Variants of Partial Learning in Inductive Inference

GAO ZIYUAN

(B.Sc (Hons.), NUS)

Supervisor: Professor Frank STEPHAN

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

National University of Singapore
Department of Mathematics

2012

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

Gao Ziyuan

26 April 2012

Acknowledgements

I would like to thank my supervisor Professor Frank Stephan for introducing me to Inductive Inference, suggesting many interesting problems to work on, and giving me the opportunity to be both his student and coauthor. I am grateful to him for the multitude of ideas he taught and inspired me with during our weekly discussion weekings, his advice on how to conduct independent research as well on numerous other practical issues such as career and scholarship choices, his regular feedback and suggestions for improvements in both the style and mathematical content of this thesis as it was being written, and his kind permission for me to present our joint paper at LATA 2012.

I would like to thank my family for their invaluable support throughout my academic experience, allowing me to work on this thesis with a calmness and peace of mind. I am grateful to them for always supporting and encouraging me to pursue my interests.

I would like to thank my friends for their kind words of encouragement and emotional support; after our regular meetings, I could always continue work on this thesis with a renewed sense of vigour and energy.

Contents

1 Summary

This thesis studies several variants of partial learning under the framework of inductive inference. In particular, the following learning criteria are examined: confident partial learning, partial conservative learning, essentially class consistent partial learning, and iterative learning. Consistent partial learning of recursive functions is classified according to the mode of data presentation; the two main types of data texts considered are canonical text and arbitrary text. The issue of consistent partial learning from incomplete texts is also given a brief treatment towards the end of the report. A further research direction taken up in this report is the investigation of the additional learning power conferred by oracles. It is shown that certain conditions on the computational degrees of oracles enable all recursive functions to be confidently partially learnt. Similarly, it is proved that all PA-complete oracles are computationally strong enough to permit the essentially consistent inference of all recursive functions. Another question particularly relevant in the effort to construct class separation examples of various learning criteria is whether there is always a uniform effective procedure to find a recursive function that is not learnt by a learner according to some criterion. The present work tries to address this question for the case of confident partial learning and consistent partial learning.

2 Introduction

This project has grown out of an attempt to systematically characterize the nature of partial learning, a generalisation of the traditional models of learning in inductive inference. Whilst the usual criteria of learning success, such as explanatory and behaviourally correct learning, do permit a large class of languages to be learnt, there are many natural examples that fail to be identifiable in the limit, even in the broadest sense of semantic convergence. The reasons for their unlearnability are not due to any lack of computational ability of the learner; indeed, even with the additional learning power conferred by any oracle, there is no recursive learner that can always converge in the limit to a correct guess on a text for any member set in the class of all finite sets plus one infinite set. The problem is due to a mix of factors. One reason is the structural nature of the class of languages; another reason may be that the learning success requirements imposed are too stringent. To enrich the classes of languages that are, in some tenable sense, learnable, one may attempt to loosen the restrictions for learning success. Various approaches devoted to this aim can be found in the inductive inference literature. Feldman [6], for example, showed that a decidable rewriting system (drs) is always learnable from positive information sequences in a certain restricted sense. Partial learning is another such proposal to overcome the deficiency of learning in the limit. Unfortunately, it has already been noted by Osherson, Stob and Weinstein [24] that the class of all r.e. sets is partially learnable. Similarly, the class of all co-r.e. sets is also partially learnable. In order to capture a more balanced sense of partial learnability, one may therefore require a careful calibration of learning success requirements, such as may be obtained by imposing additional learning contraints.

This work is organized into two main sections: the partial learnability of r.e. and co-r.e. languages, and the partial learnability of recursive functions. Confidence is shown to be a fairly strong restriction on partial learnability: even the class of all cofinite sets is not confidently partially learnable; neither is the class consisting of the unions of all finite sets with any nonrecursive set. This observation also extends to the learning of recursive functions, as may be noted from the fact that even behaviourally correct learnability is insufficient to guarantee confident partial learnability in this case. Furthermore, several theorems illuminate the role that padding, an occasionally useful tool in Recursion Theory, plays in the construction of confident partial learners. In particular, one result states that vacillatory learnability (whereby a learner is permitted to oscillate infinitely often between finitely many different correct indices) implies confident partial learnability when the hypothesis space is taken to be the standard universal numbering of all r.e. languages, or that of all partial-recursive functions. Since padding is a technique dependent on the nature of the numbering with respect to which a learner specifies its conjecture, it may be natural to inquire how the results on confident partial learnability vary with the choice of a learner's hypothesis space. To shed some light on this question, we construct an example of a uniformly r.e. class of languages which is vacillatorily learnable but not confidently partially learnable with respect to the given class numbering. It is, however, still possible to recover from this negative result a weaker connection between the two forms of learning: a later theorem demonstrates that, with respect to any general uniformly r.e. hypothesis space of languages, explanatory learnability implies confident partial learnability.

A further theme studied in this work is the additional learning power conferred by oracles. We study this problem from the viewpoints of both confident and consistent partial learnability. We suggest certain sufficient conditions on the computational degrees of oracles that permit the confident partial learnability of all recursive functions. Conversely, various necessary conditions on the computational degrees of oracles relative to which REC is confidently partially learnable are proposed. A weaker version of consistent partial learnability - essentially consistent partial learnability, according to which a learner must be consistent on cofinitely many data inputs - is introduced. It is shown that all PA-complete oracles are strong enough to allow all recursive functions to be essentially consistently partially learnable. This theorem may be viewed in contrast with the results obtained in [13], in which the authors fully characterise the computational degrees of oracles relative to which REC is consistently partially learnable. We conclude the section on consistent partial learning of recursive functions by considering a scenario in which the the learner has to infer recursive extensions of functions presented as incomplete texts. The final section deals with the notion of iterative learning, also known as memory-limited learning. In this setting, a learner has to base its conjecture only on the current input data and its last hypothesis. The requirements of iterative function learning appear to be quite exacting: it is shown that there are explanatorily learnable classes of recursive functions which are not iteratively learnable.

2.1 Notation

The set of natural numbers is denoted by $\mathbb{N}$, that is, $\mathbb{N} = \{0, 1, 2, \ldots\}$. All "numbers" in this project refer to natural numbers. The abbreviation r.e. shall be used for the term "recursively enumerable." A universal numbering of all partial-recursive functions is fixed as $\varphi_0, \varphi_1, \varphi_2, \ldots$. Given a set S, $\overline{S}$ denotes the complement of S, and S^* denotes the set of all finite sequences in S. Let $W_0, W_1, W_2, \ldots$ be a universal numbering of all r.e. sets, where W_e is the domain of φ_e. $\langle x, y\rangle$ denotes Cantor's pairing function, given by $\langle x, y\rangle = \frac{1}{2}(x+y)(x+y+1)+y$. $W_{e,s}$ is an approximation to W_e; without loss of generality, $W_{e,s} \subseteq \{0, 1, \ldots, s\}$, and $\{\langle e, x, s\rangle : x \in W_{e,s}\}$ is primitive recursive. $\varphi_e(x)\uparrow$ means that $\varphi_e(x)$ remains undefined; $\varphi_{e,s}(x)\downarrow$ means that $\varphi_e(x)$ is defined, and that the computation of $\varphi_e(x)$ halts within s steps. $\mathbb{K}$ denotes the diagonal halting problem. The jump of a set A is denoted by A'; that is, $A' = \{e : \varphi_e^A(e)\downarrow\}$. For any two sets A and B, $A \oplus B = \{2x : x \in A\} \cup \{2y+1 : y \in B\}$. Analogously, $A \oplus B \oplus C = \{3x : x \in A\} \cup \{3y+1 : y \in B\} \cup \{3z+2 : z \in C\}$. The class of all recursive functions is denoted by REC; the class of all $\{0,1\}$-valued recursive functions is denoted by $REC_{0,1}$. For any two partial-recursive functions f and g, $f =_* g$ denotes that for cofinitely many x, $f(x)\downarrow = g(x)\downarrow$.

For any $\sigma, \tau \in (\mathbb{N} \cup \{\#\})^*, \sigma \preceq \tau$ if and only if $\sigma = \tau$ or τ is an extension of σ, $\sigma \prec \tau$ if and only if σ is a proper prefix of τ, and $\sigma(n)$ denotes the element in the nth position of σ, starting from $n = 0$. Given a number a and some fixed $n \geq 1$, denote by a^n the finite sequence $a \ldots a$, where a occurs n times. a^0 denotes the empty string. The concatenation of two strings σ and τ shall be denoted by $\sigma\tau$, and occasionally by $\sigma \circ \tau$.

2.2 Definitions

The main references on Recursion Theory consulted over the course of this project were [23], [25], and [27]. The notions of *partial-recursive functions* and *recursively enumerable sets* form the theoretical backbone of the present work. These are defined formally as follows.

Definition 1 The class of *partial-recursive functions* is the smallest class $\mathcal{C}$ of functions from $\mathbb{N}^n$ (with parameter $n \in \mathbb{N}$) to $\mathbb{N}$ such that

- The function mapping any input in $\mathbb{N}^n$ to some constant m is in $\mathcal{C}$;
- The successor function S given by $S(x) = x + 1$ is in $\mathcal{C}$;
- For every n and every $m \in \{1, 2, \ldots, n\}$, the function mapping $(x_1, x_2, \ldots, x_n)$ to x_m is in $\mathcal{C}$;
- For any functions $f : \mathbb{N}^n \to \mathbb{N}$ and $g_1, \ldots, g_n : \mathbb{N}^m \to \mathbb{N}$ in $\mathcal{C}$, the function mapping $(x_1, x_2, \ldots, x_m)$ to $f(g_1(x_1, x_2, \ldots, x_m), g_2(x_1, x_2, \ldots, x_m), \ldots, g_n(x_1, x_2, \ldots, x_m))$ is in $\mathcal{C}$;
- If $g : \mathbb{N}^{n+2} \to \mathbb{N}$ and $h : \mathbb{N}^n \to \mathbb{N}$ are functions in $\mathcal{C}$, then there is a function $f : \mathbb{N}^{n+1} \to \mathbb{N}$ in $\mathcal{C}$ with $f(x_1, x_2, \ldots, x_n, 0) = h(x_1, x_2, \ldots, x_n)$ and $f(x_1, x_2, \ldots, x_n, S(x_{n+1})) = g(x_1, x_2, \ldots, x_n, x_{n+1}, f(x_1, x_2, \ldots, x_n, x_{n+1}))$;
- If $f : \mathbb{N}^{n+1} \to \mathbb{N}$ is a function in $\mathcal{C}$, the function $\mu y(f(x_1, \ldots, x_n, y) = 0)$, which takes the value z if $f(x_1, \ldots, x_n, y)$ is defined for all $y \leq z$ and $f(x_1, \ldots, x_n, y) > 0$ for $y < z$ and $f(x_1, \ldots, x_n, z) = 0$, and is undefined if no such z can be found, is in $\mathcal{C}$.

Definition 2 A function is *recursive* if it is defined on the whole domain $\mathbb{N}^n$ and *partial-recursive*. A set A is *recursively enumerable* if it is the range of a partial-recursive function. A set A is *recursive* if there is a recursive function f with $f(x) = 1$ for $x \in A$ and $f(x) = 0$ for $x \notin A$. A set A is *1-generic* if for all recursively enumerable sets $B \subseteq \{0,1\}^*$ there exists an n such that either $A(0) \circ A(1) \circ \ldots \circ A(n) \in B$ or no extension of $A(0) \circ A(1) \circ \ldots \circ A(n)$ belongs to B. More generally, a set A is *n-generic* if for every Σ^0_n set $W \subseteq \{0,1\}^*$ there is an m such that either $A(0) \circ A(1) \circ \ldots \circ A(n) \in W$ or no extension of $A(0) \circ A(1) \circ \ldots \circ A(n)$ belongs to W.

Remark 3 The abbreviation r.e. shall be used for the term "recursively enumerable." Given a partial-recursive function φ_e, one can simulate the computation of $\varphi_e(x)$ for a number s of computation steps. Then $\varphi_{e,s}(x)$ is defined if the computation halts within s steps; otherwise $\varphi_{e,s}(x)$ is undefined. Similarly, given a recursively enumerable set A, one can simulate the enumeration process of A for s computation steps, and denote by A_s the set all of elements of A that are enumerated within s steps.

Depending on the context, a *numbering* is either a uniformly r.e. family $\{L_i\}_{i \in \mathbb{N}}$ of subsets of $\mathbb{N}$, or a uniformly co-r.e. family $\{L_i\}_{i \in \mathbb{N}}$ of subsets of $\mathbb{N}$, or a family $\{\phi_i\}_{i \in \mathbb{N}}$ of partial-recursive functions from $\mathbb{N}$ to $\mathbb{N}$ such that $\langle i, x \rangle \to \phi_i(x)$ is partial-recursive. We shall fix a universal numbering $\varphi_0, \varphi_1, \varphi_2, \ldots$ of all partial-recursive functions, and a universal numbering $W_0, W_1, W_2, \ldots$ of all r.e. sets, where W_e is the domain of φ_e. By means of Cantor's pairing function, strings over a countable alphabet can be coded as natural numbers; for mathematical convenience, this work usually regards a class of languages as a set of natural numbers. $\mathbb{K}$, the diagonal halting problem, denotes the set $\{e : e \in W_e\}$, which is also equal to $\{e : \varphi_e(e) \text{ is defined}\}$.

Definition 4 Let $\mathcal{C}$ be a class of recursive, recursively enumerable, or co-recursively enumerable sets. A *text* T_L for some L in $\mathcal{C}$ is a map $T_L : \mathbb{N} \to L \cup \{\#\}$ such that $\text{range}(T_L) = L$. $T_L[n]$ denotes the string $T_L(0) \circ T_L(1) \circ \ldots \circ T_L(n)$. A learner is a recursive function $M : (\mathbb{N} \cup \{\#\})^* \to \mathbb{N}$. The main learning criterion studied in the report is *partial learning*; this notion, together with various learning constraints and other learning success criteria, are defined as follows.

i. M is said to *partially* learn $\mathcal{C}$ if, for each L in $\mathcal{C}$, and any corresponding text T_L for L, there is exactly one index e such that $M(T_L[k]) = e$ for infinitely many k, and this e satisfies $L = W_e$.

ii. M is said to *explanatorily* (Ex) learn $\mathcal{C}$ if, for each L in $\mathcal{C}$, and any corresponding text T_L for L, there is a number n for which $L = W_{M(T_L[j])}$ whenever $j \geq n$, and for any $k \geq j$, $M(T_L[k]) = M(T_L[j])$.

iii. M is said to *behaviourally correctly* (BC) learn $\mathcal{C}$ if, for each L in $\mathcal{C}$, and any corresponding text T_L for L, there is a number n for which $L = W_{M(T_L[j])}$ whenever $j \geq n$.

iv. M is said to *vacillatorily* (Vac) learn $\mathcal{C}$ if it BC learns $\mathcal{C}$ and outputs on every text T_L for each L in $\mathcal{C}$ only finitely many different indices.

v. M is said to *partially conservatively* learn $\mathcal{C}$ if it partially learns $\mathcal{C}$ and outputs on every text T_L for each L in $\mathcal{C}$ exactly one index e with $L \subseteq W_e$.

vi. M is said to *confidently partially* learn $\mathcal{C}$ if it partially learns $\mathcal{C}$ and, for every set L and every text T_L for L, outputs on T_L exactly one index infinitely often.

Definition 5 The definitions for learning of recursive functions proceed in parallel fashion; here we distinguish between learning from *canonical texts* and *arbitrary texts*. Let $\mathcal{C}$ be a class of recursive functions. The *canonical text* T_f^{can} for some f in $\mathcal{C}$ is the map $T_f^{can} : \mathbb{N} \to \mathbb{N}$ such that $T_f^{can}(n) = f(n)$ for all n. $T_f^{can}[n]$ denotes the string $T_f^{can}(0) \circ T_f^{can}(1) \circ \ldots \circ T_f^{can}(n)$. An *arbitrary text* T_f for some f in $\mathcal{C}$ is a map $T_f : \mathbb{N} \to \text{graph}(f)$ such that $T_f(\mathbb{N}) = \text{graph}(f)$. $T_f[n]$ denotes the string $T_f(0) \circ T_f(1) \circ \ldots \circ T_f(n)$. In contrast to canonical texts, the pairs $\langle x, f(x) \rangle$ in graph(f) may appear in any order. The learning success criteria are first defined with respect to learning from canonical texts.

i. M is said to *partially* ($Part^{can}$) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, there is exactly one index e such that $M(T_f^{can}[k]) = e$ for infinitely many k, and this e satisfies $f = \varphi_e$.

ii. M is said to *explanatorily* (Ex^{can}) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, there is a number n for which $f = \varphi_{M(T_f^{can}[j])}$ whenever $j \geq n$, and for any $k \geq j$, $M(T_f^{can}[k]) = M(T_f^{can}[j])$.

iii. M is said to *behaviourally correctly* (BC^{can}) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, there is a number n for which $f = \varphi_{M(T_f^{can}[j])}$ whenever $j \geq n$.

iv. M is said to *vacillatorily* (Vac^{can}) learn $\mathcal{C}$ if it BC^{can} learns $\mathcal{C}$ and outputs on the canonical text for each f in $\mathcal{C}$ only finitely many different indices.

v. M is said to *confidently partially* ($ConfPart^{can}$) learn $\mathcal{C}$ if it partially learns $\mathcal{C}$

from canonical text and outputs on every infinite sequence exactly one index infinitely often.

vi. M is said to *essentially class consistently partially* ($EssClassConsPart^{can}$) learn $\mathcal{C}$ if it partially learns $\mathcal{C}$ from canonical text and, for each f in $\mathcal{C}$, $\varphi_{M(T_f^{can}[n])}(m)\downarrow= f(m)$ holds whenever $m \leq n$ for cofinitely many n.

The analagous learning criteria defined in the context of identification with respect to arbitrary text are as follows.

i. M is said to *partially* ($Part^{arb}$) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, and any corresponding text T_f for f, there is exactly one index e such that $M(T_f[k]) = e$ for infinitely many k, and this e satisfies $f = \varphi_e$.

ii. M is said to *explanatorily* (Ex^{arb}) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, and any corresponding text T_f for f, there is a number n for which $f = \varphi_{M(T_f[j])}$ whenever $j \geq n$, and for any $k \geq j$, $M(T_f[k]) = M(T_f[j])$.

iii. M is said to *behaviourally correctly* (BC^{arb}) learn $\mathcal{C}$ if, for each f in $\mathcal{C}$, and any corresponding text T_f for f, there is a number n for which $f = \varphi_{M(T_f[j])}$ whenever $j \geq n$.

iv. M is said to *vacillatorily* (Vac^{arb}) learn $\mathcal{C}$ if it BC^{arb} learns $\mathcal{C}$ and outputs on every text T_f for each f in $\mathcal{C}$ only finitely many different indices.

v. M is said to *confidently partially* ($ConfPart^{arb}$) learn $\mathcal{C}$ if it $Part^{arb}$ learns $\mathcal{C}$ and outputs on every infinite sequence exactly one index infinitely often.

vi. M is said to *essentially class consistently partially* ($EssClassConsPart^{arb}$) learn $\mathcal{C}$ if it $Part^{arb}$ learns $\mathcal{C}$ and, for each f in $\mathcal{C}$, and any corresponding text T_f for f, $\varphi_{M(T_f[n])}(m) \downarrow = f(m)$ holds whenever $\langle m, f(m) \rangle \in \{T_f(k) : k \leq n\}$ for cofinitely many n.

On occasion, the present work also studies the question of partial learnability under the setting of any general hypothesis space. The learning success criteria are extended in a natural way; the subsequent definition carries out this generalisation for confident partial learning.

Definition 6 Let $\mathcal{L} = \{A_0, A_1, A_2, \ldots\}$ be a uniformly recursively enumerable family, and let $\mathcal{H} = \{B_0, B_1, B_2, \ldots\} \supseteq \mathcal{L}$. $\mathcal{L}$ is said to be confidently partially learnable using the hypothesis space $\mathcal{H}$ if there is a confident partial recursive learner M such that for all A_i, M outputs on a text for A_i exactly one index j infinitely often and j satisfies $B_j = A_i$.

Blum and Blum [3] introduced the notion of a *locking sequence* for explanatory learning, whose existence is a necessary criterion for a learner to successfully identify the language or recursive function generating the text seen. With a slight modification, one can adapt this concept to the partial learning model.

Definition 7 Let M be a recursive learner and $\mathcal{L}$ be a set partially learnt by M. Then there is a finite sequence σ of elements in $\mathcal{L} \cup \{\#\}$ such that

- $W_{M(\sigma)} = \mathcal{L}$;
- For all finite sequences τ of elements in $\mathcal{L} \cup \{\#\}$, there is an $\eta \in (L \cup \{\#\})^*$ such that $M(\sigma \circ \tau \circ \eta) = M(\sigma)$.

This σ shall be called a *locking sequence* for $\mathcal{L}$.

2.3 Tools from Recursion Theory

The present section summarises the results in Recursion Theory that are most frequently applied in the following work.

Theorem 8 *(Substitution theorem, or s-m-n theorem) For all m, n, a partial function $f(e_1, \ldots, e_m, x_1, \ldots, x_n)$ is partial recursive if and only if there is a recursive function g such that*

$$\forall e_1, \ldots, e_m, x_1, \ldots, x_n [f(e_1, \ldots, e_m, x_1, \ldots, x_n) = \varphi_{g(e_1,\ldots,e_m)}(\langle x_1, \ldots, x_n \rangle)].$$

Theorem 9 *(Padding lemma) There is a recursive function pad satisfying $\varphi_{pad(e)} = \varphi_e$, and $pad(e) > e$ for all e.*

Theorem 10 *(Kleene's second recursion theorem, or fixed-point theorem) Given any recursive function f, there are infinitely many e with $\varphi_{f(e)} = \varphi_e$.*

3 Partial Learning of Classes of R.e. Languages

The point of departure is the following result noted by Osherson, Stob and Weinstein [24], that the class of all r.e. sets is partially learnable. The proof can be extended to show that the class of all co-r.e. sets is also partially learnable, as is the class of all recursive functions. This theorem motivates the search for a more restrictive criterion of partial learning.

Theorem 11 *The class of all r.e. sets is partially learnable.*

Proof. Let $F_0, F_1, F_2, \ldots$ be a Friedberg numbering of all r.e. sets. One can define a recursive learner M that outputs, on any text $T(0) \circ T(1) \circ T(2) \circ \ldots$, an index e at least n times if and only if there is a stage $s > n$ such that $F_{e,s}(x) = T_s(x)$ for all $x \leq n$, where $T_s = \{T(0), T(1), \ldots, T(s)\} - \{\#\}$. By the s-m-n theorem, there is a recursive function g such that $F_d = W_{g(d)}$ for all d. A new recursive learner N can subsequently be defined to translate the indices output by M into indices from the default hypothesis space $\{W_0, W_1, W_2, \ldots\}$, by setting N to conjecture $g(e)$ just if M outputs e. The one-one numbering property of $F_0, F_1, F_2, \ldots$ implies that if T were the text for some r.e. language L, then there is exactly one index e satisfying $\forall x \leq n[F_e(x) = T_s(x)]$ for infinitely many n and s. This establishes that N is a partial learner of all r.e. languages, as required. ♦

3.1 Confident Partial Learning

The first learning constraint proposed here as a means of sharpening partial learnability is that of *confidence*. This notion is mentioned peripherally in [12] and [22], appearing within exercises in the textbooks cited. As defined earlier, a recursive learner is *confident* just if it outputs on each text for every set L exactly one index infinitely often. The next result, that the class of all cofinite sets is not confidently partially learnable, is proved in [9], and it shows that this additional learning requirement does in fact restrict the scope of partial learnability.

Theorem 12 *[9] The class of all cofinite sets is not confidently partially learnable.*

To bridge the gap between partial learning and the more traditional learning success criteria of explanatory and behaviourally correct learning, it is shown next that one can also construct a behaviourally correctly learnable class of r.e. languages which is not confidently partially learnable.

Theorem 13 *There is a uniformly r.e. class of languages which is behaviourally correctly learnable but not confidently partially learnable.*

Proof 1. Let $\mathcal{C}$ be the class $\{\{e\} \oplus (W_e \cup D) : e \in \mathbb{N} \wedge D \text{ is a finite set}\}$. A behaviourally correct learner for $\mathcal{C}$ may be defined as follows: on reading the input σ with $|\sigma| = n+1$ and $\mathrm{range}(\sigma) = \{2e\} \cup \{2x_1+1, 2x_2+1, \ldots, 2x_k+1\}$, M conjectures an r.e. index for the set $\{e\} \oplus (W_e \cup \{x_0, x_1, \ldots, x_k\})$; otherwise, M outputs a default index 0. For any given set $\{e\} \oplus (W_e \cup D)$ in $\mathcal{C}$, every text for this set must eventually contain the number $2e$ as well as the set $\{2y+1 : y \in D\}$. Consequently, M will always converge semantically to an index of the set to be learnt.

Next, assume by way of contradiction that N confidently partially learns $\mathcal{C}$. Fix any number e such that W_e is coinfinite, and using the oracle $\mathbb{K}'$, choose a subsequence $a_0, a_1, a_2, \ldots$ of $\mathbb{N} - W_e$ which satisfies the following two properties for all n:

- $a_{n+1} > a_n$;
- $a_{n+1} > \varphi_s^{\mathbb{K}}(a_0, a_1, \ldots, a_n)$,

for all $s \leq n$ such that $\varphi_s^{\mathbb{K}}(a_0, a_1, \ldots, a_n)$ is defined.

Put $L = \{e\} \oplus (\mathbb{N} - \{a_0, a_1, a_2, \ldots\})$. By the confidence of N, there is an index d and a finite sequence $\sigma \in (L \cup \{\#\})^*$ such that for all $\tau \in (L \cup \{\#\})^*$, there is an $\eta \in (L \cup \{\#\})^*$ such that $N(\sigma \circ \tau \circ \eta) = d$.

Claim 14 *There is a number n such that for all $k > n$, there is a $\tau_k \in (\{e\} \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_k\}))^*$ for which, given any $\gamma \in (\{e\} \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_k\}))^*$, there exists some $\eta \in (\{e\} \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_k\}))^*$ with $N(\sigma \circ \tau_k \circ \gamma \circ \eta) = d$.*

There is a partial $\mathbb{K}$-recursive function which evaluates the maximum value of any sequence $\tau_k \in (\{e\} \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_k\}))^*$ such that for all $\eta \in (\{e\} \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_k\}))^*$, it holds that $N(\sigma \circ \tau_k \circ \eta) \neq d$, if such a sequence τ_k does in fact exist. Let $\varphi_s^{\mathbb{K}}(a_0, a_1, \ldots, a_k)$ be this value whenever it is defined; by the choice of a_{k+1}, one has that $a_{k+1} > \varphi_s^{\mathbb{K}}(a_0, a_1, \ldots, a_k)$ for all $k \geq s$. As a consequence, for all $n \geq s$, τ_n cannot exist, for otherwise $\tau_n \in (L \cup \{\#\})^*$, and so by the locking property of σ, there is a sequence $\eta \in (L \cup \{\#\})^*$ for which $N(\sigma \circ \tau_n \circ \eta) = d$, contrary to the definition of τ_n. This establishes the claim.

Hence by the claim, there are at least two different finite sets F and G, for example $\{a_0, a_1, \ldots, a_s\}$ and $\{a_0, a_1, \ldots, a_{s+1}\}$, both of which are disjoint to W_e, and two strings $\sigma_F \in (\{e\} \oplus (\mathbb{N} - F))^*, \sigma_G \in (\{e\} \oplus (\mathbb{N} - G))^*$, as well as an index

d, such that for every $\tau_F \in (\{e\} \oplus (\mathbb{N} - F))^*$ and for every $\tau_G \in (\{e\} \oplus (\mathbb{N} - G))^*$ there is an $\eta_F \in (\{e\} \oplus (\mathbb{N} - F))^*$ with $N(\sigma_F \circ \tau_F \circ \eta_F) = d$ and there is an $\eta_G \in (\{e\} \oplus (\mathbb{N} - G))^*$ with $N(\sigma_G \circ \tau_G \circ \eta_G) = d$.

If, on the other hand, W_e were cofinite, then for every finite set F disjoint to W_e, $\{e\} \oplus (\mathbb{N} - F)$ is equal to $\{e\} \oplus (W_e \cup H)$ for some finite subset H. Since N confidently partially learns the set $\{e\} \oplus (W_e \cup H)$, it outputs on every text for this set exactly one index of the set infinitely often, so that the finite sets F and G as constructed above cannot exist. Hence it would follow that $\{e : W_e \text{ is coinfinite}\}$ is Turing reducible to $\mathbb{K}'$; denoting by $D_0, D_1, D_2, \ldots$ a canonical numbering of all finite sets, this reducibility may be realised by the Σ^0_3 formula

$$\begin{aligned}
&e \in \{c : W_c \text{ is coinfinite}\} \Leftrightarrow \exists\langle d,i,j\rangle \exists\sigma_i \exists\sigma_j \forall s \forall\tau_i \forall\tau_j \exists\eta_i \exists\eta_j [(i \neq j \\
&\wedge (D_i \cup D_j) \cap W_{e,s} = \emptyset \wedge \sigma_i \circ \tau_i \in ((\{e\} \oplus (\mathbb{N} - D_i)) \cup \{\#\})^* \\
&\wedge \sigma_j \circ \tau_j \in ((\{e\} \oplus (\mathbb{N} - D_j)) \cup \{\#\})^*) \Rightarrow (\eta_i \in ((\{e\} \oplus (\mathbb{N} - D_i)) \cup \{\#\})^* \\
&\wedge \eta_j \in ((\{e\} \oplus (\mathbb{N} - D_j)) \cup \{\#\})^* \wedge N(\sigma_i \circ \tau_i \circ \eta_i) = d \wedge N(\sigma_j \circ \tau_j \circ \eta_j) = d)],
\end{aligned}$$

which contradicts the known fact that it is Π^0_3-complete. ♦

Proof 2. Let A be any r.e. but nonrecursive set. We shall show that the uniformly r.e. class $\mathcal{C} = \{A \cup D : D \text{ is finite}\}$ is behaviourally correctly learnable but not confidently partially learnable. As the argument is based on the nonrecursiveness of A, it may be assumed without any loss of generality that A is the diagonal halting problem $\mathbb{K}$. A behaviourally correct learner for $\mathcal{C}$ may be defined as follows: on reading the input $\sigma = a_0 \circ a_1 \circ \ldots \circ a_n$, the learner M outputs an r.e. index for $\mathbb{K} \cup \{a_0, a_1, \ldots, a_n\} - \{\#\}$. If $a_0 \circ a_1 \circ a_2 \circ \ldots$ were a text for the set $\mathbb{K} \cup D$, then there is a sufficiently long prefix $a_0 \circ a_1 \circ \ldots \circ a_n$ of the text such that $D \subseteq \{a_0, a_1, \ldots, a_n\} - \{\#\}$, and consequently M will converge semantically to an index for $\mathbb{K} \cup D$.

