Learning Languages in a Union

Sanjay Jain [a,1] Yen Kaow Ng [b] Tiong Seng Tay [c]

[a] *School of Computing, National University of Singapore, Singapore 119260*

[b] *Kyushu Institute of Technology, Graduate School of Computer Science and Systems Engineering, Iizuka 820, Japan*

[c] *Department of Mathematics, National University of Singapore, Singapore 119260*

Abstract

In inductive inference, a machine is given words of a language (a recursively enumerable set in our setting) and the machine is said to identify the language if it correctly names the language. In this paper we study identifiability of classes of languages where the unions of up to a fixed number (n say) of languages from the class are provided as input. We distinguish between two different scenarios: in one scenario, the learner need only to name the language which results from the union; in the other, the learner must individually name the languages which make up the union (we say that the unioned language is *discerningly* identified). We define three kinds of identification criteria based on this and by the use of some classes of disjoint languages, demonstrate that the inferring power of each of these identification criterion decreases as we increase the number of languages allowed in the union, thus resulting in an infinite hierarchy for each identification criterion. That is, we show that for each n, there exists a class of disjoint languages where all unions of up to n languages from this class can be discerningly identified, but there is no learner which identifies every union of $n+1$ languages from this class. A comparison between the different identification criteria also yielded similar hierarchies. We give sufficient conditions for classes of languages where the unions can be discerningly identified, and characterize such discerning learnability for the indexed families. We then give naturally occurring classes of languages that witness some of the earlier hierarchical results. Finally, we present language classes which are complete with respect to weak reduction (in terms of intrinsic complexity) for our identification criteria.

Key words: Unions of languages, Inductive Inference

Email address: `sanjay@comp.nus.edu.sg` (Sanjay Jain).

[1] Supported in part by NUS grant number R252–000–127–112.

1 Introduction

A model for learning of languages (recursively enumerable sets, in our context) can be described thus. A learner is presented with all the elements, one element at a time in any order, of a language L (such a presentation is called a text for the language L). The learner, as it is receiving the data, outputs a sequence of hypotheses (grammars in our context). The learner is said to identify (learn, infer) L just in case the sequence of grammars converges to a grammar for L. A class of languages is learnable if some machine learns each language in the class. This is essentially the model of learning in the limit (called **TxtEx**) introduced by Gold [Gol67]. Note that in this setting the learner is not informed of the elements absent from the language to be learned. The aim of using only such positive data was more to address the point that children rarely, if ever, get negative data. One should also note that many natural phenomenon can be coded as languages via some coding of events into natural numbers.

In this paper, we continue a line of enquiry explored in [Wri89,SA00,GK99], where the learner is required to learn unions of languages drawn from a class of languages. What is different from previous studies is that we distinguish between two different scenarios. In one scenario, the learner is only required to name the language which results from the union (this is the case studied in the earlier enquiries); in the other, we want the learner to individually name the languages which make up the union — in a sense, the learner is discerning between the languages in the union. Our approach to the problem is motivated by the abundance of situations where learners are presented with information that is some sort of mixture. For example, children in a multi-lingual environment are frequently exposed to more than one (natural) languages at the same time, but are nonetheless able to tell what are the languages they hear; or, in a physical experiment, radiations collected by the same detector may originate from many different source processes, for which scientists are often put to the task of discerning. We hope that our study can be useful in devising mechanisms which will allow us to distinguish between languages that has to be presented as a mixture.

A technical question arises from this new approach to the problem. In the course of identifying the languages which make up a union, what happens when there are two (or more) possible sets of languages from the class which unions to the same language? Should the learner be required to name both possibilities, or should the learner be allowed to choose any one? Or perhaps such a situation should be simply declared unlearnable? We formalize different identification criteria based on these considerations.

Our results show that in general, the inferring power of learners lessens when more languages are allowed in the union, and moreover, a few of the hierarchies are witnessed by classes of *disjoint* languages. More precisely, for each n, there exists a class of disjoint languages where all the unions of up to n languages from this class can be *discerningly identified*, but there is no learner that can identify every union of $n+1$ languages from this class. We also noticed hierarchies between each of the different identification criteria, and made attempts at the conditions under which disjointness is sufficient for learnability under the new identification criteria.

In our attempt to characterize these identification criteria, we discovered two sufficient conditions for classes of languages where the unions can be discerningly identified. We demonstrate that one of these conditions is difficult to be further relaxed, by showing how some weaker conditions are insufficient to hold up the same results. We also characterize discerning identifiability for the *indexed families* [Ang80].

We note that there are naturally occurring classes of languages which hold up the hierarchies discussed. For example, a class formed using translations of a 'unit' simplex can be used to form a hierarchy (based on n) for the discerning learnability of unions of up to n languages. A modification of this class is used to form a hierarchy for the non-discerning case.

Finally, we give natural classes of languages which are complete with respect to weak reduction in terms of *intrinsic complexity* [FKS95] for the identification criteria we defined.

2 Notation and Preliminaries

Any unexplained recursion-theoretic notation is from [Rog67]. N denotes the set of natural numbers. N^+ denotes the set of positive integers. Let *rat* denote the set of non-negative rational numbers. $\emptyset$, $\in$, $\subset$, $\subseteq$, $\supset$, $\supseteq$ respectively denote empty set, element of, proper subset, subset, proper superset and superset. max(.), min(.) respectively denote maximum and minimum of a set, where by convention $\max(\emptyset) = 0$ and $\min(\emptyset) = \infty$. Cardinality of a set S is denoted by $card(S)$. We write $card(S) < \infty$, or $card(S) \leq *$, just in case S has finite but unbounded cardinality. $D_0, D_1, \ldots$ stand for a computable sequence of all finite sets [Rog67].

$\langle \cdot, \cdot \rangle$ stands for an arbitrary, computable bijective mapping from $N \times N$ onto N. For all x and y, $\pi_1(\langle x, y \rangle) = x$ and $\pi_2(\langle x, y \rangle) = y$. We assume without loss of generality that $\langle \cdot, \cdot \rangle$ is monotonically increasing in both of its arguments. $\langle \cdot, \cdot \rangle$ can be extended to n-tuples in a natural way (including $n = 1$, where $\langle x \rangle$ may be taken to be x). Projection functions $\pi_1^n, \ldots, \pi_n^n$ corresponding to n-

tuples can be defined similarly. Due to the above isomorphism between N^n and N, we often identify the tuple $(x_1, \ldots, x_n)$ with $\langle x_1, \ldots, x_n \rangle$. The quantifiers $\stackrel{\infty}{\forall}$, $\stackrel{\infty}{\exists}$ and $\exists!$ denote, for all but finitely many, there exists infinitely many and there exists a unique, respectively.

A computable numbering is a partial computable function from $N \times N$ to N. The symbol ψ ranges over computable numberings. We denote by ψ_i, the partial function, $\lambda x.\psi(i, x)$. Thus ψ_i denotes the partial function computed by the program with index i in the numbering ψ. Ψ denotes an arbitrary Blum [Blu67] complexity measure for ψ. W_i^ψ denotes $domain(\psi_i)$. W_i^ψ is, then, the recursively enumerable (r.e.) set/language ($\subseteq N$) accepted (or equivalently, generated) by the ψ-program i. We also say that i is a ψ-grammar for W_i^ψ. $W_{i,s}^\psi$ denotes the set $\{x \leq s \mid \Psi_i(x) \leq s\}$. We say that numbering ψ is reducible to numbering ψ' (written $\psi \preceq \psi'$) if and only if there exists a recursive function h such that $(\forall i)[\psi'_{h(i)} = \psi_i]$. In this case we say that h witnesses that $\psi \preceq \psi'$. An acceptable numbering is a computable numbering to which every computable numbering can be reduced. The symbol φ denotes a standard acceptable numbering [Rog67] and the symbol Φ denotes an arbitrary fixed Blum complexity measure for the φ-system [Blu67]. In this paper we abbreviate W_i^φ to W_i, and $W_{i,s}^\varphi$ to $W_{i,s}$.

$\mathcal{E}$ denotes the class of all r.e. languages. $\mathcal{R}$ denotes the set of all recursive functions, that is total computable functions. Symbol L, with or without decorations, ranges over $\mathcal{E}$. The symbol $\mathcal{L}$, with or without decorations, ranges over subsets of $\mathcal{E}$. K denotes the diagonal halting problem set, that is, $K = \{x \mid x \in W_x\}$. (K is a recursively enumerable, non-recursive set.) FIN denotes the class $\{D \subset N \mid D \text{ is finite}\}$. $INIT$ denotes the class $\{\{x \in N \mid x \leq n\} \mid n \in N\}$.

A class $\mathcal{L}$ of r.e. languages is said to be *recursively enumerable* [Rog67] if there is $S \in \mathcal{E}$ such that $\mathcal{L} = \{W_i \mid i \in S\}$. For each non-empty, recursively enumerable class of languages $\mathcal{L}$, there exists a total recursive function f such that $\mathcal{L} = \{W_{f(i)} \mid i \in N\}$. $\mathcal{L}$ is said to be 1–1 recursively enumerable if and only if (i) $\mathcal{L}$ is finite or (ii) there exists a recursive function f such that $\mathcal{L} = \{W_{f(i)} \mid i \in N\}$ and $W_{f(i)} \neq W_{f(j)}$, for $i \neq j$. In this latter case we say that $W_{f(0)}, W_{f(1)}, \ldots$ is a 1–1 recursive enumeration of $\mathcal{L}$.

We say that a family of recursive languages $\{L_0, L_1, \ldots\}$ is an indexed family [Ang80] iff there exists a recursive function f such that $f(i, x) = 1$, if $x \in L_i$; $f(i, x) = 0$, if $x \notin L_i$.

A partial function d from N to N is said to be partial limiting recursive, if and only if there exists a recursive function F from $N \times N$ to N such that for all x, $d(x) = \lim_{y\to\infty} F(x, y)$. Here if $d(x)$ is not defined then $\lim_{y\to\infty} F(x, y)$ must also be undefined. A partial limiting recursive function d is called (total) limiting recursive, if d is total. $\downarrow$ denotes defined or converges. $\uparrow$ denotes undefined or diverges.

We now present concepts from language learning theory. The next definition introduces the concept of a *sequence* of data.

Definition 1 [Gol67]

(a) A *sequence* σ is a mapping from an initial segment of N into $(N \cup \{\#\})$. The empty sequence is denoted by λ.

(b) The *content* of a sequence σ, denoted $content(\sigma)$, is the set of natural numbers in the range of σ.

(c) The *length* of σ, denoted by $|\sigma|$, is the number of elements in σ. So, $|\lambda| = 0$.

(d) For $n \leq |\sigma|$, the initial sequence of σ of length n is denoted by $\sigma[n]$. So, $\sigma[0] = \lambda$.

(e) For any two sequences σ and τ, the result of concatenating τ to the end of σ is written $\sigma\tau$.

Intuitively, #'s represent pauses in the presentation of data. We let σ and τ, with or without decorations, range over finite sequences. SEQ denotes the set of all finite sequences.

Definition 2 [Gol67]

(a) A *text* T for a language L is a mapping from N into $(N \cup \{\#\})$ such that L is the set of natural numbers in the range of T.

(b) The *content* of a text T, denoted by $content(T)$, is the set of natural numbers in the range of T; that is, the language which T is a text for.

(c) $T[n]$ denotes the finite initial sequence of T with length n.

We let T, with or without decorations, range over texts. We let $\mathcal{T}$ range over sets of texts.

Definition 3 [Gol67] An *inductive inference machine* (**IIM**) is an algorithmic device which computes a mapping from SEQ into N.

We let M, with or without decorations, range over the **IIM**s. $M(T[n])$ is interpreted as the grammar (index for an accepting program) conjectured by the machine M on the initial sequence $T[n]$. We say that M converges on T to i (written $M(T)\downarrow = i$) iff for all but finitely many n, $M(T[n]) = i$.

Gold [Gol67] introduced the following language learning criterion known as

TxtEx-identification.

Definition 4 [Gol67]

(a) M ***TxtEx**-identifies* a text T just in case there exists $i \in N$ such that $W_i = \mathit{content}(T)$, and $M(T)\downarrow = i$.

