Non U-Shaped Vacillatory and Team Learning

Lorenzo Carlucci[1,⋆], John Case[2,⋆⋆], Sanjay Jain[3,⋆⋆⋆], and Frank Stephan[4,†]

[1] Department of Computer and Information Sciences, University of Delaware, Newark, DE 19716-2586,USA and Dipartimento di Matematica, Università di Siena, Pian dei Mantellini 44, Siena, Italy
carlucci5@unisi.it

[2] Department of Computer and Information Sciences, University of Delaware, Newark, DE 19716-2586,USA
case@cis.udel.edu

[3] School of Computing, 3 Science Drive 2, National University of Singapore, Singapore 117543
sanjay@comp.nus.edu.sg

[4] School of Computing and Department of Mathematics, National University of Singapore, 3 Science Drive 2, Singapore 117543
fstephan@comp.nus.edu.sg

Abstract. *U-shaped learning behaviour* in cognitive development involves learning, unlearning and relearning. It occurs, for example, in learning irregular verbs. The prior cognitive science literature is occupied with how humans do it, for example, general rules versus tables of exceptions. This paper is mostly concerned with whether U-shaped learning behaviour may be *necessary* in the abstract mathematical setting of inductive inference, that is, in the computational learning theory following the framework of Gold. All notions considered are learning from text, that is, from positive data. Previous work showed that U-shaped learning behaviour is necessary for behaviourally correct learning but not for syntactically convergent, learning in the limit (= explanatory learning). The present paper establishes the necessity for the whole hierarchy of classes of vacillatory learning where a behaviourally correct learner has to satisfy the additional constraint that it vacillates in the limit between at most k grammars, where $k \geq 1$. Non U-shaped vacillatory learning is shown to be restrictive: Every non U-shaped vacillatorily learnable class is already learnable in the limit. Furthermore, if vacillatory learning with the parameter $k = 2$ is possible then non U-shaped behaviourally correct learning is also possible. But for $k = 3$, surprisingly, there is a class witnessing that this implication fails.

1 Introduction and Motivation

U-shaped learning is a learning behaviour in which the learner first learns the correct behaviour, then abandons the correct behaviour and finally returns to

⋆ Supported in part by NSF grant number NSF CCR-0208616.
⋆⋆ Supported in part by NSF grant number NSF CCR-0208616.
⋆⋆⋆ Supported in part by NUS grant number R252–000–127–112.
† Supported in part by NUS grant number R252–000–212–112.

S. Jain, H.U. Simon, and E. Tomita (Eds.): ALT 2005, LNAI 3734, pp. 241–255, 2005.

the correct behaviour once again. This pattern of learning behaviour has been observed by cognitive and developmental psychologists in a variety of child development phenomena, such as language learning [6, 15, 22], understanding of temperature [22], understanding of weight conservation [5, 22], object permanence [5, 22] and face recognition [7].

The case of language acquisition is paradigmatic. In the case of the past tense of english verbs, it has been observed that children learn correct syntactic forms (call/called, go/went), then undergo a period of overregularization in which they attach regular verb endings such as 'ed' to the present tense forms even in the case of irregular verbs (break/breaked, speak/speaked) and finally reach a final phase in which they correctly handle both regular and irregular verbs. The irregular verb examples of U-shaped learning behaviour has figured so prominently in the so-called "Past Tense Debate" in cognitive science that competing models of human learning are often judged on their capacity for modeling the U-shaped learning phenomenon [15, 19, 23].

The prior literature is typically concerned with modeling *how* humans achieve U-shaped behaviour. Recently, Baliga, Case, Merkle, Stephan and Wiehagen [1] looked at abstract mathematical models which give some indication *why* humans exhibit this seemingly inefficient behaviour. Is it a mere harmless evolutionary accident or is it *necessary* for full human learning power? Specifically, are there some learning tasks for which U-shaped behaviour is logically necessary? In the present paper we continue this line of work.

In order to explain our results, we have to be a bit more formal. Although we refer to Section 2 below for an explanation of the mathematical terms used, we summarize for the reader's convenience the basics of inductive inference, that is, Gold's formal model of language learning from positive data [13].

The learning task is given by a subclass $\mathcal{C}$ of the class of all recursively enumerable (r.e.) subsets of the natural numbers which are indexed by natural numbers in an acceptable way [16, Section II.5]. The learner is then required to learn all the languages in the class $\mathcal{C}$. Here, a learner $\mathbf{M}$ learns a language L if it produces, in parallel to reading a text for L (that is, an infinite sequence of all elements of L in arbitrary order), a sequence $e_0, e_1, \ldots$ of hypotheses such that almost all of these hypotheses are the same index/grammar of the set L. This criterion is called **TxtEx** and stands for "explanatory learning from text" (see [13]).

The above criterion has been relaxed to **TxtBc** [9, 17] where it is only required that almost all e_n are grammars for L, but each e_n can be different from all previous ones. This criterion **TxtBc** is more general than **TxtEx** since languages have infinitely many grammars and the equality problem of the grammars is undecidable.

$\mathbf{TxtFex}_b$ [8] is the intermediate criterion where the learner succeeds iff there is an n such that $\{e_n, e_{n+1}, \ldots\}$ is actually a finite set of up to b correct grammars; the learner is then said to *vacillate* between these grammars. The criteria $\mathbf{TxtFex}_1, \mathbf{TxtFex}_2, \ldots, \mathbf{TxtFex}_*$ form a proper hierarchy between **TxtEx** and **TxtBc**.

Within this paper we continue the investigation of these standard criteria with the requirement that the learner is *non U-shaped* (which would require that e_{n+1} generates the language to be learnt whenever e_n does).

Baliga, Case, Merkle, Stephan and Wiehagen [1] initiated the Gold style learning theoretic study of U-shaped learning behaviour and showed that it is circumventable for **TxtEx**-learning, see Theorem 5. In contrast to this, Fulk, Jain and Osherson's proof of [12, Theorem 4] shows that U-shaped learning behaviour is necessary for the full learning power of **TxtBc**-learning. We show in Theorem 7 below that U-shaped learning behaviour is also necessary for full learning power for the whole hierarchy of the learning criteria $\mathbf{TxtFex}_b$ *strictly* between **TxtEx** and **TxtBc**. While Case [8] proved that the $\mathbf{TxtFex}_b$ criteria form a hierarchy of more and more powerful learning criteria, Theorem 7 of the present paper shows that *non U-shaped* $\mathbf{TxtFex}_b$-learners are not more powerful than **TxtEx**-learners. In other words, there are classes of languages that can be $\mathbf{TxtFex}_n$-identified, for $n > 1$, but these learners *must* be U-shaped on some texts.

