Characterizing language identification in terms of computable numberings

Sanjay Jain

Department of Information Systems and Computer Science

National University of Singapore

Singapore 0511

Republic of Singapore

sanjay@iscs.nus.sg

Arun Sharma

School of Computer Science and Engineering

The University of New South Wales

Sydney, NSW 2052

Australia

arun@cse.unsw.edu.au

March 11, 2007

Abstract

Identification of programs for computable functions from their graphs and identification of grammars (r. e. indices) for recursively enumerable languages from positive data are two extensively studied problems in the recursion theoretic framework of inductive inference.

In the context of function identification, Freivalds, Kinber, and Wiehagen have shown that only those collections of functions, $\mathcal{S}$, are identifiable in the limit for which there exists a 1–1 computable numbering ψ and a discrimination function d such that (a) for each $f \in \mathcal{S}$, the number of indices i such that $\psi_i = f$ is exactly one and (b) for each $f \in \mathcal{S}$, there are only finitely many indices i such that f and ψ_i agree on the first $d(i)$ arguments.

A similar characterization for language identification in the limit has turned out to be difficult. A partial answer is provided in this paper. Several new techniques are introduced which have found use in other investigations on language identification.

1 Introduction

Recursive function theory provides a suitable framework for theoretical studies in machine learning. Identification in the limit of programs for computable functions from their graphs and identification in the limit of grammars (r. e. indices) for recursively enumerable languages from positive data are two extensively studied problems in this framework. We informally describe these problems.

A machine **M** is said to identify a computable function f in the limit just in case **M**, fed a graph of f, one ordered pair at a time, conjectures a sequence of computer programs that converges to a correct program for f. A collection of functions, $\mathcal{S}$, is said to be identifiable just in case there exists a machine that identifies each function in $\mathcal{S}$.

A machine **M** is said to identify a grammar[1] for a recursively enumerable language L just in case **M**, fed all (and only) the elements of L in any order, conjectures a sequence of grammars that converges to a correct grammar for L. A collection of languages, $\mathcal{L}$, is said to be identifiable just in case there exists a machine that identifies each language in $\mathcal{L}$. Studies about language identification turn out to be more complex than studies about function identification because the learning machine is only told about what is in the language (positive information) and is not told about what is not in the language (negative information). It should be noted that in the context of function identification, a machine can eventually determine if an ordered pair

[1]By a grammar for an r.e. language, we mean an acceptor.

belongs or does not belong to the function.

Freivalds, Kinber, and Wiehagen [7] gave an interesting characterization of identifiable collections of functions in terms of 1-1 computable numberings. We first present some notation about computable numberings.

A computable numbering is a computable function of two arguments. Suppose ψ is a computable numbering. Then $\lambda x.\psi(i,x)$ is often denoted by ψ_i. Intuitively, ψ_i denotes the the partial function computed by ψ-program i or equivalently the i-th partial function in the numbering ψ.

A computable numbering ψ is said to be 1-1 just in case $(\forall i,j \mid i \neq j)[\psi_i \neq \psi_j]$. W_i^ψ denotes the domain of ψ_i.

Freivalds, Kinber, and Wiehagen showed that only those collections of functions, $\mathcal{S}$, are identifiable in the limit for which there exists a 1-1 computable numbering ψ and a discrimination function d such that (a) for each $f \in \mathcal{S}$, the number of indices i such that $\psi_i = f$ is exactly one and (b) for each $f \in \mathcal{S}$, there are only finitely many indices i such that f and ψ_i agree on the first $d(i)$ arguments.

A similar characterization, for language identification in the limit, has turned out to be difficult. In this paper a partial answer, informally described below, is provided.

A collection of r. e. languages, $\mathcal{L}$, is identifiable in the limit just in case for some identifiable $\mathcal{L}' \supseteq \mathcal{L}$, there exists a computable numbering ψ and a discrimination function d such that ψ satisfies the following requirements:

(a) for every infinite language L in $\mathcal{L}'$, the number of indices i such that $W_i^\psi = L$ is exactly one;

(b) for every infinite language L *not* in $\mathcal{L}'$, the number of indices i such that $W_i^\psi = L$ is zero;

(c) for every finite language L in $\mathcal{L}'$, the number of indices i such that $W_i^\psi = L$ is at least one; and

(d) for every finite language L, the number of indices i such that $W_i^\psi = L$ is finite.

The definition of the discrimination function d turns out to be somewhat more complex. Using similar techniques, we also give a characterization of language identification from additional information.

An important contribution of this paper is that the techniques introduced have been found to be useful in other investigations about language identification; for example, see [5, 15].

As already acknowledged, the characterizations of language identification presented here were motivated by related results of Freivalds, Kinber, and Wiehagen [7] in the context of function identification. The only other characterization of language identification from positive data that we know appears in [14]. However, a number of characterizations have appeared in the literature for identification from positive data of indexed families of recursive languages; in particular we would like to direct the reader to the work of Angluin [1], Kapur [16], Lange and Zeugmann [18, 19], Lange, Zeugmann, and Kapur [20], Mukouchi, [22, 21], and Mukouchi and Arikawa [23].

We now proceed formally. In Section 2, we present some recursion theoretic notation and in Section 3, we present notions from inductive inference literature. Section 4 contains the main characterization result and Section 5 contains a characterization of language identification with additional information.

2 Notation

Any unexplained recursion theoretic notation is from [26]. The symbol N denotes the set of natural numbers, $\{0, 1, 2, 3, \ldots\}$. The symbol N^+ denotes the set of positive natural numbers, $\{1, 2, 3, \ldots\}$. Unless otherwise specified, a, b, i, j, k, l, m, n, r, s, t, x, y, z, with or without decorations[2], range over N. The symbol N_m denotes the set $\{x \in N \mid x \leq m\}$. Symbols

[2] Decorations are subscripts, superscripts and the like.

$\emptyset$, $\subseteq$, $\subset$, $\supseteq$, and $\supset$ denote empty set, subset, proper subset, superset, and proper superset, respectively. The symbols B, C, D, S, X, with or without decorations, range over sets. We denote by D_x the finite set whose canonical index is x [26]. According to this convention $D_0 = \emptyset$. The cardinality of a finite set S is denoted by $\text{card}(S)$. The maximum and minimum of a set are denoted by $\max(\,)$, $\min(\,)$, respectively. By convention $\max(\emptyset) = 0$ and $\min(\emptyset) = \infty$. For any set A, 2^A denotes the power set of A.

The symbols p, q range over partial recursive functions and the symbols c, d, f, g, range over total recursive functions. The set of all total recursive functions of one variable is denoted by $\mathcal{R}$. For $n \in N^+$, $\mathcal{R}^n$ denotes the set of total recursive functions of n variables. For a partial recursive function p, $\text{domain}(p)$ denotes the domain of p and $\text{range}(p)$ denotes the range of p. We write $p(x)\downarrow$ just in case $x \in \text{domain}(p)$, otherwise we write $p(x)\uparrow$.

A language is a subset of N. L ranges over recursively enumerable (r.e.) languages. The collection of all r.e. languages is denoted by $\mathcal{E}$. $\mathcal{L}$ and $\mathcal{C}$, with or without decorations, range over subsets of $\mathcal{E}$.

A programming system (also called computable numbering) is a partial computable function from N^2 to N. The symbol ψ ranges over computable numberings. In this paper, by numbering we mean computable numbering. 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 complexity measure for ψ. We say that numbering ψ is reducible to numbering ψ' (written $\psi \leq \psi'$) iff there exists a recursive function h such that $(\forall i)[\psi_i = \psi'_{h(i)}]$. In this case we say that h witnesses that $\psi \leq \psi'$. An acceptable numbering (acceptable programming system) is a computable numbering to which every computable numbering can be reduced. The symbol φ denotes a standard acceptable programming system (also referred to as standard acceptable numbering) [25, 26]. W_i^ψ denotes the recursively enumerable set $\{x \mid x \in \text{domain}(\psi_i)\}$. We say that i is a ψ-grammar for W_i^ψ. If i is a φ-grammar for L, then we sometimes just say that i is a grammar for L. $W_{i,s}^\psi$ denotes the set $\{x \mid x \leq s \wedge \Psi(x) \leq s\}$.

$\mathcal{C}_\psi = \{W_i^\psi \mid i \in N\}$. For ease of notation, we may omit φ, the standard acceptable programming system, from W_i^φ. $\mathrm{MinProg}_\psi(f) = \min(\{i \mid \psi_i = f\})$. $\mathrm{MinProg}_\psi(f)$, thus denotes the minimal program for f, if any, in the ψ programming system. $\mathrm{MinGram}_\psi(L) = \min(\{i \mid W_i^\psi = L\})$, denotes the minimal grammar for L in the ψ programming system.

$\langle \cdot, \cdot \rangle$ stands for an arbitrary, one to one, computable mapping from N^2 onto N. [26]. Corresponding projection functions are π_1 and π_2. $(\forall i, j \in N)[\pi_1(\langle i, j\rangle) = i$ and $\pi_2(\langle i, j\rangle) = j$ and $\langle \pi_1(x), \pi_2(x)\rangle = x]$. Similarly, $\langle i_1, i_2, \ldots, i_n\rangle$ denotes a computable one to one mapping from N^n onto N. *Remark:* We sometimes abuse notation and write, $\langle \ldots, S, \ldots\rangle$ to mean $\langle \ldots, x, \ldots\rangle$, where $D_x = S$. This is for simplicity of presentation and it will be made clear when we resort to such an interpretation.

