Some Recent Results in U-Shaped Learning

Sanjay Jain*[1] and Frank Stephan**[2]

[1] Department of Computer Science, National University of Singapore, Singapore 117543, Republic of Singapore.
sanjay@comp.nus.edu.sg
[2] Department of Computer Science and Department of Mathematics, National University of Singapore, Singapore 117543, Republic of Singapore.
fstephan@comp.nus.edu.sg

Abstract. U-shaped learning deals with a learner first having the correct hypothesis, then changing it to an incorrect hypothesis and then relearning the correct hypothesis. This phenomenon has been observed by psychologists in various studies of children development. In this survey talk, we will discuss some recent results regarding U-shaped learning and related crtieria.

1 Language Learning

A language is a set of sentences using words over an alphabet. Sentences and words over an alphabet can be encoded into natural numbers. Thus, one may model a language as a subset of N, the set of natural numbers. Consider the following model of learning a language. A learner, over time, receives one by one elements of the language, in arbitrary order. As the learner is receiving the data, it conjectures a sequence of grammars, $g_0, g_1, \ldots$, potentially describing the input language. One may consider the learner to be successful, if this sequence of conjectures eventually stabilizes to a grammar g (i.e., beyond certain point all its conjectures are the grammar g), and this grammar g is a indeed a grammar for the input language. In our model, we take the learner to be computable. This criteria of success originated with Gold [18], and is refered to as **TxtEx** learning (**Txt** stands for text, and **Ex** stands for explanatory learning). Note here that the learner only gets data about what is in the language, and is not told about what is not in the language. Thus, such kind of learning is often call learning from positive data.

It is not so interesting to consider learning of just one language, as a learner which just outputs the grammar for the single language, will ofcourse be able to learn it. Thus, one usually considers learnability of a class $\mathcal{L}$ of languages, where the learner is required to learn all the langauges L in the class, from all possible texts for the language L (here a text for L is presentation of all and only the elements of L, in arbitrary order). This model of learning was first introduced by

* Supported in part by NUS grant number R252-000-127-112.
** Supported in part by NUS grant number R252-000-212-112.

Gold [18], and has then been explored by various researchers (see for example, [31, 24, 20, 14, 22, 19]).

Since Gold, various authors have considered extensions and restrictions of the above model. Some of the important extensions are as follows. We first consider behaviourally correct learning in which one requires that the learner semantically converge to the correct hypothesis rather than the syntactic convergence as required in explanatory learning. A learner is said to **TxtBc**-identify a language L, iff given as input any text for L, the learner outputs an infinite sequence of conjectures, all but finitely many of which are grammars for the language L. **Bc** here stands for behaviourally correct learning. Thus, in the scenerio described above, for all but finitely many n, g_n is a grammar for L. This model of learning was first considered for function learning by [4] and for language learning by [13, 23].

Another model of learning, called vacillatory learning, can be described as follows: not only are the conjectures of the learner almost always correct, but eventually the conjectures come only from a finite set S. The learner is said to $\mathbf{TxtFex}_n$-learn the language L if this set S is of size atmost n. The learner is said to $\mathbf{TxtFex}_*$-learn the language L, if we just require the set S to be finite. This model of learning was introduced by [12]. Intuitively, we (eventually) allow vacillation among at most n correct hypothesis of the language. It can be shown that $\mathbf{TxtEx} = \mathbf{TxtFex}_1 \subset \mathbf{TxtFex}_2 \ldots \subset \mathbf{TxtFex}_* \subset \mathbf{TxtBc}$.

We now provide the formal definition of above criteria of learning. We first formally define the notion of sequence of data presented to the learner.

Definition 1. (a) A *finite sequence* σ is a mapping from an initial segment of N into $N \cup \{\#\}$. An *infinite sequence* is a mapping from N into $N \cup \{\#\}$.

(b) The content of a finite or infinite sequence σ, denoted by content(σ), is the set of natural numbers occurring in σ.

(c) The length of a sequence σ, denoted by $|\sigma|$, is the number of elements in the domain of σ.

(d) An infinite sequence T is a *text for* L iff $L = \text{content}(T)$.

(e) $T[n]$ denotes the initial segment of T of length n.

(f) For $L \subseteq N$, SEG(L) denotes the set of all finite sequences σ such that content$(\sigma) \subseteq L$.

We now define the three criteria of learning presented above.

