



Sacred Heart
UNIVERSITY

Sacred Heart University
DigitalCommons@SHU

School of Computer Science & Engineering
Faculty Publications

School of Computer Science and Engineering

11-2007

Learning Multiple Languages in Groups

Sanjay Jain
National University of Singapore

Efim Kinber
Sacred Heart University

Follow this and additional works at: https://digitalcommons.sacredheart.edu/computersci_fac



Part of the [Computer Sciences Commons](#)

Recommended Citation

Jain, S., & Kinber, E. (2007). Learning multiple languages in groups. *Theoretical Computer Science*, 387(1), 67-76. Doi: 10.1016/j.tcs.2007.07.025

This Peer-Reviewed Article is brought to you for free and open access by the School of Computer Science and Engineering at DigitalCommons@SHU. It has been accepted for inclusion in School of Computer Science & Engineering Faculty Publications by an authorized administrator of DigitalCommons@SHU. For more information, please contact ferribyp@sacredheart.edu, lysobeyb@sacredheart.edu.

Learning multiple languages in groups

Sanjay Jain^{a,*}, Efim Kinber^b

^a *School of Computing, National University of Singapore, Singapore 117590, Singapore*

^b *Department of Computer Science, Sacred Heart University, Fairfield, CT 06432-1000, USA*

Abstract

We consider a variant of Gold's learning paradigm where a learner receives as input n different languages (in the form of one text where all input languages are interleaved). Our goal is to explore the situation when a more "coarse" classification of input languages is possible, whereas more refined classification is not. More specifically, we answer the following question: under which conditions, a learner, being fed n different languages, can produce m grammars covering all input languages, but cannot produce k grammars covering input languages for any $k > m$. We also consider a variant of this task, where each of the output grammars may not cover more than r input languages. Our main results indicate that the major factor affecting classification capabilities is the difference $n - m$ between the number n of input languages and the number m of output grammars. We also explore the relationship between classification capabilities for smaller and larger groups of input languages. For the variant of our model with the upper bound on the number of languages allowed to be represented by one output grammar, for classes consisting of disjoint languages, we found complete picture of relationship between classification capabilities for different parameters n (the number of input languages), m (number of output grammars), and r (bound on the number of languages represented by each output grammar). This picture includes a combinatorial characterization of classification capabilities for the parameters n, m, r of certain types.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Computational learning theory; Learning in the limit; Classification; Learning multiple languages

1. Introduction

In this paper, we continue a line of research where the learner is required to learn unions of different concepts [7]. This situation occurs, for example, when children living in a multi-lingual environment learn several languages simultaneously. In this case, in ideal, children are able to learn each individual language. A more complex case is the problem of multi-layered classification, where descriptions of families of objects on a higher level are rather "coarse", while descriptions on a lower level are much more specific/refined. An example of this type of classification is the theory of species: as a result of learning process (inductive synthesis of concepts from examples), life can be classified as just animals and plants, or, more specifically, as families—fishes, birds, mammals, etc., or, even more specifically as different species, etc. An important issue here is that a more coarse classification is typically easier to achieve than a more refined one. For example, an alien civilization, learning life on Earth (from examples), most likely, will have

* Corresponding address: Department of Computer Science, School of Computing, National University of Singapore, Singapore 117590, Singapore. Tel.: +65 6874 7842; fax: +65 6779 4580.

E-mail addresses: sanjay@comp.nus.edu.sg (S. Jain), kinbere@sacredheart.edu (E. Kinber).

much more difficulty describing different birds, than making distinction between birds and fishes. A child, learning classical music, has much easier time determining if a piece uses 3/4 or 4/4 time signature, than telling apart waltz, mazurka, or polonaise (each of them uses time signature 3/4).

Our goal in this paper is to determine if, and under which circumstances, a more coarse classification, as a result of learning process, is possible, whereas a more refined classification is not. More specifically, we explore the following general situation: under which circumstances, a learner, facing a union of n languages on the input, is able to learn descriptions of m (larger) groups of languages from the union, but is not able to learn descriptions of k (smaller) groups of input languages for $k > m$. For example, we would like to find out when a learner, facing a union of 6 languages on the input, can learn descriptions of 3 groups of languages, but cannot learn descriptions for each individual language. We are also interested in situations when learning larger groups of languages can be easier than smaller ones.

To model the process of learning, we employ the well-known Gold's learning paradigm [5]: the learner receives all members of the union of languages in arbitrary order and produces a sequence of descriptions (grammars) that stabilizes to a correct description. This model is known in literature as **TextEx** (where Ex stands for “explanatory learning”). Exploration of this model provided a robust advice to cognition theory (see, for example, [14]). We consider also a popular variant of this model, **TextBc**, (introduced in [3,10]) where a learner produces a sequence of conjectures, almost all of which are correct descriptions of the target language (but not necessarily the same—BC here stands for “behaviourally correct learning”).

Among several papers in this line of research, the closest to our inquiry is the paper [7], where the authors primarily explore the issues of learnability of larger unions of languages versus smaller unions of languages from the same families. In particular, they define the concept of *discerning* learnability, when a learner is required to learn each member of the union, and compare this notion with the situation when the learner may provide one description for the whole union. Relevant results from [7] can be viewed as the first step in our line of research.

Our main results can be summarized as follows. On one hand, if $n - m > n' - m'$ then there exists a class of languages such that it is possible to learn unions of n languages from this class in m groups, but it is not possible to learn unions of n' languages from this class in m' groups (**Theorem 7**). That is, the difference between the number of input languages in the union and the number of conjectures ultimately produced by the learner is the major factor affecting learning capabilities. On the other hand, if a family consists of only disjoint languages, then, if it is possible to learn unions of n languages in m groups, then it is possible to learn unions of $n - 1$ languages in $m - 1$ groups. For example, learnability of unions of any 6 disjoint languages in 3 groups implies learnability of any union of 5 languages in 2 groups (**Theorem 11**). We also extend our results to the case when the number of languages in learned groups is bounded (in the general case, the learner, when required to produce at least 3 groups for the union of 6 input languages, can include 4 languages into one group and only one language into each of the remaining two groups); the corresponding results are presented in **Corollaries 16 and 17**, and **Theorem 18**. The last result of this paper, **Theorem 20**, presents a combinatorial characterization (when the language classes consist of disjoint languages) for the remaining cases not solved by **Corollaries 16 and 17**, and **Theorem 18** (i.e., the circumstances under which learnability of unions of n languages in m groups describing at most r languages in each of the groups implies learnability for other corresponding parameters n', m', r' , where $n' \leq n, r \leq r'$ and $n' - m' < n - m$).