Next, it shall be demonstrated that $\mathcal{C}$ is not confidently partially learnable. Assume by way of contradiction that N were a confident partial learner of $\mathcal{C}$. A $\mathbb{K}'$-recursive text, together with a subsequence $\{x_0, x_1, x_2, \ldots\}$ of $\mathbb{N} - \mathbb{K}$, are constructed inductively as follows:

- Since N confidently partially learns $\mathcal{C}$, a locking sequence $\sigma_0 \in (\mathbb{K} \cup \{\#\})^*$ for $\mathbb{K}$ may be found using the oracle $\mathbb{K}'$. Furthermore, suppose that N outputs the index e_0 for $\mathbb{K}$ infinitely often; σ_0 may then be chosen so that for all $\tau \in (\mathbb{K} \cup \{\#\})^*$, $N(\sigma_0 \circ \tau) \geq e_0$. By again accessing the oracle $\mathbb{K}'$, a search is then run for a number $y \in \mathbb{N} - \mathbb{K}$ such that $N(\sigma_0 \circ y) \geq e_0$, and for all $\tau \in (\mathbb{K} \cup \{\#\})^*$, $N(\sigma_0 \circ y \circ \tau) \geq e_0$. Such a y must always exist: for, suppose on the contrary that for all $y \in \mathbb{N} - \mathbb{K}$, either $N(\sigma_0 \circ y) < e_0$ holds, or there is a string $\tau \in (\mathbb{K} \cup \{\#\})^*$ for which $N(\sigma_0 \circ y \circ \tau) < e_0$. By the choice of σ_0, $N(\sigma_0 \circ y) \geq e_0$ and $N(\sigma_0 \circ y \circ \tau) \geq e_0$ for all $y \in \mathbb{K}$ and $\tau \in (\mathbb{K} \cup \{\#\})^*$. Hence one obtains an effective decision procedure for determining whether or not any given number is contained in $\mathbb{K}$, via the condition $y \notin \mathbb{K} \Leftrightarrow N(\sigma_0 \circ y) < e_0 \vee \exists \tau \in (\mathbb{K} \cup \{\#\})^* [N(\sigma_0 \circ y \circ \tau) < e_0]$, which is a contradiction. Hence the search for such a y will eventually terminate successfully; now set $x_0 = y$.

- At stage $n+1$, suppose that $x_0, x_1, \ldots, x_n$, as well as $\sigma_0, \sigma_1, \ldots, \sigma_n$ have been selected. In addition, suppose that for all $k \leq n$, N outputs the index e_k for $\mathbb{K} \cup \{x_0, \ldots, x_{k-1}\}$ infinitely often after it is fed with the locking sequence $\sigma_0 \circ x_0 \circ \ldots \circ \sigma_k$. Assume as the inductive hypothesis that $N(\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \ldots \circ \sigma_n \circ x_n) \geq e_n$, and that for all $\tau \in (\mathbb{K} \cup \{\#\})^*$, $N(\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \ldots \circ \sigma_n \circ x_n \circ \tau) \geq$

e_n. As N confidently partially learns $\mathbb{K} \cup \{x_0, x_1, \ldots, x_n\}$, there is a string $\tau \in (\mathbb{K} \cup \{\#\})^*$ and an r.e. index $e_{n+1} > e_n$ for $\mathbb{K} \cup \{x_0, x_1, \ldots, x_n\}$ such that $N(\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \ldots \circ \sigma_n \circ x_n \circ \tau \circ \eta) \geq e_{n+1}$ for all $\eta \in (\mathbb{K} \cup \{x_0, x_1, \ldots, x_n\} - \{\#\})^*$. This string τ may be found using the oracle $\mathbb{K}'$; one then sets $\sigma_{n+1} = \tau$. By an argument analogous to that of the base step of the construction, one may consult the oracle $\mathbb{K}'$ to find a number $y \in \mathbb{N} - \mathbb{K} - \{x_0, x_1, \ldots, x_n\}$ so that $N(\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \ldots \circ \sigma_n \circ x_n \circ \sigma_{n+1} \circ y) \geq e_{n+1}$, and for all $\gamma \in (\mathbb{K} \cup \{\#\})^*$, it holds that $N(\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \ldots \circ \sigma_{n+1} \circ y \circ \gamma) \geq e_{n+1}$. Setting $x_{n+1} = y$, this completes the recursion step.

It follows from the above construction that $e_0, e_1, e_2, \ldots$ is a strictly monotone increasing sequence, so that for every number e, there is an n sufficiently large so that $N(\gamma) > e$ for all $\gamma \preceq \sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \sigma_2 \circ x_2 \circ \ldots$ with $|\gamma| > n$. This means that N does not output any index infinitely often on the text $\sigma_0 \circ x_0 \circ \sigma_1 \circ x_1 \circ \sigma_2 \circ x_2 \circ \ldots$, contradicting the hypothesis that N is a confident learner. ♦

In spite of the preceding negative examples, there may still be a fair abundance of confidently partially learnable classes of languages. As demonstrated in [9], the class of all closed sets of Noetherian $\mathbb{K}$-r.e. matroids is confidently partially learnable. Furthermore, Gold's example [10], consisting of all finite sets and one infinite set, provides a relatively natural instance of a confidently partially learnable but not behaviourally correctly learnable class of languages.

Example 15 *The class $\mathcal{C} = \{D : D$ is finite$\} \cup \{\mathbb{N}\}$ is confidently partially learnable but not behaviourally correct learnable.*

Proof. One can define a recursive learner M that outputs, on the input $\sigma = a_0 \circ a_1 \circ a_2 \circ \ldots \circ a_n$, a fixed index of $\mathbb{N}$ if $\mathrm{range}(\sigma) - \{\#\} \neq \{a_0, a_1, a_2, \ldots, a_n\} - \{\#\}$, and a canonical index for $\mathrm{range}(\sigma) - \{\#\}$ if $\mathrm{range}(\sigma) - \{\#\} = \{a_0, a_1, a_2, \ldots, a_n\} - \{\#\}$. M then outputs a fixed index for $\mathbb{N}$ infinitely often on any input text with an infinite range; otherwise, it will output a canonical index for the finite range of the text. Hence M confidently partially learns $\mathcal{C}$. On the other hand, it can be shown [10] that $\mathcal{C}$ cannot be behaviourally correctly learnt, even with the aid of oracles. ♦

With a little diligence, it is possible to show that even for a uniformly recursive class of languages, behaviourally correct learnability does not necessarily imply confident partial learnability. Such an example is exhibited in the proof of the next theorem.

Theorem 16 *There is a uniformly recursive class of languages which is behaviourally correctly learnable but not confidently partially learnable with respect to the hypothesis space* $\{W_0, W_1, W_2, \ldots\}$.

Proof. Let $M_0, M_1, M_2, \ldots$ be an enumeration of all partial-recursive learners. The primary objective is to build a $\mathbb{K}$-recursive sequence $a_0, a_1, a_2, \ldots$ such that if the sequence is finite and equal to σ, then the learner M_{a_0} fails to learn the language $L_{\langle \sigma\tau \rangle}$ for all extensions $\tau \in \mathbb{N}^*$ of σ; and if the sequence is infinite, then there are finite sequences $\sigma_0, \sigma_1, \sigma_2, \ldots$ such that for all i, $\sigma_i \in (L_{\langle a_0, \ldots, a_i, s \rangle} \cup \{\#\})^*$ for a sufficiently large number s, and $\sigma_0 \circ \sigma_1 \circ \sigma_2 \circ \ldots$ is a text on which M_{a_0} outputs each index only finitely often. For each finite sequence $\langle a_0, a_1, \ldots, a_n, s \rangle \in \mathbb{N}^*$, the recursive set $L_{\langle a_0, a_1, \ldots, a_n, s \rangle}$ is defined in an inductive fashion as follows.

First, define an auxiliary class of finite sets $A_{n,s}$ by

$$A_{n,s}(x) = \begin{cases} 0 & \text{if } x > 3n+1 \text{ or } x \equiv 0 (\mathrm{mod}\ 3) \text{ or } x \equiv 2 (\mathrm{mod}\ 3); \\ W_s(x) & \text{if } x \le 3n+1 \text{ and } x \equiv 1 (\mathrm{mod}\ 3). \end{cases}$$

The purpose of introducing the finite sets $\{A_{n,s}\}_{n,s\in\mathbb{N}}$ is to ensure that each of the sets $L_{\langle a_0,a_1,\ldots a_n,s\rangle}$ differs from all of $W_0, W_1, \ldots, W_n$; the construction achieves this when s is sufficiently large. Next, put

$$L_{\langle a_0,s\rangle} = \begin{cases} \{a_0,t\} \oplus ((\mathbb{N} - A_{0,s}) \cap \{0,1,\ldots,t\}) \oplus (\mathbb{N} \cap \{0,1,\ldots,t\}) & \text{if } t \text{ is the first step} \\ & \text{with } t > \max\{s, a_0\} \\ & \text{such that } A_{0,t}(1) \ne A_{0,s}(1); \\ \{a_0\} \oplus (\mathbb{N} - A_{0,s}) \oplus \mathbb{N} & \text{if } A_{0,s}(1) = W_{0,s}(1). \end{cases}$$

Further, let $L_{\langle a_0\rangle} = L_{\langle a_0,0\rangle}$. Now, given the sequence $\langle a_0, a_1, \ldots, a_n, s\rangle$ with $n \ge 1$, consider the following conditions:

- for each i with $0 \le i \le n$, $x \in A_{i,s}$ if and only if $x \in W_i \cap \{0,1,\ldots,n\}$;
- there are finite sequences $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$ such that
  $\sigma_0 \in ((\{a_0\} \oplus (\mathbb{N} - A_{0,s}) \oplus \mathbb{N}) \cup \{\#\})^*$ is the first string found at step $a_1 > a_0$ with $a_1 > \max(\mathrm{range}(\sigma_0))$, and for which, whenever $\tau \in ((\{a_0\} \oplus (\mathbb{N} - A_{0,s}) \oplus \mathbb{N}) \cup \{\#\})^*$, it holds that $M_{a_0}(\sigma_0 \circ \tau) > 0$; in addition, for each i with $1 \le i \le n-1$,
  $\sigma_i \in ((\{a_0\} \oplus (\mathbb{N} - A_{i,s}) \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_{i-1}\})) \cup \{\#\})^*$ is the first string found at step $a_{i+1} > a_i$ with $a_{i+1} > \max(\mathrm{range}(\sigma_0 \circ \sigma_1 \circ \ldots \circ \sigma_i))$, and for all $\tau \in ((\{a_0\} \oplus (\mathbb{N} - A_{i,s}) \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_{i-1}\})) \cup \{\#\})^*$, one also has that

$M_{a_0}(\sigma_0 \circ \sigma_1 \circ \ldots \circ \sigma_i \circ \tau) > i.$

If both of the above conditions are satisfied, set

$$L_{\langle a_0, a_1, \ldots, a_n, s\rangle} = \{a_0\} \oplus (\mathbb{N} - A_{n,s}) \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_{n-1}\}).$$

If, on the other hand, at least one of the above conditions is not satisfied, and $t > \max\{s, a_0\}$ is the first step at which a condition is breached, set

$$L_{\langle a_0, a_1, \ldots, a_n, s\rangle} = \{a_0, t\} \oplus ((\mathbb{N} - A_{n,s}) \cap \{0, 1, \ldots, t\}) \oplus ((\mathbb{N} - \{a_0, a_1, \ldots, a_{n-1}) \cap \{0, 1, \ldots, t\}).$$

The first coordinate of $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$ has a dual role: to encode the learner M_{a_0} to be diagonalised against, as well as to prevent a finite set $L_{\langle a_0, a'_1, \ldots, a'_{n'}, s'\rangle}$ from being a proper subset of $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$ if, for the sequence $\langle a_0, a_1, \ldots, a_n, s\rangle$, there are finite sequences $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$ found at stages $a_1, a_2, \ldots, a_n$ respectively satisfying the conditions described above, so that $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$ is infinite. The second coordinate secures that $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$ differs from $W_0, W_1, \ldots, W_n$ provided s is large enough, while the last coordinate encodes the steps $a_0, a_1, a_2, \ldots$ at which the sequences $\sigma_0, \sigma_1, \sigma_2, \ldots$ are found. It follows from the construction that $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$ is finite and has an element equal to 0 modulo 3 which is greater than a_0 if and only if at least one of the above conditions fails to hold. It remains to show that the uniformly recursive class $\mathcal{C} = \{L_{\langle a_0, a_1, \ldots, a_n, s\rangle}\}_{a_0, a_1, \ldots, a_n, s \in \mathbb{N}}$ is $BC_{r.e.}$ learnable but not confidently partially learnable.

By the known characterisation of $BC_{r.e.}$ learnable uniformly recursive families [2], it suffices to demonstrate that each set in the class contains a possibly non-effective tell-tale set - that is, corresponding to each $L_{\langle a_0, a_1, \ldots, a_n, s\rangle}$, there is a finite set $H_{\langle a_0,a_1,\ldots,a_n,s\rangle} \subseteq L_{\langle a_0,a_1,\ldots,a_n,s\rangle}$ such that all $L' \in C$ for which $H_{\langle a_0,a_1,\ldots,a_n,s\rangle} \subseteq L' \subseteq L_{\langle a_0,a_1,\ldots,a_n,s\rangle}$ holds must be equal to $L_{\langle a_0,a_1,\ldots,a_n,s\rangle}$. These tell-tale sets may be observed by means of a case distinction. To begin with, consider sets of the form $L_{\langle a_0,s\rangle}$; since all finite sets are tell-tale sets of themselves, it may be assumed that $L_{\langle a_0,s\rangle} = \{a_0\} \oplus (\mathbb{N} - A_{0,s}) \oplus \mathbb{N}$. Suppose that there are sequences $\sigma_0, \sigma_1, \sigma_2, \ldots, \sigma_n, \ldots$, found at steps $a_1, a_2, a_3, \ldots, a_n, \ldots$, respectively satisfying the requirements for $L_{\langle a_0,a_1,\ldots,a_n,s\rangle}$ to be an infinite set when s is sufficiently large. The sequences $\sigma_0, \sigma_1, \sigma_2, \ldots$, together with steps $a_1, a_2, a_3, \ldots$, if they exist, are uniquely determined. Consequently, a tell-tale set for $L_{\langle a_0,s\rangle}$ is $\{a_0\} \oplus \emptyset \oplus \{a_1\}$, as every finite set contains at least two elements in the first coordinate, and so cannot be a proper subset of $\{a_0\} \oplus (\mathbb{N} - A_{0,s}) \oplus \mathbb{N}$. By the same token, if there exist at least n terms in the sequence $a_1, a_2, a_3, \ldots$, and $L_{\langle a_0,a_1,\ldots,a_n,s\rangle} = \{a_0\} \oplus (\mathbb{N} - A_{n,s}) \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_{n-1}\})$, then a tell-tale set for $L_{\langle a_0,a_1,\ldots,a_n,s\rangle}$ is $\{a_0\} \oplus \emptyset \oplus \{a_n\}$. On the other hand, if there is no n-th term in the sequence, then a tell-tale set for $\{a_0\} \oplus (\mathbb{N} - A_{n,s}) \oplus (\mathbb{N} - \{a_0, a_1, \ldots, a_{n-1}\})$ is $\{a_0\} \oplus \emptyset \oplus \emptyset$. Thus by the non-effective version of Angluin's criterion, $\mathcal{C}$ is $BC_{r.e.}$ learnable.

To complete the proof, assume by way of contradiction that M_{a_0} were a confident partial learner of the class $\mathcal{C}$. Suppose that there is an infinite sequence of strings $\sigma_0, \sigma_1, \sigma_2, \ldots$ found at steps $a_1, a_2, a_2, \ldots$ respectively, which satisfy the condition that for all i, $\sigma_i \in (L_{\langle a_0,a_1,\ldots,a_i,s\rangle} \cup \{\#\})^*$ for some s such that for each j between 0 and n, $x \in A_{j,s}$ if and only if $x \in W_i \cap \{0, 1, \ldots, n\}$; and whenever $\tau \in (L_{\langle a_0,a_1,\ldots,a_i,s\rangle} \cup \{\#\})^*$, then $M_{a_0}(\sigma_0 \circ \ldots \circ \sigma_i \circ \tau) \downarrow > i$. This would then imply that $\sigma_0 \circ \sigma_1 \circ \sigma_2 \circ \ldots$ is a text on which M_{a_0} outputs each index only finitely often, contrary to the assumption that M_{a_0} is a confident learner. Suppose, however, that only finitely many $a_0, a_1, a_2, \ldots$ exist; therefore, if a_k were the last term in this sequence, then for all $\sigma \in (L_{\langle a_0,a_1,\ldots,a_k,s\rangle} \cup \{\#\})^*$, where s is large enough so that $A_{k,t} = A_{k,s}$ whenever $t > s$, there is a sequence $\tau \in (L_{\langle a_0,a_1,\ldots,a_k,s\rangle} \cup \{\#\})^*$ so that $M_{a_0}(\sigma_0 \circ \sigma_1 \circ \ldots \circ \sigma_{k-1} \circ \sigma \circ \tau) \leq k$. Hence, since $L_{\langle a_0,a_1,\ldots,a_k,s\rangle} \notin \{W_0, W_1, \ldots, W_k\}$ and range$(\sigma_0 \circ \sigma_1 \circ \ldots \circ \sigma_{k-1}) \subset L_{\langle a_0,a_1,\ldots,a_k,s\rangle}$ by construction, there is a text for $L_{\langle a_0,a_1,\ldots,a_k,s\rangle}$ on which M_{a_0} outputs an incorrect index infinitely often, again contradicting the assumption that M_{a_0} is a confident partial learner of $\mathcal{C}$. In conclusion, the class $\mathcal{C}$ is $BC_{r.e.}$ learnable but not confidently partially learnable with respect to the hypothesis space $\{W_0, W_1, W_2, \ldots\}$. ♦

The following theorem formulates a learning criterion that may appear at first sight to be less stringent than confident partial learnability, but is in fact equivalent to it. This result is then applied in the subsequent theorem to show that every vacillatorily learnable class of r.e. languages is also confidently partially learnable.

Theorem 17 *A class $\mathcal{C}$ 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 by M infinitely often on T, then there is an index e of L with $e \leq d$.*

Proof. Suppose that there is a recursive learner M of $\mathcal{C}$ which satisfies the learning criteria laid out in the statement of the theorem. Let $pad(e, d)$ be a two-place recursive function such that $W_{pad(e,d)} = W_e$ and $pad(e, d) \neq pad(e', d')$ if $(e, d) \neq (e', d')$ for all numbers e, d, e', d'. One may define a confident partial learner N as

follows: on the input text $T = a_0 \circ a_1 \circ a_2 \circ \ldots$, N outputs $pad(e,d)$ at least n times if and only if M outputs d at least n times and there is a stage $s > n$ such that e is the minimal number not exceeding d which satisfies the condition
$\forall k \leq d[\max\{x \leq s : \forall y \leq x[y \in W_{k,s} \Leftrightarrow y \in \{a_0, a_1, \ldots, a_s\}]\}$
$\leq \max\{x \leq s : \forall y \leq x[y \in W_{e,n} \Leftrightarrow y \in \{a_0, a_1, \ldots, a_n\}]\}]$. Since M outputs exactly one index, say i, infinitely often on the text T, N also outputs infinitely often the number $pad(e,i)$, where e is the least index with $e \leq i$ such that either $W_e = \text{range}(T)$, or the minimum number x_e for which $W_e(x_e) \neq T(x_e)$ is equal to $\max\{\{x_k : k \leq i \wedge x_k = \min\{y : W_k(y) \neq T(y)\}\}\}$. For all i' different from i, N outputs $pad(k,i')$ finitely often as M outputs i' only finitely often; for each $k \neq e$ not exceeding i, there is a stage s sufficiently large so that for all subsequent stages, k will never satisfy the condition imposed on e. Hence N, on every text it is fed with, outputs exactly one index infinitely often. Furthermore, if T is a text for a language L in $\mathcal{C}$, and i is the index that M outputs infinitely often on T, then the number $e \leq i$ such that $W_e(y) = T(y)$ on the longest possible initial segment $\{0, 1, \ldots, x_k\}$ among all indices $k \leq i$ is also an index for L, that is, $W_e = L$. This establishes that N is a confident partial learner of $\mathcal{C}$. Conversely, if P were a confident partial learner of $\mathcal{C}$, then P also fulfils the learning criteria in the statement of the theorem: if P is presented with a text for some L in $\mathcal{C}$, then the index d that it outputs infinitely often satisfies $W_d = L$.

Theorem 18 *If a class $\mathcal{C}$ is vacillatorily learnable, then $\mathcal{C}$ is confidently partially learnable.*

Proof. By the criterion established in Theorem 17, it suffices to prove that if $\mathcal{C}$ were vacillatorily learnable, then there is a learner N such that N outputs on every text T exactly one index d infinitely often, and if T is a presentation of some L in $\mathcal{C}$, then d is an upper bound for an index of L. Suppose that M is a vacillatory learner of $\mathcal{C}$. Let $T = a_0 \circ a_1 \circ a_2 \circ \ldots$ be a text, and define N to be a recursive learner such that:

- N outputs the number d at least n times if and only if there is a stage $s > n$ such that $d = \max\{M(\sigma) : \sigma \preceq a_0 \circ \ldots \circ a_s\}$;
- N outputs a fixed index 0 for $\emptyset$ at least n times if and only there is a stage s at which $M(a_0 \circ \ldots \circ a_s) > n$.

If M outputs an infinite set of different indices on the text T, then N outputs 0 infinitely often, and all other indices for at most a finite number of times. If M outputs only finitely many indices $e_0, e_1, \ldots, e_n$, then N outputs $\max\{e_0, e_1, \ldots, e_n\}$ infinitely often. In addition, if T is a text for some L in $\mathcal{C}$, then M outputs only finitely many indices, so that N outputs the maximum, m, of these indices infinitely often, and there is an $e \leq m$ such that $W_e = L$. Thus N satisfies the required learning criteria, and it follows by Theorem 17 that $\mathcal{C}$ must be confidently partially learnable.

As was pointed out earlier, the union of the class of all finite sets and the class $\{\mathbb{N}\}$ is not behaviourally correctly learnable, even though both of the classes $\{D : D \text{ is finite}\}$ and $\{\mathbb{N}\}$ are explanatorily learnable. On the other hand, it is quite a curious feature of confident learning under various success criteria that it is closed under finite unions. In particular, it is shown in [27] that the union of finitely many confidently vacillatorily learnable classes is also confidently vacillatorily learnable;

the analogous result for confident behaviourally correct learning also holds true. The next theorem states that this property of confident learning also extends to partial learnability. That is to say, if $\mathcal{C}_1$ and $\mathcal{C}_2$ are confidently partially learnable classes of r.e. languages, then $\mathcal{C}_1 \cup \mathcal{C}_2$ is also confidently partially learnable. The proof illustrates a padding technique, dependent on the underlying hypothesis space of the learner, that is often applied throughout this work to construct confident partial learners.

Theorem 19 *Confident partial learning is closed under finite unions; that is, if $\mathcal{C}_1$ and $\mathcal{C}_2$ are confidently partially learnable classes, then $\mathcal{C}_1 \cup \mathcal{C}_2$ is confidently partially learnable.*

Proof 1. Let M and N be confident partial learners of the classes $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. A new confident partial learner which learns $\mathcal{C}_1 \cup \mathcal{C}_2$ may be defined as follows. There is a one-one function f such that $f(i, j, k)$ is an index of W_i if k is even, and an index of W_j if k is odd. The new learner R outputs $f(i, j, k)$ at least n times if and only if the following conditions hold:

- M outputs i at least n times;
- N outputs j at least n times;
- if $k = 0$, then for some $s > n$, $\forall x < n[W_{i,s}(x) = W_{j,s}(x)]$;
- if $k = 2o + 1$, then there is an $s > n$ such that o is the minimum value where $W_{i,s}(o) \neq W_{j,s}(o)$ and $W_{j,s}(o) = 1$ if and only if o has been observed in the input data so far;

- if $k = 2o + 2$, then there is an $s > n$ such that o is the minimum value where $W_{i,s}(o) \neq W_{j,s}(o)$ and $W_{i,s}(o) = 1$ if and only if o has been observed in the input data so far.

Consider an index of the form $f(i, j, k)$. If M outputs i finitely often, or N outputs j finitely often, then R outputs $f(i, j, k)$ only finitely often. Suppose, on the other hand, that M outputs i and N outputs j infinitely often. By the confidence of M and N, there is exactly one such pair of numbers $\langle i, j\rangle$. To show that there is exactly one value of k such that R outputs $f(i, j, k)$ infinitely often, consider first the case that $W_i = W_j$. Then for all x, there is an s such that for all $y < x$, $W_{i,s}(y) = W_{j,s}(y)$, and so in following the above algorithmic instructions, R outputs the index $f(i, j, 0)$ infinitely often. However, since for every number o there are at most finitely many s such that $W_{i,s}(o) \neq W_{j,s}(o)$, this means that R outputs an index of the form $f(i, j, 2o + 1)$ or $f(i, j, 2o + 2)$ only finitely often.

Secondly, suppose that $W_i \neq W_j$, and let o be the least number with $W_i(o) \neq W_j(o)$. There is an s sufficiently large so that for all $s' \geq s$, it holds that $W_{i,s'}(o) \neq W_{j,s'}(o)$, and hence R will output the index $f(i, j, 0)$ only finitely often. Let $f(i, j, m)$ be an index for which $m \neq o$. Then m is not the minimum value such that $W_i(m) \neq W_j(m)$; thus whenever s is large enough, either $W_{i,s}(m) = W_{j,s}(m)$ holds or there is a $k < m$ with $W_{i,s}(k) \neq W_{j,s}(k)$. For this reason, R outputs the indices $f(i, j, 2m+1)$ and $f(i, j, 2m + 2)$ finitely often. Lastly, consider the indices $f(i, j, 2o + 1)$ and $f(i, j, 2o + 2)$. Without loss of generality, assume that $W_i(o) = 1$ and $W_j(o) = 0$. If o eventually appears in the text presented, then for all large enough s, o is the minimum value that occurs in the data revealed with $W_{i,s}(o) \neq W_{j,s}(o)$, and in addition $W_{i,s}(o) = 1$, $W_{j,s}(o) = 0$; whence, R must output $f(i, j, 2o + 2)$ infinitely often and $f(i, j, 2o + 1)$ finitely often. If o never occurs in the text presented, then for all large enough s, o is the minimum value such that $W_{i,s}(o) \neq W_{j,s}(o)$, and $W_{j,s}(o) = o$, so that R outputs $f(i, j, 2o+1)$ infinitely often and $f(i, j, 2o+2)$ finitely often. This completes the case distinction and establishes that R is confident.

Suppose further that R is presented with a text for some L in $\mathcal{C}_1$. On this text, M will output exactly one index i for L infinitely often, and N will also output exactly one index j infinitely often. If $W_i = W_j$, then R will output the index $f(i, j, 0)$ infinitely often; by the definition of f, $f(i, j, 0)$ is an index for W_i and thus R confidently partially learns L. If $W_i \neq W_j$, let o be the minimum value such that $W_i(o) \neq W_j(o)$. If $o \in W_i$, then o will eventually appear in the input data and hence R will output $f(i, j, 2o + 2)$ infinitely often, which is an index for W_i by the definition of f. If $o \notin W_i$, then o will never occur in the input data and R still outputs the index $f(i, j, 2o + 2)$ infinitely often. For the case that L is in $\mathcal{C}_2$, an argument analogous to the preceding one, with the roles of M and N interchanged, may be applied. In conclusion, R confidently partially learns $\mathcal{C}_1 \cup \mathcal{C}_2$. ♦

Proof 2. Let M and N be confident partial learners of the classes $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. Now using Theorem 17, one can consturct a new learner R which outputs $\langle i, j \rangle$ at least n times iff M outputs i and N outputs j at least n times. It is directly obvious that on every text of a function, the learner R outputs exactly one index $\langle i, j \rangle$ infinitely often; this index is an upper bound of an index e of the function to be learnt whenever $i \geq e \vee j \geq e$. Hence R is a confident partial learner (in the sense of Theorem 17) of $\mathcal{C}_1 \cup \mathcal{C}_2$. ♦

With a similar aim as Theorem 17 - to compare and contrast the learning strength of confident partial learning with that of other possible learning criteria - the next theorem considers a variant of confident learning, whereby the learner is constrained to converge semantically on any given text. This, however, again does not give rise to any new learning notion, as one can show that any class of r.e. languages that is learnable according to the proposed criterion can already be confidently partially learnt. Nonetheless, the result bears out the view that confident partial learning is quite a versatile learning requirement.

Theorem 20 *A recursive learner M is said to confidently behaviourally correctly learn a class $\mathcal{C}$ if for every text T there is an r.e. language L such that M almost always outputs an index for L when it is presented with T; and if T is a text of some language L' in $\mathcal{C}$, then $L = L'$. Every confidently behaviourally correctly learnable class is confidently partially learnable.*

Proof. Let M be a confident behaviourally correct learner of the class $\mathcal{C}$. Suppose further that M never returns to an old hypothesis; that is, for all strings $\sigma \in (\mathbb{N} \cup \{\#\})^*$ and $\gamma \prec \sigma$, $M(\sigma) \neq M(\gamma)$. Owing to the padding lemma, this requirement on M may always be imposed by setting, if necessary, a new learner to conjecture an index $j > i$ such that $W_j = W_i$ if M has already hypothesised i at an earlier stage. A confident partial learner N of $\mathcal{C}$ may be defined as follows. Let $pad(e, d)$ be a recursive function with $W_{pad(e,d)} = W_e$ for all e, d.

N outputs $pad(e, d + 1)$ at least n times if and only if there is a stage $s > 2n$ such that

- $M(a_0 \circ a_1 \circ \ldots \circ a_{i+1}) = e$ for some i with $i \leq n$;

- for all $x < n$, $W_{e,s}(x) = W_{M(a_0 \circ a_1 \circ \ldots \circ a_{i+1} \circ \ldots \circ a_j),s}(x)$, where $j = i+2, i+3, \ldots, i+n+1$; in other words, $W_{e,s}$ agrees with the s-approximations of its subsequent n conjectures on all values of x below n;

- d is the minimum number such that $W_{M(a_0 \circ \ldots \circ a_i),s}(d) \neq W_{e,s}(d)$.

Furthermore, N outputs $pad(e, 0)$ at least n times if and only if there is a stage $s > n$ such that if $a_0 a_1 \ldots a_s$ is the input data, then $M(a_0) = e$, and for all $x < n$, $W_{e,s}(x) = W_{M(a_0 \circ a_1 \circ \ldots \circ a_j),s}(x)$, where $j = 1, 2, \ldots, n$.

At each stage, there are only finitely many values of $pad(e, d)$ that qualify as hypotheses for N; in addition, N may output an index different from its all preceding conjectures if no value of $pad(e, d)$ is valid. Hence N may be extended to a well-defined recursive learner.

To show that N is a confident partial learner of $\mathcal{C}$, let N be presented with any given text T, and suppose that M on T converges semantically to the r.e. set L; by the confident behaviourally correct learning property of M, such a set L must exist, and if T is a presentation of some language L' in $\mathcal{C}$, then $L = L'$. It shall be argued that N outputs exactly one index of the form $pad(e, d)$ infinitely often, and is such that $W_{pad(e,d)} = L$. Two cases are distinguished: first, when M, on the text T, outputs an index e such that $W_e \neq L$; second, when all the conjectures of M on T are semantically identical, that is, $W_e = L$ for all indices e that M outputs.