(b) M ***TxtEx**-identifies* an r.e. language L (written $L \in \mathbf{TxtEx}(M)$) just in case M **TxtEx**-identifies each text for L.

(c) M ***TxtEx**-identifies* a class $\mathcal{L}$ of r.e. languages (written $\mathcal{L} \subseteq \mathbf{TxtEx}(M)$) just in case M **TxtEx**-identifies each language from $\mathcal{L}$.

(d) $\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{TxtEx}(M)]\}$.

Note that there exists a recursive sequence $M_0, M_1, \ldots$ of total **IIM**s such that every class in **TxtEx** is **TxtEx**-identified by at least one of the machines in the sequence [OSW86]. Similarly, one can further show that there exists a recursive sequence $M_0, M_1, \ldots$ of total **IIM**s such that for any criteria $\mathcal{J}$ of inference considered in this paper, every class in $\mathcal{J}$ is $\mathcal{J}$-identified by at least one of the machines in the sequence. We assume $M_0, M_1, \ldots$ to be one such recursive sequence of total **IIM**s.

Definition 5 [BB75,Ful90] Let $L \in \mathcal{E}$, **IIM** M and $\sigma \in$ SEQ be given. σ is a *stabilizing sequence* for M on L just in case:

(a) $\mathit{content}(\sigma) \subseteq L$;

(b) for all $\tau \in$ SEQ, if $\mathit{content}(\tau) \subseteq L$, then $M(\sigma\tau) = M(\sigma)$.

It can be shown [BB75,Ful90] that for any language L, which is **TxtEx**-identified by M, there exists a stabilizing sequence for M on L. Similar result can be shown for learning of unions of languages considered below.

3 Identification of Unions of Languages

We first define the class formed by taking unions of languages.

Definition 6 [SA00] Let $k \in N^+ \cup \{*\}$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) The *union language of* $\mathcal{L}$, denoted $L_{\mathcal{L}}$, is the set $\bigcup_{L\in\mathcal{L}} L$.

(b) The class of *at most k unions of* $\mathcal{L}$, $\mathcal{L}^k = \{\ L_{\mathcal{L}'} \mid \mathcal{L}' \subseteq \mathcal{L} \ \wedge\ \mathit{card}(\mathcal{L}') \leq k\}$.

We now define the identification criterion $\mathbf{U}^k\mathbf{TxtEx}$ which requires that not only $\mathcal{L}$ but also $\mathcal{L}^k$ to be **TxtEx** learnable. The **U** in $\mathbf{U}^k\mathbf{TxtEx}$ stands for *Union.*

Definition 7 Let $k \in N^+ \cup \{*\}$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\boldsymbol{U^k TxtEx}$*-identifies* $\mathcal{L}$ just in case $\mathcal{L}^k \subseteq \mathbf{TxtEx}(M)$.

(b) $\mathbf{U}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{U}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

$\mathbf{U}^k\mathbf{TxtEx}$ coincides with the definition of "identification of unions of languages" in [Wri89,SA00].

We now define the identification criterion $\mathbf{DU}^k\mathbf{TxtEx}$ where the learner is required to identify $\mathcal{L}^k$ by individually identifying the languages in any k' ($\leq k$) languages $L_1, \ldots, L_{k'} \in \mathcal{L}$, when presented with a text for $L_1 \cup \ldots \cup L_{k'}$. The **D** in $\mathbf{DU}^k\mathbf{TxtEx}$ stands for *Discernible.*

Definition 8 Suppose $\mathcal{L} \subseteq \mathcal{E}$ and $card(\mathcal{L}) < \infty$.

(a) We say that a set of indices $\{x_1, x_2, \ldots, x_{card(\mathcal{L})}\} \subseteq N$ is a *representation index set of* $\mathcal{L}$ just in case $\{W_{x_1}, W_{x_2}, ..., W_{x_{card(\mathcal{L})}}\} = \mathcal{L}$.

(b) Let $\mathcal{I}_{\mathcal{L}} = \{I \mid I \text{ is a representation index set of } \mathcal{L}\}$.

(c) Let $\mathcal{I} = \{I \mid (\exists \mathcal{L} \subseteq \mathcal{E}, card(\mathcal{L}) < \infty)[I \in \mathcal{I}_{\mathcal{L}}]\}$.

Any representation index set $\{x_1, x_2, \ldots, x_{card(\mathcal{L})}\}$ can be represented by a natural number s where $D_s = \{x_1, x_2, \ldots, x_{card(\mathcal{L})}\}$. This representation is implicit whenever the context requires such an interpretation.

Definition 9 Let $k \in N^+ \cup \{*\}$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\boldsymbol{DU^k TxtEx}$*-identifies* $\mathcal{L}$ just in case for each $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$, for every text T for $L_{\mathcal{L}'}$, $M(T)\downarrow$ and $D_{M(T)} \in \mathcal{I}_{\mathcal{L}'}$.

(b) $\mathbf{DU}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{DU}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

For ease of notation, we sometimes let machines output finite sets directly rather than the canonical index for it (i.e., M outputs S rather than i such that $D_i = S$).

Proposition 10 *Suppose* $\mathcal{L} \subseteq \mathcal{E}$. *If there exist finite* $\mathcal{L}', \mathcal{L}'' \subseteq \mathcal{L}$, *such that* $\mathcal{L}' \neq \mathcal{L}''$ *but* $L_{\mathcal{L}'} = L_{\mathcal{L}''}$, *then* $\mathcal{L} \notin \boldsymbol{DU^k TxtEx}$ *for* $k = \max(card(\mathcal{L}'),\ card(\mathcal{L}''))$.

Proof. A text for $L_{\mathcal{L}'}$ is also a text for $L_{\mathcal{L}''}$. But $\mathcal{I}_{\mathcal{L}'} \cap \mathcal{I}_{\mathcal{L}''} = \emptyset$ and an **IIM** cannot converge to indices for both an $I' \in \mathcal{I}_{\mathcal{L}'}$ and an $I'' \in \mathcal{I}_{\mathcal{L}''}$. ∎

Definition 11 Let $\mathcal{L} \subseteq \mathcal{E}$ and $k \in N^+ \cup \{*\}$. The class of languages $\mathcal{L}^k$ is said to be *uniquely definable from* $\mathcal{L}$ just in case for all $L \in \mathcal{L}^k$, there exists a unique $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$, such that $L_{\mathcal{L}'} = L$.

We now introduce the identification criteria $\mathbf{WDU}^k\mathbf{TxtEx}$ where the complications of Proposition 10 is avoided. The learner is considered correct by simply naming any set of (up to) k languages in the class which make up the language of the input text. The **W** in **WDUTxtEx** stands for *Weakly.*

Definition 12 Let $k \in N^+ \cup \{*\}$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M ***WDU***^{k}***TxtEx****-identifies* $\mathcal{L}$ just in case for each $L \in \mathcal{L}^k$, for every text T for L, $M(T)\downarrow$, and there exists $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$ and T is a text for $L_{\mathcal{L}'}$, such that $D_{M(T)} \in \mathcal{I}_{\mathcal{L}'}$.

(b) $\mathbf{WDU}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{WDU}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

The following Proposition follows from definitions.

Proposition 13 *Suppose* $n \in N^+ \cup \{*\}$.

(a) $\boldsymbol{DU^n TxtEx} \subseteq \boldsymbol{WDU^n TxtEx} \subseteq \boldsymbol{U^n TxtEx}$.

(b) $\boldsymbol{WDU^1 TxtEx} \equiv \boldsymbol{DU^1 TxtEx} \equiv \boldsymbol{U^1 TxtEx} \equiv \boldsymbol{TxtEx}$.

(c) $\mathcal{L} \in \boldsymbol{U^n TxtEx}$ *iff* $\mathcal{L}^n \in \boldsymbol{TxtEx}$.

The following Proposition can be easily verified.

Proposition 14 *For all indexed families* $\mathcal{L}$*, for all* $n \in N^+$*,* $\mathcal{L} \in \boldsymbol{WDU^n TxtEx} \equiv \mathcal{L} \in \boldsymbol{U^n TxtEx}$.

We now state some known results from [Wri89] and [SA00].

Definition 15 [Wri89,MSW91]

(a) A class of languages $\mathcal{L}$ has *infinite elasticity* just in case there exists an infinite sequence of pairwise distinct numbers, $w_0, w_1, \ldots$, and an infinite sequence of pairwise distinct languages in $\mathcal{L}$, $A_0, A_1, \ldots$, such that for each $k \in N$, $\{w_i \mid i < k\} \subseteq A_k$, but $w_k \notin A_k$.

(b) $\mathcal{L}$ is said to have *finite elasticity* just in case $\mathcal{L}$ does not have infinite elasticity.

Theorem 16 *[Wri89] If an indexed family $\mathcal{L}$ has finite elasticity, then $\mathcal{L} \in$ $\boldsymbol{U^n\,TxtEx}$ for all $n \in N^+$.*

Definition 17 (a) *[Ang80]* A class of languages $\mathcal{L}$ has *finite thickness* just in case for each $x \in N$, $\{L \in \mathcal{L} \mid x \in L\}$ is finite.

(b) *[Kru72,SA00]* A class of languages $\mathcal{L}$ *has no infinite anti-chain with respect to set inclusion* just in case there does not exist an infinite collection of distinct languages $\{A_i \in \mathcal{L} \mid i \in N\}$, such that for all $i, j \in N$, $i \neq j$, $A_i \not\subseteq A_j$ and $A_j \not\subseteq A_i$.

Theorem 18 *[SA00] Let $\mathcal{L}$ be an indexed family with finite thickness. If $\mathcal{L}$ has no infinite anti-chain with respect to set inclusion, then $\mathcal{L} \in \boldsymbol{U^*\,TxtEx}$.*

The following generalizes Theorem 17 from [SA00] (with essentially the same proof idea).

Theorem 19 *An indexed family $\mathcal{L}$ is in $\boldsymbol{U^*\,TxtEx}$ if for each $L \in \mathcal{L}$, there exists $x_L \in L$ such that for each $L' \in \mathcal{L}$, $x_L \in L' \Rightarrow L \subseteq L'$.*

That Theorem 19 does not characterize $\mathbf{U^*TxtEx}$ identification for the indexed families can be seen from the following $\mathbf{DU^*TxtEx}$-identifiable indexed family $\mathcal{L}$. For $i \in N$, let $L_i = \{\langle i, j\rangle \mid j \in N\}$, $L'_i = \{\langle j, i\rangle \mid j \in N\}$, and let $\mathcal{L} = \{L_i \mid i \in N\} \cup \{L'_i \mid i \in N\}$. (This example was pointed out to us by an anonymous referee).

Note that these earlier results apply specifically to the indexed families. In this paper, when not stated, the results are with respect to the r. e. languages in general. In doing so, we hope to better contrast the difficulties involved in learning the unions of languages under different assumptions of hypotheses space.

4 Hierarchy Results

We now establish hierarchy results for our criteria of learning unions of languages. For the following theorem, it is interesting to note that the class witnessing the separation consists of disjoint languages. This is interesting as non-algorithmically, the distinct languages can be determined from the input text.

Theorem 20 $(\forall n \in N^+)[\boldsymbol{DU^n\,TxtEx} - \boldsymbol{U^{n+1}\,TxtEx} \neq \emptyset]$.

Proof. Let $n \in N^+$ be given. Unless stated otherwise, let e, i, j, k, with or without decorations, range over N, and S, with or without decorations, ranges over finite sets. σ and τ, with or without decorations, range over SEQ. For each $e \in N$, we will construct S_e, L_e^0, $L_e^1, \ldots, L_e^n$ where

$$L_e^0 = \{\langle e,0,0\rangle\} \cup \{\langle e,i,j\rangle \mid 1 \le i \le n,\ j \in S_e\}$$

and for $1 \le i \le n$, L_e^i satisfies the following two properties:

(1) $L_e^i = \{\langle e,i,j\rangle \mid j \in W_{\min(\{\pi_3^3(x) \mid x \in L_e^i\})}\}$.