2. Notations and preliminaries

Any unexplained recursion theoretic notation is from [13]. N denotes the set of natural numbers, $\{0, 1, 2, 3, \dots\}$. \emptyset denotes the empty set. $\subseteq, \subset, \supseteq, \supset$ respectively denote subset, proper subset, superset and proper superset. D_x denotes the finite set with (canonical) index x [13]. We sometimes identify finite sets with their canonical indices. The quantifier ‘ \forall^∞ ’ essentially from [2], means ‘for all but finitely many’.

\uparrow denotes undefined. $\max(\cdot), \min(\cdot)$ denote the maximum and minimum of a set, respectively, where $\max(\emptyset) = 0$ and $\min(\emptyset) = \uparrow$. $\text{card}(S)$ denotes the cardinality of set S . $\langle \cdot, \cdot \rangle$ stands for an arbitrary, computable, one-to-one encoding of all pairs of natural numbers onto N [13]. Similarly we can define $\langle \cdot, \dots, \cdot \rangle$ for encoding tuples of natural numbers onto N .

φ denotes a fixed *acceptable* programming system for the partial computable functions: $N \rightarrow N$ [12,13,8]. φ_i denotes the partial computable function computed by program i in the φ -system. W_i denotes $\text{domain}(\varphi_i)$. Note that all acceptable numberings are isomorphic and thus one could also define W_i to be the set generated by the i th type-0 grammar. \mathcal{E} denotes the set of all recursively enumerable (r.e.) languages. L , with or without

subscripts and superscripts, ranges over \mathcal{E} . \mathcal{L} , with or without subscripts and superscripts, ranges over subsets of \mathcal{E} . **DisjClass** = $\{\mathcal{L} \mid (\forall L, L' \in \mathcal{L})[L = L' \text{ or } L \cap L' = \emptyset]\}$, i.e, **DisjClass** is the collection of language classes which consist of disjoint languages.

We now consider some basic notions in language learning. We first introduce the concept of data that is presented to a learner. A *text* T is a mapping from N into $(N \cup \{\#\})$ (see [5]). The *content* of a text T , denoted $\text{content}(T)$, is the set of natural numbers in the range of T . T is a text for L iff $\text{content}(T) = L$. $T[n]$ denotes the initial segment of T of length n . We let T , with or without superscripts, range over texts. Intuitively, $\#$'s in the texts denote pauses in the presentation of data. For example, the only text for the empty language is just an infinite sequence of $\#$'s.

A finite sequence σ is an initial segment of a text. $\text{content}(\sigma)$, is the set of natural numbers in the range of σ . $|\sigma|$ denotes the length of σ , and if $n \leq |\sigma|$, then $\sigma[n]$ denotes the initial segment of σ of length n .

A *language learning machine* is an algorithmic device which computes a mapping from finite initial segments of texts into N . We let \mathbf{M} , with or without subscripts and superscripts, range over learning machines. We say that $\mathbf{M}(T)\downarrow = i \Leftrightarrow (\forall^\infty n)[\mathbf{M}(T[n]) = i]$.

We now introduce criteria for a learning machine to be considered *successful* on languages. Our first criteria is based on the learner, given a text for the language, converging to a grammar for the language.

Definition 1 ([5]). (a) **M TxtEx-identifies** L (written: $L \in \mathbf{TxtEx}(\mathbf{M})$) $\Leftrightarrow (\forall \text{ texts } T \text{ for } L)(\exists i \mid W_i = L)[\mathbf{M}(T)\downarrow = i]$.

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

The influence of Gold's paradigm [5] to human language learning is discussed by various authors, for example [11,14,9].

The following definition is based on the learner semantically rather than syntactically converging to the grammar(s) for the language. Here, note that equivalence of grammars is non-computable. The corresponding notion for learning functions was introduced by [1,4].

Definition 2 ([3,10]). (a) **M TxtBc-identifies** L (written: $L \in \mathbf{TxtBc}(\mathbf{M})$) $\Leftrightarrow (\forall \text{ texts } T \text{ for } L)(\forall^\infty n)[W_{\mathbf{M}(T[n])} = L]$.

(b) **TxtBc** = $\{\mathcal{L} \mid (\exists \mathbf{M})[\mathcal{L} \subseteq \mathbf{TxtBc}(\mathbf{M})]\}$.

It can be shown that **TxtEx** \subset **TxtBc** (for example, see [3,10]).

3. Learning languages in groups

Now we give the formal definition of our main learning model—under which a learner, being fed the union of n different languages, outputs in the limit at least m grammars, each representing a number of input languages, so that the union of all m grammars covers the union of all input languages.

Definition 3. (a) We say that $\mathbf{M}[m, n]$ **MultEx-identifies** \mathcal{L} , iff for all distinct languages L_1, \dots, L_n in \mathcal{L} , for all texts T for $L_1 \cup \dots \cup L_n$, there exist i_1, \dots, i_k , (where $k \geq m$) such that $\mathbf{M}(T)$ converges on T to the canonical index for the set $\{i_1, \dots, i_k\}$ and there exists a partition G_1, \dots, G_k of $\{1, 2, \dots, n\}$ such that

(i) each G_i is non-empty,

(ii) for $1 \leq j \leq k$, i_j is a grammar for $\bigcup_{r \in G_j} L_r$.

(b) $[m, n]$ **MultEx** = $\{\mathcal{L} \mid (\exists \mathbf{M})[\mathbf{M}[m, n]$ **MultEx-identifies** $\mathcal{L}]\}$.

Note that requiring $k = m$ in the above does not change the class of languages which can be $[m, n]$ **MultEx-identified**, as one can just combine $k - m + 1$ of the language groups into one. However, this may make a difference in some modifications we consider later. Our next definition is a modification of our model for behaviorally correct type of learning.

Definition 4. (a) We say that $\mathbf{M}[m, n]$ **MultBc-identifies** \mathcal{L} , iff for all distinct languages L_1, \dots, L_n in \mathcal{L} , for all texts T for $L_1 \cup \dots \cup L_n$, for all but finitely many t , there exist i_1, \dots, i_k , (where $k \geq m$) such that $\mathbf{M}(T[t]) =$ the index for $\{i_1, \dots, i_k\}$ and there exists a partition G_1, \dots, G_k of $\{1, 2, \dots, n\}$ such that

(i) each G_i is non-empty,

(ii) for $1 \leq j \leq k$, i_j is a grammar for $\bigcup_{r \in G_j} L_r$.

(b) $[m, n]$ **MultBc** = $\{\mathcal{L} \mid (\exists \mathbf{M})[\mathbf{M}[m, n]$ **MultBc-identifies** $\mathcal{L}]\}$.

