Angluin's theorem for indexed families of r.e. sets and applications*

Dick de Jongh
Institute for Logic, Language and Computation
University of Amsterdam
dickdj@fwi.uva.nl

Makoto Kanazawa
Department of Cognitive and Information Sciences
Chiba University
kanazawa@cogsci.l.chiba-u.ac.jp

Abstract

We extend Angluin's (1980) theorem to characterize identifiability of indexed families of r.e. languages, as opposed to indexed families of recursive languages. We also prove some variants characterizing conservativity and two other similar restrictions, paralleling Zeugmann, Lange, and Kapur's (1992, 1995) results for indexed families of recursive languages.

1 Introduction

A significant portion of the work of recent years in the field of inductive inference of formal languages, as initiated by Gold 1967, stems from Angluin's (1980b) theorem, which characterizes when an indexed family of recursive languages is identifiable in the limit from positive data in the sense of Gold. Up until around 1980, a prevalent view had been that inductive inference from positive data is too weak to be of much theoretical interest. This misconception was due to the negative result in Gold's original paper, which says that any class of languages that contains every finite language and at least one infinite language cannot be identifiable in the limit from positive data. Since Gold's theorem applies even to the class of regular languages, it was taken to mean that a class of languages that is identifiable from positive data must be quite impoverished. Angluin's (1980b) theorem, together with her (1980a, 1982) work on concrete classes of formal languages, showed that there are many nontrivial classes of languages identifiable from positive data.

Angluin's idea of restricting attention to indexed families of recursive languages, where membership is uniformly decidable, has proved fruitful (see, e.g., Kapur 1991 and Zeugmann and Lange 1995). For instance, Angluin's theorem has lead to easy-to-apply sufficient conditions for the identifiability of indexed families of recursive languages. Finite elasticity of Wright 1989 is such a condition (see also Shinohara 1990, Moriyama and Sato 1993, and Kanazawa 1994). Moreover, Zeugmann, Lange and Kapur (1992, 1995) strengthen Angluin's condition in various ways to obtain characterizations of conservativity and other restrictions on the behavior of the learner.

However, these developments have had little impact on research that deals with r.e., possibly non-recursive languages, where Angluin's theorem and its consequences do not apply. For instance, Osherson, Stob, and Weinstein's (1986) textbook does not even mention Angluin's theorem; they only present a watered-down version of Angluin's theorem—referring to Angluin's (1980b) paper—which characterizes non-effective identifiability in the general case. One almost gets the impression that the field is split into two more or less independent areas with different concerns, one dealing with indexed families of recursive languages and the other dealing with classes of r.e. languages in general. No doubt this is due to the fact that there has been no successful attempt to extend the scope of Angluin's theorem to cover language classes other than indexed families of recursive languages. A related effect is that people have paid little attention to the question of identifiability of indexed families of r.e. languages. Crucial examples offered in the literature dealing with r.e., possibly non-recursive languages have often been classes that are not r.e. indexable (see, e.g., Osherson, Stob, and Weinstein 1986, Jain and Sharma 1994, and Kinber and Stephan 1995).

*This paper will appear in the proceedings of COLT '96. We are grateful to Carl Smith, Takeshi Shinohara, Bas Terwijn, and Thomas Zeugmann for discussions and/or comments.

The paucity of connection between the two lines of research in inductive inference from positive data is unfortunate. In this paper, we try to remedy this situation by proving an Angluin-type theorem characterizing identifiability of indexed families of r.e. languages, where membership is uniformly semi-decidable (r.e.). The characterizing condition involves finite 'telltale' sets as in Angluin's theorem, but also 'warning' sets that have to do with the recognition of counterexamples. The theorem does not easily lead to simple sufficient conditions as in the recursive case, but we provide one such sufficient condition. In addition, we prove a variant of the theorem characterizing conservative identifiability of indexed families of r.e. languages, paralleling developments in the recursive case. An interesting fact is that one can distinguish at least three types of conservative-like behavior, which amount to an equivalent restriction in the recursive case, but not in the r.e. case. We show that the three restricted notions of identifiability are characterized by successive strengthenings of the characterizing condition.

2 Preliminaries

We use standard notations and terminology from recursion theory (see e.g., Rogers 1967). In particular, φ_i denotes the partial recursive function computed by the i-th program in some acceptable programming system, $\varphi(x)\downarrow$ means φ is defined on x, $\varphi_{i,s}$ is that finite subfunction of φ_i which is obtained by cutting off the computation of i after the first s steps, $W_i = \{\, x \mid \varphi_i(x)\downarrow \,\}$, $W_{i,s} = \{\, x \mid \varphi_{i,s}(x)\downarrow \,\}$, and $K = \{\, x \mid \varphi_x(x)\downarrow \,\}$. We have $W_i = \bigcup_{s\in\mathbf{N}} W_{i,s}$, where $\mathbf{N}$ denotes the set of natural numbers. We use capital letters $F, G, H, \ldots$ to denote total recursive functions.

We assume standard coding of pairs of natural numbers and of finite sequences of natural numbers. If σ is (the natural number representing) a finite sequence, $\mathrm{rng}(\sigma)$ is the set of elements that appear in σ, and $\mathrm{lh}(\sigma)$ denotes the length of σ. If σ and τ are finite sequences, $\sigma^\wedge\tau$ denotes the concatenation of σ and τ. Finite sets are also coded and put in one-one correspondence with $\mathbf{N}$.

2.1 Limiting recursive functions

Let ψ be a partial function on natural numbers. Then $\lim_{n\to\infty} \psi(n)$ is defined and equals x if and only if $\exists m \forall n \geq m (\psi(n) \simeq x)$. A partial function $\varphi\colon \mathbf{N} \to \mathbf{N}$ is called *limiting recursive* if there is a total recursive function $F\colon \mathbf{N}^2 \to \mathbf{N}$ such that $\varphi(x) \simeq \lim_{n\to\infty} F(x,n)$. Obviously, the class of partial recursive functions is included in the class of limiting recursive functions. As Gold (1965) shows, the limiting recursive functions are precisely the functions whose graph is Σ^0_2 in the arithmetical hierarchy, or, equivalently, the functions recursive in $\mathbf{0}'$. Clearly, the limiting recursive functions are closed under composition. If a function φ defined by $\varphi(x) \simeq \lim_{n\to\infty} \psi(x,n)$ is total, then φ is limiting recursive whenever ψ is partial recursive (Gold 1965). When there is a total limiting recursive function φ such that $y_i = \varphi(i)$, we say that y_i is *computable in the limit from* i.

2.2 Indexed families of r.e. languages

We can represent strings of symbols by natural numbers, so languages are sets of natural numbers. We call an infinite sequence $L_0, L_1, L_2, \ldots$ of r.e. languages *uniformly r.e.* if the set $\{\, \langle i,x\rangle \mid x \in L_i \,\}$ is r.e. It is easy to see that an infinite sequence $L_0, L_1, L_2, \ldots$ is uniformly r.e. if and only if there is a total recursive function F such that $L_i = W_{F(i)}$ for all i.

A class $\mathcal{L}$ of r.e. languages is said to be an *indexed family of r.e. languages* if there exists a uniformly r.e. sequence $L_0, L_1, L_2, \ldots$ (*an r.e. indexing of* $\mathcal{L}$) such that $\mathcal{L} = \{\, L_i \mid i \in \mathbf{N} \,\}$. The indexed families of r.e. languages are precisely the non-empty *r.e. indexable* classes of r.e. languages in Osherson, Stob and Weinstein's (1986) terminology. Each total recursive function F determines an indexed family of r.e. languages, namely, $\{\, W_{F(i)} \mid i \in \mathbf{N} \,\}$.

An indexed family of r.e. languages can also be named by an index for a partial recursive function. Let $\mathbf{G}$ be a total recursive function with the following property that $W_{\mathbf{G}(e,i)} = \{\, x \mid \varphi_e(i,x) \simeq 1 \,\}$. We write $L_{e,i}$ for $W_{\mathbf{G}(e,i)}$, and $\mathcal{L}_e$ for $\{\, L_{e,i} \mid i \in \mathbf{N} \,\}$. For each $e \in \mathbf{N}$, $\mathcal{L}_e$ is a family of r.e. languages indexed by $L_{e,0}, L_{e,1}, L_{e,2}, \ldots$. We refer to e as an index of $\mathcal{L}_e$. If φ_e is total and takes its values in $\{0,1\}$, then e is an index for the characteristic function of the recursive set $\{\, \langle i,x\rangle \mid x \in L_{e,i} \,\}$, and each $L_{e,i}$ is recursive. In that case, $\mathcal{L}_e$ is an *indexed family of recursive languages* (Angluin 1980b), and $L_{e,0}, L_{e,1}, L_{e,2}, \ldots$ is a *recursive indexing* of it. When we refer to an indexed family of r.e. languages $\mathcal{L}$, we often write $\mathcal{L} = \{\, L_i \mid i \in \mathbf{N} \,\}$, assuming $L_0, L_1, L_2, \ldots$ is an indexing of $\mathcal{L}$.

2.3 Identification of indexed families

By a recursive learning function, we mean a unary partial recursive function whose domain is recursive. (Such a function may be thought of as a total recursive function that takes its values in $\mathbf{N} \cup \{\text{undefined}\}$.) Thus, when we define a function φ by $\varphi(x) \simeq \mu i \leq F(x)\, A(x,i)$, where F is a total recursive function and A is a recursive predicate, we let $\varphi(x)$ be defined if and only if $\exists i \leq F(x)\, A(x,i)$.

As usual (see Osherson, Stob and Weinstein 1986) an infinite sequence $t = t_0, t_1, t_2, \ldots$ such that $\{ t_n \mid n \in \mathbf{N} \} = L$ is said to be a *text* for L; $\bar{t}_n$ denotes the finite sequence $t_0, t_1, \ldots, t_{n-1}$. A learning function φ is said to *converge* on t to i if $\lim_{n\to\infty} \varphi(\bar{t}_n) \simeq i$, φ is said to *identify* an r.e. set L if, for every text t of L, there is an i such that $L = W_i$ and φ converges on t to i, and φ is said to identify a class of r.e. sets $\mathcal{L}$ if φ identifies every member of $\mathcal{L}$. If $\mathcal{L} = \mathcal{L}_e$, we say that a learning function φ *identifies $\mathcal{L}$ with respect to e (w.r.t. e)* if and only if φ identifies $\mathcal{L}$ and $\mathrm{rng}(\varphi) \subseteq \{ \mathbf{G}(e,i) \mid i \in \mathbf{N} \}$.