$\overset{\infty}{\forall}$ and $\overset{\infty}{\exists}$ respectively denote 'for all but finitely many' and 'there exist infinitely many'.

3 Preliminaries

In this section, we briefly describe notions and results from formal language learning theory literature. We first introduce a notion that facilitates discussion about elements of a language being fed to a machine.

A *sequence* σ is a mapping from an initial segment of N into $(N \cup \{\#\})$. The *content* of a sequence σ, denoted content(σ), is the set of natural numbers in the range of σ. The *length* of σ, denoted by $|\sigma|$, is the number of elements in σ.

Intuitively, #'s represent pauses in the presentation of data. We let σ and τ, with or without decorations, range over finite sequences. For $n \leq |\sigma|$, $\sigma[n]$ denotes the finite initial segment of σ with length n. The result of concatenating τ onto the end of σ is denoted by $\sigma \diamond \tau$. We say that $\sigma \subseteq \tau$ just in case σ is an initial segment of τ, that is, $|\sigma| \leq |\tau|$ and $\sigma = \tau[|\sigma|]$. SEQ denotes the set of all finite sequences. The set of all finite sequences of natural numbers and #'s, SEQ, can be coded onto N. This coding assigns a canonical index to each member of SEQ. We will abuse the notation somewhat, as a reference to σ will mean both the sequence and its canonical index.

Definition 1 A *language learning machine* computes a computable mapping from SEQ into N.

We let $\mathbf{M}$, with or without decorations, range over learning machines.

A *text* is a mapping from N into $(N \cup \{\#\})$. The *content* of a text T, denoted content(T), is the set of natural numbers in the range of T. A text T is *for* L iff content$(T) = L$.

Intuitively, a text for a language is an enumeration or sequential presentation of all the objects in the language with the #'s representing pauses in the listing or presentation of such objects. For example, the only text for the empty language is just an infinite sequence of #'s.

We let T, with or without decorations, range over texts. $T[n]$ denotes the finite initial sequence of T with length n. The reader should note that $T[n]$ does not contain $T(n)$, the $(n+1)^{th}$ element of T. Hence, domain$(T[n]) = \{x \mid x < n\}$. We say that $\sigma \subset T$ just in case σ is an initial segment of T, that is, $\sigma = T[|\sigma|]$.

We next present Gold's [11] criteria for successful identification of languages. First, we spell out what it means for a learning machine on a text to converge in the limit.

Suppose $\mathbf{M}$ is a learning machine and T is a text. $\mathbf{M}(T)\downarrow$ (read: $\mathbf{M}(T)$ *converges*) $\iff$ $(\exists i)(\overset{\infty}{\forall} n)\ [\mathbf{M}(T[n]) = i]$. If $\mathbf{M}(T)\downarrow$, then $\mathbf{M}(T)$ is defined as the unique i such that $(\overset{\infty}{\forall} n)[\mathbf{M}(T[n]) = i]$; otherwise, we say that $\mathbf{M}(T)$ *diverges* (written: $\mathbf{M}(T)\uparrow$).

Definition 2 [11]

(a) $\mathbf{M}$ $\mathbf{TxtEx}$-*identifies* L (written: $L \in \mathbf{TxtEx}(\mathbf{M})$) $\iff$ $(\forall$ texts T for $L)(\exists i)[W_i = L \wedge \mathbf{M}(T)\downarrow \wedge \mathbf{M}(T) = i]$.

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

In the above **TxtEx** stands for explanatory identification from texts. The notation in the above definition is from [6]. For a survey of work on Gold's paradigm of language identification, the reader is referred to [2, 24, 17, 4].

Our results build on the notion of stabilizing and locking sequences for learning machines on languages and also employ the notions of order independent and rearrangement independent learning machines. We now present these notions.

We first define order independence (slightly differently from that in [3]) and rearrangement independence.

Definition 3 (a) A learning machine $\mathbf{M}$ is *order-independent* just in case $(\forall$ texts $T, T')[\text{content}(T) = \text{content}(T') \Rightarrow \mathbf{M}(T) = \mathbf{M}(T')]$.

(b) [9, 10] A learning machine $\mathbf{M}$ is *rearrangement-independent* just in case $(\forall \sigma_1, \sigma_2)[[\text{content}(\sigma_1) = \text{content}(\sigma_2)\ \wedge\ |\sigma_1| = |\sigma_2|] \Rightarrow \mathbf{M}(\sigma_1) = \mathbf{M}(\sigma_2)]$.

We next describe the technical notions of stabilizing and locking sequences.

Definition 4 (a) [9, 10] σ is a **TxtEx**-*stabilizing sequence for* $\mathbf{M}$ *on* L just in case $\text{content}(\sigma) \subseteq L$ and $(\forall \sigma')[[\sigma \subseteq \sigma'\ \wedge\ \text{content}(\sigma') \subseteq L] \Rightarrow \mathbf{M}(\sigma') = \mathbf{M}(\sigma)]$.

(b) [3, 24] σ is a **TxtEx**-*locking sequence for* $\mathbf{M}$ *on* L just in case σ is a **TxtEx**-stabilizing sequence for $\mathbf{M}$ on L and $W_{\mathbf{M}(\sigma)} = L$.

We often refer to **TxtEx**-locking sequence by just locking sequence. The following lemma due to L. Blum and M. Blum is a useful tool for our purposes.

Lemma 1 [3, 24] *If* $\mathbf{M}$ **TxtEx**-*identifies* L*, then there is a* **TxtEx**-*locking sequence for* $\mathbf{M}$ *on* L.

The following lemma due to M. Fulk relates order independence, rearrangement independence, and locking sequences.

Lemma 2 [9, 10] *From any learning machine* $\mathbf{M}$ *one may effectively construct* $\mathbf{M}'$ *such that all the following conditions hold.*

(a) $\mathbf{TxtEx}(\mathbf{M}) \subseteq \mathbf{TxtEx}(\mathbf{M}')$.

(b) $\mathbf{M}'$ *is order independent.*

(c) $\mathbf{M}'$ *is rearrangement independent.*

(d) For every $L \in \mathcal{E}$*, if for some text* T *for* L*,* $\mathbf{M}'(T)\downarrow$ *and* $W_{\mathbf{M}'(T)} = L$*, then* $\mathbf{M}'$ $\mathbf{TxtEx}$*-identifies* L*.*

(e) If there is a $\mathbf{TxtEx}$*-locking sequence for* $\mathbf{M}'$ *on* L*, then* $L \in \mathbf{TxtEx}(\mathbf{M}')$*.*

(f) If $L \in \mathbf{TxtEx}(\mathbf{M}')$*, then all texts for* L *contain a* $\mathbf{TxtEx}$*-locking sequence for* $\mathbf{M}'$ *on* L*.*

If a collection of r.e. languages, $\mathcal{L}$, is $\mathbf{TxtEx}$-identified by a machine $\mathbf{M}$, then using Lemma 2, we can say without loss of generality that $\mathcal{L}$ is $\mathbf{TxtEx}$-identified by a rearrangement independent and order independent machine $\mathbf{M}'$. Thus we will usually be dealing with rearrangement independent machines only, and often refer to a sequence σ by $\langle x, l\rangle$ (or, abusing notation slightly, as $\langle \text{content}(\sigma), l\rangle$) where $D_x = \text{content}(\sigma)$ and $l = |\sigma|$.

Lemma 1 states that if $\mathbf{M}$ $\mathbf{TxtEx}$-identifies L, then there is a $\mathbf{TxtEx}$-locking sequence for $\mathbf{M}$ on L. For rearrangement independent machines, we can thus define the least locking sequence as the least number $\langle x, l\rangle$, such that $\langle D_x, l\rangle$ is a locking sequence for $\mathbf{M}$ on L (note that $\langle D_x, l\rangle$ represents the sequence σ such that $D_x = \text{content}(\sigma)$ and $l = |\sigma|$).

4 A Characterization of TxtEx

In this section we characterize $\mathbf{TxtEx}$ in terms of computable numberings.

Definition 5 Let $a \in N$. A finite set D is said to be *a-consistent* with $L \in \mathcal{E}$ just in case $D \subseteq L$ and $D \cap N_a = L \cap N_a$.

Intuitively, $D \subseteq L$ is a-consistent with L iff for each $i \leq a$, $i \in L \iff i \in D$.

Definition 6 A finite set D is *a-partial consistent* with $L \in \mathcal{E}$ just in case D is $\min(\{\max(D), a\})$-consistent with L.

Definition 7 ψ is *effectively subdiscrete* for $\mathcal{L}$ just in case the following conditions are satisfied.

