

# Partial Learning of Recursively Enumerable Languages<sup>\*</sup>

Ziyuan Gao<sup>1</sup>, Frank Stephan<sup>2</sup> and Sandra Zilles<sup>3</sup>

<sup>1</sup> Department of Computer Science  
University of Regina, Regina, SK, Canada S4S 0A2  
Email: [gao257@cs.uregina.ca](mailto:gao257@cs.uregina.ca)

<sup>2</sup> Department of Mathematics and Department of Computer Science  
National University of Singapore, Singapore 119076  
Email: [fstephan@comp.nus.edu.sg](mailto:fstephan@comp.nus.edu.sg)

<sup>3</sup> Department of Computer Science  
University of Regina, Regina, SK, Canada S4S 0A2  
Email: [zilles@cs.uregina.ca](mailto:zilles@cs.uregina.ca)

**Abstract.** This paper studies several typical learning criteria in the model of partial learning of r.e. sets in the recursion-theoretic framework of inductive inference. Its main contribution is a complete picture of how the criteria of confidence, consistency and conservativeness in partial learning of r.e. sets separate, also in relation to basic criteria of learning in the limit. Thus this paper constitutes a substantial extension to prior work on partial learning. Further highlights of this work are very fruitful characterisations of some of the inference criteria studied, leading to interesting consequences about the structural properties of the collection of classes learnable under these criteria. In particular a class is consistently partially learnable iff it is a subclass of a uniformly recursive family.

## 1 Introduction

Identification in the limit from positive examples, as introduced by Gold [10], models learning as a process in which a learner is presented an infinite sequence of data items belonging to a target, say an r.e. language  $L$  or a graph of a recursive function. The learner processes the data one by one, making a conjecture about the target  $L$  in every step. Successful learning of the target  $L$  requires the learner, on any infinite input sequence containing all and only the data items contained in  $L$  (called a text for  $L$ ), to return a sequence of hypotheses that stabilises on a single correct hypothesis describing  $L$ . In most variations of the model, Gödel numbers are used as hypotheses.

This model is rather restrictive: for example, there is no learner that identifies every regular language in the limit; more generally, no class of languages containing an infinite set  $S$  and all its finite subsets is identifiable in the limit [10].

---

<sup>\*</sup> F. Stephan was partially supported by NUS grant R252-000-420-112; S. Zilles was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Intuitively, for every potential learner  $M$  there is a valid data sequence for  $S$  that forces  $M$  to conjecture finite subsets of  $S$  infinitely often, thus changing its mind infinitely often and failing to stabilise on one conjecture.

To overcome this difficulty, Osherson, Stob and Weinstein introduced the model of partial learning [14], in which the sequence of hypotheses is no longer required to converge in the limit (syntactically or semantically). Instead, a learner  $M$  is considered successful for a target  $L$ , if, on any text for  $L$ ,  $M$  returns a sequence of conjectures that contains only one hypothesis infinitely often; this hypothesis must describe  $L$ . Osherson, Stob and Weinstein proved that this relaxation of Gold’s model allows for the identification of the class of all r.e. languages.

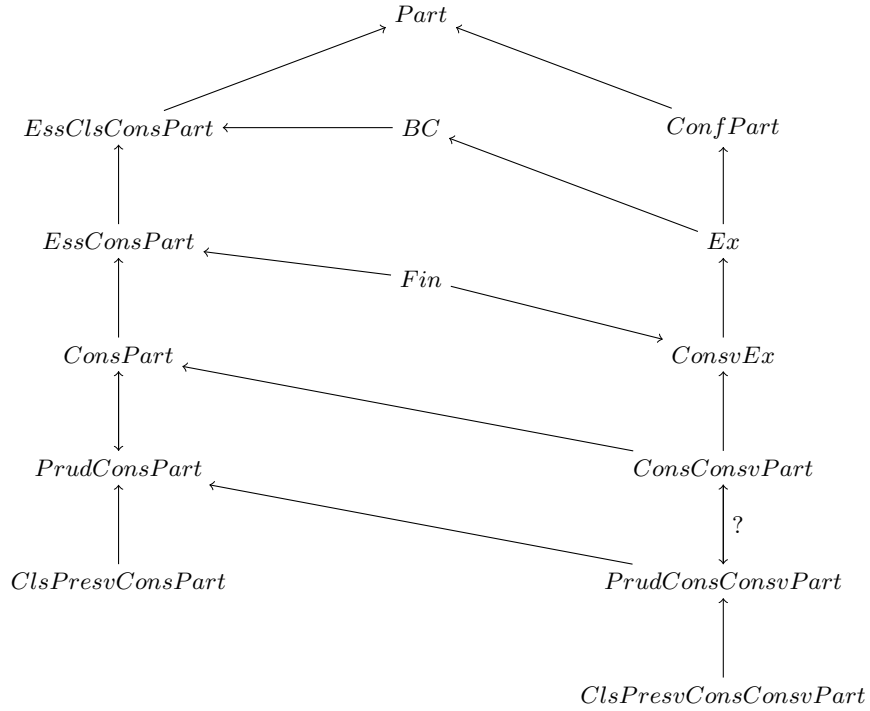
Recently, the model of partial learning has been studied in depth, namely in combination with typical learning criteria that restrict the behaviour of learners. An intuitive such criterion is consistency as introduced by Bārzdīņš [3], which requires the learner to always return conjectures for sets that contain all the examples presented in the text. The consistency requirement imposes a strong restriction on learners both in the context of learning in the limit [3] and in the context of partial learning [12]. Gao and Stephan [9] therefore introduced the notion of essential consistency, which allows the learner to be inconsistent finitely many times. For learning recursive functions, they proved this model to be less restrictive than consistent partial learning, and even more powerful than behaviourally correct learning, a version of learning in the limit in which only a semantic, not a syntactic convergence of the hypothesis sequence is required.

A criterion often considered together with or in contrast to consistency is that of conservativeness: a conservative learner identifying a class of languages in the limit is allowed to change a hypothesis on a valid text segment for a target language  $L$  only if that hypothesis is inconsistent with the text segment [1]. This definition does not transfer to the case of partial learning, where a correct hypothesis can be “suspended” infinitely often. Gao, Jain and Stephan [7] adapted the model of conservativeness to partial learning, by requiring that a conservative partial learner

- (i) outputs only one correct hypothesis (namely the one that is output infinitely often) and
- (ii) does not overgeneralise the target (when outputting incorrect hypotheses).