(2) $\min(\{\pi_3^3(x) \mid x \in L_e^i\}) > \max(S_e)$.

Let $\mathcal{L} = \{L_e^0, L_e^1, ..., L_e^n \mid e \in N\}$. It is clear that for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$. We shall first show that $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$ (for any choice of S_e, $L_e^0, \ldots, L_e^n$ satisfying the above properties). We define an auxiliary recursive function $g : N^3 \mapsto N$ as follows. For each e and j,

$$W_{g(e,0,j)} = \{\langle e,0,0\rangle\} \cup \{\langle e,i,k\rangle \mid 1 \le i \le n \wedge k \in D_j\}$$

and for each $i \ge 1$ and e, j,

$$W_{g(e,i,j)} = \{\langle e,i,k\rangle \mid k \in W_j\}.$$

Now $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$ is witnessed by following M. For each text T and each $m \in N$,

$M(T[m])$:
- $S \leftarrow \emptyset$.
- $A \leftarrow \{e \mid (\exists w \in \mathit{content}(T[m]))[\pi_1^3(w) = e]\}$.
- For each $e \in A$ do
  - $B \leftarrow \mathit{content}(T[m])$.
  - If $\langle e,0,0\rangle \in \mathit{content}(T[m])$ then
    - $C \leftarrow \{j \mid (\forall i \mid 1 \le i \le n)[\langle e,i,j\rangle \in \mathit{content}(T[m])]\}$.
    - Let j be such that $D_j = C$.
    - $S \leftarrow S \cup \{g(e,0,j)\}$.
    - $B \leftarrow B - W_{g(e,0,j)}$.
  - For $i \leftarrow 1$ to n do
    - If there exists j_0 such that $\langle e,i,j_0\rangle \in B$, then
      - For minimum such j_0, let $S \leftarrow S \cup \{g(e,i,j_0)\}$.
  - EndFor
- EndFor
- Output S.

End

It is easy to verify that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. We now show that $\mathcal{L} \notin \mathbf{U}^{n+1}\mathbf{TxtEx}$, for some appropriate choice of L_e^i, for each e, i. For each e here is the construction to show that M_e does not $\mathbf{U}^{n+1}\mathbf{TxtEx}$-identify $\mathcal{L}$. By

Kleene's Recursion Theorem [Rog67] there exists an index e' such that $W_{e'}$ may be defined in stages $s = 0, 1, 2 \ldots$, as below. For each s, $W_{e'}^s$ denotes the finite portion of $W_{e'}$ enumerated just before stage s.

Stage 0: Let σ^1 be such that $content(\sigma^1) = \{\langle e, 0, 0\rangle\} \cup \{\langle e, j, e'\rangle \mid 1 \leq j \leq n\}$. Let $W_{e'}^1 = \{e'\}$. Go to stage 1.

Stage s: Search for $\tau \supseteq \sigma^s$, such that $content(\tau) \subseteq content(\sigma^s) \cup \{\langle e, i, j\rangle \mid 1 \leq i \leq n \ \wedge\ j > \max(W_{e'}^s)\}$, and $M_e(\sigma^s) \neq M_e(\tau)$.

If and when such τ is found, enumerate $\{j \mid (\exists i', 1 \leq i' \leq n)[\langle e, i', j\rangle \in content(\tau)]\}$ into $W_{e'}$.

Let σ^{s+1} be an extension of τ such that $content(\sigma^{s+1}) = \{\langle e, 0, 0\rangle\} \cup \{\langle e, i, j\rangle \mid 1 \leq i \leq n \wedge j \in W_{e'}$ enumerated up to now$\}$.

Go to stage $s + 1$.

If the search for τ failed at any stage s, then let $L_e^0 = content(\sigma^s)$ and let $e'' > \max(W_{e'}^s)$ be such that $\min(W_{e''}) = e''$ (by Kleene's Recursion Theorem [Rog67], such an e'' exists). For each $i \in N$, $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, j\rangle \mid j \in W_{e''}\}$. Since stage s does not succeed, M_e does not **TxtEx**-identify at least one of L_e^0 and $(L_e^0 \cup \bigcup_{i=1}^n L_e^i)$.

If the search is successful at all stages, then let $L_e^0 = \{\langle e, 0, 0\rangle\}$ and for each $i \in N$, $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, x\rangle \mid x \in W_{e'}\}$. Now, M_e fails to converge on the input $\bigcup_{s \in N^+} \sigma^s$, a text for $L_e^0 \cup \bigcup_{i=1}^n L_e^i$.

Theorem follows from above analysis. ∎

Corollary 21 *For all $n \in N^+$.*

(a) $\boldsymbol{U}^{n+1}\boldsymbol{TxtEx} \subset \boldsymbol{U}^n\boldsymbol{TxtEx}$.

(b) $\boldsymbol{DU}^{n+1}\boldsymbol{TxtEx} \subset \boldsymbol{DU}^n\boldsymbol{TxtEx}$.

(c) $\boldsymbol{WDU}^{n+1}\boldsymbol{TxtEx} \subset \boldsymbol{WDU}^n\boldsymbol{TxtEx}$.

The following Theorem gives a finer separation in the above hierarchy.

Theorem 22 *For all $n \in N^+$.*

(a) $(\boldsymbol{WDU}^*\boldsymbol{TxtEx} \cap \boldsymbol{DU}^n\boldsymbol{TxtEx}) - \boldsymbol{DU}^{n+1}\boldsymbol{TxtEx} \neq \emptyset$.

(b) $(\boldsymbol{U}^*\boldsymbol{TxtEx} \cap \boldsymbol{WDU}^n\boldsymbol{TxtEx}) - \boldsymbol{WDU}^{n+1}\boldsymbol{TxtEx} \neq \emptyset$.

Proof. (a) The case of $n = 1$ can be easily shown using the class FIN, which is in $\mathbf{DU^1TxtEx}$ but not in $\mathbf{DU^2TxtEx}$. Note that FIN also belongs to $\mathbf{WDU^*TxtEx}$, as $FIN^* = FIN$.

We now consider $n > 1$. For $i < n$, let $L_i = \{2i, 2i+1\}$. Let $L_n = \{2i \mid i < n\}$. Let $\mathcal{L} = \{L_i \mid i \leq n\}$. It is easy to see that $\mathcal{L} \in \mathbf{WDU^*TxtEx}$. Since, for $i < n$, $2i + 1$ belongs only to language L_i from $\mathcal{L}$, and L_n is not contained in any $n - 1$ languages from $\{L_0, L_1, \ldots, L_{n-1}\}$, any collection of up to n languages from $\mathcal{L}$ can be individually recognized from a text for their union. Thus, $\mathcal{L} \in \mathbf{DU^nTxtEx}$. However, $L_0 \cup \ldots \cup L_{n-1} = L_0 \cup \ldots \cup L_{n-1} \cup L_n$. Thus, by Proposition 10, $\mathcal{L}$ is not in $\mathbf{DU^{n+1}TxtEx}$.

(b) We observe that while a $\mathbf{UTxtEx}$ learner, in learning a class of languages $\mathcal{L}$, is allowed to conjecture languages outside of $\mathcal{L}$, a $\mathbf{WDUTxtEx}$ learner is allowed to conjecture only languages in $\mathcal{L}$. The following proof exploits this weakness in $\mathbf{WDUTxtEx}$ identification.

For each $e \in N$, let $L_e = \{\langle e, 0\rangle, \langle e, 1\rangle, \ldots, \langle e, n\rangle\}$, and let T_e be a canonical text for L_e. Note that each L_e contains $n + 1$ elements, and for any e and e' where $e \neq e'$, $L_e \cap L_{e'} = \emptyset$. We now define a class $\mathcal{L} \notin \mathbf{DU^nTxtEx}$. For each $e \in N$, we let L_e be in $\mathcal{L}$ if M_e does not converge on T_e to an index for L_e, otherwise we let the $n + 1$ languages, $\{x\}$ where $x \in L_e$, be in $\mathcal{L}$.

Since $\mathcal{L}$ contains only finite sets, it is in $\mathbf{U^*TxtEx}$. To $\mathbf{WDU^nTxtEx}$-identify $\mathcal{L}$, it suffices that a learner, on an input sequence σ, find all the non-empty sets $S^i = \{x \mid x \in \mathit{content}(\sigma) \wedge \pi_1(x) = i\}$ where $i \in N$, and include L_i in its conjecture if $\mathit{card}(S^i) = n + 1$, or otherwise include the sets $\{\{x\} \mid x \in S^i\}$ to its conjecture. Since for all $e \in N$, M_e fails to converge on the text T_e to a representation index set for up to $n+1$ languages from $\mathcal{L}$, which union to L_e, we have that $\mathcal{L} \notin \mathbf{WDU^{n+1}TxtEx}$. ∎

Corollary 23 $(\forall n \in N, n \geq 2)[\boldsymbol{DU^n TxtEx} \subset \boldsymbol{WDU^n TxtEx} \subset \boldsymbol{U^n TxtEx}]$.

The following two results (Proposition 24 and Theorem 25) were pointed out to us by an anonymous referee.

Proposition 24 $\bigcap_{n \in N^+} \boldsymbol{U^n TxtEx} - \boldsymbol{U^* TxtEx} \neq \emptyset$.

Proof. Let $\mathcal{L} = \{\{0\}, \{1\}, \{2\}, \ldots\} \cup \{N\}$. That for any $n \in N^+$, $\mathcal{L} \in \mathbf{U^nTxtEx}$ is witnessed by a learner which, on any sequence σ, conjectures representation index for $\{\{x\} \mid x \in \mathit{content}(\sigma)\}$ if $\mathit{card}(\mathit{content}(\sigma)) \leq n$, and conjectures representation index for $\{N\}$ otherwise. However, the unbounded finite unions of $\mathcal{L}$ is equivalent to the class $FIN \cup \{N\}$ which is not learnable [Gol67], and hence $\mathcal{L} \notin \mathbf{U^*TxtEx}$. ∎

Theorem 25 *Suppose $\mathcal{L}$ is a recursively enumerable class of disjoint languages. Then, for all $n > 2$, $\mathcal{L} \in$ **U**^{2}**TxtEx** iff $\mathcal{L} \in$ **DU**^{n}**TxtEx**.*

Proof. Note that $\emptyset \notin \mathcal{L}$. If $\mathcal{L}$ is finite, then clearly the theorem holds. So suppose $\mathcal{L}$ is infinite and f is a recursive function such that $\mathcal{L} = \{W_{f(i)} \mid i \in N\}$.

Suppose M witnesses that $\mathcal{L} \in \mathbf{U}^2\mathbf{TxtEx}$. Let $\sigma_0, \sigma_1, \ldots$ denote a recursive enumeration of all finite sequences. Let G be a recursive function such that, for all finite sets D, $W_{G(D)} = \bigcup_{i \in D} W_{f(i)}$. (For ease of notation, for any set $\{\ldots\}$ we write $G(\{\ldots\})$ as simply $G(\ldots)$.) Let P be a function such that $P(\sigma, i, t)$ is true iff *content*$(\sigma) \subseteq W_{i,t}$ and for all τ such that *content*$(\tau) \subseteq W_{i,t}$, $|\tau| \leq t - |\sigma|$, $M(\sigma) = M(\sigma\tau)$. Here note that $P(\sigma, i, t)$ can be effectively determined for $t \in N$ (for $t = \infty$, $P(\sigma, i, t)$ cannot be effectively determined in general; however if *content*(σ) is known to be subset of W_i, then whether $P(\sigma, i, \infty)$ is false can be determined in r.e. sense, that is, if $P(\sigma, i, \infty)$ is false, then one can find a witness to this effectively). We assume without loss of generality that for all i, m, $P(\#^m, f(i), \infty)$ is false (note that $\emptyset \notin \mathcal{L}$).

We now define M' which $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. M' defined may not be total. However, for any text T for a language $L \in \mathcal{L}^n$, it will be defined on all initial segments of T. On input $T[m]$, M' behaves as follows. It first finds a finite set X_m of cardinality at most n, and an $s_m \geq m$ such that

(a) *content*$(T[m]) \subseteq \bigcup_{i \in X_m} W_{f(i),s_m}$ but *content*$(T[m]) \not\subseteq \bigcup_{i \in X_m - \{j\}} W_{f(i),s_m}$ for all $j \in X_m$.