What if we consider the more liberal criterion **TxtBc**? Our Theorem 18 strengthens the collapse result of Theorem 7 considerably by showing that there are classes in $\mathbf{TxtFex}_3$ that cannot be **TxtBc**-learned by a non U-shaped learner. This means that U-shaped learning behaviour cannot be dispensed with for learning such classes, even if we only require behavioural convergence and permit convergence to possibly infinitely many syntactically different correct hypotheses. By contrast, one of our main results, Theorem 17, shows that every class of languages that can be $\mathbf{TxtFex}_2$-identified can be **TxtBc**-identified by a non U-shaped learner. Hence, for only this early stage of the hierarchy, the cases in which $\mathbf{TxtFex}_2$-identification necessitates U-shaped learning behaviour can be circumvented by shifting to **TxtBc**-identification.

A further interesting aspect is that this paper gives a close relation between vacillatory learning and team learning. Theorem 2 gives the basic connection: A class is $\mathbf{TxtFex}_b$-learnable iff there is a team of b learners where all team-members converge on every text of a language to be learned and at least one of the team-members has to be correct. Furthermore, in Sections 3 and 5 some general inclusions for non U-shaped team learning are established.

We note that the relevance of Gold style learning to cognitive science has been supported in the cognitive science literature, for *example*, in [14, 18]. The publications [8, 24] critically discuss the relevance of the well studied criterion **TxtBc** to human learning; in order to avoid a mere interpolation of the data, one might want that a learner does not generate larger and larger hypotheses but tries to concentrate on finding a few correct ones. When **TxtEx**-learning is impossible, this can only be done by vacillating between some few hypotheses. Case [8] formalized this approach by introducing the criteria $\mathbf{TxtFex}_b$ for small, feasibly sized $b > 1$. Case also argues that these criteria *may* better fit the human case than the **TxtEx** criterion. Certainly, then, the new results of the present paper, regarding, for example, the $\mathbf{TxtFex}_3$ criterion are of interest for cognitive science, and *may* inform regarding the human case.

2 Preliminaries

$\mathbb{N}$ denotes the set of natural numbers, $\{0,1,2,\ldots\}$. card(D) denotes the cardinality of a set D. card$(D) \leq *$ means that card(D) is finite. The symbol $*$ is used to denote the "finite with no preassigned bound". The symbols $\subseteq, \subset, \supseteq, \supset$ respectively denote the subset, poper subset, superset and proper superset relation between sets. The quantifiers $\forall^\infty$ and $\exists^\infty$ mean "for all but finitely many" and "there exists infinitely many", respectively.

A pair $\langle \cdot, \cdot \rangle$ stands for an arbitrary, computable one-to-one encoding of all pairs of natural numbers onto $\mathbb{N}$ [20]. Similarly we can define $\langle \cdot, \ldots, \cdot \rangle$ for encoding n-tuples of natural numbers, for $n > 1$, onto $\mathbb{N}$.

φ denotes a fixed *acceptable* programming system for the partial-recursive functions [20]. φ_e denotes the partial-recursive function computed by the program with code number e in the φ-system. We will unambiguously refer to programs using their code number in the φ-system. W_e denotes domain of φ_e. $W_{e,s}$ denotes W_e enumerated within s computation stages [4]. For our purposes, we need $W_{e,s}$ to satisfy the following additional constraints, which can be easily ensured using standard techniques: (a) $W_{e,s} \subseteq \{0, \ldots, s-1\}$, (b) $\{(x, s) : x \in W_{e,s}\}$ is primitive recursive for all e and (c) for every primitive-recursive enumeration A_s of some set A with $A_0 = \emptyset \wedge (\forall s)\,[A_s \subseteq A_{s+1} \subseteq \{0, \ldots, s\}]$ there is an index e with $(\forall s)\,[W_{e,s} = A_s]$; furthermore, e can be computed from an index of the enumeration for A_s. Any unexplained recursion-theoretic notions are from [16, 20].

We now introduce the basic definitions of Gold-style computational learning theory.

A *sequence* σ is a mapping from an initial segment of $\mathbb{N}$ into $\mathbb{N} \cup \{\#\}$. An *infinite sequence* is a mapping from $\mathbb{N}$ into $\mathbb{N} \cup \{\#\}$. The content of a finite or infinite sequence σ is the set of natural numbers occurring in σ and is denoted by content(σ). The length of a sequence σ is the number of elements in the domain of σ and is denoted by $|\sigma|$. For a subset L of $\mathbb{N}$, seg(L) denotes the set of sequences σ with content$(\sigma) \subseteq L$. An infinite sequence T is a *text for* L iff $L = \text{content}(T)$.

Intuitively, a *text* for a language L is an infinite stream or sequential presentation of *all* the elements of the language L in any order with the #'s representing pauses in the presentation of the data. For example, the only text for the empty language is an infinite sequence of #'s. Furthermore, $T[n]$ denotes the first n elements of a text $T : \mathbb{N} \to \mathbb{N} \cup \{\#\}$.

A learner will map sequences from $(\mathbb{N} \cup \{\#\})^*$ to hypotheses. These are represented by natural numbers and interpreted as codes for programs in the φ-system. $\mathbf{M}$, with possible superscripts and subscripts, is intended to range over language learning machines.

Definition 1. [1, 2, 8, 9, 10, 13, 17] A *language learning machine* $\mathbf{M}$ is a computable mapping from seg$(\mathbb{N})$ into $\mathbb{N}$. $\mathbf{M}$ **TxtBc**-learns a class $\mathcal{L}$ of r.e. languages iff for every $L \in \mathcal{L}$ and every text T for L, almost all hypotheses $\mathbf{M}(T[n])$ are indices for the language L to be learned.

A **TxtBc**-learner $\mathbf{M}$ for $\mathcal{L}$ is a $\mathbf{TxtFex}_*$-learner iff for every $L \in \mathcal{L}$ and every text T for L the set $\{M(T[n]) : n \in \mathbb{N}\}$ is finite.

A $\mathbf{TxtFex}_*$-learner $\mathbf{M}$ for $\mathcal{L}$ is a $\mathbf{TxtFex}_b$-learner for a $b \in \{1, 2, \ldots\}$ iff there are for every $L \in \mathcal{L}$ and every text T for L at most b indices which $\mathbf{M}$ outputs infinitely often, that is, $|\{e : (\exists^\infty n)\,[e = \mathbf{M}(T[n])]\}| \leq b$.

A **TxtBc**-learner $\mathbf{M}$ for $\mathcal{L}$ is a **TxtEx**-learner iff for every $L \in \mathcal{L}$ and every text T for L almost all hypotheses $\mathbf{M}(T[n])$ are the same grammar for L.

A **TxtBc**-learner $\mathbf{M}$ for $\mathcal{L}$ is non U-shaped iff for every $L \in \mathcal{L}$ and every text T for L there are no three numbers k, m, n such that $k < m < n$ and $W_{\mathbf{M}(T[k])} = L$, $W_{\mathbf{M}(T[m])} \neq L$ and $W_{\mathbf{M}(T[n])} = L$. Furthermore, **NUShTxtBc**-learners, $\mathbf{NUShTxtFex}_b$-learners and **NUShTxtEx**-learners (for a class $\mathcal{L}$) are those learners which are non U-shaped and at the same time a **TxtBc**-learner, $\mathbf{TxtFex}_b$-learner and **TxtEx**-learner, respectively (for $\mathcal{L}$).