Our first result demonstrates that, for some families of languages, one can **Bc**-learn each individual language from the input union of n languages, while no **Ex**-learner can converge to a correct single grammar representing the whole union.

Theorem 5. For $1 \leq m$ and $1 \leq n$, $[n, n]\mathbf{MultBc} - [1, m]\mathbf{MultEx} \neq \emptyset$.

Moreover, this separation can be witnessed by a class in **DisjClass**.

Proof. Let \mathcal{L} be a class of languages in **TxtBc** – **TxtEx** (such a class exists by results from [4,3]). Let $\text{cyl}_L^i = \{\langle i, x \rangle \mid x \in L\}$. It is easy to verify that for all i , one can find an $L_i \in \mathcal{L}$ such that \mathbf{M}_i does not **TxtEx**-identify $\bigcup_{i * m \leq j < (i+1) * m} \text{cyl}_{L_i}^j$ (otherwise, one can easily show that $\mathcal{L} \in \mathbf{TxtEx}$). Let $\mathcal{L} = \{\text{cyl}_{L_i}^j \mid i \in N, i * m \leq j < (i + 1) * m\}$. It now follows that $\mathcal{L} \notin [1, m]\mathbf{MultEx}$. On the other hand, $\mathcal{L} \in [n, n]\mathbf{MultBc}$ easily follows as for any S_1, \dots, S_n in the class \mathcal{L} , from a text for $S_1 \cup \dots \cup S_n$, one can easily obtain texts for S_1, \dots, S_n and then use **TxtBc**-learning procedure for each of them. ■

Now we show that, if the number of the languages in the union given to an **Ex**-learner is smaller than the number of languages given to a **Bc**-learner, then the reverse of the above result also holds: for some family of languages, an **Ex**-learner can correctly infer grammars for each individual language from a smaller input union, while **Bc**-learnability of even one grammar covering the whole larger union of languages is impossible.

Theorem 6. Suppose $1 \leq n$. $[n, n]\mathbf{MultEx} - [1, n + 1]\mathbf{MultBc} \neq \emptyset$.

Moreover, this separation can be witnessed by a class in **DisjClass**.

Proof. The proof is a slight modification of the construction in [7] to show that there exists a $\mathcal{L} \in \mathbf{DisjClass}$ separating $\mathbf{DU}^n\mathbf{TxtEx}$ and $\mathbf{U}^{n+1}\mathbf{TxtEx}$.

Let $n > 0$ be given. For each $e \in N$, we will construct a finite set S_e , and languages L_e^0, \dots, L_e^n where

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

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

(1) $\{z \mid (\exists x, y)[\langle x, y, z \rangle \in L_e^i]\} \neq \emptyset$, and $\min(\{z \mid (\exists x, y)[\langle x, y, z \rangle \in L_e^i]\}) > \max(S_e)$.

(2) $L_e^i = \{\langle e, i, j \rangle \mid j \in W_{\min(\{z \mid (\exists x, y)[\langle x, y, z \rangle \in L_e^i]\})}\}$.

Let $\mathcal{L} = \{L_e^0, L_e^1, \dots, L_e^n \mid e \in N\}$. Clearly, for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$. We shall first show that $\mathcal{L} \in [n, n]\mathbf{MultEx}$ (for any choice of S_e, L_e^0, \dots, L_e^n satisfying the above properties).

By s-m-n theorem there exists a recursive function g such that, for $1 \leq i \leq n$, and $e, j \in N$,

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

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

Now $\mathcal{L} \in [n, n]\mathbf{MultEx}$ is witnessed by following **M**. For each text T and each $m \in N$,

M($T[m]$)

Let $S = \emptyset$.

Let $A = \{e \mid (\exists y, z)[\langle e, y, z \rangle \in \text{content}(T[m])]\}$.

For each $e \in A$ do

Let $B = \text{content}(T[m])$.

If $\langle e, 0, 0 \rangle \in \text{content}(T[m])$ then

(* This takes care of L_e^0 being part of input. *)

Let $C = \{j \mid (\forall i \mid 1 \leq i \leq n)[\langle e, i, j \rangle \in \text{content}(T[m])]\}$.

Let j be such that $D_j = C$.

Let $S = S \cup \{g(e, 0, j)\}$.

Let $B = B - W_{g(e,0,j)}$.

Endif

For $i = 1$ to n do

(* This takes care of L_e^i being part of input. *)

If there exists j such that $\langle e, i, j \rangle \in B$, then

For minimum such j , let $S = S \cup \{g(e, i, j)\}$.

EndFor

EndFor

Output (index for) S .

End

It is easy to verify that $\mathbf{M}[n, n]$ **MultEx**-identifies \mathcal{L} . We now show that $\mathcal{L} \notin [1, n+1]$ **MultBc**, for some appropriate choice of L_e^i , for each e, i . For each e here is the construction to show that \mathbf{M}_e does not $[1, n+1]$ **MultBc**-identify $\{L_e^0, \dots, L_e^n\}$. By Kleene's Recursion Theorem [13] there exists an index $e' > 0$ such that $W_{e'}$ may be defined in stages as follows. For each s , $W_{e'}^s$ denotes the finite portion of $W_{e'}$ enumerated just before stage s . Initially, enumerate e' in $W_{e'}$ (thus $W_{e'}^0 = \{e'\}$). Let σ^0 be such that $\text{content}(\sigma^0) = \{(e, 0, 0)\} \cup \{(e, j, e') \mid 1 \leq j \leq n\}$. Go to stage 0.

Stage s :

Search for a $\tau \supseteq \sigma^s$ and $x \in N$, such that

(i) $\text{content}(\tau) \subseteq \text{content}(\sigma^s) \cup \{(e, i, j) \mid 1 \leq i \leq n \wedge j > \max(W_{e'}^s)\}$,

(ii) $x > \max(\{y \mid (\exists i : 1 \leq i \leq n)[\langle e, i, y \rangle \in \text{content}(\tau)]\})$, and

(iii) $W_{\mathbf{M}_e(\tau)}$ contains $\langle e, i, x \rangle$, for some $i, 1 \leq i \leq n$.

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

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

Go to stage $s+1$.

End Stage s .

If the search for τ failed at any stage s , then let $L_e^0 = \text{content}(\sigma^s)$ and let $e'' > \max(W_{e'}^s)$ be such that $W_{e''} = \{x \mid x \geq e''\}$ (by Kleene's Recursion Theorem [13], 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, \mathbf{M}_e does not **TextBc**-identify $(L_e^0 \cup \bigcup_{i=1}^n L_e^i)$ (since search in stage s did not succeed, on any finite sequence τ extending σ^s such that $\text{content}(\tau) \subseteq L_e^0 \cup \bigcup_{i=1}^n L_e^i$, $W_{\mathbf{M}(\tau)}$ has finite intersection with $\{\langle e, i, x \rangle \mid 1 \leq i \leq n, x \in N\}$).

