Intrinsic Complexity of Partial Learning

Sanjay Jain[*1] and Efim Kinber[**2]

[1] School of Computing, National University of Singapore, Singapore 117417. Email: sanjay@comp.nus.edu.sg

[2] Department of Computer Science, Sacred Heart University, Fairfield, CT 06432-1000, U.S.A. Email: kinbere@sacredheart.edu

Abstract. A partial learner in the limit [16], given a representation of the target language (a text), outputs a sequence of conjectures, where one correct conjecture appears infinitely many times and other conjectures each appear a finite number of times. Following [5] and [14], we define intrinsic complexity of partial learning, based on reducibilities between learning problems. Although the whole class of recursively enumerable languages is partially learnable (see [16]) and, thus, belongs to the complete learnability degree, we discovered a rich structure of incomplete degrees, reflecting different types of learning strategies (based, to some extent, on topological structures of the target language classes). We also exhibit examples of complete classes that illuminate the character of the strategies for partial learning of the hardest classes.

1   Introduction

In his seminal paper [8], E. M. Gold introduced the framework for algorithmic learning of languages in the limit from their representations (texts), which became the standard for exploring learnability of languages in the limit (see, for example [16]). In this model (we will refer to it as **TxtEx**), a learner outputs an infinite sequence of conjectures stabilizing on a correct grammar for the target language. However, Gold himself was the first one to notice that the **TxtEx** model has a strong limitation: whereas the class of all finite languages is easily learnable within this framework, no class $\mathcal{L}$ containing just one infinite language and all its finite subsets is **TxtEx**-learnable. In particular, the class of all regular languages cannot be learnt in the limit just from positive data. To capture the extent to which the aforementioned class $\mathcal{L}$ would still be learnable from positive data, Osherson, Stob and Weinstein [16] introduced the concept of *partial learning* in the limit: a learner outputs an infinite sequence of conjectures, where one correct grammar of the target language occurs infinitely many times, whereas all other conjectures occur at most a finite number of times. The aforementioned class $\mathcal{L}$ containing an infinite recursive language L and all its finite subsets is easily learnable in this model by a simple strategy that, every time when a new datum appears on the input, conjectures a grammar for L, and conjectures

some standard code for the input seen so far, otherwise. Yet, as it was noted in [16], partial learning, without any other constraints, is very powerful: the whole class of all recursively enumerable languages turns out to be partially learnable — albeit by a much more complex strategy than the one trivially learning the aforementioned class $\mathcal{L}$.

Partial learning, under various natural constraints, has attracted a lot of attention recently (see, for example, [15, 9, 6, 7, 10]). Though partial learning can be done for the whole class of recursively enumerable sets, partial learning with constraints gives interesting results. Although partial learning does not seem to be as natural as Gold's classical model of inductive inference, one can hope that partial learning strategies for important classes of languages (like the class of regular languages) not learnable within Gold's framework can shed new light on the general problem of learnability of such classes (and, perhaps, their important subclasses) from positive data and, possibly, additional information. For example, if a relatively simple partial learning strategy for the class of regular languages from positive data is found, one can try to look at what kind of reasonable additional information could be sufficient for converting such partial strategy to a more realistic learning strategy for this class (perhaps, it could be different from Angluin's classical strategy for learning regular languages from membership queries and counterexamples to conjectures [2]). We hope that our paper can be a start for this line of research.

One of the potential issues is to understand exactly how partial learning happens and what is involved in it, as, at no particular instant, one can say what is the current "planned" hypothesis of the learner. To understand more about partial learning, we consider reductions between different learning problems (classes of languages). Reductions gave an interesting structure in explanatory learning (see [5, 14, 13, 12, 4]), and we hope to be able to understand much more about partial learning using reductions between different classes, which would, in some sense, highlight the ease/difficulty of partial learning of various subsets of the full class of all recursively enumerable languages.

Thus, our main goal in the current research is to find natural, yet non-**TxtEx**-learnable, classes (in particular, *indexed* classes [1], with decidable membership problem) and — whenever it would be possible — corresponding natural partial learning strategies that would be simpler than that for the class of all recursively enumerable languages. The concept of *reducibility* between partial learnability problems that we introduce in this paper is based on similar models defined first for learning in the limit of classes of recursive functions in [5] and then, for **TxtEx**-learnability in [14] (see also [4] for a related but different concept of complexity of learning).

A partial learnability problem (a class of languages $\mathcal{L}$) is reducible to another partial learning problem (a class of languages $\mathcal{L}'$) if there exist two computable operators, Θ and Ψ, such that (a) Θ translates every text for a language in $\mathcal{L}$ to a text for a language in $\mathcal{L}'$ and (b) Ψ translates every sequence of conjectures where a grammar for a language $L' \in \mathcal{L}'$ occurs infinitely many times and all other conjectures occur at most a finite number of times (we will call such sequences

of conjectures *admissible*) back to an admissible sequence of conjectures for the language $L \in \mathcal{L}$ such that some text for L is translated by Θ to a text for L'. We make a distinction between *strong* reducibility, where Θ translates every text for the same language in $\mathcal{L}$ to a text for the same language in $\mathcal{L}'$ and *weak* reducibility, where Θ may translate different texts of the same language $L \in \mathcal{L}$ to texts of different languages in $\mathcal{L}'$. Based on this concept of reducibility, one can naturally define degrees of learnability and the complete degree (which contains the class of all recursively enumerable languages).

Firstly, we found two relatively simple and transparent classes that are complete for weak and, respectively, strong reducibilities — these classes illuminate the character of the partial learning strategies for the hardest problems. We also show (Theorem 13) that the class of all recursive languages is not strongly complete. In particular, it means that all indexed classes, including the class of all regular languages, are not strongly complete.

A major accomplishment of our research is the discovery of a rich structure of incomplete classes under the degree of the class of all regular languages — based on a number of classes representing certain natural partial learning strategies. In particular, we define the class **iCOINIT**, which contains an infinite chain $L_1, L_2, \ldots$ of infinite recursive subsets of a recursive infinite language, and all their finite subsets, where, for every i, $L_{i+1} \subset L_i$. The natural strategy to learn this class, when choosing an infinite language as its conjecture, immediately finds out an upper bound on the possible number of infinite languages that may be conjectured in the future. We also define the counterpart of **iCOINIT**, the class **iINIT**, which also contains an enumerable, but indefinitely growing chain of infinite recursive languages and all their finite subsets. The natural learning strategy for **iINIT**, when choosing an infinite language as its conjecture, also faces a bound on the number of infinite languages that can be conjectured in the future, but unlike the case of **iCOINIT**, this bound is not known to the learner. We show that **iCOINIT** is weakly reducible to **iINIT** (Theorem 15), yet it is not strongly reducible to **iINIT** (Theorem 16); also, **iINIT** is not even weakly reducible to **iCOINIT** (Theorem 17).