Another criterion that was previously adapted from the limit-learning case to the partial learning case is that of confidence. In the classical setting, a confident learner will produce a sequence of hypotheses that stabilises, even if the input text is not for a set in the target class. In the context of partial learning, the natural adaptation studied by Gao, Stephan, Wu and Yamamoto [8] was to require that the learner output only one hypothesis infinitely often, even on texts for languages outside the target class.

The main contribution of this paper is a complete study of how the criteria of confidence, consistency and conservativeness in partial learning of r.e. sets separate, also in relation to basic criteria of learning in the limit. In particular, it is determined for any pair of criteria of interest whether or not there are classes



**Fig. 1.** Learning hierarchy

learnable under one but not under the other criterion. This goes far beyond the results from previous work on partial learning, which focussed either on learning recursive functions or on only one of the special criteria addressed here.

Many of our results are similar to those in the case of learning recursive functions, but there are also some differences.

Interestingly, most of the separations proved in this paper are already witnessed by uniformly recursive families of sets, which means that the overall hierarchy of inference criteria obtained would be very similar when restricting the study to such families.

Further highlights of the present paper are characterisations of the collection of all classes of r.e. sets that are confidently partially learnable and of the collection of all classes of r.e. sets that are consistently partially learnable. The former has the interesting consequence that the union of two confidently partially learnable classes is again confidently partially learnable. The latter demonstrates that a class of r.e. sets is consistently partially learnable if and only if it is contained in a uniformly recursive family; furthermore the conservative partial learner can always be made prudent [14] that is, can be constrained to outputting only hypotheses describing sets it can identify, on any input.

The hierarchy diagram in Figure 1 summarises most of the results of this paper. The inference criteria are abbreviated; *BC* is for behaviourally correct learning [5], *Ex* for learning in the limit, *Part* for partial learning, *Conf* for confidence, *Consv* for conservativeness, *ClsCons* for class-consistency (where the learner is required to be consistent only on valid texts for potential targets), *Cons* for global consistency (where the learner is required to be consistent on all input texts), *EssClsCons* and *EssCons* for the “essential” versions of the two consistency models, *Prud* for prudence, and *ClsPresv* for the class-preserving versions of a model, requiring the learner to return only hypotheses that represent sets from the target class, on any input. A directed arc from criterion *A* to criterion *B* means that the collection of classes learnable under model *A* is contained in that learnable under model *B*. If there is no path from *A* to *B*, then the collection of classes learnable under model *A* is not contained in that learnable under model *B*. Due to space limitations, some proofs are missing in this version of the paper.

## 2 Preliminaries

**Notation 1.** The notation and terminology from recursion theory adopted in this paper follows in general the book of Rogers [16]. Background on inductive inference can be found in [11]. The symbol  $\mathbb{N}$  denotes the set of natural numbers,  $\{0, 1, 2, \dots\}$ . Let  $\varphi_0, \varphi_1, \varphi_2, \dots$  denote a fixed acceptable numbering [16] of all partial-recursive functions over  $\mathbb{N}$ . Given a set  $S$ ,  $S^*$  denotes the set of all finite sequences in  $S$ . One defines the  $e$ -th r.e. set  $W_e$  as  $\text{dom}(\varphi_e)$ . This paper fixes a one-one padding function  $\text{pad}$  with  $W_{\text{pad}(e,d)} = W_e$  for all  $e, d$ . Furthermore,  $\langle x, y \rangle$  denotes Cantor’s pairing function, given by  $\langle x, y \rangle = \frac{1}{2}(x+y)(x+y+1)+y$ . A triple  $\langle x, y, z \rangle$  denotes  $\langle \langle x, y \rangle, z \rangle$ . The notation  $\eta(x) \downarrow$  means that  $\eta(x)$  is defined, and  $\eta(x) \uparrow$  means that  $\eta(x)$  is undefined.

Turing reducibility is denoted by  $\leq_T$ ;  $A \leq_T B$  holds if  $A$  can be computed via a machine which uses  $B$  as an oracle; that is, it can give information on whether or not  $x$  belongs to  $B$ .  $A \equiv_T B$  means that  $A \leq_T B$  and  $B \leq_T A$  both hold, and  $\{A : A \equiv_T B\}$  is called the Turing degree of  $B$ . For any partial-recursive function  $g$ ,  $\text{graph}(g) = \{\langle x, y \rangle : g(x) \downarrow = y\}$ . The symbol  $\mathbb{K}$  denotes the diagonal halting problem  $\{e : \varphi_e(e) \downarrow\}$ .

For any two sets  $A$  and  $B$ ,  $A \oplus B = \{2x : x \in A\} \cup \{2y + 1 : y \in B\}$ . Analogously,  $A \oplus B \oplus C = \{3x : x \in A\} \cup \{3y + 1 : y \in B\} \cup \{3z + 2 : z \in C\}$ .

For any  $\sigma, \tau \in (\mathbb{N} \cup \{\#\})^*$ ,  $\sigma \preceq \tau$  if and only if  $\sigma = \tau$  or  $\tau$  is an extension of  $\sigma$ ,  $\sigma \prec \tau$  if and only if  $\sigma$  is a proper prefix of  $\tau$ , and  $\sigma(n)$  denotes the element in the  $n$ th position of  $\sigma$ , starting from  $n = 0$ . The concatenation of two strings  $\sigma$  and  $\tau$  shall be denoted by  $\sigma \circ \tau$ ; for convenience, and whenever there is no possibility of confusion, this is occasionally denoted by  $\sigma\tau$ . Let  $\sigma[n]$  denote the sequence  $\sigma(0) \circ \sigma(1) \circ \dots \circ \sigma(n-1)$ . The length of  $\sigma$  is denoted by  $|\sigma|$ .

The learnability notions investigated in the present paper are built on the main learning paradigms from positive data – explanatory learning, behaviourally correct learning, and partial learning. Explanatory learning, or “learning in the

limit”, was introduced by Gold [10] to model the process of language acquisition. This model was later generalised by Bärzdiņš [3] and Case [5]; in their proposed model, known as behaviourally correct learning, the learner is required to almost always output a correct hypothesis of the input language, although it is permitted to output syntactically different hypotheses. Osherson, Stob and Weinstein [14] then extended the criterion of behaviourally correct learnability to partial learnability, according to which the learner must output exactly one correct index of the language infinitely often and output any other conjecture only finitely often. In addition, one can specify various constraints on the learner; the following definition imposes a restriction on the hypothesis space of the learner.

**Definition 2.**  $M$  is said to *class-preservingly* (*ClsPresv*) learn  $\mathcal{C}$  if it learns  $\mathcal{C}$  from text with respect to a hypothesis space  $\{H_0, H_1, H_2, \dots\}$  such that  $\mathcal{C} = \{H_0, H_1, H_2, \dots\}$ .

