceq \sigma$,

$$W_{M(\alpha)} \neq^a [e, e + i + |W_e|].$$

Since M identifies $[e, e + i + |W_e|]$, there exists $\tau \in [e, \infty)^{<\omega}$ such that

$$W_{M(\sigma^\frown \tau)} =^a [e, e + i + |W_e|].$$

In particular, $M(\sigma^\frown \tau) \notin \{M(\alpha) : \alpha \preceq \sigma\}$. Thus, $\neg Q(e)$ holds. This yields a contradiction since it shows that INF $= Q$ and Q is Σ_2^0.

In the proof that follows, $\mathcal{L}_x$ and $\mathcal{F}$ are as in Example 3.4. The proof is based on ideas which arose from correspondence between the first author and Frank Stephan.

Theorem 3.5 *For $a \in \omega$ and $b \in \omega \cup \{*\}$, the index set $FEXL_b^a$ is Σ_4^0-hard.*

Proof In what follows, "learnable" means "TxtFex$_b^a$-learnable" for fixed $a \in \omega$ and $b \in \omega \cup \{*\}$. No distinction need be made as the argument is identical for all such a and b.

Fix a Σ_4^0 predicate P and a total computable function $f : \omega^2 \to \omega$ such that for each $n \in \omega$,

$$P(n) \implies (\forall^\infty x)(f(n, x) \in \text{COINF})$$

and

$$\neg P(n) \implies (\forall x)(f(n, x) \in \text{COF}).$$

Also, assume that f is strictly increasing in the second coordinate, i.e.,

$$f(n, 0) < f(n, 1) < \dots$$

for all $n \in \omega$.

Fix $n \in \omega$. Given $x, e, i, p \in \omega$, define $R_{x,e,i,p}^n \subseteq \omega$ as follows. First of all, assume $p \in W_{f(n,x)}$; otherwise, $R_{x,e,i,p}^n$ will be undefined. Let $I_{p,s}$ be the largest interval contained in $W_{f(n,x),s}$ with $p \in I_{p,s}$.

$$R_{x,e,i,p}^n = \left\{ x : (\exists s)(e \leq x \leq e + i + \min\{|I_{p,s}|, |W_{e,s}|\} \right\}$$

Further, define $\mathcal{R}_{x,e,p}^n = \{R_{x,e,i,p}^n : (\exists s)(i \leq |I_{p,s}|)\}$. Now let

$$\mathcal{F}_n = \bigcup \left\{ \mathcal{R}_{x,e,p}^n : x \in \omega \wedge f(n, x) \leq e < f(n, x + 1) \wedge p \in W_{f(n,x)} \right\}.$$

Since $\mathcal{F}_n$ is u.c.e., the proof will therefore be completed by showing that $\mathcal{F}_n$ is learnable iff $P(n)$.

Recall now the families $\mathcal{L}_e = \{L_{e,i} : i \in \omega\}$ from Example 3.4 above. Also, let $A_{n,x,e}$ be the set of those members of $\mathcal{F}_n$ whose minimum element is e. Although the dependence of $A_{n,x,e}$ on x is redundant, including it makes some formulas simpler to write.

Claim 1. If $f(n, x) \leq e < f(n, x + 1)$, $i \in \omega$ and $W_{f(n,x)}$ is cofinite, then $R_{x,e,i,p}^n = L_{e,i}$ for all but finitely many $p \in W_{f(n,x)}$.

To see why this is so, fix $p \in W_{f(n,x)}$ such that p is contained in an infinite interval of $W_{f(n,x)}$. Note that all but finitely many $p \in W_{f(n,x)}$ are contained in such an infinite interval. For such p, it follows from the definition of $R_{x,e,i,p}^n$ that

$$R_{x,e,i,p}^n = [e, e + i + |W_e|] = L_{e,i}$$

if W_e is finite, and

$$R_{x,e,i,p}^n = [e, \infty) = L_{e,i}$$

if W_e is infinite. This establishes the claim.

Claim 2. If $W_{f(n,x)}$ is coinfinite, $R^n_{x,e,i,p}$ is finite for all e, i, p.

This follows from the fact that the minimum in the definition of $R^n_{x,e,i,p}$ will be bounded by the length of the interval in $W_{f(n,x)}$ which contains p.