We also introduce the class **iRINIT**, which contains an infinitely growing chain of recursive languages and all their finite subsets, yet, unlike the case of **iINIT**, the enumeration of members of the chain is based on the set of all rational numbers between 0 and 1. In particular, for any two infinite languages $L, L' \in$ **iRINIT**, $L \subset L'$, there is another language in **iRINIT** between L and L'. We show that **iRINIT** is not weakly complete (Corollary 19), yet all variants and generalizations formed using **iINIT** and **iCOINIT** (as defined in this paper) are strongly reducible to **iRINIT** (Theorem 26), and **iRINIT** is strongly reducible to none of them (Theorem 26). On the other hand, **iRINIT** itself turns out to be weakly reducible to **iINIT** (Theorem 22). **iRINIT** is also strictly under the degree of all regular languages (Theorem 18 and Theorem 24).

We also define a variant **iCOINIT**$_k$ of **iCOINIT**, which, in addition to every infinite language L in the chain, contains also all languages extending L by at most k additional elements. A natural strategy learning an infinite target lan-

guage $L \in \mathbf{iCOINIT}_k$, when a new datum appears on the input, first conjectures an infinite language $M \in \mathbf{iCOINIT}$, and when up to k new elements $x \notin M$ appear on the input, conjectures appropriate finite variants of M, before moving to the next M' in the chain when the number of new data not in M exceeds k. Similarly the variant $\mathbf{iINIT}_k$ is defined for $\mathbf{iINIT}$. Interestingly, though, all classes $\mathbf{iCOINIT}_k, k \geq 1$ turn out to be strongly reducible to $\mathbf{iCOINIT}_1$ and, respectively, all classes $\mathbf{iINIT}_k, k \geq 1$ are strongly reducible to $\mathbf{iINIT}_1$. Yet, surprisingly, $\mathbf{iINIT}_1$ is not strongly reducible to $\mathbf{iINIT}$ and $\mathbf{iCOINIT}_1$ is not even weakly reducible to $\mathbf{iCOINIT}$ (see Theorem 25). All these classes, though, are weakly reducible to $\mathbf{iINIT}$ (as $\mathbf{iRINIT}$ is weakly reducible to $\mathbf{iINIT}$, see above).

Lastly, based on similar multidimensional classes of languages defined in [13], one can define classes of "multidimensional" languages, where partial learning of one "dimension" aids in learning next "dimension". For example, one can, using cylindrification, define the class $(\mathbf{iINIT}, \mathbf{iCOINIT})$, where the conjecture that is output infinitely many times for the first "dimension" can be used to partially learn the second "dimension". We have extended this idea to any arbitrary sequence Q of $\mathbf{iINIT}$ and $\mathbf{iCOINIT}$ and have shown that if a sequence Q is a proper subsequence of Q', then the class corresponding to Q is strongly reducible to the one corresponding to Q', but not vice versa. Due to space constraints, results on multi-dimensional languages are not described in this paper but will be given in the full paper.

Our result on the incompleteness of any indexed class suggests that there may exist natural, relatively simple, strategies that can partially learn an indexed class. This can shed a new light on the potential of learnability of many important classes of languages from positive data.

2 Preliminaries

Any unexplained recursion theoretic notation is from [17]. N denotes the set of natural numbers $\{0, 1, 2, \ldots\}$. A language is any subset of N. We let $\emptyset, \subseteq, \subset, \supseteq, \supset$ denote empty set, subset, proper subset, superset and proper superset respectively. $A \mathbf{\Delta} B$ denotes the symmetric difference of sets A and B, that is $A \mathbf{\Delta} B = (A - B) \cup (B - A)$. $\overline{L} = N - L$ denotes the complement of L. We let $\mathrm{card}(S)$ denote the cardinality of a set S. For $S \subseteq N$, let $\max(S), \min(S)$ respectively denote maximum and minimum of a set S, where $\min(\emptyset) = \infty$ and $\max(\emptyset) = 0$. We sometimes use sets of rational numbers. In this case, we use $\max(S)$ to denote the least upper bound of the rational numbers in the set S.

A finite set $S \subseteq N$ can be coded as $\mathrm{code}(S) = \sum_{x \in S} 2^x$. D_i denotes the finite set A with $\mathrm{code}(A) = i$.

φ denotes a fixed standard acceptable numbering [17]. φ_i denotes the i-th program in the acceptable numbering φ. Let $W_i = \mathrm{domain}(\varphi_i)$. Thus, W_i is the language/set enumerated by the i-th grammar in the acceptable programming system $W_0, W_1, \ldots$. Let Φ be a Blum complexity measure [3] for the φ

programming system. Let

$$\varphi_{i,s}(x) = \begin{cases} \varphi_i(x), & \text{if } x < s \text{ and } \Phi_i(x) < s; \\ \uparrow, & \text{otherwise} \end{cases}$$

Let $W_{i,s} = \text{domain}(\varphi_{i,s})$. Intuitively, $W_{i,s+1} - W_{i,s}$ can be thought of as the elements enumerated by W_i in $(s+1)$-th step. For purposes of this paper, one can assume without loss of generality that $W_{i,s+1} - W_{i,s}$ contains at most one element.

$\mathcal{R}$ denotes the class of all recursive languages. $\mathcal{E}$ denotes the class of all recursively enumerable sets.

An indexed family is a family $(L_i)_{i \in N}$ of languages such that there exists a recursive function f uniformly deciding the membership question for L_i, that is, for all i, x, $f(i,x) = 1$ iff $x \in L_i$.

Let $\langle \cdot, \cdot \rangle$ denote a fixed recursive bijection from $N \times N$ to N. $\langle \cdot, \cdot \rangle$ can be extended to pairing of n-ary tuples by taking $\langle x_1, x_2, \ldots, x_n \rangle = \langle x_1, \langle x_2, \ldots, x_n \rangle \rangle$. For notation convenience we let $\langle x \rangle = x$. Let $\pi_i^n(\langle x_1, x_2, \ldots, x_n \rangle) = x_i$, where we drop the superscript in case $n = 2$.