Our notion of identification of an indexed family of r.e. languages with respect to an index of it is equivalent to the following definition of Angluin's (1980b), which was meant to apply to the recursive case, but naturally extends to the r.e. case. Let $L_0, L_1, L_2, \ldots$ be a uniformly r.e. sequence of languages. A learning function φ is said to identify $L_0, L_1, L_2, \ldots$ if for every i and every text t for L_i, φ converges on t to some j such that $L_j = L_i$. If $\mathcal{L} = \mathcal{L}_e$ and φ identifies $L_{e,0}, L_{e,1}, L_{e,2}, \ldots$, $\lambda\sigma\, \mathbf{G}(e, \varphi(\sigma))$ identifies $\mathcal{L}$ with respect to e. Also, if ψ identifies $\mathcal{L}$ with respect to e, then $\lambda\sigma\, \mu i(\psi(\sigma) = \mathbf{G}(e,i))$ identifies $L_{e,0}, L_{e,1}, L_{e,2}, \ldots$. Note that both translations preserve recursivity of the learning functions.

Following Lange and Zeugmann's (1993) treatment of identification of indexed families of recursive languages, we distinguish three senses in which a learning function may be said to identify an indexed family of r.e. languages $\mathcal{L} = \mathcal{L}_e$, given by an index e:

(E) φ identifies $\mathcal{L}$ with respect to e, the given index. This corresponds to what Lange and Zeugmann (1993) call *exact learning*.

(P) φ identifies $\mathcal{L}$ with respect to some f such that $\mathcal{L} = \mathcal{L}_f$. This corresponds to what Lange and Zeugmann (1993) call *class-preserving learning*. It's the same thing as saying that φ is prudent and exactly identifies $\mathcal{L}$, in the terminology of Osherson, Stob, and Weinstein (1986).

(C) φ identifies $\mathcal{L}$ in the usual sense. This is what Lange and Zeugmann (1993) call *class-comprising learning*. φ need not be prudent nor exactly identify $\mathcal{L}$ in Osherson, Stob, and Weinstein's (1986) sense.

Identification in the first sense is really a relation between a learning function and an indexing of a family of r.e. languages. The other two notions do not care which indexing is used to specify the family, as long as it is an indexed family. Corresponding to the three senses of identification, there are three senses in which an indexed family of r.e. languages (given by a certain index) is identifiable by a recursive function. More restricted notions of identifiability can also be considered in three forms, by restricting φ to appropriate spaces of learning functions.

3 Angluin's theorem

As Angluin (1980b) points out, a major difficulty in trying for identification from only positive data is the problem of 'overgeneralization'. If ever the learner makes a guess that is overly general, i.e., conjectures a proper superset of the true answer, then no counterexample will occur that will tell the learner to abandon that guess. Angluin's (1980b) condition characterizing when an indexed family of recursive languages is identifiable is stated in terms of the notion of a finite *telltale set*. A telltale set for language L in the family can be thought of as sufficient evidence to conjecture L.

Definition 1 Let $\mathcal{L}$ be a class of languages, and let L be a language in $\mathcal{L}$. A finite set D is a *telltale set* for L in $\mathcal{L}$ if the following condition holds:

$$D \subseteq L \wedge \forall L' \in \mathcal{L}(D \subseteq L' \to L' \not\subset L).$$

As Angluin puts it, the point of the telltale set is that once all its elements have appeared, the learner need not fear overgeneralization in guessing L. Even if the true answer is not L, it cannot be a proper subset of L, so the learner will eventually see a conflict between the data and L, and can change its guess. Angluin's condition characterizing identifiability of an indexed family $\mathcal{L}$ of recursive languages requires that for each language L in $\mathcal{L}$, there is a telltale set for L in $\mathcal{L}$. It moreover requires that there is an effective way of enumerating, given an index for a L in $\mathcal{L}$, some telltale set for L, but here it is instructive to focus on the first aspect of the condition.

The existence of a telltale set for each language exactly characterizes *non-effective identifiability*, i.e., identifiability by an arbitrary, possibly non-recursive, learning function (Osherson, Stob, and Weinstein 1986).

Theorem 2 *Let $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ be any class of r.e. languages. Then $\mathcal{L}$ is non-effectively identifiable if and only if for each $i \in \mathbf{N}$, there is a telltale set D_i for L_i in $\mathcal{L}$.*

For the sufficiency half of the proof, suppose that for each i, D_i is a telltale set for L_i in $\mathcal{L}$. Consider:

$$\varphi(\sigma) \simeq \mu i(D_i \subseteq \mathrm{rng}(\sigma) \wedge \mathrm{rng}(\sigma) \subseteq L_i). \qquad (1)$$

For simplicity, we interpret φ's conjecture i to mean the language L_i, rather than W_i. The strategy of φ is to conjecture the first language L_i, if any, such that L_i is consistent with the data seen so far and the elements of its telltale set D_i have already appeared in the data. To see that φ identifies $\mathcal{L}$, take any $L \in \mathcal{L}$, and let t be a text for L. Let i be the least natural number such that $L = L_i$. We show that φ converges on t to i. Note that for every $j < i$, either $D_j \not\subseteq L$ or $L \not\subseteq L_j$. Thus, we can find m such that:

(i) $D_i \subseteq \mathrm{rng}(\bar{t}_m)$,

(ii) for every $j < i$ such that $D_j \subseteq L$, $\mathrm{rng}(\bar{t}_m) \not\subseteq L_j$.

It is then easy to see that for all $n \geq m$, $\varphi(\bar{t}_n) \simeq i$.

The necessity half is proved using Blum and Blum's (1975) Locking Sequence Lemma (see also Osherson, Stob, and Weinstein 1986).

Definition 3 A finite sequence σ is said to be a *locking sequence* for r.e. set L and learning function φ if:

(i) $\mathrm{rng}(\sigma) \subseteq L$, (ii) $\varphi(\sigma)\downarrow$, (iii) $L = W_{\varphi(\sigma)}$,

(iv) $\forall\tau(\mathrm{rng}(\tau) \subseteq L \rightarrow \varphi(\sigma) \simeq \varphi(\sigma^\wedge\tau))$.

Lemma 4 (Blum and Blum) *If a learning function φ identifies a language L, there is a locking sequence for φ and L. Moreover, if φ is recursive, there is a limiting recursive function γ such that for all i, if φ identifies W_i, then $\gamma(i)$ is the least locking sequence for φ and W_i.*

The necessity half of Theorem 2 is then an easy consequence of the following observation.

Lemma 5 *Let φ identify $\mathcal{L}$, and let σ be locking for φ and L. Then* $\mathrm{rng}(\sigma)$ *is a telltale set for L in $\mathcal{L}$.*

Let us now consider Angluin's theorem in its original 'effective' form. Note that the learning function (1) is in general not recursive. There are three reasons for this: (i) the definition contains an unbounded μ-operator, (ii) D_i may not be effectively found from i, and (iii) the condition $\mathrm{rng}(\sigma) \subseteq L_i$ may not be decidable. Angluin's restriction to indexed families of recursive languages may be understood as a way of making (1) recursive. In an indexed family of recursive languages the condition $\mathrm{rng}(\sigma) \subseteq L_i$ becomes decidable. The problems (i) and (ii) can also be overcome by modifying the definition (1) slightly. Note that since $\mathcal{L}$ is an indexed family, an r.e. index of L_i can be computed from i. Then the second part of the Locking Sequence Lemma (Lemma 4) together with Lemma 5 implies that for $\mathcal{L}$ to be identifiable by a recursive function, there must be a limiting recursive function δ such that for each i, $\delta(i)$ is a telltale set for L_i in $\mathcal{L}$. Given such δ, we can rewrite (1) as follows:

$$\varphi(\sigma) \simeq \mu i(\delta(i) \subseteq \mathrm{rng}(\sigma) \wedge \mathrm{rng}(\sigma) \subseteq L_i) \qquad (2)$$

The learning function defined this way may still not be recursive, but this definition is in fact good enough, by the following lemma:

Lemma 6 *Let φ be a learning function defined by $\varphi(\sigma) \simeq \mu i\, A(\sigma, i, \alpha(i))$, where A is a recursive predicate and α is a limiting recursive function. Then there is a recursive learning function ψ such that for any text t, if $\lim_{n\to\infty} \varphi(\bar{t}_n)\downarrow$, then $\lim_{n\to\infty} \varphi(\bar{t}_n) \simeq \lim_{n\to\infty} \psi(\bar{t}_n)$.*

Proof. Let F be a total recursive function such that for all i, $\alpha(i) \simeq \lim_{n\to\infty} F(i,n)$. Define ψ as follows:

$$\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)\, A(\sigma, i, F(i, \mathrm{lh}(\sigma))).$$

We show that ψ satisfies the required property. Suppose that φ converges on text t to index i. Then there is an m such that for all $n \geq m$, $A(\bar{t}_n, i, \alpha(i))$ and for all $j < i$, $\neg A(\bar{t}_n, j, \alpha(j))$. Take $m' \geq m$ such that for all $j \leq i$ and for all $n \geq m'$, $F(j,n) = \alpha(j)$. Then for all $n \geq m'$, we have $\varphi(\bar{t}_n) \simeq \psi(\bar{t}_n)$. As φ, ψ converges on t to i. ⊣

The learning function in (2) satisfies the conditions in Lemma 6, so there is a recursive learning function that does the job of (2) if δ is limiting recursive. It is useful to see how such a recursive learning function can be explicitly defined. Let $D_i = \delta(i) = \lim_{n\to\infty} D_i^n$, where D_i^n is computable from n and i. Then an application of the proof of Lemma 6 to the learning function in (2) gives:

$$\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(D_i^{\mathrm{lh}(\sigma)} \subseteq \mathrm{rng}(\sigma) \wedge \mathrm{rng}(\sigma) \subseteq L_i) \quad (3)$$

With the assumptions at hand, (3) is a recursive learning function which identifies $\mathcal{L}$. To sum up the foregoing, we have

Theorem 7 (Angluin) *An indexed family of recursive languages $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ is identifiable by a recursive function if and only if there is a limiting recursive function δ with $D_i = \delta(i)$ a telltale set for L_i in $\mathcal{L}$.*

Our statement of Angluin's theorem is slightly different from hers. Angluin (1980b) requires the set $\{ \langle i, x\rangle \mid x \in D_i \}$ to be r.e. This strengthening does not add anything, however. If D_i^n is as above, then $D_i' = (\bigcup_n D_i^n) \cap L_i$ is also a telltale set for L_i in $\mathcal{L}$. Clearly, $\{ \langle i, x\rangle \mid x \in D_i' \}$ is r.e.

