Learning without Coding

Sanjay Jain[a,1], Samuel E. Moelius III[b,*], Sandra Zilles[c,2]

[a]*Department of Computer Science, National University of Singapore,*
Singapore 117417, Republic of Singapore
[b]*IDA Center for Computing Sciences,*
17100 Science Drive, Bowie, MD 20715-4300, USA
[c]*Department of Computer Science, University of Regina,*
Regina, Saskatchewan, Canada S4S 0A2

Abstract

Iterative learning is a model of language learning from positive data, due to Wiehagen. When compared to a learner in Gold's original model of language learning from positive data, an iterative learner can be thought of as *memory-limited*. However, an iterative learner can memorize *some* input elements by *coding* them into the syntax of its hypotheses. A main concern of this paper is: to what extent are such coding tricks *necessary*?

One means of preventing *some* such coding tricks is to require that the hypothesis space used be free of redundancy, i.e., that it be 1-1. In this context, we make the following contributions. By extending a result of Lange & Zeugmann, we show that many interesting and non-trivial classes of languages can be iteratively identified using a Friedberg numbering as the hypothesis space. (Recall that a Friedberg numbering is a 1-1 effective numbering of all computably enumerable sets.) An example of such a class is the class of pattern languages over an arbitrary alphabet. On the other hand, we show that there exists an iteratively identifiable class of languages that can*not* be iteratively identified using *any* 1-1 effective numbering as the hypothesis space.

We also consider an iterative-like learning model in which the computational component of the learner is modeled as an *enumeration operator*, as opposed to a partial computable function. In this new model, there are no hypotheses, and, thus, no syntax in which the learner can encode what elements it has or has not yet seen. We show that there exists a class of languages that *can* be identified under this new model, but that can*not* be iteratively identified. On the other hand, we show that there exists a class of languages that can*not* be identified under this new model, but that *can* be iteratively identified using a Friedberg numbering as the hypothesis space.

*corresponding author

Email addresses: `sanjay@comp.nus.edu.sg` (Sanjay Jain), `semoeli@super.org` (Samuel E. Moelius III), `zilles@cs.uregina.ca` (Sandra Zilles)

---

1. Introduction

Iterative learning (**It**-learning, Definition 1(a)) is a model of language learning from positive data, due to Wiehagen [Wie76]. Like many models based on positive data, the **It**-learning model involves a learner that is repeatedly fed elements drawn from $\{\#\}$ and from some unknown target language $L \subseteq \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers, $\{0, 1, 2, ...\}$.[3] After being fed each such element, the learner outputs a hypothesis (provided that the learner does not diverge). The learner is said to *identify* the target language L iff there is some point from whence on the learner outputs always the same hypothesis, and that hypothesis corresponds to L. Furthermore, the learner is said to identify a *class* of languages $\mathcal{L}$ iff the learner identifies *each* $L \in \mathcal{L}$ when fed the elements of L (and possibly $\#$).

In the **It**-learning model, the learner itself is modeled as a triple.

- The first element of the triple is a two-place partial computable function, whose arguments are, respectively, the learner's most recently output hypothesis, and the next input element.

- The second element of the triple is a preliminary hypothesis, i.e., the hypothesis output by the learner before being fed any input.

- The third element of the triple is a *hypothesis space*. The *hypothesis space* determines the language that corresponds to each of the learner's hypotheses. Formally, a hypothesis space is a numbering $(X_j)_{j \in \mathbb{N}}$ of some collection of subsets of $\mathbb{N}$, and that is *effective* in the sense that the two-place predicate $\lambda j, x . [x \in X_j]$ is partial computable.[4]

It-learning is a special case of Gold's original model of language learning from positive data [Gol67]. In Gold's original model, the learner is provided access to all previously seen input elements, in addition to the next input element. In this sense, a learner in Gold's model can be thought of as *memorizing* all previously seen input elements. When compared to learners in Gold's model, iterative learners are restricted in terms of the classes of languages that they can identify.[5] In this sense, the *memory-limited* aspect of iterative learners is a *true* restriction, and *not* a mere superficial difference in definitions.