Let  $\mathcal{C}$  be a class of r.e. languages. Throughout this paper, the mode of data presentation is that of a *text*, by which is meant an infinite sequence of natural numbers and the  $\#$  symbol. Formally, a *text*  $T_L$  for some  $L$  in  $\mathcal{C}$  is a map  $T_L : \mathbb{N} \rightarrow \mathbb{N} \cup \{\#\}$  such that  $L = \text{content}(T_L)$ ; here  $T_L[n]$  denotes the sequence  $T_L(0) \circ T_L(1) \circ \dots \circ T_L(n-1)$  and the content of a text  $T$ , denoted  $\text{content}(T)$ , is the set of numbers in the range of  $T$ . Analogously, for a finite sequence  $\sigma$ ,  $\text{content}(\sigma)$  is the set of numbers in the range of  $\sigma$ . In the following definitions,  $M$  is a recursive function mapping  $(\mathbb{N} \cup \{\#\})^*$  into  $\mathbb{N} \cup \{?\}$ ; the  $?$  symbol permits  $M$  to abstain from conjecturing at any stage.

- Definition 3.** (I) [14]  $M$  *partially* (*Part*) learns  $\mathcal{C}$  if, for every  $L$  in  $\mathcal{C}$  and each text  $T_L$  for  $L$ , there is exactly one index  $e$  such that  $M(T_L[k]) = e$  for infinitely many  $k$ ; furthermore, if  $M$  outputs  $e$  infinitely often on  $T_L$ , then  $L = W_e$ .
- (II) [10]  $M$  *explanatorily* (*Ex*) learns  $\mathcal{C}$  if, for every  $L$  in  $\mathcal{C}$  and each text  $T_L$  for  $L$ , there is a number  $n$  for which  $L = W_{M(T_L[n])}$  and, for any  $j \geq n$ ,  $M(T_L[j]) = M(T_L[n])$ .
- (III) [5]  $M$  *behaviourally correctly* (*BC*) learns  $\mathcal{C}$  if, for every  $L$  in  $\mathcal{C}$  and each text  $T_L$  for  $L$ , there is a number  $n$  for which  $L = W_{M(T_L[j])}$  whenever  $j \geq n$ .
- (IV) [14]  $M$  is *prudent* if it learns the class  $\{W_{M(\sigma)} : \sigma \in (\mathbb{N} \cup \{\#\})^*, M(\sigma) \neq ?\}$ . In other words,  $M$  learns every set it conjectures.

As a point of departure, the following theorem establishes that the learning criterion of partial learning is quite powerful.

**Theorem 4 (Osherson, Stob and Weinstein [14]).** *The class of all r.e. sets is partially learnable.*

### 3 Confident Partial Learning

Gao, Stephan, Wu and Yamamoto [8] introduced the notion of confident partial learning, by naturally generalising the constraint that the learner must, with

respect to the convergence criterion considered, single out a hypothesis on every possible text for every possible, even non-r.e., language.

**Definition 5 (Gao, Stephan, Wu and Yamamoto [8]).**  $M$  is said to *confidently partially (ConfPart) learn*  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text and outputs on every infinite sequence exactly one index infinitely often.

Confidence is a proper restriction on a partial learner in the sense that the class of all r.e. sets is no longer partially learnable if the learner is required to be confident. This is witnessed even by a class that can be learned behaviourally correctly, the corresponding result for recursive functions can be carried over by considering the graphs.

**Theorem 6 (Gao and Stephan [9]).** *There is a class of recursive sets that is behaviourally correctly learnable, but not confidently partially learnable.*

Furthermore, one can show that Gold's class containing one infinite set and all its finite subsets [10] is confidently partially learnable but not behaviourally correctly learnable.

**Theorem 7 (Gao, Stephan, Wu and Yamamoto [8]).** *There is a uniformly recursive family of sets that is confidently partially learnable, but not behaviourally correctly learnable.*

By contrast, every class that is explanatorily learnable is also confidently partially learnable. This holds true, even when the *Ex*-learner is allowed to converge to an index of a set that disagrees with the target set on at most one number, which is Case and Smith's criterion  $Ex^1$  of learning with at most one anomaly [6].

**Theorem 8.** *If a class of r.e. sets is explanatorily learnable with at most one anomaly, then it is also confidently partially learnable.*

The following characterisation can be brought over from function learning to language learning.

**Theorem 9 (Gao and Stephan [9]).** *A class  $\mathcal{C}$  of r.e. sets is confidently partially learnable if and only if there is a recursive learner  $M$  such that*

- $M$  outputs on each text exactly one index infinitely often;
- if  $T$  is a text for a language  $L$  in  $\mathcal{C}$  and  $d$  is the index output infinitely often by  $M$  on  $T$ , then there is some  $e \leq d$  with  $W_e = L$ .

**Corollary 10.** *If  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are two classes of r.e. sets, both of which are confidently partially learnable, then their union  $\mathcal{C}_1 \cup \mathcal{C}_2$  is also confidently partially learnable.*

## 4 Essentially Consistent Partial Learning

Consistency [3, 4, 17] is a quite natural condition which postulates that every conjecture should at least enumerate all the data seen so far. Consistency is known to be restrictive, for both explanatory and partial learning. In the present section, consistency is weakened to essentially global consistent and essentially class consistent learning which generalise the learnability without making the criterion so strong that it permits to learn the class RE of all r.e. languages. Note that essentially class consistent learning is a restriction only for partial learning, while it would be automatically implied by the criteria of explanatory and behaviourally correct learning.

**Definition 11.** Let  $\mathcal{C}$  be a class of r.e. languages and  $M$  be a recursive learner.

- (I)  $M$  is said to *essentially globally consistently partially* (*EssConsPart*) learn  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text, and for each language  $L$  and every text  $T$  for  $L$ ,  $\text{content}(T[n]) \subseteq W_{M(T[n])}$  holds for cofinitely many  $n$ .
- (II)  $M$  is said to *essentially class consistently partially* (*EssClsConsPart*) learn  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text, and for each language  $L$  in  $\mathcal{C}$  and every text  $T$  for  $L$ ,  $\text{content}(T[n]) \subseteq W_{M(T[n])}$  holds for cofinitely many  $n$ .
- (III)  $M$  is *consistent* (*Cons*) if for all  $\sigma \in (\mathbb{N} \cup \{\#\})^*$ ,  $\text{content}(\sigma) \subseteq W_{M(\sigma)}$ .
- (IV) For any text  $T$ ,  $M$  is *consistent on*  $T$  if  $\forall n > 0 [\text{content}(T[n]) \subseteq W_{M(T[n])}]$ .
- (V)  $M$  is said to be *class consistent* (*ClsCons*) if it is consistent on each text for every  $L$  in  $\mathcal{C}$ .
- (VI)  $M$  is said to *consistently partially* (*ConsPart*) learn  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text and is consistent.
- (VII)  $M$  is said to *class consistently partially* (*ClsConsPart*) learn  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text and is class consistent.