1. $\mathcal{L} \subseteq \mathcal{C}_\psi$.
2. $(\forall L \in \mathcal{C}_\psi)[L$ is infinite $\Rightarrow \mathrm{card}(\{i \mid W_i^\psi = L\}) = 1]$.
3. $(\forall L \notin \mathcal{C}_\psi)[L$ is infinite $\Rightarrow \mathrm{card}(\{i \mid W_i^\psi = L\}) = 0]$.
4. $(\forall L \in \mathcal{E})[L$ is finite $\Rightarrow \mathrm{card}(\{i \mid W_i^\psi = L\}) < \infty]$.
5. $\exists d \in \mathcal{R}$ such that both 5a and 5b below hold:

   5a. $(\forall L \in (\mathcal{L} - \{\emptyset\}))(\exists k)[(W_k^\psi = L) \wedge (D_{d(k)} = L \cap N_{\max(D_{d(k)})})]$

   5b. $(\forall L \in (\mathcal{L} - \{\emptyset\}))(\exists n_L \in L)$

   $[\mathrm{card}(\{j \mid [(L \cap N_{n_L}) \subseteq W_j^\psi] \wedge [D_{d(j)}$ is n_L-partial consistent with $L]\}) < \infty]$

Note that for the numbering ψ to be effectively subdiscrete, ψ is nearly 1–1 (it may contain more than one grammar for finite sets)[3]. The recursive function d acts as a discrimination function. Consider any $L \in \mathcal{L} - \{\emptyset\}$. At most finitely many grammars, j, in the ψ numbering satisfy

$$[(L \cap N_{n_L}) \subseteq W_j^\psi] \wedge [D_{d(j)} \text{ is } n_L\text{-partial consistent with } L]$$

Note that grammar k from clause 5a, does satisfy this constraint. Intuitively, this means that for some ψ grammar k for L, $D_{d(k)}$ contains exactly the elements of L up to $\max(D_{d(k)})$. Also, for each $L \in \mathcal{L}$, there is an $n_L \in L$, such that, for all but finitely many i, if W_i^ψ contains all the elements of L up to n_L, then $D_{d(i)}$ is not n_L-partial consistent with L. Thus, in some sense,

[3] In fact, Clause 3 in Definition 7 is included only to emphasize that the numbering is 1-1 for infinite languages.

$D_{d(i)}$'s act as discriminating sets. To search for a ψ grammar for $L \in \mathcal{L}$, d can be used to narrow down the search to finitely many grammars. We call the numbering *effectively subdiscrete* for this reason.

Definition 8 *(a)* $\mathcal{L}$ is *effectively subdiscrete* $\iff (\exists\psi)[\psi$ is effectively subdiscrete for $\mathcal{L}$].

(b) $\mathbf{Esd} = \{\mathcal{L} \subseteq \mathcal{E} \mid \mathcal{L}$ is effectively subdiscrete$\}$.

The following theorem shows that the classes **TxtEx** and **Esd** are exactly the same.

Theorem 1 $\mathbf{TxtEx} = \mathbf{Esd}$.

PROOF. We first prove that $\mathcal{L} \in \mathbf{TxtEx} \Rightarrow (\exists\psi)[\psi$ is effectively subdiscrete for $\mathcal{L}]$. For ease of presentation, we give a numbering, ψ, which may contain infinitely many grammars for $\emptyset$. This numbering, ψ, can easily be modified to give a numbering, ψ', which contains only finitely many grammars for $\emptyset$. To see this, assume without loss of generality that $\mathrm{card}(\mathcal{L})$ is infinite. We then construct a numbering ψ' from ψ as follows. Consider an enumeration of grammars for the nonempty sets in $\mathcal{C}_\psi$, $i_0, i_1, i_2, i_3, \ldots$, such that each ψ-grammar for a nonempty set appears exactly once. Let $\psi'_j = \psi_{i_j}$. Similarly, the discrimination function d presented below can also be suitably modified for the new numbering as $d'(j) = d(i_j)$. Now, clearly $\mathcal{C}_{\psi'} = \mathcal{C}_\psi - \{\emptyset\}$. A grammar for $\emptyset$ can also be added if $\emptyset \in \mathcal{L}$.

Suppose $\mathcal{L} \subseteq \mathbf{TxtEx}(\mathbf{M})$, where $\mathbf{M}$ is rearrangement independent and order independent. We further assume that for all σ such that $\mathrm{content}(\sigma) = \emptyset$, $W_{\mathbf{M}(\sigma)} = \emptyset$. Note that if $\mathbf{M}$, does not satisfy this property, then we can easily modify $\mathbf{M}$ to satisfy this property. In the sequel whenever finite sets, $S, S_m, S', \ldots$, appear in $\langle\cdot,\cdot,\cdot\rangle$, we will interpret them as canonical indices for $S, S_m, S', \ldots$.

We now describe the idea behind the construction of ψ. To construct ψ, we try to construct exactly one grammar for every language $L \in \mathbf{TxtEx}(\mathbf{M})$ (this is not fully successful). We would also like to ensure some properties for this grammar so that the discrimination function d can be constructed.

To associate one grammar with $L \in \mathbf{TxtEx}(\mathbf{M})$, we use the least locking sequence $\langle S, l\rangle$ for $\mathbf{M}$ on L. Note that for $\langle S, l\rangle$ to be a locking sequence for $\mathbf{M}$ on L, $W_{\mathbf{M}(\langle S,l\rangle)} = L$ (this helps us determine the language L with which a sequence might be associated). In other words, we wish to associate $\langle S, l\rangle$ with $W_{\mathbf{M}(\langle S,l\rangle)}$, if $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$.

Note however that it cannot be determined effectively if $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$. Hence, we need to constrain the languages enumerated by grammars corresponding to $\langle S, l\rangle$, such that $\langle S, l\rangle$ is not the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$. We address this problem in two ways based on the two reasons due to which $\langle S, l\rangle$ may not be the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$.

First, there might be a smaller locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$. For this reason we attach with each $\langle S, l\rangle$ which is the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$, an "evidence" that smaller sequences are not locking sequences for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$. This is done by attaching the set

$$S_m = W_{\mathbf{M}(\langle S,l\rangle)} \cap [\bigcup_{\langle S',l'\rangle \leq \langle S,l\rangle} S']$$

with $\langle S, l\rangle$. Now, for any $\langle S', l'\rangle < \langle S, l\rangle$, to prove that $\langle S', l'\rangle$ is not a locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$, we just need to check that either $S' \not\subseteq S_m$, or there exists an extension $\langle S'', l''\rangle$ of $\langle S', l'\rangle$ such that $S'' \subseteq W_{\mathbf{M}(\langle S,l\rangle)}$, and $\mathbf{M}(\langle S', l'\rangle) \neq \mathbf{M}(\langle S'', l''\rangle)$. This check is r.e. in nature (and is done in Step (1c) in the construction of ψ below). This S_m also helps in defining the discrimination function.

The other reason due to which $\langle S, l\rangle$ may not be the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$ is that $\langle S, l\rangle$ itself may not be a locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$. Also, note that the attaching of S_m with $\langle S, l\rangle$ additionally introduces the need for verifying that S_m is indeed $W_{\mathbf{M}(\langle S,l\rangle)} \cap [\bigcup_{\langle S',l'\rangle \leq \langle S,l\rangle} S']$. These two issues are addressed in Step 2 of the construction of ψ. If $\langle S, l\rangle$ is not a locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$ or $S_m \neq W_{\mathbf{M}(\langle S,l\rangle)} \cap [\bigcup_{\langle S',l'\rangle \leq \langle S,l\rangle} S']$, then we make the grammar associated with $\langle S, S_m, l\rangle$, enumerate a finite language in Step 3. In doing this we take care to satisfy clause 4 in the definition of effectively subdiscrete and also maintain certain other properties useful in describing the discrimination function, d.

Based on the above description, we define the following technical notion that facilitates the description of our proof. Suppose $k = \langle S, S_m, l\rangle$.

(a) We say that k is *nice* if

(i) $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$, and

(ii) $S_m = (W_{\mathbf{M}(\langle S,l\rangle)} \cap N_{m_0})$, where $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$.

(b) We say that k is *nice for L* if k is nice and $W_{\mathbf{M}(\langle S,l\rangle)} = L$.

We now define ψ as follows:

Definition of W_i^ψ

Suppose $i = \langle S, S_m, l\rangle$. Let $j = \mathbf{M}(\langle S, l\rangle)$, and $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$.

(*Remark: In Step 1 below we attempt to check if i satisfies certain properties of being nice. In Step 2 we check for the remaining properties. If i is nice, then we enumerate W_j^φ (i.e., $W_i^\psi = W_j^\varphi$). If i is not nice, then either W_i^ψ is empty, or, in Step 3, we make W_i^ψ equal to a finite set which is different from each $W_{i'}^\psi$ such that i' is not nice.*)

1. Check the following four conditions

(a) $S \subseteq S_m \subseteq W_j^\varphi$.

(b) $\max(S_m) \leq m_0$.

(c) $(\forall \langle S', l'\rangle < \langle S, l\rangle)[(S' \not\subseteq S_m) \ \vee \ (\exists S'', l'')[(S' \subseteq S'' \subseteq W_j^\varphi) \ \wedge \ (l' + \mathrm{card}(S'' - S') \leq l'') \ \wedge \ (\mathbf{M}(\langle S', l'\rangle) \neq \mathbf{M}(\langle S'', l''\rangle))]]$.

Remark: The above step checks that each $\langle S', l'\rangle < \langle S, l\rangle$ is not a locking sequence for $\mathbf{M}$ on W_j^φ.

(d) $(\forall \langle S', l'\rangle)[[(S \subseteq S' \subseteq S_m) \wedge (l + \mathrm{card}(S' - S) \leq l' \leq l + i + \mathrm{card}(S_m))] \Rightarrow [\mathbf{M}(\langle S', l'\rangle) = j]]$.

If any of the above conditions fails to hold, then let $W_j^\psi = \emptyset$.