Claim 3. If W_e is infinite and $W_{f(n,x)}$ is cofinite, then $\mathcal{A}_{n,x,e}$ is a finite - hence learnable - family.

Indeed, in this case, claim 1 implies that, for all but finitely many $p \in W_{f(n,x)}$,

$$R^n_{x,e,i,p} = L_{e,i} = [e, \infty).$$

Hence $\mathcal{A}_{n,x,e}$ consists of the infinite set $[e, \infty)$ together with finitely many finite sets.

Towards the end of showing that $\mathcal{F}_n$ is learnable iff $P(n)$, suppose first that $P(n)$ holds. It follows from the choice of f that $W_{f(n,x)}$ is coinfinite for all but finitely many x, i.e.,

$$H = \{x \in \omega : W_{f(n,x)} \text{ is cofinite}\}$$

is a finite set. Consequently, the set

$$G = \{\langle x, e \rangle : x \in H \land e \in \text{INF}\}$$

is also finite. Let

$$\mathcal{B} = \bigcup\{\mathcal{A}_{n,x,e} : \langle x, e \rangle \in G\}$$

Since H is finite, claim 3 implies that $\mathcal{B}$ is a finite set as well. Hence, $\mathcal{B}$ is learnable, say by a total machine M. It follows from the definition of $\mathcal{F}_n$ that

$$\mathcal{F}_n \subseteq \mathcal{B} \cup \mathcal{C}.$$

where $\mathcal{C}$ is a collection of finite sets. Therefore, if N is a total machine which learns all finite sets, $\mathcal{F}_n$ is learned by the machine $\widehat{M}$ with

$$\widehat{M}(\sigma) = \begin{cases} M(\sigma) & \text{if } (\exists x \in H)\big(\langle x, \min(\text{content}(\sigma)) \rangle \in G\big), \\ N(\sigma) & \text{otherwise.} \end{cases}$$

for each string $\sigma \in \omega^{<\omega}$.

Finally, suppose that $\neg P(n)$. In this case, $W_{f(n,x)}$ is cofinite for every $x \in \omega$ and hence, for each pair e, i, $R^n_{x,e,i,p} = L_{e,i}$ all but finitely many p. It follows that $\mathcal{F}_n$ contains the non-learnable family $\mathcal{F}$ from Example 3.4. In particular, $\mathcal{F}_n$ is not learnable. This complete the proof. □

4 Unbounded vacilation and error

4.1 Upper bounds

An upper bound for $FEXTL_*^*$ is obtained in a manner similar to $FEXTL_b^a$. The only distinction between the descriptions is $W_e = W_{e'}$ has a Π_2^0 description whereas $W_e =^* W_{e'}$ has a Σ_3^0 description. The defining arithmetic formula for $FEXTL_*^*$ is as follows.

$$(\exists a)(\forall k, i)\Big((M_a \text{ is total}) \wedge (\phi_k \text{ is total}) \wedge (\phi_k \text{ enumerates } L_i) \to \psi(k, a, i)\Big) \quad (1)$$

where we define $\psi(k, a, i)$ to be

$$(\exists s, c)(\forall t \geq s)$$
$$\Big(M_a(\phi_k \upharpoonright t) \in D_c \wedge \forall x, y \in D_c(W_x = W_y)\Big) \wedge (\forall n)\Big(W_{M_a(\phi_k \upharpoonright n)} =^* L_i$$
$$\vee (\exists m)(\forall u \geq m)\big(M_a(\phi_k \upharpoonright m) \neq M_a(\phi_k \upharpoonright n)\big)\Big).$$

The last formula is Π_4^0. Thus, (1) is Σ_5^0 and characterizes TxtFex-learning because, for any family coded by a number e which satisfies formula (1), there is a learner whose hypotheses converge to correct hypotheses on every computable enumeration and if e fails to satisfy formula (1), every learner must fail for some set in the family.

We have, therefore, exhibited a Σ_5^0 description of $FEXTL_*^*$. Notice the weakening of $FEXTL_*^*$ to $FEXL_*^*$ must also have a Σ_5^0 description. It is meaningful to study both $FEXTL_*^*$ and $FEXL_*^*$ as $FEXTL_*^* \neq FEXL_*^*$, a result due to the first author that appears in [2].