Let φ be an acceptable [26] programming system, and φ_i denote the function computed by the i-th program in this system. Let $W_i = \text{domain}(\varphi_i)$. Then W_i can be viewed as the recursively enumerable language accepted/generated by the i-th grammar in the φ-system.

Let $\mathcal{E}$ denote the class of all recursively enumerable languages.

Definition 2. [4, 12–14, 18, 23]

(a) A *language learning machine* $\mathbf{M}$ is a (possibly partial) comptuable mapping from SEG(N) into N.

(b) $\mathbf{M}$ **TxtEx**-identifies a language L, iff for all texts T for L, there exists a grammar e such that $W_e = L$ and for all but finitely many n, $\mathbf{M}(T[n]) = e$.

(c) **M** **TxtBc**-identifies a language L, iff for all texts T for L, for all but finitely many n, $W_{\mathbf{M}(T[n])} = L$.

(d) For $b \in N$, **M** $\mathbf{TxtFex}_b$-identifies a language L, iff for all texts T for L, there exists a set S of size at most b such that (i) for each $i \in S$, $W_i = L$ and (ii) for all but finitely many n, $\mathbf{M}(T[n]) \in S$. If we only require the above set S to be finite, then we say that **M** $\mathbf{TxtFex}_*$-identifies L.

(e) For $\mathbf{J} \in \{\mathbf{TxtEx}, \mathbf{TxtBc}, \mathbf{TxtFex}_b\}$, we say that **M** **J**-identifies a class $\mathcal{L}$ of languages if it **J**-identifies each $L \in \mathcal{L}$.

(f) For $\mathbf{J} \in \{\mathbf{TxtEx}, \mathbf{TxtBc}, \mathbf{TxtFex}_b\}$, we define the criteria $\mathbf{J} = \{\mathcal{L} \mid (\exists \mathbf{M})[\mathbf{M}\ \mathbf{J}\text{-identifies } \mathcal{L}]\}$.

It is known that $\mathbf{TxtEx} = \mathbf{TxtFex}_1 \subset \mathbf{TxtFex}_2 \subset \cdots \subset \mathbf{TxtFex}_* \subset \mathbf{TxtBc}$ and $\mathcal{E} \notin \mathbf{TxtBc}$.

2 U-shaped Behaviour

A U-shaped learning behaviour is one in which a learner first learns a correct grammar, then changes its mind to an incorrect grammar and then comes back to a correct grammar. It other words, it involves learning, unlearning and relearning. This learning behaviour has been observed by cognitive and developmental psychologists in various child development phenomena, such as language learning [6, 21, 28], understanding of temperature [28, 29], understanding of weight conservation [5, 28], object permanence [5, 28] and face recognition [7]. For example in language learning during the process of learning past tense of english verbs, children first 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. U-shaped learning behaviour has figured so prominently in the so-called "Past Tense Debate" in cognitive science that models of human learning are often judged on their capacity for modeling the U-shaped learning phenomenon [21, 25, 30].

In this paper we will illustrate some of the recent results which have been obtained regarding necessity of U-shaped learning (rather than just that it happens in humans due to some peculiarity in evolution). We will also discuss some of the related models of learning behaviour which are similar to U-shaped behaviour. Most of the results of this paper are from [1, 2, 8–10].

Before formally discussing U-shaped behaviour, let us first consider the related notion of decisive learning. A learner is said to be decisive if it never returns to an abandoned conjecture. Note that, using a padding function, one can always make newer conjectures syntactically different from previous conjectures. Thus what is more interesting is that we require the learner not to semantically return to any abandoned hypothesis. Formally,

Definition 3. [22] (a) A learner **M** is said to be *decisive* on a text T iff there does not exist m, n, t such that $m < n < t$, $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])}$, but $W_{\mathbf{M}(T[m])} \neq W_{\mathbf{M}(T[n])}$.

(b) A learner $\mathbf{M}$ is *decisive* on L, iff it is decisive on each text T for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner **DecJ**-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, $\mathbf{M}$ is decisive on L and $\mathbf{J}$-identifies L.

Osherson, Stob and Weinstein [22] asked the natural question whether decisiveness is restrictive. Fulk, Jain and Osherson [17] answered this question for behaviourally correct learning and Baliga, Case, Merkle and Stephan [1] for explanatory learning. Actually, both results can be subsumed in one theorem which then also covers the case of vacillatory learning as it is between explanatory and behaviourally correct learning.