---

[3] The symbol '$\#$' is pronounced "pause". The inclusion of $\#$ in the model allows the target language L to be empty, i.e., in such a case, the learner is repeatedly fed $\#$.

[4] *Not-necessarily-effective* hypothesis spaces have also been considered [dBY10]. However, such hypothesis spaces are not needed herein. For the remainder, we use the terms *hypothesis space* and *effective numbering* interchangeably.

[5] Many variants of the **It**-learning model have been considered, and have also been shown to be restricted in this sense [LZ96, CCJS07, JLMZ10].

This does not however mean that iterative learners are *memory-less*. In particular, an iterative learner can memorize *some* input elements by employing *coding tricks*, which we define (informally) as follows.

- A *coding trick* is any use by an iterative learner of the syntax of a hypothesis to determine what elements that learner has or has not yet seen.

The following is an example. Suppose that an iterative learner $(M, p, (X_j)_{j \in \mathbb{N}})$ identifies a class of languages $\mathcal{L}$. Further suppose that one desires a learner that identifies the class $\mathcal{L}'$, where

$$\mathcal{L}' = \mathcal{L} \cup \{L \cup \{0\} \mid L \in \mathcal{L}\}. \tag{1}$$

Such a learner $(M', p', (Y_k)_{k \in \mathbb{N}})$ may be obtained as follows. Let $(Y_k)_{k \in \mathbb{N}}$ be such that, for each j:

$$Y_{2j} = X_j; \qquad Y_{2j+1} = X_j \cup \{0\}. \tag{2}$$

Then, let M' be such that, for each $x \in (\mathbb{N} \cup \{\#\}) - \{0\}$:

$$
\begin{aligned}
M'(2j, \quad x) &= 2M(j, x); & M'(2j, \quad 0) &= 2M(j, 0) + 1; \\
M'(2j+1, x) &= 2M(j, x) + 1; & M'(2j+1, 0) &= 2M(j, 0) + 1.
\end{aligned}
\tag{3}
$$

It is easily seen that $(M', 2p, (Y_k)_{k \in \mathbb{N}})$ iteratively identifies $\mathcal{L}'$. Intuitively, M' simulates M, while using the least-significant bit of each hypothesis to *encode* whether or not M' has seen a 0. (Note the switch from even to odd hypotheses in the upper-right of (3).) Further note that, if $\mathcal{L}$ already contains languages for which 0 is a member, then there is *redundancy* in the hypothesis space $(Y_k)_{k \in \mathbb{N}}$. In particular, if $0 \in X_j$, then $Y_{2j} = Y_{2j+1}$. For such hypotheses, the least-significant bit affects *only* their syntax, and *not* their semantics.

This example demonstrates how coding tricks can at least *facilitate* the identification of a class of languages. A main concern of this paper is: to what extent are such coding tricks *necessary*?

One approach to preventing *some* such coding tricks is to require that the hypothesis space be free of redundancy, i.e., that it be 1-1. One means of doing this is to require that the hypothesis space be a *Friedberg numbering* [Fri58, Kum90]. A *Friedberg numbering* is a 1-1 effective numbering of all computably enumerable (ce) subsets of $\mathbb{N}$. The use of such numberings as hypothesis spaces was considered by Jain & Stephan [JS08].[6] They observed, for example, that $\mathcal{F}in$, the collection of all finite subsets of $\mathbb{N}$, can*not* be iteratively identified using a Friedberg numbering as the hypothesis space [JS08, Remark 25]. For the remainder, to **FrIt**-*identify* a class of languages $\mathcal{L}$ shall mean to iteratively identify $\mathcal{L}$ using a Friedberg numbering as the hypothesis space (see Definition 1(b)).