The criteria **TxtBc**, $\mathbf{TxtFex}_b$, **TxtEx**, **NUShTxtBc**, $\mathbf{NUShTxtFex}_b$, **NUShTxtEx** are the sets consisting of all those classes which are learnable by a learner satisfying the respective above defined requirements.

The historically most important learning criterion is **TxtEx** where the learner has to converge syntactically to a single index of the language to be learned [13]. **TxtEx** stands for "explanatory identification from text". Intuitively, the notion **TxtBc** captures what could be called learning in the most general sense, where "Bc" stands for "behaviourally correct" identification. In this, the learner outputs correct grammars almost always. A class $\mathcal{L}$ of r.e. languages is $\mathbf{TxtFex}_b$ identified by a machine $\mathbf{M}$ iff when $\mathbf{M}$ is given as input *any* listing T of *any* $L \in \mathcal{L}$, it outputs a sequence of grammars such that, past some point in this sequence, no more than b *syntactically* different grammars occur and each of them is a grammar for L. **TxtFex** stands for 'finite explanatory identification from text'. $\mathbf{TxtFex}_1$ is equivalent to **TxtEx**. Osherson and Weinstein [17] first studied the case with $b = *$, later Case [8] studied the whole hierarchy with $b \in \mathbb{N}^+$.

We say σ is a **TxtEx**-*stabilizing-sequence* [11] for a learner $\mathbf{M}$ on a set L iff $\sigma \in \mathrm{seg}(L)$ and $\mathbf{M}(\sigma\tau) = \mathbf{M}(\sigma)$ for all $\tau \in \mathrm{seg}(L)$. Furthermore a **TxtEx**-stabilizing-sequence σ is called a **TxtEx**-*locking-sequence* [3] for $\mathbf{M}$ on L iff $W_{\mathbf{M}(\sigma)} = L$. Note that stabilizing and locking sequence definitions can be generalized to other learning criteria such as **TxtFex** and **TxtBc**; we often drop "**TxtEx**" (respectively, "**TxtBc**", "**TxtFex**") from "**TxtEx**-stabilizing-sequence" and "**TxtEx**-locking-sequence", when it is clear from context.

Smith [21] studied learning by teams of machines, and, we can show vacillatory learning can be characterized by teams as follows.

Theorem 2. *A class $\mathcal{L}$ has a $\mathbf{TxtFex}_b$-learner iff there is a team of b machines $\mathbf{N}_1, \ldots, \mathbf{N}_b$ such that for every $L \in \mathcal{L}$ and for every text T for L, each machine $\mathbf{N}_a$ converges to a single index e_a and at least one of these indices e_a is an index for L.*

Case [8] showed $\mathbf{TxtFex}_1 \subset \mathbf{TxtFex}_2 \subset \ldots \subset \mathbf{TxtFex}_*$, as stated in the following Theorem.

Theorem 3. [8] *For* $b \in \{1, 2, \ldots\}$, $\mathcal{H}_b = \{W_e : e \in W_e \wedge |W_e \cap \{0, \ldots, e\}| \leq b+1\}$ *is in* $\mathbf{TxtFex}_{b+1} - \mathbf{TxtFex}_b$ *and* $\mathcal{H}_* = \{W_e : W_e \neq \emptyset \wedge e \leq \min(W_e)\}$ *is in* $\mathbf{TxtFex}_* - \bigcup_{b\in\{1,2,\ldots\}} \mathbf{TxtFex}_b$.

Proposition 4. $\mathbf{NUShTxtBc} \not\subseteq \mathbf{TxtFex}_*$.

The next theorem is from [1] and states that being non U-shaped is not restrictive for **TxtEx**-learning. Its proof can also be obtained by letting $b = 1$ in Theorem 12 below. Its fundamental equality will be extended to all classes $\mathbf{NUShTxtFex}_b$ in Theorem 7.

Theorem 5. [1] $\mathbf{NUShTxtEx} = \mathbf{TxtEx}$.

Hence, for **TxtEx**, U-shaped behaviour is not necessary for full learning power. By contrast, an easy adaptation of the proof of Theorem 4 in [12] shows that, for **TxtBc**, U-shaped behaviour is necessary for full learning power.

Theorem 6. [1, 12] $\mathbf{NUShTxtBc} \subset \mathbf{TxtBc}$.

3 Non U-Shaped Vacillatory Learning

We show that the $\mathbf{TxtFex}_b$-hierarchy collapses if U-shaped behaviour is forbidden.

Theorem 7. $\mathbf{NUShTxtFex}_* \subseteq \mathbf{TxtFex}_1$.

Proof. Let $\mathcal{L} \in \mathbf{NUShTxtFex}_*$ and let $\mathbf{M}$ be a learner witnessing this fact. We define a new learner $\mathbf{N}$ witnessing that $\mathcal{L} \in \mathbf{TxtFex}_1$ as follows.

On a text T, $\mathbf{N}$ keeps a list of all of $\mathbf{M}$'s conjectures in order of appearance and without repetitions and outputs the most recent entry in the list.

If T is a text for a language $L \in \mathcal{L}$, then $\mathbf{M}$ outputs on T only finitely many different hypotheses and at least one of them, say e, infinitely often, and $W_e = L$. Furthermore, $\mathbf{N}$ converges to that hypothesis e' which goes into the list last. If $e = e'$ then $\mathbf{N}$ is correct on T. If $e \neq e'$ then $\mathbf{M}$ has output e' the first time after having already output e at least once. Since $\mathbf{M}$ is not U-shaped and e is correct, so is e'. Thus $\mathbf{N}$ is correct on T again. ∎

Theorems 5 and 7 give $\mathbf{NUShTxtFex}_1 = \mathbf{TxtFex}_1$ and $\mathbf{NUShTxtFex}_* \subseteq \mathbf{NUShTxtFex}_1$. Furthermore, the definition of $\mathbf{NUShTxtFex}_b$ immediately gives $\mathbf{NUShTxtFex}_1 \subseteq \mathbf{NUShTxtFex}_b \subseteq \mathbf{NUShTxtFex}_*$. Thus all these criteria coincide.

Corollary 8. $(\forall b \in \{1, 2, \ldots, *\})\,[\mathbf{NUShTxtFex}_b = \mathbf{TxtFex}_1]$.

The result $\mathbf{NUShTxtFex}_1 = \mathbf{TxtFex}_1$ stands in contrast to the fact that $\mathbf{TxtFex}_1 \subset \mathbf{TxtFex}_2 \subset \ldots \subset \mathbf{TxtFex}_*$. Thus we have that the following inclusions are proper.

Corollary 9. $(\forall b \in \{2, 3, \ldots, *\})\,[\mathbf{NUShTxtFex}_b \subset \mathbf{TxtFex}_b]$.