4 Angluin's theorem for indexed families of r.e. sets

It is easy to see that one half of Angluin's theorem holds of indexed families of r.e. languages: if an indexed family of r.e. languages $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ is identifiable by a recursive function, then there is a limiting recursive function that maps each i to a telltale set for L_i in $\mathcal{L}$. The other direction fails in the following example.

Example 8 Let $\mathcal{L} = \{ K \cup \{i\} \mid i \in \mathbf{N} \}$. For each i, $\{i\}$ is a telltale set for $K \cup \{i\}$, which is trivially computable from i. However, $\mathcal{L}$ is not identifiable by a recursive function (Osherson, Stob, and Weinstein 1986).

The reason that Angluin's condition is not sufficient in the case of identification of indexed families of r.e. languages is that the learning function defined in (3) is no longer guaranteed to be recursive since the second conjunct $\mathrm{rng}(\sigma) \subseteq L_i$ may not be decidable. Informally speaking, in the case of indexed families of recursive languages, any conflict between the data and the conjecture can be recognized by the learner, but in the r.e. case, a special condition must obtain to enable the learner to detect such conflicts. We characterize this additional condition in terms of the presence of *warning sets*.

Definition 9 A finite set D is a *warning set* for L if $D \not\subseteq L$.

Definition 10 Let $\mathcal{L}$ be a class of languages, and let L be a language. A set E of finite sets is a *collection of warning sets* for L in $\mathcal{L}$ if the following conditions hold:

(i) $\forall D(D \in E \to D \not\subseteq L)$. ($E$ must consist of warning sets for L.)

(ii) $\forall L' \in \mathcal{L}(L' - L \neq \emptyset \to \exists D(D \in E \wedge D \subseteq L')$. (Each $L' \in \mathcal{L}$ must contribute to E a warning set for L if it can.)

A collection of warning sets is in general an infinite set of finite sets. Note that a collection of warning sets always exists, unlike telltale sets. That is why it doesn't figure in the watered-down version of Angluin's theorem characterizing non-effective identifiability of arbitrary language classes.

Let $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ be a class of languages. Define a (possibly non-recursive) learning function φ as follows:

$$\varphi(\sigma) \simeq \mu i(D_i \subseteq \mathrm{rng}(\sigma) \wedge \neg\exists D(D \in E_i \wedge D \subseteq \mathrm{rng}(\sigma))) \quad (4)$$

The above definition replaces the condition $\mathrm{rng}(\sigma) \subseteq L_i$ in (1) by $\neg\exists D(D \in E_i \wedge D \subseteq \mathrm{rng}(\sigma))$. If D_i is a telltale set for L_i in $\mathcal{L}$ and E_i is a collection of warning sets for L_i in $\mathcal{L}$ for every i, then we can show that φ identifies $\mathcal{L}$. In fact, a slightly weaker condition on E_i turns out to be sufficient.

Lemma 11 *Let φ be as defined in* (4). *Suppose that for each $i \in \mathbf{N}$, D_i is a telltale set for L_i in $\mathcal{L}$, and E_i is a collection of warning sets for L_i in $\{ L_j \mid D_i \subseteq L_j \}$. Then φ identifies $\mathcal{L}$.*

Proof. Let t be a text for $L \in \mathcal{L}$ with i least such that $L_i = L$. We show that φ converges on t to i. Note that for every $j < i$, if $D_j \subseteq L$, $L - L_j \neq \emptyset$ by the property of telltale sets. For any such j, there is a $D \subseteq L$ such that $D \in E_j$. Thus, we can find an m such that:

(i) $D_i \subseteq \mathrm{rng}(\bar{t}_m)$,

(ii) for every $j < i$, if $D_j \subseteq L$, then for some $D \in E_j$, $D \subseteq \mathrm{rng}(\bar{t}_m)$.

Then it is easy to see that for all $n \geq m$, $\varphi(\bar{t}_n) \simeq i$. $\dashv$

In our theorem characterizing when an indexed family of r.e. languages is identifiable, we demand that E_i be an r.e. set of finite sets whose r.e. index is computable in the limit from i. This, together with the assumption on D_i as before, turns out to be necessary and sufficient.

In the necessity half of Angluin's theorem, telltale sets were taken to be the ranges of locking sequences, and this works for the r.e. case as well. As for warning sets, we can take them to be the ranges of *unlocking sequences*.

Definition 12 A finite sequence τ is an *unlocking sequence* for φ and σ if and only if $\varphi(\sigma) \not\simeq \varphi(\sigma^\wedge\tau)$.

Lemma 13 *Let σ be a locking sequence for φ and L.*

(i) *If τ is an unlocking sequence for φ and σ, then* $\mathrm{rng}(\tau) \not\subseteq L$.

(ii) *If* $\mathrm{rng}(\sigma) \subseteq L' \neq L$ *and φ identifies L', then there is an unlocking sequence τ for φ and σ such that* $\mathrm{rng}(\tau) \subseteq L'$.

We are now ready to state our main theorem.

Theorem 14 *An indexed family of r.e. languages $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ is identifiable by a recursive function if and only if there are a limiting recursive function δ mapping each i to a finite set D_i and a limiting recursive function ε mapping i to an index e_i for an r.e. set $E_i = W_{e_i}$ of nonempty finite sets, satisfying:*

(i) $\forall i(D_i$ *is a telltale set for* L_i *in* $\mathcal{L})$;

(ii) $\forall i \forall D(D \in E_i \to D \not\subseteq L_i)$; *and*

(iii) $\forall i, j(D_i \subseteq L_j \wedge L_j - L_i \neq \emptyset \to \exists D(D \in E_i \wedge D \subseteq L_j))$.

Note that the conditions (ii) *and* (iii) *say that E_i is a collection of warning sets for L_i in $\{ L_j \mid D_i \subseteq L_j \}$.*

Proof. ($\Rightarrow$). Suppose φ is a recursive learning function that identifies $\mathcal{L}$. Since there is a uniform enumeration procedure for all L_i, the Locking Sequence Lemma gives a limiting recursive function γ such that for all i, $\gamma(i)$ is a locking sequence for φ and L_i. Define the limiting recursive functions δ and ε by

$$\delta(i) = \mathrm{rng}(\gamma(i)),$$
$$W_{\varepsilon(i)} = \{ \mathrm{rng}(\tau) \mid \varphi(\gamma(i)) \not\simeq \varphi(\gamma(i)^\wedge\tau) \}.$$

By Lemma 5, $D_i = \delta(i)$ satisfies (i). Lemma 13 implies that $E_i = W_{\varepsilon(i)}$ satisfies (ii) and (iii).

($\Leftarrow$). Define a learning function φ as follows:

$$\varphi(\sigma) \simeq \mu i(\delta(i) \subseteq \mathrm{rng}(\sigma) \wedge \neg\exists D(D \in W_{\varepsilon(i)} \wedge D \subseteq \mathrm{rng}(\sigma))) \tag{5}$$

Let e be such that $L_i = L_{e,i}$. By Lemma 11, $\lambda\sigma\, \mathbf{G}(e, \varphi(\sigma))$ identifies $\mathcal{L}$ (with respect to e). This learning function is not in general recursive, but a slight complication of Lemma 6 shows that it can be turned into a recursive learning function which identifies $\mathcal{L}$.

Lemma 15 *Let φ be a learning function defined by $\varphi(\sigma) \simeq \mu i(A(\sigma, i, \alpha(i)) \wedge \neg\exists y\, B(\sigma, i, \beta(i), y))$, where α and β are limiting recursive functions, A and B are recursive predicates, and B is 'persistent' in the first argument in the following sense:*

$$\forall\sigma\forall\tau\forall i\forall x\forall y(B(\sigma, i, x, y) \rightarrow B(\sigma^\wedge\tau, i, x, y)).$$

Then there is a recursive ψ such that for any text t, if $\lim_{n\to\infty} \varphi(\bar{t}_n)\downarrow$, then $\lim_{n\to\infty} \varphi(\bar{t}_n) \simeq \lim_{n\to\infty} \psi(\bar{t}_n)$.

Proof. Let total recursive F and G be such that for all i, $\alpha(i) \simeq \lim_{n\to\infty} F(i, n)$ and $\beta(i) \simeq \lim_{n\to\infty} G(i, n)$. We show that the recursive function ψ defined as follows satisfies the required property.

$$\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(A(\sigma, i, F(i, \mathrm{lh}(\sigma))) \wedge \neg\exists y \leq \mathrm{lh}(\sigma)\, B(\sigma, i, G(i, \mathrm{lh}(\sigma)), y)).$$

Suppose that φ converges on text t to index i. Then there is an m such that for all $n \geq m$, the following hold: $A(\bar{t}_n, i, \alpha(i)) \wedge \neg\exists y\, B(\bar{t}_n, i, \beta(i), y)$, and moreover for all $j < i$, either $\neg A(\bar{t}_n, i, \alpha(i))$ or $\exists y\, B(\bar{t}_n, i, \beta(i))$. For $j < i$, we distinguish two cases:

Case 1. For all $n \geq m$, $\neg A(\bar{t}_n, j, \alpha(j))$.

Case 2. For some $n \geq m$, $\exists y\, B(\bar{t}_n, j, \beta(j), y)$.

Let $m' \geq m$ be large enough to insure that:

(i) for all $j \leq i$ and for all $n \geq m'$, $F(j, n) = \alpha(j)$ and $G(j, n) = \beta(j)$,

(ii) for all $j < i$ such that Case 2 obtains, there are $n, y \leq m'$ such that $B(\bar{t}_n, j, \beta(j), y)$.