Theorem 4. [1] $\mathbf{TxtEx} \not\subseteq \mathbf{DecBc}$.

The class $\mathcal{L}_{\mathbf{Ex}}$ which witnesses above theorem can be defined as follows.

Let K denote the halting problem. Let $\mathbf{M}_0, \mathbf{M}_1, \ldots$ be recursive enumeration of all learning machines. Then one constructs K-recursive sequences $e_0, e_1, \ldots$ and $\sigma_0, \sigma_1, \ldots$ such that

- for all x, σ_x is a finite sequence and $\{y \mid y < x\} \subseteq \text{content}(\sigma_x) \subset W_{\mathbf{M}_{e_x}(\sigma_x)} \subseteq \{y \mid y \neq x\}$;
- for all e, if $\mathbf{M}_e$ **TxtBc**-identifies infinitely many cosingleton sets and does not conjecture N on any input, then there is an x with $e_x = e$.

Then, $\mathcal{L}_{\mathbf{Ex}} = \{\text{content}(\sigma_x) \mid x \in N\} \cup \{W_{\mathbf{M}_{e_x}(\sigma_x)} \mid x \in N\}$, can be used to show Theorem 4.

Interestingly, if one considers second-time decisive, where a learner is not allowed to return to a twice abandoned hypothesis, then it is not restrictive in the context of explanatory learning. However, this notion is still restrictive for vacillatory learning.

We now formally consider U-shaped behaviour of a learner.

Definition 5. [2]

(a) A learner $\mathbf{M}$ is *non U-shaped* on a text T, iff there do not exist m, n, t such that $m < n < t$, and $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])} = \text{content}(T)$, but $W_{\mathbf{M}(T[n])} \neq \text{content}(T)$.

(b) A learner $\mathbf{M}$ is non U-shaped on a language L if it is non U-shaped on each text for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner $\mathbf{M}$ **NUShJ**-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, it is non U-shaped on L and $\mathbf{J}$-identifies L.

One can define the class **NUShJ** similarly. Intriguingly, unlike the decisive case, non U-shaped learning does not hurt for explanatory learning.

Theorem 6. [1] $\mathbf{NUShEx} = \mathbf{TxtEx}$.

However, non U-shapedness does restrict behaviourally and vacillatory learning as it is easy to see that the example $\mathcal{L}_{Bc} \in \mathbf{TxtBc} - \mathbf{DecBc}$ of Fulk, Jain and Osherson [17] is also not **NUShBc**-identifiable.

Theorem 7. [1, 17] **NUShBc** $\subset$ **TxtBc**.

Theorem 8. [9] *Suppose* $b \in \{1, 2, \ldots, *\}$. *Then* $\mathbf{NUShFex}_b \subset \mathbf{TxtFex}_b$.

In fact, non U-shaped requirement puts severe constraints on vacillatory learning as [9] showed that $\mathbf{NUShFex}_* = \mathbf{NUShTxtEx} = \mathbf{TxtEx}$. Thus, any vacillatory learner for learning the classes in $\mathbf{TxtFex}_b - \mathbf{TxtEx}$, for $b \geq 2$, necessarily exhibits U-shaped behaviour!

An interesting question is whether non U-shaped requirement can be circumvented for classes in $\mathbf{TxtFex}_b - \mathbf{TxtEx}$ if one allows behaviourally correct learning. Here surprisingly one can circumvent non U-shaped behaviour for classes in $\mathbf{TxtFex}_2$, but not necessarily for classes in $\mathbf{TxtFex}_3$!

Theorem 9. [9] $\mathbf{TxtFex}_2 \subseteq \mathbf{NUShTxtBc}$.

Theorem 10. [9] $\mathbf{TxtFex}_3 \not\subseteq \mathbf{NUShTxtBc}$.

The class witnessing the above theorem can be constructed as follows. Let $\langle \cdot, \cdot \rangle$ denote a comptuable 1–1 pairing function from $N \times N$ to N. $\langle \cdot, \cdot \rangle$ can be extended to triple (and n-tuples) by using $\langle x, y, z \rangle = \langle x, \langle y, z \rangle \rangle$.