Corollaries 8 and 9 show that U-shaped learning behaviour is *necessary* for the full learning power of $\mathbf{TxtFex}_b$-identification for $b > 1$ in a strong sense: if U-shaped learning behaviour is forbidden, the hierarchy collapses to $\mathbf{TxtFex}_1$. Hence, $\mathbf{TxtFex}_*$-learnability of any class in $(\mathbf{TxtFex}_* - \mathbf{TxtFex}_1)$ *requires* U-shaped learning behaviour.

Corollary 10. *Let $b \in \{1, 2, \ldots, *\}$ and $\mathcal{H}_b$ be the class from Theorem 3. Then any machine witnessing $\mathcal{H}_b \in \mathbf{TxtFex}_{b+1}$ necessarily employs U-shaped learning behaviour on $\mathcal{H}_b$.*

A non U-shaped learner does not make a mind change from a correct to an incorrect hypothesis since it cannot learn the set otherwise. This property is enforced on all machines for the case of team learning.

Definition 11. A team learning a class $\mathcal{L}$ is *non U-shaped* iff no machine in the team on any text for any language $L \in \mathcal{L}$ ever makes a mind change from an index for L to an index for a different language. In particular, the class $\mathcal{L}$ is in $[a, b]\mathbf{NUShTxtEx}$ iff there are b machines such that on any text for any language in $\mathcal{L}$ at least a machines converge to an index for that language and no machine makes a mind change from a correct to an incorrect hypothesis. For any learning criterion I, $[a, b]I$ is the corresponding team variant of this criterion.

The next result shows that Theorem 2 can be extended such that every class in $\mathbf{TxtFex}_b$ is learnable by a non U-shaped team. So the restriction $\mathbf{NUShTxtFex}_b = \mathbf{TxtFex}_1$ is caused by the fact that the hypothesis of the learner have to be brought into an ordering and cannot be done in parallel as in the case of the team below. Actually Theorem 12 enables us to achieve more properties of the team than that it is just non U-shaped.

Theorem 12. *Let $b \in \mathbb{N}^+$ and $\mathcal{L} \in \mathbf{TxtFex}_b$. Then there is a team of b learners $\mathbf{M}_1, \ldots, \mathbf{M}_b$ such that for all $L \in \mathcal{L}$ and all texts T for L there is an $n \in \mathbb{N}$ such that,*

(1) *$T[n]$ is a stabilizing sequence of all members of the team on L, in particular, $\mathbf{M}_a(T[m]) = \mathbf{M}_a(T[n])$ for all $m \geq n$;*
(2) *there is an $a \in \{1, \ldots, b\}$ such that $\mathbf{M}_a(T[n])$ is an index for L;*
(3) *if $a \in \{1, \ldots, b\}$ and $\mathbf{M}_a(T[m])$ is an index for L then $m \geq n$.*

In particular, $\mathbf{M}_1, \ldots, \mathbf{M}_b$ $[1, b]\mathbf{NUShTxtEx}$-learns $\mathcal{L}$.

Proof. By Theorem 2 there is a team $\mathbf{N}_1, \ldots, \mathbf{N}_b$ of $\mathbf{TxtEx}$-learners for $\mathcal{L}$ such that for every $L \in \mathcal{L}$ and every text T for L, every machine converges on T to some hypothesis and at least one of these hypotheses is an index for L.

The basic idea of the proof is to search for a $\sigma \in \mathrm{seg}(L)$, which is a $\mathbf{TxtEx}$-stabilizing-sequence for each member of the team $\mathbf{N}_1, \ldots, \mathbf{N}_b$ on L. Additionally, we will also find a maximal set $D \subseteq \{1, \ldots, b\}$ such that σ is a stabilizing sequence for each $\mathbf{N}_i$, $i \in \{1, \ldots, b\}$ on $W_{\mathbf{N}_j(\sigma)}$, $j \in D$. Before such σ, D, is obtained, we will make sure that the output of $\mathbf{M}_a$ below is not a grammar for

L. Once such σ, D is obtained, we will have that the learners $\mathbf{M}_a$ do not change their hypothesis and one of them correctly outputs a grammar for L. We now proceed formally.

Let E to be an infinite recursive set such that $E \cup \tilde{E} \notin \mathcal{L}$ for all finite sets $\tilde{E}$. Such an E can be defined as follows. Let $\mathbf{M}$ be a $\mathbf{TxtFex}_b$-learner for $\mathcal{L}$. If $\mathbb{N} \notin \mathcal{L}$, then let $E = \mathbb{N}$. If $\mathbb{N} \in \mathcal{L}$, then there exists a $\mathbf{TxtFex}_b$-locking sequence σ' for $\mathbf{M}$ on $\mathbb{N}$. Now we can take E to be any infinite and coinfinite recursive set such that $\text{content}(\sigma') \subseteq E$.

Let $\sigma \sqsubseteq \tau$ denote that $\text{content}(\sigma) \subseteq \text{content}(\tau)$ and $|\sigma| \leq |\tau|$. Furthermore, let T_e be the canonical text for W_e, that is, T_e is the text generated by some standard enumeration of W_e.

As long as the content of the input is $\emptyset$ or no σ is found in the algorithm below, all machines $\mathbf{M}_1, \ldots, \mathbf{M}_b$ output the least index of $\emptyset$. The σ searched for on input $\tau = T[t]$ has to satisfy the following conditions:

(a) $\sigma \sqsubseteq T[t]$ and $\text{content}(\sigma) \neq \emptyset$;
(b) $\mathbf{N}_a(\sigma\eta) = \mathbf{N}_a(\sigma)$ for all $a \in \{1, \ldots, b\}$ and $\eta \sqsubseteq T[t]$.

Once having σ, this is only replaced by a σ' on a future input $T[t']$ iff σ' but not σ satisfies (a) and (b) with respect to $T[t']$ (if there are several choices to replace σ, the first one with respect to some fixed recursive enumeration of $\text{seg}(\mathbb{N})$ is taken). Having σ, define D as follows.

(c) $D = \{a \in \{1, ..., b\} : (\forall \eta \sqsubseteq T_{\mathbf{N}_a(\sigma)}[t])\, (\forall c \in \{1, ..., b\})\, [\mathbf{N}_c(\sigma\eta) = \mathbf{N}_c(\sigma)]\}$.

Having σ and D, $\mathbf{M}_a(\tau) = F(\sigma, D, a)$ where $W_{F(\sigma,D,a)}$ is the set of all x for which there is an s such that the conditions (d) and either (e) or (f) below hold.

(d) $a \in D$ and $\sigma \sqsubseteq \text{content}(T_{\mathbf{N}_a(\sigma)}[s])$;
(e) $x \in \text{content}(T_{\mathbf{N}_a(\sigma)}[s])$ and $\mathbf{N}_c(\sigma\eta) = \mathbf{N}_c(\sigma)$ for all $c \in \{1, \ldots, b\}$, $d \in D$ and $\eta \sqsubseteq T_{\mathbf{N}_d(\sigma)}[s]$;
(f) $x \in E$ and $\mathbf{N}_c(\sigma\eta) \neq \mathbf{N}_c(\sigma)$ for some $c \in \{1, \ldots, b\}$, $d \in D$ and $\eta \sqsubseteq T_{\mathbf{N}_d(\sigma)}[s]$.