For the first case, suppose that $p = \max\{e : W_{M(T[e])} \neq L\}$; here $T[e]$ denotes the sequence of the first $e + 1$ data bits of T. Let $h = M(T[p + 1])$; h is the first conjecture of M from which point onwards it converges semantically to L. Then $W_{M(T[p+k])} = L$ for all $k \geq 1$, and there is a minimum value d such that

$W_{M(T[p])}(d) \neq L(d)$. Hence for all n, there is a stage $s > 2n$ such that whenever $x < n$ and $1 \leq j \leq n$, then $W_{p,s}(x) = W_{M(T[e+j]),s}(x)$; furthermore, d is the least number such that $W_{M(T[p]),s}(d) \neq W_{h,s}(d)$. As a consequence of the first condition defined on N, N outputs the index $pad(h, d+1)$ infinitely often.

Next, consider any index g that M conjectures before it outputs h, that is, $g = M(T[k])$ for some $k \leq p$. Since, by assumption, all the indices that M outputs on T are different, $g \neq h$. There is a subsequent conjecture of M, say $M(T[k+l])$, such that $W_{M(T[k+l])} \neq W_g$. It follows that if e is the least number for which $W_{M(T[k+l])}(e) \neq W_g(e)$, then for all large enough s, $W_{M(T[k+l]),s}(e) \neq W_g(e)$, and thus for any value of x, $pad(g, x+1)$ fails to qualify as a valid conjecture of N at almost all stages.

Now let g' be any index that M conjectures after it outputs h; $g' = M(T[p+k+1])$ for some k. Then $W_{M(T[p+k])} = W_{g'} = L$, that is, there is no minimum number d' such that $W_{M(T[p+k])}(d') \neq W_{g'}(d')$; whence, every index of the form $pad(g', x)$ is output only finitely often.

In regard to the second case: as $W_{M(T[k])} = L$ for all k, there are no numbers d', k, such that $W_{M(T[k+1])}(d') \neq W_{M(T[k])}(d')$, so that the first condition defined on N occurs at most finitely often. This means that every index of the form $pad(g', x+1)$, where g' is a conjecture of M on T, is output only finitely often. On the other hand, since $W_{M(T[0])} = W_{M(T[k])}$ for all k, there is for every n an $s > n$ such that $W_{M(T[0]),s}(x) = W_{M(T[k]),s}(x)$ whenever $x < n$ and $k \leq n$. Hence N outputs $pad(M(T[0]), 0)$ infinitely often.

This completes the case distinction and establishes that N is a confident partial learner of $\mathcal{C}$, as claimed. ♦

The fact that the Padding Lemma, satisfied by any acceptable numbering of all r.e. sets, is used in a crucial way for some of the preceding proofs, raises the question of how confident partial learnability varies with the choice of a learner's hypothesis space. To emphasise the connection between these two aspects of learning, the next series of results show that certain analogues of earlier theorems fail to hold under the setting of more general hypothesis spaces where the technique of padding may not be applicable, as would be the case if, for example, the learner fixes a Friedberg numbering as its hypothesis space.

Theorem 21 *The class $\mathcal{C} = \{\{e\} \oplus W_e : W_e$ is cofinite$\}$ of recursive sets is explanatorily learnable with respect to r.e. indices but is not confidently partially learnable with respect to co-r.e. indices.*

Proof. On the input data σ, an explanatory learner outputs an r.e. index for $\{e\} \oplus W_e$ for the first e such that $2e \in \text{range}(\sigma)$; if no such number e exists, then the learner outputs 0. Now assume by way of contradiction that there were a confident partial co-r.e. learner M of the class $\mathcal{C}$. By the confidence of M, for every number e there is a sequence $\sigma \in ((\{e\} \oplus W_e) \cup \{\#\})^*$ and an index d with $M(\sigma) = d$ such that for all $\tau \in ((\{e\} \oplus W_e) \cup \{\#\})^*$ there is an $\eta \in ((\{e\} \oplus W_e) \cup \{\#\})^*$ for which $M(\sigma\tau\eta) = d$. This sequence σ and index d may be found using the oracle $\mathbb{K}'$. Suppose first that W_e were cofinite. Since M confidently partially learns $\{e\} \oplus W_e$, one has that $|W_d| < \infty$, and for all numbers x, $x \in W_e$ holds if and only if $x \notin W_d$ holds as well. The latter condition may be checked by means of the oracle $\mathbb{K}'$. Suppose, on the other hand, that W_e were coinfinite. Then, either $|W_d|$ is infinite, or there must exist an x such that $x \notin W_e \cup W_d$. This case distinction shows that $\{e : W_e \text{ is cofinite}\}$ is Turing reducible to $\mathbb{K}'$, a contradiction to the established fact that it is Σ_3^0-complete. In conclusion, the class $\mathcal{C}$ is not confidently partially learnable with respect to co-r.e. indices. ♦

Theorem 22 *There are uniformly r.e classes $\mathcal{L}_1, \mathcal{L}_2$, such that $\mathcal{L}_1$ and $\mathcal{L}_2$ are confidently partially learnable using $\mathcal{L}_1$ and $\mathcal{L}_2$ as hypothesis spaces respectively, but $\mathcal{L}_1 \cup \mathcal{L}_2$ is not confidently partially learnable using itself as a hypothesis space.*

Proof. Let $\mathcal{L}_1 = \{U_{\langle d,e,0\rangle} = \{\langle d,e,x\rangle : x \in W_d\} : d,e \in \mathbb{N}\}$, and $\mathcal{L}_2 = \{U_{\langle d,e,1\rangle} = \{\langle d,e,x\rangle : x \in W_e\} : d,e \in \mathbb{N}\}$. Each of $\mathcal{L}_1$ and $\mathcal{L}_2$ is confidently partially learnable using itself as a hypothesis space: a confident partial learner for $\mathcal{L}_1$ outputs $\langle d,e,0\rangle$ if $\langle d,e,x\rangle$, where x is any number, is the first triple that the data reveals, while a confident partial learner for $\mathcal{L}_2$ outputs $\langle d,e,1\rangle$ upon witnessing the same data; otherwise, if no number occurs in the data, then the learners output a default index ?. Now assume by way of contradiction that $\mathcal{L}_1 \cup \mathcal{L}_2$ were confidently partially learnable using $\mathcal{L}_1 \cup \mathcal{L}_2$ as the hypothesis space; let M be such a recursive learner. Fix any index d of $\mathbb{K}$. It shall be shown next that there is an algorithm using the oracle $\mathbb{K}$ for deciding whether or not any given r.e. set W_e is equal to $\mathbb{K}$. Let e be any given number; now generate an infinite text $T = \langle d,e,x_0\rangle \circ \langle d,e,x_1\rangle \circ \langle d,e,x_2\rangle \circ \ldots$ for $U_{\langle d,e,0\rangle}$, where $x_0, x_1, x_2, \ldots$ is a one-one enumeration of $\mathbb{K}$. By accessing the oracle $\mathbb{K}$, run a search for the first $x_i \in \mathbb{K}$ such that one of the following conditions holds:

1. There is a $y \leq x_i$ with $y \in \mathbb{K} - W_e$ or $y \in W_e - \mathbb{K}$;
2. There is no sequence $\sigma \in ((U_{\langle d,e,0\rangle} \cap U_{\langle d,e,1\rangle}) \cup \{\#\})^*$ such that $M(\langle d,e,x_0\rangle \circ$

$\ldots \circ \langle d,e,x_i\rangle \circ \sigma) = \langle d,e,0\rangle$;

3. There is no sequence $\sigma \in ((U_{\langle d,e,1\rangle} \cap U_{\langle d,e,0\rangle}) \cup \{\#\})^*$ such that $M(\langle d,e,x_0\rangle \circ \ldots \circ \langle d,e,x_i\rangle \circ \sigma) = \langle d,e,1\rangle$.

If $W_e \neq \mathbb{K}$, then there is a y and an x_i with $y \leq x_i$ for which either $y \in \mathbb{K} - W_e$ or $y \in W_e - \mathbb{K}$ holds; thus condition 1. would eventually be satisfied. If, on the other hand, $W_e = \mathbb{K}$, then $U_{\langle d,e,0\rangle} = U_{\langle d,e,1\rangle}$, so that T is also a text for $U_{\langle d,e,1\rangle}$; indeed, $U_{\langle d,e,0\rangle}$ and $U_{\langle d,e,1\rangle}$ are the only two r.e. sets in $\mathcal{L}_1 \cup \mathcal{L}_2$ for which T is a text. By the confidence of M, M outputs exactly one of the two indices - $\langle d,e,0\rangle$ or $\langle d,e,1\rangle$ - infinitely often on the text T. If M outputs $\langle d,e,0\rangle$ infinitely often, then condition 3. would be satisfied at some stage; if it outputs $\langle d,e,1\rangle$ infinitely often, then condition 2. would eventually hold. Hence the above decision procedure using the oracle $\mathbb{K}$ is effective. One can then conclude that if condition 1. holds, then $W_e \neq \mathbb{K}$; and if either condition 2. or 3. is satisfied, then $W_e = \mathbb{K}$. In other words, the index set $\{e : W_e = \mathbb{K}\}$ is Turing reducible to $\mathbb{K}$, which is impossible since $\{e : W_e = \mathbb{K}\}$ has the Turing degree of $\mathbb{K}'$. In conclusion, the class $\mathcal{L}_1 \cup \mathcal{L}_2$ is not confidently partially learnable using itself as a hypothesis space. ♦

Theorem 23 *The uniformly r.e. class* $\mathcal{C} = \mathcal{L}_1 \cup \mathcal{L}_2$, *where* $\mathcal{L}_1 = \{L_e = \{e + x : x \leq |W_e|\} : e \in \mathbb{N}\}$ *and* $\mathcal{L}_2 = \{H_e = \{e + x : x \in \mathbb{N}\} : e \in \mathbb{N}\}$ *is vacillatorily learnable, but not confidently partially learnable using the hypothesis space* $\{L_0, H_0, L_1, H_1, L_2, H_2, \ldots\}$.

Proof. A behaviourally correct learner of $\mathcal{C}$ may perform as follows: on the input σ with minimum number e and maximum number $e+a$, the learner checks if $|W_{e,|\sigma|}| \geq a$. If so, then it conjectures L_e; otherwise, it outputs H_e.

On the other hand, if $\mathcal{C}$ were confidently partially learnable by a recursive learner M, then, for any given number e, one may enumerate a default text $T(0) \circ T(1) \circ T(2) \circ \ldots$ for L_e, and use the oracle $\mathbb{K}$ to search for the first number k such that for all $\sigma \in (L_e \cup \{\#\})^*$, M does not conjecture one of the sets L_e, H_e on the input $T(0) \circ T(1) \circ \ldots \circ T(k) \circ \sigma$. By the confidence of M, such a number k must always exist. If k is found such that M does not conjecture L_e for all inputs $T(0) \circ T(1) \circ T(2) \circ \ldots \circ T(k) \circ \sigma$ such that $\sigma \in (L_e \cup \{\#\})^*$, then it may be concluded that W_e is infinite. Otherwise, if H_e is the set that M eventually rejects, then it may be tested, again by means of the oracle $\mathbb{K}$, whether or not there exists a $\tau \in (H_e \cup \{\#\})^*$ for which M conjectures H_e on the input $T(0) \circ T(1) \circ \ldots \circ T(k) \circ \tau$. If such a τ exists, then one may conclude that W_e is finite; if, however, no such τ can be found, then W_e must be infinite. Hence $\{e : |W_e| = \infty\}$ is Turing reducible to $\mathbb{K}$, which is impossible since it has the same Turing degree as $\mathbb{K}'$. In conclusion, $\mathcal{C}$ is not confidently partially learnable. ♦

Fortunately, not all of the relations established hitherto between confident partial learning and other learning criteria with respect to the default hypothesis space $\{W_0, W_1, W_2, \ldots\}$ are lost when considering more general hypothesis spaces; if the learner's hypothesis space is uniformly r.e., one can show that a weaker version of Theorem 18, that explanatory learnability implies confident partial learnability, is preserved.

Theorem 24 *Let $\mathcal{C} = \{L_0, L_1, L_2, \ldots\}$ be a uniformly r.e. class that is explanatorily learnable. Then $\mathcal{C}$ is confidently partially learnable with respect to the hypothesis space $\{L_0, L_1, L_2, \ldots\}$.*

Proof. Assume that M is an explanatory learner of $\mathcal{C}$ with respect to a uniformly r.e. hypothesis space $\{H_0, H_1, H_2, \ldots\}$. Then there exists a uniformly $\mathbb{K}$-recursive family of finite sequences $\sigma_0, \sigma_1, \sigma_2, \ldots$ such that for each e,

- range$(\sigma_e) \subseteq L_e$;
- for all $\tau \in (L_e \cup \{\#\})^*$, $M(\sigma_e\tau) = M(\sigma_e)$.

One can define a new learner N as follows: on the input η, N outputs the least $e \leq |\eta|$ such that range$(\sigma_{e,|\eta|}) \subseteq$ range(η), where $\sigma_{e,s}$ denotes the sth approximation to σ_e, and for all τ satisfying $|\tau| \leq |\eta|$ and range$(\tau) \subseteq$ range(η), $M(\sigma_{e,|\eta|}\tau) = M(\sigma_{e,|\eta|})$. If such a number e does not exist, then N outputs the default index 0.

Claim 25 *If N outputs on a text T an index e infinitely often, then M converges to an index i with respect to its hypothesis space $\{H_0, H_1, H_2, \ldots\}$ on the text $\sigma_e \circ T(0) \circ T(1) \circ T(2) \circ T(3) \circ \ldots$, and if T were a text for some language L in $\mathcal{C}$, then $L_e = H_i = L$.*

Suppose that N outputs the index e infinitely often, and let n be sufficiently large so that $\sigma_{e,s} = \sigma_e$ for all $s > n$. Then e is an index for which range$(\sigma_e) \subseteq$ range(T). Furthermore, for all τ such that τ is a prefix of T, $M(\sigma_e\tau) = M(\sigma_e)$. Hence M converges on the text $\sigma_e \circ T(0) \circ T(1) \circ T(2) \circ T(3) \circ \ldots$ to some fixed index i. Suppose further that T were a text for some L_a in $\mathcal{C}$. Then, since M explanatorily learns L_a, there is a least number e for which M converges to some fixed index on $\sigma_e \circ T$, and is such that $L_e = L_a$. Moreover, since σ_e is a locking sequence for L_e (and thus also for L_a), this means that for all $\tau \in (L_a \cup \{\#\})^*$, $M(\sigma_e\tau) = M(\sigma_e)$. Hence N explanatorily learns $\mathcal{C}$ using the hypothesis space $\{L_0, L_1, L_2, \ldots\}$. This establishes the claim.

The confident partial learner P is now defined by setting P to output e at least n times if and only if N outputs e at least n times, and to output the default index 0 at least n times if N makes at least n mind changes. P is indeed confident: if there is a least index e such that M converges to some index i on the text $\sigma_e \circ T$, then P converges in the limit to e; if, on the other hand, no such index e exists, then N will continue searching for a larger index at every stage that satisfies the required condition that $M(\sigma_k\tau) = M(\sigma_k)$ for all $\tau \in (\text{range}(T) \cup \{\#\})^*$, and consequently outputs the default index 0 infinitely often. Finally, since N explanatorily learns $\mathcal{C}$ with respect to the hypothesis space $\{L_0, L_1, L_2, \ldots\}$, it follows that P also explanatorily learns $\mathcal{C}$ using the same hypothesis space. ♦

3.2 Partial Conservative Learning

Conservativeness is a learnability constraint that has been studied fairly extensively in the inductive inference literature, especially in the setting of indexed families [1, 15]. In the remainder of this section, we consider the notion of *partial conservativeness* in language learning; in brief, this is partial learning combined with the constraint that if a learner outputs e infinitely often on a text for some target language L, then none of its other conjectures on this text can contain L as a subset. In the first place, it is observed that Gold's class does not satisfy this learning criterion.

Theorem 26 *The class* $\mathcal{C} = \{\mathbb{N}\} \cup \{F : F$ *is finite*$\}$ *is not partially conservatively learnable.*

Proof. Assume by way of contradiction that M were a recursive partially conservative learner of $\mathcal{C}$. Since M learns $\mathbb{N}$, there is a sequence $a_0 \circ a_1 \circ \ldots \circ a_n \in (\mathbb{N} \cup \{\#\})^*$ such that $M(a_0 \circ a_1 \circ \ldots \circ a_n) = e$ for some e with $\mathbb{N} = W_e$. Then $a_0 \circ a_1 \circ \ldots \circ a_n$ is the initial segment of a text for the finite set $\{a_0, a_1, \ldots, a_n\} - \{\#\}$, but since M outputs an index e with $\mathbb{N} = W_e \supset \{a_0, a_1, \ldots, a_n\} - \{\#\}$, M cannot be a partially conservative learner of $\mathcal{C}$. ♦

Theorem 27 *Let $\{\varphi_{f(0)}, \varphi_{f(1)}, \varphi_{f(2)}, \ldots\}$ be a Friedberg numbering of all partial-recursive functions. Consider the set $\mathcal{C} = \{\varphi_{f(e)} : \varphi_{f(e)}$ is recursive$\}$ of recursive functions, and build the class of graphs $\mathcal{G} = \{\{\langle x, y\rangle : \varphi_{f(e)}(x) \downarrow= y\} : \varphi_{f(e)} \in \mathcal{C}\}$. Then $\mathcal{G}$ is partially conservatively learnable but neither confidently partially learnable nor behaviourally correctly learnable.*

Proof. First, a partially conservative learner M may be programmed to work as follows: on the input $\sigma = \langle x_0, y_0\rangle \circ \langle x_1, y_1\rangle \circ \ldots \langle x_n, y_n\rangle$, M searches for the least $e \leq n$ such that $\varphi_{f(e),n}(x_i) \downarrow= y_i$ for $i = 0, 1, \ldots, n$, and conjectures $g(e)$ for which $W_{g(e)} = \{\langle x, y\rangle : x \in \mathbb{N} \wedge \varphi_e(x) \downarrow= y\}$; if e does not exist, then M outputs $\max\{M(\tau) : \tau \prec \sigma\}$ if $|\sigma| > 1$, and an index for $\emptyset$ if $|\sigma| = 1$. M as defined must be a partial learner of $\mathcal{G}$, for if it were presented with a text of the graph of some $\varphi_{f(e)}$ in $\mathcal{C}$, then, due to the one-one numbering property of $\{\varphi_{f(0)}, \varphi_{f(1)}, \varphi_{f(2)}, \ldots\}$, $graph(\varphi_{f(e)}) \subseteq \{\langle x, y\rangle : \varphi_{f(d)}(x) \downarrow= y\}$ holds if and only if $d = e$. Consequently, M must output $g(e)$ infinitely often, and every other index $g(d)$ with $d \neq e$ only finitely often. Furthermore, M is also partially conservative: for every $d \neq e$, there is a number x such that either $\varphi_{f(d)}(x) \uparrow$, or $\varphi_{f(d)}(x) \downarrow\neq \varphi_{f(e)}(x)$. This implies that for every $d \neq e$, $W_{g(e)} \not\subset W_{g(d)}$, so that M is partially conservative. Thus $\mathcal{G}$ is partially conservatively learnable.

That $\mathcal{G}$ is not, however, confidently partially learnable, follows from Theorems 32 and 4.1. Alternatively, one can argue as follows. Assume by way of contradiction that $\mathcal{G}$ were confidently partially learnable via a recursive learner M. By the confidence of M, one may find a finite sequence $\alpha = \langle 0, y_0 \rangle \circ \langle 1, y_1 \rangle \circ \ldots \circ \langle n, y_n \rangle$ such that, for some unique index e, $M(\alpha) = e$, and for each $\sigma \in (\mathbb{N} \cup \{\#\})^*$ of the form $\sigma = \langle n+1, z_{n+1} \rangle \circ \ldots \circ \langle n+k, z_{n+k} \rangle$, there is a sequence $\tau \in (\mathbb{N} \cup \{\#\})^*$ of the form $\tau = \langle n+k+1, z_{n+k+1} \rangle \circ \ldots \circ \langle n+k+i, z_{n+k+i} \rangle$ with $M(\alpha \circ \sigma \circ \tau) = e$. A new recursive function g may now be defined inductively as follows.

- Set $g(i) = y_i$ for all $i \leq n$.
- Assume that $g(x)$ has been defined for all $x \leq k$ with $k \geq n$. Run a search for a sequence of the form $\langle k+1, z_{k+1} \rangle \circ \ldots \circ \langle k+l, z_{k+l} \rangle$ such that $M(\langle 0, g(0) \rangle \circ \langle 1, g(1) \rangle \circ \ldots \circ g(k) \circ \langle k+1, z_{k+1} \rangle \circ \ldots \circ \langle k+l, z_{k+l} \rangle) = e$; since $\langle 0, g(0) \rangle \circ \ldots \langle n, g(n) \rangle = \alpha$ is a locking sequence for M corresponding to the index e, the search must eventually terminate successfully. Set $g(k+j) = z_{k+j}$ for $j = 1, \ldots, l$, and $g(k+l+1) = \varphi_{e'}(k+l+1) + 1$ if W_e is the graph of a recursive function $\varphi_{e'}$; otherwise, $g(k+l+1)$ remains undefined until the next stage.

If W_e is not the graph of a recursive function, then
$W_e \neq \{\langle x, y \rangle : x \in \mathbb{N} \wedge g(x) \downarrow = y\}$; M, however, outputs e infinitely often on the text $\langle 0, g(0) \rangle \circ \langle 1, g(1) \rangle \circ \langle 2, g(2) \rangle \circ \ldots$, and so it cannot confidently partially learn the graph of g. In the case that W_e were the graph of some recursive function $\varphi_{e'}$, then, since g is defined to be such that $\langle k, g(k) \rangle \neq \langle k, \varphi_{e'}(k) \rangle$ for infinitely many k, $W_e \neq \{\langle x, y \rangle : x \in \mathbb{N} \wedge g(x) \downarrow = y\}$ still holds, and thus M fails to confidently

partially learn the graph of g. This contradiction establishes that $\mathcal{G}$ is not confidently partially learnable.

Lastly, assume towards a contradiction that N were a behaviourally correct learner of $\mathcal{G}$. Now, given any number e, one may check relative to the oracle $\mathbb{K}$ whether or not φ_e is recursive via the following decision procedure.

1. At stage s, determine whether $\varphi_e(x)$ is defined for all $x \leq s$. If there is an $x \leq s$ for which $\varphi_e(x) \uparrow$, then φ_e is not recursive. Otherwise, proceed to the next step.

2. Check via $\mathbb{K}$ whether or not there exists a $\tau \in (graph(\varphi_e) \cup \{\#\})^*$ such that for some $\langle x, y \rangle \in W_{N(\sigma \circ \tau)}$, where $\sigma = \langle 0, \varphi_e(0) \rangle \circ \ldots \circ \langle s, \varphi_e(s) \rangle$, $\langle x, y \rangle \in W_{\sigma \circ \tau}$ and $\varphi_e(x) \downarrow \neq y$. If so, proceed to the next stage and return to Step 1. ; otherwise, it may be concluded that φ_e is a total recursive function.

If φ_e were a total recursive function, then N must behaviourally correct learn the graph of φ_e, that is, there is a locking sequence σ for which the condition in Step 2. does not hold. Thus the assumption that $\mathcal{G}$ is BC learnable yields a decision procedure relative to $\mathbb{K}$ for the Π^0_2 set $\{e : \varphi_e \text{ is recursive}\}$, a contradiction. ♦

The next theorem succinctly characterises the oracles relative to which a class of infinite languages is partially conservatively learnable. The hypothesis that all the languages in the class be infinite cannot, however, be dropped, as will be shown in the subsequent result.

Theorem 28 *Let $\mathcal{C}$ be a class of infinite r.e. sets. Then the following three conditions are equivalent.*

(i) $\mathcal{C}$ is partially conservatively learnable;

(ii) $\mathcal{C}$ has an $Ex[\mathbb{K}]$ learner using $\mathbb{K}$-r.e. indices;

(iii) $\mathcal{C}$ has an $Ex[\mathbb{K}]$ learner using r.e. indices.

Proof. Suppose first that $\mathcal{C}$ is $Ex[\mathbb{K}]$ learnable, and let M be an explanatory learner of $\mathcal{C}$ that outputs $\mathbb{K}$-r.e. indices. Assume further that M never repeats a hypothesis e if its subsequent conjecture differs from e; that is, if M outputs e, e' at stages s and $s+1$ respectively, where $e \neq e'$, then M thenceforth does not output e. On the text $T = a_0 \circ a_1 \circ a_2 \circ \ldots$, simulate the learner M, and let f be a recursive function such that for each number e that M outputs on T and all e', n, if σ_e is the shortest prefix of T for which $M(\sigma_e) = e$,

$$W_{f(e',e,v_0,\ldots,v_n,s_0,\ldots,s_n)} = \begin{cases} W_{e'} \cap \{0,1,\ldots,t\} & \text{if } t \text{ is the least number such that} \\ & t > \max(s_0,\ldots,s_n) \wedge \exists i[1 \leq i \leq n \\ & \wedge (W_{e',t}(i) \neq v_i \\ & \text{or } W_{e',t}(i) = 1 \wedge W_{e,t}^{\mathbb{K}_t}(i) = 0)]; \\ W_{e'} \cap \{0,1,\ldots,s\} & \text{if } s \text{ is the least number such that} \\ & \forall u > s[\exists \tau \in (W_{e,u}^{\mathbb{K}_u} \cup \{\#\})^*[M(\sigma_e \circ \tau) \neq e]]; \\ W_{e'} & \text{otherwise.} \end{cases}$$

The first of the above three cases is always assigned priority over the remaining ones; the second case applies only if no t satisfying the condition in the first case is found. If M does not output d on T, then set $W_{f(i,d,v_0,\ldots,v_n,s_0,\ldots,s_n)} = \emptyset$ for all $i, n, v_0, \ldots, v_n, s_0, \ldots, s_n$. Construct a padding function pad for which $W_{pad(e',e,v_0,\ldots,v_n,s_0,\ldots,s_n)} = W_{e'}$, and for all e', e, n, k with $k \leq n$, $pad(e', e, v_0, \ldots, v_k, s_0, \ldots, s_k) =$

$pad(e', d, v'_0, \ldots, v'_n, s'_0, \ldots, s'_n)$ if and only if $e = d$ and for all i such that $1 \leq i \leq k$, $v_i = v'_i$, and if $v_i = v'_i = 1$, then $s_i = s'_i$. Build a new learner P as follows: P outputs $pad(f(e', e, v_0, \ldots, v_n, s_0, \ldots, s_n), e, v_0, \ldots, v_n, s_0, \ldots, s_n)$ exactly once if and only if the conditions listed below hold:

1. M outputs e at least n times;
2. there is a stage $s > n$ for which $\forall i \leq n[W_{e',s}(i) = v_i]$;
3. for all $1 \leq i \leq n$, if $v_i = 1$, then $W_{e,s_i}^{\mathbb{K}_{s_i}}(i) = 1$;
4. for all $1 \leq i \leq n$, if $v_i = 0$, then there is a stage $t_i \geq n$ for which $\varphi_{e,t_i}^{\mathbb{K}_{t_i}}(i)\uparrow$.

It shall be shown that P is partially conservative, and if M converges to some e on T such that $W_e^{\mathbb{K}}$ is r.e., then P outputs an index e' infinitely often if and only if $W_{e'} = W_e^{\mathbb{K}}$ and P outputs e' at least once. Suppose that M does converge to e on the text T, that T is a presentation of some $\mathcal{L}$ in $\mathcal{C}$, and that $W_e^{\mathbb{K}}$ is an r.e. set. If M conjectures d at some stage with $d \neq e$, then it outputs d only finitely often, so that by condition 1., P outputs all indices of the form $pad(f(e', d, v_0, \ldots, v_n, s_0, \ldots, s_n), e, v_0, \ldots, v_n, s_0, \ldots, s_n)$ with $d \neq e$ for at most a finite number of times. To prove the partial conservativeness of P, suppose first that $\mathcal{L} \subset W_d^{\mathbb{K}}$. Since M is an $Ex[\mathbb{K}]$ learner of $\mathcal{L}$, and M never re-issues a hypothesis d if it conjectures an index different from d at a later stage, this implies that there is a sequence $\tau \in (W_d^{\mathbb{K}} \cup \{\#\})^*$ such that $M(\sigma_d \circ \tau) \neq d$, where σ_d is the shortest prefix of T with $M(\sigma_d) = d$. This corresponds to the second case in the construction of f, and so $W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$ must be finite. Hence, as $\mathcal{L}$ is infinite, $\mathcal{L}$ cannot be a proper subset of $W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$. Next, consider the case that $\mathcal{L} \not\subseteq W_d^{\mathbb{K}}$, that is, there is an $x \in \mathcal{L} - W_d^{\mathbb{K}}$. From the first condition in the construction of f, it follows that if $W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$ is infinite, then it is a subset of $W_d^{\mathbb{K}}$. Consequently, if $W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$ is infinite, then there is an $x \in \mathcal{L} - W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$. Thus, the hypothesis that $\mathcal{L}$ is infinite again leads to the conclusion that $\mathcal{L} \not\subset W_{pad(f(e',d,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$. Furthermore, for all indices of the form $pad(f(e', e, v_0, \ldots, v_n, s_0, \ldots, s_n), e, v_0, \ldots, v_n, s_0, \ldots, s_n)$, the construction of f gives that every r.e. set $W_{pad(f(e',e,v_0,\ldots,v_n,s_0,\ldots,s_n),e,v_0,\ldots,v_n,s_0,\ldots,s_n)}$ is either finite, or a subset of $W_e^{\mathbb{K}} = \mathcal{L}$. This completes the verification that P is a partial conservative learner.

Now let e' be an r.e. index with $W_{e'} = W_e^{\mathbb{K}}$. There is an infinite sequence of values $s_0, s_1, s_2, \ldots$ such that for all i, $W_{e',s_i}(i) = W_{e'}(i)$, and if $W_{e',s_i}(i) = 1$, then $W_{e,t}^{\mathbb{K}_t}(i) = 1$ whenever $t \geq s_i$. Thus $W_{pad(f(e',e,W_{e'}(0),\ldots,W_{e'}(n),s_0,\ldots,s_n),e,W_{e'}(0),\ldots,W_{e'}(n),s_0,\ldots,s_n)} = W_{e'}$ for the values of s_i in the above sequence. In addition, it may be observed that the set of values $\{e', e, W_{e'}(0), \ldots, W_{e'}(n), s_0, \ldots, s_n\}$ satisfies conditions 1. to 4. for all n, so that P outputs every index $pad(f(e', e, W_{e'}(0), \ldots, W_{e'}(n), s_0, \ldots, s_n), e, W_{e'}(0), \ldots, W_{e'}(n), s_0, \ldots, s_n)$ exactly once. As pad is defined to be such that $pad(f(e', e, W_{e'}(0), \ldots, W_{e'}(n), s_0, \ldots, s_n), e, W_{e'}(0), \ldots, W_{e'}(n), s_0, \ldots, s_n)$ $= pad(f(e', e, W_{e'}(0), \ldots, W_{e'}(k), s_0, \ldots, s_k), e, W_{e'}(0), \ldots, W_{e'}(k), s_0, \ldots, s_k)$ for all n, k, it follows that P outputs a single index for $W_{e'}$ infinitely often.