Let $L_{i,j} = \{\langle i, j, k \rangle \mid k \in N\}$, $I_{i,j} = W_i \cap L_{i,j}$ and $J_{i,j} = W_j \cap L_{i,j}$ for $i, j \in N$. Then,

$$\mathcal{L} = \{L_{i,j} \mid i, j \in N\} \cup \{I_{i,j}, J_{i,j} \mid i, j \in N \wedge I_{i,j} \subset J_{i,j} \wedge |I_{i,j}| < \infty\}$$

witnesses the separation in the above theorem.

On the other hand, it can be shown that there are classes which can be learnt in non U-shaped manner in behaviourally correct model, but which cannot be learned, even U-shapedly, in vacillatory learning model. An example is the class of the graphs of those functions f for which $\varphi_{f(0)}$ is defined at almost all inputs and f is a total extension of $\varphi_{f(0)}$.

Theorem 11. **NUShBc** $\not\subseteq$ $\mathbf{TxtFex}_*$.

3 Consistent Learning

Consistency requires that the hypothesis output at any stage by the learner contains the input seen until then.

Definition 12. [3] (a) **M** is *consistent* on L, iff for all texts T for L, for all n, $\mathrm{content}(T[n]) \subseteq W_{\mathbf{M}(T[n])}$.

(b) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner **M** **ConsJ**-identifies $\mathcal{L}$, iff **M** is consistent on each $L \in \mathcal{L}$, and it **J**-identifies L.

One may combine consistency also with the decisive or non U-shapedness requirement considered earlier.

Intuitively consistency is a very natural expectation. However for explantory learning it is a severe restriction, as **ConsEx** $\subset$ **TxtEx**. Interestingly, every class in **ConsEx** can also be decisively learnt (while preserving consistency).

Theorem 13. [10] **ConsEx** = **DecConsEx**.

On the other hand one can show that

Theorem 14. [10] **DecEx** $\not\subseteq$ **ConsEx**.

Note that every behaviourally correct learner can be trivially made consistent, by just patching the input. Thus,

Proposition 15. **NUShBc** $\subset$ **ConsBc**.

As mentioned above, every behaviourally correct learner can be trivially made consistent, by patching the input. However this patching may not preserve non U-shapedness. A more involved construction can be used to show the following.

Theorem 16. [10] **NUShBc** = **NUShConsBc**.

4 Team Learning

Smith [27] studied learning by teams of machines. Intuitively, a team of machines is successful in learning, if some predetermined number of members of the team are successful in learning.

Definition 17. [27] A class $\mathcal{L}$ is in $[m, n]$**TxtEx** iff there is a team, $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_n$, of n machines such that for all $L \in \mathcal{L}$, for every text T for L, at least m of the n machines in the team converge on T to a grammar for L.

A non U-shaped learner does not make a mind change from a correct hypothesis to an incorrect one. For learning by a team in non U-shaped manner, we require such a property from each member of the team.

Definition 18. [9] A class $\mathcal{L}$ is in $[m, n]$**NUShEx** iff there are n machines such that on any text for any language L in $\mathcal{L}$

(a) at least m machines in the team converge to a grammar for L and

(b) no machine in the team makes a mind change from a grammar for L to a grammar for some other language.

The following result characterizes vacillatory learning in terms of teams.

Theorem 19. [9] *$\mathcal{L} \in \mathbf{TxtFex}_n$ iff there exists a team $\mathbf{M}_1, \ldots, \mathbf{M}_n$ of n machines such that*

(a) *$\mathbf{M}_1, \ldots, \mathbf{M}_n$ witness that $\mathcal{L} \in [1, n]\mathbf{NUShEx}$, and*

(b) *each of $\mathbf{M}_1, \ldots, \mathbf{M}_n$ converge on each text for every $L \in \mathcal{L}$.*

However note that $\mathbf{TxtFex}_m \subset [1, m]\mathbf{NUShEx}$, and thus we cannot drop the requirement (b) from above characterization. Here is another result:

Theorem 20. [9] *For $m \geq 1$, $\mathbf{TxtFex}_m \subseteq [2, m+1]\mathbf{NUShEx}$.*

It follows that $[2,3]\mathbf{NUShEx} \not\subseteq \mathbf{NUShEx}$. For team learning, it can be shown that non U-shapedness is a restriction.

Theorem 21. [9] *For* $m \geq 2$, $[1,m]\mathbf{NUShEx} \subset [1,m]\mathbf{TxtEx}$.

However, one can mitigate the above by considering larger teams as follows.