One can generalise these notions correspondingly for learners recursive relative to an oracle.

**Example 12.** The class of r.e. languages  $\mathcal{C} = \{\mathbb{K} \cup D : D \text{ is finite}\} \cup \{\mathbb{N}\}$  is essentially class consistently partially learnable relative to an oracle  $A$  only if  $\mathbb{K} \leq_T A$ .

**Proof.** Let  $A$  be any oracle such that there is an  $A$ -recursive essentially class consistent partial learner  $M$  of the class  $\mathcal{C}$ . Then, as  $\mathbb{N} \in \mathcal{C}$  and  $M$  is essentially class consistent, there is a  $\sigma \in \mathbb{N}^*$  such that for all  $\tau \in \mathbb{N}^*$ ,  $\text{range}(\sigma \circ \tau) \subseteq W_{M(\sigma \circ \tau)}$ . Fixing any such  $\sigma$ , one can build a text  $T$  for  $\mathbb{K} \cup \text{range}(\sigma)$  as follows.

1. For all  $x < |\sigma|$ ,  $T(x) = \sigma(x)$ .
2. At stage  $s$ , let  $a_s$  be the last position on which  $T$  has been defined up to the present stage. Let  $b_s = \min((\mathbb{K} \cup \text{range}(\sigma)) - \text{range}(T[a_s + 1]))$  and  $F_s = \{M(T[k]) : k \leq a_s + 1 \wedge W_{M(T[k])} = \mathbb{K} \cup \text{range}(\sigma)\}$ . Search noneffectively for an  $x_s$  such that  $x_s \in \mathbb{K} \cup \text{range}(\sigma)$  and the condition  $\{M(T[a_s + 1] \circ x_s), M(T[a_s + 1] \circ x_s \circ b_s)\} \cap F_s = \emptyset$  holds. If such an  $x_s$  is found, set  $T(a_s + 1) = x_s$  and  $T(a_s + 2) = b_s$ .

There must be a stage  $s$  at which the search for an  $x_s$  fails to terminate successfully. For, by the construction of  $T$ , if the stages proceed through infinitely often, then  $M$  on every text segment of  $T$  outputs an index different from all of its prior correct conjectures, contradicting the fact that it partially learns  $\mathcal{C}$ . Thus there is a stage  $s$  such that whenever  $x \in \mathbb{K} \cup \text{range}(\sigma)$ , then  $M(T[a_s + 1] \circ x) \in F_s \vee M(T[a_s + 1] \circ x \circ b_s) \in F_s$  holds. The global consistency of  $M$  on any text extension of  $\sigma$  gives the condition that  $x \notin \mathbb{K} \cup \text{range}(\sigma) \Rightarrow M(T[a_s + 1] \circ x) \notin F_s \wedge M(T[a_s + 1] \circ x \circ b_s) \notin F_s$ . Noting that  $a_{s+1}$  and  $b_s$  are fixed numbers,  $F_s$  is a fixed finite set and  $\sigma$  is a fixed string, one therefore has the reduction  $x \in \mathbb{K} \cup \text{range}(\sigma) \Leftrightarrow \{M(T[a_s + 1] \circ x), M(T[a_s + 1] \circ x \circ b_s)\} \cap F_s \neq \emptyset$ , which shows that  $\mathbb{K} \leq_T A$ , as required. ■

**Corollary 13.** *The class RE of all r.e. sets is essentially class consistently partially learnable relative to an oracle A iff  $\mathbb{K} \leq_T A$ .*

**Example 14.** The class  $\mathcal{C}$  in Example 12 is not essentially class consistently partially learnable with finitely many queries to any oracle.

**Theorem 15.** *Every behaviourally correctly learnable class of r.e. languages is essentially class consistently partially learnable.*

**Proof.** Let  $\mathcal{C}$  be a class of r.e. languages that is behaviourally correctly learnable via a recursive learner  $M$ . On text  $T = a_0 \circ a_1 \circ a_2 \circ \dots$ , let  $e_0, e_1, e_2, \dots$  be a one-one enumeration of all the distinct conjectures of  $M$ . Define a new learner  $N$  as follows: on text  $a_0 \circ a_1 \circ a_2 \circ \dots$ ,  $N$  outputs for each  $i$  the conjecture  $e_i$  at least  $n$  times iff there is a stage  $s > n$  such that  $\forall x < n [x \in \{a_0, a_1, \dots, a_s\} \Leftrightarrow x \in W_{e_i, s}]$  holds. Since  $M$  is a BC-learner of  $\mathcal{C}$ , it outputs on a text for any  $L \in \mathcal{C}$  only finitely many incorrect conjectures; so there is a stage  $s$  after which  $N$  only outputs indices of  $L$ . Furthermore,  $N$  infinitely often conjectures every correct index output by  $M$ . Let  $d_0, d_1, d_2, \dots$  be the sequence of conjectures of  $N$  on some text  $T$ . One can define a learner  $N'$  which outputs on  $T$  the index  $\text{pad}(d_i, m_i)$  for each conjecture  $d_i$  of  $N$  on  $T$ , where  $m_i = |\{k < i : d_k < d_i\}|$ . By construction, if  $d_m$  is the minimum correct index among all of  $N$ 's conjectures, there is a unique number  $k$  such that  $N'$  outputs  $\text{pad}(d_m, k)$  infinitely often, while every other index is output only finitely often; thus  $N'$  essentially class consistently partially learns  $\mathcal{C}$ . ■

**Corollary 16.** *Essentially class consistent partial learning is not closed under finite unions; that is, there are classes of r.e. languages  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , each of which is essentially class consistently partially learnable, such that  $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$  is not essentially class consistently partially learnable.*

In prior work [9], it was shown that essentially globally consistent partial learning of recursive functions is closed under finite unions. Theorem 17 establishes the analogue of this result for the case of learning r.e. languages.