4.2 Lower bounds

Lemma 4.1 *Let $M = M_m$ be a computable learning machine and W_e a c.e. set. There is a family $\mathcal{F}_{m,e}$, uniformly computable in both m and e, such that:*

(1) If W_e is coinfinite, then $\mathcal{F}_{m,e}$ is not TxtFex$_*^*$-learnable by M, but the family is TxtFext$_*^*$-learnable.
(2) If W_e is cofinite, then $\mathcal{F}_{m,e}$ is uniformly TxtFext$_*^*$-learnable in both m and e.

Proof Fix a machine $M = M_m$ and a c.e. set W_e. We will construct a family

$$\mathcal{F}_{m,e} = \Big\{A, F_0^0, \ldots, F_{k_0}^0, F_0^1, \ldots, F_{k_1}^1, \ldots\Big\}$$

in stages. Each set in $\mathcal{F}_{m,e}$ will have two columns ($\{\langle 0, x \rangle : x \in \omega\}$ and $\{\langle 1, x \rangle : x \in \omega\}$) reserved for markers. Every set in $\mathcal{F}_{m,e}$ contains $\langle 0, \langle m, 0 \rangle \rangle$ and $\langle 0, \langle e, 1 \rangle \rangle$ and A contains the marker $\langle 0, \langle 0, 3 \rangle \rangle$ as well. Unless otherwise indicated, any action during the construction is performed on the complement of the reserved columns. We identify this complement with ω as it is a computable copy. At any stage of the construction,

there will be at most one $n \in \omega$ for which the sets $F_0^n, \ldots, F_{k_n}^n$ will be actively involved in the construction. When there is such a value n, we call $F_0^n, \ldots, F_{k_n}^n$ the *active sets* and n the *active index*. We also maintain an associated function r_n which stores information about that those sets.

Stage 0: Search for the least string σ on which M outputs a hypothesis. Set σ_0 equal to σ and enumerate content(σ_0) into A.

Stage s+1: Let $\sigma_0 \prec \ldots \prec \sigma_s$ be the sequence of strings passed to the current stage from stage s. If there is a currently active index n with active sets $F_0^n, \ldots, F_{k_n}^n$, then for $j \in W_{e,s+1}$, we enumerate $r_n(j), \ldots, r_n(j) + k_n$ into all of the active sets. We then consider four cases depending on the status of two parameters. First, the existence of an active index. Second, the availability, within computational bounds, of an extension $\alpha \succeq \sigma_s$ on which M outputs a hypothesis different from all the hypothesis it has output on prefixes of σ_s. We only consider the finite set of strings $S = \{\sigma_s{}^\frown\tau : (|\tau| < s + 1) \wedge (\text{content}(\tau) \subset s + 1)\}$ in our search for α.

Case 1: Suppose there is no active index, but there is an extension $\alpha \in S$, of σ_s such that $M(\alpha) \notin \{M(\tau) : \tau \preceq \sigma_s\}$. We pick the least such α. Set $\sigma_{s+1} = \alpha{}^\frown\beta$, where β is an increasing enumeration of $\{x : x \leq \max(\text{content}(\alpha))\}$, and enumerate content($\sigma_{s+1}$) into A.

Case 2: Next, consider the case where there is neither an active index nor an $\alpha \in S$ such that $M(\alpha) \notin \{M(\tau) : \tau \preceq \sigma_s\}$. Let $n \in \omega$ be least such that n has never been the active index before, set $k_n = |\{M(\tau) : \tau \preceq \sigma_s\}|$ and set $F_0^n, \ldots, F_{k_n}^n$ to be the active sets. Set $\sigma_{s+1} = \sigma_s$ and enumerate content(σ_{s+1}) into each of $F_0^n, \ldots, F_{k_n}^n$. Pick the least $q \in \omega$ such that no number greater than $(k_n + 1)q$ has appeared in the construction so far. Enumerate $(k_n + 1)q + i$ into F_i^n for each $i \leq k_n$.