Let $pad(\cdot, \cdot)$ be a 1–1 recursive function, increasing in both its arguments, such that for all i and j, $\varphi_{pad(i,j)} = \varphi_i$. Note that there exists such a padding function pad (see [17]).

$RAT_{0,1}$ denotes the set of rational numbers between 0 and 1 (both inclusive). Let ntor be a recursive bijection from N to $RAT_{0,1}$. Let rton be the inverse of ntor. Left r.e. real means a real number which is approximable from below using rational numbers enumerated by a recursive procedure. That is, a real number r is called a left r.e. real iff there exists a recursive function f mapping N to the set of rational numbers such that: for all i, $f(i) \leq f(i+1)$, and $\lim_{i \in N} f(i) = r$.

We now give some concepts from language learning theory. Let $\#$ be a special pause symbol. A *finite sequence* σ is a mapping from an initial segment of N to $(N \cup \{\#\})$. Let Λ denote the empty sequence. SEQ denotes the set of all finite sequences. A *text* is a mapping from N to $(N \cup \{\#\})$. Let $|\sigma|$ denote the length of sequence σ. Let $T[n]$ denote the initial segment of length n of the text T. For $n \leq |\sigma|$, $\sigma[n]$ denotes the initial segment of length n of the sequence σ. The concatenation of sequences σ and τ is denoted by $\sigma \diamond \tau$. The content of T, denoted content(T), is the set of the numbers in the range of T, that is, $\{T(n) : n \in N\} - \{\#\}$. Let content$(\sigma)$ be defined similarly. We say that T is a text for a language L iff content$(T) = L$.

A *language learning machine* is a partial computable function which maps SEQ to N. We let $\mathbf{M}$, with or without decorations, range over learning machines.

Definition 1. [16]

(a) $\mathbf{M}$ **Part**-learns L iff for all texts T for L,
   (i) for all n, $\mathbf{M}$ is defined on $T[n]$,
   (ii) there exists a unique p such that $p = \mathbf{M}(T[n])$ for infinitely many n, and
   (iii) for p as in (ii) above, $W_p = L$.

(b) **M Part**-learns a class $\mathcal{L}$ iff it **Part**-learns each $L \in \mathcal{L}$.

(c) **Part** $= \{\mathcal{L} : (\exists \mathbf{M})[\mathbf{M}\ \mathbf{Part}\text{-learns}\ \mathcal{L}\}$.

It can be shown that $\mathcal{E}$ is **Part**-learnable [16]. If **M Part**-learns a class $\mathcal{L}$ then we say that **M** witnesses **Part**-learnability of $\mathcal{L}$. If an infinite sequence $p_0 p_1 \ldots$ satisfies the following two requirements:

(i) there exists a unique p such that $p = p_n$ for infinitely many n, and

(ii) for p as in (i) above, $W_p = L$,

then, we say that the sequence $p_0 p_1 \ldots$ witnesses **Part**-learnability of L.

An *enumeration operator* (or just operator) Θ is an algorithm mapping from SEQ to SEQ such that for all $\sigma, \tau \in SEQ$, if $\sigma \subseteq \tau$, then $\Theta(\sigma) \subseteq \Theta(\tau)$. We let $\Theta(T) = \bigcup_{n \in N} \Theta(T[n])$.

We further assume that $\lim_{n \to \infty} |\Theta(T[n])| = \infty$, that is texts are mapped to texts by the operator Θ. Note that any operator Θ can be modified to satisfy the above property without violating the content of its output on infinite texts.

We will also use Θ as an operator on languages (rather than individual texts representing them, as above). Note that, in general, for different texts T, T' of a language L, Θ may produce texts $\Theta(T)$ and $\Theta(T')$ of different languages. Thus, we define $\Theta(L)$ as a collection of languages: $\Theta(L) = \{\text{content}(\Theta(T)) : T$ is a text for $L\}$, and, accordingly, the image $\Theta(\mathcal{L}) = \bigcup_{L \in \mathcal{L}} \Theta(L)$. In the special case (important for our strong reductions, defined below), when $\Theta(L)$ is a singleton $\{L'\}$, we abuse notation and say simply $\Theta(L) = L'$. (Note that if $\Theta(L) = \{L'\}$, then $L' = \bigcup_{\sigma:\text{content}(\sigma) \subseteq L} \text{content}(\Theta(\sigma))$).

We let Θ and Ψ range over operators, where for ease of notation, we assume that for Ψ the input and output sequences contain only elements of N (and thus do not contain #). We view Ψ as mapping sequences of grammars to sequences of grammars. Again, as in the definition of operator Θ, we assume that Ψ maps infinite sequences to infinite sequences. This can be easily done without changing the set of grammars which appear infinitely often in the sequence.

The following two definitions are based on the corresponding reductions for explanatory function learning [5] and explanatory language learning [14]. In these definitions, we view operators Θ as mapping texts to texts, as well as mapping languages to collections of languages (as discussed above).

Definition 2. We say that $\mathcal{L} \leq_{\mathbf{Part}}^{\mathrm{Weak}} \mathcal{L}'$ iff there exist operators Θ and Ψ such that

(a) for all $L \in \mathcal{L}$, $\Theta(L) \subseteq \mathcal{L}'$.

(b) for all $L \in \mathcal{L}$, for all $L' \in \Theta(L)$, if $p_0 p_1 \ldots$ is a sequence witnessing **Part**-learnability of L', then $\Psi(p_0 p_1 \ldots)$ witnesses **Part**-learnability of L.

Intuitively, Θ reduces a text for a language $L \in \mathcal{L}$ to a text for a language $L' \in \mathcal{L}'$. Ψ then converts sequences witnessing **Part**-learnability of L' to sequences witnessing **Part**-learnability of L.

However, as we noted above, different texts for L may be mapped by Θ to texts for different languages in $\mathcal{L}'$. If we require that the mapping should be to the texts of the same language, then we get strong reduction.

Definition 3. We say that $\mathcal{L} \leq^{Strong}_{\mathbf{Part}} \mathcal{L}'$, iff there exist operators Θ and Ψ such that

(a) Θ and Ψ witness that $\mathcal{L} \leq^{Weak}_{\mathbf{Part}} \mathcal{L}'$ and
(b) for all $L \in \mathcal{L}$, $\mathrm{card}(\Theta(L)) = 1$.