**Theorem 17.** *Essentially globally consistent partial learning is closed under finite unions; that is, if  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are both essentially globally consistently partially learnable, then  $\mathcal{L}_1 \cup \mathcal{L}_2$  is essentially globally consistently partially learnable.*



**Proof.** Assume that  $M_1$  and  $M_2$  are two *EssConsPart*-learners. Now make  $M_3$  from  $M_1$  as follows: If  $M_1$  on input  $\sigma$  conjectures  $e$  then count the number of times which  $M_3$  has conjectured  $e$  on prefixes  $\tau \prec \sigma$ ; let  $m$  be this number. If now for all  $x < m$  it holds that  $\text{range}(\sigma)(x) = W_{e,|\sigma|}(x)$  then let  $M_3(\sigma) = e$  else let  $M_3(\sigma)$  be an index  $d$  of  $\text{range}(\sigma)$  which in addition satisfies  $d \geq |\sigma|$ , such an index can be found by the padding lemma.

It is easy to see that  $M_3$  is recursive. Furthermore, if  $M_3$  outputs an index  $e$  infinitely often then  $W_e$  is equal to the language to be learnt. On the other hand, if  $M_1$  outputs an index  $e$  infinitely often and  $W_e$  is equal to the language to be learnt then one can show by induction that  $e$  is output infinitely often; if  $M_3$  outputs  $e$  at least  $m$  times and  $\sigma$  is a sufficiently long prefix of the text with  $M_1(\sigma) = e$  then  $\text{range}(\sigma)(x) = W_{e,|\sigma|}(x)$  for all  $x < m$  and therefore  $M_3$  will also output  $e$  for the  $m + 1$ -st time. Furthermore, whenever  $M_3(\sigma) \neq M_1(\sigma)$  then  $M_3(\sigma)$  is consistent. Hence it follows that  $M_3$  is an *EssConsPart*-learner for the class of sets learnt by  $M_1$ . One can make a similar learner  $M_4$  out of  $M_2$ .

Now  $M_5(\sigma) = \min\{M_3(\sigma), M_4(\sigma)\}$  is a further learner; as  $M_3$  and  $M_4$  are consistent for almost all prefixes of a given text, so is  $M_5$ . Furthermore, the least index  $e$  output on a given text infinitely often by either  $M_3$  or  $M_4$  is also output infinitely often by  $M_5$ . Hence  $M_5$  outputs on every language learnt by either learner at least one index infinitely often and every infinitely often output index is correct. Following the usual padding construction [9], one can modify  $M_5$  to a further learner  $M_6$  which is also essentially partially consistent and whenever  $M_5$  outputs at least one index infinitely often, then  $M_6$  outputs a padded version of the least such index infinitely often. Hence  $M_6$  is an *EssConsPart*-learner which learns every language learnt either by  $M_1$  or by  $M_2$ . ■

As shown in [7, Theorem 24], every consistently partially learnable class of r.e. languages is contained in a uniformly recursive family of languages. The following theorem establishes a strong converse of this result, showing that every subclass of a uniformly recursive family may even be prudently consistently partially learnt. This provides a complete characterisation of all consistently partially learnable classes of languages.

**Theorem 18.** *The following statements are equivalent for a class  $\mathcal{C}$  of r.e. sets.*

- (I)  $\mathcal{C}$  is a subclass of a uniformly recursive family;
- (II)  $\mathcal{C}$  is *ConsPart*-learnable;
- (III)  $\mathcal{C}$  is *PrudConsPart*-learnable;
- (IV)  $\mathcal{C}$  is *PrudConsPart*-learnable using a uniformly recursive hypothesis space.

**Proof.** First the implication from the first statement to the last is shown. Let the class be contained in the class-comprising hypothesis space  $L_0, L_1, \dots$  which is also uniformly recursive and in addition one-one. Furthermore, assume that the hypothesis space contains all cofinite sets (in order to always have sufficiently many hypotheses to choose from).

Given any text  $T$ , on input  $T(0)T(1)\dots T(n)$ , the learner determines the least pair  $\langle i, j \rangle$  such that

$$L_i \cap \{0, 1, \dots, j\} \subseteq \{T(0), T(1), \dots, T(n)\} \subseteq L_i \cup \{\#\}.$$

and  $j \leq n$  and the learner has conjectured  $L_i$  exactly  $j$  times before on inputs  $T(0)T(1)\dots T(m)$  with  $m < n$ . Having  $\langle i, j \rangle$ , the learner conjectures  $L_i$ .

Note that no wrong set is conjectured infinitely often: if  $j \in L_i - \text{range}(T)$  or  $T(j) \notin L_i \cup \{\#\}$  then the pair  $\langle i, j \rangle$  will never qualify and therefore  $L_i$  will be conjectured at most  $j$  times. Furthermore, if  $L_i$  is the language to be learnt then each pair  $\langle i, j \rangle$  will qualify from that point onwards where  $L_i$  has been conjectured  $j$  times and where all the members of  $L_i \cap \{0, 1, \dots, j\}$  have been observed in the input; as there are only finitely many smaller pairs which will be dealt with in only finitely many steps, the learner will eventually address the pair  $\langle i, j \rangle$  and conjecture  $L_i$  again. It is easy to see that the learner is consistent. Furthermore, for every  $n$  there is a cofinite set such that its members below  $n$  are exactly those which appear in  $\{T(0), T(1), \dots, T(n)\}$  and therefore, for every input  $T(0)T(1)\dots T(n)$  there is a pair  $\langle i, j \rangle$  which qualifies so that the learner is total.

The implication from the fourth to the third and from the third to the second statements are obvious; the implication from the second to the first statement has been established in prior work [7, Theorem 24] and this completes the proof. ■

**Example 19.** Let a class  $\mathcal{C}$  contain the set of all pairs of natural numbers plus, for each  $x$ , the following set:  $L_x = \{\langle x, y \rangle : \forall z < y [\langle x, z \rangle \in L_x]\}$  and the  $x$ -th machine  $M_x$  outputs on the sequence  $\langle x, 0 \rangle \langle x, 1 \rangle \dots \langle x, y \rangle$  either ? or an index  $e$  such that  $W_e$  contains some pair  $\langle x', y' \rangle$  with  $x' \neq x$ . Then the class  $\mathcal{C}$  is *PrudConsPart*-learnable but not *ClsPresvClsConsPart*-learnable.

One implication of Theorem 18 is that every consistently partially learnable class contains only recursive languages. This characterisation, however, does not extend to the notion of essentially consistent partial learnability, as the following example demonstrates.

**Theorem 20.** *The class  $\mathcal{C} = \{\mathbb{K}\} \cup \{D : D \text{ is finite}\}$  is *PrudEssConsPart*-learnable.*

It was shown in earlier work [8] that the class of all cofinite sets is not confidently partially learnable. As this class is uniformly recursive, it follows from Theorem 18 that it is *PrudConsPart* learnable. Thus the criterion of *PrudConsPart* learnability does not imply confident partial learnability in general.