---

[6]Freivalds, et al. [FKW82] considered the use of Friedberg numberings as hypothesis spaces in the context of *function* learning.

Our first main result is to show that, despite this observation of Jain & Stephan, many interesting and non-trivial classes can be **FrIt**-identified. More specifically, we extend a result of Lange & Zeugmann [LZ96, Theorem 12] by showing that, for each class $\mathcal{L}$, if there exists a single hypothesis space witnessing that $\mathcal{L}$ is both uniformly decidable and computably finitely thick, then $\mathcal{L}$ can be **FrIt**-identified (Theorem 6). By comparison, Lange & Zeugmann showed that such a class can be **It**-identified.

We delay the definitions of the terms *uniformly decidable* and *computably finitely thick* to Section 3. In the meantime, however, we mention one significant application of our result. A *pattern language* [Ang80] is a type of language with applications to molecular biology (see, e.g., [SSS^{+}94]). Furthermore, the pattern languages naturally form classes that are **It**-identifiable by Lange & Zeugmann's result,[7] and, thus, are **FrIt**-identifiable, by ours.

We briefly recall the definition of a pattern language. Suppose that Σ is an alphabet, i.e., a non-empty, finite set of symbols. A *pattern* over Σ is a finite string whose symbols are drawn from Σ, and from some infinite collection of variables. The *language determined by* a pattern p (over Σ) is the set of all strings that result by substituting some non-empty string (over Σ) for each variable in p. A *pattern language* over Σ is any language determined by a pattern over Σ. $\mathcal{P}at^{\Sigma}$ denotes the class of pattern languages over Σ.

For example, suppose that $\Sigma = \{0, 1\}$, and that p is the pattern $x0x1y$, where x and y are variables. Then, $\mathcal{P}at^{\Sigma}$ includes the language determined by p, which, in turn, includes the following strings. (To lessen the burden upon the reader, we have underlined in each string a 0 and 1 that may be regarded as part of the original pattern.)

0000010	0001010	0001110	1001010	1011010	1011110
0000011	0001011	0001111	1001011	1011011	1011111
0001000	0001100	0100110	1011000	1011100	1101110
0001001	0001101	0100111	1011001	1011101	1101111

On the other hand, the language determined by p includes no other strings of length 7.

As the reader may have already noticed, if one's intent is simply to eliminate redundancy in the hypothesis space, then to require that the hypothesis space be a Friedberg numbering is really overkill. That is because to require that the hypothesis space be a Friedberg numbering is to require that it be free of redundancy *and* that it represent all of the ce sets.

Thus, we consider two variants of **FrIt**-learning, which we call *injective iterative learning* (**InjIt**-learning, Definition 1(c)) and *class injective iterative learning* (**ClassInjIt**-learning, Definition 1(d)). In the **InjIt**-learning learning model, the hypothesis space is required to be free of redundancy (i.e., be 1-1), but need not represent all of the ce sets.[8] The **ClassInjIt**-learning model is

---

[7]The pattern languages were first shown to be **It**-identifiable by Lange & Wiehagen [LW91].

[8]The use of 1-1 hypothesis spaces was also considered in [BCJS10] in the context of learning

similar; however, for a learner to **ClassInjIt**-identify a class of languages $\mathcal{L}$, it is additionally required the learner's hypothesis space represent $\mathcal{L}$ *exactly*. Clearly, every class of languages that can be **FrIt**-identified can be **InjIt**-identified, and, similarly, every class of languages that can be **ClassInjIt**-identified can be **InjIt**-identified. On the other hand, *Fin* can be **ClassInjIt**-identified, but, as per Jain & Stephan's observation mentioned above, *Fin* can*not* be **FrIt**-identified. Furthermore, if one lets K be the diagonal halting problem [Rog67], then

$$\overline{\mathcal{K}} \stackrel{\text{def}}{=} \left\{ \{x\} \mid x \in \overline{K} \right\} \tag{4}$$

can be **FrIt**-identified, but since no hypothesis space represents $\overline{\mathcal{K}}$ exactly, $\overline{\mathcal{K}}$ can*not* be **ClassInjIt**-identified.