(b) For all $i \leq \max(X_m)$, $W_{f(i),s_m} \cap W_{f(j),s_m} \neq \emptyset$, for at most one $j \in X_m$;

(c) For $i \in X_m$, let r_i be least number such that $P(\sigma_{r_i}, f(i), s_m)$. Then, $P(\sigma_{r_i}, G(i,j), \infty)$ is false, for all $i, j \in X_m$, $i \neq j$. (Note that M **TxtEx**-identifies $\mathcal{L}^2$ and *content*$(\sigma_{r_i}) \subseteq W_{f(i)} \subseteq W_{G(i,j)}$, thus if $P(\sigma_{r_i}, G(i,j), \infty)$ is false, then one can eventually detect this.)

Then, M' outputs k such that $D_k = \{M(\sigma_{r_i}) \mid i \in X_m\}$.

We now show that M' $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. Fix D of cardinality at most n. Let $L = \bigcup_{i \in D} W_{f(i)}$, and T be a text for L. Here without loss of generality assume that $W_{f(i)} \neq W_{f(j)}$, for distinct $i, j \in D$.

We first claim that for all m, above process indeed finds some X_m satisfying (a), (b) and (c). This is so, since one could take $X_m = D$, with r_i to be such that $\sigma_{r_i} = \#^{m+1}$.

For $i \in D$, let r_i denote the minimal number such that σ_{r_i} is a stabilizing sequence for M on $W_{f(i)}$ (that is, $P(\sigma_{r_i}, f(i), \infty)$ is true, and $P(\sigma_j, f(i), \infty)$ is false for all $j < r_i$).

Now let t be a large enough number such that the following conditions hold:

(d) for all $i \in D$, $[\bigcup_{j \le r_i} content(\sigma_j) \cap W_{f(i)}] \subseteq W_{f(i),t} \cap content(T[t])$.

(e) for all $i \in D$, $j < r_i$, $P(\sigma_j, f(i), t)$ is false (thus, using (d), either $content(\sigma_j) \not\subseteq W_{f(i)}$ or there exists a τ of length at most $t - |\sigma_j|$ such that $content(\tau) \subseteq W_{f(i),t}$ and $M(\sigma_j) \neq M(\sigma_j \tau)$).

Now for all $m \ge t$, since every language in $\mathcal{L}$ is disjoint, by (d) above, for each $i \in D$, X_m does contain an i' such that $W_{f(i)} = W_{f(i')}$ (recall our assumption that σ_{r_i} is not of empty content). Also, for all $j \in X_m$, there must be a $j' \in D$ such that $W_{f(j)} = W_{f(j')}$ (since otherwise one could drop such j from X_m, see condition (a)). Furthermore, by (b) above, X_m does contain a unique i' such that $W_{f(i),t} \cap W_{f(i'),s_m} \neq \emptyset$. As $content(T[m])$ contains $W_{f(i),t}$, the i' above also satisfies $W_{f(i),t} \subseteq W_{f(i'),s_m}$. Thus, by (c) and (e) above, $r_{i'}$ as found on input $T[m]$, must be same as r_i.

Now suppose X_m also contains an i'' such that $i' \neq i''$ and $W_{f(i'')} = W_{f(i)}$. But then, $P(\sigma_{r_{i'}}, G(i', i''), \infty)$ is true, and thus (c) would not hold.

It follows that $\{\sigma_{r_{i'}} \mid i' \in X_m\} = \{\sigma_{r_i} \mid i \in D\}$, and thus, using **TxtEx**-identification of $\mathcal{L}$ by M, we have that M' converges on T to a representation index for $\{W_{f(i)} \mid i \in D\}$. Thus, M' $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. ∎

Theorem 25 shows that no recursively enumerable class of disjoint languages is in $\mathbf{U}^n\mathbf{TxtEx} - \mathbf{U}^{n+1}\mathbf{TxtEx}$ for any $n \ge 2$. (Note that the class of disjoint languages used in the proof of Theorem 20 is not recursively enumerable.) The following example (suggested by Frank Stephan) shows that the same is not true for the case of $n = 1$.

Example 26 Let

$$L_{x,0} = \{\langle x, 0\rangle\} \cup \{\langle x, y\rangle \mid (\forall z \in N \mid z \le y)[\varphi_x(z)\downarrow]\}$$

$$L_{x,y+1} = \begin{cases} \{\langle x, y+1\rangle\}, & \text{if } \langle x, y+1\rangle \not\in L_{x,0}; \\ L_{x,0}, & \text{otherwise.} \end{cases}$$

Let $\mathcal{L} = \{L_{x,i} \mid x, i \in N\}$. Clearly, $\mathcal{L} \in \mathbf{TxtEx}$. Now suppose there exists M such that $\mathcal{L}^2 \in \mathbf{TxtEx}(M)$, then $(\exists \sigma \in \mathrm{SEQ} \mid content(\sigma) \subseteq L_{x,0})$ $[(\forall \tau \in \mathrm{SEQ} \mid content(\tau) \subseteq \{\langle x, i\rangle \mid i \in N\})[M(\sigma) = M(\sigma\tau)]] \Leftrightarrow \varphi_x \in \mathcal{R}$. The condition on the left hand side is Σ_2 to check. However, the set $\{x \mid \varphi_x$ is recursive$\}$ is not Σ_2 (see [Rog67]), a contradiction. ∎

We note that the class of languages in Example 26 is not 1–1 recursively enumerable. As will be shown by Corollary 28, for a 1–1 recursively enumerable classes of languages, disjointness is a sufficient condition for the class to be in $\mathbf{DU^*TxtEx}$.

5 Sufficient Conditions For DUTxtEx Identification

In this section we consider some sufficient conditions for learning unions of languages.

5.1 Functions That Enumerate Distinguishing Elements

Let recursively enumerable $\mathcal{L} \subseteq \mathcal{E}$ be given. Suppose for all $L \in \mathcal{L}$, there is an effective procedure to enumerate an element which is uniquely in L, that is, no other language in $\mathcal{L}$ contains this element. Can we then identify the union of every finite collection of languages drawn from $\mathcal{L}$? An answer is attempted in the following Theorem.

Theorem 27 *Let $\mathcal{L}$ be a 1–1 recursively enumerable class of languages as witnessed by the computable numbering ψ. If there exists a limiting recursive function d and total recursive F for which $d(i) = \lim_{t\to\infty} F(i,t)$ such that*

(a) for all $i \in N$, $d(i) \in W_i^{\psi}$,

(b) for all $i, j \in N$, $d(i) \in W_j^{\psi} \Rightarrow i = j$, and

(c) for all $j \in N$, $card(range(F) \cap W_j^{\psi}) < \infty$.

*Then $\mathcal{L} \in$ **DU* TxtEx**.*

Proof. Let $\mathcal{L}$, ψ, F, d be as in Theorem. Let recursive function h witness that $\psi \preceq \varphi$. Unless stated otherwise, we let i, j, with or without decorations, range over N. Define M as follows, such that for each text T and for each $m \in N$,

$M(T[m])$:
    $S \leftarrow \emptyset$.
    For $i = 0$ to $m-1$ do
        If $[F(i,m) \in content(T[m]) \cap W^{\psi}_{i,m}]$
            and $\neg[(\exists i', j')[i' < m \ \wedge\ j' < m \wedge i' \neq j'$
                $\wedge\ F(i,m) \in W^{\psi}_{i',m} \cap W^{\psi}_{j',m}]]$
            Then let $S \leftarrow S \cup \{h(i)\}$.
    EndFor

Output S.
End

We claim that M $\mathbf{DU^*TxtEx}$-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$ be a finite collection of languages. Let D be such that $\{W_i^\psi \mid i \in D\} = \mathcal{L}'$. Let T be a text for $L = \bigcup_{i\in D} W_i^\psi$, and let $A = \text{range}(F) \cap L$. By clause (c) in the Theorem, $card(A) < \infty$. Intuitively, A contains all the potential "distinguishing element"s M will encounter during the identification process. Since D and A are finite, there exists $n \in N$ so large that

(1) For all $i \in D$, $(\forall t \in N, t > n)$ $[F(i,t) = d(i) \wedge d(i) \in content(T[t]) \cap W_{i,t}^\psi]$.

(2) For all $x \in A - \{d(k) \mid k \in D\}$, $(\forall n' \in N, n' > n)$ $[(\exists j \in N - D)[x \in W_{j,n'}^\psi]$ $\Rightarrow (\exists i', j' < n)$ $[i' \neq j' \wedge x \in W_{i',n'}^\psi \cap W_{j',n'}^\psi]]$.

Clause (1) ensures that all $i \in D$ will eventually be output by M. Clause (2) ensures that all grammars $j \notin D$, which enumerate some element in A are excluded from consideration (note that every element in A is enumerated by some grammar in D).

Hence for all $n' \in N$, where $n' > n$, $i \in D$ if and only if $i \in S$ output by $M(T[n'])$. It follows that M $\mathbf{DU^*TxtEx}$-identifies $\mathcal{L}$. ▮

Corollary 28 *Let $\mathcal{L}$ be a class of languages for which there exists a 1–1 numbering and that*

(a) $\emptyset \notin \mathcal{L}$, and

(b) for all $L, L' \in \mathcal{L}$, $L \neq L' \Rightarrow L \cap L' = \emptyset$.

*Then $\mathcal{L} \in$ **DU* TxtEx**.*

Proof. Let $\mathcal{L} = \{W_i^\psi \mid i \in N\}$, where ψ is a 1–1 numbering for $\mathcal{L}$. For $i, t \in N$, let $F(i,t) = \min(W_{i,t}^\psi)$ and let $d(i) = \lim_{t\to\infty} F(i,t)$. Clearly, (a) for all $i \in N$, $d(i) \in W_i^\psi$, (b) for all $i, j \in N$, $d(i) \in W_j^\psi \Rightarrow i = j$, and (c) for all $j \in N$, $card(range(F) \cap W_j^\psi) < \infty$. Thus d fulfills all the conditions for Theorem 27. It follows that $\mathcal{L} \in \mathbf{DU^*TxtEx}$. ▮

Corollary 29 *Let $\mathcal{L}$ be an indexed family of recursive languages such that*

(a) $\emptyset \notin \mathcal{L}$, and

(b) for all $L, L' \in \mathcal{L}$, $L \neq L' \Rightarrow L \cap L' = \emptyset$.

*Then $\mathcal{L} \in$ **DU* TxtEx**.*

In Theorem 27, some weaker conditions for (a) and (b) may not be sufficient, even if we require d to be recursive. For instance, if we have only the following conditions (where the requirement (b) is relaxed):

(a) for all $i \in N$, $d(i) \in W_i^\psi$, and

(b) for all $i \in N$, $card(\{j \in N \mid d(i) \in W_j^\psi\}) < \infty$,

(c) d is recursive.

Then identifiability for $\mathcal{L}^2$ cannot be guaranteed, as the following example shows.

Example 30 For $i \in N$, let

$$L_0 = \{\langle 0, x\rangle \mid x \in N\} \cup \{\langle 1, x\rangle \mid x \in K\}$$

$$L_{i+1} = \begin{cases} \{\langle 0, i+1\rangle\} \cup \{\langle 1, i\rangle\} \cup \{\langle 2, i\rangle\}, & \text{if } i \in K; \\ \{\langle 0, i+1\rangle\} \cup \{\langle 1, i\rangle\}, & \text{otherwise.} \end{cases}$$

Let $\mathcal{L} = \{L_i \mid i \in N\}$ and define d such that for all $x \in N$, $d(x) = \langle 0, x\rangle$. It is easy to verify that (a) $\mathcal{L}$ is 1–1 recursively enumerable, (b) $\mathcal{L} \in \mathbf{TxtEx}$, (c) $\mathcal{L}^2$ is uniquely definable from $\mathcal{L}$, and (d) d satisfies all the conditions given above for $\mathcal{L}$. However, for all $k \in N$, the language $\{\langle 0, i\rangle \mid i \in N\} \cup \{\langle 1, x\rangle \mid x \in K \cup \{k\}\}$ is in $\mathcal{L}^2$, hence (using Proposition 4.7 in [JORS99]) $\mathcal{L}^2$ is not in **TxtEx**. ∎

A similar weakening of these conditions, where instead of a single unique element d is required to name only a set of elements which is unique to each language in the class, as in the following:

(a) for all $i \in N$, $D_{d(i)} \subseteq W_i^\psi$,

(b) for all $i, j \in N$, $D_{d(i)} \subseteq W_j^\psi \Rightarrow i = j$.