Remark: Note that if the above conditions hold then it can be verified. If it cannot be verified whether the above conditions hold or not, then, by default, W_j^ψ will be empty. Also note that Condition (d), above partially check if $\langle S, l\rangle$ is a locking sequence for $\mathbf{M}$ on $W_{\mathbf{M}(\langle S,l\rangle)}$, and is used to make the definition of the discrimination function d work and it will be used in the proof of Claim 4.

2. Enumerate more and more elements of $W_j^\varphi - (N_{m_0} - S_m)$ until one of the following two conditions hold

   (a) $W_j^\varphi \cap (N_{m_0} - S_m) \neq \emptyset$, or

   (b) $(\exists \langle S', l'\rangle)[(S \subseteq S' \subseteq W_j^\varphi) \wedge (l + \mathrm{card}(S' - S) \leq l') \wedge (\mathbf{M}(\langle S', l'\rangle) \neq \mathbf{M}(\langle S, l\rangle))]$.

   in which case go to Step 3.

   Remark: In this Step we have tried to check if $\langle S, l\rangle$ is indeed a locking sequence and if $S_m = L \cap N_{m_0}$. Note that W_i^ψ does not enumerate any $x \in N_{m_0} - S_m$.

3. Output $\{1 + m_0\} \cup \{1 + x \mid m_0 < x \wedge (\exists k < i)[x \in W_k^\psi \wedge$ the procedure for W_k^ψ reaches Step 3 $]\}$.

   Note: This step ensures that all nonempty W_k^ψ, such that the procedure for W_k^ψ reaches Step 3 are finite and distinct.

End of definition of W_i^ψ

Claim 1 $(\forall L \in \mathbf{TxtEx}(\mathbf{M}))(\exists i)[W_i^\psi = L \wedge i$ *is nice for* $L]$.

PROOF. Suppose $L \in \mathbf{TxtEx}(\mathbf{M})$. Suppose $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on L and $S_m = L \cap N_{m_0}$, where $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$. Let $i = \langle S, S_m, l\rangle$. Clearly i is nice. Now, since all the conditions checked in Step 1 hold and the procedure never leaves Step 2, $W_i^\psi = L$. $\Box$

Remark: Note that this claim implies property 1 in the definition of effective subdiscreteness.

Claim 2 *Let* $X = \{i \mid i$ *is not nice and* $W_i^\psi \neq \emptyset\}$. *Then* *(a)* $(\forall i \in X)[\mathrm{card}(W_i^\psi) < \infty]$, *and*

(b) $(\forall i, i' \in X)[i = i' \vee W_i^\psi \neq W_{i'}^\psi]$.

PROOF. Let $i = \langle S, S_m, l \rangle$. Let $j = \mathbf{M}(\langle S, l \rangle)$. We will first show that if i is not nice, then either W_i^ψ is empty or the procedure of W_i^ψ reaches Step 3.

So suppose i is not nice. Then one of the following holds:

(A) $\langle S, l \rangle$ is not the least locking sequence for $\mathbf{M}$ on $W_{i_0}^\varphi$;

(B) $S_m \neq W_{i_0}^\varphi \cap N_{m_0}$, where $m_0 = \max(\{\max(S') \mid \langle S', l' \rangle \leq \langle S, l \rangle\})$.

(A) is equivalent to the conjunction of the following two conditions.

(A1) $\langle S, l \rangle$ is not a locking sequence for $\mathbf{M}$ on W_j^φ,

(A2) There exists $\langle S', l' \rangle < \langle S, l \rangle$, such that $\langle S', l' \rangle$ is a locking sequence for $\mathbf{M}$ on $\mathbf{M}(\langle S, l \rangle)$.

If (B) does not hold, but (A2) does, then by Step 1c in the construction of W_i^ψ, we have $W_i^\psi = \emptyset$.

If $\max(S_m) > \max(\{\max(S') \mid \langle S', l' \rangle \leq \langle S, l \rangle\})$, then due to Step (1b) in the construction of W_i^ψ, we have $W_i^\psi = \emptyset$.

If $\max(S_m) \leq \max(\{\max(S') \mid \langle S', l' \rangle \leq \langle S, l \rangle\})$, and (A1) or (B) hold, then either W_i^ψ is $\emptyset$ due to Step 1 of the construction or Step 2 would succeed in finding that i is not nice and thus W_i^ψ would reach Step 3.

Now (a) and (b) follow easily using induction on all j, such that W_j^ψ reaches Step 3. □

Remark: Note that this claim implies clauses (2), (3) and (4) in the definition of effective subdiscreteness since there is a unique nice i for each $L \in \mathbf{TxtEx}(\mathbf{M})$.

Claim 3 *Let* $i = \langle S, S_m, l \rangle$. *Let* $m_0 = \max(\{\max(S') \mid \langle S', l' \rangle \leq \langle S, l \rangle\})$. *Then* $W_i^\psi \cap (N_{m_0} - S_m) = \emptyset$.

PROOF. Obvious by Steps 2 and 3 in the construction. □

Note that the above claim says that all elements in W_i^ψ, that are less than or equal to m_0, are in S_m.

We now define d which will satisfy clause (4) in the definition of effective subdiscreteness. For $i = \langle S, S_m, l\rangle$, let $d(i) = x$, where $D_x = S_m$.

Claim 4 *d satisfies clauses 5a and 5b in the definition of effective subdiscreteness.*

Proof. For $L \in \mathcal{L} - \{\emptyset\}$, let $i = \langle S, S_m, l\rangle$ be such that i is nice and $W_i^\psi = L$ (such an i exists by proof of Claim 1). Clearly $D_{d(i)} = S_m = L \cap N_{\max(S_m)}$. Hence clause 5a in the definition of effective subdiscreteness is satisfied.

Let $n_L = \max(S_m)$. Let $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$. We now show that there are only finitely many j such that j satisfies

$$[(L \cap N_{n_L}) \subseteq W_j^\psi] \wedge [D_{d(j)} \text{ is } n_L\text{-partial consistent with } L]$$

This would prove the claim.

Let $j = \langle S', S'_m, l'\rangle$. Let $m'_0 = \max(\{\max(S'') \mid \langle S'', l''\rangle \leq \langle S', l'\rangle\})$. Clearly j satisfies at least one of the following:

(A) $j \leq l$,

(B) $m'_0 < m_0$,

(C) $S'_m \not\subseteq L$,

(D) $m'_0 \geq m_0$ and $S_m \not\subseteq S'_m$,

(E) (None of the above) $j > l$ and $m'_0 \geq m_0$ and $S_m \subseteq S'_m \subseteq L$.

If j satisfies (C), then $D_{d(j)}$ is not n_L-partial consistent with L. If j satisfies (D), then the following four conditions hold:

$L \cap N_{m_0} = S_m$,

$L \cap N_{n_L} = S_m$ (By definition of n_L and S_m),

$m'_0 \geq m_0 \geq \max(S_m) = n_L$,

$W_j^\psi \cap (N_{m'_0} - S'_m) = \emptyset$ (By Claim 3),

Hence there exists $x \in S_m - W_j^\psi$. Thus $L \cap N_{n_L} \not\subseteq W_j^\psi$.

If j satisfies (E), then at least one of the checks in Steps 1c and 1d in the construction of W_j^ψ will not succeed and, thus, $W_j^\psi = \emptyset$.

Since there are only finitely many j satisfying cases (A) or (B) we have that there are only finitely many j which satisfy

$$[(L \cap N_{n_L}) \subseteq W_j^\psi] \wedge [D_{d(j)} \text{ is } n_L\text{-partial consistent with } L].$$

Thus d satisfies clause (5a) and (5b) in the definition of effective subdiscreteness. □

Claim 1 implies clause (1), Claim 2 implies clauses (2)–(4), Claim 4 implies clause (5) in the definition of effective subdiscreteness. This proves one direction of the theorem.

We now prove that $(\exists\psi)[\psi$ is effectively subdiscrete for $\mathcal{L}] \Rightarrow \mathcal{L} \in \mathbf{TxtEx}$. Let d be as claimed in the definition of effective subdiscreteness. We now describe a machine $\mathbf{M}$ that **TxtEx**-*identifies* $\mathcal{L} - \{\emptyset\}$ (clearly, this implies $\mathcal{L} \in \mathbf{TxtEx}$). Let c be a recursive function reducing ψ to φ.

Definition of $\mathbf{M}(T[n])$

1. (Here the machine tries to find the j's which satisfy

$$[(L \cap N_{n_L}) \subseteq W_j^\psi] \wedge [D_{d(j)} \text{ is } n_L\text{-partial consistent with } L]$$

   However, since the machine does not know n_L, this is not completely possible. So, for each guess s for n_L, $\mathbf{M}$ collects $j \leq s$ which satisfy the above. We will show later that this suffices).

   For $s \leq n$, let $B_s = \{j \mid j \leq s \wedge (\text{content}(T[n]) \cap N_s \subseteq W_{j,n}^\psi) \wedge \wedge D_{d(j)} \text{ is } s\text{-partial consistent with content}(T[n])\}$.

   Let $B = \bigcup_{s \leq n} B_s$.

2. **if** $(\exists j \in B)[W_{j,n}^\psi = \text{content}(T[n])]$,

   then Output $c(j)$ for minimum such j.

   else

Let $s_0 = \max(\{s \mid (\exists j \in B)[W^{\psi}_{j,n} \supseteq \text{content}(T[s]) \wedge\ W^{\psi}_{j,s} \subseteq \text{content}(T[n])]\})$.