Case 3: Let n be the active index and $F_0^n, \ldots, F_{k_n}^n$ be the active sets. Also, suppose $\alpha \in S$ is least such that $M(\alpha) \notin \{M(\tau) : \tau \preceq \sigma_s\}$. Set $\sigma_{s+1} = \alpha{}^\frown\beta$, where β is an increasing enumeration of $\{x : x \leq \max(\text{content}(\alpha))\}$, and enumerate content($\sigma_{s+1}$) into A. Next, we "cancel" the active index and active sets. Specifically, we enumerate the marker element $\langle 1, \langle n, i \rangle \rangle$ into F_i^n for each $i \leq k_n$ and mark the sets as inactive.

Case 4: Finally, assume there is a currently active index n and active sets $F_0^n, \ldots, F_{k_n}^n$, but no $\alpha \in S$ such that $M(\alpha) \notin \{M(\tau) : \tau \preceq \sigma_s\}$. Pick the least $q \in \omega$ such that no number greater than $(k_n + 1)q$ has appeared in the construction so far. Enumerate $(k_n + 1)q + i$ into F_i^n for each $i \leq k_n$ and set $r_n(i + 1) = (k_n + 1)q$, where i is the greatest value for which $r_n(i)$ is defined.

To verify that the above construction produces a family with the desired properties, we must verify three statements:

(1) $\mathcal{F}_{m,e}$ is TxtFext$_*^*$-learnable for all M and e.
(2) If W_e is coinfinite, then M does not TxtFex$_*^*$-learn $\mathcal{F}_{m,e}$.
(3) If W_e is cofinite, then there is a machine, computable from m and e, that TxtFext$_*^*$-learns $\mathcal{F}_{m,e}$.

If a natural number remains the active index cofinitely, then $\mathcal{F}_{m,e}$ is a finite family. If no such number exists, then every finite set has a unique marker by which it can be

identified and the only infinite set is A. In either case, the family is learnable and we conclude that the first statement is true.

To prove the second statement, we must again consider two cases. Suppose W_e is coinfinite. If $n \in \omega$ remains the active index cofinitely, then M outputs the same hypothesis on all extensions of a finite partial enumeration whose content is contained in each of $F_0^n, \ldots, F_{k_n}^n$. Thus, there are $k_n + 1$ enumerations, one for each of $F_0^n, \ldots, F_{k_n}^n$, on which M converges to the same hypothesis. The pairwise symmetric difference of any two active sets, however, is infinite. If no pair remains active infinitely, then there must be an infinite number of stages during the construction at which $\sigma_0 \prec \sigma_1 \prec \ldots$ are found such that $M(\sigma_{s+1}) \notin M(\alpha) \notin \{M(\tau) : \tau \preceq \sigma_s\}$ and A is enumerated by $f(n) = \sigma_n(n)$. In either case, M fails to TxtFex$_*^*$-learn $\mathcal{F}_{m,e}$.

Finally, we must exhibit a machine that can TxtFext$_*^*$-learn all possible families $\mathcal{F}_{m,e}$ where $e \in \text{COF}$. In particular, a machine that can TxtFex$_*^*$-learn the following possibly non-*u.c.e.* family as well as every subfamily:

$$\mathcal{G} = \bigcup_{e \in \text{COF}, m \in \omega} \mathcal{F}_{m,e}.$$

Since W_e is cofinite for all the families under consideration, observe that $\mathcal{F}_{m,e}$ consists of a (possibly infinite) number of finite sets and a finite number (A or $F_0^n, \ldots, F_{k_n}^n$) that are cofinite in the complement of the marker columns. Fix codes a_0, a_1 and $a_{m,e}$ such that $W_{a_0} = \emptyset$, $W_{a_1} = \omega \setminus \{\langle x, y \rangle : (x = 0 \lor x = 1) \land y \in \omega\}$ and $W_{a_{m,e}}$ is the set $A \in \mathcal{F}_{m,e}$. For notational ease, let $C(k) = \{\langle k, x \rangle : x \in \omega\}$. Define $N_{m,e}$ by

$$N_{m,e}(\sigma) = \begin{cases} a_0 & \text{if content}(\sigma) \cap C(1) \neq \emptyset, \\ a_{m,e} & \text{if } \langle 0, \langle 0, 3 \rangle \rangle \in \text{content}(\sigma), \\ a_1 & \text{otherwise.} \end{cases}$$

Further, define a machine N by

$$N(\sigma) = \begin{cases} N_{m,e} & \text{if } \langle 0, \langle m, 0 \rangle \rangle, \langle 0, \langle e, 1 \rangle \rangle \in \text{content}(\sigma), \\ 0 & \text{otherwise.} \end{cases}$$

To prove that N learns $\mathcal{G}$, select an arbitrary $D \in \mathcal{G}$. Let e and m be the codes such that $D \in \mathcal{F}_{m,e}$ for cofinite W_e.