(c) d is recursive.

then such a function will also fail to guarantee that $\mathcal{L}^2 \in \mathbf{TxtEx}$, as demonstrated by the following example.

Example 31 For $i \in N$, let

$$L_0 = \{\langle 0, 0\rangle\} \cup \{\langle 1, x\rangle \mid x \in N\}$$

$$L_1 = \{\langle 1,1\rangle\} \cup \{\langle 0,x\rangle \mid x \in N\} \cup \{\langle 2,x\rangle \mid x \in K\}$$

$$L_{i+2} = \begin{cases} \{\langle 0,i+2\rangle\} \cup \{\langle 1,x\rangle \mid x \in N\} \cup \{\langle 2,i\rangle\} \cup \{\langle 3,i\rangle\}, & \text{if } i \in K; \\ \{\langle 0,i+2\rangle\} \cup \{\langle 1,x\rangle \mid x \in N\} \cup \{\langle 2,i\rangle\}, & \text{otherwise.} \end{cases}$$

Let $\mathcal{L} = \{L_i \mid i \in N\}$ and define d such that for all $x \in N$, $d(x) = \{\langle 0,x\rangle, \langle 1,x\rangle\}$. It is easy to verify that $\mathcal{L}$ is a 1–1 recursively enumerable class of languages in **TxtEx** where all the languages in $\mathcal{L}^2$ are uniquely definable from $\mathcal{L}$, and that d satisfies all the conditions above for $\mathcal{L}$. However, for all $k \in N$, the language $\{\langle 0,i\rangle, \langle 1,i\rangle \mid i \in N\} \cup \{\langle 2,x\rangle \mid x \in K \cup \{k\}\}$ is in $\mathcal{L}^2$, thus (using Proposition 4.7 in [JORS99]) $\mathcal{L}^2$ is not in **TxtEx**. ∎

In contrast to Theorem 27, the following characterizes $\mathbf{DU^*TxtEx}$ learning for the indexed families.

Proposition 32 *An indexed family $\mathcal{L}$ is in* $\boldsymbol{DU^*TxtEx}$ *iff for each language $L \in \mathcal{L}$, for every finite $\mathcal{L}' \subseteq \mathcal{L}$, $L \subseteq L_{\mathcal{L}'}$ iff $L \in \mathcal{L}'$.*

Proof. ($\Rightarrow$) Suppose by way of contradiction that $L \in \mathcal{L}$ and $\mathcal{L}' \subseteq \mathcal{L}$ is finite and $L \subseteq L_{\mathcal{L}'}$ but $L \notin \mathcal{L}'$. Thus, $\{L\} \cup \mathcal{L}' \neq \mathcal{L}'$ but $L_{\{L\}\cup\mathcal{L}'} = L_{\mathcal{L}'}$. By an argument similar to that for Proposition 10, $\mathcal{L} \notin \mathbf{DU^*TxtEx}$.
($\Leftarrow$) Let $L_0, L_1, \ldots$ be a 1–1 enumeration of $\mathcal{L}$ (there exists such enumeration for all indexed families). Let $D_0, D_1, \ldots$ be a 1–1 enumeration of all the finite subsets of N such that $D_i \subseteq D_j$ implies $i \leq j$ (the definition of $D_0, D_1, \ldots$ in [Rog67] fulfills this requirement). Let $\mathcal{L}_i = \{L_j \mid j \in D_i\}$. Note that by hypothesis, $L_{\mathcal{L}_i} \nsubseteq L_{\mathcal{L}_j}$, for $j < i$. Let M be defined as follows: $M(T[n]) = D_i$, where $i = \min(\{j \mid \mathit{content}(T[n]) \subseteq L_{\mathcal{L}_j}\})$ (that is, the smallest j where $L_{\mathcal{L}_j}$ is consistent with the input). To see that M $\mathbf{DU^*TxtEx}$-learns $\mathcal{L}$, suppose T is a text for $L_{\mathcal{L}_i}$. For each $j < i$, let x_{ij} be an element in $L_{\mathcal{L}_i} - L_{\mathcal{L}_j}$ (such x_{ij} exists because of the conditions required of $\mathcal{L}$ and $D_0, D_1, \ldots$). Let n be such that $\{x_{ij} \mid j < i\} \subseteq \mathit{content}(T[n])$. It follows that $M(T[m]) = D_i$, for $m \geq n$. ∎

It is worth noting that the condition in Proposition 32 has similarities with the notion of *compactness with respect to containment* discussed in [ASO94].

From Proposition 32, we see that if each language in a given indexed family contains even just one unique word, then the class would be $\mathbf{DU^*TxtEx}$-learnable. Regrettably, none of the indexed families studied in [SA00] fulfill this condition. In some sense, this suggests that the requirement of $\mathbf{DU^*TxtEx}$ may be too restrictive for the learning of the *pattern languages*, especially when we put this in contrast to the learnability results for **WDU*TxtEx** (Theorem 18 and 19).

5.2 Restrictions On Structures Of Languages

Theorem 33 *Given $n \in N^+$. Let $\mathcal{L}$ be a class of languages such that*

(a) every language in $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$,

(b) for all $L \in \mathcal{L}$, card$(\{L' \in \mathcal{L} \mid L' \cap L \neq \emptyset\}) < \infty$,

(c) there exists a computable numbering ψ for $\mathcal{L}$ such that:
(c.1) for all $L \in \mathcal{L}$, card$(L) = \infty \Rightarrow$ card$(\{i \mid W_i^\psi = L\}) = 1$;
(c.2) for all $L \in \mathcal{L}$, card$(L) < \infty \Rightarrow$ card$(\{i \mid W_i^\psi = L\}) < \infty$.

*Then $\mathcal{L} \in$ **DU^n TxtEx**.*

Proof. Let $n \in N^+$ be given. Let $\mathcal{L}$ be as in Theorem. Unless stated otherwise, we let i, j, k, m, n, with or without decorations, range over N. We let A and B, with or without decorations, range over FIN. Let h witness that $\psi \preceq \varphi$. Define **IIM** M as follows such that for each text T,

$M(T[m])$:
  Let $\mathcal{C}^m = \{i \mid i \leq m \ \wedge \ W_{i,m}^\psi \cap content(T[m]) \neq \emptyset\}$.
  Let $Candidates^m = \{S \subseteq \mathcal{C}^m \mid card(S) \leq n\}$.
  Let $s_0 = \max(\{s \in N \mid (\exists S \in Candidates^m)[\ \bigcup_{i \in S} W_{i,s}^\psi \subseteq content(T[m]) \wedge \bigcup_{i \in S} W_{i,m}^\psi \supseteq content(T[s])]\})$.
  Output $\{h(i) \mid i \in D_{k_0}\}$, where $k_0 = \min(\{k \mid D_k \in Candidates^m \wedge \bigcup_{i \in D_k} W_{i,s_0}^\psi \subseteq content(T[m]) \wedge \bigcup_{i \in D_k} W_{i,m}^\psi \supseteq content(T[s_0])\})$.
End

Intuitively, M outputs the seemingly best grammar set in $Candidates^m$ which describes the input text. We claim that M **DU^{n}TxtEx**-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq n$. Let B be such that $card(B) \leq n$, and $\{W_i^\psi \mid i \in B\} = \mathcal{L}'$. Let T be a text for $L = \bigcup_{i \in B} W_i^\psi$. We divide B into two groups, $B_1 = \{i \in B \mid W_i^\psi \text{ is finite}\}$ and $B_2 = \{i \in B \mid W_i^\psi \text{ is infinite}\}$. By the requirement of ψ, for each $i \in B_1$, there exist only finitely many j such that $W_i^\psi = W_j^\psi$, and for each $i \in B_2$, for all $j \neq i$, $W_i^\psi \neq W_j^\psi$. Let $\mathcal{A} = \{A \mid \bigcup_{i \in A} W_i^\psi = \bigcup_{i \in B_1} W_i^\psi\}$. Since $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$, the only sets of languages which are capable of generating L are $\{B_2 \cup A \mid A \in \mathcal{A}\}$. Let $CorrectInd = \{B_2 \cup A \mid A \in \mathcal{A}\}$.

Let $\mathcal{C}' = \{i \mid W_i^\psi \cap content(T) \neq \emptyset\}$. Since each language in $\{W_i^\psi \mid i \in B\}$ intersects with only finitely many other languages in $\mathcal{L}$, $\mathcal{C}'$ is finite. It is easy to verify that there exists n_0 such that for all $n' > n_0$, $\mathcal{C}^{n'} = \mathcal{C}'$.

Let $Candidates' = Candidates^{n_0}$. Clearly, for all $n' > n_0$, $Candidates^{n'} = Candidates'$.

Let $n_1 > n_0$ be so large that

$$(\forall i \in \mathcal{C}')[card(W_i^\psi) < \infty \Rightarrow [W_i^\psi = W_{i,n_1}^\psi \wedge W_i^\psi \cap content(T[n_1]) = W_i^\psi \cap content(T)]]$$

Let $n_2 > n_1$ be so large that

$$\neg[(\exists B' \in Candidates' - CorrectInd)[(\bigcup_{i \in B'} W_{i,n_2}^\psi \subseteq L) \wedge (\bigcup_{i \in B'} W_i^\psi \supseteq content(T[n_2]))]]$$

Let $n_3 > n_2$ be so large that

$$[\bigcup_{i \in B} W_{i,n_2+1}^\psi \subseteq content(T[n_3]) \wedge \bigcup_{i \in B} W_{i,n_3}^\psi \supseteq content(T[n_2+1])]$$

Clearly, for all $n' > n_3$, $\{D \in Candidates' \mid \bigcup_{i \in D} W_{i,n_2+1}^\psi \subseteq content(T[n'])$ $\wedge \bigcup_{i \in D} W_{i,n'}^\psi \supseteq content(T[n_2+1])\} = CorrectInd$. Hence for all $n' > n_3$, M outputs $\{h(i) \mid i \in D_{k_0}\}$, for $k_0 = \min(\{k \mid D_k \in CorrectInd\})$. It follows that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. ∎

Corollary 34 *Fix $n \in N^+$. Let $\mathcal{L} = \{L_i \mid i \in N\}$ be a 1–1 recursively enumerable class of languages where*

(a) every language in $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$.

(b) for all $i \in N$, $card(\{j \mid L_i \cap L_j \neq \emptyset\}) < \infty$.

*Then $\mathcal{L} \in$ **DU^n TxtEx**.*

The conditions in Theorem 33 are not necessary — this can be shown using $TRANSIM_n$ (see Section 6.1), which is 1–1 recursively enumerable but every language in the class intersects with infinitely many other languages within the class.

The following characterizes $\mathbf{DU}^n\mathbf{TxtEx}$ learning for the indexed families.

Proposition 35 *For any $n \in N$, an indexed family $\mathcal{L}$ is in **DU^n TxtEx** iff* *(a) every language in $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$, and*

*(b) $\mathcal{L} \in$ **U^nTxtEx**.*

Proof. ($\Rightarrow$) follows from definition. We show ($\Leftarrow$). Let $L_0, L_1, \ldots$ be a 1–1 enumeration of $\mathcal{L}$. Let $X_0, X_1, \ldots$ be a 1–1 enumeration of all the finite subsets of N of size at most n, such that $X_i \subseteq X_j$ implies $i \leq j$ (the definition of $D_0, D_1, \ldots$ in [Rog67] can be easily adapted to fulfill this requirement). Let $\mathcal{L}_i = \{L_j \mid j \in X_i\}$. Let $L'_i = L_{\mathcal{L}_i}$. Now, $\mathcal{L}^n = \{L'_0, L'_1, \ldots\}$ is an indexed family in **TxtEx** (by hypothesis (b)). Thus, there is a learner which on a text for any language L'_i, outputs in the limit index i (see the proof of learning the indexed families via the finite tell-tales in [Ang80]). From i we can obtain the constituent languages in L'_i. By hypothesis (a), these languages are the only possible sets of languages in $\mathcal{L}$ that unions to L'_i, and hence $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$. ∎

Here it is interesting to mention that Wright showed that classes of indexed families which have finite elasticity (such as pattern languages, see [Wri89,MSW91] for details), belong to $\mathbf{U}^n\mathbf{TxtEx}$ for all n.