Output $c(j_0)$, where $j_0 = \min(\{j \mid j \in B\ \wedge\ W^{\psi}_{j,n} \supseteq \text{content}(T[s_0]) \wedge\ W^{\psi}_{j,s_0} \subseteq \text{content}(T[n])\})$.

endif

(Intuitively, $\mathbf{M}$ here outputs the seemingly best grammar in B, for the input language.)

End of definition of $\mathbf{M}$

Claim 5 $\mathbf{M}$ $\mathbf{TxtEx}$-*identifies* $\mathcal{L} - \{\emptyset\}$.

Proof. For any $L \in \mathcal{L} - \{\emptyset\}$, let n_L be as in the definition of effective subdiscreteness. Suppose T is a text for L. Let *Candidates* be the finite set of j's which satisfy

$$[(L \cap N_{n_L}) \subseteq W^{\psi}_j]\ \wedge\ [D_{d(j)} \text{ is } n_L\text{-partial consistent with } L].$$

Let B^n_s denote B_s constructed by $\mathbf{M}$ on input $T[n]$, and B^n denote B constructed by $\mathbf{M}$ on input $T[n]$.

Let $n_1 \geq n_L$ be so large that $L \cap N_{n_L} \subseteq \text{content}(T[n_1])$.

Hence, for all $n \geq s \geq n_1$, if $j \in B^n_s$, then $L \cap N_{n_L} \subseteq W^{\psi}_j$ and $D_{d(j)}$ is n_L-partial consistent with L (since $s \geq n_L$). Therefore, for all $n \geq s \geq n_1$, $B^n_s \subseteq \mathit{Candidates}$. Thus, for all $n \geq n_1$, $B^n \subseteq \mathit{Candidates}\ \cup N_{n_1}$.

Let j_0 denote the grammar k, as claimed in clause 5a in the definition of effective subdiscreteness and let $D_{d(j_0)} \subseteq \text{content}(T[n_2])$, where $n_2 \geq \max(\{n_1, j_0\})$.

Let n_3 be such that $L \cap N_{n_2} \subseteq W^{\psi}_{j_0,n_3}$.

Therefore, we have that, for all $n \geq \max(\{n_3, n_2\})$, $j_0 \in B^n_{n_2}$. Hence for all $n \geq \max(\{n_3, n_2\})$, $j_0 \in B^n$.

Let $C = (\mathit{Candidates}\ \cup N_{n_1}) \cap \{j \mid (\exists s \geq j)[(L \cap N_s \subseteq W^{\psi}_j) \text{ and } D_{d(j)} \text{ is } s\text{-partial consistent with } L]\}$.

It is easy to see that

$$\lim_{n\to\infty} B^n = C$$

Let n_4 be such that $(\forall n \geq n_4)[B^n = C]$.

We now consider the following two cases.

Case 1: L is infinite.

In this case j_0 is the only element in C, such that $W^{\psi}_{j_0} = L$ (by clause (2) in the definition of effective subdiscreteness).

Let n_5 be so large that

$\neg[(\exists j \in (C - \{j_0\})) \mid [(W^{\psi}_{j,n_5} \subseteq L) \wedge (W^{\psi}_{j} \supseteq \mathrm{content}(T[n_5]))]]$.

Let n_6 be so large that, $(W^{\psi}_{j_0,n_5+1} \subseteq \mathrm{content}(T[n_6])) \wedge (\mathrm{content}(T[n_5+1]) \subseteq W^{\psi}_{j_0,n_6})$.

Clearly, such n_5, n_6 exist. Now for $n \geq \max(\{n_2, n_3, n_4, n_5, n_6\})$ j_0 will be output at Step 2 of the procedure for **M** on input $T[n]$. Therefore, **M** **TxtEx**-*identifies* L.

Case 2: L is finite.

Let $j_0' = \min(\{j \mid j \in C \wedge W^{\psi}_{j} = L\})$. Now for sufficiently large n, **M** on input $T[n]$, will output j_0' at Step 2 of the procedure. Hence, **M** **TxtEx**-*identifies* L.

From the two cases it follows that **M** **TxtEx**-identifies $\mathcal{L} - \{\emptyset\}$. □

This proves Theorem 1. ∎

5 A Characterization of TxtEx with Additional Information

The result of the previous section presented a characterization of language identification from positive data in terms of computable numberings. According to this characterization a collection of r.e. languages is identifiable if and only if there exists a computable numbering that has exactly one index for all the infinite languages in the class and finitely many indices for any finite language in the class, with the additional requirement that a suitable discrimination procedure exist. In this section we present a similar characterization for language identification from positive data in the presence of additional information, a notion more general than **TxtEx**-identification. In Section 5.1, we describe this general notion and in Section 5.2, we present the characterization.

5.1 Identification with Additional Information

In **TxtEx**-identification, the only information provided to a learning agent is the positive data about the language. Motivated by the work of Freivalds and Wiehagen [8] in the context of function identification, Jain and Sharma [13] considered identification paradigms that allowed the learner to have knowledge of an upper bound on the minimal index grammar for the language being learned. See also [12] for another notion of additional information.

To formally consider this paradigm, it is technically expedient to treat learning machines as computing recursive functions of two arguments, viz., additional information and finite initial sequence of a text for the language being learned. From the context, it will be clear when we are discussing learning with additional information as opposed to learning without additional information.

$\mathbf{M}(b,\sigma)$ denotes the output of $\mathbf{M}$ on input σ with additional information b. $\mathbf{M}(b,T)\downarrow= i$ just in case $(\stackrel{\infty}{\forall} n)[\mathbf{M}(b,T[n]) = i]$. We write $\mathbf{M}(b,T)\downarrow$ just in case $(\exists i)[\mathbf{M}(b,T)\downarrow= i]$.

Definition 9 [13]

(a) $\mathbf{M}$ **TxtBex**-*identifies* $L \in \mathcal{E}$ (written: $L \in \mathbf{TxtBex}(\mathbf{M})$) just in case $(\forall T \text{ for } L)(\forall b \geq \mathrm{MinGram}_{\varphi}(L))(\exists i)[W_i^{\varphi} = L \wedge \mathbf{M}(b,T)\downarrow = i]$.

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

Intuitively, machine $\mathbf{M}$ **TxtBex**-identifies a language L if $\mathbf{M}$, fed b, an upper bound on the minimal grammar for L, and a text for L, converges in the limit to a grammar for L. If we further require that the grammar converged to in the limit be the same for any upper-bound, we get a new language identification paradigm described below.

Definition 10 [13]

(a) **M** **TxtUniBex**-*identifies* $L \in \mathcal{E}$ (written: $L \in \mathbf{TxtUniBex}(\mathbf{M})$) just in case $(\exists i \mid W_i^{\varphi} = L)(\forall T \text{ for } L)(\forall b \geq \mathrm{MinGram}_{\varphi}(L))[\mathbf{M}(b, T)\!\downarrow = i]$.

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

We refer the reader to Jain and Sharma [13] for an extensive discussion of the two paradigms introduced above. Note that a counterpart of Lemmas 1 and 2 can easily be obtained for **TxtUniBex**-identification.

The relationship between the paradigms introduced so far is summarized below (see [13]).

$$\mathbf{TxtEx} \subset \mathbf{TxtUniBex} \subset \mathbf{TxtBex} \subset 2^{\mathcal{E}}$$

To illustrate the techniques presented in the previous section we next give a characterization of **TxtUniBex**.

5.2 Characterization of TxtUniBex

We now introduce the notion of weak effectively subdiscrete numbering. This notion is used to characterize **TxtUniBex**.

Definition 11 ψ is *weak effectively subdiscrete* for $\mathcal{L}$ iff the following four conditions are satisfied.

1. $\mathcal{L} \subseteq \mathcal{C}_{\psi}$.
2. $(\forall L \in \mathcal{L})[L \text{ is infinite} \Rightarrow \mathrm{card}(\{i \mid W_i^{\psi} = L\}) = 1]$.
3. $(\forall L \in \mathcal{L})[L \text{ is finite} \Rightarrow \mathrm{card}(\{i \mid W_i^{\psi} = L\}) < \infty]$.

4. $(\exists d \in \mathcal{R}^2)(\forall L \in (\mathcal{L} - \{\emptyset\}))(\forall b > \text{MinGram}_\varphi(L))$ the following two conditions are satisfied

   4a. $(\exists k)[(W_k^\psi = L) \wedge (\exists \langle m, s\rangle)[(\langle m, s\rangle \in W_{d(k,b)}) \wedge (D_s = L \cap N_m)]]$.

   4b. $(\exists n_L \in L)[\text{card}(\{j \mid (\exists \langle m, s\rangle \in W_{d(j,b)})[D_s \text{ is } \min(\{m, n_L\})\text{-consistent with } L \text{ and } (L \cap N_{n_L} \subseteq W_j^\psi)]\}) < \infty]$

This definition is similar to the definition of effective subdiscreteness. Here d, does not directly give a canonical index for a discriminating finite set. Intuitively, in this case, d gives a gammar for a set of numbers, at least one of which codes the discriminating finite set. Also note the restriction in clauses (2) and (3) to the languages in $\mathcal{L}$. It should be noted that this notion is weaker than the notion of effectively subdiscrete because Clauses 1–3 in the definition hold only for the languages in the class and the discrimination function d does not directly give the canonical index for discriminating set.