If the search is successful at all stages, then let $L_e^0 = \{(e, 0, 0)\}$ and, for $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, x \rangle \mid x \in W_{e'}\}$. Now, \mathbf{M}_e fails to **TextBc**-identify $\bigcup_{s \in N} \sigma^s$, a text for $L_e^0 \cup \bigcup_{i=1}^n L_e^i$ (as in each stage a τ is found, with $\sigma^s \subseteq \tau \subseteq \sigma^{s+1}$, such that $\mathbf{M}(\tau)$ enumerates an element $\langle e, i, x \rangle$ (for some $i, 1 \leq i \leq n$, see clause (iii) above) which is not in $L_e^0 \cup \bigcup_{i=1}^n L_e^i$).

Theorem follows from the above analysis. ■

Yet another type of situation when **Ex**-learnability of unions in groups may be possible but **Bc**-learnability might be not is presented in the following result: for some class of languages, if the difference $n - m$ between the size n of the union of languages and the number of learned groups m is greater than $n' - m'$ (where $m' \geq 2$), then an **Ex**-learner can infer grammars for m groups representing an input union of n languages, while **Bc**-learning of any union of n' input languages in m' groups is not possible.

Theorem 7. Suppose $1 \leq m \leq n, 2 \leq m' \leq n'$, and $n - m > n' - m'$. Then, $[m, n]$ **MultEx** – $[m', n']$ **MultBc** $\neq \emptyset$.

Moreover, this separation can be witnessed by a class in **DisjClass**.

Proof. For $i \in N$, define

(i) for $x \leq n - m$, let

$$X_{i, s_0, s_1, \dots, s_{i-1}}^x = \{\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2x \rangle, \langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2x + 1 \rangle\}.$$

(ii) for $x < n - m, i \in N$, let

$$Y_{i, s_0, s_1, \dots, s_{i-1}}^x = \{\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2x + 1 \rangle, \langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2x + 2 \rangle\};$$

$$Y_{i, s_0, s_1, \dots, s_{i-1}}^{n-m} = \{\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2(n - m) + 1 \rangle, \langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 0 \rangle\}.$$

(iii) for $x < n' - (n - m + 1)$, let $Z_{i, s_0, s_1, \dots, s_{i-1}}^x = \{\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, 2n + x \rangle\}$.

Note that $\bigcup_{x \leq n-m} X_{i,s_0,s_1,\dots,s_{i-1}}^x = \bigcup_{x \leq n-m} Y_{i,s_0,s_1,\dots,s_{i-1}}^x$ (this will be utilized for diagonalization against \mathbf{M}_i).

For any binary sequence of s_i 's, let $\mathcal{L}_{s_0,s_1,\dots} = \{X_{i,s_0,s_1,\dots,s_{i-1}}^x \mid i \in N, x \leq n-m, s_i = 0\} \cup \{Y_{i,s_0,s_1,\dots,s_{i-1}}^x \mid i \in N, x \leq n-m, s_i = 1\} \cup \{Z_{i,s_0,s_1,\dots,s_{i-1}}^x \mid i \in N, x < n' - (n-m+1)\}$.

Claim 8. $\mathcal{L}_{s_0,s_1,\dots} \in [m, n]\mathbf{MultEx}$, for any fixed binary values of s_i 's.

Proof. Fix s_0, s_1, \dots . Consider any text T for $L_1 \cup \dots \cup L_n$ being given as input. Let $i = \max(\{i' \mid (\exists x, y)\langle i', x, y \rangle \in \text{content}(T)\})$. Let s'_0, \dots, s'_{i-1} be such that, for some y , $\langle i, \langle s'_0, s'_1, \dots, s'_{i-1} \rangle, y \rangle \in \text{content}(T)$. Note that it must be the case $s'_w = s_w$, for $w < i$. Thus, i and s_0, \dots, s_{i-1} can be determined in the limit.

Now the learner (in the limit) outputs the index for set S defined as follows:

- (A) For $j < i, x \leq n-m$, if $s_j = 0$ and $X_{j,s_0,\dots,s_{j-1}}^x \subseteq \text{content}(T)$, then S contains a grammar for $X_{j,s_0,\dots,s_{j-1}}^x$.
- (B) For $j < i, x \leq n-m$, if $s_j = 1$ and $Y_{j,s_0,\dots,s_{j-1}}^x \subseteq \text{content}(T)$, then S contains a grammar for $Y_{j,s_0,\dots,s_{j-1}}^x$.
- (C) For $j \leq i, x < n' - (n-m+1)$, if $Z_{j,s_0,\dots,s_{j-1}}^x \subseteq \text{content}(T)$, then S contains a grammar for $Z_{j,s_0,\dots,s_{j-1}}^x$.
- (D) S contains a grammar for $\bigcup_{x \leq n-m} X_{i,s_0,s_1,\dots,s_{i-1}}^x \cap \text{content}(T)$ (assuming this set is non-empty).

It is easy to verify that the above method outputs a set of at least m grammars which partition the input languages (the only case of more than one language in the input being combined to form a single grammar is via case (D) above). Claim follows. \square

Claim 9. There exist values of s_0, s_1, \dots such that $\mathcal{L}_{s_0,s_1,\dots} \notin [m', n']\mathbf{TextBc}$.

Proof. For each $i \in N$, define s_i inductively as follows.

Let T be a text for $\bigcup_{x \leq n-m} X_{i,s_0,\dots,s_{i-1}}^x \cup \bigcup_{x < n' - (n-m+1)} Z_{i,s_0,\dots,s_{i-1}}^x$.

Note that if \mathbf{M}_i , on text T , infinitely often outputs an index for a set which contains a grammar enumerating all of $\bigcup_{x \leq n-m} X_{i,s_0,\dots,s_{i-1}}^x$, then \mathbf{M}_i does not $[m', n']\mathbf{MultBc}$ -identify the input (since it combines $n-m+1$ languages in the same group). In this case one can choose s_i arbitrarily.

On the other hand, suppose for all but finitely many t , $\mathbf{M}_i(T[t])$ is an index for a set which contains at least two grammars which enumerate part of $\bigcup_{x \leq n-m} X_{i,s_0,\dots,s_{i-1}}^x$. Then, there must exist a $w \leq 2(n-m)$, such that $\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, w \rangle$ and $\langle i, \langle s_0, s_1, \dots, s_{i-1} \rangle, w+1 \rangle$ end up being enumerated by different grammars in the index set output by $\mathbf{M}_i(T[t])$, for infinitely many t . Now consider the following cases.