6 Natural Class that Witnesses the Hierarchies of UTxtEx and DUTxtEx

In this section we give two natural classes of languages which give rise to our hierarchy results. We first describe an indexed family that give rise to the **DUTxtEx** hierarchy.

6.1 The Class $TRANSIM_n$

Let RAT_n be the set of all the points in an $(n-1)$-dimensional space with only rational valued coordinates. Let $\mathrm{coderat}_n(\cdot)$ be an effective bijective mapping from RAT_n to N. Fix $n \in N^+$, $n \geq 2$. Let $v_1, v_2, \ldots, v_{n-1}$ be unit vectors along each axis of an $(n-1)$-dimensional space. Let O denote the origin. Let $\Gamma_n = \{\sum_{i=1}^{n-1} \epsilon_i v_i \mid \epsilon_i \in rat\}$.

For each simplex [Cox63] G, let $V(G)$ denote the vertices of G, and $P(G)$ denote the set of points in the simplex G. For $(n-1)$-dimensional simplex G, points X in $P(G)$ satisfy n linear equations $\overrightarrow{\nu_k} \cdot X \leq b_k$, $k = 1, 2, \ldots, n$ where for each k, the coefficient b_k and the vector $\overrightarrow{\nu_k}$ can be obtained by solving $n-1$ linear equations, each formed by substituting in the equation a vertex of G [Cox63]. Intuitively, the inequality for each k represents a bounding hyperplane for the polytope, where each vector $\overrightarrow{\nu_k}$ is the outward normal for the bounding hyperplane. For any vector Γ, we let $G+\Gamma$, denote the simplex formed by translating each point in G by Γ.

For a simplex G in $(n-1)$-dimensional space, let $Lang(G) = \{\mathrm{coderat}_n(X) \mid X \in P(G) \cap RAT_n\}$. For Λ a set of simplexes, let $\mathcal{L}(\Lambda) = \{Lang(G) \mid G \in \Lambda\}$.

We fix a simplex $\mathbb{G}_n$ with n vertices in $(n-1)$-dimensional space, with vertices at $O, v_1, v_2, \ldots, v_{n-1}$. (For $n=2$, the vertices are at O and v_1.) Let $\mathbb{A}_n = \{\mathbb{G}_n + \Gamma \mid \Gamma \in \Gamma_n\}$. Let $TRANSIM_n = \mathcal{L}(\mathbb{A}_n) = \{Lang(G) \mid G \in \mathbb{A}_n\}$.

We now give some properties of $\mathbb{A}_n$ (and hence $TRANSIM_n$) which we shall use to demonstrate our hierarchy results.

Lemma 36 *Let $G_i = \mathbb{G}_n + \epsilon v_i$. Suppose $\epsilon < 1/n$, and S is the simplex with vertices at O and $1+\epsilon v_i$, for $1 \le i \le n-1$. Then, $P(S) = P(\mathbb{G}_n) \cup \bigcup_{1\le i\le n-1} P(G_i)$.*

Proof. The points $X = (x_1, x_2, \ldots, x_{n-1})$ in $P(\mathbb{G}_n)$ satisfy the equations:

(E1) $x_j \ge 0$, for $1 \le j \le n-1$, and

(E2) $\sum_{j=1}^{n-1} x_j \le 1$.

The points $X = (x_1, x_2, \ldots, x_{n-1})$ in G_i satisfy the equations:

(E3.1.i) $x_j \ge 0$ (for $1 \le j \le n-1$, $i \ne j$),

(E3.2.i) $x_i \ge \epsilon$, and

(E4) $\sum_{j=1}^{n-1} x_j \le 1+\epsilon$.

The points $X = (x_1, x_2, \ldots, x_{n-1})$ in S satisfy the equations:

(E5) $x_j \ge 0$ (for $1 \le j \le n-1$), and

(E6) $\sum_{j=1}^{n-1} x_j \le 1+\epsilon$.

Note that any point X satisfying (E1) and (E2) also satisfies (E5) and (E6). Similarly, any point X satisfying (E3.1.i), (E3.2.i) and (E4) also satisfies (E5) and (E6). Thus, $P(S) \supseteq P(\mathbb{G}_n) \cup \bigcup_{1\le i\le n-1} P(G_i)$.

Now suppose $X = (x_1, \ldots, x_{n-1}) \in P(S)$. Thus X satisfies (E5) and (E6). If X additionally satisfies (E2), then clearly, $X \in P(G_0)$. If not, then there must exist an x_i, such that $x_i \ge 1/(n-1) > 1/n$. Thus, X satisfies, (E3.1.i), (E3.2.i) and (E4), and thus $X \in P(G_i)$. $\square$

Lemma 37 *Let $n \geq 2$. There exist distinct simplexes $G_0, G_1, \ldots, G_n \in \mathbb{A}_n$, and a constant $\xi > 0$, $\xi \in rat$ such that: $(\forall \delta \in rat \mid \delta \leq \xi)[P(G_0 + \delta v_1) \subseteq \bigcup_{j=1}^{n} P(G_j)]$.*

Proof. Let $G_0 = \mathbb{G}_n + (\frac{1}{n+2})v_1$. For $1 \leq i \leq n-1$, let $G_i = \mathbb{G}_n + (\frac{1}{n+1})v_i$. Let $G_n = \mathbb{G}_n$. Let $\xi = \frac{1}{n+1} - \frac{1}{n+2}$. Now the Lemma follows by using Lemma 36. ∎

Lemma 38 *Let $\Lambda, \Lambda' \subset \mathbb{A}_n$ where $\max(\{card(\Lambda), card(\Lambda')\}) \leq n$. Then*

(a) $\bigcup_{G \in \Lambda} V(G) \subseteq \bigcup_{G \in \Lambda'} P(G)$, and

(b) $\bigcup_{G \in \Lambda'} V(G) \subseteq \bigcup_{G \in \Lambda} P(G)$

if and only if $\Lambda = \Lambda'$.

Proof. We first show the following two claims.

Claim 39 *Suppose $G = \mathbb{G}_n + \Gamma$, where $\Gamma = \sum_{i=1}^{n-1} a_i v_i$, and $a_i \geq 0$. The vertices in $V(G)$ are thus: $A_0 = (a_1, a_2, \ldots, a_{n-1})$ and, for $1 \leq i \leq n-1$, $A_i = (r_{i,1}, r_{i,2}, \ldots, r_{i,n-1})$, where $r_{i,i} = a_i + 1$, and $r_{i,j} = a_j$, for $i \neq j$.*

Then, for any $\Gamma' = \sum_{i=1}^{n-1} a'_i v_i$, where $a'_i \geq 0$, (a) and (b) hold, where $G' = \mathbb{G}_n + \Gamma'$:

(a) If $A_0 \in P(G')$, then

> *(a.1) for $1 \leq k \leq n-1$, $a_k \geq a'_k$, and*
>
> *(a.2) $\sum_{k=1}^{n-1} a_k \leq 1 + \sum_{k=1}^{n-1} a'_k$.*

(b) For $1 \leq i \leq n-1$, if $A_i \in P(G')$, then

> *(b.1) $a_k \geq a'_k$, for $1 \leq k \leq n-1, k \neq i$, and*
>
> *(b.2) $a_i + 1 \geq a'_i$, and*
>
> *(b.3) $\sum_{k=1}^{n-1} a_k \leq \sum_{k=1}^{n-1} a'_k$.*

Proof. Follows by noting that points $X = (x_1, \ldots, x_{n-1})$ in $P(G')$ must satisfy the equations:

$x_k \geq a'_k$, for $1 \leq k \leq n-1$, and

$\sum_{k=1}^{n-1} x_k \leq 1 + \sum_{k=1}^{n-1} a'_k$. □

Claim 40 *Suppose $G = \mathbb{G}_n + \Gamma$ and $G' = \mathbb{G}_n + \Gamma'$, where $\Gamma, \Gamma' \in \Gamma_n$. If $\Gamma \neq \Gamma'$, then $card(P(G') \cap V(G)) \leq 1$.*

Proof. Suppose $\Gamma = \sum_{i=1}^{n-1} a_i v_i$, and $\Gamma' = \sum_{i=1}^{n-1} a'_i v_i$. Thus vertices of G are: $A_0 = (a_1, a_2, \ldots, a_{n-1})$ and, for $1 \leq i \leq n-1$, $A_i = (r_{i,1}, r_{i,2}, \ldots, r_{i,n-1})$, where $r_{i,i} = a_i + 1$, and $r_{i,j} = a_j$, for $i \neq j$.

Now if $P(G')$ contains A_0 and A_i, for some i, $1 \leq i \leq n-1$, then by Claim 39(a.1) and (b.3), we have that $a_k = a'_k$, for $1 \leq k \leq n-1$, and thus $\Gamma = \Gamma'$.

If $P(G')$ contains A_i and A_j, for some distinct i, j, $1 \leq i < j \leq n-1$, then by using Claim 39 (b.1) (with values i and j for i as in Claim 39(b)), as well as using Claim 39(b.3), we get $a_k = a'_k$, for $1 \leq k \leq n-1$, and thus $\Gamma = \Gamma'$.

Claim follows. $\square$

We now prove Lemma 38. Now suppose $\Lambda, \Lambda' \subseteq \mathbb{A}_n$ such that $\max(card(\Lambda), card(\Lambda')) \leq n$. Suppose $\Lambda \neq \Lambda'$. Then there exists a $G \in \mathbb{A}_n$ which belongs to $(\Lambda - \Lambda') \cup (\Lambda' - \Lambda)$. Without loss of generality suppose $G \in \Lambda - \Lambda'$. Then, by Claim 40, each element of Λ' can contain at most one vertex of G. Thus, Λ' must contain exactly n simplexes, each containing one vertex of G.

Suppose $\Gamma = \sum_{i=1}^{n-1} a_i v_i$ is such that $G = \mathbb{G}_n + \Gamma$. Thus vertices of G are: $A_0 = (a_1, a_2, \ldots, a_{n-1})$ and, for $1 \leq i \leq n-1$, $A_i = (r_{i,1}, r_{i,2}, \ldots, r_{i,n-1})$, where $r_{i,i} = a_i + 1$, and $r_{i,j} = a_j$, for $i \neq j$.

For $0 \leq i \leq n-1$, suppose G_i is the simplex in Λ' which includes A_i. Suppose $G_i = \mathbb{G}_n + \Gamma_i$, where $\Gamma_i = \sum_{k=1}^{n-1} b^i_k v_k$.

Let $B_0 = (b^0_1, b^0_2, \ldots, b^0_{n-1})$. For $1 \leq i \leq n-1$, define $B_i = (w_{i,1}, w_{i,2}, \ldots, w_{i,n-1})$, where $w_{i,j} = b^i_j$, for $i \neq j$, and $w_{i,i} = b^i_i + 1$. Note that B_i is a vertex of G_i.

We claim that (C1) and (C2) below hold.

(C1) For $0 \leq i \leq n-1$, $B_i \notin P(G)$.

(C2) For any $G' \in \mathbb{A}_n$, $P(G')$ contains at most one of B_i, $0 \leq i \leq n-1$.