Definition 12

(a) $\mathcal{L}$ is *weak effectively subdiscrete* $\iff$ $(\exists\psi)[\psi$ is weak effectively subdiscrete for $\mathcal{L}$].

(b) $\mathbf{Wesd} = \{\mathcal{L} \subseteq \mathcal{E} \mid \mathcal{L}$ is weak effectively subdiscrete $\}$.

Theorem 2 **TxtUniBex** = **Wesd**.

Proof. The proof proceeds along similar lines as the proof of Theorem 1.

We first prove that $\mathcal{L} \in$ **TxtUniBex**$\Rightarrow$ $(\exists\psi)[\psi$ is weak effectively subdiscrete for $\mathcal{L}]$. For ease of presenting the proof we give a numbering which may contain infinitely many grammars for $\emptyset$. This numbering can be modified to give a numbering which contains finite number of grammars for $\emptyset$, as explained in the proof of Theorem 1.

Suppose $\mathcal{L} \subseteq$ **TxtUniBex**(**M**), where **M** is rearrangement independent and order independent. We further assume that for all b, σ such that content$(\sigma) = \emptyset$, $W_{\mathbf{M}(b,\sigma)} = \emptyset$. Note that this can easily be ensured.

The proof of this theorem is quite similar to the proof of Theorem 1. An analogous definition of nice in this case needs introduction of the additional information. The construction of ψ is similar, except for taking care of this additional information. The construction of d is different, since not all the properties of ψ hold as before (since $\mathbf{M}$ only **TxtUniBex**-identifies $\mathcal{L}$).

We now introduce a technical notion that facilitates the description of our proof. Suppose $k = \langle S, S_m, l, j\rangle$.

(a) We say that k is *nice* iff

 (i) $W_j^\varphi \in \mathcal{L}$,

 (ii) $\mathbf{M}(j, \langle S, l\rangle) = j$,

 (iii) $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on W_j^φ with j as the additional information and

 (iv) $S_m = (W_j^\varphi \cap N_{m_0})$, where $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$.

(b) We say that $k = \langle S, S_m, l, j\rangle$ is *nice for L* if k is nice and $W_{\mathbf{M}(j,\langle S,l\rangle)} = L$.

We define ψ as follows (this is very similar to the corresponding ψ in proof of Theorem 1).

Definition of W_i^ψ

Let $i = \langle S, S_m, l, j\rangle$. Let $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$.

In Step 1 we attempt to check if i satisfies certain properties of being nice. In Step 2 we check for the remaining properties. If i is nice, then we enumerate W_j^φ (i.e., $W_i^\psi = W_j^\varphi$). If i is not nice, then either W_i^ψ is empty or $W_i^\psi \notin \mathcal{L}$ or in Step 3 we make W_i^ψ a finite set different from all W_j^ψ such that j is not nice.

1. Check the following five conditions

 (a) $\mathbf{M}(j, \langle S, l\rangle) = j$.

 (b) $S \subseteq S_m \subseteq W_j^\varphi$.

(c) $\max(S_m) \leq m_0$.

(d) $(\forall \langle S', l' \rangle < \langle S, l \rangle)[S' \not\subseteq S_m \ \vee \ (\exists S'', l'')[(S' \subseteq S'' \subseteq W_j^{\varphi}) \ \wedge \ (l' + \mathrm{card}(S'' - S') \leq l'') \ \wedge \ (\mathbf{M}(j, \langle S', l' \rangle) \neq \mathbf{M}(j, \langle S'', l'' \rangle))]]$

Remark: The above step checks that $\langle S', l' \rangle < \langle S, l \rangle$ are not a locking sequence for $\mathbf{M}$ on W_j^{φ}, with additional information j.

(e) $(\forall \langle S', l' \rangle)[[(S \subseteq S' \subseteq S_m) \ \wedge \ (l + \mathrm{card}(S' - S) \leq l' \leq (l + i + \mathrm{card}(S_m)))] \Rightarrow \mathbf{M}(j, \langle S', l' \rangle) = j]$.

If any one of the above conditions fail to hold, then let $W_j^{\psi} = \emptyset$.

Remark: Note that if the above conditions hold we can verify the fact. If we cannot verify whether or not a condition holds, then by default W_j^{ψ} will be empty. Step 1e partially checks whether $\langle S, l \rangle$ is a locking sequence for $\mathbf{M}$ on W_j^{φ}, with additional information j – this part is needed for proving that the discrimination function d works.

2. Enumerate elements of $W_j^{\varphi} - (N_{m_0} - S_m)$ until it is found that one of the following conditions hold:

$W_j^{\varphi} \cap (N_{m_0} - S_m) \neq \emptyset$ or

$(\exists \langle S', l' \rangle)[(S \subseteq S' \subseteq W_j^{\varphi}) \ \wedge \ (l' \geq l + \mathrm{card}(S' - S)) \ \wedge \ (\mathbf{M}(j, \langle S', l' \rangle) \neq \mathbf{M}(j, \langle S, l \rangle))]$

in which case go to 3.

Remark: In this step we have tried to check if $\langle S, l \rangle$ is indeed a locking sequence with additional information j and if $S_m = L \cap N_{m_0}$. Also note that W_i^{ψ} does not output any $x \in N_{m_0} - S_m$.

3. Output $\{1 + m_0\} \cup \{1 + x \mid m_0 < x \ \wedge \ (\exists k < i)[x \in W_k^{\psi} \wedge$ the procedure for W_k^{ψ} reaches Step 3 $]\}$.

Note: This step ensures that all W_k^{ψ}, such that procedure for W_k^{ψ} reaches Step 3, are finite and distinct.

End of definition of W_i^{ψ}

Proof of the fact that ψ satisfies clauses (1)–(3) of the definition of weak effective subdiscreteness follows along the same lines as the corresponding proofs in Theorem 1. Verification for clause (4) is different.

Claim 6 $(\forall L \in \mathcal{L})(\exists i)[W_i^\psi = L \ \wedge \ i$ *is nice for* $L]$.

PROOF. Consider any $L \in \mathcal{L}$. Let a_L be such that $W_{a_L} = L$ and is the grammar output by $\mathbf{M}$ in the limit for any text for L and additional information $b \geq \mathrm{MinGram}_\varphi(L)$. Let $i = \langle S, S_m, l, a_L\rangle$ be such that $\langle S, l\rangle$ is the least locking sequence for $\mathbf{M}$ on L with additional information a_L and $S_m = (L \cap N_{m_0})$, where $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$. Clearly i is nice and since all the conditions checked in Step 1 hold and the procedure never leaves Step 2, $W_i^\psi = L$. □

Remark: Note that this claim implies clause (1) in the definition of weak effective subdiscreteness.

Claim 7 *Let* $X = \{i \mid i$ *is not nice* $\ \wedge \ W_i^\psi \neq \emptyset \ \wedge \ W_i^\psi \in \mathcal{L}\}$ *Then*

(a) $(\forall i \in X)[\mathrm{card}(W_i^\psi) < \infty]$, *and*

(b) $(\forall i, i' \in X)[i = i' \ \vee \ W_i^\psi \neq W_{i'}^\psi]$.

PROOF. Let $i = \langle S, S_m, l, j\rangle$. We will first show that if i is not nice, then either W_i^ψ is empty or $W_i^\psi \notin \mathcal{L}$ or the procedure of W_i^ψ reaches Step 3.

So suppose i is not nice. Then one of the following four conditions holds:

(A) $W_j^\varphi \notin \mathcal{L}$,

(B) $\mathbf{M}(j, \langle S, l\rangle) \neq j$,

(C) $\langle S, l\rangle$ is not the least locking sequence for $\mathbf{M}$ on W_j^φ with additional information j,

(D)$S_m \neq (W_j^\varphi \ \cap \ N_{m_0})$, where $m_0 = \max(\{\max(S') \mid \langle S', l'\rangle \leq \langle S, l\rangle\})$.

(C) is equivalent to the conjunction of the following two conditions.

(C1) $\langle S, l\rangle$ is not a locking sequence for $\mathbf{M}$ on W_j^φ with additional information j.

(C2) There is a $\langle S', l'\rangle < \langle S, l\rangle$ which is a locking sequence for $\mathbf{M}$ on W_j^φ with additional information j.

Below let $m_0 = \max(\{\max(S') \mid \langle S', l' \rangle \leq \langle S, l \rangle\})$.

If (A) holds, then either $W_i^\psi = \emptyset$ due to Step 1, or $W_i^\psi = W_j^\varphi \notin \mathcal{L}$ due to Step 2, or W_i^ψ reaches Step 3.

If (D) does not hold but one of (B) or (C2) holds, then by Steps 1a and 1d in the construction of W_i^ψ, we have $W_i^\psi = \emptyset$.

If $\max(S_m) > m_0$, then due to Step 1c in the construction W_i^ψ, we have that $W_i^\psi = \emptyset$.

If (D) or (C1) holds and $\max(S_m) \leq m_0$, then either W_i^ψ is $\emptyset$ due to Step 1 of the construction or the violation of (C1) and (D) would be detected in Step 2 and thus W_i^ψ reaches Step 3.

Theorem now follows using induction on i, such that W_i^ψ reaches Step 3. $\square$

Remark: Note that this claim implies clauses (2) and (3) in the definition of weak effective subdiscreteness since there is a unique nice i for each $L \in \mathcal{L}$.