Suppose, on the other hand, that e'' were an r.e. index such that $W_{e''} \neq W_e^{\mathbb{K}}$. First, assume that for some i, $W_{e''}(i) = 1$ but $W_e^{\mathbb{K}}(i) = 0$. Therefore condition 3. does not hold at infinitely many stages, and so for all s_i, P outputs indices of the form $pad(f(e'', e, v_0, \ldots, v_n, s_0, \ldots, s_i, \ldots, s_n), e, v_0, \ldots, v_n, s_0, \ldots, s_i, \ldots, s_n)$ only finitely often. Second, assume that for some i, $W_{e''}(i) = 0$ but $W_e^{\mathbb{K}}(i) = 1$. As a consequence, there is a sufficiently large stage s so that for all $u > s$, $\varphi_{e,u}^{\mathbb{K}_u}(i) \downarrow$, implying that condition 4. fails to hold for indices of the form $pad(f(e'', e, W_{e''}(0), \ldots, W_{e''}(n), s_0, \ldots, s_n), e, W_{e''}(0), \ldots, W_{e''}(n), s_0, \ldots, s_n)$ whenever $n > s$. Hence P outputs indices of the form $pad(f(e'', e, v_0, \ldots, v_n, s_0, \ldots, s_n), e, v_0, \ldots, v_n, s_0, \ldots, s_n)$ only finitely often. Therefore P is a partial conservative learner that outputs at least one r.e. index e' with $W_{e'} = \mathcal{L}$ infinitely often, and if $W_{e''} \neq \mathcal{L}$, then P outputs e'' only finitely often. It remains to construct a recursive learner N which, in addition to being partially conservative, outputs exactly one correct index infinitely often if T were a presentation of some $\mathcal{L}$ in $\mathcal{C}$. This may be done by considering another padding function pad_1, where $pad_1(j, t)$ is an index for W_j, simulating the learner P, and setting N to output $pad_1(j, t)$ at least n times if and only if there is a stage $s \geq t$ such that P outputs j at least n times and t is the last stage at which P outputs some index i with $i < j$ up to stage t. N is then the desired partial conservative learner of $\mathcal{C}$.

For the converse direction of the proof, suppose that M is a partial conservative learner of $\mathcal{C}$. To construct a new $Ex[\mathbb{K}]$ learner N, let N be fed with the input $\sigma = a_0 \circ a_1 \circ \ldots \circ a_n$; N identifies via the oracle $\mathbb{K}$ the least member e of $\{M(\tau) : \tau \preceq a_0 \circ a_1 \circ \ldots \circ a_n\}$ for which $\text{range}(\sigma) - \{\#\} \subseteq W_e$.

N then outputs the index e', where $W_{e'}^{\mathbb{K}} = W_e$ if there exists a least number e which satisfies the preceding condition, and $W_{e'}^{\mathbb{K}} = \emptyset$ if such a number e cannot be found. Suppose that N is presented with a text $T = a_0 \circ a_1 \circ a_2 \circ \ldots$ for some $\mathcal{L} \in \mathcal{C}$. Since M partially conservatively learns $\mathcal{L}$, it outputs on T exactly one index

e with $W_e = \mathcal{L}$ infinitely often, and for all other indices $d \neq e$ that it outputs, $\mathcal{L} \not\subseteq W_d$. Let σ be the shortest prefix of T such that $M(\sigma) = e$. For each proper prefix τ of σ, there is a sufficiently long segment $a_0 \circ a_1 \circ \ldots \circ a_s$ of T such that $\{a_0, a_1, \ldots, a_s\} - \{\#\} \not\subseteq W_\tau$, and so the required condition is not met. On the other hand, as range$(T) - \{\#\} = W_e$, the index e is a valid candidate at every stage, implying that N will converge to a unique index e' with $W_{e'}^{\mathbb{K}} = W_e$ in the limit. Hence N is an $Ex[\mathbb{K}]$ learner of $\mathcal{C}$, as was to be shown. In conclusion, a class $\mathcal{C}$ of infinite sets is partially conservatively learnable if and only if it is $Ex[\mathbb{K}]$ learnable. ♦

The example furnished below shows that in the above theorem, the condition that the class of languages to be learnt must be infinite is indeed a necessary hypothesis. Further, the subsequent example gives that partial conservative learnability is weaker than learnability relative to oracles whose degrees are Turing above $\mathbb{K}$.

Theorem 29 *The class* $\mathcal{C} = \{\{e+x : x \in \mathbb{N}\} : e \in \mathbb{N}\} \cup \{\{e+x : x \leq d\} : e \in \mathbb{K} - \mathbb{K}_d\}$ *is explanatorily learnable but not partially conservatively learnable.*

Proof. A programme for an explanatory learner M of $\mathcal{C}$ is as follows: on the input σ with $e = \min(\{x : x \in \text{range}(\sigma)\})$ and $e + d = \max(\{x : x \in \text{range}(\sigma)\})$, M conjectures an index for the set $\{e + x : x \in \mathbb{N}\}$ if $e \notin \mathbb{K}_{|\sigma|}$ or if $e \in \mathbb{K}_d$, and an index for the set $\{e + x : x \leq d\}$ if $e \in \mathbb{K}_{|\sigma|} - \mathbb{K}_d$. Suppose that M is fed with a text for the set $\{e + x : x \in \mathbb{N}\}$. If $e \notin \mathbb{K}$ then M will always output an index for the correct set. If $e \in \mathbb{K}_{s+1} - \mathbb{K}_s$, then M will converge to a correct index once the element $e + s + 1$ occurs in a segment of the text of length at least s. On the other hand, if M processes a text of the set $\{e + x : x \leq d\}$ with $e \in \mathbb{K}_s - \mathbb{K}_d$ for some

$s > d$, then it will also converge to a correct index after the sth stage.

For the sake of a contradiction, suppose that N were a partial conservative learner of $\mathcal{C}$. Define a recursive function f by setting $f(e)$ to be the first number d found such that $\{e, e+1, \ldots, e+d+1\} \subseteq W_{N(e \circ e+1 \circ \ldots \circ e+d)}$. Since N learns the set $\{e+x : x \in \mathbb{N}\}$, such a number d must exist, and so f is a recursive function. Furthermore, owing to the partial conservativeness of N, it follows that $e \in \mathbb{K}$ holds if and only if $e \in \mathbb{K}_{f(e)}$. This provides a recursive procedure for the halting problem, which is a contradiction. Thus N cannot be a partial conservative learner of $\mathcal{C}$, as required. ♦.

Theorem 30 *The class of infinite sets* $\mathcal{C} = \{\{e\} \oplus (W_e \cup D) : D$ *is finite and* W_e *is cofinite*$\}$ $\cup \{\{e\} \oplus \mathbb{N} : e \in \mathbb{N}\}$ *is* $Ex[\mathbb{K}']$ *learnable but not partially conservatively learnable.*

Proof. An $Ex[\mathbb{K}']$ learner M may be programmed as follows: on the input σ, if $2e$ is the minimum even number in the range of σ, M checks relative to the oracle $\mathbb{K}'$ whether or not there is a minimum $x < |\sigma|$ such that the Π^0_2 condition $\forall y > x \exists s[y \in W_{e,s}]$ holds. If such a number x does not exist, M conjectures the set $\{e\} \oplus \mathbb{N}$; if x is the minimum such number, then M again accesses $\mathbb{K}'$ to determine the finite set $D_\sigma = \{z \leq x : z \in \mathrm{range}(\sigma) - W_e\}$, and conjectures the set $\{e\} \oplus (W_e \cup D_\sigma)$. Otherwise, if no such e is found, M outputs a default index 0.

Suppose that M is presented with a text T for the set $\{e\} \oplus \mathbb{N}$. First, assume that W_e is cofinite. Then there is a least number x such that for all $y > x$, y is contained in W_e. Further, for a sufficiently long segment σ of the text, $\{z \leq x : z \notin W_e\} \subseteq \mathrm{range}(\sigma)$ and $|\sigma| > x$ both hold. Hence M will converge on T to a fixed index for the set $\{e\} \oplus \mathbb{N}$. Secondly, assume that W_e is coinfinite. In this case, the condition $\forall y > x \exists s[y \in W_{e,s}]$ fails to hold for all x, and so M will conjecture the set $\{e\} \oplus \mathbb{N}$ on all segments of T. Next, suppose that M is fed with a text T' for the set $\{e\} \oplus (W_e \cup D)$, where W_e is cofinite and D is finite. Let x be the minimum number such that for all $y \geq x$, $y \in W_e$ holds. Then, upon witnessing a segment σ of T' with $|\sigma| \geq x$ which contains all the elements of D, M will thenceforth always conjecture a fixed index for $\{e\} \oplus (W_e \cup D)$. Therefore M is an $Ex[\mathbb{K}']$ learner of $\mathcal{C}$, as required.

On the other hand, assume for the sake of a contradiction that N were a partial conservative learner of $\mathcal{C}$. Fix any number e, and load the text $2e \circ 1 \circ 3 \circ 5 \circ \ldots \circ (2n+1) \circ \ldots$ into N. Since N partially learns the set $\{e\} \oplus \mathbb{N}$, there is a least number k such that N outputs an index for $\{e\} \oplus \mathbb{N}$ on the segment $2e \circ 1 \circ \ldots \circ 2k+1$; moreover, one can search for k by means of the oracle $\mathbb{K}'$. One may subsequently check relative to $\mathbb{K}'$ whether or not $\forall z > k \exists s[z \in W_{e,s}]$ holds. If it does hold, then W_e is cofinite; otherwise, W_e must be coinfinite, for if W_e were cofinite and $z > k$ were a number such that $z \notin W_e$, then the segment $2e \circ 1 \circ \ldots \circ 2k+1$ may be extended to a text for $\{e\} \oplus (W_e \cup \{0, 1, \ldots, k\})$, and since N outputs an index for some set of which $\{e\} \oplus (W_e \cup \{0, 1, \ldots, k\})$ is a proper subset, this implies that N cannot partially conservatively learn $\{e\} \oplus (W_e \cup \{0, 1, \ldots, k\})$, contrary to hypothesis. Thus the initial assumption would lead to a decision procedure relative to $\mathbb{K}'$ for the Π^0_3-complete set $\{e : W_e \text{ is coinfinite}\}$, a contradiction. In conclusion, $\mathcal{C}$ is not partially conservatively learnable, as required. ♦

As a conclusion to the present section, the last result shows that Theorem 28 does not hold generally for every hypothesis space.

Theorem 31 *The class of infinite sets* $\mathcal{D} = \{\{e\} \oplus \{0, 1, \ldots, d\} \oplus \mathbb{N} : e \in \mathbb{K} - \mathbb{K}_d\} \cup \{\{e\} \oplus \mathbb{N} \oplus \mathbb{N} : e \in \mathbb{N}\}$ *is explanatorily learnable but not partially conservatively learnable using* $\mathcal{D}$ *as the hypothesis space.*

Proof. An explanatory learner M may work as follows: on the input σ with $3e = \min(\{3x : 3x \in \text{range}(\sigma)\})$ and $\{3x + 1 : x \leq d\} \subseteq \text{range}(\sigma)$, M conjectures the set $\{e\} \oplus \{0, 1, \ldots, d\} \oplus \mathbb{N}$ if $e \in \mathbb{K}_{|\sigma|}$, and conjectures $\{e\} \oplus \mathbb{N} \oplus \mathbb{N}$ if $e \notin \mathbb{K}_{|\sigma|}$, or if the number e does not exist, or if there is no number $3x + 1 \in \text{range}(\sigma)$. An argument analogous to that in the preceding claim shows that $\mathcal{D}$ cannot be partially conservatively learnt using $\mathcal{D}$ as the hypothesis space: otherwise, if N were a partial conservative learner, one may define a recursive function f which, on input e, searches for the first number d such that $\{3e\} \cup \{3x + 1 : x \leq d + 1\} \subseteq W_{N(3e \circ 1 \circ 2 \circ 4 \circ 5 \circ \ldots \circ 3d+1 \circ 3d+2)}$. Due to the condition that N only outputs indices of sets in $\mathcal{D}$, it must hold that if d is the first such number found, then $\{e\} \oplus \{0, 1, \ldots, d + 1\} \oplus \mathbb{N} \subseteq W_{N(3e \circ 1 \circ 2 \circ 4 \circ 5 \circ \ldots \circ 3d+1 \circ 3d+2)}$. Therefore, by the conservativeness of N, $e \in \mathbb{K}$ holds if and only if $e \in \mathbb{K}_d$, a contradiction. ♦

4 Partial Learning of Classes of Recursive Functions

4.1 Confident Partial Learning

This section deals with partial learning of recursive functions. In a manner of speaking, a text for a recursive function, whether canonical or arbitrary, conveys more information than that for a language, since the learner progressively gains knowledge about the graph of the target recursive function as well as its complement.

That vacillatory learnability generally implies explanatory learnability in the case of learning recursive functions but not for language learning, as proved in Theorem 41, lends some weight to this heuristic observation. Nonetheless, a few of the relations between confident partial learning and other learning success criteria that have been established so far in the context of language learning also hold for recursive function learning. To exemplify this point, the section's first theorem gives an example of a behaviourally correctly learnable class of recursive functions which is not confidently partially learnable.

Theorem 32 *There is a behaviourally correctly learnable class of recursive functions which is not confidently partially learnable.*

Proof 1. Let $\sigma_0, \sigma_1, \ldots$ be an enumeration of all binary strings. Define, for each $e \in \mathbb{N}$, the Π^1_0 class $C_e = \{A \subseteq \mathbb{N} : \forall x \in W_e \exists y[\sigma_x(y) \neq A(y)]\}$. Set

$$\mathcal{F} = \{B \subseteq \mathbb{N} : \exists e \forall y \leq e \forall z \exists A \in C_e[B(y) = 0 \wedge B(e+1) = 1 \wedge B(z+e+2) = A(z) \wedge A \text{ is isolated}]\}.$$

It shall be shown that $\mathcal{F}$ is behaviourally correctly learnable but not confidently partially learnable. A behaviourally correct learner M may perform as follows: on the input σ, M first identifies the number e such that $0^e \circ 1 \preceq \sigma$; if no such e exists, M outputs 0. Otherwise, let $\sigma = 0^e \circ 1 \circ \tau$; M then outputs the index i for which

$$\varphi_i(x) = \begin{cases} \sigma(x) & \text{if } x \leq |\sigma| - 1; \\ \eta(x) & \text{if } \tau \preceq \eta \wedge \forall \theta \in \{0,1\}^*[\theta \preceq \eta \wedge \\ & \sigma_x = (1-\theta(0)) \circ (1-\theta(1)) \circ \ldots \circ (1-\theta(|\theta|-1)) \Rightarrow x \in W_e]. \end{cases}$$

Suppose that M is fed with a text for B, which is of the form $0^e \circ 1 \circ A$, where A

is an isolated member of C_e. There is a binary string σ_x such that A is the unique member of C_e which extends σ_x. This means that for all $\sigma_x \circ \eta \preceq A$, if $\sigma_y = \sigma_x \circ \eta \circ o$, where $o \in \{0,1\}$, then $y \in W_e \Leftrightarrow A(|\sigma_x| + |\eta|) = 1 - o$. Thus when a sufficiently long segment of the text is revealed to M, of which σ_x is a prefix, M will converge semantically to a correct index for the characteristic function of B.

Assume now by way of contradiction that N were a confident partial learner of $\mathcal{F}$. For each $e \in \mathbb{N}$, an r.e. set $W_{f(e)}$ shall be built so that there are only finitely many infinite branches A with A in $C_{f(e)}$, and N outputs some index d infinitely often on at least two of these branches subjoined to the string $0^{f(e)} \circ 1$. $W_{f(e)}$ is constructed in stages, according to the following algorithm.

- At stage 0, set $W_{f(e),0} = \emptyset$.
- At stage $s+1$, put
  $S_*^{s+1} = \{0,1\}^{s+1} - \{\sigma \in \{0,1\}^* : \exists \tau \preceq \sigma[\tau \in W_{f(e),s}]\}$, where
  $\tau \in W_{f(e),s}$ denotes that if $\sigma_x = \tau$, then $x \in W_{f(e),s}$. Let
  $S_*^{s+1} = \{\eta_0, \eta_1, \ldots, \eta_n\}$, where
  $N(0^e \circ 1 \circ \eta_0) \leq N(0^e \circ 1 \circ \eta_1) \leq \ldots \leq N(0^e \circ 1 \circ \eta_n)$.
- For $m = 0, 1, \ldots, n$, determine whether there exists a shortest prefix τ of η_m such that the number of prefixes θ of τ for which $\theta \circ 0$ and $\theta \circ 1$ are each extended by some element of S_*^{s+1} is equal to $N(0^e \circ 1 \circ \eta_m) + 2$. If such a τ exists, remove all η_k with $k > m$ such that $\tau \preceq \eta_k$ from S_*^{s+1}; denote the new set of strings by S^{s+1}, and proceed to the next value of m. Otherwise, proceed to the next value of m.
- Put all strings removed from S_*^{s+1} during the preceding steps into $W_{f(e),s}$.

By Kleene's Recursion Theorem, there is an e for which $W_e = W_{f(e)}$. Fix any such number e. Consider the set of binary strings $S = \bigcup_{s \in \mathbb{N}} S^{s+1}$: by the above construction, $\sigma \notin S \Rightarrow \exists \sigma_x[\sigma \preceq \sigma_x \wedge x \in W_{f(e)}]$, so that by the first step of the algorithm, $\sigma\tau \notin S$ for all $\sigma, \tau \in \{0,1\}^*$. This means that S is a recursive tree whose infinite branches are the set elements of $C_{f(e)}$. Furthermore, as $W_{f(e),0} = \emptyset$, both $\eta_0 \circ 0$ and $\eta_0 \circ 1$ are contained in S_*^2, where η_0 is as defined in the second step of the algorithm at stage 1. It thus follows inductively that the set S_*^{s+1} is nonempty for all $s \in \mathbb{N}$, so that S must be an infinite tree. Consequently, by König's Lemma, S contains at least one infinite branch, say A.

Suppose that N is fed with a text for the recursive function represented by $0^e \circ 1 \circ A$. By the confidence of N, there is an index d and infinitely many prefixes σ of A such that $N(0^e \circ 1 \circ \sigma) = d$. As each number $e < d$ is output only finitely often, $N(0^e \circ 1 \circ \sigma) \geq d$ for almost all prefixes σ of A. Moreover, one may argue by induction that there are at least $d+1$ different infinite branches A' that branch off from A, as follows. Let τ be a prefix of A such that $N(0^e \circ 1 \circ \tau \circ A(|\tau|) \ldots A(|\tau|+k)) \geq d$ for all $k \geq 0$. Assume first that there are at least $d+1$ prefixes $\theta_0, \theta_1, \ldots, \theta_d, \ldots$ of τ such that for all i, $\theta_i \circ 0$ and $\theta_i \circ 1$ are each extended by an element of $S_*^{|\tau|}$. From the second step of the algorithm at stage $|\tau|$, it follows that $d+1$ strings in $S_*^{|\tau|}$ that contain $\theta_0, \theta_1, \ldots, \theta_d$ as prefixes are preserved in $S^{|\tau|}$, and if σ_k is such a string, then $\sigma_k \circ 0$ and $\sigma_k \circ 1$ are both contained in $S_*^{|\tau|+1}$. Therefore at stages $|\tau|, |\tau|+1, |\tau|+2, \ldots$, there are at least $d+1$ strings in $S^{|\tau|}, S^{|\tau|+1}, S^{|\tau|+2}, \ldots$ respectively, such that each of these strings is a segment of a unique infinite branch. Hence there are at least $d+1$ different infinite paths branching off from A. If, on the other hand, there are less than $d+1$ prefixes θ of τ for which $\theta \circ 0$ and $\theta \circ 1$ are each extended by a string in $S_*^{|\tau|}$, then the second step of the algorithm for τ will be skipped, and $\tau \circ 0, \tau \circ 1$ proceed accordingly to the next stage $|\tau| + 1$. This process will continue until there is a stage $k > |\tau|$ with at least $d + 1$ strings of length k branching off from A; one can now follow the argument of the preceding case to conclude that there must be at least $d + 1$ different infinite branches that share a common prefix with A.

Now let α be a prefix of A such that $|\alpha|$ is the first stage at which $S_*^{|\alpha|}$ contains at least $d+2$ prefixes $\tau_0, \tau_1, \ldots, \tau_{d+1}$ branching off from A and $N(0^e \circ 1 \circ \alpha) = d$. By the second step of the algorithm, the string in $S_*^{|\alpha|}$ extending τ_{d+1} will be removed at the end of stage $|\alpha|$, so that $S^{|\alpha|}$ is left with exactly $d + 1$ strings that branch off from A. This implies that every infinite branch of S is isolated; that is, for each infinite branch A of S, there is a prefix σ_A of A such that A is the unique branch of S extending σ_A. There can only be finitely many isolated infinite branches of S; denote these branches by $A_0, A_1, \ldots, A_l$. Let p be the maximum number that N outputs infinitely often on each of the canonical texts for $0^e \circ 1 \circ A_0, 0^e \circ 1 \circ A_1, \ldots, 0^e \circ 1 \circ A_l$, and the corresponding infinite branch be A_i. By the argument in the preceding paragraph, there are at least $p+1$ different infinite paths that branch off from A_i; as a consequence, there is a number $q \leq p$ such that N outputs q infinitely often on the canonical texts for at least two of the sets amongst $0^e \circ 1 \circ A_0, 0^e \circ 1 \circ A_1, \ldots, 0^e \circ 1 \circ A_l$. Thus N fails to learn the class $\mathcal{F}$, a contradiction. ♦

The second proof provides yet another example of a behaviourally correctly learnable class of recursive functions which is not confidently partially learnable from canonical text; moreover, the proof suggests a necessary condition on the computational power of confident learners that can partially learn all recursive functions. An indispensable ingredient in the proof is the existence of a *low*, *PA-complete* set, which was first proved by Jockush and Soare [14] as a corollary of a more general result on $\prod_1^0$ classes. The relevant properties of such a set utilised in the proof, together with other related concepts, are briefly reviewed below.

Definition. A class of sets is a $\prod_1^0$ *class* if it is the set of infinite branches of some infinite recursive binary tree. If P is a recursive predicate, then the class of sets A such that $(\forall x)P(c_A(x))$ is a $\prod_1^0$ class.

Shoenfield [26] showed that, for any consistent axiomatizable theory T_1, the set A of complete extensions of T_1 which have the same symbols as T_1 is non-empty, and that every $\alpha \in A$ can be written in the form $(\forall x)R(gn(\alpha(x)))$ with R recursive; here $gn(\alpha(x))$ denotes the Gödel number of $\alpha(x)$. In other words, by the above definition, the set of complete extensions of a given consistent theory is a nonempty $\prod_1^0$ class. Conversely, Jockusch and Soare [14], as well as Hanf [11], showed that the class of degrees of members of a given $\prod_1^0$ class coincides with the class of degrees of complete extensions of some finitely axiomatizable first-order theory; a set which falls within the latter class is known as *PA-complete*. An equivalent definition of a set A being PA-complete, which is explicitly applied in the next proof of Theorem 32, is that given any partial-recursive and $\{0,1\}$-valued function ψ, one can compute relative to A a total extension Ψ of ψ.

Definition. A set A is *low* if $A' \equiv_T \mathbb{K}$.

The specific result of Jockusch and Soare required for the proof of the subsequent theorem is the following.

Theorem 33 *[14] Any consistent axiomatizable theory (in particular, Peano Arithmetic (P.A.)) has a complete extension of degree whose jump is* $\mathbb{K}'$.

To put Theorem 33 in another way: there exists a low, PA-complete set.

Proof 2. The class of recursive functions

$\mathcal{C} = \{f : f$ is recursive and $\{0,1\}$-valued $\wedge\ \exists e[|\overline{W}_e| < \infty \wedge f(e+1) = 1$
$\wedge\, \forall x \leq e[f(x) = 0] \wedge f =_* \varphi_e]\}$

is behaviourally correctly learnable but not confidently partially learnable.

A behaviourally correct learner M outputs a default index 0 until it witnesses the first number e such that $f(x) = 0$ for all $x \leq e$ and $f(e+1) = 1$; subsequently, on the input $\sigma = 0^e \circ 1 \circ f(e+2) \circ \ldots \circ f(e+k)$, it conjectures the index i with

$$\varphi_i(x) = \begin{cases} \sigma(x) & \text{if } x < |\sigma|; \\ \varphi_e(x) & \text{if } x \geq |\sigma|. \end{cases}$$

Suppose that M is fed with the canonical text for a recursive function f from the class to be learnt. Let e be the index such that $f(e+1) = 1$ and $f(x) = 0$ for all $x \leq e$, and n be the least number with $\varphi_e(x)\downarrow = f(x)$ for all $x > n$. The preceding algorithm ensures that if M witnesses a segment of the text with length at least $\max(e+1, n)$, then it will output a correct index for f. Hence M is indeed a BC learner of $\mathcal{C}$.

Assume by way of contradiction that one may define a recursive confident partial learner N of the class $\mathcal{C}$. It shall be shown that this implies the existence of a $\mathbb{K}'$-recursive procedure for deciding whether $d \in \{e : W_e$ is cofinite$\}$ for any given d, contradicting the known fact that the latter set is Σ^0_3-complete. First, let g be a recursive function for which $\varphi_{g(d)}$ is defined in stages as follows:

- Set $\varphi_{g(d),0}(x)\uparrow$ for all x. Initialise the markers $a_0, a_1, a_2, \ldots$ by setting $a_{i,0} = \langle i, 0\rangle + d + 1$ for $i \in \mathbb{N}$.

- At stage $t+1$, consider the markers $a_{0,t}, a_{1,t}, a_{2,t}, \ldots, a_{t,t}$ with $a_{i,t} = \langle i, r\rangle + d + 1$, and perform the following: if neither $\varphi_{g(d),t}$ nor $\varphi_{i,t}$ is defined on the input $\langle i, j\rangle + d + 1$ for $j \in \{0, 1, \ldots, t+1\} - \{r\}$, set $\varphi_{g(d)}(\langle i, j\rangle + d + 1) = 0$; if $\varphi_{i,t}(\langle i, r\rangle + d + 1)$ is defined but $\varphi_{g(d)}(\langle i, r\rangle + d + 1)$ is not defined, then set $\varphi_{g(d)}(\langle i, r\rangle + d + 1) = 1 - \varphi_{i,t}(\langle i, r\rangle + d + 1)$.

  Furthermore, update $a_{i,t+1} = \langle i, t+1\rangle + d + 1$ if and only if $r \leq t$ and $|\{0, 1, \ldots, r\} - W_{d,t}| < i$.

  Let $\varphi_{g(d),t+1}(x) = \varphi_{g(d),t}(x)$ for all x with $\varphi_{g(d),t}(x)\downarrow$.

It shall be shown that the partial-recursive function $\varphi_{g(d)}$ as defined above possesses the following properties:

1. If W_d is cofinite, then there is an i_0 for which the markers $a_{i,t}$ move infinitely often if and only if $i \geq i_0$, so that $W_{g(d)}$ is also cofinite.

2. If W_d is coinfinite, then the markers $a_{i,t}$ move only finitely often, and there is no total recursive function extending $\varphi_{g(d)}$.

1. follows because if W_d is cofinite, and $|\overline{W}_d| = k$, then for all $i > k$ and each r, there is a t large enough so that $|\{0, 1, \ldots, r\} - W_{d,t}| < i$. This means that for all $i > k$, the markers $a_{i,t}$ move infinitely often. Moreover, this implies that $W_{g(d)}$ is cofinite, for each stage ensures that $\varphi_{g(d)}$ is defined on all inputs $\langle i, j\rangle + d + 1$ for which $j < r$, and since $a_{i,t}$ is shifted to $\langle i, r\rangle + d + 1$ for arbitrarily large values of r for all $i > k$, $\varphi_{g(d)}$ eventually becomes defined on all inputs $\langle i, j\rangle + d + 1$ for $i > k$ and

$j \in \mathbb{N}$. For $i \leq k$, suppose that the markers $a_0, a_1, \ldots, a_k$ settle down permanently on the values $\langle 0, r_0 \rangle + d + 1, \langle 1, r_1 \rangle + d + 1, \ldots, \langle k, r_k \rangle + d + 1$ respectively; by the algorithm, while $\varphi_{g(d)}$ remains undefined on all of these inputs, $\varphi_{g(d)}$ is, however, defined for all $\langle i, j \rangle + d + 1$ with $i \leq k$ and $j > r_i$. Thus $W_{g(d)}$ is indeed cofinite.

On the other hand, if W_d were coinfinite, then for each fixed i there are r, t sufficiently large so that $|\{0, 1, \ldots, r\} - W_{d,t}| \geq i$. At stage $t + 1$, each marker $a_i = \langle i, r \rangle + d + 1$ is updated to a new value $\langle i, t + 1 \rangle + d + 1$ with $t + 1 > r$ if $|\{0, 1, \ldots, r\} - W_{d,t}| < i$; for this reason, there will eventually be a stage s at which $|\langle 0, 1, \ldots, u\} - W_{d,s}| \geq i$, when $a_{i,s} = \langle i, u \rangle + d + 1$, and the inequality would continue to hold at all subsequent stages, in turn implying that the value of a_i will be permanently fixed as this value. Furthermore, if φ_i were a total function, then there will be a stage s' at which $\varphi_{i,s'}(\langle i, u \rangle + d + 1)$ is defined, and the algorithm would secure that $\varphi_{g(d)}(\langle i, u \rangle + d + 1)$ differs from the value of $\varphi_{i,s'}(\langle i, u \rangle + d + 1)$. Therefore there cannot be a total recursive function extending $\varphi_{g(d)}$.

Now let A be a PA-complete set which is low, that is, every partial-recursive $\{0, 1\}$ function may be extended to an A-recursive function, and, in addition, $A'' \equiv_T \mathbb{K}'$. Furthermore, let $\varphi^A_{f(d)}$ be a uniformly A-recursive extension of the partial-recursive function $\varphi_{g(d)}$ such that $\varphi^A_{f(d)}$ is $\{0, 1\}$-valued. There is a further recursive function h for which

$W^A_{h(d,e)} = \{n : N$ outputs e at least n times on the text $0^{g(d)} \circ 1 \circ \varphi^A_{f(d)}(g(d) + 2)$ $\circ \varphi^A_{f(d)}(g(d) + 3) \circ \ldots\}$. Owing to the confidence of N, one can determine by means of the oracle A'' the unique e such that $W^A_{h(d,e)}$ is infinite.