A related notion that has been considered is *order independent learning* [BB75, Ful90]. A learner is said to *order independently identify* a language L iff there exists some hypothesis j for L such that, when fed *any* text for L, there is some point from whence on the learner outputs only j. In effect, such a learner may benefit from redundancy in the hypothesis space *in the near term*, but it cannot so benefit *in the limit*. For the remainder, to **OrdIndIt**-*identify* a class of languages $\mathcal{L}$ shall mean to iteratively identify $\mathcal{L}$ order independently (see Definition 1(e)). Clearly, every class of languages that can be **InjIt**-identified can be **OrdIndIt**-identified. Interestingly, we show that the converse also holds (Theorem 15). Thus, for each class $\mathcal{L}$, if there exists a learner for $\mathcal{L}$ that does not benefit from redundancy in the hypothesis space in the limit, then there exists a learner for $\mathcal{L}$ that does not benefit from redundancy in the hypothesis space *whatsoever*.

Our next model, which we call *extensional iterative learning* (**ExtIt**-learning, Definition 1(f)), also tries to limit the extent to which a learner can benefit from redundancy in the hypothesis space. The approach differs considerably from that of **OrdIndIt**-learning, however. For a learner to **ExtIt**-identify a class of languages, it is required that, when presented with equivalent hypotheses and identical input elements, the learner must produce equivalent hypotheses. More formally: suppose that, for some class of languages $\mathcal{L}$, the following conditions are satisfied.

- σ_0 is a non-empty sequence of elements drawn from $\{\#\}$ and from some language in $\mathcal{L}$.

- σ_1 is another non-empty sequence of elements drawn from $\{\#\}$ and from some (possibly distinct) language in $\mathcal{L}$.

- When fed all but the last elements of σ_0 and σ_1, respectively, the learner outputs hypotheses for the same language (though these hypotheses may differ syntactically).

- The last elements of σ_0 and σ_1 are identical.

---

certain *specific* classes of languages.

Then, for the learner to **ExtIt**-identify $\mathcal{L}$, it is required that:

- When fed *all* of σ_0 and σ_1, respectively, the learner outputs hypotheses for the same language (though these hypotheses may differ syntactically).

Clearly, if a learner identifies a class of languages using a 1-1 hypothesis space, then that learner satisfies the just above requirement. Thus, every class of languages that can be **InjIt**-identified can be **ExtIt**-identified. On the other hand, we show that there exists a class of languages that *can* be **ExtIt**-identified, but that can*not* be **InjIt**-identified (Theorem 16).

Before introducing our final model, let us recall the definition of an *enumeration operator* [Rog67, §9.7]. For now, we focus on enumeration operators of a particular type. A more general definition is given in Section 2.1.

Let $\mathcal{P}(\mathbb{N})$ be the powerset of $\mathbb{N}$, i.e., the collection of all subsets of $\mathbb{N}$. Let $\langle \cdot, \cdot \rangle$ be any pairing function, i.e., a computable, 1-1, onto function of type $\mathbb{N}^2 \to \mathbb{N}$ [Rog67, page 64]. Let $\hat{\#} = 0$, and, for each $x \in \mathbb{N}$, let $\hat{x} = x + 1$. Let $(D_j)_{j \in \mathbb{N}}$ be any 1-1, canonical enumeration of *Fin*.