We now give the construction of d and the proof that it satisfies clause 4 in the definition of weak effectively subdiscrete.

Let $d(i, b)$ be the index for the following program (note that this index can be effectively found from i and b; we give the program as an enumerator). Let $i = \langle S, S_m, l, j \rangle$.

Definition of $W_{d(i,b)}$

1. Check whether $W_i^\psi \neq \emptyset$ and $S_m \subseteq W_i^\psi$. If so, then let $z_0 = \min(\{t \mid W_{i,t}^\psi \supseteq S_m\})$ and proceed to Step 2.

   (Note that otherwise $W_{d(i,b)} = \emptyset$.)

2. Search for $z_1 > z_0$ such that $W_{i,z_1}^\psi \supseteq S_m$ and $\mathbf{M}(b, \langle W_{i,z_1}^\psi, \mathrm{len} \rangle) = j$, where $\mathrm{len} = 2 * (i + \mathrm{card}(W_{i,z_1}^\psi) + z_1)$.

   If and when such a z_1 is found, let $\mathrm{len} = 2 * (i + \mathrm{card}(W_{i,z_1}^\psi) + z_1)$, $m = \max(\{\{i\} \cup W_{i,z_1}^\psi\})$ and s be such that $D_s = W_{i,z_1}^\psi$. Enumerate $\langle m, s \rangle$ in $W_{d(i,b)}$ and proceed to Step 3.

3. Dovetail Steps 3.1 and 3.2 until the search in one of them succeeds. If Step 3.1 succeeds before Step 3.2 does (if ever), then go to Step 3.3. If Step 3.2 succeeds before Step 3.1

does (if ever), then go to Step 3.4.

3.1. Search for x, z_2 such that $x \in ((W^{\psi}_{i,z_2} \cap N_m) - D_s)$.

3.2. Search for $\langle S', l' \rangle$ such that $(S' \supseteq W^{\psi}_{i,z_1})$ and $(l' \geq len + \mathrm{card}(S' - W^{\psi}_{i,z_1}))$ and $[\mathbf{M}(b, \langle S', l' \rangle)] \neq j$.

3.3. Let x, z_2 be as found in Step 3.1.

if $(\forall D, l)[[(W^{\psi}_{i,z_1} \subseteq D \subseteq W^{\psi}_{i,z_2}) \wedge (\mathrm{len} + \mathrm{card}(D - W^{\psi}_{i,z_1}) \leq l \leq 2 * (z_2 + i + \mathrm{card}(W^{\psi}_{i,z_2})))] \Rightarrow [\mathbf{M}(b, \langle D, l \rangle) = j]]$.

then

Enumerate $\langle m, s' \rangle$ in $W_{d(i,b)}$, where $D_{s'} = D_s \cup \{x\}$.

Let $s = s'$ and Go to Step 3.

else Let $z_0 = z_2$ and Go to Step 2.

endif

3.4. Let $z_0 = z_1 + 1$. Go to Step 2.

End of definition of $W_{d(i,b)}$

Claim 8 *d defined above satisfies clauses 4a and 4b in the definition of weak effective subdiscreteness.*

Proof. Let $L \in \mathcal{L} - \{\emptyset\}$ and $b \geq \mathrm{MinGram}_{\varphi}(L)$.

We first show that d satisfies clause 4a in the definition of weak effective subdiscreteness. Let $i_L = \langle S_L, S_{L_m}, l_L, j_L \rangle$ be such that i_L is nice and $W^{\psi}_{i_L} = L$ (there exists one as shown in proof of Claim 6).

We claim that for $b > \mathrm{MinGram}_{\varphi}(L)$, $W_{d(i_L,b)}$ is finite. To see this, let $\langle S, l \rangle$ be a locking sequence for $\mathbf{M}$ on L with additional information b. Let $n > l_L$ be so large that $W^{\psi}_{i_L,n} \supseteq S$. Consider the execution of the enumerator for $W_{d(i_L,b)}$ described above.

(a) Each execution of Step 2 increases the value of z_1 by at least 1.

(b) All executions of Step 2 are followed by execution of Step 3.

(c) Step 3 can be executed only finitely many times before Step 2 is executed again.

(d) Step 2 follows Step 3 only if it has been verified in Step 3 that $\langle W^{\psi}_{i_L,z_1}, \mathrm{len}\rangle$, where $\mathrm{len} = 2*(i+z_1+\mathrm{card}(W^{\psi}_{i,z_1}))$ is not a locking sequence for $\mathbf{M}$ on L with additional information b.

Thus for $W_{d(i_L,b)}$ to be infinite $\langle W^{\psi}_{i_L,z_1}, \mathrm{len}\rangle$ must not be a locking sequence for $\mathbf{M}$ on L with additional information b for infinitely many z_1. But this is not true for $z_1 > n$ (since $n > l$ and $W^{\psi}_{i,n} \supseteq S$, which implies that $\langle W^{\psi}_{i_L,z_1}, \mathrm{len}\rangle$ is a locking sequence for $\mathbf{M}$ on L with additional information b). Thus $W_d(i_L, b)$ is finite.

Let the last element in the order of enumeration of $W_{d(i_L,b)}$ as described above be $\langle m_L, s_L\rangle$. Now, clause (4a) in the definition of weak effective subdiscreteness is satisfied by taking $k = i_L$, since $W^{\psi}_{i_L} = L$, and $D_{s_L} = L \cap N_{m_L}$ (otherwise search in Step 3.1. would succeed).

We now show that d satisfies clause (4b) in the definition of weak effective subdiscreteness. Let $n_L = \max(D_{s_L})$. Also, for some l (say l_{s_L}), $\langle D_{s_L}, l\rangle$ is a locking sequence for $\mathbf{M}$ on L with additional information b, since otherwise Step 3.2 would succeed and $\langle m_L, s_L\rangle$ would not be the last element enumerated in $W_{d(i_L,b)}$.

We now show that, only finitely many i can satisfy the following

$$(\exists \langle m, s\rangle \in W_{d(i,b)})[D_s \text{ is } \min(\{m, n_L\})\text{-consistent with } L \wedge (L \cap N_{n_L} \subseteq W^{\psi}_i)]$$

This would prove the claim

Consider any $i = \langle S, S_m, l, j\rangle \geq \max(\{n_L, l_{s_L}\})$ such that $\langle S, l\rangle > \langle S_L, l_L\rangle$. This assumption is fine, since there are only finitely many $i = \langle S, S_m, l, j_L\rangle$ such that $\langle S, l\rangle \leq \langle S_L, l_L\rangle$ and $S_m \subseteq \bigcup_{\langle S', l'\rangle \leq \langle S, l\rangle} S'$.

We will show that for such i, there does not exists a $\langle m, s\rangle \in W_{d(i,b)}$ satisfying

$[D_s$ is $\min(\{m, n_L\})$-consistent with $L \wedge (L \cap N_{n_L} \subseteq W^{\psi}_j)]$.

Suppose by way of contradiction that, $\langle m, s\rangle \in W_{d(i,b)}$ satisfies [D_s is $\min(\{m, n_L\})$-consistent with $L \wedge (L \cap N_{n_L} \subseteq W^{\psi}_j)$].

We have $m \geq i \geq n_L$ (Step 2 of the procedure makes $m \geq i$). Therefore D_s is n_L-consistent with L. Also, $\mathbf{M}(b, \langle D_s, 2 * (\mathrm{card}(D_s) + i + z'_1)\rangle) = j$, where z'_1 is the value of z_1 when $\langle m, s\rangle$ is enumerated in $W_{d(i,b)}$ (since $W_{d(i,b)}$ enumerates $\langle m, s\rangle$ only if this condition is satisfied.) Therefore, since $\langle L \cap N_{n_L}, l_{s_L}\rangle$ is a locking sequence for $\mathbf{M}$ on L with additional information b, we have $\mathbf{M}(b, \langle D_s, 2 * (\mathrm{card}(D_s) + i + z'_1)\rangle) = j_L$. Thus $j = j_L$.

Now if $S_m \supseteq S_L$, then by the construction of ψ, $W^{\psi}_i = \emptyset$ (check performed in Step 1d in the definition of W^{ψ}_i). But, if $S_m \not\supseteq S_L$, then $S_L \not\subseteq W^{\psi}_i$ (by the construction of ψ, see remark at the end of Step 2 and the assumption that $\langle S, l\rangle \geq \langle S_L, l_L\rangle$); and hence $S_L \not\subseteq D_s$. This contradicts the fact that D_s is n_L-consistent with L (since $n_L \geq \max(S_L)$ by Step 1 in the construction of d). This proves the claim. $\Box$

The above claims show that $\mathcal{L} \in$ **TxtUniBex**$\Rightarrow (\exists\psi)[\psi$ is weak effectively subdiscrete for $\mathcal{L}]$.

We now show that $(\exists\psi)[\psi$ is weak effectively subdiscrete for $\mathcal{L}] \Rightarrow \mathcal{L} \in$ **TxtUniBex**.

Let c be the function reducing ψ to φ. Let d be as claimed in the definition of weak effective subdiscreteness. Define $\mathbf{M}$ on additional information b as follows

Definition of $\mathbf{M}(b, T[n])$

1. Here we are trying to find i's such that some $\langle m, s\rangle \in W_{d(i,b)}$ satisfies [D_s is $\min(\{m, n_L\})$-consistent with $L \wedge (L \cap N_{n_L} \subseteq W^{\psi}_i)$], where n_L is as in clause 4b of the definition of weak effectively subdiscrete. However, since $\mathbf{M}$ does not know n_L, this is not completely possible. Thus $\mathbf{M}$, for each guess r for n_L, collects all $i \leq r$, satisfying the above. We will see later that this suffices.