Theorem 22. [9] *For* $m \geq 1$ *and* $n \geq m$, $[m,n]\mathbf{TxtEx} \subseteq [m, m+n]\mathbf{NUShEx}$.

In particular, we have the following hierarchy for non U-shaped team learning.

Theorem 23. [9] *For* $n \geq 1$, $[1,n]\mathbf{NUShEx} \subset [1, n+1]\mathbf{NUShEx}$.

5 Some Related Criteria

U-shaped learning can be seen as a special case of more general situation where a learner abandons an hypothesis and comes back to it later. One may put different requirements on which type of abandoned conjectures a machine may return to. Non-U-shaped learning concerns the situation when the learner is not allowed to return to abandoned correct conjectures. As a dual, one can consider the situation when a learner is not allowed to return to abandoned wrong conjectures. When a learner returns to correct conjecture, one may view this as being dictated by the requirements of learning the input – however, returning to wrong conjectures seems to put in unnecessary ineffeciency in the learner. Examples of this kind of apparently inefficient behaviour have been documented by developmental psychologists in the context of infants' face recognition. For example, it has been shown that children exhibit an "inverted-U-shaped" (wrong-correct-wrong) learning curve for recognition of inverted faces and an "N-shaped" (wrong-correct-wrong-correct) learning curve for recognition of upright faces [15, 16]. Formally one can define non-return to wrong hypothesis as follows.

Definition 24. [10]

(a) We say that $\mathbf{M}$ is *decisive on wrong conjectures* (abbreviated *Wr-decisive*) on text T, if there do not exist any m, n, t such that $m < n < t$, and $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])} \neq \text{content}(T)$ and $W_{\mathbf{M}(T[m])} \neq W_{\mathbf{M}(T[n])}$.

(b) We say that $\mathbf{M}$ is Wr-decisive on L if $\mathbf{M}$ is Wr-decisive on each text for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner $\mathbf{M}$ $\mathbf{WrDJ}$-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, $\mathbf{M}$ is Wr-decisive on L and $\mathbf{J}$-identifies L.

One can similarly define the class $\mathbf{WrDJ}$.

Interestingly for explanatory learning Wr-decisive learning coincides with decisive learning. Thus, if one could learn a class by not returning to wrong conjectures, then one may as well learn the class without returning to any conjectures, correct or wrong.

Theorem 25. [10] $\mathbf{WrDEx} = \mathbf{DecEx}$.

As a corollary we have that restricting return to wrong conjectures does hurt exaplanatory learnability.

For vacillatory learning, non-return to wrong conjectures forces its collapse to **WrDEx** = **DecEx**.

Theorem 26. [10] $\mathbf{WrDFex}_* = \mathbf{DecEx}$.

Thus, for both explanatory and vacillatory learning, allowing return to wrong conjectures is more crucial than allowing return to correct conjectures. On the other hand for behaviourally correct learning, these two notions are incomparable!

Theorem 27. [10] $\mathbf{WrDBc} \not\subseteq \mathbf{NUShBc}$ *and* $\mathbf{NUShBc} \not\subseteq \mathbf{WrDBc}$.

In fact, **WrDBc** does not even contain **TxtEx**.

As mentioned earlier, inverted-U-shaped learning (wrong-correct-wrong sequence of conjectures) has also attracted attention of from psychologists for face recognition by children. This leads us to the following definition.

Definition 28. [10]

(a) We say that **M** is *non inverted-U-shaped* on text T, if there do not exist any m, n, t such that $m < n < t$, $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])} \neq W_{\mathbf{M}(T[n])} = \mathrm{content}(T)$.

(b) We say that **M** is non inverted-U-shaped on L if **M** is non inverted-U-shaped on each text for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner **M** **NInvUJ**-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, **M** is non inverted-U-shaped on L and it **J**-identifies L.

Note that this definition does not rule out all wrong-correct-wrong sequences of conjectures but only returning to an equivalent wrong hypothesis after having conjectured a correct one. So every non U shaped learner is also non inverted-U-shaped but the converse does not hold. For that reason, **NInvUEx** = **TxtEx** follows directly from **NUShEx** = **TxtEx**. Furthermore, one can show that non inverted-U-shaped learning is also not restrictive for behaviourally correct learning which stands in contrast to the fact that non U-shaped behaviourally correct learning is restrictive.