It is easy to see that the sets $W_{F(\sigma,D,a)}$ are uniformly recursively enumerable and thus the specified function F can be taken to be recursive. Thus also the learners $\mathbf{M}_1, \ldots, \mathbf{M}_b$ are recursive. Verification of properties (1), (2) and (3) is omitted. ∎

The next result gives a further inclusion between vacillatory learning and team-learning.

Theorem 13. $\mathbf{TxtFex}_b \subseteq [2, b+1]\mathbf{NUShTxtEx}$ *for all* $b \in \{1, 2, 3, \ldots\}$.
$\mathbf{TxtFex}_b \subset [1, b]\mathbf{NUShTxtEx}$ *for all* $b \in \{2, 3, \ldots\}$.

As $\mathbf{TxtFex}_2$-learning is more general than $\mathbf{TxtEx}$-learning, one gets the following corollary.

Corollary 14. $[2, 3]\mathbf{NUShTxtEx} \not\subseteq \mathbf{NUShTxtEx}$.

4 Vacillatory Versus Behaviourally Correct Learning

As every $\mathbf{NUShTxtFex}_b$-learner can be turned into a $\mathbf{NUShTxtEx}$-learner identifying the same class, the restriction to *non* U-shaped learning without loss of learning power is only possible in the least class $\mathbf{TxtFex}_1$ of the $\mathbf{TxtFex}_b$ hierarchy. But, the next, quite surprising result shows that in the case of $\mathbf{TxtFex}_2$ one can avoid U-shaped learning behaviour if one gives up the constraint that the learner has to vacillate between *finitely* many indices. That is, $\mathbf{TxtFex}_2 \subseteq \mathbf{NUShTxtBc}$. In Theorem 16 it is shown that there is a uniform learner $\mathbf{U}$ which is given a set E of up to 2 indices and $\mathbf{NUShTxtBc}$-identifies every $\{W_e : e \in E\}$ such that every hypothesis is a subset of an W_e with $e \in E$. Then this result is combined with Theorem 12 to show the inclusion $\mathbf{TxtFex}_2 \subseteq \mathbf{NUShTxtBc}$. But before turning to Theorem 16, the following auxiliary proposition gives a method to enforce that the sets $W_{i'}, W_{j'}$ mentioned there are represented by corresponding approximations W_i, W_j which differ from one another at all *relevant* stages of their enumerations.

Proposition 15. *Given a set $F = \{i', j'\}$ one can compute a set $G(F) = \{i, j\}$ such that*

- *For all s, either $W_{i,s} \cup W_{j,s} = \emptyset$ or $W_{i,s} \neq W_{j,s}$;*
- *$W_i \subseteq W_{i'}$ and if $W_{i'}$ is infinite then $W_i = W_{i'}$;*
- *$W_j \subseteq W_{j'}$ and if $W_{j'}$ is infinite then $W_j = W_{j'}$;*
- *$\{W_{i'}, W_{j'}\} \subseteq \{W_i, W_j\}$.*

This also holds with $j' = i'$ for the case that $F = \{i'\}$.

Theorem 16. *There is a learner $\mathbf{U}$ such that for all r.e. sets L, H and every set F of indices for L, H with $|F| \leq 2$, $\mathbf{U}$ $\mathbf{NUShTxtBc}$-identifies $\{L, H\}$ using the additional information F. Furthermore, for every $\sigma \in seg(\mathbb{N})$,*

(1) *$W_{\mathbf{U}(F,\sigma)} \subseteq L$ or $W_{\mathbf{U}(F,\sigma)} \subseteq H$;*
(2) *if $L = H$ and L is infinite then $W_{\mathbf{U}(F,\sigma)} \in \{\emptyset, L\}$.*

Note that $L = H$ is explicitly permitted.

Given F, let $G(F)$ be as in Proposition 15. Figure 1 gives the algorithm witnessing this inclusion. We omit the proof that it works.

Theorem 17. *Every $\mathbf{TxtFex}_2$-learnable class is $\mathbf{NUShTxtBc}$-learnable.*

Proof. Given a $\mathbf{TxtFex}_2$-learnable class $\mathcal{L}$, there is by Theorem 2 a pair of two learners $\mathbf{N}_1, \mathbf{N}_2$ which converge on every language from $\mathcal{L}$ and $[1,2]\mathbf{TxtEx}$-identify $\mathcal{L}$. Obtain from these two learners the team $\mathbf{M}_1, \mathbf{M}_2$ as done in the proof of Theorem 12. Let F be as defined in the proof of Theorem 12. Let σ_τ, D_τ be the values of σ, D computed on input τ by the algorithm for $\mathbf{M}_1, \mathbf{M}_2$ in the proof of Theorem 12. Note that one can, for each input σ_τ, D_τ, check in the limit whether $W_{F(\sigma_\tau, D_\tau, a)}$ enumerates some elements using part (f) of the algorithm in the proof of Theorem 12. Let $\mathbf{U}$ be as in Theorem 16.

Uniform non U-shaped Behaviourally Correct Learner U
Parameter: F. Input: σ. Output: k, specified implicitly.
Algorithm to enumerate $W_k = \bigcup_r W_{k,r}$.

(Start) Let $u = |\sigma|$, $C = \text{content}(\sigma)$ and $s = 0$.
Let $W_{k,t} = \emptyset$ for all $t < |\sigma|$.
If $C = \emptyset$ or $W_{e,u} = \emptyset$ for all $e \in G(F)$, Then go to (Empty).
Let $\tau = \sigma[|\sigma| - 1]$.
Select i, j, x such that
(a) $\{i, j\} = G(F)$;
(b) $x = \min((W_{i,u} - W_{j,u}) \cup (W_{j,u} - W_{i,u}))$;
(c) $x \in W_{i,u} \Leftrightarrow x \in C$.
Go to (Branch).
(Branch) If $C \cup W_{i,s} \subseteq W_{\mathbf{U}(F,\tau),u}$ and $(W_{i,u} - W_{i,|\sigma|}) \cup (W_{j,u} - W_{j,|\sigma|}) \neq \emptyset$ Then go to (Copy) Else go to (Enum).
(Enum) Let t be the maximal element of $\{s, \ldots, u\}$ such that one of the following conditions holds:
(Min) $t = s$;
(Equal) $C \subseteq W_{i,t} \cup W_{j,t} \subset W_{i,u} \cap W_{j,u}$;
(Inf) $C \subseteq W_{i,t} \subset W_{i,u} \wedge (\forall y \leq x)\,[y \notin (W_{i,u} - W_{i,|\sigma|}) \cup (W_{j,u} - W_{j,|\sigma|})]$;
(Diff) $C \subseteq W_{i,t} \subset W_{i,u}$ and $W_{j,u} = W_{i,s}$;
(Sub) $C \subseteq W_{i,t} \subseteq W_{\mathbf{U}(F,\tau),u}$;
(Exact) $C = W_{i,t}$ and $t = |\sigma|$.
Let $W_{k,u} = W_{i,t}$, update $s = t$, $u = u + 1$ and go to (Branch).
(Copy) Let $W_{k,u} = W_{\mathbf{U}(F,\tau),u}$, update $u = u + 1$ and go to (Copy).
(Empty) Let $W_{k,u} = \emptyset$, update $u = u + 1$ and go to (Empty).