**Corollary 21.** *The class of all cofinite sets is *ConsPart*-learnable but not confidently partially learnable.*

**Theorem 22.** *There is a confidently partially learnable class of recursive languages which is not essentially class consistently partially learnable.*

**Proof.** Let  $M_0, M_1, M_2, \dots$  be an enumeration of all partial-recursive learners. For each  $\sigma \in \mathbb{N}^*$  and  $i \in \mathbb{N}$ , let  $A_{\langle \sigma, i \rangle}$  denote the set  $\{\langle \sigma, i, k \rangle : k \in \mathbb{N}\}$  and define an r.e. language  $L_{\langle \sigma, i \rangle}$  in stages as follows. The construction proceeds by trying to build a text for  $L_{\langle \sigma, i \rangle}$  on which  $M_i$  either never outputs any index infinitely often, or is inconsistent at infinitely many stages.  $\tau_0 = \sigma$  is an initial approximation

to this text; at stage  $s + 1$ , one defines a further approximation  $\tau_{s+1}$  based on the outputs of  $M_i$  on some potential extensions of  $\tau_s$ . For bookkeeping, define approximations  $B_0, B_1, B_2, \dots$  to an auxiliary r.e. set  $B$ ;  $B$  records the numbers that must not be added into  $L_{\langle \sigma, i \rangle}$  in order to maintain the inconsistency of  $M_i$  on some earlier constructed text segment.

1. Let  $L_{\langle \sigma, i \rangle, 0} = \text{range}(\sigma), \tau_0 = \sigma$  and  $B_0 = \emptyset$ .
2. At stage  $s + 1$ , search for either (I) the first  $w \in A_{\langle \sigma, i \rangle} \cap \{w : w > \max(B_s \cup L_{\langle \sigma, i \rangle, s})\}$  such that  $M_i(\tau_s \circ w) \downarrow \notin \{M_i(\gamma) : \gamma \preceq \tau_s\}$ , or (II) the first pair  $x, y$  with  $x \neq y, \{x, y\} \subseteq A_{\langle \sigma, i \rangle} \cap \{w : w > \max(B_s \cup L_{\langle \sigma, i \rangle, s})\}$ , so that for some  $e$ ,
  - $M_i(\tau_s \circ x) \downarrow = M_i(\tau_s \circ y) \downarrow = e$ ;
  - $x \in W_e \vee y \in W_e$  holds.

In case (I), let  $L_{\langle \sigma, i \rangle, s+1} = L_{\langle \sigma, i \rangle, s} \cup \{w\}$ ,  $\tau_{s+1} = \tau_s \circ w$  and  $B_{s+1} = B_s$ . In case (II), let  $z$  be the first element in  $\{x, y\}$  that  $W_e$  enumerates and let  $z'$  be the other element of  $\{x, y\}$ . Then set  $L_{\langle \sigma, i \rangle, s+1} = L_{\langle \sigma, i \rangle, s} \cup \{z'\}$ ,  $\tau_{s+1} = \tau_s \circ z'$  and  $B_{s+1} = B_s \cup \{z\}$ .

Let  $L_{\langle \sigma, i \rangle} = \bigcup_{s \in \mathbb{N}} L_{\langle \sigma, i \rangle, s}$  and define  $\mathcal{C}_1 = \{L_{\langle \sigma, i \rangle} : \sigma \in \mathbb{N}^* \wedge i \in \mathbb{N}\}$ ,  $\mathcal{C} = \mathcal{C}_1 \cup \{\mathbb{N}\}$ . Then  $\mathcal{C}$  is confidently partially learnable. The class  $\mathcal{C}_1$  is confidently partially learnable: the subclass of all  $L_{\langle \sigma, i \rangle}$  which are infinite may be explanatorily learnt via a learner which, on a text  $T$ , converges to an index for  $L_{\langle \sigma, i \rangle}$  in the case that almost all members of  $\text{range}(T)$  are contained in  $A_{\langle \sigma, i \rangle}$  and outputs a default index infinitely often otherwise; the subclass of all  $L_{\langle \sigma, i \rangle}$  which are finite may also be explanatorily learnt by a learner which, on a given text segment  $T[n]$ , outputs a canonical index for  $\text{range}(T[n])$ . Therefore each of these two subclasses of  $\mathcal{C}_1$  is confidently partially learnable, and so the union  $\mathcal{C}_1 \cup \{\mathbb{N}\}$  is confidently partially learnable as well.

Next, assume for a contradiction that  $M_n$  essentially class consistently partially learns the class  $\mathcal{C}$ . Since  $M_n$  must also essentially class consistently partially learn  $\mathbb{N}$ , there must exist some  $\sigma \in \mathbb{N}^*$  such that for all  $\tau \in \mathbb{N}^*$ ,  $\text{range}(\sigma \circ \tau) \subseteq W_{M_n(\sigma \circ \tau)}$ . Fix such a  $\sigma$ . By the construction of  $L_{\langle \sigma, n \rangle}$ , there is a text  $T$  for  $L_{\langle \sigma, n \rangle}$  such that on every text segment of  $T$ ,  $M_n$  either outputs a conjecture different from all of its previous ones, or it outputs an index  $e$  such that for some  $y$ ,  $y \in W_e - L_{\langle \sigma, n \rangle}$ , that is, the index is incorrect. Consequently,  $M_n$  cannot be an essentially class consistent partial learner of  $\mathcal{C}$ . ■

**Corollary 23.** *There is an explanatorily learnable class of recursive languages which is not essentially globally consistently partially learnable.*

**Proof.** Let  $\mathcal{C}'_1 = \{L_{\langle \sigma, i \rangle} : \sigma \in \mathbb{N}^* \wedge i \in \mathbb{N} \wedge |L_{\langle \sigma, i \rangle}| = \infty\}$ , where  $L_{\langle \sigma, i \rangle}$  is as defined in Theorem 22. As was argued in the proof of Theorem 22,  $\mathcal{C}'_1$  is explanatorily learnable. Suppose, however, that it were essentially globally consistently partially learnt by some recursive learner  $M_n$ . Then there is some  $\sigma \in \mathbb{N}^*$  such that for all  $\tau \in \mathbb{N}^*$ ,  $\text{range}(\sigma \circ \tau) \subseteq W_{M_n(\sigma \circ \tau)}$ . Thus the language  $L_{\langle \sigma, n \rangle}$  is infinite and contained in  $\mathcal{C}'_1$ , but Theorem 22 shows that there is a text for  $L_{\langle \sigma, n \rangle}$  on which  $M_n$  almost always either outputs an incorrect hypothesis, or outputs a hypothesis different from all its prior ones. Hence  $\mathcal{C}'_1$  is not essentially globally consistently partially learnable. ■

## 5 Conservative Partial Learning

Angluin [1] introduced the notion of *conservativeness* in the model of explanatory learning and she gave sufficient conditions for an indexed family of nonempty recursive languages to be inferable by a conservative learner [1, Theorem 5]. Subsequent studies on conservative learning in the case of uniformly r.e. classes as well as indexed families [13, 18] yielded fairly succinct characterisations of this learning criterion. In prior work [7], the notion of conservativeness was adapted to the model of partial learning; in this modified version of conservative learning, the learner is required to output exactly one correct index of the input language  $L$  infinitely often and it cannot conjecture any proper superset of  $L$ .