For $n \geq m'$, $A(\bar{t}_n, i, F(i, n)) \wedge \neg\exists y \leq n\, B(\bar{t}_n, i, G(i, n), y)$ holds by (i), and for all $j < i$ such that Case 1 obtains, $\neg A(\bar{t}_n, j, F(j, n))$. Conditions (i) and (ii) and the persistence of B ensure that for all $j < i$ such that Case 2 obtains, for all $n \geq m'$, $\exists y \leq n\, B(\bar{t}_n, j, G(j, n), y)$. Therefore, for all $n \geq m'$, $\psi(\bar{t}_n) \simeq i$. ⊣

We go back to the proof of Theorem 14. The learning function φ in (5) does not quite have the form required by Lemma 15, but it is easy to rewrite it in such a way that Lemma 15 can apply to it:

$$\varphi(\sigma) \simeq \mu i(\delta(i) \subseteq \mathrm{rng}(\sigma) \wedge \neg\exists n\exists D(D \in W_{\varepsilon(i),n} \wedge D \subseteq \mathrm{rng}(\sigma)) \tag{6}$$

Note that $\exists D(D \in W_{x,n} \wedge D \subseteq \mathrm{rng}(\sigma))$ is a recursive predicate persistent in σ. ⊣

Let us see how the proof of Lemma 15 gives an explicit definition of a recursive learning function ψ that identifies $\mathcal{L}$. Let D_i and e_i be as in Theorem 14, and let D_i^n and e_i^n, computable from i and n, be such that $\lim_{n\to\infty} D_i^n = D_i$ and $\lim_{n\to\infty} e_i^n = e_i$. Write E_i^n for $W_{e_i^n, n}$. Then an application of Lemma 15 to φ in (5) gives:

$$\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(D_i^{\mathrm{lh}(\sigma)} \subseteq \mathrm{rng}(\sigma) \wedge \neg\exists D(D \in E_i^{\mathrm{lh}(\sigma)} \wedge D \subseteq \mathrm{rng}(\sigma))) \tag{7}$$

Thus, $\lambda\sigma\, \mathbf{G}(e, \varphi(\sigma))$ is a recursive learning function which identifies $\mathcal{L}$ with respect to e.

As in the case of Angluin's theorem, Theorem 14 implies that if its characterizing condition holds of one indexing of a family, it holds of any other indexing of it. Note also that $\lambda\sigma\, \mathbf{G}(e, \psi(\sigma))$ defined by (7) identifies $\mathcal{L}$ with respect to e. So we have

Corollary 16 *An indexed family of r.e. languages $\mathcal{L}_e$ is identifiable if and only if it is identifiable w.r.t. e.*

Thus, the three senses for recursive identifiability of an indexed family of r.e. languages amount to the same thing. Nevertheless, it is important to distinguish (E), (P), and (C), as they can lead to different notions of identifiability when learning functions are required to behave in a certain restricted way. One such restriction is conservativity, which will be discussed in Section 6. The situation is parallel to the case of indexed families of recursive languages.[1]

5 Finite fatness

Angluin's theorem has led to the discovery of a number of sufficient conditions for identifiability of indexed families of recursive languages that are easy to apply and useful (Angluin 1980b, Wright 1989).

Definition 17 A class of languages $\mathcal{L}$ is of *finite thickness* if for every finite $D \neq \emptyset$, $\{ L \in \mathcal{L} \mid D \subseteq L \}$ is finite.

[1]Another easy corollary to Theorem 14 is that an indexed family consisting solely of infinite r.e. languages is identifiable by a recursive function if and only if it is identifiable by a set-driven recursive function (Osherson, Stob, and Weinstein 1986).

Definition 18 (i) A class of languages $\mathcal{L}$ is said to have *infinite elasticity* if there are an infinite sequence $s_0, s_1, s_2, \ldots$ of natural numbers and an infinite sequence $L_0, L_1, L_2, \ldots$ of languages in $\mathcal{L}$ such that for all $n \in \mathbf{N}$, $s_n \notin L_n$ and $\{s_0, \ldots, s_n\} \subseteq L_{n+1}$.

(ii) A class $\mathcal{L}$ of languages is said to have *finite elasticity* if it does not have infinite elasticity.

Finite thickness implies finite elasticity. Extending Angluin's result about finite thickness, Wright (1989) proves the following useful theorem, explicitly relying on Angluin's theorem.[2]

Theorem 19 (Wright) *Let $\mathcal{L}$ be an indexed family of recursive languages that has finite elasticity. Then $\mathcal{L}$ is identifiable by a recursive function.*

The indexed family of r.e. languages in Example 8 has finite elasticity; so it follows that finite elasticity is not a sufficient condition for identifiability of indexed families of r.e. languages. As the following example shows, even finite thickness is not sufficient in the r.e. case.[3]

Example 20 The languages $L_{2j} = \{\, \langle j, x\rangle \mid x \in \mathbf{N} \,\}$ and $L_{2j+1} = \{\, \langle j, x\rangle \mid \forall y \leq x\ \varphi_j(y)\!\downarrow \,\}$ determine an indexed family of r.e. languages $\mathcal{L} = \{\, L_i \mid i \in \mathbf{N} \,\}$ with finite thickness. Suppose $\mathcal{L}$ is identifiable by a recursive function φ. We get a contradiction. Let δ be the limiting recursive 'telltale function' given by Theorem 14. Then we have $L_{2j} = L_{2j+1}$ if and only if $\delta(2j) \subseteq L_{2j+1}$. The latter condition is r.e. in δ, and hence it is Σ^0_2. This means that $\{\, j \mid L_{2j} = L_{2j+1} \,\}$ is Σ^0_2. However, $L_{2j} = L_{2j+1}$ if and only if $W_j = \mathbf{N}$, and $\{\, j \mid W_j = \mathbf{N} \,\}$ is known to be Π^0_2-complete, a contradiction.

Note that the two families in Examples 8 and 20 fail to be identifiable for different reasons. The former lacks a limiting recursive 'warning function', and the latter lacks a limiting recursive telltale function.

Definition 21 A class of r.e. languages $\mathcal{L}$ is said to be *r.e. indexable without repetition* if it has an r.e. indexing $L_0, L_1, L_2, \ldots$ such that $i \neq j$ implies $L_i \neq L_j$.

It is not really hard to see that a class $\mathcal{L}$ is r.e. indexable without repetition if it has an indexing $L_0, L_1, L_2, \ldots$ such that $\{\, \langle i, j\rangle \mid L_i \neq L_j \,\}$ is r.e.

It can be shown that if a class of r.e. languages has finite thickness and is moreover r.e. indexable without repetition, it is identifiable by a recursive function. In fact, if a class is r.e. indexable without repetition, a condition much weaker than finite thickness is sufficient for identifiability by a recursive function.

[2] Wright's original defective definition of finite elasticity was later corrected by Motoki, Shinohara and Wright (1991).

[3] This example is due to Bas Terwijn.

Definition 22 A class of languages $\mathcal{L}$ is said to have *finite fatness* if every $L \in \mathcal{L}$ has a *strict telltale set* D in $\mathcal{L}$, i.e., for all $L' \in \mathcal{L}$, $D \subseteq L'$ implies $L \subseteq L'$), and, moreover each $L \in \mathcal{L}$ is contained in only finitely many other languages in $\mathcal{L}$.

It is easy to see that a class of languages has finite fatness if and only if for each $L \in \mathcal{L}$ there exists a finite subset D of L such that $\{\, L' \in \mathcal{L} \mid D \subseteq L' \,\}$ is finite.

Theorem 23 *Let $\mathcal{L}$ be a class of r.e. languages which is r.e. indexable without repetition. If $\mathcal{L}$ has finite fatness, it is identifiable by a recursive function.*

Proof. Let $L_0, L_1, L_2, \ldots$ be an r.e. indexing of $\mathcal{L}$ such that for $i \neq j$, $L_i \neq L_j$. Let F be a total recursive function such that $L_i = W_{F(i)}$. We show that $\mathcal{L}$ satisfies the condition in Theorem 14 with respect to the indexing $L_0, L_1, L_2, \ldots$.

First, we show a limiting recursive procedure computing a finite subset D^0_i of L_i such that $\{\, j \mid D^0_i \subseteq L_j \,\}$ is finite. The fact that the predicate $P(i,n)$ defined by

$$P(i,n) \Leftrightarrow |\{\, j \mid W_{F(i),n} \subseteq W_{F(j)} \,\}| \geq n$$

is r.e. makes $m_i = \mu n\, \neg P(i,n)$, which always exists because $\mathcal{L}$ has finite fatness, computable in the limit from i. Also, $D^0_i = W_{F(i),m_i}$ is computable in the limit from i, as is the finite set $J_i = \{\, j \mid j \neq i \wedge D^0_i \subseteq L_j \,\}$. Next, note that for $i \neq j$,

$$d_{ij} = \mu x (x \in (L_i - L_j) \cup (L_j - L_i))$$

always exists by assumption. Moreover, d_{ij} is computable in the limit given i and j with $i \neq j$. To see this, let

$$d^n_{ij} \simeq \mu x (x \in (W_{F(i),n} - W_{F(j),n}) \cup (W_{F(j),n} - W_{F(i),n})).$$

Then $d_{ij} = \lim_{n\to\infty} d^n_{ij}$. Finally, we show that a telltale set D_i for L_i in $\mathcal{L}$ and a finite set E_i such that

- $E_i \cap L_i = \emptyset$, and
- $i \neq j \wedge D_i \subseteq L_j \rightarrow E_i \cap L_j \neq \emptyset$

are both computable in the limit from i, D^0_i, and J_i. Since the limiting recursive functions are closed under composition, the two sets are computable in the limit from i. D_i and $\{\, \{x\} \mid x \in E_i \,\}$, which is also computable in the limit from i, satisfy the conditions in Theorem 14.