For ease of notation, when considering strong reductions, as discussed above, we consider Θ as directly mapping languages to languages, rather than considering it as a mapping from languages to a set containing just one language.

We say that $\mathcal{L} <^{Weak}_{\mathbf{Part}} \mathcal{L}'$ if $\mathcal{L} \leq^{Weak}_{\mathbf{Part}} \mathcal{L}'$ but $\mathcal{L}' \not\leq^{Weak}_{\mathbf{Part}} \mathcal{L}$. Similarly, $\mathcal{L} \equiv^{Weak}_{\mathbf{Part}}$ if $\mathcal{L} \leq^{Weak}_{\mathbf{Part}} \mathcal{L}'$ and $\mathcal{L}' \leq^{Weak}_{\mathbf{Part}} \mathcal{L}$.

Similarly, we can define $\mathcal{L} <^{Strong}_{\mathbf{Part}} \mathcal{L}'$ and $\mathcal{L} \equiv^{Strong}_{\mathbf{Part}} \mathcal{L}'$.

Definition 4. We say that $\mathcal{L}$ is $\leq^{Weak}_{\mathbf{Part}}$-complete if

(a) $\mathcal{L} \in \mathbf{Part}$ and
(b) For all $\mathcal{L}' \in \mathbf{Part}$, $\mathcal{L}' \leq^{Weak}_{\mathbf{Part}} \mathcal{L}$.

$\leq^{Strong}_{\mathbf{Part}}$-completeness can be defined similarly.

We now define some languages and classes which are often used in the paper. We used the names **iINIT**, **iCOINIT**, **iRINIT** for the classes defined below as the infinite languages in these classes are obtained by cylindrification of the languages in $INIT$, $COINIT$ and $RINIT$ used in the literature (the class $INIT$ contains languages $\{1, 2, \ldots, i\}$ and $COINIT$ contains languages $\{i, i + 1, i + 2, \ldots\}$; $RINIT$ is similar to $INIT$ and contains, for each $r \in R_{0,1}$, the language having (the representatives of) rational numbers below r). Additionally the classes **iINIT**, **iCOINIT**, **iRINIT** contain all the finite languages. For $i \in N, r \in RAT_{0,1}$,

(a) $INIT_i = \{\langle x, y \rangle : x, y \in N \text{ and } x \leq i\}$,
(b) $COINIT_i = \{\langle x, y \rangle : x, y \in N \text{ and } x \geq i\}$,
(c) $RINIT_r = \{\langle x, y \rangle : x, y \in N \text{ and } \mathrm{ntor}(x) \leq r\}$,
(d) $INIT_{i,s} = D_s \cup INIT_i$,
(e) $COINIT_{i,s} = D_s \cup COINIT_i$,
(f) $\mathcal{FIN} = \{L : L \text{ is finite}\}$,
(g) **iINIT** $= \{INIT_i : i \in N\} \cup \mathcal{FIN}$,
(h) **iCOINIT** $= \{COINIT_i : i \in N\} \cup \mathcal{FIN}$,
(i) **iRINIT** $= \{RINIT_r : r \in RAT_{0,1}\} \cup \mathcal{FIN}$,
(j) **iINIT**$_k$ $=$ **iINIT** $\cup \{INIT_{i,s} : \mathrm{card}(D_s) \leq k, i \in N\}$,
(k) **iCOINIT**$_k$ $=$ **iCOINIT** $\cup \{COINIT_{i,s} : \mathrm{card}(D_s) \leq k, i \in N\}$.

A natural partial learning strategy for the languages in **iINIT** is as follows: when a new datum $\langle j, x \rangle$, where j is larger than all m for all pairs $\langle m, y \rangle$ seen so far, appears on the input, the learner, for the first time, outputs the conjecture $INIT_j$. Now, as long as no new (not previously seen) datum appears on the input, the learner conjectures the finite set representing the input seen so far; if a new datum (not previously seen) from $INIT_j$ appears on the input, the learner repeats the conjecture $INIT_j$. This continues as long as no datum outside

$INIT_j$ is seen. Clearly, the correct language $INIT_j$ or some finite input set is the only one that will be conjectured infinite number of times. A similar strategy works for **iRINIT**.

For **iCOINIT** a similar strategy chooses a new infinite conjecture $COINIT_j$ when a new pair $\langle j, x \rangle$, where j is smaller than all m for all pairs $\langle m, y \rangle$ seen so far, appears on the input. Otherwise, the strategy is identical to the one for **iINIT**.

For **iINIT**$_k$ the above **iINIT**-learning strategy can be adjusted as follows: the learner keeps track of the smallest j such that the set $E = \{\langle x, y \rangle : x > j$ and $\langle x, y \rangle$ is seen in the input so far$\}$ has at most k elements. Then, the strategy for learning **iINIT**$_k$ is similar to that of **iINIT**, except that whenever the strategy for **iINIT** outputs an infinite conjecture, strategy for **iINIT**$_k$ outputs an infinite conjecture for $INIT_{j,s}$, where $D_s = E$, where j, E are as described above. Similar modification to the strategy for **iCOINIT** works for **iCOINIT**$_k$.

3   Basic Properties of Reductions

In this section, we establish a number of technical facts used in many proofs of our results.

Lemma 5. *Suppose Θ witnesses part (a) of Definition 2 for $\mathcal{L} \leq_{\mathbf{Part}}^{Weak} \mathcal{L}'$. Suppose F_1, F_2 are computable functions such that, for any $L \in \mathcal{L}$ and $L' \in \Theta(L)$, the following three properties hold:*

  (i)   *if L' is finite, then $F_1(L')$ is a grammar for L.*
  (ii) *if L' is infinite and p is a grammar for L', then $\lim_{t \to \infty} F_2(p, t)$ exists and is a grammar for L.*
  (iii) *if L' is infinite, then for any sequence of finite sets $S_1, S_2, \ldots$ such that $S_0 \subset S_1 \subset S_2 \ldots$ and $\bigcup_{i \in N} S_i = L'$, for all t, for all but finitely many t', $F_1(S_t) \neq F_1(S_{t'})$.*

*Then, there exists a Ψ such that, for all $L \in \mathcal{L}$, for any sequence of grammars $p_0 p_1 \ldots$ witnessing **Part**-learnability of $L' \in \Theta(L)$, $\Psi(p_0 p_1 \ldots) = q_0 q_1 \ldots$ witnesses **Part**-learnability of L.*

The above lemma is useful in simplifying the construction of Ψ in many of the proofs: we can just give the relevant F_1 and F_2.