If W_d were cofinite, then, as was shown above, $\varphi_{g(d)}$ is also cofinite, and so $\varphi^A_{f(d)}$ is a total recursive extension of $\varphi_{g(d)}$, that is, $\varphi_{g(d)} =_* \varphi^A_{f(d)}$. Therefore N learns

the recursive function generating the text
$0^{g(d)} \circ 1 \circ \varphi^A_{f(d)}(g(d)+2) \circ \varphi^A_{f(d)}(g(d)+3) \circ \ldots$, and consequently $\varphi_e(x) = \varphi^A_{f(d)}(x)$ for all $x \geq g(d)+2$.

However, if W_d were coinfinite, it follows from the construction of $\varphi_{g(d)}$ that there is no total recursive function extending $\varphi_{g(d)}$, giving that $\varphi_e \neq \varphi^A_{f(d)}$, or more specifically, there is an $x \geq g(d)+2$ such that either $\varphi_e(x)\uparrow$ or $\varphi_e(x)\downarrow \neq \varphi^A_{f(d)}(x)\downarrow$.

Hence W_d is cofinite if and only if for all $x \geq g(d)+2$, $\varphi_e(x)\downarrow = \varphi^A_{f(d)}(x)\downarrow$. As this condition may be checked using the oracle A'', and A'' is Turing equivalent to $\mathbb{K}'$, it may be concluded that $\{d : W_d \text{ is cofinite}\} \equiv_T \mathbb{K}'$, which is the desired contradiction. Therefore the class $\mathcal{C}$ cannot be confidently partially learnt. ♦

A review of the second proof of Theorem 32 produces the following corollary. This may be a first step towards characterising the Turing degrees of oracles relative to which all recursive functions can be confidently partially learnt.

Theorem 34 *There is a behaviourally correctly learnable class $\mathcal{C} \subseteq REC_{0,1}$ such that $\mathcal{C}$ is confidently partially learnable relative to B only if $B'' \geq_T \mathbb{K}''$.*

Proof. Consider the class

$$\mathcal{C} = \{f : f \text{ is recursive and } \{0,1\}\text{-valued } \wedge\ \exists e[|\overline{W}_e| < \infty \wedge f(e+1) = 1$$
$$\wedge\, \forall x \leq e[f(x) = 0] \wedge f =_* \varphi_e]\}$$

which was demonstrated to be behaviourally correctly learnable but not confidently partially learnable in the second proof of Theorem 32. In the proof that $\mathcal{C}$ is not confidently partially learnable, it was seen in the last paragraph that there is a low, PA-complete set A such that for all d, W_d is cofinite if and only if there is an A-recursive total extension $\varphi^A_{f(d)}$ of the partial-recursive function $\varphi_{g(d)}$, and a confident partial learner N that outputs e infinitely often on the text $0^{g(d)} \circ 1 \circ \varphi^A_{f(d)}(g(d)+2) \circ \varphi^A_{f(d)}(g(d)+3) \circ \ldots$, such that for all $x \geq g(d)+2$, $\varphi_e(x)\downarrow = \varphi^A_{f(d)}(x)\downarrow$. Suppose that the confident partial learner N is endowed with an oracle B. This implies that the index e that N outputs infinitely often on the text $0^{g(d)} \circ 1 \circ \varphi^A_{f(d)}(g(d)+2) \circ \varphi^A_{f(d)}(g(d)+3) \circ \ldots$ may be determined relative to the oracle B'', since the condition $\forall s \exists s' > s[N(0^{g(d)} \circ 1 \circ \varphi^A_{f(d)}(g(d)+2) \circ \ldots \circ \varphi^A_{f(d)}(g(d)+s')) = e]$ is B''-recursive. Moreover, as $A'' \equiv_T \mathbb{K}'$, it can be checked relative to $\mathbb{K}'$ whether or not $\varphi_e(x)\downarrow = \varphi^A_{f(d)}(x)$ holds for all $x \geq g(d)+2$. Therefore $\{d : W_d \text{ is cofinite}\} \leq_T \mathbb{K}' \oplus B''$, and as $\mathbb{K}' \leq_T B''$, one has $\{d : W_d \text{ is cofinite}\} \leq B''$. Finally, from the fact that $\{d : W_d \text{ is cofinite}\} \equiv_T \mathbb{K}''$, it may be concluded that $\mathbb{K}'' \leq_T B''$, as was to be shown. ♦

To complement Theorem 32, we now show that, similar to the case of language learning, behaviourally correct learning of recursive functions is not a more severe criterion than confident partial learning. Thus, both of these learnability criteria have incomparable learning strengths.

Theorem 35 *There is a class of recursive functions which is confidently partially learnable but not behaviourally correctly learnable with respect to a canonical text.*

Proof 1. Consider the class of recursive functions

$$\mathcal{C} = \{f : \forall x[f(0)\downarrow \wedge \varphi_{f(0)}(x)\downarrow = f(x)]\} \cup \{f : \forall x[f(x)\downarrow \wedge \exists y \forall z > y[f(z) = 0]]\};$$

the class $\mathcal{C}$ is the union of all self-describing recursive functions together with all recursive functions that are almost everywhere equal to 0. A confident partial learner M of $\mathcal{C}$ may be defined as follows: on the input $f(0)\circ f(1)\circ\ldots\circ f(n)$, M distinguishes two cases:

- There exists a minimum number k such that for all x with $k \leq x \leq n$, $f(x) = 0$. M then conjectures an index i for which

$$\varphi_i(y) = \begin{cases} f(y) & \text{if } y < k; \\ 0 & \text{if } y \geq k. \end{cases}$$

- For all x with $0 \leq x \leq n$, there is a $k > x$ and $k \leq n$ for which $f(k) \neq 0$. M then conjectures the index $f(0)$.

To verify that M is a confident partial learner of $\mathcal{C}$, suppose first that M is fed with the canonical text $f(0) \circ f(1) \circ f(2) \circ f(3) \circ \ldots$ for a total function f such that there is a minimum number k with $f(x) = 0$ whenever $x > k$. In accordance with the learning algorithm, M then converges syntactically to an index i for the recursive function φ_i that is equal to $f(x)$ for all $x \leq k$, and equal to 0 for all $x > k$. Secondly, suppose that $f(x) = \varphi_{f(0)}(x)$ for all x, and, in addition, there are infinitely many x with $f(x) \neq 0$. This implies that the second case in the learning algorithm holds infinitely often, so that the learner M will output $f(0)$ infinitely often, and every other index only finitely often. Furthermore, M is confident on every text, as it will output the index $f(0)$ infinitely often if $f(x) \neq 0$ for almost all x; otherwise, if there exists a minimum number k for which $f(x) = 0$ whenever $x > k$, then M converges syntactically to an index i such that $\varphi_i(x) = f(x)$ for all $x \leq k$, and $\varphi_i(x) = 0$ for all $x > k$. Hence M is a confident partial learner of $\mathcal{C}$.

Next, assume by way of contradiction that N were a BC-learner of $\mathcal{C}$. For each number e, one may construct a recursive function $\varphi_{g(e)}$ in stages as follows.

- Set $\varphi_{g(e)}(0) = e$.
- At stage $s+1$, assume inductively that $\varphi_{g(e)}(x)$ has been defined for all $x \leq k$. Let $\sigma_s = \varphi_{g(e)}(0) \circ \varphi_{g(e)}(1) \circ \ldots \circ \varphi_{g(e)}(k)$. Run a search for a pair of numbers p_{s+1}, q_{s+1}, such that $\varphi_{N(\sigma_s \circ 0^{p_{s+1}} \circ 1 \circ 0^{q_{s+1}})}(|\sigma_s|+p_{s+1}) \neq \varphi_{N(\sigma_s \circ 0^{p_{s+1}})}(|\sigma_s|+p_{s+1})$. Then define $\varphi_{g(e)}(x) = 0$ if $|\sigma_s| \leq x \leq |\sigma_s| + p_{s+1} - 1$ or $|\sigma_s| + p_{s+1} + 1 \leq x \leq |\sigma_s| + p_{s+1} + q_{s+1} - 1$, and $\varphi_{g(e)}(|\sigma_s| + p_{s+1}) = 1$. This condition imposes the requirement that $\varphi_{g(e)}$ be defined so that N makes a semantic mind change between the stages where it has seen the text segments $\sigma_s \circ 0^{p_{s+1}}$ and $\sigma_s \circ 0^{p_{s+1}} \circ 1 \circ 0^{q_{s+1}}$.

Since N BC-learns every recursive function which is almost everywhere equal to 0, the inductive step in the construction of $W_{g(e)}$ always terminates successfully. For, given any text segment σ_s at stage $s+1$, there is a number p_{s+1} such that $\varphi_{N(\sigma_s \circ 0^{p_{s+1}})}(x) = 0$ for all $x \geq |\sigma_s|$; fixing any such number p_{s+1}, it follows along an analogous line of reasoning that there is another number q_{s+1} for which $\varphi_{N(\sigma_s \circ 0^{p_{s+1}} \circ 1 \circ 0^{q_{s+1}})}(x) = 1$ when $x = |\sigma_s| + p_{s+1}$. Thus N makes a semantic mind change between the text segments $\sigma_s \circ 0^{p_{s+1}}$ and $\sigma_s \circ 0^{p_{s+1}} \circ 1 \circ 0^{q_{s+1}}$, as required.

Owing to Kleene's Recursion Theorem, there are infinitely many indices e such that $\varphi_{g(e)} = \varphi_e$. Fix any such number e. As a consequence of the inductive step in the construction of $\varphi_{g(e)}$, there are infinitely many y for which $\varphi_{N(\varphi_{g(e)}(0) \circ \varphi_{g(e)}(1) \circ \ldots \circ \varphi_{g(e)}(y))}(x) \neq$

$\varphi_{g(e)}(x)$ for some number x. This in turn implies that N cannot BC-learn the self-describing recursive function φ_e, a contradiction. ♦

Proof 2. Blum and Blum's Non-Union Theorem [3] provides classes $\mathcal{C}_1$ and $\mathcal{C}_2$ which are explanatory learnable while their union is not behaviourally correctly learnable. By Theorem 18 the two classes are confidently partially learnable and by Theorem 19 their union $\mathcal{C}_1 \cup \mathcal{C}_2$ is confidently partially learnable as well. ♦

Theorem 32 demonstrates that the class of all total recursive functions is not confidently partially learnable. Nonetheless, there is a less restrictive notion of confident partial learning, somewhat analogous to a blend of behaviourally correct learning and partial learning, that permits the class of all recursive functions to be learnt. This notion of learning is spelt out in the following theorem.

Theorem 36 *There is a recursive learner M such that on every function f there is exactly one partial-recursive function Ψ for which M outputs an index infinitely often, and $f = \Psi$ whenever f is recursive.*

Proof. Let the input function f be presented as a canonical text
$T = f(0) \circ f(1) \circ f(2) \circ f(3) \ldots$; on this text, the recursive learner M performs the following instructions.

1. M outputs e at least n times if and only if there is a stage $s > n$ such that $\varphi_{e,s}(x) \downarrow = f(x)$ for all $x \leq \max(e, n)$.
2. For each number e, suppose $n \geq e$ is found at some stage s so that $\varphi_{e,s}(x) = f(x)$ whenever $x \leq n$. M then outputs an index $g(e, n)$ for the partial-recursive

function $\varphi_{g(e,n)}$ defined by

$$\varphi_{g(e,n)}(x) = \begin{cases} \uparrow & \text{if } \forall d \leq e \exists y \leq n+1[\varphi_d(y)\uparrow \vee \varphi_d(y)\downarrow \neq f(y)]; \\ \varphi_d(x) & \text{if } d \text{ is the least number satisfying } d \leq e \text{ and} \\ & \forall y \leq n+1[\varphi_d(y)\downarrow = f(y)]. \end{cases}$$

It shall be shown that M satisfies the learning criteria specified in the theorem. First, suppose that f is a recursive function. If $\varphi_e \neq f$ and $W_e \neq \emptyset$, then there is a least x_0 such that $\varphi_e(x_0)\uparrow$ or $\varphi_e(x_0)\downarrow \neq f(x_0)$. By the requirements of 1. and 2., this means that every index d with $\varphi_e = \varphi_d$ is output only finitely often. Moreover, whenever $p > x_0$ is an index for φ_e, the condition in 1. that $\varphi_p(x)\downarrow = f(x)$ for all $x \leq p$ guarantees that M does not output p. Hence the partial-recursive function φ_e is conjectured only finitely often. If $W_e = \emptyset$, then, since there is a least index p such that $\varphi_p(x)\downarrow = f(x)$ for all x, the definition of $g(e,n)$ in 2. and the requirement of 1. together ensure that the partial-recursive function φ_e is conjectured for at most a finite number of times. Furthermore, by the requirement of 1., every index e with $f = \varphi_e$ is output infinitely often. Next, suppose that f is not equal to any total recursive function. The output criteria of M specified in 1. alone then gives that for every partial-recursive function φ_e, M outputs an index for φ_e only finitely often. In addition, according to the output criteria of 2., every partial-recursive function which is defined on at least one input is conjectured by M only finitely often. On the other hand, as there are infinitely many numbers d such that $\varphi_d(0)\downarrow = f(0)$, and - owing to the nonrecursiveness of f - for every such d there is a maximum input x such that for some $e \leq d$ and all $y \leq x$, $\varphi_e(y)\downarrow = f(y)$, it follows from 2. that M outputs an index for the partial-recursive function which is everywhere undefined infinitely often. This establishes that M fulfils the learning specifications of the theorem, as required. ♦

The next lemma, in whose proof the padding property of the default hypothesis space $\{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ is pivotal, will be applied in the subsequent theorem.

Lemma 37 *For every A''-recursive function $F^{A''}$, there is an A-recursive function f^A such that for all numbers d, if $F^{A''}(d) = e$, then there is a unique number e' for which there are infinitely many t with $f^A(d,t) = e'$ and $\varphi_e = \varphi_{e'}$.*

Proof. Given that $F^{A''} \leq_T A''$, there exists a sequence of A-recursive approximations $\{f_{i,j}\}_{i,j \in \mathbb{N}}$ such that for all numbers e, $\exists i \forall i' \geq i \exists j \forall j' \geq j [f_{i,j}(e) = F^{A''}(e)]$ holds. One may define an A-recursive function G which satisfies $G(e,t) = pad(e,i)$, for all t, where i is the minimal number for which $\forall i' \geq i \exists j \forall j' \geq j [f_{i',j'}(e) = F^{A''}(e)]$. The A-recursive function G may be constructed in stages as follows. First, let $a_{e,0}, a_{e,1}, a_{e,2}, \ldots$ be an A-recursive sequence in which $pad(d,i)$ occurs at least n times if and only if for all $i' \in \{i, i+1, \ldots, i+n\}$, there are n numbers j' such that $f_{i',j'}(e) = d$. This condition ensures that $pad(d,i)$ occurs in $a_{e,0}, a_{e,1}, a_{e,2}, \ldots$ infinitely often if and only if $d = F^{A''}(e)$, although there still exist $i' > i$ such that $pad(d,i')$ is output infinitely often in the constructed sequence. Next, build a new A-recursive sequence $a'_{e,0}, a'_{e,1}, a'_{e,2}, \ldots$ in which $pad(d,i,s)$ occurs n times if and only if there is a stage $t \geq s$ such that s is the least stage where some number $pad(d,i')$ with $i' < i$ occurs in the sequence $a_{e,0}, a_{e,1}, a_{e,2}, \ldots$ up to stage t and $pad(d,i)$ occurs there at least n times before stage t. This procedure selects the minimal value of i such that $pad(d,i)$ occurs infinitely often in the sequence $a_{e,0}, a_{e,1}, a_{e,2}, \ldots$ constructed above. Subsequently, one may produce a two-valued A-recursive function

G by setting $G(e,t) = a'_{e,t}$ for all such sequences $a'_{e,0}, a'_{e,1}, a'_{e,2}, \ldots$ constructed for each e. By the above construction, the A-recursive function G satisfies the condition that for all e, there is exactly one index e' with $G(e,t) = e'$ for infinitely many t, and, in addition, there is a fixed number i such that $e' = pad(F^{A''}(e), i)$. This establishes the claim. ♦

Having established a necessary condition on the computational power of confident learners that can learn REC, one may hope for an analogous sufficient condition. By means of the above lemma, the theorem below proposes several oracle conditions that, when taken together, enable REC to be confidently partially learnt.

Theorem 38 *If B is low, PA-complete and $A \geq_T B$, $A'' \geq_T \mathbb{K}''$, then there is an A-recursive confident partial learner for REC.*

Proof. The class of all recursive $\{0,1\}$-valued functions, $REC_{0,1}$, is explanatorily learnable by a learner M which outputs B-recursive indices. First, one may construct a numbering $\{\varphi^B_{h(0)}, \varphi^B_{h(1)}, \ldots\}$ of $\{0,1\}$-valued B-recursive functions such that $REC_{0,1} \subset \{\varphi^B_{h(0)}, \varphi^B_{h(1)}, \ldots\}$, and for all e and each input x,

$$\varphi^B_{h(e)}(x) = \begin{cases} 0 & \text{if } \varphi_e(x) \downarrow = 0; \\ 1 & \text{if } \varphi_e(x) \downarrow > 0; \end{cases}$$

as B is PA-complete, there is a B-recursive function g such that each partial B-recursive function $\varphi^B_{h(e)}$ may be extended to a total $\{0,1\}$-valued function $\varphi^B_{g(e)}$. Without loss of generality, assume that $g(d_k) \geq d_k$. The explanatory learner M may be defined by setting M to conjecture, on the input $f(0) \circ f(1) \circ \ldots \circ f(n)$, the least index $g(e)$ for which $\varphi^B_{g(e)}(x) = f(x)$ for all $x \leq n$. Next, let $g(d_0), g(d_1), g(d_2), \ldots$ be the hypotheses issued by M when it is learning some $f \in REC_{0,1}$; according to the learning algorithm of M described above, $d_k = \min\{d : \forall x \leq k[\varphi^B_{g(d)}(x) = f(x)]\}$. Define the B'''-recursive function $F^{B'''}$ by

$$F^{B'''}(g(d_k)) = \begin{cases} e & \text{if } e \text{ is the minimal index with } \varphi_e = \varphi^B_{g(d_k)}; \\ 0 & \text{if there is no index } e \text{ with } \varphi_e = \varphi^B_{g(d_k)}. \end{cases}$$

The B'''-recursive function $F^{B'''}$ produces a new confident partial learner that outputs partial-recursive indices. If there is indeed a recursive $\{0,1\}$-valued function φ_e upon which the text is based, then $F^{B'''}$ outputs the minimal index of φ_e infinitely often; if, on the other hand, no such φ_e exists, then $F^{B'''}$ outputs 0 infinitely often. In either case, all the remaining indices are output only finitely often, and therefore $F^{B'''}$ may be used to construct a confident partial learner. Furthermore, since $B''' \leq_T A''$ by assumption, it follows that $F^{B'''} = F^{A''}$. One can now define a confident partial A-recursive learner N: by means of the claim proved earlier, there is an A-recursive function $f^A(d,t)$ such $f^A(d,t)$ outputs a unique index e' with $\varphi_{e'} = \varphi_{F^{A''}(d)}$ for infinitely many t. N may be set to output $f^A(g(d_k), t)$ if and only if M outputs $g(d_k)$ for the t-th time.

If there is a number e such that $F^{A''}(g(d_k)) = e$ holds for infinitely many k, then e is a partial-recursive index for the recursive $\{0,1\}$-valued function f generating the text revealed to N. In addition, every other index in the range of $F^{A''}(g(d_k))$ is output for only finitely many k. Correspondingly, N outputs a single r.e. index e' for f infinitely often; for each of the other numbers a in the range of $F^{A''}$, as there are only finitely many stages t at which M hypothesises $g(d_k)$ if $a = F^{A''}(g(d_k))$,

$f^A(g(d_k), t)$ is output for finitely many t. This establishes that N is an A-recursive confident partial learner of $REC_{0,1}$.

One can further generalise the preceding result to construct a learner P that confidently partially learns REC relative to A. There is a uniformly B-recursive numbering $B_0, B_1, B_2, \ldots$ such that for all $x \in \mathbb{N}$, if $\varphi_e(x) \downarrow$, then $\langle x, \varphi_e(x)\rangle \in B_e$. Furthermore, on the text $f(0) \circ f(1) \circ f(2) \circ \ldots$, one can find in the limit the least index e such that $\langle x, f(x)\rangle \in B_e$ for all x if such an e does exist. Consider the B'''-recursive function $F^{B'''}$ defined by the condition that $F^{B'''}(e) = e'$ if e' is the least index of a recursive function $\varphi_{e'}$ such that $\langle x, \varphi_{e'}(x)\rangle \in B_e$ for all x, and $F^{B'''}(e) = 0$ whenever such a recursive function $\varphi_{e'}$ does not exist. The function $F^{B'''}$ produces a new confident partial learner Q of REC that outputs r.e. indices. By applying the above claim again, and following an argument exactly analogous to the case of learning $REC_{0,1}$, Q may be simulated to construct an A-recursive learner P of REC, as required. ⧫

The condition that the double jump of the oracle be Turing above $\mathbb{K}''$ is not, however, sufficient for confidently partially learning REC, as the following theorem demonstrates.

Theorem 39 *There is a set A with $A'' \geq_T \mathbb{K}''$ such that A is 2-generic and $REC_{0,1}$ is not confidently partially learnable relative to A.*

Proof. The proof of this result is based on the existence of a 2-generic set A such that $\mathbb{K}'' \leq_T \mathbb{K}' \oplus A$, so that A is *high*$_2$, that is, $A'' \geq_T \mathbb{K}''$. It shall be shown that $REC_{0,1}$ is not confidently partially learnable relative to any such set A. Fix such a set A, as well as a $\{0,1\}$-valued total function f which is 2-generic relative to A; one then has that $A \oplus \{\langle x,y\rangle : y = f(x)\}$ is also 2-generic.

Assume towards a contradiction that M^A were a confident partial learner of $REC_{0,1}$. By the confidence of M^A, it must output some index, say e, infinitely often on the canonical text for f, where f was chosen as above. Then there are prefixes α of $A(0)\circ A(1)\circ A(2)\circ\ldots$ and σ of $f(0)\circ f(1)\circ f(2)\ldots$ for which $\forall\beta\forall\tau\exists\gamma\exists\eta[M^{\alpha\circ\beta\circ\gamma}(\sigma\circ\tau\circ\eta) = e]$ holds. This property of M^A follows from the 2-genericity of $A \oplus \{\langle x,y\rangle : y = f(x)\}$; for, assuming that the prefixes α,σ do not exist, consider the Π^0_1 set of binary strings

$$W = \{\beta\oplus\theta : \forall\gamma\in\{0,1\}^*\forall\tau\in\mathbb{N}^*\forall x,y,z[\theta\in\{0,1\}^*\wedge|\theta| = |\beta|$$
$$\wedge\,(\theta(\langle x,y\rangle) = \theta(\langle x,z\rangle) = 1 \Leftrightarrow y = z)\wedge((\max(\{p:\exists q[\langle p,q\rangle < |\beta|]\})$$
$$< |\tau| \wedge (\tau(x) = y \Leftrightarrow \theta(\langle x,y\rangle) = 1)) \Rightarrow (M^{\beta\circ\gamma}(\tau)\neq e))]\},$$

where the join of two strings $\beta\oplus\theta$ is defined to be the string ξ of length $2\max(|\beta|,|\theta|)$ such that $\xi(2x) = \beta(x)$, $\xi(2x+1) = \theta(x)$ whenever $\beta(x),\theta(x)$ are defined; otherwise, $\xi(2x) = \xi(2x+1) = 0$. By assumption, for all m,n there exist extensions $A[n]\circ\beta$ and $f[m]\circ\tau$ of $A[n]$ and $f[m]$ respectively such that for any strings $\gamma\in\{0,1\}^*,\eta\in\mathbb{N}^*$, $M^{A[n]\circ\beta\circ\gamma}(f[m]\circ\tau\circ\eta)\neq e$. The constant m and string τ may be chosen so that $\max(\{p:\exists q[\langle p,q\rangle < |A[n]\circ\beta|]\}) < |f[m]\circ\tau|$, implying that $(A[n]\circ\beta)\oplus\theta\in W$, where θ is a binary string of length $|A[n]\circ\beta|$ with $\theta(\langle x,y\rangle) = 1$ if and only if $y = (f[m]\circ\tau)(x)$ and $\theta(\langle x,y\rangle) = \theta(\langle x,z\rangle) = 1$ if and only if $y = z$. Moreover, there cannot exist an n such that, if θ is a binary string of length $n+1$ representing the characteristic function of the set $\{\langle x,y\rangle\leq n : y = f(x)\}$, then $A[n]\oplus\theta\in W$. For, by the hypothesis that M^A outputs e infinitely often on the canonical text for f, there must exist $\beta\in\{0,1\}^*$ and $\tau\in\mathbb{N}^*$ satisfying $\max(\{p:\exists q[\langle p,q\rangle < |A[n]|]\}) < |\tau|$,

$\tau(x) = y$ if and only if $\theta(\langle x, y\rangle) = 1$, and $M^{A[n]\circ\beta}(\tau) = e$; this would thus contradict the condition for $A[n] \oplus \theta$ to be in W. The preceding two conclusions contradict the 2-genericity of $A \oplus \{\langle x, y\rangle : y = f(x)\}$, which means that the prefixes α and σ with the required properties must exist. Now fix the two prefixes α and σ.

The proof proceeds next by constructing two different $\{0, 1\}$-valued recursive functions, f_0 and f_1, such that M^A outputs e infinitely often on the canonical texts for f_0 and f_1. Let f_0 and f_1 be defined as follows.

- At the initial stage, put $f_0(x) = \sigma(x)$ for all $x < |\sigma|$, and $f_0(|\sigma|) = 0$; $f_1(x) = \sigma(x)$ for all $x < |\sigma|$, and $f_1(|\sigma|) = 1$. Let $\sigma_{0,0} = \sigma \circ 0$ and $\sigma_{1,0} = \sigma \circ 1$.
- At stage $s + 1$, consider all 2^{s+1} binary strings of length $s + 1$; call them $\beta_0, \beta_1, \ldots, \beta_{2^s}$. Search for a sequence of binary strings $\tau_{0,s,0}, \tau_{0,s,1}, \ldots, \tau_{0,s,2^{s+1}}$ with $\tau_{0,s,0} = \sigma_{0,s}$, and for $k = 0, 1, \ldots, 2^s$, $\tau_{0,s,k+1}$ is a proper extension of $\tau_{0,s,k}$ such that $M^{\alpha\circ\beta_k\circ\gamma_k}(\tau_{0,s,k+1}) \downarrow= e$ for some $\gamma_k \in \{0, 1\}^*$. Similarly, find a sequence of binary strings $\tau_{1,s,0}, \tau_{1,s,1}, \ldots, \tau_{1,s,2^{s+1}}$ with $\tau_{1,s,0} = \sigma_{1,s}$, and for $k = 0, 1, \ldots, 2^s$, there is a $\delta_k \in \{0, 1\}^*$ such that $\tau_{1,s,k} \prec \tau_{1,s,k+1}$ and $M^{\alpha\circ\beta_k\circ\delta_k}(\tau_{1,s,k+1}) \downarrow= e$. Let $\sigma_{0,s+1} = \tau_{0,s,2^{s+1}}$ and $\sigma_{1,s+1} = \tau_{1,s,2^{s+1}}$. By the properties of α and σ, the chains of string extensions $\{\tau_{0,s,1}, \tau_{0,s,2}, \ldots, \tau_{0,s,2^{s+1}}\}$, $\{\tau_{1,s,1}, \tau_{1,s,2}, \ldots, \tau_{1,s,2^{s+1}}\}$, as well as the strings γ_k, δ_k must exist, since it may be assumed inductively that σ is a prefix of both $\tau_{0,s,k}$ and $\tau_{1,s,k}$ for $k = 0, 1, \ldots, 2^s$.

  Set $f_0(x) = \sigma_{0,s+1}(x)$ for all $x \in \mathrm{dom}(\sigma_{0,s+1})$ if $f_0(x)$ is not already defined. Likewise, set $f_1(x) = \sigma_{1,s+1}(x)$ for all $x \in \mathrm{dom}(\sigma_{1,s+1})$ if $f_1(x)$ has not been defined.

It shall be shown that for infinitely many s and binary strings γ_k found in the algorithm at stage $s+1$, if $\alpha \circ \beta_k$ is a prefix of $A(0) \circ A(1) \circ A(2) \circ \ldots$, then $A(0) \circ A(1) \circ A(2) \circ \ldots$ also extends $\alpha \circ \beta_k \circ \gamma_k$. Assume for the sake of a contradiction that there is an s_0 such that for all stages $s+1 > s_0$, whenever $\alpha \circ \beta_k$ is a prefix of $A(0) \circ A(1) \circ A(2) \circ \ldots$, then the string γ_k found with $M^{\alpha \circ \beta_k \circ \gamma_k}(\tau_{0,s,k+1}) \downarrow = e$ fails to satisfy the condition that $A(0) \circ A(1) \circ A(2) \circ \ldots$ extends $\alpha \circ \beta_k \circ \gamma_k$. Consider the Σ^0_1 set U consisting of all binary strings $\alpha \circ \beta_k \circ \gamma_k$ such that γ_k is the first string found at stage $s+1$ for which $M^{\alpha \circ \beta_k \circ \gamma_k}(\tau_{0,s,k+1}) \downarrow = e$. For all n, there is a stage $s+1 > s_0$ at which $\alpha \circ \beta_k = A(0) \circ A(1) \circ A(2) \circ \ldots \circ A(n)$ for some β_k, and by assumption the string $\alpha \circ \beta_k \circ \gamma_k$ in U is not a prefix of $A(0) \circ A(1) \circ A(2) \circ \ldots$; this contradicts the 2-genericity of A. Hence there are infinitely many stages s at which $M^{A(0) \circ A(1) \circ \ldots \circ A(k)}(\tau_{0,s,n}) = e$ for some numbers k, n, and so M outputs e infinitely often on the canonical text for f_0 when it has access to the oracle A. An argument exactly analogous to the preceding one, with δ_k in place of γ_k and $\tau_{1,s,k+1}$ in place of $\tau_{0,s,k+1}$, establishes that M, with access to the oracle A, also outputs e infinitely often on the canonical text for f_1. These two conclusions contradict the fact that M must confidently partially learn both the recursive functions f_0 and f_1, since f_0 and f_1 differ on the argument $|\sigma|$, and yet M outputs the same index infinitely often on their respective canonical texts. In conclusion, $REC_{0,1}$ is not confidently partially learnable relative to A. ♦

A possible further question to consider is whether confidence and behaviourally correct learnability, when imposed all at once on a class of recursive functions, can secure explanatory learnability; a negative answer to this is provided in the next result.

Theorem 40 *The class* $\mathcal{C} = \{f : f \text{ is recursive } \wedge \forall x[f(x)\downarrow = \varphi_{f(0)}(x)\downarrow]\}$ $\cup \{f : f \text{ is recursive } \wedge f(0)\downarrow \wedge \exists p \forall x[\varphi_{f(0)}(x)\uparrow \leftrightarrow x = p \wedge \forall y \neq p[f(y)\downarrow = \varphi_{f(0)}(y)\downarrow]]\}$ *is behaviourally correctly learnable and confidently partially learnable, but not explanatorily learnable.*

Proof. A behaviourally correct learner M may be programmed as follows: on input σ, M conjectures an index for the partial-recursive function

$$\varphi_i(x) = \begin{cases} \sigma(x) & \text{if } x < |\sigma|; \\ \varphi_{\sigma(0)}(x) & \text{if } x \geq |\sigma|. \end{cases}$$

That M behaviourally correctly learns $\mathcal{C}$ is justified by the observation that for every recursive function f in $\mathcal{C}$, f is almost everywhere equal to $\varphi_{f(0)}$. Hence, on the canonical text for any $f \in \mathcal{C}$, M will converge semantically to a correct index.