Fig. 1. Algorithm to enumerate $W_{\mathbf{U}(F,\sigma)}$ from Theorem 16

Now one builds the following new learner $\mathbf{U}'(\tau)$ which outputs the least index of $\emptyset$ if $\text{content}(\tau) = \emptyset$ or, in the definition of $\mathbf{M}_1, \mathbf{M}_2$ on input τ, the search for σ satisfying (a), (b) (as in proof of Theorem 12) fails. Otherwise $U'(\tau)$ is defined as follows.

$$W_{\mathbf{U}'(\tau)} = W_{\mathbf{U}(\{e_1,e_2\},\tau)} \text{ where, for } a = 1, 2,$$
$$W_{e_a} = \begin{cases} W_{F(\sigma_\tau,D_\tau,a)} & \text{if (f) is not used for } W_{F(\sigma_\tau,D_\tau,1)} \text{ or } W_{F(\sigma_\tau,D_\tau,2)}; \\ W_{F(\sigma_\tau,D_\tau,1)} \cup W_{F(\sigma_\tau,D_\tau,2)} & \text{if (f) is used for } W_{F(\sigma_\tau,D_\tau,1)} \text{ or } W_{F(\sigma_\tau,D_\tau,2)}. \end{cases}$$

Let $L \in \mathcal{L}$ and T be a text for L. Let n be the first number where $L \in \{W_{\mathbf{M}_1(T[n])}, W_{\mathbf{M}_2(T[n])}\}$. Then, for any $m \geq n$ and any $a \in \{1, 2\}$, $W_{\mathbf{M}_a(T[m])} = W_{\mathbf{M}_a(T[n])}$ and $W_{\mathbf{M}_a(T[m])}$ is not of the form $E \cup \tilde{E}$, where E is as introduced in the proof of Theorem 12 and $\tilde{E}$ is finite. Then $\mathbf{U}$ is fed with the same parameter set $\{e_1, e_2\}$ for all $m \geq n$ and one of the e_1, e_2 enumerates L. Thus $\mathbf{U}'$ **TxtBc**-learns L on T.

It remains to show that $\mathbf{U}'$ is non U-shaped on T. This is clearly true if L is the empty set. So assume $L \neq \emptyset$. Consider any m with $W_{\mathbf{U}'(T[m])} = L$. If case (f) of the algorithm for $F(\sigma_{T[m]}, D_{T[m]}, 1)$ or $F(\sigma_{T[m]}, D_{T[m]}, 2)$ applies, then W_{e_1} and W_{e_2} are the same infinite set $E \cup \tilde{E}$ for some finite set $\tilde{E}$. It follows by the additional property (2) of $\mathbf{U}$ in Theorem 16 that $\mathbf{U}'(T[m])$ either outputs an index for the empty set or for $E \cup \tilde{E}$; both sets are different from L, thus case (f) does not apply. Hence, $T[m]$ is a stabilizing sequence for both $\mathbf{M}_1, \mathbf{M}_2$ on those sets $W_{\mathbf{M}_1(T[m])}, W_{\mathbf{M}_2(T[m])}$ which are not empty. Since one of these is a superset of L by the additional property (1) of $\mathbf{U}$ in Theorem 16, it follows that $\mathbf{M}_1, \mathbf{M}_2$ do not change mind on T beyond $T[m]$. For $a = 1, 2$ the parameter e_a is defined as above for $\tau = T[m]$ and it holds that $W_{e_a} = W_{\mathbf{M}_a(T[m])}$. Thus $\mathbf{U}'(T[o])$ coincides with $\mathbf{U}(\{e_1, e_2\}, T[o])$ for all $o \geq m$ and $\mathbf{U}$ with the parameter set $\{e_1, e_2\}$ is non U-shaped on the text T for L. The same holds for $\mathbf{U}'$. Thus $\mathbf{U}'$ **NUShTxtBc**-learns $\mathcal{L}$. ∎

From Theorem 7 it is already known that, for all $b > 1$, U-shaped learning behaviour is necessary for $\mathbf{TxtFex}_b$ identification of any class in $\mathbf{TxtFex}_b - \mathbf{TxtFex}_1$. Theorem 18 strengthens this result by showing that, for some classes of languages in $\mathbf{TxtFex}_b$ for $b > 2$, the necessity of U-shaped behaviour cannot be circumvented by allowing infinitely many correct grammars in the limit, that is, by shifting to the more liberal criterion of **TxtBc**-identification. This is one of the rare cases in inductive inference where the containment in a class defined without numerical parameters holds for level 2 but not for level 3 and above of a hierarchy. The proof is a diagonalization proof reminiscent of the proof of Theorem 4 in [12].

Theorem 18. $\mathbf{TxtFex}_3 \not\subseteq \mathbf{NUShTxtBc}$.

Proof. Let $L_{i,j} = \{\langle i, j, k\rangle : k \in \mathbb{N}\}$, $I_{i,j} = W_i \cap L_{i,j}$ and $J_{i,j} = W_j \cap L_{i,j}$ for $i, j \in \mathbb{N}$. The class $\mathcal{L} = \{L_{i,j} : i, j \in \mathbb{N}\} \cup \{I_{i,j}, J_{i,j} : i, j \in \mathbb{N} \wedge I_{i,j} \subset J_{i,j} \wedge |I_{i,j}| < \infty\}$ witnesses the separation.

To see that $\mathcal{L}$ is in $\mathbf{TxtFex}_3$, consider the following machines $\mathbf{N}_I, \mathbf{N}_J, \mathbf{N}_L$ which initially output indices of the empty set. Each of them waits for the first tuple of the form $\langle i, j, k\rangle$ for some k to come up in the input. From then on, $\mathbf{N}_I$ outputs an index for $I_{i,j}$ forever, $\mathbf{N}_J$ an index for $J_{i,j}$ forever and $\mathbf{N}_L$ an index for $L_{i,j}$ forever. So, for every $i, j \in \mathbb{N}$, $\mathbf{N}_I$ learns the set $I_{i,j}$, $\mathbf{N}_J$ the set $I_{i,j}$ and $\mathbf{N}_L$ the set $L_{i,j}$. The class $\mathcal{L}$ is learnable by a team of three machines which converge on every text for every language in $\mathcal{L}$ to some index. It follows from Theorem 2 that $\mathcal{L}$ is in $\mathbf{TxtFex}_3$.