An *enumeration operator* of type $\mathcal{P}(\mathbb{N}) \times (\mathbb{N} \cup \{\#\}) \to \mathcal{P}(\mathbb{N})$ is a mapping that is algorithmic in the following precise sense. To each enumeration operator Θ (of the given type), there corresponds a ce set H, such that, for each $X \subseteq \mathbb{N}$ and $x \in \mathbb{N} \cup \{\#\}$,

$$\Theta(X, x) = \big\{ y \mid \langle j, \langle \hat{x}, y \rangle \rangle \in H \ \wedge \ D_j \subseteq X \big\}. \tag{5}$$

Thus, given an *enumeration of* X, and given x, one can enumerate $\Theta(X, x)$ in the following manner.

- Enumerate H. For each element of the form $\langle j, \langle \hat{x}, y \rangle \rangle \in H$, if ever the finite set D_j appears in the enumeration of X, then list y into $\Theta(X, x)$.

Enumeration operators exhibit certain notable properties, including *monotonicity*. Intuitively, this means that an enumeration operator can tell from its set argument X what elements are in X, but it cannot tell from X what elements are in the *complement of* X. More is said about the properties of enumeration operators in Section 2.1.

The final model that we consider is called *iterative learning by enumeration operator* (**EOIt**-learning, Definition 1(g)). As the name suggests, the computational component of the learner is modeled as an enumeration operator, as opposed to a partial computable function. Specifically, the learner is modeled as a *pair*, where:

- The first element of the pair is an enumeration operator of type $\mathcal{P}(\mathbb{N}) \times (\mathbb{N} \cup \{\#\}) \to \mathcal{P}(\mathbb{N})$, whose arguments are, respectively, the learner's most recently output *language*, and the next input element.

- The second element of the pair is the learner's preliminarily output *language*, i.e., the language output by the learner before being fed any input. (We require that this preliminary language be ce.)

Thus, there are no hypotheses in this model. Since there are no hypotheses, there is no *syntax* in which the learner can encode what elements it has or has not yet seen.

The expulsion of hypotheses from the model has an additional consequence, and that is that the success criterion has to be adjusted. Specifically, we say that a learner in this model *identifies* a language L iff when fed the elements of L (and possibly #), there is some point from whence on the learner outputs only the *language L*. The success criterion for identifying a *class* of languages is adjusted similarly. This more liberal approach to language identification, in some sense, gives an advantage to learners in this model. In particular, there exists a class of languages that *can* be **EOIt**-identified, but that can*not* be **It**-identified (Proposition 20).

Interestingly, there also exists a class of languages that can*not* be **EOIt**-identified, but that *can* be **FrIt** and **ClassInjIt**-identified (Theorem 24). To help to see why, consider the following two scenarios. First, suppose that $(\mathcal{M}, X)$ is a learner in the enumeration operator model, and that Y is its most recently output language. Then, since $\mathcal{M}$ is an enumeration operator, $\mathcal{M}$ can tell from Y what elements are in Y, but it cannot tell from Y what elements are in the *complement of Y*. Next, consider the analogous situation for a conventional iterative learner. That is, suppose that $(M, p, (X_j)_{j \in \mathbb{N}})$ is such a learner, and that j is its most recently output hypothesis. Then, in many cases, M *can* tell from j what elements are in the complement of X_j. In this sense, one could say that a hypothesis implicitly encodes *negative* information about the language that it represents. (In fact, this phenomenon can clearly be seen in the proof of Theorem 24 below.)

A question to then ask is: is this a *coding trick*, i.e., is it the case that *every* learner that operates on hypotheses (as opposed to languages) is employing coding tricks? At present, we do not see a clear answer to this question. Thus, we leave it as a subject for further study.

The main points of the preceding paragraphs are summarized in Figure 1. The remainder of this paper is organized as follows. Section 2 covers preliminaries. Section 3 presents our results concerning uniformly decidable and computably finitely thick classes of languages. Section 4 presents our results concerning Friedberg, injective, class injective, order independent, and extensional iterative learning (**FrIt**, **InjIt**, **ClassInjIt**, **OrdIndIt**, and **ExtIt**-learning, respectively). Section 5 presents our results concerning iterative learning by enumeration operator (**EOIt**-learning).