Furthermore, $\mathcal{C}$ is confidently partially learnable via the following algorithm: on input σ, the learner P identifies the least number $x_0 < |\sigma|$ such that $\varphi_{\sigma(0),|\sigma|}(x_0)\uparrow$; if $x_0 > y$ for some y such that $\varphi_{\sigma(0),|\sigma|-1}(y)\uparrow$, P first conjectures $\varphi_{\sigma(0)}$ one time, and then outputs an index for the partial-recursive function φ_i which was defined above for the behaviourally correct learner M. If no such y exists, P outputs j, where

$$\varphi_j(x) = \begin{cases} \sigma(x_0) & \text{if } x = x_0; \\ \varphi_{\sigma(0)}(x) & \text{if } x \neq x_0. \end{cases}$$

For the remaining case that $\varphi_{\sigma(0),|\sigma|}(x)\downarrow$ whenever $x < |\sigma|$, P conjectures a fixed index for $\varphi_{\sigma(0)}$.

If P is fed with a text for some $f \in \mathcal{C}$ such that $\varphi_{f(0)}(p)\uparrow$, then there is a stage s

from which point onwards p will always remain as the least input on which $\varphi_{\sigma(0)}$ is undefined, and P will converge syntactically to a correct index for f; namely, that for the partial-recursive function φ_i with $\varphi_i(x) = f(p)$ if $x = p$, and $\varphi_i(x) = \varphi_{f(0)}(x)$ for all other values of x. If P is presented with a text for some $f \in \mathcal{C}$ with $\varphi_{f(0)}$ total, then it will conjecture $\varphi_{f(0)}$ infinitely often, and output every other index for at most a finite number of times. Thus P confidently partially learns $\mathcal{C}$.

Assume towards a contradiction that N were an explanatory learner of the class $\mathcal{C}$. Applying Kleene's Recursion Theorem, there is an index e such that $\varphi_e(0) = e$, and for $x > 0$, $\varphi_e(x)$ is defined inductively as follows. Let k be the least value on which φ_e has not been defined; then $\varphi_e(x) = 0$ for all $x > k$ if, given any number s, $N(\varphi_e(0)\circ\varphi_e(1)\circ\ldots\circ\varphi_e(k-1)\circ t\circ 0^s) \leq k$ whenever $t \leq s$. Otherwise, let s be the first number found such that for some least $n \leq s$, $N(\varphi_e(0)\circ\varphi_e(1)\circ\ldots\circ\varphi_e(k-1)\circ n\circ 0^s) > k$ holds; then set $\varphi_e(k) = n$ and $\varphi_e(k+i) = 0$ for all i with $1 \leq i \leq s$.

First, suppose that φ_e as defined above is total. This means, in particular, that $\varphi_e \in \mathcal{C}$; however, since N outputs arbitrarily large indices on the canonical text for φ_e, it cannot be an explanatory learner of $\mathcal{C}$. Secondly, suppose that $\varphi_e(x)$ is undefined if and only if $x = k$, and for all $x > k$, $\varphi_e(x) \downarrow= 0$. By the construction of φ_e, this implies that for all numbers s and $t \leq s$, $N(\varphi_e(0) \circ \varphi_e(1) \circ \ldots \circ \varphi_e(k-1) \circ t \circ 0^s) \leq k$. Now one may choose a number a sufficiently large so that for all $l \leq k$, either $\varphi_l(k) \uparrow$ or $a > \varphi_l(k) \downarrow$ holds. Consequently, there is a recursive function $f \in \mathcal{C}$ defined by

$$f(x) = \begin{cases} a & \text{if } x = k; \\ \varphi_e(x) & \text{if } x \neq k. \end{cases}$$

As N outputs at least one index $l \leq k$ infinitely often on the canonical text for f, but $f(k)$ is chosen so that either $\varphi_l(k)\uparrow$ or $\varphi_l(k)\downarrow< f(k)$, N fails to explanatorily correctly learn $\mathcal{C}$, a contradiction. This case distinction establishes that $\mathcal{C}$ is not explanatorily learnable. ♦

It may be asked whether the preceding result can be sharpened by identifying non-explanatorily learnable classes that are not only behaviourally correctly learnable but even vacillatorily learnable. This, however, is not possible, as every vacillatorily learnable class of recursive functions is already explanatorily learnable.

Theorem 41 *If a class $\mathcal{C}$ of recursive functions is vacillatorily learnable, then it is explanatorily learnable.*

Proof. Let $\mathcal{C}$ be a class of recursive functions such that M is a vacillatory recursive learner of $\mathcal{C}$. An algorithm for an explanatory learner N is as follows: on input $\sigma = f(0)\circ f(1)\circ\ldots\circ f(n)$, let $e_0, e_1, \ldots, e_n$ be all the hypotheses issued by M on the initial segments of σ. Choose the subset $S = \{e_{i_0}, \ldots, e_{i_k}\}$ of $\{e_0, e_1, \ldots, e_n\}$ such that for all $e_{i_j}\in S$, $\varphi_{e_{i_j},n}$ is consistent with all the data seen so far; that is, for all $x \leq n$, either $\varphi_{e_{i_j},n}(x)\uparrow$ or $\varphi_{e_{i_j},n}(x)\downarrow= f(x)$. N then conjectures the index d satisfying

$$\varphi_d(x) = \begin{cases} \varphi_{e_{i_j}}(x) & \text{if } e_{i_j} \text{ is the first number found in } S \text{ such that } \varphi_{e_{i_j}}(x)\downarrow; \\ \uparrow & \text{if } \varphi_{e_{i_j}}(x)\uparrow \text{ for all } e_{i_j}\in S. \end{cases}$$

Suppose N is fed with the canonical text for some $f\in\mathcal{C}$. Since M vacillatorily learns $\mathcal{C}$, it conjectures only finitely many different hypotheses on any text for f. Consequently, at a sufficiently large stage, the set S identified at every step of the above algorithm contains only all the hypotheses of M consistent with f. In addition,

S must contain a correct index for f in the limit. Therefore N explanatorily learns every $f \in \mathcal{C}$. ♦

We now address a different sort of question in partial learning: can one always uniformly extend the recursive functions confidently partially learnt by some recursive learner to a class of partial-recursive functions so that every recursive function in this class is also confidently partially learnable? The following theorem gives an affirmative answer.

Theorem 42 *If a class $\mathcal{C}$ of recursive functions is confidently partially learnable, then there is a one-one numbering $f_0, f_1, f_2, \ldots$ of partial-recursive functions such that*

- $\mathcal{C} \subseteq \{f_0, f_1, f_2, \ldots\}$;
- *each f_i has either a finite or a cofinite domain;*
- *the subclass of all recursive functions in $\{f_0, f_1, f_2, \ldots\}$ is confidently partially learnable with respect to the hypothesis space $\{f_0, f_1, f_2, \ldots\}$.*

Proof. Let $\mathcal{C}$ be a class of recursive functions that is confidently partially learnt by the recursive learner M. Now define a numbering $f_0, f_1, f_2, \ldots$ of partial-recursive functions according to the following steps.

1. For each sequence $\sigma \in \mathbb{N}^*$, determine whether or not $M(\sigma) \neq M(\tau)$ for all $\tau \prec \sigma$. If so, then define f_σ according to Step 2.; otherwise, f_σ is defined according to Step 3.

2. Let $f_\sigma(x) = \sigma(x)$ for all $x < |\sigma|$, and for all $y \geq |\sigma|$,

$$f_\sigma(y) = \begin{cases} \varphi_{M(\sigma)}(y) & \text{if } \exists \eta \in \mathbb{N}^*[M(\sigma \circ \eta) = M(\sigma) \wedge y < |\sigma \circ \eta| \\ & \wedge \forall z < |\sigma \circ \eta|[\varphi_{M(\sigma)}(z) \downarrow = (\sigma \circ \eta)(z)]]; \\ \uparrow & \text{otherwise.} \end{cases}$$

3. Put

$$f_\sigma(x) = \begin{cases} \sigma(x) & \text{if } x < |\sigma|; \\ \uparrow & \text{if } x = |\sigma|; \\ 0 & \text{if } x > |\sigma|. \end{cases}$$

First, it is shown that $\mathcal{C} \subseteq \{f_0, f_1, f_2, \ldots\}$. Let g be any recursive function in $\mathcal{C}$. As M confidently partially learns g, there is a shortest sequence σ with $g(x) = \sigma(x)$ for all $x \in \mathrm{dom}(\sigma)$ and $g = \varphi_{M(\sigma)}$, such that M outputs on the canonical text $g(0) \circ g(1) \circ g(2) \circ \ldots$ the index $M(\sigma)$ infinitely often. Thus the Σ_1^0 condition defining f_σ in Step 2. is satisfied for all numbers y, giving that $f_\sigma = g$. Moreover, if $M(\sigma) \neq M(\tau)$ for all $\tau \prec \sigma$, then by Step 2. f_σ is either total or has finite domain; otherwise, the construction of f_σ in Step 3. ensures that the domain of f_σ is cofinite.

In addition, the numbering is one-one: for any $\sigma, \tau \in \{0,1\}^*$, if $\sigma \not\preceq \tau$ and $\tau \not\preceq \sigma$, then, since $\sigma \preceq f_\sigma(0) \circ f_\sigma(1) \circ \ldots$ and $\tau \preceq f_\tau(0) \circ f_\tau(1) \circ \ldots$, f_σ and f_τ must differ on at least one input. Suppose, on the other hand, that $\sigma \prec \tau$ holds. Consider the following case distinction. (1) If Step 2. applies to both σ and τ, then $M(\sigma) \neq M(\tau)$, so that by the confidence of M, σ and τ cannot both be extended to a common infinite sequence on which M outputs two different numbers infinitely often. Hence $f_\sigma \neq f_\tau$. (2) If Step 3. applies to σ, then it also applies to τ. Consequently, $f_\sigma(|\sigma|) \uparrow$ but $f_\tau(|\sigma|) = \tau(|\sigma|)$, and so $f_\sigma \neq f_\tau$ again holds. (3) If Steps

2. and 3. apply to σ and τ respectively, then f_σ is either total or has finite domain, while f_τ remains undefined on one input and has infinite domain. Therefore $f_\sigma \neq f_\tau$ still holds. This completes the case distinction, and shows that $\{f_0, f_1, f_2, \ldots\}$ is a one-one numbering. To produce a new confident partial learner N of all recursive functions in $\mathcal{C}$ using $\mathcal{C}$ itself as a hypothesis space, suppose that N is fed with the text segment σ; it then chooses the shortest $\tau \preceq \sigma$ with $M(\tau) = M(\sigma)$ and outputs τ. On any input text $a_0 \circ a_1 \circ a_1 \circ \ldots$, M outputs exactly one index e infinitely often, and if η is the shortest prefix of the given text with $M(\eta) = e$, then N outputs η infinitely often, and all other indices only finitely often. If g is any recursive function in $\{f_0, f_1, f_2, \ldots\}$, then there is a unique segment $\sigma \prec g(0) \circ g(1) \circ g(2) \circ \ldots$ such that Step 2. applies to σ, and the Σ^0_1 criteria defining f_σ is fulfilled for all inputs y. Therefore $g = \varphi_{M(\sigma)}$, and since $\varphi_{M(\tau)}(x) = \tau(x)$ for all prefixes τ of $\varphi_{M(\sigma)}(0) \circ \varphi_{M(\sigma)}(1) \circ \varphi_{M(\sigma)}(2) \circ \ldots$, N outputs σ infinitely often. This establishes all the properties of the numbering $\{f_0, f_1, f_2, \ldots\}$ in the claim. ♦

The example given below shows that one cannot in general obtain a uniformly recursive class of functions covering all the recursive functions confidently partially learnt by a recursive learner.

Example 43 *Consider the class $\mathcal{C} = \{f : \forall x[f(x)\downarrow = \varphi_{f(0)}(x)\downarrow]\}$ of self-describing functions. $\mathcal{C}$ is confidently partially learnable, but there is no numbering of recursive functions $f_0, f_1, f_2, \ldots$ such that $\mathcal{C} \subseteq \{f_0, f_1, f_2, \ldots\}$.*

Proof. Suppose for the sake of a contradiction that there exists a numbering $f_0, f_1, f_2, \ldots$ of recursive functions such that $\mathcal{C} \subseteq \{f_0, f_1, f_2, \ldots\}$. Now define a family of recursive functions as follows. For any given number e, let

$$g(e,x) = \begin{cases} e & \text{if } x = 0; \\ f_{x-1}(x) + 1 & \text{if } x > 0. \end{cases}$$

Since $f_0, f_1, f_2, \ldots$ is a numbering of recursive functions, each function $g(e,x)$ for a fixed e is recursive. By the s-m-n theorem, there is a recursive function h with $\varphi_{h(e)}(x) \downarrow = g(e,x) \downarrow$ for all x. Further, it follows from Kleene's Recursion Theorem that $\varphi_{h(e)} = \varphi_e$ for some e. Then $\varphi_{h(e)} \in \mathcal{C}$ for this e and $\varphi_e(x+1) = f_x(x+1)+1 > f_x(x+1)$ for all x. Hence the assumption that $\mathcal{C} \subseteq \{f_0, f_1, f_2, \ldots\}$ is wrong. ♦

4.2 Consistent Partial Learning

The present section considers a weakened notion of consistency in partial learning, namely, *essential class consistency.* Under this learning paradigm, the learner is permitted to be inconsistent on finitely many data inputs. First, we review the original notion of class consistent partial learning introduced in [13] with some examples.

Example 44 *The class of self-describing functions* $\mathcal{C} = \{f : \forall x[f(x) \downarrow = \varphi_{f(0)}(x) \downarrow]\}$ *is class consistently explanatorily learnable but not consistently explanatorily learnable.*

Theorem 45 *There is a class of recursive functions which is confidently explanatorily learnable but not class consistently partially learnable.*

Proof 1. The class $\mathcal{C} = \{f : f \text{ is recursive} \wedge (m = \min(\text{range}(f)) \rightarrow \forall x[f(x) \downarrow = \varphi_m(x) \downarrow])\}$ is confidently explanatorily learnable but not class consistently partially learnable.

An explanatory learner M of $\mathcal{C}$ may be programmed as follows: on input σ with $e = \min(\text{range}(\sigma))$, M outputs e. If M is presented with the canonical text $f(0) \circ f(1) \circ f(2) \circ \ldots$ for some $f \in \mathcal{C}$ such that $e = \min(\text{range}(f))$, then M will always correctly conjecture the recursive function $f = \varphi_e$ once e appears in the text. Hence M is a confident explanatory learner of $\mathcal{C}$.

Now assume by way of contradiction that N were a class consistent partial learner of $\mathcal{C}$. The following claim is first established.

Claim 46 *For any number e, there are sequences σ_1, σ_2 which satisfy the following conditions.*

- *$range(\sigma_1) \cup range(\sigma_2) \subseteq \{e, e+1, e+2, \ldots\}$;*
- *$\exists x[\sigma_1(x)\downarrow \neq \sigma_2(x)\downarrow]$;*
- *$N(\sigma_1) = N(\sigma_2)$.*

Suppose to the contrary that there exists a number e_0 such that for all σ_1, σ_2 with $\sigma_1(x)\downarrow \neq \sigma_2(x)\downarrow$ for some x and $\text{range}(\sigma_1) \cup \text{range}(\sigma_2) \subseteq \{e_0, e_0+1, e_0+2, \ldots\}$, the condition $N(\sigma_1) \neq N(\sigma_2)$ holds. Consequently, there is a recursive function f such that for all $e < e_0$, $\varphi_{f(e)} = \varphi_{f(e_0)}$, and for all $e \geq e_0$, $\varphi_{f(e)}$ is defined inductively by

$$\varphi_{f(e)}(x) = \begin{cases} e & \text{if } x = 0; \\ \min(\{y : N(\varphi_{f(e)}(0) \circ \varphi_{f(e)}(1) \circ \ldots \circ \varphi_{f(e)}(x-1) \circ y) > e + x\}) & \text{if } x > 0. \end{cases}$$

Owing to the initial assumption that for all σ_1, σ_2 with $\mathrm{range}(\sigma_1) \cup \mathrm{range}(\sigma_2) \subseteq \{e_0, e_0+1, e_0+2, \ldots\}$, $|\sigma_1| = |\sigma_2|$, and $\sigma_1 \neq \sigma_2$, it holds that $N(\sigma_1) \neq N(\sigma_2)$, every partial-recursive function $\varphi_{f(e)}$ is total. By Kleene's Recursion Theorem, there exists an $i \geq e_0$ for which $\varphi_{f(i)} = \varphi_i$. Then $\varphi_i \in \mathcal{C}$ for this i, but since N outputs on the canonical text for φ_i each index only finitely often, it cannot partially learn φ_i. This establishes the claim.

Applying the claim, one may find two-place recursive functions g, h which perform the following instructions. On input (x, y), g and h search for the first two finite sequences $\sigma_{x,y,1}, \sigma_{x,y,2}$ which fulfil the criteria laid out in the subclaim with $e = \max(\{x, y\})$. Then g and h are programmes such that

$$\varphi_{g(x,y)}(z) = \begin{cases} \sigma_{x,y,1}(z) & \text{if } z < |\sigma_{x,y,1}|; \\ x & \text{if } z \geq |\sigma_{x,y,1}|, \end{cases}$$

$$\varphi_{h(x,y)}(z) = \begin{cases} \sigma_{x,y,2}(z) & \text{if } z < |\sigma_{x,y,2}|; \\ y & \text{if } z \geq |\sigma_{x,y,2}|. \end{cases}$$

By the choice of $\sigma_{x,y,1}$ and $\sigma_{x,y,2}$, the learner N must be inconsistent on at least one of these two sequences, that is, there is a $j \in \{1, 2\}$ for which either $\varphi_{M(\sigma_{x,y,j})}$ is undefined on some input $z < |\sigma_{x,y,i}|$, or $\varphi_{M(\sigma_{x,y,j})}(z) \downarrow \neq \sigma_{x,y,j}(z) \downarrow$. Furthermore, by the Double Recursion Theorem, there exist numbers a, b for which $\varphi_{g(a,b)} = \varphi_a$ and $\varphi_{h(a,b)} = \varphi_b$. For this pair of values (a, b), $\varphi_a \in \mathcal{C}$ and $\varphi_b \in \mathcal{C}$; on the other hand, since N is inconsistent on at least one of the canonical texts for φ_a and φ_b, N cannot be a class consistent partial learner of $\mathcal{C}$. In conclusion, $\mathcal{C}$ is confidently explanatorily learnable but not class consistently partially learnable. ♦

Proof 2. The class $\mathcal{L} = \{f : f \text{ is recursive} \wedge f = \varphi_{f(0)} \wedge \forall x[f(x) > 0]\} \cup \{f : f \text{ is recursive } \wedge \exists x \forall y[f(y) = 0 \leftrightarrow y \geq x]\}$ is confidently explanatorily learnable but not class consistently partially learnable.

Consider a recursive learner N that, on input σ, outputs a fixed index for $\varphi_{\sigma(0)}$ if $\min(\text{range}(\sigma)) > 0$; otherwise, if $m = \min(\{y : \sigma(y) = 0\})$, it outputs a programme for the recursive function f given by $f(x) = \sigma(x)$ if $x < m$, and $f(x) = 0$ if $x \geq m$. N is then a confident explanatory learner of $\mathcal{L}$. Assume that M were a class consistent partial learner of $\mathcal{L}$. Let $F(x) = \max(\{s \geq 1 : \sigma \in \{1, 2, \ldots, x\}^{\{1,2,\ldots,x\}} \wedge \forall y \in \text{dom}(\sigma)[\varphi_{M(\sigma),s}(y) \downarrow \wedge \varphi_{M(\sigma),s-1}(y) \uparrow]\})$. F is recursive: firstly, every finite sequence may be extended to a recursive function f that is almost everywhere equal to zero, so that $f \in \mathcal{L}$. Therefore the class consistency of M implies that for every $\sigma \in \{1, 2, \ldots, x\}^{\{1,2,\ldots,x\}}$, $\varphi_{M(\sigma)}(y)$ is defined for all $y \in \text{range}(\sigma)$. Now let g be a self-describing recursive function such that for all $x > 0$,
$g(x) \in \{1, 2, \ldots, x\} - \{\varphi_{0,F(x)}(x), \varphi_{1,F(x)}(x), \ldots, \varphi_{x-2,F(x)}(x)\}$. If M were presented with the canonical text $T_g = g(0) \circ g(1) \circ g(2) \circ \ldots$, then for every prefix
$\sigma = g(0) \circ g(1) \circ g(2) \circ \ldots \circ g(x)$ of T_g, $M(\sigma) \notin \{0, 1, \ldots, x-2\}$ holds; otherwise, by the construction of g, $\varphi_{M(\sigma),F(x)}(x) \downarrow = \varphi_{M(\sigma)}(x) \neq g(x)$, contradicting the class consistency of M. Hence M outputs each index only finitely often on T_g, and consequently does not class consistently learn $\mathcal{L}$. ♦

Whilst class consistency is a fairly natural learning constraint in inductive inference of recursive functions, the next theorem shows that it cannot in general guarantee that a class is also confidently partially learnable. However, it is presently unknown whether this theorem remains true when the condition of class consistency is replaced with general consistency.

Theorem 47 *There is a class of recursive functions which is class consistently partially learnable but not confidently partially learnable.*

Proof. The following example essentially modifies the construction of the programme $g(d)$ in Theorem 4.1 so that a subclass of $\mathcal{C}$ may be class consistently partially learnable. For each number d, let $g(d)$ be a programme for a partial-recursive function $\varphi_{g(d)}$ which is defined as follows.

- Set $\varphi_{g(d),s}(0) = d$ for all s.
- Initialize the markers $a_0, a_1, a_2, \ldots$ by setting $a_{i,0} = \langle i, 0\rangle + 1$ for $i \in \mathbb{N}$.
- At stage $s+1$, consider each marker $a_{i,s} = \langle i, r\rangle + 1$ such that $a_{i,s} \leq s+1$, and execute the following instructions in succession. Set $\varphi_{g(d),s+1}(x) = 0$ for all $x = \langle i, j\rangle + 1 \leq s+1$ such that $j \neq r$ if $\varphi_{g(d),s}$ is not already defined on x. Next, check whether $\varphi_{i,s+1}(a_{i,s}) \downarrow \in \{0,1\}$ holds; if so, let $\varphi_{g(d),s+1}(a_{i,s}) = 1 - \varphi_{i,s+1}(a_{i,s})$ if $\varphi_{g(d)}$ is not already defined on the input $a_{i,s}$. Now, for each i such that $\langle i, m\rangle + 1 \leq s+1$ for some m, let $u = \max(\{m : \langle i, m\rangle + 1 \leq s+1\})$. Associate the marker $a_{i,s+1}$ with $\langle i, u+1\rangle + 1$ if at least one of the following two conditions applies; otherwise, let $a_{i,s+1} = a_{i,s}$.

  1. There is a $j < i$ with $\langle j, m\rangle + 1 \leq s+1$ for some m such that $a_{j,s+1} \neq a_{j,s}$.
  2. If $a_{i,s} = \langle i, r\rangle + 1$, then the inequality $|\{0, 1, \ldots, r\} - W_{d,s+1}| < i$ holds.

Let $\mathcal{C} = \{f : W_d \text{ is cofinite } \wedge f \text{ is a total recursive extension of } \varphi_{g(d)}\}$. One may prove the following properties of the partial-recursive function $\varphi_{g(d)}$.

- If W_d is cofinite, then all the markers a_i with $i \leq |\overline{W}_d|$ settle down permanently, while all the markers a_j with $j > |\overline{W}_d|$ move infinitely often, so that $W_{g(d)}$ is cofinite.

- If W_d is coinfinite, then each of the markers a_i is eventually fixed permanently, so that $W_{g(d)}$ is coinfinite; moreover, there is no total recursive function extending $\varphi_{g(d)}$.

First, suppose that W_d is cofinite. Then for all $i \leq |\overline{W}_d|$, there is a sufficiently large stage $s+1$ for which $|\{0, 1, \ldots, r\} - W_{d,s'}| \geq i$ holds if $a_{i,s'} = \langle i, r\rangle + 1$ and whenever $s' \geq s+1$. Hence condition 2. for the marker a_i to move almost always fails. Furthermore, condition 1. is fulfilled only finitely often. This can be seen by induction on the indices of all markers a_j: for $j = 0$, the marker a_0 can only be moved if condition 2. is satisfied, and, as argued above, this can only happen finitely often. For $j > 0$, the marker a_j can only be moved due to condition 1. if some marker a_k with $k < j$ is moved; by the inductive assumption, all markers a_k such that $k < j$ are moved only finitely often, so that in the limit, the movement of a_j is contingent only on condition 2. Therefore a_i is permanently associated to some fixed value after a large enough stage. On the other hand, if $i > |\overline{W}_d|$, then $a_{i,s}$ satisfies condition 2. at infinitely many stages s, implying that the marker a_i moves infinitely often. One may note further that whenever a marker a_i is moved at some stage $s+1$ from $\langle i, r\rangle + 1$ to $\langle i, u+1\rangle + 1$, where $u = \max(\{m : \langle i, m\rangle + 1 \leq s+1\})$, then $\varphi_{g(d)}(\langle i, r\rangle + 1)$ is assigned the value 0 at a subsequent stage. In particular, this implies that $\varphi_{g(d)}$ is defined on all inputs $\langle i, j\rangle + 1$ with $i > |\overline{W}_d|$, and thus $W_{g(d)}$ is cofinite.

Secondly, suppose that W_d is coinfinite. As was argued in the preceding paragraph, only condition 2. may effect a shift in the marker a_0, and since W_d is coinfinite, this condition can only be satisfied finitely often; it then follows by induction on the indices of the markers that for each marker, a movement due to condition 1. happens for at most a finite number of times. Owing to the fact that W_d is coinfinite, a marker meets condition 2. finitely often, and therefore it must settle down permanently on a fixed value after a sufficiently large stage. For each i, let $a_i = \lim_{s \to \infty} a_{i,s}$. By the construction of $\varphi_{g(d)}$, $\varphi_{g(d)}(a_i)$ is defined if and only if $\varphi_i(a_i) \downarrow \in \{0, 1\}$, in which case it is equal to $1 - \varphi_i(a_i)$. Hence any total extension of $\varphi_{g(d)}$ cannot be a recursive function.

Now it is shown that $\mathcal{C}$ is class consistently partially learnable. First, define a recursive learner N as follows. On input $\sigma = d \circ f(1) \circ \ldots \circ f(n)$, N first identifies the maximum i, if it exists, such that $a_{j,n} = a_{j,n+1}$ for all $j \leq i$. If no such i exists, N outputs an index for a partial-recursive function ϕ such that $\phi(x) = f(x)$ for all $x \leq n$, and $\phi(x) \uparrow$ for all $x > n$. Otherwise, it conjectures the programme e for which

$$\varphi_e(x) = \begin{cases} f(m) & \text{if } \exists t[m = \langle k, t \rangle + 1 \leq n \wedge \varphi_{g(d),n}(m) \uparrow \text{ and } k \leq i]; \\ \varphi_{g(d)}(x) & \text{otherwise.} \end{cases}$$

Suppose that N processes a text for some recursive function $f \in \mathcal{C}$, so that $W_{f(0)}$ is cofinite. Consider an input sequence $\sigma = d \circ f(1) \circ \ldots \circ f(n)$. If there is a least i such that $a_{i,n} \neq a_{i,n+1}$ and $\langle i, m \rangle + 1 \leq n$ for some m, then by condition 1. above, all markers $a_{j,n}$ with $j \geq i$ and $\langle j, l \rangle + 1 \leq n$ for some l will be moved to a new position $\langle j, u \rangle + 1$ for which $u = \max\{m : \langle i, m \rangle + 1 \leq n + 1\}$. Hence $\varphi_{g(d)}$ will be defined on all inputs $\langle j, m \rangle + 1 \leq n$ such that $j \geq i$. This in turn implies that N is class consistent.

Next, one shows that N has the following learning characteristic: it outputs incorrect indices only finitely often, and it outputs at least one correct index infinitely often. Let $\sigma = d \circ f(1) \circ \ldots \circ f(n)$ with $i = \max\{j : \forall k \leq j[a_{j,n} = a_{j,n+1}]\}$ be a given input sequence. For a case distinction, suppose first that $i > |\overline{W}_d|$. Then, since $W_{g(d)}$ is cofinite and $\varphi_{g(d)}$ is undefined only for values of the form $\langle j, m \rangle + 1$ with $j \leq |\overline{W}_d| < i$, there is a sufficiently large stage after which N patches all the undefined places of $\varphi_{g(d)}$ with the correct values of the input function. Secondly, suppose that $i \leq |\overline{W}_d|$. As was demonstrated above, each of the markers a_j with $j \leq |\overline{W}_d|$ is fixed after a large enough number of computation steps; whence, from this stage onwards, $i \geq |\overline{W}_d|$. Since the marker a_j with $j = |\overline{W}_d| + 1$ moves infinitely often, one concludes that i must be equal to $|\overline{W}_d|$ at infinitely many stages. This establishes the learning property of N claimed at the beginning.

Finally, a class consistent learner M may be built from N as follows: whenever N outputs the sequence of conjectures $e_0, e_1, e_2, \ldots, e_n, \ldots$, M, for each e_n, outputs the index $pad(e_n, k_n)$, where pad is a padding function with $\varphi_{pad(e,d)} = \varphi_e$ for all e, d, and $k_n = |\{m \leq n : e_m < e_n\}|$. Then M outputs exactly one correct index for the input function infinitely often, and it is also class consistent. In conclusion, $\mathcal{C}$ is class consistently partially learnable. The proof that $\mathcal{C}$ is not confidently partially learnable is exactly similar to that in Theorem 4.1: assuming the contrary, one can obtain a $\mathbb{K}'$ procedure for the deciding the set $\{d : W_d \text{ is cofinite}\}$, a contradiction. ♦

Definition. A recursive learner M is *essentially class consistent* if and only if for each canonical text T_f corresponding to some $f \in \mathcal{C}$, where $\mathcal{C}$ is a class of recursive functions to be learnt, $\varphi_{M(T_f(0) \circ T_f(1) \circ \ldots \circ T_f(n))}(m) \downarrow = T_f(m)$ holds whenever $m \leq n$ for almost all n.