Case 1: Suppose that, during the construction of $\mathcal{F}_{m,e}$, no natural number is the active index cofinitely. In this case, every member of $\mathcal{F}_{m,e}$ is marked, either with a marker in $C(1)$ or with $\langle 0, \langle 0, 3 \rangle \rangle$. Thus, N succeeds in TxtFex$_*^*$-learning $\mathcal{F}_{m,e}$.

Case 2: Suppose, on the other hand, that n remains active cofinitely during the construction. Let $F_0^n, \ldots, F_{k_n}^n$ be the sets that are active cofinitely. If $D = A$, then $\langle 0, \langle 0, 3 \rangle \rangle \in D$ and cofinitely often $N_{m,e}$ hypothesizes $a_{m,e}$. Every other finite set contains a unique marker and is hence TxtEx*-learnable by $N_{m,e}$. Finally, if D is one of the active sets, then $D =^* \omega \setminus (C(0) \cup C(1))$. No initial segment of any enumera-

tion of D contains either a marker in $C(1)$ or the marker $\langle 0, \langle 0, 3 \rangle \rangle$. Thus, $N_{m,e}$ again succeeds in TxtFex$_*^*$-learning the set. □

Theorem 4.2 *FEXL$_*^*$ and FEXTL$_*^*$ are both Σ_5^0-hard*

Proof We wish to reduce an arbitrary Σ_5^0 predicate $P(e)$ to FEXL$_*^*$ and FEXTL$_*^*$. For an arbitrary Σ_4^0 predicate $Q(e)$, there is a Σ_2^0 predicate, $R(e, x, y)$, such that the following representation can be made:

$$
\begin{aligned}
Q(e) &\leftrightarrow (\exists a)(\forall b)(R(e, a, b)) \\
&\leftrightarrow (\exists \langle a, s \rangle)[((\forall b)(R(e, a, b))) \wedge ((\forall a' < a)(\exists s' \leq s)(\neg R(e, a', s'))) \\
&\quad \wedge ((\exists a' < a)(\forall s' < s)(R(e, a', s')))] \\
&\leftrightarrow (\exists! \langle a, s \rangle)[((\forall b)(R(e, a, b))) \wedge ((\forall a' < a)(\exists s' \leq s)(\neg R(e, a', s'))) \\
&\quad \wedge ((\exists a' < a)(\forall s' < s)(R(e, a', s')))].
\end{aligned}
$$

Since the predicate

$$
\begin{aligned}
&((\forall b)(R(e, a, b))) \wedge ((\forall a' < a)(\exists s' \leq s)(\neg R(e, a', s'))) \\
&\wedge ((\exists a' < a)(\forall s' < s)(R(e, a', s')))
\end{aligned}
$$

is Π_3^0, for a suitable computable function g,

$$
Q(e) \to (\exists! x)(g(e, x) \in \text{COINF})
$$

and

$$
\neg Q(e) \to (\forall x)(g(e, x) \in \text{COF}).
$$

Applying the above to $P(e)$, the arbitrary Σ_5^0 predicate under consideration, we may define a computable function f such that

$$
\begin{aligned}
P(e) \to (\exists x)[&(\forall x' > x)(\forall y)(f(e, x', y) \in \text{COF}) \\
&\wedge (\forall x' \leq x)(\exists^{\leq 1} y)(f(e, x', y) \in \text{COINF})]
\end{aligned}
$$

and

$$
\neg P(e) \to (\forall x)[(\exists! y)(f(e, x, y) \in \text{COINF})].
$$

We will now define a family $\mathcal{G}_e$ from e such that $\mathcal{G}_e$ will be learnable if and only if $P(e)$. Define

$$
\mathcal{G}_e = \bigcup_{x, y \in \omega} \mathcal{F}_{x, f(e, x, y)}.
$$

Case 1: Suppose $\neg P(e)$. Then for every x, there is a y for which $f(e, x, y) \in$ COINF. From this we conclude that for each computable learner, M coded by m, there is a y such that $f(e, m, y) \in$ COINF. $\mathcal{G}_e$ contains a subfamily, $\mathcal{F}_{m, f(e,m,y)}$, that M cannot TxtFex$^*_*$-learn. Thus, $\mathcal{G}_e$ is not TxtBC-learnable.