So it remains to show that $\mathcal{L}$ is not in **NUShTxtBc**, that is, to show that any given **TxtBc**-learner for $\mathcal{L}$ is U-shaped on some text for some language in $\mathcal{L}$. Given the learner $\mathbf{M}$, one defines the following function F by an approximation from below.

$$F_s(i,j) = \begin{cases} F_{s-1}(i,j) & \text{if } s > 0 \text{ and} \\ & W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F_{s-1}(i,j)\rangle),s} \subseteq L_{i,j}; \\ k & \text{otherwise where } k \text{ is the first number} \\ & \text{found with } k > F_t(i+j)+s, \text{ for all } t < s, \text{ and} \\ & \{\langle i,j,0\rangle, \langle i,j,1\rangle, ..., \langle i,j,k\rangle\} \subset W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,k\rangle)}. \end{cases}$$

Since $\langle i,j,0\rangle, \langle i,j,1\rangle, \ldots$ is a text for $L_{i,j}$ and $\mathbf{M}$ **TxtBc**-learns $L_{i,j}$, almost all hypotheses $\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,k\rangle)$ are indices for $L_{i,j}$. Thus the k is always found in the second part of the definition of F_s and F_s is well-defined. Furthermore, if $F_{s-1}(i,j)$ is sufficiently large, the condition

$$W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F_{s-1}(i,j)\rangle),s} \subseteq L_{i,j}$$

holds for all s and thus $F_s(i,j) = F_{s-1}(i,j)$. So the limit $F(i,j)$ of all $F_s(i,j)$ exists and is approximated from below. By considering the first s where $F(i,j) = F_s(i,j)$ and the fact that it is then no longer updated, one has

$$\{\langle i,j,0\rangle, \langle i,j,1\rangle, \ldots, \langle i,j,F(i,j)\rangle\} \subset W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F(i,j)\rangle)} \subseteq L_{i,j}.$$

Now there are r.e. sets W_a, W_b such that

$$W_a = \{\langle i,j,l\rangle : i,j \in \mathbb{N} \wedge l \in \{0,1,\ldots,F(i,j)\}\},$$
$$W_b = \{\langle i,j,l\rangle : i,j \in \mathbb{N} \wedge (\exists t > l)\,[\langle i,j,l\rangle \in W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F_t(i,j)\rangle),t}]\}.$$

Now fix the parameters i,j such that $i = a$ and $j = b$; the cases where $i \neq a$ or $j \neq b$ are not important in the considerations below.

Assume that $\langle i,j,l\rangle \in W_b$ using a parameter t with $F_t(i,j) \neq F(i,j)$. Let s be the first stage with $F_s(i,j) = F(i,j)$; note that $s > t$. Then by the definition of F_s, $F(i,j) = F_s(i,j) > s > t$ and $\{\langle i,j,0\rangle, \langle i,j,1\rangle, \ldots, \langle i,j,F_s(i,j)\rangle\} \subset W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F_s(i,j)\rangle)}$. So $\langle i,j,l\rangle$ is in $W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F(i,j)\rangle)}$ as well. Thus $\{\langle i,j,0\rangle, \langle i,j,1\rangle, \ldots, \langle i,j,F(i,j)\rangle\} = W_i \cap L_{i,j} = I_{i,j} \subset J_{i,j} = W_j \cap L_{i,j} = W_{\mathbf{M}(\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F(i,j)\rangle)}$ and $I_{i,j}$ is finite. Hence $I_{i,j}, J_{i,j} \in \mathcal{L}$.

Now consider a text T for $J_{i,j}$ formed as follows. Let σ be the sequence $\langle i,j,0\rangle\, \langle i,j,1\rangle \,\ldots\, \langle i,j,F(i,j)\rangle$. Note that $\mathbf{M}(\sigma)$ outputs an index for $J_{i,j}$. Let $\tau = \sigma\#^r$, for some r, be such that $\mathbf{M}(\tau)$ is an index for $I_{i,j}$. Note that there exists such τ since $\mathbf{M}$ **TxtBc**-identifies $I_{i,j}$. Let T be a text for $J_{i,j}$ starting with τ. Now $\mathbf{M}$ on T has to output an index $J_{i,j}$ beyond τ. Hence, $\mathbf{M}$ is U-shaped on text T, and thus $\mathbf{M}$ is not a **NUShTxtBc**-learner for $\mathcal{L}$. Since $\mathbf{M}$ was chosen arbitrarily, $\mathcal{L}$ is not **NUShTxtBc**-learnable. ∎

Since $\mathbf{TxtFex}_3 \subset \mathbf{TxtFex}_4 \subset \ldots \subset \mathbf{TxtFex}_*$, one immediately gets the following corollary.

Corollary 19. $(\forall b \in \{3,4,\ldots,*\})\,[\mathbf{TxtFex}_b \not\subseteq \mathbf{NUShTxtBc}]$.

A further corollary is that the counterpart of Theorem 16 does not hold for sets of three indices. Indeed, if such an algorithm would exist, then one could

NUShTxtBc-learn $\mathcal{L}$ from Theorem 18 by conjecturing $\emptyset$ until the first triple $\langle i, j, k \rangle$ comes up and then simulating the uniform learner with a set of three indices for the sets $I_{i,j}, J_{i,j}, L_{i,j}$ from then on without changing this parameter set anymore. But Theorem 18 clearly showed that such a learner does not exist.

Corollary 20. *No machine uniformly* **NUShTxtBc***-learns* $\{W_e : e \in F\}$ *with* F *as additional information where* F *is a set of* 3 *indices.*

5 Teams Revisited

Classes in $\mathbf{TxtFex}_2$ are in **TxtBc** and in $[1,2]$**NUShTxtEx**. The next proposition shows that one cannot weaken the condition of being in $\mathbf{TxtFex}_2$ to the combination of the two consequences in Theorem 17. Furthermore the condition that the team members converge on every text for a language in $\mathcal{L}$ is essential in Theorem 2.

Proposition 21. *The class* $\mathcal{L}$ *from Theorem 18 is* $[1,2]$**NUShTxtEx***-learnable and* $\mathbf{TxtFex}_3$*-learnable but it is not* **NUShTxtBc***-learnable.*

Remark 22. $\mathbf{TxtFex}_* \not\subseteq [1,b]\mathbf{TxtEx}$ for all $b \in \mathbb{N}^+$, as witnessed by $\mathcal{H}_*$. Note that by Proposition 21 it can be that a class in $\mathbf{TxtFex}_{b+1} - \mathbf{TxtFex}_b$ is already $[1,b]$**TxtEx**-learnable.

A further interesting question is whether one can at least obtain non U-shaped team learning for arbitrary team learnable classes. This is true for $[1,1]$**TxtEx** by Theorem 12 but it fails for $[1,2]$**TxtEx**-learning.

Theorem 23. *For all* $b \in \{2,3,\ldots\}$*,* $[1,b]\mathbf{NUShTxtEx} \subset [1,b]\mathbf{TxtEx}$*.*
For all a, b *with* $1 \leq a \leq b$*,* $[a,b]\mathbf{TxtEx} \subseteq [a,a+b]\mathbf{NUShTxtEx}$*.*

6 Conclusion

The following results were obtained.