Theorem 29. [10] $\mathbf{NInvUEx} = \mathbf{TxtEx}$ *and* $\mathbf{NInvUBc} = \mathbf{TxtBc}$.

On the other hand,

Theorem 30. [10] $\mathbf{NInvUFex}_* = \mathbf{NInvUEx} = \mathbf{TxtEx} \subset \mathbf{TxtFex}_*$.

Overgeneralization is one of the crucial concerns in language learning from positive data: as a learner is not getting negative data, a learner may not be able to restrict its conjecture after overgeneralizing, based on data available. So an interesting variant to consider is whether one can avoid return to overgeneralized conjectures. We consider two variants here.

Definition 31. [10]

(a) We say that $\mathbf{M}$ is *decisive on overinclusive conjectures* (abbreviated *OI-decisive*) on text T, if there do not exist m, n, t such that $m < n < t$, $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])} \not\subseteq \text{content}(T)$ and $W_{\mathbf{M}(T[m])} \neq W_{\mathbf{M}(T[n])}$.

(b) We say that $\mathbf{M}$ is OI-decisive on L if $\mathbf{M}$ is OI-decisive on each text for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner $\mathbf{M}$ $\mathbf{OIDJ}$-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, $\mathbf{M}$ is OI-decisive on L and it $\mathbf{J}$-identifies L.

Definition 32. [10]

(a) We say that $\mathbf{M}$ is *decisive on overgeneralizing conjectures* (abbreviated *OG-decisive*) on text T, if there do not exist m, n, t such that $m < n < t$, $W_{\mathbf{M}(T[m])} = W_{\mathbf{M}(T[t])} \supset \text{content}(T)$ and $W_{\mathbf{M}(T[m])} \neq W_{\mathbf{M}(T[n])}$.

(b) We say that $\mathbf{M}$ is OG-decisive on L if $\mathbf{M}$ is OG-decisive on each text for L.

(c) Suppose $\mathbf{J} \in \{\mathbf{Ex}, \mathbf{Fex}_b, \mathbf{Bc}\}$. A learner $\mathbf{M}$ $\mathbf{OGDJ}$-identifies $\mathcal{L}$, iff for each $L \in \mathcal{L}$, $\mathbf{M}$ is OG-decisive on L and it $\mathbf{J}$-identifies L.

For explanatory and behaviourally correct learning both the above forms are not restrictive. For vacillatory learning, $\mathbf{OIDFex}_*$ coincides with $\mathbf{OIDEx}$ and is thus restrictive.

Theorem 33. [10]

(a) $\mathbf{OIDEx} = \mathbf{TxtEx}$ *and* $\mathbf{OGDEx} = \mathbf{TxtEx}$.

(b) $\mathbf{OIDBc} = \mathbf{TxtBc}$ *and* $\mathbf{OGDBc} = \mathbf{TxtBc}$.

(c) $\mathbf{OIDFex}_* = \mathbf{TxtEx} \subset \mathbf{TxtFex}_*$.

While these results fit into what was already observed for the notion of non inverted-U-shaped learning, forbidding return to overgeneralized hypothesis is a bit different. It is restrictive for vacillatory learning but not as much as some of the other constraints we have discussed above.

Theorem 34. [10]

(a) $\mathbf{TxtFex}_2 \not\subseteq \mathbf{OGDFex}_*$.

(b) *For* $m \geq 1$, $\mathbf{OGDFex}_{m+1} \not\subseteq \mathbf{TxtFex}_m$.

6 Acknowledgements

This survey was based on work of several people, including Ganesh Baliga, Lorenzo Carlucci, John Case, Mark Fulk, Efim Kinber, Wolfgang Merkle, Daniel Osherson and Rolf Wiehagen, who coauthored some of the papers on the topic with us. The reader may refer to [1, 2, 8–10] for more detailed discussion and proofs of the results presented here.