Define D^n_i as follows:

$$D^n_i = D^0_i \cup \{\, d^n_{ij} \mid j \in J_i \wedge d^n_{ij} \in W_{F(i),n} - W_{F(j),n} \,\}.$$

D_i^n is computable from i, n, D_i^0, J_i. Since $\lim_{n\to\infty} d_{ij}^n$ exists for $j \in J_i$ and J_i is finite, it is not hard to see that $D_i = \lim_{n\to\infty} D_i^n$ exists and

$$D_i = D_i^0 \cup \{ d_{ij} \mid j \in J_i \wedge d_{ij} \in L_i - L_j \}.$$

If $D_i^0 \subseteq L_j \subset L_i$, then $d_{ij} \in L_i - L_j$, so $d_{ij} \in D_i$ and $D_i \not\subseteq L_j$. This shows that D_i is a telltale set for L_i in $\mathcal{L}$, as desired.

The set E_i can be defined as the limit of E_i^n, defined below:

$$E_i^n = \{ d_{ij}^n \mid j \in J_i \wedge d_{ij}^n \in W_{F(j),n} - W_{F(i),n} \},$$

which is computable from J_i, i, n. Again, it is not hard to see that $E_i = \lim_{n\to\infty} E_i^n$ exists and

$$E_i = \{ d_{ij} \mid j \in J_i \wedge d_{ij} \in L_j - L_i \}.$$

Clearly, $E_i \cap L_i = \emptyset$. Note that for $j \in J_i$, either $d_{ij} \in D_i$ or $d_{ij} \in E_i$, but not both. If $i \neq j$ and $D_i \subseteq L_j$, then $j \in J_i$ and $d_{ij} \notin D_i$, so $d_{ij} \in E_i$. Thus, E_i satisfies the required properties. $\dashv$

Obviously, finite thickness implies finite fatness. Neither finite fatness nor finite elasticity implies the other.[4] We cannot replace finite fatness by finite elasticity in Theorem 23, as the following example shows.

Example 24 Let

$$L_0 = \{ \langle 0, x\rangle \mid x \in K \},$$
$$L_{i+1} = \begin{cases} \{ \langle 0, x\rangle \mid x \in K \cup \{i\} \} & \text{if } i \notin K, \\ \{ \langle 0, x\rangle \mid x \in K \} \cup \{\langle 1, i\rangle\} & \text{if } i \in K. \end{cases}$$

Thus, $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ is a family of r.e. languages r.e. indexable without repetition. This $\mathcal{L}$ has finite elasticity, but not finite fatness, and is not identifiable by a recursive function since it contains an isomorphic copy of the family in Example 8.

6 Conservativity and related restrictions

Zeugmann, Lange, and Kapur (1992, 1995) show a way of modifying Angluin's characterizing condition to characterize conservative identifiability of indexed families of recursive languages. Here, we modify the condition in Theorem 14 to characterize conservative identifiability of indexed families of r.e. languages. Lange and Zeugmann (1993) show in the case of indexed families of recursive languages that the three notions of conservative identifiability defined in terms of (E), (P), and (C) are all non-equivalent. The same holds in the case of indexed families of r.e. languages.

Here we will concentrate on the first notion of conservative identifiability defined in terms of (E). It admits of a natural characterizing condition, which immediately yields a characterization for the (P) version. So far, we have not been able to find a characterization for the (C) version.

An intriguing fact about the r.e. case is that we can discern at least three types of conservative-like behavior, which yield equivalent notions of identifiability in the recursive case but are distinguished from each other in the r.e. case. They are characterized by successive strengthenings of the condition in Theorem 14.

Definition 25 A learning function φ is said to be[5] (i) *non-contracting*, (ii) *conservative*, (iii) *strictly conservative*, respectively, if and only if

(i) $\forall \sigma, \tau (\varphi(\sigma)\downarrow \wedge \varphi(\sigma^\wedge\tau)\downarrow \rightarrow W_{\varphi(\sigma)} \not\supset W_{\varphi(\sigma^\wedge\tau)})$

(ii) $\forall \sigma, x (\varphi(\sigma)\downarrow \wedge \mathrm{rng}(\sigma^\wedge\langle x\rangle) \subseteq W_{\varphi(\sigma)} \rightarrow \varphi(\sigma^\wedge\langle x\rangle) \simeq \varphi(\sigma))$

(iii) $\forall \sigma, x (\varphi(\sigma)\downarrow \wedge x \in W_{\varphi(\sigma)} \rightarrow \varphi(\sigma^\wedge\langle x\rangle) \simeq \varphi(\sigma))$.

Note that these restrictions all imply that the learner never *overgeneralizes* on a text for a language it identifies. That is, if φ satisfies one of the three restrictions and if φ identifies L, then for all texts t for L, it is never the case that $W_{\varphi(\bar{t}_n)} \supset L$.

A strictly conservative learning function is conservative, so trivially, identifiability by a strictly conservative recursive function implies identifiability by a conservative recursive function. We will show that identifiability by a conservative recursive function w.r.t. the given index implies identifiability by a non-contracting recursive function w.r.t. the same index. Neither of these implications can be reversed.

6.1 Characterizations

We start in the middle, with conservative identifiability. Recall Zeugmann, Lange, and Kapur's (1992, 1995) characterization of conservative identifiability in the recursive case, which we state as follows:

Theorem 26 (Zeugmann, Lange, and Kapur) *Let e be an index for a total recursive function, and*

[4] Finite elasticity does imply that each language in the class has a strict telltale set.

[5] There are weaker versions of these definitions, where φ is required to behave in the prescribed way only on texts for languages in the class to be identified. The proofs of all theorems below go through with this modification, except for that of Theorem 29. Jain and Sharma (1994) show in the case of the (C) version of conservative identifiability the two variants to be equivalent .

write L_i for $L_{e,i}$. The indexed family of recursive languages $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$ is identifiable w.r.t. e by a conservative recursive function if and only if there is a partial recursive function δ such that $\{ L_i \mid \delta(i)\downarrow \} = \mathcal{L}$ and $\delta(i)$, if defined, is a telltale set for L_i in $\mathcal{L}$.

In the recursive case, the same condition characterizes both non-contracting and strictly conservative identifiability, the three restrictions amount to the same thing:

Theorem 27 *Let $\mathcal{L} = \mathcal{L}_e$ be an indexed family of recursive languages, and let φ be a non-contracting recursive learning function that identifies $\mathcal{L}$ with respect to e. Then there is a strictly conservative recursive learning function ψ that identifies $\mathcal{L}$ with respect to e.*

PROOF. Since $\mathcal{L}$ is a family of *recursive* languages, we can assume without loss of generality that for all σ, if $\varphi(\sigma)\downarrow$, $\mathrm{rng}(\sigma) \subseteq W_{\varphi(\sigma)}$. Define ψ as follows:

$$\psi(\sigma^\frown\langle x\rangle) \simeq \begin{cases} \psi(\sigma) & \text{if } \psi(\sigma)\downarrow \text{ and } x \in W_{\psi(\sigma)}, \\ \varphi(\sigma^\frown\langle x\rangle) & \text{otherwise.} \end{cases}$$

This definition makes ψ strictly conservative. Like φ, ψ is consistent in the sense that if $\varphi(\sigma)\downarrow$, $\mathrm{rng}(\sigma) \subseteq W_{\varphi(\sigma)}$, and ψ is also recursive since $\mathcal{L}_e$ is an indexed family of recursive languages and $\mathrm{rng}(\psi) \subseteq \mathrm{rng}(\varphi) \subseteq \{ \mathbf{G}(e,i) \mid i \in \mathbf{N} \}$. Let t be a text for $L \in \mathcal{L}$. Since φ identifies $\mathcal{L}$, φ must converge to an index i such that $L = W_i$. Let m be such that $\forall n \geq m(\varphi(\bar{t}_n) \simeq i)$. We show that ψ also converges to an index for L on t. First, note that for all n, $\psi(\bar{t}_n) \simeq \varphi(\bar{t}_{n'})$ for some $n' \leq n$. Note also that since φ is non-contracting, for all n, $W_{\varphi(\bar{t}_n)} \not\supset L$. It follows that for all n, $W_{\psi(\bar{t}_n)} \not\supset L$ holds as well. There are two cases to consider.

CASE 1. $\psi(\bar{t}_{n-1}) \not\simeq \psi(\bar{t}_n)$ for some $n \geq m$. Then it must be that $\psi(\bar{t}_n) \simeq \varphi(\bar{t}_n) \simeq i$. It follows that for all $n' \geq n$, $\psi(\bar{t}_{n'}) \simeq i$.

CASE 2. There is an $m' < m$ such that for all $n \geq m'$, $\psi(\bar{t}_n) \simeq \psi(\bar{t}_{m'})$. Since $\varphi(\bar{t}_n)\downarrow$ for all $n \geq m$, it must be that $\psi(\bar{t}_{m'})\downarrow$. Since ψ is consistent, $L \subseteq W_{\psi(\bar{t}_{m'})}$. Since $L \not\subset W_{\psi(\bar{t}_{m'})}$, we have $W_{\psi(\bar{t}_{m'})} = L$. So ψ converges to an index for L on t.

Since $\mathrm{rng}(\psi) \subseteq \mathrm{rng}(\varphi)$, ψ identifies $\mathcal{L}$ w.r.t. e. ⊣

Our characterization of conservative identifiability in the r.e. case is a natural analogue of Zeugmann, Lange, and Kapur's theorem.

Theorem 28 *Let $\mathcal{L} = \mathcal{L}_e$ be an indexed family of r.e. languages, and write L_i for $L_{e,i}$. $\mathcal{L}$ is identifiable with respect to e by a conservative recursive function if and only if there are partial recursive functions δ and ε with a common domain I, and δ maps $i \in I$ to a finite set D_i, and ε maps $i \in I$ to an index e_i for an r.e. set $E_i = W_{e_i}$ of nonempty finite sets, satisfying:*

(i) $\{ L_i \mid i \in I \} = \mathcal{L}$;

(ii) $\forall i \in I(D_i$ *is a telltale set for* L_i *in* $\mathcal{L})$;

(iii) $\forall i \in I \forall D(D \in E_i \to D \not\subseteq L_i)$;

(iv) $\forall i \in I \forall j(D_i \subseteq L_j \wedge L_j - L_i \neq \emptyset \to \exists D(D \in E_i \wedge D \subseteq L_j))$.

PROOF. ($\Rightarrow$). The idea is essentially the same as Zeugmann, Lange, and Kapur's (1992, 1995) proof. Suppose φ is a conservative recursive function and φ identifies $\mathcal{L}$ with respect to e. The crucial observation is

- For all $i \in \mathbf{N}$, if $\varphi(\sigma) \simeq \mathbf{G}(e,i)$ and $\mathrm{rng}(\sigma) \subseteq L_i$, then σ is a locking sequence for φ and L_i.