- $\mathbf{TxtFex}_b \subset [1,b]\mathbf{NUShTxtEx}$ for all $b \in \{2,3,\ldots\}$.
- $\mathbf{TxtFex}_1 = \mathbf{TxtEx} = \mathbf{NUShTxtEx} = [1,1]\mathbf{NUShTxtEx}$, see also [1].
- $[1,b]\mathbf{NUShTxtEx} \subset [1,b]\mathbf{TxtEx}$, for all $b \in \{2,3,\ldots\}$.
- $\mathbf{NUShTxtFex}_b = \mathbf{NUShTxtEx}$ for all $b \in \{1,2,\ldots,*\}$.
- $\mathbf{TxtFex}_2 \subseteq \mathbf{NUShTxtBc}$.
- $\mathbf{TxtFex}_3 \not\subseteq \mathbf{NUShTxtBc}$.

These results and the facts known from previous work [1, 8] are summarized in Figure 2. Single-headed arrows in the diagram denote proper inclusions. Double-headed arrows denote equality. All transitive closures of the inclusions displayed are valid and no other inclusions hold between language learning criteria in the diagram.

$$
\begin{tikzcd}
\mathbf{TxtBc} & \mathbf{TxtFex}_* \arrow[l] & & \\
& \vdots \arrow[u] & \vdots & \vdots \\
& \mathbf{TxtFex}_5 \arrow[u] \arrow[r] & [1,5]\mathbf{NUShTxtEx} \arrow[r] & [1,5]\mathbf{TxtEx} \\
& \mathbf{TxtFex}_4 \arrow[u] \arrow[r] & [1,4]\mathbf{NUShTxtEx} \arrow[u] \arrow[r] & [1,4]\mathbf{TxtEx} \arrow[u] \arrow[ul] \\
& \mathbf{TxtFex}_3 \arrow[u] \arrow[r] & [1,3]\mathbf{NUShTxtEx} \arrow[u] \arrow[r] & [1,3]\mathbf{TxtEx} \arrow[u] \arrow[ul] \\
\mathbf{NUShTxtBc} \arrow[uuuuu] & \mathbf{TxtFex}_2 \arrow[u] \arrow[l] \arrow[r] & [1,2]\mathbf{NUShTxtEx} \arrow[u] \arrow[r] & [1,2]\mathbf{TxtEx} \arrow[u] \arrow[ul] \\
\mathbf{NUShTxtFex}_b \arrow[u] \arrow[r, leftrightarrow] & \mathbf{TxtFex}_1 \arrow[u] \arrow[r, leftrightarrow] & [1,1]\mathbf{NUShTxtEx} \arrow[u] \arrow[r, leftrightarrow] & [1,1]\mathbf{TxtEx} \arrow[u] \arrow[ul]
\end{tikzcd}
$$

Fig. 2. Summary of the results for $b \in \{1, 2, 3, 4, 5, *\}$

We note that our proof that $\mathbf{TxtFex}_3 \not\subseteq \mathbf{NUShTxtBc}$ intriguingly features learning finite tables versus general rules, but does *not*, as might be expected from some models of the human case of U-shaped learning, feature, among other things, learning an incorrect general rule followed by learning a general rule augmented by a correcting finite table. This difference may be significant or, more likely, nothing more than an artifact of our particular proof. Not explored herein, but very interesting to investigate in the future, are complexity-issues of U-shaped learning.

References

[1] Ganesh Baliga, John Case, Wolfgang Merkle, Frank Stephan and Rolf Wiehagen. *When unlearning helps.* http://www.cis.udel.edu/˜case/papers/decisive.ps, Manuscript, 2005. Preliminary version of the paper appeared at ICALP, Springer LNCS 1853:844–855, 2000.

[2] Janis Bārzdiņš. Two theorems on the limiting synthesis of functions. In *Theory of Algorithms and Programs, vol. 1*, pages 82–88. Latvian State University, 1974. In Russian.

[3] Lenore Blum and Manuel Blum. Towards a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

[4] Manuel Blum. A machine independent theory of the complexity of the recursive functions. *Journal of the Association for Computing Machinery* 14:322–336, 1967.

[5] T. G. R. Bower. Concepts of development. In *Proceedings of the 21st International Congress of Psychology*. Presses Universitaires de France, pages 79–97, 1978.
[6] Melissa Bowerman. Starting to talk worse: Clues to language acquisition from children's late speech errors. In S. Strauss and R. Stavy, editors, *U-Shaped Behavioral Growth*. Academic Press, New York, 1982.
[7] Susan Carey. Face perception: Anomalies of development. In S. Strauss and R. Stavy, editors, *U-Shaped Behavioral Growth*, Developmental Psychology Series. Academic Press, pages 169–190, 1982.
[8] John Case. The power of vacillation in language learning. *SIAM Journal on Computing*, 28(6):1941–1969, 1999.
[9] John Case and Chris Lynes. Machine inductive inference and language identification. In M. Nielsen and E. M. Schmidt, editors, *Proceedings of the 9th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science 140, pages 107–115. Springer-Verlag, 1982.
[10] John Case and Carl H. Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
[11] Mark Fulk. Prudence and other conditions on formal language learning. *Information and Computation*, 85:1–11, 1990.
[12] Mark Fulk, Sanjay Jain and Daniel Osherson. Open problems in "Systems That Learn". *Journal of Computer and System Sciences*, 49:589–604, 1994.
[13] E. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
[14] David Kirsh. PDP learnability and innate knowledge of language. In S. Davis, editor, *Connectionism: Theory and Practice*, pages 297–322. Oxford University Press, 1992.
[15] Gary Marcus, Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen and Fei Xu. *Overregularization in Language Acquisition*. Monographs of the Society for Research in Child Development, vol. 57, no. 4. University of Chicago Press, 1992. Includes commentary by Harold Clahsen.
[16] Piergiorgio Odifreddi. *Classical Recursion Theory*. North Holland, Amsterdam, 1989.
[17] Daniel Osherson and Scott Weinstein. Criteria of language learning. *Information and Control*, 52:123–138, 1982.
[18] Steven Pinker. Formal models of language learning. *Cognition*, 7:217–283, 1979.
[19] Kim Plunkett and Virginia Marchman. U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, 38(1):43–102, 1991.
[20] Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, New York, 1967. Reprinted, MIT Press, 1987.
[21] Carl H. Smith. The power of pluralism for automatic program synthesis. *Journal of the Association of Computing Machinery*, 29:1144–1165, 1982.
[22] Sidney Strauss and Ruth Stavy, editors. *U-Shaped Behavioral Growth*. Developmental Psychology Series. Academic Press, 1982.
[23] Niels A. Taatgen and John R. Anderson. Why do children learn to say broke? A model of learning the past tense without feedback. *Cognition*, 86(2):123–155, 2002.
[24] Kenneth Wexler. On extensional learnability. *Cognition*, 11:89–95, 1982.