To see (C1), note that if $B_0 \in P(G)$, then by using the assumption $A_0 \in P(G_0)$ and Claim 39 (a.1), we get $a_k \leq b^0_k \leq a_k$, for $1 \leq k \leq n-1$. Thus $a_k = b^0_k$, $1 \leq k \leq n-1$, and $G = G_0$. On the otherhand if for some i, $1 \leq i \leq n-1$, $B_i \in P(G)$, then by using the assumption $A_i \in P(G_i)$ and Claim 39 (b.1) and (b.3), we get (I) $a_k \leq b^i_k \leq a_k$, for $1 \leq k \leq n-1$, $k \neq i$, and (II) $\sum_{k=1}^{n-1} a_k \leq \sum_{k=1}^{n-1} b^i_k \leq \sum_{k=1}^{n-1} a_k$. Thus $a_k = b^i_k$, and $G = G_i$. However, for $0 \leq i \leq n-1$, $G_i \neq G$. Thus we immediately get (C1).

For (C2) suppose by way of contradiction that G' contains B_i and B_j, where $0 \leq i < j \leq n-1$. Let $\Gamma' = \sum_{k=1}^{n-1} a'_k v_k$ be such that $G' = \mathbb{G}_n + \Gamma'$.

Case 1: $i = 0$.

In this case we have, $a'_k \leq b^0_k \leq a_k$, for $1 \leq k \leq n-1$, (as $A_0 \in P(G_0)$, and $B_0 \in P(G')$, and using (a.1) of Claim 39). Furthermore, $\sum_{k=1}^{n-1} a_k \leq \sum_{k=1}^{n-1} b^j_k \leq \sum_{k=1}^{n-1} a'_k$ (as $A_j \in P(G_j)$, and $B_j \in P(G')$, and using (b.3) of Claim 39). It follows that $a'_k = a_k$, for $1 \leq k \leq n-1$, and thus $G' = G$. A contradiction to (C1).

Case 2: i, j are not 0.

In this case we have, $a'_k \leq b^i_k \leq a_k$, for $1 \leq k \leq n-1$, $i \neq k$; $a'_k \leq b^j_k \leq a_k$, for $1 \leq k \leq n-1$, $j \neq k$ (as $A_i \in P(G_i)$ and $B_i \in P(G')$, $A_j \in P(G_j)$ and $B_j \in P(G')$, using (b.1) of Claim 39). Furthermore, $\sum_{k=1}^{n-1} a_k \leq \sum_{k=1}^{n-1} b^j_k \leq \sum_{k=1}^{n-1} a'_k$ (as $A_j \in P(G_j)$, and $B_j \in P(G')$, and using (b.3) of Claim 39). It follows that $a'_k = a_k$, for $1 \leq k \leq n-1$, and thus $G' = G$. A contradiction to (C1).

It follows that (C2) holds. As Λ contains at most n simplexes, (C1) and (C2) imply that $\bigcup_{S \in \Lambda'} V(S) \not\subseteq \bigcup_{S \in \Lambda} P(S)$, contradicting the hypothesis of Lemma. Thus, we must have $\Lambda = \Lambda'$. ∎

Proposition 41 *For $n \in N^+$, $n \geq 2$, (a) $TRANSIM_n \in$ **DU^nTxtEx**.*

*(b) $TRANSIM_n \notin$ **DU^{n+1}TxtEx**.*

Proof. (a) $TRANSIM_n \in$ **DU**^{n}**TxtEx** is seen by M below, where for each text T and each $m \in N$,

$M(T[m])$:
  Let $S^m \leftarrow \{A \mid \mathrm{coderat}_n(A) \in content(T[m])\}$.
      (This step converts input into coordinates.)
  If there exists a collection $\Lambda^m \subset \mathbb{A}_n$ of at most n simplexes such that
    (A) $\bigcup_{G \in \Lambda^m} V(G) \subseteq S^m$, and
    (B) $S^m \subseteq \bigcup_{G \in \Lambda^m} P(G)$.
      (Note that since S^m is finite, this check is recursive.)
  Then, pick (lexicographically least) such Λ^m and output a
    (standard) representation index for
    $\mathcal{L}(\Lambda^m) = \{Lang(G) \mid G \in \Lambda^m\}$.
  Otherwise, output 0.
End

To see that M **DU**^{n}**TxtEx**-identifies $TRANSIM_n$, let $\Lambda' \subset \mathbb{A}_n$ be any collection of at most n simplexes from $\mathbb{A}_n$. Suppose T is a text for $\mathcal{L}(\Lambda')$, and consider the outputs of M on $T[m]$. By Lemma 38, it is easy to verify that for any m such that all of $\bigcup_{G \in \Lambda'} V(G)$ has appeared in S^m, the only set of (at most n) simplexes (from $\mathbb{A}_n$) that can fulfill conditions (A) and (B) in the definition of M is Λ'. Thus, for all but finitely many m, $\Lambda^m = \Lambda'$. Thus M, given a text for $L_{\mathcal{L}(\Lambda')}$, outputs a representation index for $\mathcal{L}(\Lambda')$ in the limit. Thus, $TRANSIM_n \in \mathbf{DU}^n\mathbf{TxtEx}$.

(b) By Lemma 37, there exist distinct $G_0, G_1, \ldots, G_n \in \mathbb{A}_n$ such that $P(G_0) \subseteq \bigcup_{1\leq i\leq n} P(G_i)$. Let $\Lambda = \{G_1, G_2, \ldots, G_n\}$ and $\Lambda' = \{G_0, G_1, G_2, \ldots, G_n\}$. Now $\mathcal{L}(\Lambda) \neq \mathcal{L}(\Lambda')$ but $L_{\mathcal{L}(\Lambda)} = L_{\mathcal{L}(\Lambda')}$. Hence by Proposition 10, $TRANSIM_n \notin \mathbf{DU}^{n+1}\mathbf{TxtEx}$. ∎

6.2 The Class $ExtTRANSIM_n$

We now define the class $ExtTRANSIM_n$ based on $TRANSIM_n$, which witnesses that $\mathbf{DU}^n\mathbf{TxtEx} - \mathbf{U}^{n+1}\mathbf{TxtEx} \neq \emptyset$.

Let $PRIMES$ be the set of all the prime numbers and $p_0, p_1, \ldots$ be an enumeration of $PRIMES$ in ascending order. Let ψ be a computable numbering for which $(\forall i \in N)[W^{\psi}_{p_i} = W_i]$.

For $a \in rat$, let $h(a)$ denote the denominator of a in reduced form. Clearly, h is a recursive function. For $G \in \mathbb{A}_n$, suppose $\Gamma \in \Gamma_n$ is such that $G = \mathbb{G}_n + \Gamma$. Then, let $X_1(G) = \Gamma \cdot v_1$, and $L_G = \{\langle 0, x\rangle \mid x \in Lang(G)\} \cup \{\langle 1, y\rangle \mid y \in W^{\psi}_{h(X_1(G))}\}$. Finally, $ExtTRANSIM_n = \{L_G \mid G \in \mathbb{A}_n\}$.

Theorem 42 $(\forall n \in N^+, n \geq 2)[ExtTRANSIM_n \in \boldsymbol{DU^n TxtEx} - \boldsymbol{U^{n+1} TxtEx}]$.

Proof. To see that $ExtTRANSIM_n \in \mathbf{DU}^n\mathbf{TxtEx}$, consider each $L \in ExtTRANSIM^n_n$ as consisting of two parts, $A = \{x \mid \langle 0, x\rangle \in L\}$ and $B = \{x \mid \langle 1, x\rangle \in L\}$. Now, $L = L_{\{L_G \mid G\in\Lambda\}}$ iff $A = L_{\{Lang(G)\mid G\in\Lambda\}}$. Furthermore, a grammar for L_G can be obtained effectively from a grammar for $Lang(G)$. Thus, using Proposition 41(a), it follows that $ExtTRANSIM_n \in \mathbf{DU}^n\mathbf{TxtEx}$.

We now show that $ExtTRANSIM_n \notin \mathbf{U}^{n+1}\mathbf{TxtEx}$. Let $G_0, G_1, G_2, \ldots, G_n \in \mathbb{A}_n, \xi \in rat, \xi > 0$ be such that for all $\delta \in rat$, $\delta \leq \xi$, $P(G_0 + \delta v_1) \subseteq \bigcup_{i=1}^{n} P(G_i)$. Such $G_0, G_1, \ldots, G_n, \xi$ exist by Lemma 37. Let $\Lambda = \{G_1, G_2, \ldots, G_n\}$. Without loss of generality, we assume that in the programming system ψ, $W^{\psi}_{h(X_1(G_i))} = \emptyset$, for $1 \leq i \leq n$.

Let $\Lambda' = \{G_0 + \alpha v_1 \mid 0 \leq \alpha \leq \xi \ \wedge \ \alpha \in rat\}$. Let $\mathcal{L}' = \{L_{G'} \cup \bigcup_{G\in\Lambda} L_G \mid G' \in \Lambda'\}$.

Clearly, for all but finitely many primes p, $(\exists\, l \in N \mid l$ is co-prime with $p)[X_1(G_0) < \frac{l}{p} < X_1(G_0) + \xi]$. Thus, the set $\{W^{\psi}_{h(z)} \mid z \in rat,\ X_1(G_0) < z < X_1(G_0) + \xi\}$ includes all the r.e. languages. Furthermore, for each $z \in rat$ such that $X_1(G_0) < z < X_1(G_0) + \xi$, there exists a language in $\mathcal{L}'$ which differs from $\bigcup_{G\in\Lambda} L_G$ by the set $\{\langle 1, y\rangle \mid y \in W^{\psi}_{h(z)}\}$. Thus if $\mathcal{L}'$ is in **TxtEx**, then $\mathcal{E} \in$ **TxtEx**. Since $\mathcal{E} \notin$ **TxtEx** [Gol67], it follows that $\mathcal{L}' \notin$ **TxtEx**. Since $\mathcal{L}' \subseteq ExtTRANSIM_n^{n+1}$, $ExtTRANSIM_n \notin \mathbf{U}^{n+1}\mathbf{TxtEx}$. ∎

Note that the diagonalization in the above Theorem holds even against non-computable learners, as $\mathcal{E} \notin$ **TxtEx** even for non-computable learners [Gol67].

7 Intrinsic Complexity

The concept of *intrinsic complexity* [FKS95,JS96] is an attempt to describe the relative hardness of identifying a class of languages under the requirement given by an identification criterion. The idea is to reduce the task of identifying a class of languages to the task of identifying another class. To be able to reduce the identification of $\mathcal{L}$ to that of identifying $\mathcal{L}'$, we should be able to transform admissible texts T for languages in $\mathcal{L}$ to admissible texts T' for languages in $\mathcal{L}'$ and further transform sequences of conjectures witnessing identification of T' into sequences of conjectures witnessing identification of T. We refer the reader to [FKS95,JS96] for more discussion on intrinsic complexity.

A sequence $i_0 i_1 \ldots$ is said to be **TxtEx**-admissible for a text T of language L, iff for all but finitely many n, $W_{i_n} = L = \mathit{content}(T)$. One can similarly define admissible sequences for a text for unions of languages, as follows:

(a) A sequence $i_0 i_1 \ldots$, which converges to i, is said to be $\mathbf{U}^n\mathbf{TxtEx}$-admissible for a text T iff $W_i = \mathit{content}(T)$.

(b) A sequence $i_0 i_1 \ldots$, which converges to i, is said to be $\mathbf{WDU}^n\mathbf{TxtEx}$-admissible (with respect to a class $\mathcal{L}$) for a text T iff $\bigcup_{j\in D_i} W_j = \mathit{content}(T)$, $\mathit{card}(D_i) \le n$, and $\{W_j \mid j \in D_i\} \subseteq \mathcal{L}$.

For $\mathbf{DU}^n\mathbf{TxtEx}$-admissible sequence, we additionally require that $\{W_j \mid j \in D_i\}$ be the unique subset of $\mathcal{L}$ which unions to $\mathit{content}(T)$.

An *enumeration operator* (or just operator), Θ, is an algorithmic mapping from SEQ into SEQ such that for all σ, $\tau \in$ SEQ, if $\sigma \subseteq \tau$, then $\Theta(\sigma) \subseteq \Theta(\tau)$. We further assume that for all texts T, $\lim_{n\to\infty} |\Theta(T[n])| = \infty$. By extension, we think of Θ as also defining a mapping from texts to texts such that $\Theta(T) = \bigcup_{n\in N} \Theta(T[n])$.