In particular when considered together with consistency, conservativeness turned out to rule out many irregularities although not all which can arise from Pitt's delaying trick [4, 15]. When it comes to partial learning, the combination of consistency and conservativeness also reduces the learning power and brings the criterion down to conservative explanatory learning, as shown below. The present section aims to shed further light on the nature of conservative partial learning alone as well as on its combination with consistency.

**Definition 24 (Gao, Jain and Stephan [7]).** A recursive learner  $M$  is said to *conservatively partially* (*ConsvPart*) learn  $\mathcal{C}$  if it partially learns  $\mathcal{C}$  from text and outputs on each text for every  $L$  in  $\mathcal{C}$  exactly one index  $e$  with  $L \subseteq W_e$ .

The first example notes that confident partial learnability does not imply conservative partial learnability in general.

**Example 25.**  $\{D : D \text{ finite}\} \cup \{\mathbb{N}\}$  is confidently partially learnable, but not conservatively partially learnable.

One can in fact construct an explanatorily learnable class of languages that is not conservatively partially learnable, as the following theorem demonstrates. By Theorem 8, the class given in Example 26 is also confidently partially learnable.

**Theorem 26.** *There is a uniformly recursive family of sets that is explanatorily learnable, but not conservatively partially learnable.*

**Proof.** Define an indexed family which contains for every  $e$

- the set  $\{e, e + 1, e + 2, \dots\}$  and
- the first set of the form  $\{e, e + 1, \dots, e + t\}$  found such that the  $e$ -th learner conjectures on the input  $e e + 1 e + 2 \dots e + t$  a set containing  $e, e + 1, \dots, e + t, e + t + 1$ ; if no such  $t$  exists then no finite set with minimum  $e$  is in the class.

It is easy to see that the resulting family can be made uniformly recursive and that none of the learners *ConsvPart*-learn this family. Furthermore, an explanatory learner would find in the limit the least element  $e$  in the text. In the case that a set of the form  $\{e, e + 1, \dots, e + t\}$  is added in the family and the text does not contain any element larger than  $t$  then the learner converges to an

index of this set else the learner converges to the index of  $\{e, e + 1, \dots\}$  which, without loss of generality, comes first in the indexed family while the index of  $\{e, e + 1, \dots, e + t\}$  is the second index with least element  $e$  (if any). ■

For completeness, the next theorem states that conservative partial learnability does not imply confident partial learnability or behaviourally correct learnability in general. Gao, Jain and Stephan [7, Example 9] have proven that the class of graphs of all recursive functions witnesses this separation.

**Theorem 27 (Gao, Jain and Stephan [7]).** *There is a class of infinite recursive sets that is conservatively partially learnable, but neither confidently partially learnable nor behaviourally correctly learnable.*

Consistent partial learning has been studied previously mainly in the context of learning recursive functions [12] and it turned out that for the case of learning recursive functions from arbitrary texts, consistent partial learnability is equivalent to explanatory learnability. The next theorem provides an analogue of this result for the case of learning r.e. languages, showing that consistency, when enforced together with partial conservativeness, is no less stringent than explanatory learnability.

**Theorem 28.** *If a class  $\mathcal{C}$  of r.e. languages is  $ConsConsvPart$ -learnable, then  $\mathcal{C}$  is  $Ex$ -learnable by a learner which does not output any index for a proper superset of a given target language  $L \in \mathcal{C}$  on every text for  $L$ .*

**Theorem 29.** *If a class  $\mathcal{C}$  of r.e. languages is  $Ex$ -learnable by a learner which does not output any index for a proper superset of a given target language  $L \in \mathcal{C}$  on every text for  $L$ , then  $\mathcal{C}$  is  $ConsvEx$ -learnable.*

The next corollary is a consequence of Theorems 28 and 29.

**Corollary 30.** *If a class  $\mathcal{C}$  of r.e. languages is  $ConsConsvPart$ -learnable, then  $\mathcal{C}$  is  $ConsvEx$ -learnable.*

**Example 31.** The class  $\{\mathbb{K}\}$  is finitely learnable but not  $ConsConsvPart$ -learnable. The class  $\mathcal{F}$  of all finite languages is  $ConsConsvPart$ -learnable but not finitely learnable.

**Theorem 32.** *There exists a uniformly recursive class of languages which is  $PrudConsvBC$ -learnable as well as  $EssConsPrudConsvPart$ -learnable but neither  $ConsvPart$ -learnable with respect to a class-preserving hypothesis space nor explanatorily learnable.*

This section concludes with some results on partially conservative learning with respect to uniformly recursive families. In particular, these observations illustrate the connection between partial learning and learning in the limit (in both the syntactic as well as semantic sense).

**Theorem 33.** *If a uniformly recursive family  $\mathcal{C}$  is  $ConsvPart$ -learnable, then  $\mathcal{C}$  is behaviourally correctly learnable.*

**Proof.** Let  $M$  be a recursive *ConsvPart*-learner of  $\mathcal{C}$ , and  $T$  be a text for any language  $L_e$  in  $\mathcal{C}$ . As  $M$  *ConsvPart*-learns  $L_e$ , there is a number  $n$  sufficiently large so that  $W_{M(T[n])} = L_e$ . Let  $H_e = \text{range}(T[n])$ . Since  $M$  is partially conservative,  $L_d = L_e$  holds for every  $L_d$  in  $\mathcal{C}$  with  $H_e \subseteq L_d \subseteq L_e$ , for otherwise one may build a text for  $L_d$  extending  $T[n]$  on which  $M$  outputs a proper superset of  $L_d$ . Hence there is a family of finite tell-tale sets for  $\mathcal{C}$ . As shown in [2, Section 3.2, Corollary 3], a uniformly recursive class non-effectively satisfying Angluin's tell-tale condition is *BC*-learnable, and therefore  $\mathcal{C}$  is *BC*-learnable. ■