2. Preliminaries

Computability-theoretic concepts not covered below are treated in [Rog67].

$\mathbb{N}$ denotes the set of natural numbers, $\{0, 1, 2, \ldots\}$. Lowercase math-italic letters (e.g., a, j, x), with or without decorations, range over elements of $\mathbb{N}$, unless stated otherwise. Uppercase italicized letters (e.g., A, J, X), with or without decorations, range over subsets of $\mathbb{N}$, unless stated otherwise. For each
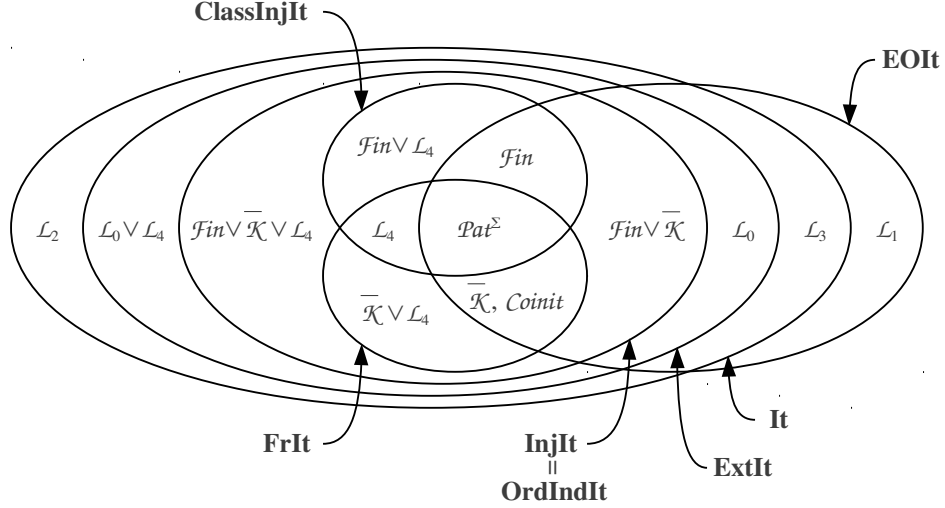
Figure 1: A summary of main results. $\mathcal{P}\!at^\Sigma$ is the class of pattern languages over Σ, where Σ is an arbitrary alphabet. $\mathcal{F}\!in$ is the collection of all finite subsets of $\mathbb{N}$. $\overline{\mathcal{K}}$ is defined as $\{\{x\} \mid x \in \overline{K}\}$, where K is the diagonal halting problem. $\mathcal{C}\!oinit$ is defined as $\{\mathbb{N} + e \mid e \in \mathbb{N}\}$. The remaining classes are defined in the proofs of the following results: $\mathcal{L}_0$, Theorem 16; $\mathcal{L}_1$, Proposition 20; $\mathcal{L}_2$, Theorem 22; $\mathcal{L}_3$, Theorem 23; and $\mathcal{L}_4$, Theorem 24. The operation '$\vee$' is defined in (6).

X and y, $X + y \overset{\text{def}}{=} \{x + y \mid x \in X\}$. For each non-empty X, $\min X$ denotes the minimum element of X. $\min \emptyset \overset{\text{def}}{=} \infty$. For each non-empty, finite X, $\max X$ denotes the maximum element of X. $\max \emptyset \overset{\text{def}}{=} -1$. For each non-empty, finite X, $X^- \overset{\text{def}}{=} X - \{\max X\}$.