Proposition 6. *There exists an operator Ψ such that for any sequence $q_0 q_1 \ldots$, $\Psi(q_0 q_1 \ldots) = q_0' q_1', \ldots$ such that:*

(a) *at most one grammar appears in $q_0' q_1' \ldots$ infinitely often,*
(b) *if q is the least grammar which appears infinitely often in $q_0 q_1 \ldots$, then q' appears infinitely often in $q_0' q_1' \ldots$, where $W_{q'} = W_q$.*

Proof. Let $q_i' = pad(q_i, j)$, where $j = \text{card}(\{i' : i' \le i \text{ and } q_{i'} < q_i\})$. It is easy to verify that the above sequence satisfies the requirements of the proposition. ∎

The above proposition is useful to simplify some of the constructions for Ψ in our proofs.

Proposition 7. *Suppose $\mathcal{L} \le_{\mathbf{Part}}^{Weak} \mathcal{L}'$ as witnessed by Θ and Ψ. Then, for all distinct $L, L' \in \mathcal{L}$, $\Theta(L) \cap \Theta(L') = \emptyset$.*

Proposition 8. *For any operator Θ, if $L \subseteq L'$, $\Theta(L) = \{X\}$ and $\Theta(L') = \{X'\}$, then $X \subseteq X'$.*

Proposition 9. *Suppose L is infinite and $\mathcal{L}$ contains L and all finite subsets of L. Suppose further that $\mathcal{L} \le_{\mathbf{Part}}^{Weak} \mathcal{L}'$ as witnessed by Θ and Ψ. Then, for all finite sets S such that $S \subseteq L'$ for some $L' \in \Theta(L)$, there exists an infinite superset of S in $\Theta(L)$ (in particular, $\Theta(L)$ contains an infinite language).*

4 Complete Classes and the Class $\mathcal{R}$

As $\mathcal{E} \in \mathbf{Part}$, we trivially have that $\mathcal{E}$ is $\le_{\mathbf{Part}}^{Strong}$-complete. The following results give some simple classes which are complete.

Let $\mathbf{iCOINIT}_* = \{COINIT_i \cup A : i \in N \text{ and } A \text{ is finite}\}$. We first show that every text for every recursively enumerable language can be appropriately "encoded" as a text for some language in $\mathbf{iCOINIT}_*$ — thus showing that $\mathbf{iCOINIT}_*$ is weakly complete.

Theorem 10. $\mathbf{iCOINIT}_*$ *is $\le_{\mathbf{Part}}^{Weak}$-complete.*

Proof. To show that $\mathcal{E} \le_{\mathbf{Part}}^{weak} \mathbf{iCOINIT}_*$, define Θ and Ψ as follows.

Suppose T is a given text. Let $C_{p,T} = \max(\{t : W_{p,t} \subseteq \text{content}(T) \text{ and } \text{content}(T[t]) \subseteq W_p\})$. Note that $C_{p,T}$ can be approximated from below (that is, there exists a recursive function f such that $f(p, T[n]) \le f(p, T[n+1]) \le C_{p,T}$ and $\lim_{n \to \infty} f(p, T[n]) = C_{p,T}$.

Define Θ as follows. $\Theta(T) = T'$ such that $\text{content}(T') = \{\langle q, x \rangle : (\exists q' \le q)[x \le C_{q',T}]\}$.

Now, suppose p is the least grammar for W_p, T is a text for W_p, and $T' = \Theta(T)$. Then, it is easy to verify that $\text{content}(T') = COINIT_{p,s}$, for some s, as $C_{p,T} = \infty$, but $C_{p',T} < \infty$ for $p' < p$.

We define Ψ' as follows.

$\Psi'(p_0 p_1 \ldots) = q_0 q_1 \ldots$, where

$q_i = pad(j, p_i)$, where, for some x, $\langle j, x \rangle$ is the last new element enumerated in $W_{p_i, k}$ and k is the number of times p_i appears in $p_0 p_1 \ldots p_i$.

Claim 11. *Suppose $L \in \mathcal{E}$, and $L' \in \Theta(L)$.*

If $p_0 p_1 \ldots$ witnesses $\mathbf{Part}$-learnability of L', then $q_0 q_1 \ldots$ is a sequence satisfying:

there exists a minimal q such that q appears infinitely often in $q_0 q_1 \ldots$ and this $q = pad(j, p)$, for minimal grammar j for L and some p.

To see the claim, suppose $L' \in \Theta(L)$ and $p_0, p_1 \ldots$ is a sequence witnessing **Part**-learnability of L' and $\Psi'(p_0 p_1 \ldots) = q_0 q_1 \ldots$. Suppose p is the only grammar which appears infinitely often in $p_0 p_1 \ldots$. Then only q_i with $p_i = p$ can possibly appear infinitely often in the sequence $q_0 q_1 \ldots$ as q_i used p_i in its padding. Furthermore, $pad(j, p)$, with $j \geq \min(\{e : W_e = L\})$ appear infinitely often in the sequence $q_0 q_1 \ldots$. The claim follows. ∎

Now the theorem follows using Proposition 6.

We now consider a $\leq_{\mathbf{Part}}^{Strong}$-complete class.

Let $V(L) = \frac{1}{4} + \sum_{x \in L} 4^{-x-1}$.

Intuitively, V maps languages to real numbers where the mapping is monotonic in L. Furthermore, if $L \neq L'$ and $\min(L \mathbf{\Delta} L') \in L$, then $V(L) > V(L')$.

The reason for choosing the additive part "$\frac{1}{4}$" is just to make sure that $V(L)$ is non-zero.

Let $L_{r_0, r_1, \ldots, r_k} = \{\langle i, x \rangle : i < k \text{ and } \mathrm{ntor}(x) < r_i \text{ or } i \geq k \text{ and } \mathrm{ntor}(x) < r_k\}$.

Let $\mathcal{STRCOMP} = \{L_{r_0, r_1, \ldots, r_k} : k \in N, r_0 \leq r_1 \leq \ldots \leq r_k, \text{and } r_0, r_1, \ldots, r_k$ are left r.e. reals$\}$.

$\mathcal{STRCOMP}$ denotes "strong complete class". The languages in $\mathcal{STRCOMP}$ can be thought of as follows: in the i-th cylinder we keep rational numbers $< r_i$. The r_i's are monotonically non-decreasing left r.e. real numbers and the sequence $r_0, r_1, \ldots$ converges (that is, for some k, for all $i \geq k$, $r_i = r_k$). We suggest the reader to contrast this class with the previously defined class **iRINIT** (which, in the sequel, will be shown to be incomplete).