Define

$$\gamma(i) \simeq \begin{cases} \sigma & \text{if } \langle\sigma, s\rangle \text{ is the least pair such that} \\ & \varphi(\sigma) \simeq \mathbf{G}(e,i) \text{ and} \\ & \mathrm{rng}(\sigma) \subseteq W_{\mathbf{G}(e,i),s}, \\ \text{undefined} & \text{if there is no such pair.} \end{cases}$$

γ is partial recursive, and since φ identifies $\mathcal{L}$ with respect to e, $\{ L_i \mid i \in \mathrm{dom}(\gamma) \} = \mathcal{L}$. Then δ and ε defined by

$$\begin{aligned} \delta(i) &\simeq \mathrm{rng}(\gamma(i)), \\ W_{\varepsilon(i)} &= \{ \mathrm{rng}(\tau) \mid \varphi(\gamma(i)) \not\simeq \varphi(\gamma(i)^\frown\tau) \} \quad \text{if } \gamma(i)\downarrow \end{aligned}$$

satisfy the required properties.

($\Leftarrow$). This direction is slightly more complex than in the recursive case. Suppose that δ and ε are partial recursive functions satisfying (i)–(iv), and let k be such that $W_k = I = \mathrm{dom}(\delta) = \mathrm{dom}(\varepsilon)$. We write I^s for $W_{k,s}$. Define a learning function ψ as follows:

$$\begin{aligned} \psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(i \in I^{\mathrm{lh}(\sigma)} \wedge \delta(i) \subseteq \mathrm{rng}(\sigma) \wedge \\ \neg\exists D(D \in W_{\varepsilon(i),\mathrm{lh}(\sigma)} \wedge D \subseteq \mathrm{rng}(\sigma))). \end{aligned}$$

As in the proof of Theorem 14, one can see that $\lambda\sigma\, \mathbf{G}(e, \psi(\sigma))$ identifies $\mathcal{L}$ with respect to e, using condition (i). This learning function is not necessarily conservative, but one can define a conservative learning function in terms of it. Note that ψ never overgeneralizes: If t is a text for some $L_j \in \mathcal{L}$, then for all n, $\varphi(\bar{t}_n) \simeq i$ implies $L_j \not\subset L_i$. This is so because if $\varphi(\sigma) \simeq i$, $\mathrm{rng}(\sigma)$ must include a telltale set for L_i in $\mathcal{L}$.

Define a learning function φ by

$$\varphi(\sigma^\frown\langle x\rangle) \simeq \begin{cases} i & \text{if } \varphi(\sigma) \simeq i \text{ and} \\ & \neg\exists D(D \in W_{\varepsilon(i),\mathrm{lh}(\sigma^\frown\langle x\rangle)} \wedge \\ & D \subseteq \mathrm{rng}(\sigma^\frown\langle x\rangle)), \\ \psi(\sigma^\frown\langle x\rangle) & \text{otherwise.} \end{cases}$$

Note that the condition for the first case is decidable, since if $\varphi(\sigma) \simeq i$, then $\varepsilon(i)\downarrow$. This definition guarantees

φ to be conservative. We show that whenever ψ converges to j on text t for L_j, φ converges on t to some i such that $L_i = L_j$. It then follows that $\lambda\sigma\, \mathbf{G}(e, \varphi(\sigma))$ identifies $\mathcal{L}$ with respect to e and is conservative.

Let t be a text for some $L \in \mathcal{L}$ and suppose ψ converges on t to some j such that $L_j = L$. By the definition of φ, for all $n \in \mathbf{N}$, $\varphi(\bar{t}_n) \simeq \psi(\bar{t}_{n'})$ for some $n' \leq n$, and if $\varphi(\bar{t}_n) \not\simeq \varphi(\bar{t}_{n+1})$, then $\varphi(\bar{t}_{n+1}) \simeq \psi(\bar{t}_{n+1})$. Since ψ never overgeneralizes, nor does φ. Let m be such that for all $n \geq m$, $\psi(\bar{t}_n) \simeq j$. We distinguish two cases.

Case 1. For some $m' \geq m$, $\varphi(\bar{t}_{m'}) \simeq \psi(\bar{t}_{m'})$. Then by the definition of ψ and φ, it is easy to see that for all $n \geq m'$, $\varphi(\bar{t}_n) \simeq j$.

Case 2. There is an $m' < m$ such that $\varphi(\bar{t}_{m'}) \simeq i$ and for all $n \geq m'$, $\varphi(\bar{t}_n) \simeq i$. By the definitions of ψ and φ, $\delta(i)\downarrow$ and $\delta(i) \subseteq \mathrm{rng}(\bar{t}_{m'})$. So $D_i \subseteq L_j$. Since φ never overgeneralizes and t is a text for L_j, $L_i \not\supset L_j$. Also, for all $n \geq m'$, we have $\neg\exists D(D \in W_{\varepsilon(i),n} \wedge D \subseteq \mathrm{rng}(\bar{t}_n))$. This implies $\neg\exists D(D \in W_{\varepsilon(i)} \wedge D \subseteq L_j)$. From this and $D_i \subseteq L_j$, using condition (iv) on ε, we get $L_j - L_i = \emptyset$, that is, $L_i \supseteq L_j$. So $L_i = L_j$. ⊣

Next we turn to strict conservativity. The characterizing condition for strict conservative identifiability of course implies that for conservative identifiability.

Theorem 29 *Let $\mathcal{L} = \mathcal{L}_e$ be an indexed family of r.e. languages, and write L_i for $L_{e,i}$. $\mathcal{L}$ is identifiable with respect to e by a strictly conservative recursive function if and only if there are partial recursive functions δ and ε with a common domain I, and δ maps $i \in I$ to a finite set D_i, and ε maps $i \in I$ to an index e_i for the characteristic function of a recursive set E_i, satisfying the following conditions:*

(i) $\{ L_i \mid i \in I \} = \mathcal{L}$;

(ii) $\forall i \in I(D_i$ *is a telltale set for* L_i *in* $\mathcal{L})$;

(iii) $\forall i \in I(E_i \cap L_i = \emptyset)$;

(iv) $\forall i \in I \forall j (D_i \subseteq L_j \wedge L_j - L_i \neq \emptyset \rightarrow E_i \cap (L_j - L_i) \neq \emptyset)$.

Proof. ($\Rightarrow$). Suppose φ is a strictly conservative recursive function and φ identifies $\mathcal{L}$ with respect to e. As in the proof of Theorem 28, we can find a partial recursive function γ such that $\{ L_i \mid i \in \mathrm{dom}(\gamma) \} = \mathcal{L}$ and $\gamma(i)$ is a locking sequence for φ and L_i, if defined. As before, we can take $\delta(i) \simeq \mathrm{rng}(\gamma(i))$. To construct the required recursive set E_i, we make use of the fact that φ is strictly conservative.

Claim. Let σ be a locking sequence for φ and L_i, and $\mathrm{rng}(\sigma) \subseteq L_j \neq L_i$. Then there is a strictly increasing unlocking sequence τ for φ and σ such that $\mathrm{rng}(\tau) \subseteq L_j$.

Proof of the Claim. If L_j is infinite, let t be a strictly increasing text for L_j. Since φ must converge on $\sigma^\frown t$ to an index for L_j, there must be an m such that $\varphi(\sigma) \not\simeq \varphi(\sigma^\frown \bar{t}_m)$. If L_j is finite, let τ be the strictly increasing finite sequence such that $\mathrm{rng}(\tau) = L_j$. Let t be the text $x, x, x, \ldots$, where x is the first element of σ. Learner φ must converge on $\sigma^\frown\tau^\frown t$ to an index for L_j. If $\varphi(\sigma) \simeq \varphi(\sigma^\frown\tau)$, then φ converges on this text to an index for L_i, because $x \in L_i$ and φ is strictly conservative. But this is a contradiction, so τ must be an unlocking sequence. ⊣

Going back to the proof of Theorem 29, let

$$E_i = \{ x \mid \exists\tau(\tau^\frown\langle x\rangle \text{ is strictly increasing} \wedge \varphi(\gamma(i)) \simeq \varphi(\gamma(i)^\frown\tau) \not\simeq \varphi(\gamma(i)^\frown\tau^\frown\langle x\rangle)) \}.$$

It is not hard to see that E_i is a recursive set and an index for its characteristic function is computable from $\gamma(i)$. It remains to show that E_i satisfies (iii) and (iv).

Since φ is strictly conservative, if $x \in E_i$, $x \notin L_i$. So (iii) holds. To see that (iv) holds, assume $\gamma(i) \subseteq L_j \neq L_i$ and let τ be a shortest strictly increasing unlocking sequence for φ and $\gamma(i)$ with $\mathrm{rng}(\tau) \subseteq L_j$, which exists by the above claim. Then the last element x of τ must be in $L_j - L_i$.

($\Leftarrow$). Suppose that δ and ε are partial recursive functions satisfying (i)–(iv), and let k be such that $W_k = I = \mathrm{dom}(\delta) = \mathrm{dom}(\varepsilon)$. We write I^s for $W_{k,s}$. Let

$$\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(i \in I^{\mathrm{lh}(\sigma)} \wedge \delta(i) \subseteq \mathrm{rng}(\sigma) \wedge \neg\exists x(x \in \mathrm{rng}(\sigma) \wedge \varphi_{\varepsilon(i)}(x) = 1)).$$

Define a learning function φ as follows:

$$\varphi(\sigma^\frown\langle x\rangle) \simeq \begin{cases} i & \text{if } \varphi(\sigma) \simeq i \text{ and } \varphi_{\varepsilon(i)}(x) = 0, \\ \psi(\sigma^\frown\langle x\rangle) & \text{otherwise.} \end{cases}$$

Then $\lambda\sigma\, \mathbf{G}(e, \varphi(\sigma))$ is strictly conservative and identifies $\mathcal{L}$ with respect to e. We omit the details. ⊣

Finally, we present a condition characterizing non-contracting identifiability. The theorem implies that conservative identifiability with respect to some index implies non-contracting identifiability with respect to the same index, which is not immediately clear from the definitions.