Case 1: w is even. Let $s_i = 0$. In this case one can easily verify that \mathbf{M}_i does not $[m', n']\mathbf{MultBc}$ -identify $\{X_{i,s_0,\dots,s_{i-1}}^x \mid x \leq m-n\} \cup \{Z_{i,s_0,\dots,s_{i-1}}^x \mid x < n' - (n-m+1)\}$.

Case 2: w is odd. Let $s_i = 1$. In this case one can easily verify that \mathbf{M}_i does not $[m', n']\mathbf{MultBc}$ -identify $\{Y_{i,s_0,\dots,s_{i-1}}^x \mid x \leq m-n\} \cup \{Z_{i,s_0,\dots,s_{i-1}}^x \mid x < n' - (n-m+1)\}$. \square

Theorem follows from the above claims. \blacksquare

Following technical proposition is helpful for our results.

Proposition 10. Fix $1 \leq n$. Suppose \mathcal{L} is an infinite class in $\mathbf{DisjClass}$ such that $\mathcal{L} \in [1, n]\mathbf{MultEx}$. Let $1 \leq m \leq n$.

Then, from any distinct $L_1, \dots, L_m \in \mathcal{L}$, from a text T for $L_1 \cup \dots \cup L_m$, one can effectively find in the limit

- (a) g_1, \dots, g_{n-m} such that W_{g_i} are distinct languages in $\mathcal{L} - \{L_1, \dots, L_m\}$,
- (b) a grammar g for $L_1 \cup \dots \cup L_m$.

Note that part (b) above implies $\mathcal{L} \in [1, m]\mathbf{MultEx}$.

Proof. Let e_1, \dots, e_{2n-m} be such that $W_{e_1}, \dots, W_{e_{2n-m}}$ are distinct languages in \mathcal{L} . Let \mathbf{M} be $[1, n]\mathbf{MultEx}$ -learner for \mathcal{L} .

Now given any text T as in the hypothesis, one can effectively search for distinct $g_1, \dots, g_{n-m}, g'_1, \dots, g'_{n-m} \in \{e_1, \dots, e_{2n-m}\}$ such that

$$\left[\bigcup_{1 \leq r \leq n-m} (W_{g_r} \cup W_{g'_r}) \right] \cap \text{content}(T) = \emptyset.$$

Let i be the grammar to which \mathbf{M} converges on a text for $\text{content}(T) \cup \bigcup_{1 \leq r \leq n-m} W_{g_r}$ and i' be the grammar to which \mathbf{M} converges on a text for $\text{content}(T) \cup \bigcup_{1 \leq r \leq n-m} W_{g'_r}$.

It is now easy to verify that, g_1, \dots, g_{n-m} satisfy part (a) of the Proposition, and $L_1 \cup \dots \cup L_m = W_i \cap W_{i'}$ (which allows us to find g as required for part (b)). \blacksquare

Our next result shows that, for classes in **DisjClass**, reducing the number of languages in the input union and the number of learned groups by the same parameter does not affect **MultEx**- and **MultBc**-learnability.

Theorem 11. *Suppose $\mathcal{L} \in \mathbf{DisjClass}$. Assume $1 \leq m \leq n$, $1 \leq s \leq n$. Then (a) to (d) hold.*

- (a) *For $m \geq 2$, $\mathcal{L} \in [m, n]\mathbf{MultEx}$ implies $\mathcal{L} \in [m - 1, n - 1]\mathbf{MultEx}$.*
- (b) *For $m \geq 2$, $\mathcal{L} \in [m, n]\mathbf{MultBc}$ implies $\mathcal{L} \in [m - 1, n - 1]\mathbf{MultBc}$.*
- (c) *$\mathcal{L} \in [m, n]\mathbf{MultEx}$ implies $\mathcal{L} \in [1, s]\mathbf{MultEx}$.*
- (d) *$\mathcal{L} \in [m, n]\mathbf{MultBc}$ implies $\mathcal{L} \in [1, s]\mathbf{MultBc}$.*

Proof. We show part (a). Part (b) can be shown similarly. Part (c) follows from [Proposition 10](#)(b), and part (d) can be proved similarly.

(a) Without loss of generality assume \mathcal{L} is infinite. Suppose \mathbf{M} $[m, n]\mathbf{MultEx}$ -identifies \mathcal{L} . Define \mathbf{M}' as follows.

Suppose a text T for $L_1 \cup \dots \cup L_{n-1}$ is given as input. Let g_1 be a grammar such that $W_{g_1} \cap \text{content}(T) = \emptyset$ and $W_{g_1} \in \mathcal{L}$. Let g be a grammar for $\text{content}(T)$ (Note that by [Proposition 10](#), one can find such a g , g_1 in the limit).

Let T' be a text for $\text{content}(T) \cup W_{g_1}$. Suppose $\mathbf{M}(T')$ converges to the index for $\{i_1, \dots, i_m\}$. For $1 \leq r \leq m$, let i'_r be a grammar for $W_{i_r} \cap W_g$. Then, $\mathbf{M}'(T)$ converges to the index for $\{i'_r \mid 1 \leq r \leq m \text{ and } W_{i'_r} \neq \emptyset\}$. It is easy to verify that \mathbf{M}' $[m - 1, n - 1]\mathbf{MultEx}$ -identifies \mathcal{L} (as at most one of $W_{i'_r}$ is empty). ■

Corollary 12. *Suppose $\mathcal{L} \in \mathbf{DisjClass}$. Suppose $1 \leq m \leq n$, $1 \leq m' \leq n'$, $n' \leq n$ and $n - m \leq n' - m'$. Then,*

- (a) *$\mathcal{L} \in [m, n]\mathbf{MultEx}$ implies $\mathcal{L} \in [m', n']\mathbf{MultEx}$.*
- (b) *$\mathcal{L} \in [m, n]\mathbf{MultBc}$ implies $\mathcal{L} \in [m', n']\mathbf{MultBc}$.*

Now we will demonstrate the complexity advantages of $[1, m]\mathbf{MultEx}$ -learnability over $[m, m]\mathbf{MultEx}$ -learnability (for classes of languages which are learnable under both criteria). First we need a technical proposition.