[JS96] distinguished between two kinds of reductions, called *weak* and *strong reductions*. We consider only the former here. We extend the definition for weak reduction as follows, so that instead of just reducing the task of identifying every language in a class, $\mathcal{L}_1$ say, to tasks of identifying languages in another class $\mathcal{L}_2$, we want to reduce the task for identifying every language in $\mathcal{L}_1^n$ to tasks of identifying languages in $\mathcal{L}_2^m$, for some $m, n \in N$.

Definition 43 (Based on [FKS95,JS96]) Let $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{E}$ be given. Let $\mathcal{K}_1, \mathcal{K}_2 \in \{\mathbf{U}, \mathbf{DU}, \mathbf{WDU}\}$ and $n, m \in N^+$ be given. Let $\mathcal{T}_1 = \{T \mid T$ is a text for $L \in \mathcal{L}_1^n\}$. Let $\mathcal{T}_2 = \{T \mid T$ is a text for $L \in \mathcal{L}_2^m\}$. We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}_1^n \boldsymbol{TxtEx}, \mathcal{K}_2^m \boldsymbol{TxtEx}} \mathcal{L}_2$ just in case there exist operators Θ and Ω such that for all $T \in \mathcal{T}_1$ and for all infinite sequences of conjectures $\mathcal{G}$ the following hold:

(a) $\Theta(T) \in \mathcal{T}_2$, and

(b) if $\mathcal{G}$ is a $\mathcal{K}_2^m\mathbf{TxtEx}$-admissible sequence for $\Theta(T)$, then $\Omega(\mathcal{G})$ is a $\mathcal{K}_1^n\mathbf{TxtEx}$-admissible sequence for T.

We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}^n \boldsymbol{TxtEx}} \mathcal{L}_2$ if and only if $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}^n \boldsymbol{TxtEx}, \mathcal{K}^n \boldsymbol{TxtEx}} \mathcal{L}_2$.

Definition 44 [JS96] Let $\mathcal{J}$ be an identification criterion. Let $\mathcal{L} \subseteq \mathcal{E}$ be given.

(a) If for all $\mathcal{L}' \in \mathcal{J}$, $\mathcal{L}' \leq_{weak}^{\mathcal{J}} \mathcal{L}$, then $\mathcal{L}$ is $\leq_{\mathrm{weak}}^{\mathcal{J}}$-*hard.*

(b) If $\mathcal{L}$ is $\leq_{weak}^{\mathcal{J}}$-hard and $\mathcal{L} \in \mathcal{J}$, then $\mathcal{L}$ is $\leq_{\mathrm{weak}}^{\mathcal{J}}$-*complete.*

Theorem 45 *For all $n \in N^+$,*

(a) INIT is $\leq_{\mathrm{weak}}^{\mathbf{U}^n\mathbf{TxtEx}}$-complete.

(b) INIT is $\leq_{\mathrm{weak}}^{\mathbf{DU}^n\mathbf{TxtEx}}$-hard.

(c) INIT is $\leq_{\mathrm{weak}}^{\mathbf{WDU}^n\mathbf{TxtEx}}$-complete.

Proof. Fix $n \in N^+$. We first note that $INIT^n = INIT$.

(a) Clearly, $INIT \in \mathbf{U}^n\mathbf{TxtEx}$. Now suppose $\mathcal{L} \in \mathbf{U}^n\mathbf{TxtEx}$ is given. Since $INIT$ is $\leq_{weak}^{\boldsymbol{TxtEx}}$-complete [JS96], there exist Θ and Ω which observe that $\mathcal{L}^n \leq_{weak}^{\boldsymbol{TxtEx}} INIT$. It follows that $\mathcal{L} \leq_{weak}^{\boldsymbol{U}^n \boldsymbol{TxtEx}} INIT$. Part (a) follows.

(b), (c) can be proved using essentially the same proof as used to show that $INIT$ is $\leq_{weak}^{\boldsymbol{TxtEx}}$-complete in [JS96]. (The reduction of $\mathcal{L}$ to $INIT$ in [JS96] only used the final conjecture of the **TxtEx**-learner M on texts T for $L \in \mathcal{L}$ as a numeric value. One can do the same for the final conjecture of $\mathbf{DU}^n\mathbf{TxtEx}$-learner for $\mathcal{L}$). We omit the details. Note that $INIT \in \mathbf{WDU}^n\mathbf{TxtEx}$, but not in $\mathbf{DU}^n\mathbf{TxtEx}$. Thus, we only get the hardness result for $\mathbf{DU}^n\mathbf{TxtEx}$. ∎

It is clear that for any n, $TRANSIM_n$ fulfills the condition for Proposition 35. In fact, we shall next show that in some sense, $TRANSIM_n$ is the most difficult class to learn in $\mathbf{DU}^n\mathbf{TxtEx}$.

Let $XL(n, \epsilon) = Lang(\mathbb{G}_n) \cup \bigcup_{1 \leq i \leq n-1} Lang(\mathbb{G}_n + \epsilon v_i)$. It is easy to verify that for $\epsilon \in rat$, $XL(n, \epsilon) \in TRANSIM_n^n$.

Lemma 46 *There exists $\omega^* \in rat$, $\omega^* > 0$ such that for all $\omega, \omega' \in rat$, $0 \leq \omega \leq \omega' \leq \omega^*$, $XL(n, \omega) \subseteq XL(n, \omega')$.*

Proof. Let $\omega^* = \frac{1}{n+1}$. Lemma 46 now follows by using Lemma 36. ∎

Theorem 47 *For all $n \in N$, $n \geq 2$, $TRANSIM_n$ is $\leq_{\text{weak}}^{\mathbf{DU}^n\mathbf{TxtEx}}$-complete.*

Proof. Let $n \in N$, $n \geq 2$. Let $\langle .,. \rangle_p$ be a 1–1 pairing function with range in the prime numbers. For any $\mathcal{L} \subseteq \mathbf{DU}^n\mathbf{TxtEx}$, we construct Θ and Ω which witness that $\mathcal{L} \leq_{weak}^{\boldsymbol{DU^n TxtEx}} TRANSIM_n$. Suppose M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$, $card(\mathcal{L}') \leq n$ and let T be a text for $L_{\mathcal{L}'}$. Without loss of generality assume that $M(T[0]) \neq M(T[1])$.

Let ω^* be as in Lemma 46.

Define Θ as follows. For text T and $s \in N$,

$\Theta(T[0])$:
   Let $\omega^0 = 0$.
   Return $\#$.
End

$\Theta(T[s+1])$:
   If $M(T[s+1]) = M(T[s])$ then let $\omega^{s+1} = \omega^s$.
   Else
      Find least $m \in N$, and corresponding $l \in N$ such that,
         l is co-prime with $\langle M(T[s+1]), m \rangle_p$
         and $\omega^s \leq \frac{l}{\langle M(T[s+1]), m \rangle_p} < \omega^*$.
      Let $\omega^{s+1} = \frac{l}{\langle M(T[s+1]), m \rangle_p}$.
   Let σ be an extension of $\Theta(T[s])$ such that
      $content(\sigma) = \{x \mid x \leq s+1 \wedge x \in XL(n, \omega^{s+1})\}$.
   Return σ.
End

For any $i, j \in N$, $i < j$, clearly $\omega^i \leq \omega^j < \omega^*$. Thus, by Lemma 46, $XL(n, \omega^i) \subseteq XL(n, \omega^j)$. Hence at each stage s, $content(\Theta(T[s])) \subseteq XL(n, \omega^s)$. If M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $L_{\mathcal{L}'}$, then at some stage t, M stops changing it's mind (that is, $M(T[t]) = M(T[t'])$, for $t' \geq t$). Thus, $\Theta(T)$ is a text for the language $XL(n, \omega^t)$.

To obtain operator Ω transforming a sequence of conjectures for $XL(n, \omega^t)$ into a sequence of conjectures for $L_{\mathcal{L}'}$, observe that it is possible to restore the value $M(T)$ from a sequence of conjectures for $XL(n, \omega^t)$. Let $\mathcal{G} = \mathcal{G}(0)\mathcal{G}(1)\mathcal{G}(2)\ldots$ be an infinite sequence of conjectures, define $\Omega(\mathcal{G}) = \mathcal{G}'$ where for each s, $\mathcal{G}'(s)$ is defined as follows. Let $z_s = \max(\{X_1(\mathrm{decoderat}_n(w)) - 1 \mid w \in \bigcup_{j \in D_{\mathcal{G}(s)}} W_{j,s}\})$ (the function $X_1(p)$ denotes the coordinate in x_1 axis of the point p). Intuitively, here z_s attempts to restore the value ω^t from $\mathcal{G}(s)$, a conjecture for $XL(n, \omega^t)$. Finally, let $\mathcal{G}'(s) = \pi_1(h(z_s))$, where $h(a)$ is the denominator of rational a in reduced form. It is easy to verify that if M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies T, and $\mathcal{G}$ converges to a conjecture for *content*$(\Theta(T))$, then $\Omega(\mathcal{G})$ converges to $M(T)$.

Since $TRANSIM_n \in \mathbf{DU}^n\mathbf{TxtEx}$, Theorem follows. ∎

Similarly, one can show

Theorem 48 *For all $n \in N$, $n \geq 2$,*

(a) $TRANSIM_n$ is $\leq_{\mathrm{weak}}^{\mathbf{U}^n\mathbf{TxtEx}}$-complete.

(b) $TRANSIM_n$ is $\leq_{\mathrm{weak}}^{\mathbf{WDU}^n\mathbf{TxtEx}}$-complete.

8 Acknowledgements

We thank an anonymous Referee for helpful comments and pointing out some improvements. We would also like to thank Frank Stephan for helpful discussions and Example 26.

References

[Ang80] D. Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.

[ASO94] H. Arimura, T. Shinohara, and S. Otsuki. Finding minimal generalizations for unions of pattern languages and its application to inductive inference from positive data. In *Proc. Annual Symp. on Theoretical Aspects of Computer Sci.*, 775:649–660, 1994.

[Blu67] M. Blum. A machine-independent theory of the complexity of recursive functions. *Journal of the ACM*, 14:322–336, 1967.

[BB75] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

[Cox63] H.S.M. Coxeter. *Regular Polytopes (2nd ed.).* Methuen & Co. Ltd., London, 1963.

[FKS95] R. Freivalds, E. Kinber, and C. Smith. On the intrinsic complexity of learning. *Information and Computation*, 123(1):64–71, 1995.

[Ful85] M. Fulk. *A Study of Inductive Inference Machines.* PhD thesis, SUNY/Buffalo, 1985.

[Ful90] M. Fulk. Prudence and other conditions on formal language learning. *Information and Computation*, 85:1–11, 1990.

[GK99] S. Goldman and S. Kwek. On learning unions of pattern languages and tree patterns. *Algorithmic Learning Theory: Tenth International Conference (ALT' 99)*, 1720:347–363, 1999.

[Gol67] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[JORS99] S. Jain, D. Osherson, J. Royer, and A. Sharma. *Systems that Learn: An Introduction to Learning Theory.* MIT Press, Cambridge, Mass., second edition, 1999.

[JS96] S. Jain and A. Sharma. The intrinsic complexity of language identification. *Journal of Computer and System Sciences*, 52:393–402, 1996.

[Kru72] J. B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory (A)*, 13:297–305, 1972.

[MSW91] T. Motoki, T. Shinohara, and K. Wright. The correct definition of finite elasticity: Corrigendum to identification of unions. In L. Valiant and M. Warmuth, editors, *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, page 375. Morgan Kaufmann, 1991.

[OSW86] D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists.* MIT Press, 1986.

[Rog67] H. Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, 1967. Reprinted by MIT Press in 1987.

[SA00] T. Shinohara and H. Arimura. Inductive inference of unbounded unions of pattern languages from positive data. *Theoretical Computer Science A*, 241:191–209, 2000.

[Wri89] K. Wright. Identification of unions of languages drawn from an identifiable class. In R. Rivest, D. Haussler, and M. Warmuth, editors, *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 328–333. Morgan Kaufmann, 1989.