Theorem 30 *Let $\mathcal{L} = \mathcal{L}_e$ be an indexed family of r.e. languages, and write L_i for $L_{e,i}$. $\mathcal{L}$ is identifiable with respect to e by a non-contracting recursive function if and only if there are a partial recursive function β mapping each $i \in \mathrm{dom}(\beta)$ to a finite set B_i, a limiting recursive function δ mapping each i to a finite set D_i, and a limiting recursive function ε mapping each i to an index e_i for an r.e. set E_i of nonempty finite sets, satisfying the following properties:*

(i) $\{ L_i \mid i \in \mathrm{dom}(\beta) \} = \mathcal{L}$;

(ii) $\forall i \in \mathrm{dom}(\beta)(B_i$ *is a telltale set for* L_i *in* $\mathcal{L})$;

(iii) $\forall i(D_i$ *is a telltale set for* L_i *in* $\mathcal{L})$;

(iv) $\forall i(D \in E_i \to D \not\subseteq L_i)$;

(v) $\forall i \forall j(D_i \subseteq L_j \wedge L_j - L_i \neq \emptyset \to \exists D(D \in E_i \wedge D \subseteq L_j))$.

PROOF. ($\Rightarrow$). This direction is easy. Let φ be a non-contracting recursive function identifying $\mathcal{L}$ with respect to e. Since the conditions on δ and ε are identical to those in Theorem 14, we only need to worry about β.

Note that if $\varphi(\sigma) \simeq \mathbf{G}(e,i)$ and $\mathrm{rng}(\sigma) \subseteq L_i$, σ may not be a locking sequence for φ and L_i, but $\mathrm{rng}(\sigma)$ must be a telltale set for L_i in $\mathcal{L}$. For, if $\mathrm{rng}(\sigma) \subseteq L_j \subset L_i$ and t is a text for L_j, then for all n, $W_{\varphi(\sigma^\wedge \bar{t}_n)} \neq L_j$ because φ is non-contracting. This means that φ fails to identify L_j on $\sigma^\wedge t$, a text for L_j.

So we can define β as follows:

$$\beta(i) \simeq \begin{cases} \mathrm{rng}(\sigma) & \text{if } \langle \sigma, s\rangle \text{ is the least pair such that} \\ & \varphi(\sigma) \simeq \mathbf{G}(e,i) \text{ and} \\ & \mathrm{rng}(\sigma) \subseteq W_{\mathbf{G}(e,i),s}\,, \\ \text{undefined} & \text{if there is no such pair.} \end{cases}$$

($\Leftarrow$). Let β, δ, and ε be functions satisfying the required properties, and let k be such that $\beta = \varphi_k$. We write $\beta^s(i)$ for $\varphi_{k,s}(i)$. Let D_i^n and e_i^n, computable from i and n, be such that $\lim_{n\to\infty} D_i^n = D_i$ and $\lim_{n\to\infty} e_i^n = e_i$. Write E_i^n for $W_{e_i^n,n}$. Define

$$\begin{aligned}\psi(\sigma) \simeq \mu i \leq \mathrm{lh}(\sigma)(\beta^{\mathrm{lh}(\sigma)}(i){\downarrow} \wedge \beta(i) \subseteq \mathrm{rng}(\sigma) \wedge \\ D_i^{\mathrm{lh}(\sigma)} \subseteq \mathrm{rng}(\sigma) \wedge \\ \neg\exists D(D \in E_i^{\mathrm{lh}(\sigma)} \wedge D \subseteq \mathrm{rng}(\sigma))).\end{aligned}$$

The definition of ψ is more complex than the definition of φ in the ($\Leftarrow$) direction of the proof of Theorem 14, but it is not hard to see that on any text t for some $L \in \mathcal{L}$, φ converges to the least i such that $L = L_i$ and $\beta(i){\downarrow}$. So $\lambda\sigma\,\mathbf{G}(e, \psi(\sigma))$ identifies $\mathcal{L}$ with respect to e.

We haven't made ψ non-contracting, but with the help of β, we can weed out all offending guesses of ψ to obtain a non-contracting learning function. Define

$$\varphi(\sigma) \simeq \begin{cases} i & \text{if } \psi(\sigma) \simeq i \text{ and for all initial} \\ & \text{segments } \tau \text{ of } \sigma, \\ & \varphi(\tau) \simeq j \to \beta(j) \subseteq W_{\mathbf{G}(e,i),\mathrm{lh}(\sigma)}, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Note that the condition in the first case is decidable. If $\varphi(\sigma) \simeq j$ and $\varphi(\sigma^\wedge\tau) \simeq i$, then $\beta(j) \subseteq L_i$, so $L_i \not\subset L_j$. Thus φ is non-contracting. To see that φ identifies $\mathcal{L}$ with respect to e, let t be a text for some $L \in \mathcal{L}$, and let i be the least such that $L = L_i$ and $\beta(i){\downarrow}$. We have noted that ψ converges on t to i, so let m be such that $\forall n \geq m(\psi(\bar{t}_n) \simeq i)$. Let $J = \{ j \mid \exists n < m(\varphi(\bar{t}_n) \simeq j) \}$. By the definitions of ψ and φ, for all $j \in J$, $\beta(j){\downarrow}$ and $\beta(j) \subseteq L_i$. So we can take an $m' \geq m$ large enough that for all $j \in J \cup \{i\}$, $\beta(j) \subseteq W_{\mathbf{G}(e,i),m'}$. Then for all $n \geq m'$, $\varphi(\bar{t}_n) \simeq i$. So φ converges on t to i. $\dashv$

Since the condition in Theorem 28 obviously implies that in Theorem 30, the not immediately obvious fact follows that if $\mathcal{L}_e$ is identifiable with respect to e by a conservative recursive function, it is identifiable with respect to e by a non-contracting recursive function.

6.2 Separations

We now show that the three restrictions we have introduced indeed lead to non-equivalent notions of identifiability. Wherever possible, we will state the results in the most informative form, comparing identifiability with respect to an index by a learning function satisfying a weaker restriction with identifiability simpliciter by a learning function satisfying a stronger restriction.

Theorem 31 *There is an indexed family of r.e. languages $\mathcal{L} = \mathcal{L}_e$ which is identifiable by a recursive function (with respect to e), but not identifiable by a non-contracting recursive function.*

PROOF. Let

$$\mathcal{L} = \{ D \mid D \subseteq K \wedge D \text{ is finite} \} \cup \{ K \cup \{x\} \mid x \notin K \}.$$

To see that $\mathcal{L}$ is an indexed family, first note that the set $A = \{ D \mid D \subseteq K \wedge D \text{ is finite} \}$ is r.e. So there is a total recursive function G such that $\mathrm{rng}(G) = A$. Let F be a total recursive function that turns a finite set into an r.e. index of it, and define $L_{2i} = W_{F(G(i))}$. Then $\{ L_{2i} \mid i \in \mathbf{N} \} = \{ D \mid D \subseteq K \wedge D \text{ is finite} \}$. Next, let $x_0 \notin K$, and let H be a total recursive function such that

$$y \in W_{H(x)} \Leftrightarrow y \in K \vee y = x \vee (x \in K \wedge y = x_0).$$

Then

$$W_{H(x)} = \begin{cases} K \cup \{x\} & \text{if } x \notin K, \\ K \cup \{x_0\} & \text{if } x \in K. \end{cases}$$

If we define $L_{2i+1} = W_{H(i)}$, then $L_0, L_1, L_2, \ldots$ is a uniformly r.e. sequence and $\mathcal{L} = \{ L_i \mid i \in \mathbf{N} \}$.

The class $\mathcal{L}$ can be identified by the following strategy: On input σ, conjecture $\mathrm{rng}(\sigma)$ if all elements in $\mathrm{rng}(\sigma)$ can be shown to be in K within $\mathrm{lh}(\sigma)$ steps; otherwise, conjecture $K \cup \{x\}$, where x is the element that appears first in σ that is not shown to be in K within $\mathrm{lh}(\sigma)$ steps. By Corollary 16, $\mathcal{L}$ can be identified with respect to any index of it.

Now suppose φ is a non-contracting recursive function identifying $\mathcal{L}$. We get a contradiction. Since φ is non-contracting, if $\mathrm{rng}(\sigma) \subseteq K$, $\mathrm{rng}(\sigma) \not\subset W_{\varphi(\sigma)}$, for otherwise φ fails to identify $\mathrm{rng}(\sigma)$. On the other hand, if $x \notin K$, φ must converge to an index for $K \cup \{x\}$ on $\langle x\rangle^\wedge t$, where t is a text for K. So, for some σ, $\mathrm{rng}(\sigma) \subseteq K$ and $\mathrm{rng}(\langle x\rangle^\wedge\sigma) \subset W_{\varphi(\langle x\rangle^\wedge\sigma)}$. Thus, we have

$$\begin{aligned} x \notin K \Leftrightarrow \exists\sigma(&\mathrm{rng}(\sigma) \subseteq K \wedge \varphi(\langle x\rangle^\wedge\sigma)\downarrow \wedge \\ &\mathrm{rng}(\langle x\rangle^\wedge\sigma) \subset W_{\varphi(\langle x\rangle^\wedge\sigma)}). \end{aligned}$$

But the right-hand side is easily seen to be r.e., contradicting the fact that $\overline{K}$ is not r.e. $\dashv$

Theorem 32 *There is an indexed family of r.e. languages $\mathcal{L} = \mathcal{L}_e$ which is identifiable by a non-contracting recursive function with respect to e, but not identifiable by a conservative recursive function with respect to any index f such that $\mathcal{L} = \mathcal{L}_f$.*

PROOF. Consider the subclass $\mathcal{L} = \{\, K \cup \{x\} \mid x \notin K \,\}$ of the class defined and shown to be identifiable by a recursive learner in the proof of Theorem 31. As was shown in the proof of Theorem 31, $\mathcal{L}$ is also an indexed family, so by Corollary 16, there is a recursive learner ψ that identifies $\mathcal{L}$ with respect to some e such that $\mathcal{L} = \mathcal{L}_e$. But ψ must be non-contracting, since no two languages in $\mathcal{L}$ stand in the proper inclusion relationship.