Theorem 12. $\mathcal{STRCOMP}$ is $\leq_{\mathbf{Part}}^{Strong}$-complete.

Proof. For any index i and any language L, let

$X_{i,L} = \{\frac{x}{y} : x, y \in N, y \neq 0, \frac{x}{y} < V(W_i \cap L) \text{ and } x, y \leq \min(\{t : W_{i,t} - L \neq \emptyset\})\}$.

Intuitively, $sup(X_{i,L})$ gives a value to how much W_i and L are similar to each other:

(P1) If $W_i = L$, then $sup(X_{i,L})$ is $V(L)$ (as $\min(\{t : W_{i,t} - L \neq \emptyset\})$ is infinite).

(P2) If $W_i \neq L$, then $sup(X_{i,L}) < V(L)$. To see this note that if $L \not\subseteq W_i$, then clearly $V(W_i \cap L) < V(L)$ and thus $sup(X_{i,L}) \leq V(W_i \cap L) < V(L)$. On the other hand, if $W_i \not\subseteq L$, then $sup(X_{i,L}) < V(W_i \cap L) \leq V(L)$, as $X_{i,L}$ is a finite set and the supremum of a finite set of rational numbers $< r$ is $< r$ for any positive real number r.

Note that $X_{i,L}$ depends only on i and L and not on the particular presentation of L. This allows us to construct Θ (and corresponding Ψ) which give a strong reduction from $\mathcal{E}$ to $\mathcal{STRCOMP}$.

$\Theta(L) = \bigcup_{i \in N} [\{\langle i, y \rangle : (\exists i' \leq i)(\exists s \in X_{i',L})[\mathrm{ntor}(y) \leq s]\}]$.

Intuitively, Θ just collects all the members of $X_{i,L}$ in the i-th cylinder and then does an upward closure.

Note that $\Theta(L)$ can be enumerated from a text T for L. Moreover, $\Theta(L)$ depends only on L and not on the particular presentation T, and thus the reduction is a strong reduction.

We say that m improves at time step j in the enumeration of W_p, if, for some x, the following two conditions are satisfied:

(i) $\langle m, x \rangle$ gets enumerated at time step j in W_p.
(ii) Suppose $j' < j$ is the largest earlier time step when m improved, if any, in the enumeration of W_p (take $j' = 0$, if m did not improve earlier). Then, for any y such that $\langle m', y \rangle \in W_{p,j'}$, $\mathrm{ntor}(x) > \mathrm{ntor}(y)$.

Note that for any grammar p for $L' = L_{r_0, r_1, \ldots, r_k}$, with $r_0 \leq r_1 \leq r_2 \ldots \leq r_{k-1} < r_k$, k improves at infinitely many steps j in the enumeration of W_p, but $0, 1, \ldots, k-1$ improve only for finitely many steps j in the enumeration of W_p.

We define an operator Ψ' as follows. This can then be converted to the desired Ψ using Proposition 6. $\Psi'(p_0 p_1 \ldots) = q_0 q_1 \ldots$, where q_i is defined below. Note that if $L \in \mathcal{E}$, $\Theta(L) = L'$ and $p_0, p_1, \ldots$ witnessed **Part**-learning of L', then we want $\Psi(p_0 p_1 \ldots)$ to witness **Part**-learning of L.

Without loss of generality assume that, for any k, W_k enumerates at most one element at any step. For any fixed i, suppose p_i appears j times in $p_0 p_1 \ldots p_i$, and m_i improves at step j in the enumeration of W_{p_i} (if no such m_i exists, then take m_i to be $i+1$). Then, let $q_i = pad(m_i, p_i)$.

Now suppose $L \in \mathcal{E}$ and $L' = \Theta(L)$ and i is the minimal grammar for L. Then, by the properties (P1) and (P2), for all $j < i$, $sup(X_{i,L}) > sup(X_{j,L})$, and for all $j \geq i$, $sup(X_{i,L}) \geq sup(X_{j,L})$. Thus, L' is of the form $L_{r_0, r_1, \ldots, r_i}$, where $r_0 \leq r_1 \leq r_{i-1} < r_i$.

Now, suppose p appears infinitely often in $p_0 p_1 \ldots$, which witnesses **Part**-learnability of L'. Then for $\Psi'(p_0 p_1 \ldots) = q_0 q_1 \ldots$ only q_j with $p_j = p$ could possibly appear infinitely often in the sequence $q_0 q_1 \ldots$ as q_j used p_j in its padding. Furthermore, i is the minimal number which improves at infinitely many steps in the enumeration of W_p. It follows that $pad(i, p)$ is the minimal element which appears infinitely often in the sequence $q_0 q_1 \ldots$.

Now Ψ' can be converted to the required Ψ using Proposition 6. The theorem follows. ∎

Our next result states that the class $\mathcal{R}$ of all recursive languages is not strongly complete. In particular, this means that all indexed classes of languages, including the class of all regular languages, are incomplete. This opens a possibility of creating partial learning strategies for these classes that would be simpler than the general strategy for partial learning of all recursively enumerable languages.

Theorem 13. $\mathcal{R}$ *is not* $\leq_{\mathbf{Part}}^{Strong}$*-complete.*

Proof. Suppose Θ (and Ψ) witness that $\mathcal{E} \leq_{\mathbf{Part}}^{Strong} \mathcal{R}$. Let K denote the halting set $\{i : i \in N$ and $\varphi_i(i)\downarrow\}$. Now, by Proposition 7 and Proposition 8, for all x, $\Theta(K) \subset \Theta(K \cup \{x\})$. Let $S = \Theta(K)$. Note that, by assumption, S is recursive. Now, $x \in K$ iff $\Theta(\{x\} \cup K) \subseteq S$. As S is recursive, this would imply that $\overline{K}$ is recursively enumerable, which is in contradiction to a known fact that K is not recursively enumerable [17]. ∎

5 Relationship between iINIT, iCOINIT and iRINIT classes

In this section, we explore the relationships between the classes **iINIT**, **iCOINIT** (and some of their variants), and **iRINIT**. We also establish that their degrees are strictly under the degree of all regular languages.