Theorem 48 *Every behaviourally correctly learnable class of recursive functions is essentially class consistently partially learnable.*

Proof. Let $\mathcal{C}$ be a class of recursive functions which is behaviourally correctly learnt by a learner M. Next, define a recursive learner N as follows. On an input text $f(0) \circ f(1) \circ f(2) \circ \ldots$, simulate the learner M and observe the conjectures $e_0, e_1, e_2, \ldots$ output by M. N then outputs a conjecture e_i of M at least s times if and only if $\forall x \leq s[\varphi_{e_i,s}(x) \downarrow = f(x)]$ holds. If N is presented with the canonical text for some $f \in \mathcal{C}$, then M, being a behaviourally correct learner of $\mathcal{C}$, will output only finitely many incorrect indices. Therefore N will output each correct index infinitely often, and every incorrect index finitely often. Now one can build a further learner P: whenever N, on the input text, conjectures the sequence $d_0, d_1, d_2, \ldots$, P, for each d_n, outputs $pad(d_n, k_n)$, where pad is a padding function with $\varphi_{pad(d,k)} = \varphi_d$ for all d, k, and $k_n = |\{m \leq n : d_m < d_n\}|$. This learner P is then the required essentially class consistent partial learner of $\mathcal{C}$. ♦

Theorem 49 *The class* $\mathcal{C} = \{f : f$ *is recursive* $\wedge (\exists x \forall y[f(y+1) \downarrow = \varphi_{f(0)}(y) \downarrow \leftrightarrow y \neq x] \vee \forall y[f(y+1) \downarrow = \varphi_{f(0)}(y) \downarrow])\}$ *is essentially class consistently partially learnable but not class consistently partially learnable.*

Proof. Construct a recursive learner M as follows: on input $\sigma = f(0) \circ f(1) \circ \ldots \circ f(n)$, M identifies the least $y \leq n$ such that $\varphi_{f(0),n}(y) \uparrow$; if no such y exists, M

outputs e, where e is the programme defined by

$$\varphi_e(x) = \begin{cases} f(0) & \text{if } x = 0; \\ \varphi_{f(0)}(x-1) & \text{if } x > 0. \end{cases}$$

Otherwise, suppose that y is different from the least $z \leq n-1$ such that $\varphi_{f(0),n-1}(z)\uparrow$ if such a z exists; it then outputs e, with e defined exactly as above, and, on the subsequent input $f(0) \circ f(1) \circ \ldots \circ f(n) \circ f(n+1)$, outputs d, where

$$\varphi_d(x) = \begin{cases} f(0) & \text{if } x = 0; \\ f(y) & \text{if } x = y; \\ \varphi_{f(0)}(x-1) & \text{if } x \notin \{0, y\}. \end{cases}$$

If the last conjecture of M was d, or $n = 0$, then it outputs d on the current input $f(0) \circ f(1) \circ \ldots \circ f(n)$. It will then follow that M essentially class consistently partially learns every $f \in \mathcal{C}$.

In Theorem 40, $\mathcal{C}$ was shown to be behaviourally correctly and confidently partially learnable, but not explanatorily learnable. Now assume by way of contradiction that N were a class consistent recursive learner of $\mathcal{C}$. By Kleene's Recursion Theorem, there is a partial-recursive function φ_e defined in stages as follows: at the initial stage, the programme e searches for the first number x_0 such that either $N(e \circ x_0) > N(e)$ holds, or there is a number $y_0 > x_0$ with $N(e \circ x_0) = N(e \circ y_0)$. If the latter holds, then $\varphi_e(0)$ is left undefined, while $\varphi_e(x)\downarrow = 0$ for all $x > 0$. On the other hand, if x_0 is found such that $N(e \circ x_0) > N(e)$, then $\varphi_e(0)$ is assigned the value x_0, and the programme e proceeds with the next stage of the algorithm. At stage $s+1$, assume that $\varphi_e(x)$ has been defined if and only if

$x \le s$; the programme e then searches for the first number x_{s+1} for which either $N(e{\circ}\varphi_e(0){\circ}\ldots{\circ}\varphi_e(s){\circ}x_{s+1}) > N(\tau)$ holds for all $\tau \prec e{\circ}\varphi_e(0){\circ}\ldots{\circ}\varphi_e(s){\circ}x_{s+1}$, or for some $y_{s+1} > x_{s+1}$, $N(e{\circ}\varphi_e(0){\circ}\ldots{\circ}\varphi_e(s){\circ}x_{s+1}) = N(e{\circ}\varphi_e(0){\circ}\ldots{\circ}\varphi_e(s){\circ}y_{s+1})$. If the first case holds, then $\varphi_e(s{+}1)$ is defined to be x_{s+1}, and the algorithm proceeds to the next stage; if the second case holds, then $\varphi_e(s{+}1)$ remains undefined, and $\varphi_e(x) \downarrow= 0$ for all $x > s+1$. Suppose that the stages run through infinitely often; consequently, N outputs on the canonical text $e \circ \varphi_e(0) \circ \varphi_e(1) \circ \ldots$ for some $f \in \mathcal{C}$ each index only finitely often, and thus cannot be a class consistent learner of f. Suppose instead that a stage s is reached at which $\varphi_e(s) \uparrow$, $\varphi_e(x) \downarrow= 0$ for all $x > s$, and there are distinct numbers x_s, y_s such that $N(e \circ \varphi_e(0) \circ \ldots \circ x_s) = N(e \circ \varphi_e(0) \circ \ldots \circ y_s) = p$ for some p. Hence either $\varphi_p(s) \uparrow$ holds, or $\varphi_p(s) \downarrow$ and $\varphi_p(s)$ differs from at least one of the numbers x_s, y_s. Let f be a recursive function such that $f(0) = e$, $f(x{+}1) = \varphi_e(x)$ for all $x \neq s$, and $\varphi_p(s) \neq f(s+1) \in \{x_s, y_s\}$ if $\varphi_p(s) \downarrow$; if $\varphi_p(s) \uparrow$, then $f(s+1)$ can be arbitrarily selected. For this choice of f, $f \in \mathcal{C}$, but since N is inconsistent on the text segment $e \circ \varphi_e(0) \circ \ldots \circ \varphi_e(s-1) \circ f(s+1)$, it cannot class consistently learn f. In conclusion, $\mathcal{C}$ is not class consistently partially learnable. ♦

Theorem 50 *The class* $\mathcal{C} = \{f : f$ *is recursive* $\wedge f(0) \downarrow \wedge |\overline{W}_{f(0)}| < \infty \wedge \forall x[\varphi_{f(0)}(x) \downarrow \Rightarrow f(x) \downarrow= \varphi_{f(0)}(x) \downarrow]\}$ *is neither class consistently partially learnable nor confidently partially learnable.*

Proof. That $\mathcal{C}$ is not class consistently partially learnable follows directly from Theorem 49; that $\mathcal{C}$ is not confidently partially learnable may be shown by an argument exactly analogous to that in the second proof of Theorem 32. ♦

Theorem 51 *The class* $REC_{0,1}$ *of all* $\{0,1\}$*-valued recursive functions is not es-*

sentially class consistently partially learnable.

Proof. Suppose for the sake of a contradiction that M were a recursive essentially class consistent learner of $REC_{0,1}$. By the reductio hypothesis, one can prove the following claim.

Claim 52 *Let M be as above. Then for any binary string σ, there are string extensions $\tau_0, \tau_1 \in \{0,1\}^*$ such that $\tau_0(x) \neq \tau_1(x)$ for some $x \in dom(\tau_0 \cap \tau_1)$, and $M(\sigma \circ \tau_0) = M(\sigma \circ \tau_1)$.*

Assume that a counterexample to the claim is witnessed by the binary string σ. One may build a recursive $\{0,1\}$-valued function f in stages as follows. At the initial stage $s = 0$, let $f(x) = \sigma(x)$ for all $x \in \text{dom}(\sigma)$, and $f(|\sigma|) = 0$. At stage $s+1$, suppose that $f(x)$ has been defined for all $x \leq |\sigma| + s$. Now consider the outputs $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 0)$ and $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 1)$; by the assumed property of σ, $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 0) \neq M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 1)$. Choose $f(|\sigma| + s + 1) \in \{0,1\}$ such that $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ f(|\sigma| + s + 1)) \neq M(f(0) \circ \ldots \circ f(|\sigma| + k))$ holds for all $k \leq s$ if this is possible; otherwise, if M has already conjectured both $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 0)$ and $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ 1)$ on some prefix of $f(0) \circ \ldots \circ f(|\sigma| + s)$, assign a $\{0,1\}$ value to $f(|\sigma| + s + 1)$ so that $M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ f(|\sigma| + s + 1)) > M(f(0) \circ \ldots \circ f(|\sigma| + s) \circ (1 - f(|\sigma| + s + 1)))$.

One notes that by the construction of f, M outputs on the canonical text for f each index only finitely often. For, according to the algorithm, if $M(f(0) \circ \ldots \circ f(k)) = M(f(0) \circ \ldots \circ f(l))$ for some $l < k$, then there is a number $b < k$ distinct from l with $M(f(0) \circ \ldots \circ f(b)) = M(f(0) \circ \ldots \circ f(k-1) \circ (1 - f(k)))$ and $M(f(0) \circ \ldots \circ f(b)) < M(f(0) \circ \ldots \circ f(k))$. Consequently, by the property of σ, M cannot output $M(f(0) \circ \ldots \circ f(b))$ after processing extensions of the text segment $f(0) \circ \ldots \circ f(k)$. In particular, this means that M outputs $M(f(0) \circ \ldots \circ f(k))$ for at most $M(f(0) \circ \ldots \circ f(k))$ times. Thus M does not essentially class consistently partially learn f, and this establishes the claim.

Next, one constructs a $\{0,1\}$-valued partial- recursive function θ as follows. First, set $\theta(0) = 0$. At stage $s+1$, suppose that θ has been defined on all values up to s', and run a search for two incomparable binary strings, τ_0 and τ_1, such that $M(\theta(0) \circ \ldots \circ \theta(s') \circ \tau_0) = M(\theta(0) \circ \ldots \circ \theta(s') \circ \tau_1) = c_{s+1}$ for some number c_{s+1}, and $\varphi_{c_s+1}(x) \downarrow \in \{0,1\}$, where x is the least number such that $x \in \mathrm{dom}(\tau_0 \cap \tau_1)$ and $\tau_0(x) \neq \tau_1(x)$. Choose the binary string τ_i, $i \in \{0,1\}$, so that $\tau_i(x) = 1 - \varphi_{c_{s+1}}(x)$, and define $\theta(s' + y + 1) = \tau_i(y)$ for all $y \in \mathrm{dom}(\tau_i)$. From this construction of θ, there are two possible cases to consider.

Case (A): *Every stage terminates successfully, so that θ is total.*

It follows directly from the construction of θ that for infinitely many numbers k, there is a $b < k$ with $\theta(b) \neq \varphi_{M(\theta(0) \circ \ldots \circ \theta(k))}(b)$. Consequently, M cannot be an essentially class consistent partial learner of θ.

Case (B): *There is a stage $s+1$ at which no pair of incomparable binary strings τ_0, τ_1 can be found such that, if θ has been defined on all values up to s', then $M(\theta(0) \circ \ldots \circ \theta(s') \circ \tau_0) = M(\theta(0) \circ \ldots \circ \theta(s') \circ \tau_1) = c_{s+1}$ for some number c_{s+1}, and $\varphi_{c_s+1}(x) \downarrow \in \{0,1\}$, where x is the least number such that $x \in dom(\tau_0 \cap \tau_1)$ and $\tau_0(x) \neq \tau_1(x)$.*

One may extend θ to a $\{0,1\}$-valued total recursive function ξ as follows. First, set $\xi(y) = \theta(y)$ for all $y \leq s$. By virtue of the subclaim established above, one can successfully find at stage $t+1$ two binary strings $\tau_{0,t+1}$, $\tau_{1,t+1}$, such that $M(\xi(0) \circ \ldots \circ \xi(t') \circ \tau_{0,t+1}) = M(\xi(0) \circ \ldots \circ \xi(t') \circ \tau_{1,t+1})$ and $\tau_{0,t+1}(x) \neq \tau_{1,t+1}(x)$ for some $x \in \mathrm{dom}(\tau_{0,t+1} \cap \tau_{1,t+1})$; it is assumed that at this stage ξ has been defined up to t'. Choose the binary string $\tau_{i,t+1}$, $i \in \{0,1\}$, which is at least as long as the other, and define $\xi(t'+y+1) = \tau_{i,t+1}(y)$ for all $y \in \mathrm{dom}(\tau_{i,t+1})$. On the hypothesis of Case (B), it follows that if the binary string $\tau_{i,t+1}$ is selected at stage $t+1$, then $\varphi_{M(\xi(0)\circ\ldots\circ\xi(t')\circ\tau_{i,t+1})}(x) \uparrow$ for some $x \in \mathrm{dom}(\tau_{i,t+1})$. This implies that there are infinitely many numbers k such that $\varphi_{M(\xi(0)\circ\ldots\circ\xi(k))}(x) \uparrow$ for some $x \leq k$. Hence M is not an essentially class consistent partial learner of ξ.

In conclusion, M cannot be an essentially class consistent partial learner of $REC_{0,1}$, and so $REC_{0,1}$ is not essentially class consistently partially learnable, as required. ♦

The example furnished in the subsequent result shows that behaviourally correct learning is in fact a strictly weaker learning notion than essentially class consistent partial learning.

Theorem 53 *There is a class of recursive functions which is essentially class consistently partially learnable but not behaviourally correct learnable.*

Proof. Consider the class of recursive functions $\mathcal{C} = \{f : f \text{ is recursive} \wedge \forall x[f(x) \downarrow = \varphi_{f(0)}(x) \downarrow]\} \cup \{f : f \text{ is recursive} \wedge \forall^{\infty} x[f(x) \downarrow = 0]\}$, the union of the self-describing recursive functions with the recursive functions which are almost everywhere equal to 0. $\mathcal{C}$ is essentially class consistently partially learnable via the following algorithm: on input $f(0) \circ f(1) \circ \ldots \circ f(n)$, the learner M identifies the least $k \leq n$ such that

$f(i) = 0$ for all $k \leq i \leq n$, if such a k exists; it then outputs the programme e with

$$\varphi_e(x) = \begin{cases} f(x) & \text{if } x < k; \\ 0 & \text{if } x \geq k. \end{cases}$$

Otherwise, if no such k exists, M outputs $f(0)$. It will then follow that M is an essentially class consistent partial learner of $\mathcal{C}$. The proof that $\mathcal{C}$ is not behaviourally correctly learnable was carried out in Theorem 35. ♦

Although the specifications of an essentially class consistent partial learner may seem quite liberal, the next result demonstrates that its learning strength does not exceed that of confident partial learning.

Theorem 54 *There is a class of recursive functions which is confidently partially learnable but not essentially class consistently partially learnable.*

Proof 1. Let $M_0, M_1, M_2, \ldots$ be an enumeration of all partial-recursive learners. The following construction of a class of recursive functions which diagonalises against all essentially class consistent learners mirrors the procedure used to build the recursive functions in the preceding claim. First, for each number e, let $g(e)$ be a programme for the partial-recursive function $\varphi_{g(e)}$ which is defined as follows. One determines in the limit a sequence of strings $\sigma_{e,0}, \sigma_{e,1}, \sigma_{e,2}, \ldots$ which satisfy the following conditions for all i.

- $\sigma_{e,0} = e$;
- $\sigma_{e,i} \preceq \sigma_{e,i+1}$;

- If $\sigma_{e,i} \prec \sigma_{e,i+1}$, that is, $\sigma_{e,i+1}$ is a proper string extension of $\sigma_{e,i}$, then σ_{i+1} is the first string found such that for all $x \geq |\sigma_i|$, either $\varphi_{M_e(\sigma_{e,i+1})}(x) \downarrow \neq \sigma_{e,i+1}(x) \downarrow$ holds, or $M_e(\sigma_{e,i+1}[x]) > M_e(\tau)$ whenever $\tau \prec \sigma_{e,i+1}[x]$; here $\sigma_{e,i+1}[x]$ denotes the prefix of $\sigma_{e,i+1}$ with length $x+1$.

The partial-recursive function $\varphi_{g(e)}$ is defined by setting, for all x,
$\varphi_{g(e)}(x) = \sigma_{e,j}(x)$ whenever j is an index such that $x \in \mathrm{dom}(\sigma_{e,j})$; if no such $\sigma_{e,j}$ exists, then $\varphi_{g(e)}$ remains undefined on the input x.

Let $\mathcal{C}_1 = \{\varphi_{g(e)} : e \in \mathbb{N} \wedge \varphi_{g(e)} \text{ is total}\}$.

Secondly, for each number e and string $\eta \in \mathbb{N}^*$, one constructs inductively a sequence $\tau_{e,0}, \tau_{e,1}, \tau_{e,2}, \ldots$ of strings such that the following conditions hold for all i.

- $\tau_{e,0} = e \circ \eta$;
- $\tau_{e,i} \preceq \tau_{e,i+1}$;
- If z is the first number found such that $M_e(\tau_{e,i} \circ z) > M_e(\theta)$ for all $\theta \preceq \tau_{e,i}$, then $\tau_{e,i+1} = \tau_{e,i} \circ z$; otherwise, if (x, y) is the first pair of numbers found with $x < y$ and $M_e(\tau_{e,i} \circ x) = M_e(\tau_{e,i} \circ y)$, then $\tau_{e,i+1} = \tau_{e,i} \circ x$.

Let $h(\langle e, \sigma \rangle)$ be the programme for the partial-recursive function $\varphi_{h(\langle e,\sigma\rangle)}$ such that for all x, $\varphi_{h(\langle e,\sigma\rangle)}(x) \downarrow = \tau_{e,j}(x) \downarrow$, where j is any index with
$x \in \mathrm{dom}(\tau_{e,j})$; if no such $\tau_{e,j}$ exists, then $\varphi_{h(\langle e,\sigma\rangle)}$ remains undefined on x.

Define $\mathcal{C}_2 = \{\varphi_{h(\langle e,\eta\rangle)} : e \in \mathbb{N} \wedge \eta \in \mathbb{N}^* \wedge M_e \text{ is total}\}$.

To finish the construction, let $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$. It shall be shown that $\mathcal{C}$ is confidently partially learnable but not essentially class consistently partially learnable.

Define a recursive learner M as follows. On the input $\xi = e \circ \tau$, M simulates the programme $g(e)$ and determines the sequence $\sigma_{e,0}, \sigma_{e,1}, \ldots, \sigma_{e,|\xi|}$ constructed in the algorithm. M then carries out the first of the following instructions which applies.

1. If $\sigma_{e,|\xi|}(x)\downarrow = \xi(x)\downarrow$ for all $x \in \mathrm{dom}(\sigma_{e,|\xi|}) \cap \mathrm{dom}(\xi)$, and $\sigma_{e,|\xi|-1} \neq \sigma_{e,|\xi|}$, then M outputs the index $g(e)$.

2. If $\sigma_{e,|\xi|}(x)\downarrow = \xi(x)\downarrow$ for all $x \in \mathrm{dom}(\sigma_{e,|\xi|}) \cap \mathrm{dom}(\xi)$, but $\sigma_{e,|\xi|-1} = \sigma_{e,|\xi|}$, then M outputs the index $h(\langle e, \alpha \rangle)$, where $\alpha = \sigma_{e,|\xi|}$ if $\xi \preceq \sigma_{e,|\xi|}$, and if $\sigma_{e,|\xi|} \prec \xi$, α is the shortest string such that $\sigma_{e,|\xi|} \preceq \alpha \preceq \xi$ and $\varphi_{h(\langle e,\alpha \rangle),|\xi|} \subseteq \xi$. If such an α does not exist, M outputs $g(e)$. Furthermore, if case 2. applied at the last stage and M had output $h(\langle e, \alpha' \rangle)$ for some $\alpha' \neq \alpha$, then M conjectures $g(e)$ once before outputting $h(\langle e, \alpha \rangle)$ at the subsequent stage.

3. If $\sigma_{e,|\xi|}(x)\downarrow \neq \xi(x)\downarrow$ for some $x \in \mathrm{dom}(\sigma_{e,|\xi|}) \cap \mathrm{dom}(\xi)$, then M outputs the index $h(\langle e, \theta \rangle)$, where θ is the shortest prefix of ξ such that $\varphi_{h(\langle e,\theta \rangle),|\xi|} \subseteq \xi$. If such a prefix does not exist, or if case 3. applied at the last stage with a different $\theta' \prec \xi$ satisfying $\varphi_{h(\langle e,\theta' \rangle),|\xi|-1} \subseteq \xi[|\xi|-2]$, then M outputs $g(e)$ once before outputting $h(\langle e, \theta \rangle)$ at the subsequent stage.

Suppose that M is presented with the canonical text for $\varphi_{g(e)}$, where $\varphi_{g(e)}$ is assumed to be total. Then there are infinitely many i such that $\sigma_{e,i} \neq \sigma_{e,i+1}$; furthermore, for all x, there is a j for which $\varphi_{g(e)}(x)\downarrow = \sigma_{e,j}(x)\downarrow$. Hence case 1. applies infinitely often, and so M outputs $g(e)$ infinitely often. On the other hand, for each i, since there are only finitely many $\sigma_{e,j}$ with $\sigma_{e,i} = \sigma_{e,j}$, M conjectures each index of the form $h(\langle e, \alpha \rangle)$ only finitely often.

Suppose next that one feeds M with the canonical text for $\varphi_{h(\langle e,\eta \rangle)}$, where M_e

is total. If $\varphi_{g(e)}$ is total and $\varphi_{g(e)} = \varphi_{h(\langle e,\eta\rangle)}$, then M outputs $g(e)$ infinitely often, and each index of the form $h(\langle e,\alpha\rangle)$ only finitely often. If $\varphi_{g(e)}$ is not total but agrees with $\varphi_{h(\langle e,\eta\rangle)}$ on its whole domain, then there is a k such that $\sigma_{e,k} = \sigma_{e,l}$ whenever $k \leq l$, and so case 2. will always apply after some stage, that is, M will converge syntactically to a correct index $h(\langle e,\alpha\rangle)$ for a fixed α. Finally, if $\varphi_{g(e)}(x) \downarrow \neq \varphi_{h(\langle e,\eta\rangle)}(x) \downarrow$ for some $x \in \mathrm{dom}(\varphi_{g(e)}) \cap \mathrm{dom}(\varphi_{h(\langle e,\eta\rangle)})$, then there is a stage after which case 3. will always hold, so that M converges syntactically to a fixed correct index $h(\langle e,\theta\rangle)$. This completes the verification that M is a confident partial learner of $\mathcal{C}$.

Now assume by way of contradiction that M_d were an essentially class consistent partial learner of $\mathcal{C}$. If $\varphi_{g(d)}$ is total, then it follows from the construction of the sequence $\sigma_{d,0}, \sigma_{d,1}, \sigma_{d,2}, \ldots$ that either $M_d(\varphi_{g(d)}[n]) > M_d(\tau)$ for all $\tau \prec \varphi_{g(d)}[n]$ holds for cofinitely many n, or for infinitely many x, there is a $\sigma_{d,k}$ with $\varphi_{M_d(\sigma_{d,k})}(x) \downarrow \neq \sigma_{d,k}(x) \downarrow$. Hence M_d is not an essentially class consistent learner of $\varphi_{g(d)}$. If $\varphi_{g(d)}$ is not total, and $\sigma_{d,k} = \sigma_{d,l}$ for all $l \geq k$, then $\varphi_{h(\langle e,\sigma_{d,k}\rangle)}$ is a total function such that there are arbitrarily large x satisfying $\varphi_{M_d(\varphi_{h(\langle e,\sigma_{d,k}\rangle)}[x])}(x) \uparrow$, so M_d does not essentially class consistently learn $\varphi_{h(\langle d,\sigma_{d,k}\rangle)}$. This establishes that the class $\mathcal{C}$ is confidently partially learnable but not essentially class consistently partially learnable. ♦

Proof 2. Let $M_0, M_1, M_2, \ldots$ be a recursive enumeration of all partial-recursive learners.

For each M_e define a function $\varphi_{g(e)}$ by starting with $\sigma_{e,0} = e$ and taking $\sigma_{e,k+1}$ to be the first extension of $\sigma_{e,k}$ found such that $M_e(\sigma_{e,k+1})$ outputs an index d with

$\varphi_d(x) \downarrow \neq \sigma_{e,k+1}(x)$ for some $x < |\sigma_{e,k+1}|$. $\varphi_{g(e)}(x)$ takes as value $\sigma_{e,k}(x)$ for the first k found where this is defined.

Furthermore, for each e, k where $\sigma_{e,k}$ is defined, let $\varphi_{h(e,k)}$ be the partial recursive function ψ extending $\sigma_{e,k}$ such that for all $x \geq |\sigma_{e,k}|$, $\psi(x)$ is the least a such that either $M_e(\psi(0) \circ \psi(1) \circ \ldots \circ \psi(x-1) \circ a) > x$ or $M_e(\psi(0) \circ \psi(1) \circ \ldots \circ \psi(x-1) \circ a) = M_e(\psi(0) \circ \psi(1) \circ \ldots \circ \psi(x-1) \circ b)$ for some $b < a$.

Let $\mathcal{C}_1$ contain all those $\varphi_{g(e)}$ which are total and $\mathcal{C}_2$ contain all $\varphi_{h(e,k)}$ where M_e is total and $\varphi_{g(e)} = \sigma_{e,k}$, that is, the construction got stuck at stage k. The class $\mathcal{C}_1$ is obviously explanatorily learnable; for the class $\mathcal{C}_2$, an explanatory learner identifies first the e and then simulates the construction of $\varphi_{g(e)}$ and updates the hypothesis always to $h(e,k)$ for the largest k such that $\sigma_{e,k}$ has already been found. Hence both classes are explanatorily learnable, hence their union $\mathcal{C}$ is confidently partially learnable.

However $\mathcal{C}$ is not essentially class consistently partially learnable, as it is now shown. So consider a total learner M_e. If $\varphi_{g(e)}$ is total then M_e is inconsistent on this function infinitely often and so M_e does not essentially class consistently partially learn $\mathcal{C}$. So consider the k with $\varphi_{g(e)} = \sigma_{e,k}$. Note that the inductive definition of $\varphi_{h(e,k)}$ results in a total function. If M_e outputs on $\varphi_{h(e,k)}$ each index only finitely often, then M_e does not partially learn $\varphi_{h(e,k)}$. If M_e outputs an index d infinitely often, then for all sufficiently long $\tau \circ a \preceq \varphi_{h(e,k)}$ with $M_e(\tau \circ a) = d$ it holds that there is a $b < a$ with $M(\tau \circ b) = d$ as well. By assumption, $\sigma_{e,k+1}$ does not exist and can be neither $\tau \circ a$ nor $\tau \circ b$. Hence $\tau \circ a$ is not extended by φ_d and so M_e outputs an inconsistent index for almost all times where it conjectures d; again M_e does not essentially class consistently partially learn $\mathcal{C}$. ♦

Theorem 55 *Essentially class consistent learning is not closed under finite unions; that is, there are essentially class consistently partially learnable classes* $\mathcal{C}_1, \mathcal{C}_2$*, such that* $\mathcal{C}_1 \cup \mathcal{C}_2$ *is not essentially class consistently partially learnable.*

Proof. Take $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, where $\mathcal{C}_1$ and $\mathcal{C}_2$ are defined according to Proof 1. in the preceding theorem. $\mathcal{C}_1$ is finitely learnable, while $\mathcal{C}_2$ is behaviourally correctly learnable: on every input $\xi = e \circ \tau$, a finite learner of $\mathcal{C}_1$ may output $g(e)$, and a behaviourally correct learner of $\mathcal{C}_2$ may output $h(\langle e, \tau \rangle)$. Consequently, by Theorem 48, both $\mathcal{C}_1$ and $\mathcal{C}_2$ are essentially class consistently partially learnable. However, as was shown in Proof 1. of Theorem 54, the union $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ is not essentially class consistently partially learnable. ⧫

In [13], it is shown that REC is consistently partially learnable relative to an oracle A if and only if A is hyperimmune. The theorem below asserts that a recursive learner with access to a PA-complete oracle may essentially class consistently partially learn REC. Since the class of hyperimmune-free, PA-complete degrees is nonempty, as demonstrated in [14], one may conclude that for partial learning, essential class consistency is indeed a weaker criterion than general consistency, even when learning with oracles.

Theorem 56 *If* A *is a PA-complete set, then* $REC_{0,1}$ *is essentially class consistently partially learnable using* A *as an oracle.*

Proof. Let $\psi_0, \psi_1, \psi_2, \ldots$ be a one-one numbering of the recursive functions plus the functions with finite domain. For example, Kummer [16] provides such a numbering. Let g be a recursive function such that $\psi_e = \varphi_{g(e)}$ for all e. There is a recursive sequence $(e_0, x_0, y_0), (e_1, x_1, y_1), \ldots$ of pairwise distinct triples such that $\psi_e(x) \downarrow = y$ iff the triple (e, x, y) appears in this sequence.

On input $\sigma = f(0) \circ f(1) \circ \ldots \circ f(n)$, the learner M searches for the first $s \geq n$ such that for all $t \leq s$ either $e_t \neq e_s$ or $x_t > n$ or $y_t = f(x_t)$; that is, s is the first stage where ψ_{e_s} — to the extent it can be judged from the triples enumerated until stage s — is consistent with σ. Then M determines using the PA-complete oracle an $d \leq e_s$ such that either ψ_d extends σ or there is no $c \leq e_s$ such that ψ_c extends σ; note that in that second case the oracle can provide "any false d" below e. The learner conjectures then $g(d)$ for the index d determined this way.

If now e is the unique ψ-index of the function f to be learnt, then for all sufficiently long inputs σ, the above e_s satisfies $e_s \geq e$ as for each $d < e$ either there are only finitely many triples having d in the first component with all of them appearing before n or there is a $t \leq n$ with $e_t = d \wedge x_t \leq n \wedge y_t \neq f(x_t)$. Hence, the s selected satisfies $e_s \geq e$ and therefore the d provided satisfies that ψ_d extends σ. Furthermore, there are infinitely many n with $e_n = e$ and for those the choice is $s = n$ and, if n is sufficiently large, $d = e$. Hence the learner outputs infinitely often $g(e)$ and almost always an index $g(d)$ with $\varphi_{g(d)}$ being consistent with the input seen so far. ♦

Theorem 57 *Every class consistently partially learnable class of recursive functions can be extended to a one-one numbering of partial-recursive functions $\{f_0, f_1, f_2, \ldots\}$ such that the subclass of all recursive functions in $\{f_0, f_1, f_2, \ldots\}$ is class consistently partially learnable. The same statement holds with essentially class consistent partial learning in place of class consistent partial learning.*

Proof. Let M be a recursive class consistent learner of the class $\mathcal{C}$. For each number e, build a partial-recursive function $\varphi_{g(e)}$ with the following property: for all x, $\varphi_{g(e)}(x)\downarrow = \varphi_e(x)\downarrow$ if and only if there is a $z \geq x$ such that $\varphi_e(w)\downarrow = \varphi_{M(\varphi_e[y])}(w)\downarrow$ for all $w \leq y$ and $y \leq z$, and $M(\varphi_e[z]) = e$. If there is an x which does not fulfil the preceding condition, then $\varphi_{g(e)}$ remains undefined for all $y \geq x$. Now let $g(j(0)), g(j(1)), g(j(2)), \ldots$ be a one-one enumeration of all the indices in $I = \{g(e) : \varphi_{g(e)}(0)\downarrow\}$. Corresponding to each index $g(j(e)) \in I$, consider the sequence $pad(M(\varphi_{g(j(e))}(0)), k_0), pad(M(\varphi_{g(j(e))}[1]), k_1)$, $pad(M(\varphi_{g(j(e))}[2]), k_2), \ldots$, where k_i is the number of times that M has already output an index less than $M(\varphi_{g(j(e))}[i])$ up to the ith term of the sequence. Next, construct a class of partial-recursive functions $\{\varphi_{h(e,a)}\}$ with indices e and a in a similar manner to that of the functions $\varphi_{g(e)}$: for all x, $\varphi_{h(e,a)}(x)\downarrow = \varphi_a(x)\downarrow$ holds if and only if there is a $z \geq x$ such that $a = pad(M(\varphi_{g(j(e))}[z]), k_z)$, and for all $y \leq z$, $\varphi_{g(j(e))}(w)\downarrow = \varphi_a(w)\downarrow = \varphi_{pad(M(\varphi_{g(j(e))}[y]),k_y)}(w)\downarrow$ whenever $w \leq y$; otherwise, $\varphi_{h(e,a)}$ remains undefined for all $l \geq x$. Finally, let $h(e_0, a_0), h(e_1, a_1), h(e_2, a_2), \ldots$ be a one-one enumeration of all the indices in $I' = \{h(e,a) : \varphi_{h(e,a)}(0)\downarrow\}$.