References

1. G. Baliga, J. Case, W. Merkle, and F. Stephan. Unlearning helps. In Ugo Montanari, José D. P. Rolim, and Emo Welzl, editors, *Automata, Languages and Programming, 27th International Colloquium*, volume 1853 of *Lecture Notes in Computer Science*, pages 844–855. Springer-Verlag, 2000.
2. G. Baliga, J. Case, W. Merkle, F. Stephan, and R. Wiehagen. When unlearning helps. Manuscript, http://www.cis.udel.edu/˜case/papers/decisive.ps, 2005. Preliminary version of the paper appeared in ICALP, 2000.
3. J. Bārzdiņš. Inductive inference of automata, functions and programs. In *Int. Math. Congress, Vancouver*, pages 771–776, 1974.
4. J. 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.
5. T.G.R. Bower. Concepts of development. In *Proceedings of the 21st International Congress of Psychology*, pages 79–97. Presses Universitaires de France, 1978.
6. M. 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*. Developmental Psychology Series. Academic Press, New York, 1982.
7. S. Carey. Face perception: Anomalies of development. In S. Strauss and R. Stavy, editors, *U-Shaped Behavioral Growth*, Developmental Psychology Series. Academic Press, New York, 1982.
8. L. Carlucci, J. Case, S. Jain, and F. Stephan. Non u-shaped vacillatory and team learning. In Sanjay Jain, Hans Ulrich Simon, and Etsuji Tomita, editors, *Algorithmic Learning Theory: Sixteenth International Conference (ALT' 2005)*, volume 3734 of *Lecture Notes in Artificial Intelligence*, pages 241–255. Springer-Verlag, 2005.
9. L. Carlucci, J. Case, S. Jain, and F. Stephan. U-shaped learning may be necessary. *Journal of Computer and System Sciences*, 2005. To appear.
10. L. Carlucci, S. Jain, E. Kinber, and F. Stephan. Variations on u-shaped learning. Technical Report TRA8/05, School of Computing, National University of Singapore, 2005. Preliminary version of this paper appeared in [11].
11. L. Carlucci, S. Jain, E. Kinber, and F. Stephan. Variations on u-shaped learning. In P. Auer and R. Meir, editors, *Proceedings of the Eighteenth Annual Conference on Learning Theory*, volume 3559 of *Lecture Notes in Artificial Intelligence*, pages 382–397. Springer-Verlag, 2005.
12. J. Case. The power of vacillation in language learning. *SIAM Journal on Computing*, 28(6):1941–1969, 1999.
13. J. Case and C. 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*, volume 140 of *Lecture Notes in Computer Science*, pages 107–115. Springer-Verlag, 1982.
14. J. Case and C. Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
15. C.H. Cashon and L.B. Cohen. The construction, deconstruction and reconstruction of infant face perception. In A. Slater and O. Pascalis, editors, *The development of face processing in infancy and early childhood*, pages 55–58. NOVA Science Publishers, New York, 2003.
16. C.A. Cashon and L.B. Cohen. Beyond U-shaped development in infants' processing of faces: An information-processing account. *Journal of Cognition and Development*, 5(1):59–80, 2004.

17. M. Fulk, S. Jain, and D. Osherson. Open problems in systems that learn. *Journal of Computer and System Sciences*, 49(3):589–604, 1994.
18. E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
19. S. Jain, D. Osherson, J. Royer, and A. Sharma. *Systems that Learn: An Introduction to Learning Theory*. MIT Press, Cambridge, Mass., second edition, 1999.
20. R. Klette and R. Wiehagen. Research in the theory of inductive inference by GDR mathematicians – A survey. *Information Sciences*, 22:149–169, 1980.
21. G. Marcus, S. Pinker, M. Ullman, M. Hollander, T. Rosen, and F. 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.
22. D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, 1986.
23. D. Osherson and S. Weinstein. Criteria of language learning. *Information and Control*, 52:123–138, 1982.
24. S. Pinker. Formal models of language learning. *Cognition*, 7:217–283, 1979.
25. K. Plunkett and V. Marchman. U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, 38(1):43–102, 1991.
26. H. Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967. Reprinted by MIT Press in 1987.
27. C. Smith. The power of pluralism for automatic program synthesis. *Journal of the ACM*, 29:1144–1165, 1982.
28. S. Strauss and R. Stavy. *U-Shaped Behavioral Growth*. Developmental Psychology Series. Academic Press, New York, 1982.
29. S. Strauss, R. Stavy, and N. Orpaz. The child's development of the concept of temperature. Manuscript, Tel-Aviv University, 1977.
30. N.A. Taatgen and J.R. Anderson. Why do children learn to say broke? a model of learning the past tense without feedback. *Cognition*, 86(2):123–155, 2002.
31. K. Wexler and P. Culicover. *Formal Principles of Language Acquisition*. MIT Press, 1980.