Proposition 13. *Let $S = \{e \mid \text{card}(W_e) \leq e, \text{ and } \text{card}(W_e) \text{ is odd}\}$. Then, there does not exist an $n \in N$ and a recursive function h such that:*

- (a) *for all e , $\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) \leq n$, and*
- (b) *$\lim_{t \rightarrow \infty} h(e, t) = 1$, if $e \in S$; $\lim_{t \rightarrow \infty} h(e, t) = 0$, otherwise.*

Proof. Suppose $n \in N$ and recursive h are given such that (a) holds. Then, by Kleene's Recursion Theorem [13] there exists an $e > n$ such that W_e may be described as follows. $m \in W_e$ iff $[\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) \geq m + 1]$ or $[\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) = m \text{ and } \lim_{t \rightarrow \infty} h(e, t) = 0 \text{ and } m \text{ is even}]$ or $[\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) = m \text{ and } \lim_{t \rightarrow \infty} h(e, t) = 1 \text{ and } m \text{ is odd}]$.

As $\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) \leq n$, it is easy to verify that W_e does not contain any $x > n$. Thus $\text{card}(W_e) \leq n + 1 \leq e$. Furthermore, if $\text{card}(\{t \mid h(e, t) \neq h(e, t + 1)\}) = r$, then W_e contains all $x < r$, W_e contains no $x > r$, and

- (i) if $\lim_{t \rightarrow \infty} h(e, t) = 0$, then \mathbf{M} contains r iff r is even (thus, W_e contains odd number of elements).
- (ii) if $\lim_{t \rightarrow \infty} h(e, t) = 1$, then \mathbf{M} contains r iff r is odd (thus, W_e contains even number of elements).

It follows that (b) does not hold. ■

If $\mathbf{M}(T[r]) \neq \mathbf{M}(T[r + 1])$, then we say that \mathbf{M} made a mind change at $T[r + 1]$.

Theorem 14. *Suppose $2 \leq m$. There exists a class \mathcal{L} in $[m, m]\mathbf{MultEx}$, such that*

- (a) *For any $n \in N$, for any \mathbf{M} which $[m, m]\mathbf{MultEx}$ -identifies \mathcal{L} , \mathbf{M} makes $\geq n$ mind changes on a text for $L_1 \cup \dots \cup L_m$, for some L_1, \dots, L_m in the class.*
- (b) *Some \mathbf{M} $[1, m]\mathbf{MultEx}$ -identifies \mathcal{L} using no mind changes.*

Proof. Let $\mathcal{L}_e^1 = \{\langle e, 2x \rangle, \langle e, 2x + 1 \rangle \mid x < m\}$.

Let $\mathcal{L}_e^2 = \{\langle e, 2x + 1 \rangle, \langle e, 2x + 2 \rangle \mid x < m - 1\} \cup \{\langle e, 2m - 1 \rangle, \langle e, 0 \rangle\}$.

Note that, for all e , $\bigcup_{L \in \mathcal{L}_e^1} L = \bigcup_{L \in \mathcal{L}_e^2} L$.

Let $S = \{e \mid \text{card}(W_e) \leq e, \text{ and } \text{card}(W_e) \text{ is odd}\}$.

Let $\mathcal{L} = \bigcup_{e \in S} \mathcal{L}_e^1 \cup \bigcup_{e \notin S} \mathcal{L}_e^2$.

It is easy to verify that $\mathcal{L} \in [m, m]\mathbf{MultEx}$ (one determines in the limit the $2m$ elements that constitute the m input languages, and whether $e \in S$ or not, for each e such that, for some y , $\langle e, y \rangle$ belongs to the input text. This information is enough to determine the individual languages which constitute the input).

Furthermore, $\mathcal{L} \in [1, m]\mathbf{MultEx}$ via a learner which makes no mind changes (one just needs to wait until at least $2m$ elements appear in the input. At which point the learner can output the grammar which enumerates these $2m$ elements).

However, a learner which $[m, m]\mathbf{MultEx}$ -identifies \mathcal{L} using at most n mind changes, also gives us a method to decide S limit effectively using at most n mind changes. An impossible task by Proposition 13. ■

4. Some extensions

In this section we consider learnability of unions in groups under additional constraint: the number of languages in learned groups may be limited.

Definition 15. (a) We say that $\mathbf{M} [m, s, n]\mathbf{MultEx}$ -identifies \mathcal{L} , iff for all distinct languages L_1, \dots, L_n in \mathcal{L} , for all texts T for $L_1 \cup \dots \cup L_n$, there exist i_1, \dots, i_k , ($k \geq m$) such that $\mathbf{M}(T)$ converges on T to the index for the set $\{i_1, \dots, i_k\}$ and there exists a partition G_1, \dots, G_k of $\{1, 2, \dots, n\}$ such that

- (i) each G_i is non-empty and of size at most s ,
 - (ii) for $1 \leq j \leq k$, i_j is a grammar for $\bigcup_{r \in G_j} L_r$.
- (b) $[m, s, n]\mathbf{MultEx} = \{\mathcal{L} \mid (\exists \mathbf{M})[\mathbf{M} [m, s, n]\mathbf{MultEx}\text{-identifies } \mathcal{L}]\}$.

One can similarly define $[m, s, n]\mathbf{MultBc}$. Note that $[m, n]\mathbf{MultEx}$ is the same as $[m, s, n]\mathbf{MultEx}$, for any $s \geq n - m + 1$. Thus, often when $s \geq n - m + 1$, we just use $[m, \infty, n]\mathbf{MultEx}$ to show that there is no restriction on individual groups except the one forced by values of m, n .

We first consider some results which follow from the results/proofs of Theorems in the previous section. As a corollary to Theorem 6, we get

Corollary 16. Suppose $1 \leq n < n'$ and $1 \leq m' \leq n'$. Then $[n, 1, n]\mathbf{MultEx} - [m', \infty, n']\mathbf{MultBc} \neq \emptyset$.

Corollary 17. Suppose $1 \leq r' < r$, $1 \leq m \leq n - r + 1$, $1 \leq m' \leq n' - r' + 1$. Then, $[m, r, n]\mathbf{MultEx} - [m', r', n']\mathbf{MultBc} \neq \emptyset$.

Proof. By Proof of Theorem 7, we have that $[n - r + 1, r, n]\mathbf{MultEx} - [m', r', n']\mathbf{MultBc} \neq \emptyset$. Corollary follows. ■

Proof of Theorem 11 essentially shows the following theorem also.

Theorem 18. Suppose $n - m \leq n' - m'$, $r \leq r'$ and $n' \leq n$. Suppose $\mathcal{L} \in \mathbf{DisjClass}$. Then,

- (a) $[m, r, n]\mathbf{MultEx} \subseteq [m', r', n']\mathbf{MultEx}$.
- (b) For $n' \leq n$, $[m, r, n]\mathbf{MultEx} \subseteq [1, n', n']\mathbf{MultEx}$.