Proposition 14. *Fix $n \in N$. Suppose $\mathcal{L} \subseteq \mathcal{E}$ contains only n infinite languages. Then,*

*(a) $\mathcal{L} \leq_{\textbf{Part}}^{Strong}$ **iINIT**.*
*(b) $\mathcal{L} \leq_{\textbf{Part}}^{Strong}$ **iCOINIT**.*

First, we explore the relationship between **iINIT** and **iCOINIT**. We begin with establishing that **iCOINIT** is reducible to **iINIT**, but only weakly. Perhaps, this fact and the fact that **iINIT** is not reducible to **iCOINIT** (see below) are not surprising, as the chain of infinite languages in **iINIT** is growing indefinitely, whereas every growing chain of infinite languages in **iCOINIT** is finite.

Theorem 15. **iCOINIT** $\leq_{\textbf{Part}}^{Weak}$ **iINIT**.

Theorem 16. **iCOINIT** $\not\leq_{\textbf{Part}}^{Strong}$ **iINIT**.

On the other hand, **iINIT** is not reducible to **iCOINIT** even weakly.

Theorem 17. **iINIT** $\not\leq_{\textbf{Part}}^{Weak}$ **iCOINIT**.

The class **iRINIT** is similar to **iINIT** in that it features an infinitely growing chain of infinite languages. However, unlike **iINIT**, between any two infinite languages in **iRINIT**, there is always another language. We show that the degree of **iRINIT** is strictly above the degrees of the classes **iINIT** and **iCOINIT**. However, first we show that **iRINIT** is not even weakly complete. Let $\mathcal{REG}$ denote the class of all regular sets [11] (we assume some standard recursive bijection between strings and N, so that regular sets can be considered as subsets of natural numbers). Topologically, $\mathcal{REG}$ is much more complex than containing just one growing chain of infinite languages **iRINIT** (plus finite sets), and this translates into greater complexity of partial learning of $\mathcal{REG}$, as the following theorem indicates.

Theorem 18. $\mathcal{REG} \not\leq_{\textbf{Part}}^{Weak}$ **iRINIT**.

Proof. Suppose by way of contradiction that Θ and Ψ witness that $\mathcal{REG} \leq_{\textbf{Part}}^{\text{Weak}}$ **iRINIT**.

Inductively define σ_i as follows.

$\sigma_0 = \Lambda$.

σ_{i+1} is an extension of $\sigma_i \diamond i$ such that $\max(\{\text{ntor}(x) : \langle x, y \rangle \in \Theta(\sigma_{i+1})\}) > \max(\{\text{ntor}(x) : \langle x, y \rangle \in \Theta(\sigma_i)\})$.

If all σ_{i+1} get defined, then for $T = \bigcup_{i \in N} \sigma_i$, T is a text for N (which is regular), but $sup(\{\mathrm{ntor}(x) : \langle x, y \rangle \in \mathrm{content}(\Theta(T))\})$ does not belong to $\{\mathrm{ntor}(x) : \langle x, y \rangle \in \mathrm{content}(\Theta(T))\}$ (as if it belonged, then it would belong to $\{\mathrm{ntor}(x) : \langle x, y \rangle \in \mathrm{content}(\Theta(\sigma_j))\}$, for some j, and that would violate the definitions of σ_i's).

If some σ_{i+1} does not get defined, then let $r = \max(\{\mathrm{ntor}(x) : \langle x, y \rangle \in \mathrm{content}(\Theta(\sigma_i))\})$. Now, for all infinite regular languages L containing $\mathrm{content}(\sigma_i)$, $RINIT_r \in \Theta(L)$ (as $\Theta(L)$ contains an infinite language containing $\langle \mathrm{rton}(r), y \rangle$, for some y, and σ_{i+1} did not get defined). A contradiction to Proposition 7, as there are infinitely many (in particular at least two) infinite regular languages which contain $\mathrm{content}(\sigma_i)$. ∎

Corollary 19. **iRINIT** *is not* $\leq_{\mathbf{Part}}^{Weak}$*-complete.*

Our next result shows that both **iINIT** and **iCOINIT** are strongly reducible to **iRINIT**. Let $0 < r_0, r_1, \ldots$ in $RAT_{0,1}$ be a strictly increasing sequence of rational numbers. $\{INIT_i : i \in N\}$ can be naturally embedded into **iRINIT**, by mapping $INIT_i$ to $RINIT_{r_i}$. Note that, by our convention on coding of finite sets, $D_s \subset D_{s'}$ implies $s < s'$. For any s, let $k_s = \max(\{x : \langle x, y \rangle \in D_s\})$. Now, mapping finite sets D_s to $RINIT_{r_{k_s} + (r_{k_s+1} - r_{k_s}) * r_s}$ ensures that **iINIT** $\leq_{\mathbf{Part}}^{Strong}$ **iRINIT**. A similar method works to show that **iCOINIT** $\leq_{\mathbf{Part}}^{Strong}$ **iRINIT**.

Theorem 20. *(a)* **iINIT** $\leq_{\mathbf{Part}}^{Strong}$ **iRINIT**.
(b) **iCOINIT** $\leq_{\mathbf{Part}}^{Strong}$ **iRINIT**.

Next we show that **iRINIT** is neither strongly reducible to **iINIT**, nor even weakly reducible to **iCOINIT**. Yet, it is weakly reducible to **iINIT**. The latter fact is quite interesting: every text for a language in **iRINIT** can be encoded as a text for a language in **iINIT**, yet the corresponding languages in **iINIT** for such texts may be different for different texts of the same language in **iRINIT**.

Theorem 21. **iRINIT** $\not\leq_{\mathbf{Part}}^{Strong}$ **iINIT**.

Proof. Suppose by way of contradiction otherwise, as witnessed by Θ and Ψ. Note that, by Proposition 9, $\Theta(RINIT_{0.2})$ cannot be a finite set. Suppose $\Theta(RINIT_{0.2}) = INIT_k$.

Then for two different values of $r < 0.2$, $\Theta(RINIT_r) = INIT_i$, for same i, as for all $r < 0.2$, $\Theta(RINIT_r) \subseteq INIT_k$. A contradiction. ∎

Theorem 22. **iRINIT** $\leq_{\mathbf{Part}}^{Weak}$ **iINIT**.

Proof. Let $r_S = \max(\{\mathrm{ntor}(x) : \langle x, y \rangle \in S\})$. Define $m_{T[n]}$ as follows.