Case 2: Suppose $P(e)$ and let x_0 be such that $(\forall x \geq x_0)(\forall y)(f(e, x, y) \in$ COF$)$. Let $a_0, a_1, \ldots, a_k$ enumerate the numbers less than x_0 such that, for unique corresponding $b_0, b_1, \ldots, b_k$, we have $f(e, a_i, b_i) \in$ COINF and let K_i be a computable machine that learns $\mathcal{F}_{a_i, f(e,a,b_i)}$. The existence of such a machine is guaranteed by Lemma 4.1. Using the machine N from the proof of Lemma 4.1, define a computable machine M on input string σ by

$$M(\sigma) = \begin{cases} K_i(\sigma) & \text{if } \langle 1, \langle 0, a_i \rangle \rangle, \langle 1, \langle 1, b_i \rangle \rangle \in \text{content}(\sigma) \text{ for } i \leq k, \\ N(\sigma) & \text{otherwise.} \end{cases}$$

If an enumeration of a set in the subfamily $\mathcal{F}_{a_i, f(e,a_i,b_i)}$ is fed to M, then eventually a tag in the 1st-column will appear identifying it as such. Cofinitely often, the appropriate K_i will be used to learn the enumeration. If the enumeration is for a set from $\mathcal{F}_{x, f(e,x,y)}$ with either $x \neq a_i$ or $y \neq b_i$ for any $i \leq k$, then N will be used. From Lemma 4.1, it is known that N is capable of TxtFext$^*_*$-learning $\mathcal{F}_{x, f(x,y)}$ for any x provided that $y \in$ COF.

We conclude that both FEXL$^*_*$ and FEXTL$^*_*$ are Σ^0_5-hard. □

Note that because the family described above is in fact TxtEx*-learnable in the Σ^0_5 case, meaning that this result has a result from [3] as a corollary.

5 Conclusion

With these results, the arithmetic complexity of the core models of vacillatory learning have been established.

The effort to classify the arithmetic complexity of learning criteria was initiated by the first author as an attempt to distinguish TxtFex$^*_*$ and TxtFext$^*_*$, a problem the first author answered in [2]. As is clear from these results, although the two learning models are distinct, their complexities do not distinguish them.

References

1. Baliga, G.R., Case, J., Jain, S.: The synthesis of language learners. Inf. Comput. **152**(1), 16–43 (1999)
2. Beros, A.A.: Anomalous vacillatory learning. J. Symb. Logic **78**(04), 1183–1188 (2013)
3. Beros, A.A.: Learning theory in the arithmetic hierarchy. J. Symb. Logic **79**(03), 908–927 (2014)
4. Blum, L., Blum, M.: Toward a mathematical theory of inductive inference. Inf. Control **28**(2), 125–155 (1975)
5. Carlucci, L., Case, J., Jain, S., Stephan, F.: Non U-shaped vacillatory and team learning. In: International Conference on Algorithmic Learning Theory, pp. 241–255. Springer, Berlin (2005)

6. Case, J.: The power of vacillation in language learning. SIAM J. Comput. **28**(6), 1941–1969 (1999)
7. Case, J., Jain, S., Sharma, A.: Anomalous learning helps succinctness. Theor. Comput. Sci. **164**(1–2), 13–28 (1996)
8. Gold, E.M.: Language identification in the limit. Inf. Control **10**(5), 447–474 (1967)
9. Jain, S., Osherson, D., Royer, J.S., Sharma, A.: Systems That Learn: An Introduction to learning Theory. MIT Press, Cambridge (1999)
10. Kapur, S., Bilardi, G.: On uniform learnability of language families. Inf. Process. Lett. **44**(1), 35–38 (1992)
11. Kummer, M., Stephan, F.: On the structure of degrees of inferability. J. Comput. Syst. Sci. **52**(2), 214–238 (1996)
12. Soare, R.I.: Recursively Enumerable Sets and Degrees: A Study of Computable Functions and Computably Generated Sets. Springer, Berlin (1999)