**Theorem 34.** *If a uniformly recursive family  $\mathcal{C}$  is *ConsvPart*-learnable with respect to a class-preserving hypothesis space, as well as *Ex*-learnable with respect to a class-preserving hypothesis space, then it is *ClsPresvEx*-learnable by a learner which does not output any index for a proper superset of a given target language  $L \in \mathcal{C}$  on every text for  $L$ .*

**Proof.** Let  $M$  be a recursive *ConsvPart*-learner of the given class  $\mathcal{C}$  which uses a class-preserving hypothesis space. One may assume that  $M$  uses any general class-preserving hypothesis space. As  $\mathcal{C}$  is also explanatorily learnable, there is a uniformly r.e. family of finite tell-tale sets for  $\mathcal{C}$ . Suppose  $L_0, L_1, L_2, \dots$  is a uniformly recursive numbering of  $\mathcal{C}$ , and that  $H_0, H_1, H_2, \dots$  is the corresponding family of tell-tale sets, that is, for all  $e$ ,  $H_e \subseteq L_e$  and there is no  $d$  such that  $H_e \subseteq L_d \subset L_e$ . One can define a learner  $N$  as follows. On input  $\sigma$ ,  $N$  searches for the least  $e \leq |\sigma|$ , if such an  $e$  exists, with  $H_{e,|\sigma|} \subseteq \text{content}(\sigma) \subseteq L_e$ ; if no such  $e$  is found,  $N$  outputs ?. If  $e'$  is the least such number,  $N$  then searches for the shortest  $\tau \preceq \sigma$  such that  $H_{e',|\sigma|} \subseteq W_{M(\tau),|\sigma|} \subseteq L_{e'}$ ; if no such  $\tau$  exists,  $N$  outputs ?. If  $\tau'$  is the shortest such prefix found, then  $N$  outputs  $M(\tau')$ .

Suppose  $N$  is fed with a text  $T$  for the language  $L$  in  $\mathcal{C}$ . Since  $N$  only outputs indices conjectured by  $M$  on  $T$ , and  $M$  is a *ConsvPart*-learner of  $\mathcal{C}$ ,  $N$  never conjectures a proper superset of  $L$ . It remains to show that  $N$  explanatorily learns  $L$ . Suppose that in the numbering  $L_0, L_1, L_2, \dots$ ,  $e$  is the least index for  $L$ . There is an  $n$  sufficiently large so that for all  $k > n$ ,  $H_{e,k} = H_e$  and  $e$  is the least index not exceeding  $k$  with  $H_{e,k} \subseteq \text{content}(T[k]) \subseteq L_e$ . Furthermore, as  $M$  outputs at least one correct index for  $L$ , there is a least number  $l$  such that  $H_e \subseteq W_{M(T[l])} \subseteq L_e$ . Thus  $N$  will converge to  $M(T[l])$  in the limit, and since  $M$  only outputs indices of languages in  $\mathcal{C}$ , it follows that  $W_{M(T[l])} = L_e$ . Thus  $N$  is a class-preserving explanatory learner of  $\mathcal{C}$  which never conjectures a proper superset of any target language  $L \in \mathcal{C}$  on every text for  $L$ . ■

Theorems 29 and 34 imply the following corollary.

**Corollary 35.** *If a uniformly recursive family  $\mathcal{C}$  is *ConsvPart*-learnable with respect to a class-preserving hypothesis space, as well as *Ex*-learnable, then it is *ConsvEx*-learnable.*

## References

1. Dana Angluin. Inductive inference of formal languages from positive data. *Information and Control* 45(2) (1980): 117–135.

2. Ganesh Baliga, John Case and Sanjay Jain. The synthesis of language learners. *Information and Computation* 152 (1999): 16–43.
3. Janis Bārzdīņš. Two theorems on the limiting synthesis of functions. In *Theory of Algorithms and Programs, vol. 1*, pages 82–88. Latvian State University, 1974. In Russian.
4. John Case and Timo Kötzing. Difficulties in forcing fairness of polynomial time inductive inference. *Algorithmic Learning Theory*, Twentieth International Conference, ALT 2009, Porto, Portugal, October 3-5, 2009. Proceedings. Springer LNAI 5809 (2009): 263–277.
5. John Case and Chris Lynes. Machine inductive inference and language identification. *Proceedings of the Ninth International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science 140 (1982): 107–115.
6. John Case and Carl Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science* 25 (1983):193–220.
7. Ziyuan Gao, Sanjay Jain and Frank Stephan. On conservative learning of recursively enumerable languages. Accepted for the conference Computability in Europe (CiE 2013).
8. Ziyuan Gao, Frank Stephan, Guohua Wu and Akihiro Yamamoto. Learning families of closed sets in matroids. *Computation, Physics and Beyond; International Workshop on Theoretical Computer Science*, WTCS 2012, Springer LNCS 7160 (2012): 120–139.
9. Ziyuan Gao and Frank Stephan. Confident and consistent partial learning of recursive functions. *Algorithmic Learning Theory*, Twenty-third International Conference, ALT 2012, Lyon, France, October 2012, Proceedings. Springer LNAI 7568 (2012): 51–65.
10. E. Mark Gold. Language identification in the limit. *Information and Control* 10 (1967): 447–474.
11. Sanjay Jain, Daniel Osherson, James S. Royer and Arun Sharma. 1999. *Systems that learn: an introduction to learning theory*. MIT Press, Cambridge, Massachusetts.
12. Sanjay Jain and Frank Stephan. Consistent partial identification. COLT 2009: 135–145.
13. Dick de Jongh and Makoto Kanazawa. Angluin’s theorem for indexed families of r.e. sets and applications. *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, pages 193–204, ACM Press, 1996.
14. Daniel N. Osherson, Michael Stob and Scott Weinstein. 1986. *Systems that learn: an introduction to learning theory for cognitive and computer scientists*. Cambridge, Massachusetts.: MIT Press.
15. Leonard Pitt. Inductive inference, DFAs, and computational complexity. *Analogical and Inductive Inference, Proceedings of the Second International Workshop*, AII 1989. Springer LNAI 397 (1989): 18–44.
16. Hartley Rogers, Jr. 1987. *Theory of recursive functions and effective computability*. Cambridge, Massachusetts: MIT Press.
17. Rolf Wiehagen and Thomas Zeugmann. Learning and consistency. *Algorithmic Learning for Knowledge-Based Systems*, GOSLER Final Report, Springer LNAI 961 (1995): 1–24.
18. Thomas Zeugmann, Steffen Lange and Shyam Kapur. Characterizations of monotonic and dual monotonic language learning. *Information and Computation* 120(2) (1995): 155–173.