Now assume that there is an index f of $\mathcal{L}$ and a conservative recursive function φ that identifies $\mathcal{L}$ with respect to f. This assumption leads to a contradiction. Observe that there must be a σ such that $\mathrm{rng}(\sigma) \subseteq K$ and $\varphi(\sigma)\downarrow$. For, if not,

$$x \notin K \Leftrightarrow \exists\sigma(\mathrm{rng}(\sigma) \subseteq K \wedge \varphi(\langle x\rangle^\wedge\sigma)\downarrow),$$

and the right-hand side of this equivalence is r.e., yielding a contradiction. So take a σ with $\mathrm{rng}(\sigma) \subseteq K$ and $\varphi(\sigma)\downarrow$. Since $W_{\varphi(\sigma)} \in \mathcal{L}$, $W_{\varphi(\sigma)} = K \cup \{x\}$ for some $x \notin K$. Since $\mathrm{rng}(\sigma) \subseteq K$ and φ is conservative, for all τ such that $\mathrm{rng}(\tau) \subseteq K \cup \{x\}$, $\varphi(\sigma^\wedge\tau) \simeq \varphi(\sigma)$. Also, for all $y \notin K$, φ must converge on $\sigma^\wedge\langle y\rangle^\wedge t$ to an index for $K \cup \{y\}$, where t is a text for K. This means that for all $y \notin K \cup \{x\}$, there is an n such that $\varphi(\sigma^\wedge\langle y\rangle^\wedge\bar{t}_n)\downarrow$ and $\varphi(\sigma) \not\simeq \varphi(\sigma^\wedge\langle y\rangle^\wedge\bar{t}_n)$. So we have

$$\begin{aligned} y \notin K \cup \{x\} \Leftrightarrow \exists\tau(&\mathrm{rng}(\tau) \subseteq K \wedge \varphi(\sigma^\wedge\langle y\rangle^\wedge\tau)\downarrow \wedge \\ &\varphi(\sigma) \not\simeq \varphi(\sigma^\wedge\langle y\rangle^\wedge\tau)), \end{aligned}$$

which implies that $\overline{K \cup \{x\}}$ is r.e. But

$$y \notin K \Leftrightarrow y = x \vee y \notin K \cup \{x\},$$

and the left-hand side is not r.e., so we indeed have a contradiction. $\dashv$

Theorem 33 *There is an indexed family of r.e. languages $\mathcal{L} = \mathcal{L}_e$ which is identifiable by a conservative recursive function with respect to e but not identifiable by a strictly conservative recursive function.*

PROOF. Let A and B be nonempty r.e. sets such that $A \cap B = \emptyset$. Let $b_0, b_1, b_2, \ldots$ be a recursive enumeration of B, and define $L_0 = A$, $L_{i+1} = A \cup \{b_i\}$. Then $L_0, L_1, L_2, \ldots$ is a uniformly r.e. sequence, so there is an e such that $L_i = L_{e,i}$.

There is a conservative recursive function that identifies $\mathcal{L}_e = \{\, L_i \mid i \in \mathbf{N} \,\}$ with respect to e. Define

$$\psi(\sigma) \simeq \begin{cases} i+1 & \text{if } i = \mu j \le \mathrm{lh}(\sigma)\, b_j \in \mathrm{rng}(\sigma), \\ 0 & \text{if } \neg\exists j \le \mathrm{lh}(\sigma)\, b_j \in \mathrm{rng}(\sigma). \end{cases}$$

Then $\lambda\sigma\, \mathbf{G}(e, \psi(\sigma))$ is conservative and identifies $\mathcal{L}_e$ with respect to e. Alternatively, if we let $D_0 = \emptyset$, $D_{i+1} = \{b_i\}$, $E_0 = \{\, \{b_i\} \mid i \in \mathbf{N} \,\}$ and $E_{i+1} = \emptyset$, then D_i and an r.e. index for E_i are computable from i and satisfy the conditions in Theorem 28.

Now suppose that φ is a strictly conservative recursive function that identifies $\mathcal{L}$. Let σ be a locking sequence for φ and A. Then we have

$$\begin{aligned} &x \in A \rightarrow \varphi(\sigma) \simeq \varphi(\sigma^\wedge\langle x\rangle), \\ &x \in B \rightarrow \varphi(\sigma) \not\simeq \varphi(\sigma^\wedge\langle x\rangle). \end{aligned}$$

The second implication follows from the fact that φ is strictly conservative: if $x \in B$ and $\varphi(\sigma) \simeq \varphi(\sigma^\wedge\langle x\rangle)$, then φ fails to identify $A \cup \{x\}$. Since $\varphi(\sigma) \simeq \varphi(\sigma^\wedge\langle x\rangle)$ is decidable, this means that A and B are recursively separable. Thus, if we take recursively inseparable sets as A and B, no strictly conservative recursive function can identify $\mathcal{L}_e$.[6] $\dashv$

7 Conclusions

We hope to have shown that many research problems that have arisen out of Angluin's theorem can be extended to the realm of indexed families of r.e. languages, with added complexity and often finer distinctions.

We think that the study of identifiability of indexed families of r.e. languages, as opposed to arbitrary classes of r.e. languages, is an interesting subject worthy of further research. The notion of identification of an indexed family of r.e. languages with respect to an index of it is a natural one. In any concrete language learning problem, a class of languages is specified by an r.e. class of grammars of some kind, and the learner is supposed to output grammars in the target class. Since the universal membership problem for the type of grammar in question is at least r.e., any such problem is a problem of identifying an indexed family of r.e. languages with respect to an index of it in our sense. The notion is also plausible when we view inductive inference from positive data

[6] If A and B are recursively separable, then there is a strictly conservative recursive function that identifies $\mathcal{L}_e$ with respect to e, which is an easy application of Theorem 29.

as a model of child language acquisition. According to an influential view in linguistics, there is such a thing as 'universal grammar', which characterizes the set of all possible grammars of human languages. There must also be a universal 'interpreter' which conditions how mentally represented grammar is put to use. It is quite natural to think of the former as defining an r.e. set of grammars, and the latter as specifying, among other things, at least a semi-decision procedure for grammaticality.

A final remark is that the study of identification of indexed families of r.e. languages is not a halfway compromise between the special case of indexed families of recursive languages and the general case of arbitrary classes of r.e. languages. Firstly, by Fulk's (1990) result, any identifiable class of r.e. languages can be extended to an indexed family of r.e. languages identifiable with respect to an index of it. So, Theorem 14 is in a sense completely general. Secondly, the present perspective brings in, more often than not, new problems and subtlety. For instance, examples separating different restrictions on the learner have to be constructed more carefully when they have to be found among the r.e. indexable classes than when they can be arbitrary classes of r.e. languages.

References

Angluin, Dana. 1980a. Finding patterns common to a set of strings. *Journal of Computer System Sciences* **21**, 46–62.

Angluin, Dana. 1980b. Inductive inference of formal languages from positive data. *Information and Control* **45**, 117–135.

Blum, Leonore and Manuel Blum. 1975. Toward a mathematical theory of inductive inference. *Information and Control* **28**, 125–155.

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

Gold, E. M. 1965. Limiting recursion. *Journal of Symbolic Logic* **30**, 28–48.

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

Jain, Sanjay and Arun Sharma. 1994. On monotonic strategies for learning r.e. languages. In Setsuo Arikawa and Klaus P. Jantke, eds., *Algorithmic Learning Theory*. Proceedings, 1994. Lecture Notes in Artificial Intelligence 872. Berlin: Springer.

Kanazawa, Makoto. 1994. *Learnable Classes of Categorial Grammars*. Ph.D. dissertation, Stanford University. Available as ILLC Dissertation Series 1994–8, Institute for Logic, Language and Computation, University of Amsterdam.

Kapur, Shyam. 1991. *Computational Learning of Languages*. Ph.D. dissertation, Cornell University. Available as Technical Report 91–1234, Department of Computer Science, Cornell University.

Kinber, Efim, and Frank Stephan. 1995. Language learning from texts: Mindchanges, limited memory, and monotonicity. *Information and Computation* **123**, 224–241.

Lange, Steffen, and Thomas Zeugmann. 1993. The learnability of recursive languages in dependence on the space of hypotheses. GOSLER-Report 20/93. Fachbereich Mathematik und Informatik, TH Leipzig.

Moriyama, Takashi and Masako Sato. 1993. Properties of language classes with finite elasticity. In K. P. Jantke, S. Kobayashi, E. Tomita, and T. Yokomori, eds., *Algorithmic Learning Theory*. Proceedings, 1993. Lecture Notes in Artificial Intelligence 744. Berlin: Springer.

Motoki, T., T. Shinohara, and K. Wright. 1991. The correct definition of finite elasticity: Corrigendum to identification of unions. In *The Fourth Annual Workshop on Computational Learning Theory*, p. 375. San Mateo, Calif.: Morgan Kaufmann.

Osherson, D. N., M. Stob, and S. Weinstein. 1986. *Systems That Learn*. Cambridge, Mass.: MIT Press.

Rogers, Hartley, Jr. 1967. *Theory of Recursive Functions and Effective Computability*. Cambridge, Mass.: MIT Press (1987).

Shinohara, T. 1990. Inductive inference from positive data is powerful. In *The 1990 Workshop on Computational Learning Theory*, pp. 97–110. San Mateo, Calif.: Morgan Kaufmann.

Wright, K. 1989. Identification of unions of languages drawn from an identifiable class. In *The 1989 Workshop on Computational Learning Theory*, pp. 328–333. San Mateo, Calif.: Morgan Kaufmann.

Zeugmann, Thomas, and Steffen Lange. 1995. A guided tour across the boundaries of learning recursive languages. In Klaus P. Jantke and Steffen Lange, eds., *Algorithmic Learning for Knowledge-Based Systems*. Lecture Notes in Artificial Intelligence 961. Berlin: Springer.

Zeugmann, Thomas, Steffen Lange, and Shyam Kapur. 1992. Characterizations of class preserving monotonic and dual monotonic language learning. Technical Report IRCS–92–24. Institute for Research in Cognitive Science, University of Pennsylvania.

Zeugmann, Thomas, and Steffen Lange, and Shyam Kapur. 1995. Characterizations of monotonic and dual monotonic language learning. *Information and Computation* **120**, 155–173.