We now give a result which solves the remaining cases for the relationship between different $[m, r, n]\mathbf{MultEx}$ -learnability for classes in $\mathbf{DisjClass}$. Suppose $n' \leq n$, $r \leq r'$ and $n' - m' < n - m$: since the other cases have been handled in corollaries and theorem above. For classes in $\mathbf{DisjClass}$, the next result, Theorem 20, shows when $[m, r, n]\mathbf{MultEx}$ can be simulated by $[m', r', n']\mathbf{MultEx}$. This depends on a complex relationship between m, r, n and m', r', n' , which we express as the following property.

Intuitively, the property says that if we distribute n balls in at least $k \geq m$ non-empty boxes, such that each box has at most r balls, and then take out $n - n'$ balls, then at least m' boxes will remain non-empty.

Definition 19. Suppose $1 \leq m' \leq n'$, $1 \leq m \leq n$, $1 \leq n' \leq n$, $1 \leq r \leq r'$ and $n' - m' < n - m$.

We say that $Prop(m, r, n, m', r', n')$ holds iff for any $k \geq m$, $a_1, \dots, a_k, b_1, \dots, b_k$, if (A) to (D) hold, then (E) also holds.

- (A) $\sum_{1 \leq i \leq k} b_i = n - n'$,
- (B) $\sum_{1 \leq i \leq k} a_i = n$,
- (C) for $1 \leq i \leq k$, $0 \leq b_i \leq a_i \leq r$,
- (D) for $1 \leq i \leq k$, $1 \leq a_i \leq r$,
- (E) $\text{card}(\{i \mid 1 \leq i \leq k, a_i > b_i\}) \geq m'$.

Theorem 20. Suppose $n' \leq n, r \leq r'$ and $n' - m' < n - m$.

(a) If $\text{Prop}(m, r, n, m', r', n')$ holds then for any class $\mathcal{L} \in \text{DisjClass}$,

$$\mathcal{L} \in [m, r, n]\text{MultEx} \text{ implies } \mathcal{L} \in [m', r', n']\text{MultEx}.$$

(b) If $\text{Prop}(m, r, n, m', r', n')$ does not hold then

$$[m, r, n]\text{MultEx} - [m', r', n']\text{MultEx} \neq \emptyset.$$

Moreover, this separation can be witnessed by a class in **DisjClass**.

Proof. (a) Suppose $\text{Prop}(m, r, n, m', r', n')$ holds. Let \mathbf{M} be $[m, r, n]\text{MultEx}$ -learner for $\mathcal{L} \in \text{DisjClass}$.

If \mathcal{L} is finite, then the simulation is trivial. So assume \mathcal{L} is infinite. \mathbf{M}' behaves as follows.

Given any text T , \mathbf{M}' finds (i) $g_1, \dots, g_{n-n'}$ such that $W_{g_1}, W_{g_2}, \dots, W_{g_{n-n'}} \in \mathcal{L}$ and input text T does not contain any element of $\bigcup_{1 \leq i \leq n-n'} W_{g_i}$, and (ii) a grammar g for content(T) (note that by Proposition 10 this can be done in the limit). \mathbf{M}' runs \mathbf{M} on a text T' for content(T) $\cup \bigcup_{1 \leq i \leq n-n'} W_{g_i}$.

Suppose \mathbf{M} converges on T' to the index for set $\{i_1, \dots, i_k\}$. Then, \mathbf{M}' forms grammars $\{i'_1, \dots, i'_k\}$, such that for $1 \leq w \leq k$, $W_{i'_w} = W_{i_w} \cap W_g$. Then, \mathbf{M}' outputs the index for the set $\{i'_w \mid 1 \leq w \leq k, W_{i'_w} \neq \emptyset\}$.

Using definition of Prop, it is easy to verify that \mathbf{M}' $[m', r', n']\text{MultEx}$ -identifies \mathcal{L} .

(b) Suppose $\text{Prop}(m, r, n, m', r', n')$ does not hold. Let $a_1, \dots, a_k, b_1, \dots, b_k$, be such that (A) to (D) are satisfied but (E) does not hold in Definition 19.

The diagonalizing class \mathcal{L} will consist of $L_{j,i,w}$, for $w < a_i, j \in N, 1 \leq i \leq k$.

$L_{j,i,w}$ will satisfy the following properties. For the following, let code_j be the index for a set of grammars for the languages in $\{L_{j',i,w} \mid j' < j, 1 \leq i \leq k, w < a_i\}$.

(P1) $L_{j,i,w} \subseteq \{(j, \text{code}_j, i, x) \mid x < 2r\}$.

(P2) $\text{card}(L_{j,i,w}) \geq 2$.

(P3) $L_{j,i,w}$ are disjoint for different values of w .

We first claim that \mathcal{L} is in $[m, r, n]\text{MultEx}$, irrespective of what the exact chosen $L_{j,i,w}$ are as long as the above properties are satisfied.

On any input text, a learner can first determine the largest j and corresponding code_j such that, $\langle j, \text{code}_j, i, x \rangle$ belongs to the input text, for some i, x . Now the learner can determine (in the limit) grammars for:

(i) any language from \mathcal{L} which is of form $L_{j',i',w'}$, for some $j' < j$, and $L_{j',i',w'} \subseteq \text{content}(T)$ (this can be done using code_j).

(ii) $\text{content}(T) \cap \{(j, \text{code}_j, i, x) \mid x < 2r\}$, for each $i \in N$, such that $\text{content}(T) \cap \{(j, \text{code}_j, i, x) \mid x < 2r\} \neq \emptyset$.

The learner can then, in the limit, converge to the index for set of grammars obtained in (i) and (ii). It immediately follows that the learner $[m, r, n]\text{MultEx}$ -identifies \mathcal{L} .

We now show that for appropriate choice of $L_{j,i,w}$, $\mathcal{L} \notin [m', r', n']\text{MultEx}$.

Suppose by way of contradiction that, for some j , one cannot choose appropriate $L_{j,i,w}, 1 \leq i \leq k, w < a_i$ such that \mathbf{M}_j fails to $[m', r', n']\text{MultEx}$ -identify $\{L_{j,i,w} \mid 1 \leq i \leq k, w < a_i\}$. Let j be least such number.

Define $L_{j,i,w}$ for $1 \leq i \leq k, w < b_i$, as $\{(j, \text{code}_j, i, 2w), \langle j, \text{code}_j, i, 2w + 1 \rangle\}$. Give as input a text T to \mathbf{M}_j , where $\text{content}(T) = \{(j, \text{code}_j, i, x) \mid a_i > b_i \text{ and } 2b_i \leq x < 2r\}$. (Note that code_j is determined by languages chosen for $L_{j',i,w}$ for $j' < j$.)