(i) $m_\Lambda = \langle \mathrm{rton}(0), 0 \rangle$.
(ii) $m_{T[n+1]} = m_{T[n]}$, if $r_{\mathrm{content}(T[n+1])} = r_{\mathrm{content}(T[n])}$; otherwise $m_{T[n+1]} = \langle \mathrm{rton}(r_{\mathrm{content}(T[n+1])}), m_{T[n]} + 1 \rangle$.

Now, let $\Theta(T[n]) = \{\langle x, y\rangle : x \le m_{T[n]}, y \le \text{code}(\text{content}(T[n]))\}$.
It is easy to verify that,

(1) For a finite set L, $\Theta(L) \subseteq \{\{\langle x, y\rangle : x \le i, y \le \text{code}(L)\} : i \in N\}$.
(2) $\Theta(RINIT_r) \subseteq \{INIT_{\langle \text{rton}(r), w\rangle} : w \in N\}$.

We can define the operator Ψ for the reduction using Lemma 5, where F_1 and F_2 are defined as follows.

$F_1(S) = $ canonical grammar for D_w, where $w = \max(\{y : \langle 0, y\rangle \in S\})$.

$F_2(p, t) = $ canonical grammar for $RINIT_{\text{ntor}(j)}$, where, for some w, $\langle j, w\rangle = \max(\{x : \langle x, y\rangle \in W_{p,t}\})$.

It is now easy to verify using Lemma 5 that Θ and Ψ (as given by Lemma 5) witness that **iRINIT** $\le^{Weak}_{\mathbf{Part}}$ **iINIT**. ∎

Theorem 23. **iRINIT** $\not\le^{Weak}_{\mathbf{Part}}$ **iCOINIT**.

The next result shows that **iRINIT** is strongly reducible to $\mathcal{REG}$, the class of all regular languages. As we noted above, $\mathcal{REG}$ is not reducible to **iRINIT** (even weakly), thus, the degree of **iRINIT** is strictly below the degree of $\mathcal{REG}$.

Theorem 24. **iRINIT** $\le^{Strong}_{\mathbf{Part}}$ $\mathcal{REG}$.

Now we turn our attention to the classes **iINIT**$_k$ and **iCOINIT**$_k$. The infinite languages in these classes do not form simple strict chains, as, for every infinite language L in the chain, both classes contain its variants having up to k extra elements. Interestingly, though, it turns out that, whereas adding one such extra element to infinite languages in the chain makes the partial learning problem harder, the difficulty of the partial learning problem does not increase when more elements are added.

Theorem 25. *For all $k > 0$,*
    (a) **iINIT**$_k$ $\le^{Strong}_{\mathbf{Part}}$ **iINIT**$_1$.
    (b) **iCOINIT**$_k$ $\le^{Strong}_{\mathbf{Part}}$ **iCOINIT**$_1$.
    (c) **iINIT**$_1$ $\not\le^{Strong}_{\mathbf{Part}}$ **iINIT**.
    (d) **iCOINIT**$_1$ $\not\le^{Weak}_{\mathbf{Part}}$ **iCOINIT**.
    (e) **iINIT**$_1$ $\le^{Weak}_{\mathbf{Part}}$ **iINIT**.

Theorem 26. *For all $k > 0$,*
    (a) **iINIT**$_k$ $\le^{Strong}_{\mathbf{Part}}$ **iRINIT**.
    (b) **iCOINIT**$_k$ $\le^{Strong}_{\mathbf{Part}}$ **iRINIT**.
    (c) **iRINIT** $\not\le^{Strong}_{\mathbf{Part}}$ **iINIT**$_k$.
    (d) **iRINIT** $\not\le^{Weak}_{\mathbf{Part}}$ **iCOINIT**$_k$.

References

1. D. Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.
2. D. Angluin. Learning regular sets from queries and counter-examples. *Information and Computation*, 75:87–106, 1987.
3. M. Blum. A machine-independent theory of the complexity of recursive functions. *Journal of the ACM*, 14(2):322–336, 1967.
4. J. Case and T. Kötzing. Computability-theoretic learning complexity. *Phil. Trans. R. Soc. of London*, 370:3570–3596, 2011.
5. R. Freivalds, E. Kinber, and C. Smith. On the intrinsic complexity of learning. *Information and Computation*, 123(1):64–71, 1995.
6. Z. Gao, S. Jain, and F. Stephan. On conservative learning of recursively enumerable languages. In P. Bonizzoni, V. Brattka, and B. Löwe, editors, *9th Conference on Computability in Europe: The Nature of Computation, Logic, Algorithms, Applications*, volume 7921 of *Lecture Notes in Computer Science*, pages 181–190. Springer-Verlag, 2013.
7. Z. Gao, S. Jain, F. Stephan, and S. Zilles. A survey on recent results on partial learning. In *Proceedings of the Thirteenth Asian Logic Conference*, pages 68–92. World Scientific, 2015.
8. E. M. Gold. Language identification in the limit. *Information and Control*, 10(5):447–474, 1967.
9. Z. Gao and F. Stephan. Confident and consistent partial learning of recursive functions. In Nader Bshouty, Gilles Stoltz, Nicolas Vayatis, and Thomas Zeugmann, editors, *Algorithmic Learning Theory: 23rd International Conference (ALT' 2012)*, volume 7568 of *Lecture Notes in Artificial Intelligence*, pages 51–65. Springer-Verlag, 2012.
10. Z. Gao, F. Stephan, and S. Zilles. Combining models of approximation with partial learning. In Kamalika Chaudhuri, Claudio Gentile, and Sandra Zilles, editors, *Algorithmic Learning Theory: 26th International Conference (ALT' 2015)*, volume 9355 of *Lecture Notes in Artificial Intelligence*, pages 56–70. Springer-Verlag, 2015.
11. J. Hopcroft and J. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, 1979.
12. S. Jain, E. Kinber, C. Papazian, C. Smith, and R. Wiehagen. On the intrinsic complexity of learning recursive functions. *Information and Computation*, 184(1):45–70, July 2003.
13. S. Jain, E. Kinber, and R. Wiehagen. Language learning from texts: Degrees of intrinsic complexity and their characterizations. *Journal of Computer and System Sciences*, 63:305–354, 2001.
14. S. Jain and A. Sharma. The intrinsic complexity of language identification. *Journal of Computer and System Sciences*, 52:393–402, 1996.
15. S. Jain and F. Stephan. Consistent partial learning. In *Proceedings of the Twenty Second Annual Conference on Learning Theory*, 2009.
16. D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, 1986.
17. H. Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, 1967. Reprinted by MIT Press in 1987.