We claim that $\varphi_{h(e_0,a_0)}, \varphi_{h(e_1,a_1)}, \varphi_{h(e_2,a_2)}, \ldots$ is a one-one numbering such that the subclass of all recursive functions in this numbering is class consistently partially learnable. Consider any two distinct pairs of indices (e, a) and (d, b). Assume first that $a \neq b$. One of the following cases must hold.

Case (A): $\varphi_{h(e,a)}$ *and* $\varphi_{h(d,b)}$ *both have finite domains, up to some numbers* n_0 *and* n_1 *respectively.*

It follows from the above construction that $a = pad(M(\varphi_{g(j(e))}[n_0]), k_{n_0})$ and $b = pad((M(\varphi_{g(j(d))}[n_1]), k_{n_1}))$, but since $a \neq b$, $\varphi_{h(e,a)} \neq \varphi_{h(d,b)}$.

Case (B): *One of the partial-recursive functions, $\varphi_{h(e,a)}$ or $\varphi_{h(d,b)}$, has finite domain while the other has infinite domain, so that they cannot be equal.*

Case (C): *Both $\varphi_{h(e,a)}$ and $\varphi_{h(d,b)}$ have infinite domains.*

If $\varphi_{g(j(e))} = \varphi_{g(j(d))}$, then $\varphi_{h(e,a)}$ has infinite domain if and only if a is the minimum index that M outputs infinitely often on the canonical text for $\varphi_{g(j(e))}$; since $a \neq b$, the conclusion that $\varphi_{h(e,a)} \neq \varphi_{h(e,b)}$ again follows. Furthermore, by the consistency condition of M on the text for $\varphi_{g(j(e))}$, if $\varphi_{h(e,a)}$ has infinite domain, then $\varphi_{g(j(e))}(x) \downarrow = \varphi_a(x) \downarrow$ for all x. If $\varphi_{g(j(e))} \neq \varphi_{g(j(d))}$, then, since $\varphi_{h(e,a)}$ and $\varphi_{h(d,b)}$ both have infinite domains, one has $\varphi_{h(e,a)} = \varphi_{g(j(e))}$ and $\varphi_{h(d,b)} = \varphi_{g(j(d))}$, and therefore $\varphi_{h(e,a)} \neq \varphi_{h(d,b)}$.

This completes the verification that $\varphi_{h(e_0,a_0)}, \varphi_{h(e_1,a_1)}, \varphi_{h(e_2,a_2)}, \ldots$ is a one-one numbering. A class consistent partial learning strategy for all the recursive functions in this numbering is to output, given the data $f[n]$, the index $pad(M(f[n]), k_n)$, where k_n again denotes the number of l's such that $l \leq n$ and $M(f[l]) < M(f[n])$. An analogous proof shows that this result also holds when M is an essentially class consistent partial learner; in this case, the recursive functions in the one-one numbering will be essentially class consistently learnable. ♦

It is unknown at present whether or not the converse of Theorem 56 holds: that is, whether every oracle relative to which *REC* is essentially class consistently partially learnable must necesssarily be PA-complete. The following definition of *weak PA-completeness* proposes a streamlined alternative to PA-completeness, but no explicit construction of a set possessing the specified properties has been found so far.

Definition. A set A is *weakly PA-complete* if and only if there is an A-recursive function g^A such that for all n, indices $e_1, e_2, \ldots, e_n$, infinite recursive sets R, and all $f \in REC$, the following conditions hold.

- $f \in \{\varphi_{e_1}, \varphi_{e_2}, \ldots, \varphi_{e_n}\} \Rightarrow \exists x \in R[g^A(f(0) \circ f(1) \circ \ldots \circ f(x), e_1, e_2, \ldots, e_n) = e_i]$ for some $e_i \in \{e_1, e_2, \ldots, e_n\}$ with $f = \varphi_{e_i}$.
- For all x, $g^A(f(0) \circ f(1) \circ \ldots \circ f(x), e_1, e_2, \ldots, e_n) \in \{?, e_1, e_2, \ldots, e_n\}$, where ? is some default index.
- For all x and $\sigma \in \mathbb{N}^*$, if φ_{e_i} extends σ for some i with $1 \leq i \leq n$, and $g^A(\sigma, e_1, e_2, \ldots, e_n) = e_k$, then φ_{e_k} extends σ.

Proposition 58 *If A is hyperimmune, then A is weakly PA-complete.*

Proof. As A is hyperimmune, there is an A-recursive function h^A which is not dominated by any recursive function. Given any infinite recursive set R and recursive function $f = \varphi_{e_i}$, there is a programme $g(e_i)$ for the recursive function $\varphi_{g(e_i)}$ defined by $\varphi_{g(e_i)}(n) = \max(\{\Phi_{e_i}(y) : y \leq x_n\})$, where Φ denotes a fixed Blum complexity measure for the programming system φ, and $x_1, x_2, x_3, \ldots$ is a strictly increasing enumeration of R. Now consider the A-recursive function F^A defined by

$$F^A(\sigma(0) \circ \sigma(1) \circ \ldots \circ \sigma(x), e_1, e_2, \ldots, e_n) = \begin{cases} e_k & \text{if } k \text{ is the least number} \leq n \\ & \text{such that } \forall y \leq x[\varphi_{e_k, h^A(x)}(y) \downarrow = \sigma(y)]; \\ ? & \text{if no such } k \text{ exists.} \end{cases}$$

By the hyperimmune property of h^A, there are infinitely many numbers n such that $g^A(n) > \varphi_{g(e_i)}$. In other words, if f is a recursive function with $f = \varphi_{e_i}$

for some $e_i \in \{e_1, e_2, \ldots, e_n\}$, then there are infinitely many numbers $x_n \in R$ for which $\varphi_{e_i, g^A(n)}(y) \downarrow = f(y) \downarrow$ whenever $y \leq x_n$, so that for infinitely many $x \in R$, $F^A(f(0) \circ f(1) \circ \ldots \circ f(x), e_1, e_2, \ldots, e_n)$ is equal to some index for f contained in $\{e_1, e_2, \ldots, e_n\}$. Hence F^A satisfies the required properties for A to be weakly PA-complete. ♦

Theorem 59 *One has the m-reducibility $\{e : \varphi_e$ is total$\} \leq_m \{e : \varphi_e(0) \downarrow \wedge \forall x[\varphi_e(x) \downarrow = \varphi_{\varphi_e(0)}(x) \downarrow]\}$.*

Proof. Let g be a two-place recursive function such that for any numbers d, e, $\varphi_{g(d,e)}(0) \downarrow = d$, and for all $x > 0$, $\varphi_{g(d,e)}(x) \downarrow = 0$ iff for all $y \leq x$, $\varphi_e(y) \downarrow$. The domain of $\varphi_{g(e)}$ is thus an initial segment of $\mathbb{N}$ if φ_e is not total; otherwise the domain of $\varphi_{g(e)}$ is $\mathbb{N}$. By the generalized Recursion Theorem, there is a recursive function n such that for any e, $\varphi_{g(n(e),e)} = \varphi_{n(e)}$. Hence the required m-reducibility holds via the relation $e \in \{e : \varphi_e$ is total$\} \Leftrightarrow n(e) \in \{e : \varphi_e(0) \downarrow \wedge \forall x[\varphi_e(x) \downarrow = \varphi_{\varphi_e(0)}(x) \downarrow]\}$, and this establishes the claim. ♦.

The next question posed is whether, given any recursive learner M, there must always exist a uniform effective procedure to construct a recursive function f that M does not learn according to some stipulated criterion. An affirmative answer may offer a uniform method of constructing class separation examples for different learning criteria. The present work takes up this question in the context of confident as well as consistent partial learning of recursive functions.

Theorem 60 *There are recursive functions f and g such that for each n, if M_n is a recursive confident partial learner, and $\mathcal{C}_n$ is the class of all recursive functions that*

M_n confidently partially learns, then there is a $\sigma_n \in \mathbb{N}^*$ with either $\varphi_{f(\sigma_n)}$ recursive and $\varphi_{f(\sigma_n)} \notin \mathcal{C}_n$, or $\varphi_{g(\sigma_n)}$ recursive and $\varphi_{g(\sigma_n)} \notin \mathcal{C}_n$.

Proof. Let $\tau_0, \tau_1, \tau_2, \ldots$ be an enumeration of all sequences in $\mathbb{N}^*$. For each partial-recursive learner M_n, define $\varphi_{\tau_{k,n}}$ as follows.

- Stage 0. Set $\varphi_{f(\tau_{k,n})}(x) = \tau_k(x)$ and $\varphi_{g(\tau_{k,n})}(x) = \tau_k(x)$ for all $x < |\tau_k|$, $\varphi_{f(\tau_{k,n})}(|\tau_k|) = 0$, and $\varphi_{g(\tau_{k,n})}(|\tau_k|) = 1$.

- Stage s. Suppose that $\varphi_{f(\tau_{k,n})}$ and $\varphi_{g(\tau_{k,n})}$ have been defined up to a_s. Search, noneffectively, for string extensions $\theta_s, \eta_s \in \mathbb{N}^*$ for which $M_n(\varphi_{f(\tau_{k,n})}[a_s] \circ \theta_s) \downarrow= M_n(\varphi_{g(\tau_{k,n})}[a_s] \circ \eta_s) = M_n(\tau_k)$. Suppose that $|\theta_s| \geq |\eta_s|$. Set $\varphi_{f(\tau_{k,n})}(x) = \theta_s(x)$ for all x with $a_s < x \leq a_s + |\theta_s|$, $\varphi_{g(\tau_{k,n})}(x) = \eta_s(x)$ for all x with $a_s < x \leq a_s + |\eta_s|$, and $\varphi_{g(\tau_{k,n})}(x) = 1$ for all x with $a_s + |\eta_s| < x \leq a_s + |\theta_s|$. If $|\theta_s| < |\eta_s|$, then the roles of θ_s and η_s in the above constructions of $\varphi_{f(\tau_{k,n})}$ and $\varphi_{g(\tau_{k,n})}$ are interchanged.

Suppose that M_n is a recursive confident partial learner; this means that there is a string τ_k such that for all $\eta \in \mathbb{N}^*$, there is some $\theta \in \mathbb{N}^*$ for which $M_n(\tau_k \circ \eta \circ \theta) = M_n(\tau_k)$. Consequently, both the partial-recursive functions $\varphi_{f(\tau_{k,n})}$ and $\varphi_{g(\tau_{k,n})}$ constructed according to the above algorithm must be total. Furthermore, as $\varphi_{f(\tau_{k,n})}(|\tau_k|) \neq \varphi_{g(\tau_{k,n})}(|\tau_k|)$, but M_n outputs the same index $M_n(\tau_k)$ infinitely often on either of the canonical texts for these recursive functions, it must follow that at least one of $\varphi_{f(\tau_{k,n})}$ and $\varphi_{g(\tau_{k,n})}$ is not confidently partially learnt by M_n, and this establishes the required result. ♦

Theorem 61 *There are recursive functions f and g such that for each n, if M_n is* *a recursive consistent partial learner, and $\mathcal{C}_n$ is the class of all recursive functions that M_n consistently partially learns, then there is a $\sigma_n \in \mathbb{N}^*$ with either $\varphi_{f(\sigma_n)}$ recursive and $\varphi_{f(\sigma_n)} \notin \mathcal{C}_n$, or $\varphi_{g(\sigma_n)}$ recursive and $\varphi_{g(\sigma_n)} \notin \mathcal{C}_n$.*

Proof. Let M_n be any given partial-recursive learner. One defines a partial-recursive function $\varphi_{f(n)}$ in stages as follows.

- Stage 0. Search for a number x_0 such that $M_n(x_0)\downarrow$ and set $\varphi_{f(n)}(0) = \varphi_{g(n)}(0) = x_0$.
- Stage $s+1$. Search for either a number x_{s+1} such that $M_n(\varphi_{f(n)}[s] \circ x_{s+1})\downarrow > s$, or a pair of numbers y_{s+1}, z_{s+1} with $y_{s+1} \neq z_{s+1}$ such that $M_n(\varphi_{f(n)}[s] \circ y_{s+1})\downarrow = M_n(\varphi_{f(n)}[s] \circ z_{s+1})\downarrow$. If the first case applies, define $\varphi_{f(n)}(s+1) = \varphi_{g(n)}(s+1) = x_{s+1}$, and proceed to the next stage of the algorithm. If the second case applies, define $\varphi_{f(n)}(s+1) = y_{s+1}, \varphi_{g(n)}(s+1) = z_{s+1}$, $\varphi_{f(n)}(w) = \varphi_{g(n)}(w) = 0$ for all $w > s+1$, and terminate the algorithm.

It follows from the above construction that if M_n were a recursive consistent partial learner, then either $\varphi_{f(n)}, \varphi_{g(n)}$ are recursive functions on whose canonical texts M_n outputs each index only finitely often, or M_n is inconsistent on at least one of the canonical texts for $\varphi_{f(n)}$ and $\varphi_{g(n)}$. This establishes the required result. ♦

Theorem 62 *For every recursive function f such that $\varphi_{f(k)}$ is recursive for all k, there is an e for which M_e is a partial learner that consistently partially learns $\varphi_{f(e)}$.*

Proof. For each k, one can construct a partial learner $M_{g(k)}$ as follows. On the input $\sigma = g(0) \circ g(1) \circ \ldots \circ g(n)$, $M_{g(k)}$ first determines whether or not $\varphi_{f(k)}(x)\downarrow = g(x)$

for all $x \leq n$. If this condition holds, then $M_{g(k)}$ outputs $f(k)$. If there is a $y \leq n$ for which $\varphi_{f(k)}(y) \downarrow \neq g(y)$, $M_{g(k)}$ outputs an index for the partial-recursive function equal to $g(x)$ for all $x \leq n$, and equal to 0 on all inputs greater than n. By Kleene's Recursion Theorem, there must exist a partial learner M_e such that $M_{f(e)} = M(e)$; by the construction of $M_{f(e)}$, $M_{f(e)}$ consistently partially learns $\varphi_{f(e)}$, and so M_e also consistently partially learns $\varphi_{f(e)}$, as was required to be established. ♦

To wind up the discussion on consistent partial learning, we shall consider a learning situation in which the learner does not have access to the complete graph for some recursive function, and is instead tasked to output exactly one index infinitely often for some recursive extension of the partial-function generating the text.

Definition. An *incomplete text* for a recursive function f is an infinite sequence T in which $\langle x, f(x)\rangle$ occurs in T for cofinitely many x.

A recursive learner M consistently partially learns f from incomplete texts if and only if for all incomplete texts T_f for f and all m, $\varphi_{M(T[m])}(x) \downarrow = y$ holds whenever $\langle x, y\rangle \in \text{range}(T[m])$, and M outputs on T_f exactly one index e infinitely often such that φ_e is a recursive extension of range(T_f).

Theorem 63 *If the class $\{f : \forall x[f(x) \downarrow = \varphi_{f(0)}(x) \downarrow]\}$ of all self-describing recursive functions is class consistently partially learnable relative to the oracle A from incomplete texts, then REC is consistently partially learnable on canonical text relative to A.*

Proof. Let M^A be a recursive learner that consistently partially learns all self-describing recursive functions from incomplete texts relative to A. Define a new A-recursive learner N^A as follows: on input $\sigma = f(0) \circ f(1) \circ \ldots \circ f(n)$, N^A conjectures an index c for which

$$\varphi_c(x) = \begin{cases} f(0) & \text{if } x = 0; \\ \varphi_{M^A(f(1)\circ f(2)\circ\ldots\circ f(n))}(x) & \text{if } x \neq 0. \end{cases}$$

It shall first be shown that N^A must be consistent on all texts. Suppose that there is a number n such that $\varphi_{M^A(f(1)\circ\ldots f(n))}(k)\uparrow$ or $\varphi_{M^A(f(1)\circ\ldots f(n))}(k)\downarrow\neq f(k)$ for some k with $1 \leq k \leq n$. By Kleene's Recursion Theorem, there is an index e for which

$$\varphi_e(x) = \begin{cases} e & \text{if } x = 0; \\ f(x) & \text{if } 1 \leq x \leq n; \\ 0 & \text{if } x > n. \end{cases}$$

Then φ_e is a self-describing function, but M^A is inconsistent on an incomplete text for φ_e, a contradiction. Consequently, N^A is consistent on all texts, as claimed. Furthermore, as M^A outputs exactly one index infinitely often, N^A also outputs a single correct index on the given text for the recursive function infinitely often, giving that it is indeed a consistent partial learner of REC. ♦

Example 64 *The class $\mathcal{C} = \{f : f$ is recursive $\wedge$ $\forall^\infty x[f(x) = 0]\}$ is consistently partially learnable from incomplete texts.*

4.3 Iterative Partial Learning

The present section introduces a variant paradigm of partial learning under which a learner must base its conjecture only upon the current input data and its last hypothesis. Such a learner may also be termed "memory-limited" [22], the condition reflecting a constraint that is quite likely faced when dealing with the practical realities of language acquisition. Although a memory-limited learner may attempt to encode all the input data revealed so far into its last conjecture, the success of this strategy is contingent on the learner's own consistency, as the subsequent results demonstrate. A view suggested by the learning relations obtained below is that iterative learning may be less flexible compared to the other learning criteria defined so far.

Definition. An *iterative* learner is a partial-recursive function $M : (\mathbb{N} \cup \{\emptyset\}) \times \mathbb{N} \to \mathbb{N}$.

Let M be an iterative learner, and f be a given recursive function. Abbreviate the pair $\langle n, f(n) \rangle$ as $f(n)$. Define $M_f : \mathbb{N}^* \times \mathbb{N} \to \mathbb{N}$ recursively as follows:

- $M_f(\emptyset, f(0)) = M(\emptyset, f(0))$;
- $M_f(f[0], f(1)) = M(M_f(\emptyset, f(0)), f(1))$;
- $M_f(f[n+1], f(n+2)) = M(M_f(f[n], f(n+1)), f(n+2))$.

M is said to *partially learn* f if there is exactly one index e such that $\varphi_e = f$ and $M_f(f[k], f(k+1)) = e$ for infinitely many k.

Theorem 65 *Every consistently partially learnable class of recursive functions is consistently partially learnable by an iterative learner.*

Proof. Let $\mathcal{C}$ be a class of recursive functions which is consistently partially learnt by M. Define an iterative learner N as follows. First, let $N(\emptyset, f(0)) = M(f(0))$, $N(\emptyset, f(n)) = 0$, and $N(p, f(0)) = 0$ for all $p \in \mathbb{N}$ and $n > 0$. Secondly, given $k \in \mathbb{N}$,

N, on the input $(k, f(n+1))$, waits until the computations of $\varphi_k(0), \varphi_k(1), \ldots, \varphi_k(n)$ converge. N then outputs $M(\varphi_k(0) \circ \varphi_k(1) \circ \ldots \circ \varphi_k(n) \circ f(n+1))$. Since M is a consistent partial learner of $\mathcal{C}$, it follows that for all $f \in \mathcal{C}$, $\varphi_{N_f(f[n], f(n+1))}(x) \downarrow = f(x) \downarrow$ for all $x \leq n+1$; thus N codes the inputs $f(0), f(1), \ldots, f(n+1)$ into its current conjecture. Therefore N will output the same sequence of conjectures that M outputs on the canonical text $f(0) \circ f(1) \circ f(2) \circ \ldots$, implying that it also consistently partially learns $\mathcal{C}$. ♦

Theorem 66 *There is a class of recursive functions which is partially learnable by a total iterative learner but not behaviourally correctly learnable.*

Proof. Consider the class of recursive functions $\mathcal{C} = \{f : f \text{ is recursive } \wedge \exists a \exists^\infty k [f = \varphi_a \wedge f(k) = a \wedge (\forall b \neq a) |\{y : f(y) = b\}| < \infty]\}$. An iterative learning strategy is to output e on both of the inputs $(\emptyset, e)$, (k, e) for all $e, k \in \mathbb{N}$. As any $f \in \mathcal{C}$ outputs exactly one index for itself infinitely often, it follows that this algorithm guarantees that $\mathcal{C}$ is partially learnt. Now assume for a contradiction that some recursive learner N behaviourally correctly learns $\mathcal{C}$. By Kleene's Recursion Theorem, one can construct a recursive function φ_e as follows: at stage s, suppose that $\varphi_e(x) \downarrow$ for all $x < a_s$; run a search for a sequence $\sigma \in \mathbb{N}^*$ so that range$(\sigma) \subseteq \{m+1, m+2, m+3, \ldots\}$, where $m = \max(\{\varphi_e(x) : x < a_s\})$, and $\varphi_{N(\varphi_e(0) \circ \ldots \circ \varphi_e(a_s - 1) \circ \sigma)}(a_s + |\sigma|) \downarrow$. Then let $\varphi_e(a_s + x) = \sigma(x)$ for all $x < |\sigma|$, $\varphi_e(a_s + |\sigma|) = \varphi_{N(\varphi_e(0) \circ \ldots \circ \varphi_e(a_s - 1) \circ \sigma)}(a_s + |\sigma|) + 1$, and $\varphi_e(a_s + |\sigma| + 1) = e$. Every stage of this algorithm must terminate: for, assuming that the contrary holds at stage s, one can build another recursive function $\varphi_b \in \mathcal{C}$ such that if $p = \max(\{\varphi_b(x) : x < a_s\})$, then $b > p$ and $\varphi_b(x) = b$ for all $x \geq a_s$; in addition, $N_{\varphi_b[z]}(z+1) \uparrow$ for all $z \geq a_s$, implying that N fails to behaviourally correctly learn φ_b. Thus $\varphi_e \in \mathcal{C}$, but by direct construction, N does not converge to a correct hypothesis on the canonical text $\varphi_e(0) \circ \varphi_e(1) \circ \varphi_e(2) \circ \ldots$; this is the desired contradiction. ⧫

Theorem 67 *There is a class of recursive functions which is explanatorily learnable by a total iterative learner but not class consistently partially learnable.*

Proof. Let $\mathcal{C}$ be the class of recursive functions $\{f : f \text{ is recursive } \wedge$ $(m = \min(\text{range}(f)) \Rightarrow \forall x[f(x) \downarrow = \varphi_m(x) \downarrow])\}$, which was considered in the second proof of Theorem 45. It was shown (loc cit) that $\mathcal{C}$ is not class consistently partially learnable. $\mathcal{C}$, however, is explanatorily learnable by a total iterative learner: for any $e, d \in \mathbb{N}$, an iterative learner N, on the input $(\emptyset, e)$, may output e; on the input (d, e), N outputs $\min(\{d, e\})$. Consequently, on the canonical text for any $f \in \mathcal{C}$, N will converge in the limit to the minimum number in the range of f, which by the definition of $\mathcal{C}$ is an index for f. ⧫

Theorem 68 *There is a class of recursive functions which is explanatorily learnable but not partially learnable by an iterative learner.*

Proof. Consider the class $\mathcal{C} = \{f : f \text{ is recursive } \wedge \exists k > 0 \forall x[\varphi_{f(0)}(k) \uparrow$ $\wedge (x \neq k \Rightarrow \varphi_{f(0)}(x) \downarrow = f(x) \downarrow)]$. An explanatory learning strategy is as follows: on the input $f[n]$, the learner N searches for the least $x_s > 0$ such that $\varphi_{f(0),n}(x_s) \uparrow$; it then hypothesizes the index e with $\varphi_e(x_s) = f(x_s)$ and $\varphi_e(y) = \varphi_{f(0)}(y)$ for all $y \neq x_s$. Assume towards a contradiction that M were an iterative partial learner of $\mathcal{C}$. By Kleene's Recursion Theorem, there is a programme e for the partial-recursive function φ_e defined as follows.

- At the initial stage, set $\varphi_e(0) = e$.

- At stage $s+1$, suppose first that $\varphi_{e,s}$ has been defined on all $x \leq s$. Now one runs a search until either a number a_s is found such that $M_{\varphi_{e,s}}(\varphi_{e,s}[s], a_s) > M_{\varphi_{e,s}}(\varphi_{e,s}[k], \varphi_{e,s}(k+1))$ for all $k < s$, or there are distinct numbers b_s, c_s satisfying $M_{\varphi_{e,s}}(\varphi_{e,s}[s], b_s) = M_{\varphi_{e,s}}(\varphi_{e,s}[s], c_s)$. In the former case, $\varphi_e(s+1)$ is left undefined but one stores the value a_s for future use; the algorithm then proceeds to the next stage $s+2$. In the latter case, $\varphi_e(s+1)$ is also undefined, and $\varphi_e(y) \downarrow= 0$ for all $y > s+1$; the algorithm is then terminated.

- Secondly, suppose that $\varphi_{e,s}$ has been defined on $\{x : x \leq s\} - \{k\}$. There is a value a_k associated to the undefined position k; one then temporarily assigns the value a_k to $\varphi_e(k)$, and searches for either a number a_s or a pair of distinct numbers b_s, c_s satisfying exactly the same properties formulated in the preceding case. If the number a_s is found, $\varphi_e(k)$ is still left undefined, and $\varphi_e(s+1) \downarrow= a_s$; one then proceeds to the next stage $s+2$. If the pair of numbers b_s, c_s is found, then $\varphi_e(k)$ is assigned the value a_k, $\varphi_e(s+1) \uparrow$, and $\varphi_e(y) \downarrow= 0$ for all $y > s+1$; after which, the algorithm terminates.

In the first place, suppose that the algorithm terminates at some stage $s+1$. This occurs if and only if there is a pair of distinct numbers b_s, c_s so that $M_{\varphi_{e,s}}(\varphi_{e,s}[s], b_s) = M_{\varphi_{e,s}}(\varphi_{e,s}[s], c_s)$. Let f_0 and f_1 be recursive functions such that $f_i(x) \downarrow= \varphi_e(x) \downarrow$ for all $x \neq s+1$ and $i \in \{0,1\}$; furthermore, $f_0(s+1) = b_s$ and $f_1(s+1) = c_s$. Then $f_0, f_1 \in \mathcal{C}$, but since M outputs the same index infinitely often on the canonical texts for both of these functions, it cannot iteratively partially learn at least one of f_0, f_1. In the second place, suppose that the algorithm never terminates. Then φ_e is undefined on exactly one place k, and there is a value a_k associated to this position. Let f be the recursive function in $\mathcal{C}$ equal to φ_e on all inputs except k, and $f(k) = a_k$. Since M outputs a strictly increasing sequence of conjectures on the canonical text for f, it does not fulfil the requirements of a partial learner. Therefore $\mathcal{C}$ is not iteratively partially 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] Lenore Blum and Manuel Blum. Towards a mathematical theory of inductive inference. *Information and Control* 28 (1975): 125-155.

[4] Lorenzo Carlucci, John Case, and Sanjay Jain. Learning correction grammars. COLT 2007: 203-217.

[5] John Case, Sanjay Jain, and Arun Sharma. On learning limiting programs. COLT 1992: 193-202.

[6] Jerome Feldman. Some decidability results on grammatical inference and complexity. *Information and Control* 20 (1972): 244-262.

[7] Rusins Freivalds, Efim Kinber and Rolf Wiehagen. Inductive inference and computable one-one numberings. *Zeitschrift fuer mathematische Logik und Grundlagen der Mathematik* 28 (1982): 463-479.

[8] Mark A. Fulk. Prudence and other conditions on formal language learning. *Information and Computation* 85(1) (1990): 1-11.

[9] 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.

[10] Mark Gold. Language identification in the limit. *Information and Control* 10 (1967): 447-474.

[11] William Hanf. The Boolean algebra of logic. *Bulletin of the American Mathematical Society* 81 (1975): 587-589.

[12] Sanjay Jain, Daniel Osherson, James S. Royer and Arun Sharma. 1999. *Systems that learn: an introduction to learning theory.* Cambridge, Massachusetts.: MIT Press.

[13] Sanjay Jain and Frank Stephan. Consistent partial identification. COLT 2009: 135-145.

[14] Carl G. Jockusch, Jr and Robert I. Soare. $\prod_1^0$ classes and degrees of theories. *Transactions of the American Mathematical Society* 173 (1972): 33-56.

[15] Steffen Lange, Thomas Zeugmann and Shyam Kapur. Characterizations of monotonic and dual monotonic language learning. *Information and Computation* 120(2) (1995): 155-173.

[16] Martin Kummer. Numberings of $\mathcal{R}_1 \cup F$. *Computer Science Logic 1988, Springer Lecture Notes in Computer Science* 385 (1989): 166-186.

[17] Steffen Lange and Thomas Zeugmann. Language learning in dependence on the space of hypotheses. COLT 1993: 127-136.

[18] Steffen Lange and Thomas Zeugmann. A guided tour across the boundaries of learning recursive languages. GOSLER Final Report 1995: 190-258.

[19] Steffen Lange, Thomas Zeugmann, and Shyam Kapur. Monotonic and dual monotonic language learning. *Theoretical Computer Science* 155(2) (1996): 365-410.

[20] Steffen Lange and Thomas Zeugmann. Set-driven and rearrangement-independent learning of recursive languages. *Mathematical Systems Theory* 29(6) (1996): 599-634.

[21] Steffen Lange, Thomas Zeugmann, and Sandra Zilles. Learning indexed families of recursive languages from positive data: a survey. *Theoretical Computer Science* 397(1-3) (2008): 194-232.

[22] Eric Martin and Daniel N. Osherson. 1998. *Elements of scientific inquiry.* Cambridge, Massachusetts.: MIT Press.

[23] Piergiorgio Odifreddi. 1989. *Classical recursion theory, studies in logic and the foundations of mathematics*, volume 125. North-Holland, Amsterdam: Elsevier Science Publishing Co.

[24] 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.

[25] Hartley Rogers, Jr. 1987. *Theory of recursive functions and effective computability*. Cambridge, Massachusetts: MIT Press.

[26] Joseph R. Shoenfield. Degrees of models. *Journal of Symbolic Logic* 25 (1960): 233-237.

[27] Frank Stephan. *Recursion theory*. Manuscript, 2009.