   For $r \leq n$, let $B_r = \{j \mid j \leq r \wedge (\exists\langle m, s\rangle \in W_{d(j,b),n})[D_s$ is $\min(\{m, r\})$-consistent with $\mathrm{content}(T[n]) \wedge (T[n] \cap N_r \subseteq W^{\psi}_{j,n})]\}$.

Let $B = \bigcup_{r \leq n} B_r$.

2. **if** $(\exists j \in B)[W^{\psi}_{j,n} = \text{content}(T[n])]$

then Output $c(j)$ for minimum such j.

else

Let $r_0 = \max(\{r \mid (\exists j \in B)[(W^{\psi}_{j,n} \supseteq \text{content}(T[r])) \wedge (W^{\psi}_{j,r} \subseteq \text{content}(T[n]))]\})$.

Output $c(j_0)$, where $j_0 = \min(\{j \mid [j \in B] \wedge [(W^{\psi}_{j,n} \supseteq \text{content}(T[r_0])]) \wedge [(W^{\psi}_{j,r_0} \subseteq \text{content}(T[n]))]\})$.

endif

(Intuitively, **M** here tries to output the seemingly best grammar in B, for the input language.)

End of definition of **M**

Claim 9 **M TxtUniBex**-*identifies* $\mathcal{L} - \{\emptyset\}$. *Thus* $\mathcal{L} \in$ **TxtUniBex**.

PROOF. Let $L \in \mathcal{L}$, $b \geq \text{MinGram}_{\varphi}(L)$ and T be a text for L. Let n_L be as in the definition of weak effective subdiscreteness. Let *Candidates* be the finite set of i's satisfying $(\exists \langle m, s\rangle \in W_{d(i,b)})$ $[D_s$ is $\min(\{m, n_L\})$-consistent with $L \wedge L \cap N_{n_L} \subseteq W^{\psi}_{j}]$. Let B^n_r denote B_r as computed by **M** on input $T[n]$. Let B^n denote B as computed by **M** on input $T[n]$.

Let $n_1 \geq n_L$ be such that $L \cap N_{n_L} \subseteq \text{content}(T[n_1])$. Hence, for all $n \geq r \geq n_1$, $j \in B^n_r$ implies that $j \in \mathit{Candidates}$. Therefore, for all $n \geq n_1$, $B^n \subseteq \mathit{Candidates} \cup N_{n_1}$.

Let j_0 be the grammar k as claimed in clause (4a) in the definition of weak effective subdiscreteness and let n_2 be so large that, for $\langle m, s\rangle$ claimed in clause (4a) in the definition of weak effective subdiscreteness, $D_s \subseteq \text{content}(T[n_2])$ and $\langle m, s\rangle \in W_{d(j_0,b),n_2}$, where $n_2 \geq j_0$. Also, let n_3 be so large that $L \cap N_{n_2} \subseteq W^{\psi}_{j_0,n_3}$. Therefore, $(\forall n \geq \max(\{n_3, n_2\}))[j_0 \in B^n]$.

Let $C = (\mathit{Candidates} \cup N_{n_1}) \cap \{i \mid (\exists \langle m, s\rangle \in W_{d(i,b)})(\exists r \geq i)[D_s$ is $\min(\{m, r\})$-consistent with $L \wedge (L \cap N_r \subseteq W^{\psi}_{i})]\}$.

Clearly,

$$\lim_{n\to\infty} B^n = C.$$

Let n_4 be such that for all $n > n_4$, $B^n = C$.

We now consider the following two cases.

Case 1: L is infinite.

In this case j_0 is the only element in C such that $W^{\psi}_{j_0} = L$ (by clause 2 in the definition of weak effective subdiscreteness).

Let n_5 be so large that

$\neg[(\exists j \in (C - \{j_0\}))[(W^{\psi}_{j,n_5} \subseteq L) \wedge (W^{\psi}_{j} \supseteq \text{content}(T[n_5]))]]$.

Let n_6 be so large that $[(W^{\psi}_{j_0,n_5+1} \subseteq \text{content}(T[n_6])) \wedge (\text{content}(T[n_5+1]) \subseteq W^{\psi}_{j_0,n_6})]$.

Clearly, such n_5, n_6 exist. Now for $n \geq \max(\{n_2, n_3, n_4, n_5, n_6\})$, **M** with additional information b, on input $T[n]$, will output j_0. Therefore, **M TxtUniBex**-identifies L.

Case 2: L is finite.

Let $j_0' = \min(\{j \mid j \in C \wedge W^{\psi}_{j} = L\})$. Now for $n \geq \{n_2, n_3, n_4\}$, **M**, with additional information b, on input $T[n]$, will output j_0' at Step 2 of the procedure. Hence, **M TxtUniBex**-*identifies* L.

From the above two cases it follows that $(\mathcal{L} - \{\emptyset\}) \in$ **TxtUniBex**(**M**). $\square$

This proves Theorem 2. ∎

6 Summary

We characterized **TxtEx** and **TxtUniBex**. We summarize our results below.

$$\mathbf{TxtEx} = \mathbf{Esd} \subset \mathbf{TxtUniBex} = \mathbf{Wesd} \subset \mathbf{TxtBex} \subset 2^{\mathcal{E}}.$$

We feel that one of the main contributions of this paper are the techniques developed to deal with language identification. As already noted, we have used these techniques in other investigations of language learning [5, 15].

7 Acknowledgements

We are grateful to John Case and Mark Fulk for several suggestions. We would also like to thank Zuzana Dobes, Lata Narayanan and Rajeev Raman for helpful discussions. Research was partly supported by the National Science Foundation Grants CCR 832-0136, CCR 871-3846 and by the Australian Research Council Grant Number A49530274.

References

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

[2] D. Angluin and C. Smith. A survey of inductive inference: Theory and methods. *Computing Surveys*, 15:237–289, 1983.

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

[4] J. Case. Learning machines. In W. Demopoulos and A. Marras, editors, *Language Learning and Concept Acquisition*. Ablex Publishing Company, 1986.

[5] J. Case, S. Jain, and A. Sharma. Vacillatory learning of nearly minimal size grammers. *Journal of Computer and System Sciences*, 49(2):189–207, October 1994.

[6] J. Case and C. Lynes. Machine inductive inference and language identification. In M. Nielsen and E. M. Schmidt, editors, *Proceedings of the 9th International Colloquium on Automata, Languages and Programming*, pages 107–115. Springer-Verlag, 1982. Lecture Notes in Computer Science 140.

[7] R. Freivalds, E. B. Kinber, and R. Wiehagen. Connections between identifying functionals, standardizing operations, and computable numberings. *Zeitschr. j. math. Logik und Grundlagen d. Math. Bd.*, 30:145–164, 1984.

[8] R. Freivalds and R. Wiehagen. Inductive inference with additional information. *Electronische Informationverarbeitung und Kybernetik*, 15:179–195, 1979.

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

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

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

[12] S. Jain and A. Sharma. Learning in the presence of partial explanations. *Information and Computation*, 95-2:162–191, 1991.

[13] S. Jain and A. Sharma. Learning with the knowledge of an upper bound on program size. *Information and Computation*, 102–1:118–166, 1993.

[14] S. Jain and A. Sharma. Characterizing language learning by standardizing operations. *Journal of Computer and System Sciences*, 49–1:96–107, 1994.

[15] S. Jain and A. Sharma. Prudence in vacillatory language identification. *Mathematical Systems Theory*, 1994. To Appear.

[16] S. Kapur. Uniform characterizations of various kinds of language learning. In *Proceedings of the Fourth International Workshop on Algorithmic Learning Theory, Lecture Notes in Artificial Intelligence 744*. Springer, 1993.

[17] R. Klette and R. Wiehagen. Research in the theory of inductive inference by GDR mathematicians – A survey. *Information Sciences*, 22:149–169, 1980.

[18] S. Lange and T. Zeugmann. Learning recursive languages with bounded mind changes. Technical Report 16/92, GOSLER-Report, FB Mathematik und Informatik, TH Lepzig, 1992.

[19] S. Lange and T. Zeugmann. Types of monotonic language learning and their characterization. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, Pennsylvania*, pages 377–390. ACM Press, 1992.

[20] S. Lange, T. Zeugmann, and S. Kapur. Monotonic and dual monotonic language learning. *Theoretical Computer Science A*, 1995. To appear.

[21] Y. Mukouchi. Characterization of pattern languages. In *Proceedings of the 2nd Workshop on Algorithmic Learning Theory*, pages 93–104, 1991.

[22] Y. Mukouchi. Characterization of finite identification. In *Proceedings of the Third International Workshop on Analogical and Inductive Inference, Dagstuhl Castle, Germany*, pages 260–267, October 1992.

[23] Y. Mukouchi and S. Arikawa. Inductive inference machines that can refute hypothesis spaces. Technical Report RIFIS-TRCS-67, RIFIS, Kyushu University, 1993.

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

[25] H. Rogers. Gödel numberings of partial recursive functions. *Journal of Symbolic Logic*, 23:331–341, 1958.

[26] H. Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw Hill, New York, 1967. Reprinted, MIT Press 1987.