Suppose \mathbf{M}_j on T converges to the index set $\{s_1, \dots, s_{k'}\}$. Now $W_{s_w}, 1 \leq w \leq k'$ must be non-empty, and $k' \geq m'$ (otherwise clearly, one can choose appropriate $L_{j,i,w}$ such that \mathbf{M}_j fails to $[m', r', n']\text{MultEx}$ -identify $\{L_{j,i,w} \mid 1 \leq i \leq k, w < a_i\}$). Furthermore note that there cannot be an $i, 1 \leq i \leq k, a_i > b_i$, such that for two distinct w and w' , $W_{s_w}, W_{s_{w'}}$ intersect with $\{(j, \text{code}_j, i, x) \mid 2b_i \leq x < 2r\}$ (since otherwise, one may take L_{j,i,b_i} to contain one element from both $W_{s_w}, W_{s_{w'}}$, and other $L_{j,i,w}$ appropriately, to contradict \mathbf{M}_j $[m', r', n']\text{MultEx}$ -identifying \mathcal{L}). Now for $i, 1 \leq i \leq k$, such that $a_i > b_i$, for $b_i \leq w < a_i$, define $L_{j,i,w}$ such that each $L_{j,i,w}$ contains at least 2 elements and $\bigcup_{b_i \leq w < a_i} L_{j,i,w} = \{(j, \text{code}_j, i, x) \mid 2b_i \leq x < 2r\}$.

Now each i in $\{i \mid a_i > b_i\}$ can be mapped to a $w, 1 \leq w \leq k'$ such that W_{s_w} contains $\{(j, \text{code}_j, i, x) \mid 2b_i \leq x < 2r\}$. Thus, $\{i \mid a_i > b_i\} \geq k' \geq m'$. A contradiction.

This completes the proof of the theorem. ■

Corollary 21. Suppose $n - m > n' - m'$, and $n' \geq n - m + \lceil \frac{n-m}{r-1} \rceil$. Then, $[m, r, n]\text{MultEx} - [m', \infty, n']\text{MultEx} \neq \emptyset$.

Proof. Let $a_1, a_2, \dots, a_m, b_1, \dots, b_m$, be defined as follows.

$$a_i = r, b_i = 0, \text{ for } 1 \leq i \leq \lfloor \frac{n-m}{r-1} \rfloor.$$

If $\frac{n-m}{r-1}$ is not an integer, then let $a_{\lceil \frac{n-m}{r-1} \rceil} = 1 + (n-m) - (\lfloor \frac{n-m}{r-1} \rfloor * (r-1))$, $b_{\lceil \frac{n-m}{r-1} \rceil} = 0$.

$$a_i = 1, b_i = 0, \text{ for } \lceil \frac{n-m}{r-1} \rceil < i \leq m - (n - n').$$

$$a_i = 1, b_i = 1, \text{ for } m - (n - n') < i \leq m.$$

Now, as $m - (n - n') < m'$, and $n' \geq n - m + \lceil \frac{n-m}{r-1} \rceil$, we immediately have that $Prop(m, r, n, m', r', n')$ cannot hold (as there are $< m'$ w 's in $\{1, \dots, m\}$ such that $a_w - b_w > 0$). Corollary follows. ■

5. Conclusions

In this paper we explored relationships between a more coarse and a more refined classification from the standpoint of computability. Some of our main results (for example, Theorems 11, 18 and 20(a)) worked on the assumption that underlying targets of classification were pairwise distinct. While it is true for many cognitive classification tasks, there are classification problems where such an assumption cannot be made. For example, when one wants to classify all classical music pieces as being in major or minor, this can probably be done, relatively easily, for everything written within Western musical tradition before the 20th century. However, this has changed by impressionism, introduction of the atonal scale, etc. in the 20th century. Many pieces written by contemporary composers often alternate between major and minor several times, which makes the task of classifying such pieces much harder. To model a situation of this kind, one has to lift the requirement of classification targets being pairwise distinct, or, at least, replace it by a much weaker requirement allowing intersections of classification targets being finite. It would be interesting to explore the issues discussed in the paper in such a setting. We leave it for a future research.

Acknowledgements

Preliminary version of this paper appeared as [6]. We thank the anonymous referees of ALT 2005 and this journal for the several helpful comments. The first author was supported in part by NUS grant numbers R252-000-127-112, R252-000-212-112 and R252-000-308-112.

References

- [1] J. Bärzdīņš, Two theorems on the limiting synthesis of functions, in: Theory of Algorithms and Programs, vol. 1, Latvian State University, 1974, pp. 82–88 (in Russian).
- [2] M. Blum, A machine-independent theory of the complexity of recursive functions, Journal of the ACM 14 (1967) 322–336.
- [3] J. Case, C. Lynes, Machine inductive inference and language identification, in: M. Nielsen, E.M. Schmidt (Eds.), Proceedings of the 9th International Colloquium on Automata, Languages and Programming, in: Lecture Notes in Computer Science, vol. 140, Springer-Verlag, 1982, pp. 107–115.
- [4] J. Case, C. Smith, Comparison of identification criteria for machine inductive inference, Theoretical Computer Science 25 (1983) 193–220.
- [5] E.M. Gold, Language identification in the limit, Information and Control 10 (1967) 447–474.
- [6] S. Jain, E. Kinber, Learning multiple languages in groups, in: Sanjay Jain, Hans Ulrich Simon, Etsuji Tomita (Eds.), Algorithmic Learning Theory: Sixteenth International Conference, ALT' 2005, in: Lecture Notes in Artificial Intelligence, vol. 3734, Springer-Verlag, 2005, pp. 256–268.
- [7] Sanjay Jain, Yen Kaow Ng, Tiong Seng Tay, Learning languages in a union, Journal of Computer and System Sciences 73 (2007) 89–108.
- [8] M. Machtey, P. Young, An Introduction to the General Theory of Algorithms, North-Holland, New York, 1978.
- [9] D. Osherson, M. Stob, S. Weinstein, Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists, MIT Press, 1986.
- [10] D. Osherson, S. Weinstein, Criteria of language learning, Information and Control 52 (1982) 123–138.
- [11] S. Pinker, Formal models of language learning, Cognition 7 (1979) 217–283.
- [12] H. Rogers, Gödel numberings of partial recursive functions, Journal of Symbolic Logic 23 (1958) 331–341.
- [13] H. Rogers, Theory of Recursive Functions and Effective Computability, McGraw-Hill, 1967 (Reprinted by MIT Press in 1987).
- [14] K. Wexler, P. Culicover, Formal Principles of Language Acquisition, MIT Press, 1980.