$\mathcal{P}(\mathbb{N})$ denotes the powerset of $\mathbb{N}$, i.e., the collection of all subsets of $\mathbb{N}$. $\mathcal{P}(\mathbb{N})^m$ denotes the collection of all tuples of length m whose elements are drawn from $\mathcal{P}(\mathbb{N})$. Uppercase calligraphic letters (e.g., $\mathcal{L}$, X, $\mathcal{Y}$), with or without decorations, range over subsets of $\mathcal{P}(\mathbb{N})$, unless stated otherwise. For each X and $\mathcal{Y}$,

$$X \vee \mathcal{Y} \overset{\text{def}}{=} \{2X \mid X \in X\} \cup \{2Y + 1 \mid Y \in \mathcal{Y}\}. \tag{6}$$

$\mathcal{F}\!in$ denotes the collection of all finite subsets of $\mathbb{N}$. $(D_j)_{j \in \mathbb{N}}$ denotes a 1-1, canonical enumeration of $\mathcal{F}\!in$.

$\langle \cdot, \cdot \rangle$ denotes any fixed pairing function, i.e., a computable, 1-1, onto function of type $\mathbb{N}^2 \to \mathbb{N}$ [Rog67, page 64]. For each x, $\langle x \rangle \overset{\text{def}}{=} x$. For each $x_0, ..., x_{n-1}$, where $n > 2$, $\langle x_0, ..., x_{n-1} \rangle \overset{\text{def}}{=} \langle x_0, \langle x_1, ..., x_{n-1} \rangle \rangle$.

$\mathbb{N}_\# \overset{\text{def}}{=} \mathbb{N} \cup \{\#\}$. The function $\lambda x \in \mathbb{N}_\#.\hat{x}$ is such that $\hat{\#} = 0$, and, for each $x \in \mathbb{N}$, $\hat{x} = x + 1$. A *text* is a total function of type $\mathbb{N} \to \mathbb{N}_\#$. For each text t and $i \in \mathbb{N}$, $t[i]$ denotes the initial segment of t of length i. For each text t, $\text{content}(t) \overset{\text{def}}{=} \{t(i) \mid i \in \mathbb{N}\} - \{\#\}$. For each text t and $L \subseteq \mathbb{N}$, t is a text *for L* $\overset{\text{def}}{\Longleftrightarrow} \text{content}(t) = L$.

Seq denotes the set of all initial segments of texts. Lowercase Greek letters (e.g., ρ, σ, τ), with or without decorations, range over elements of Seq, unless stated otherwise. λ denotes the empty initial segment (equivalently, the everywhere divergent function). For each σ, $|\sigma|$ denotes the length of σ (equivalently, the size of the domain of σ). For each σ and $i \leq |\sigma|$, $\sigma[i]$ denotes the initial segment of σ of length i. For each σ, $\mathrm{content}(\sigma) \overset{\text{def}}{=} \{\sigma(i) \mid i < |\sigma|\} - \{\#\}$. For each σ and τ, $\sigma \cdot \tau$ denotes the concatenation of σ and τ. For each $\sigma \in \mathrm{Seq} - \{\lambda\}$:

$$\sigma^{-} \quad \overset{\text{def}}{=} \quad \sigma[|\sigma| - 1]; \tag{7}$$

$$\mathrm{last}(\sigma) \quad \overset{\text{def}}{=} \quad \sigma(|\sigma| - 1). \tag{8}$$

For each L and $\mathcal{L}$, $\mathrm{Txt}(L)$, $\mathrm{Txt}(\mathcal{L})$, $\mathrm{Seq}(L)$, and $\mathrm{Seq}(\mathcal{L})$ are defined as follows.

$$\mathrm{Txt}(L) \quad = \quad \{\, t \mid t \text{ is a text for } L \,\}. \tag{9}$$

$$\mathrm{Txt}(\mathcal{L}) \quad = \quad \{\, t \mid (\exists L \in \mathcal{L})[t \in \mathrm{Txt}(L)] \,\}. \tag{10}$$

$$\mathrm{Seq}(L) \quad = \quad \{\, \sigma \mid \mathrm{content}(\sigma) \subseteq L \,\}. \tag{11}$$

$$\mathrm{Seq}(\mathcal{L}) \quad = \quad \{\, \sigma \mid (\exists L \in \mathcal{L})[\sigma \in \mathrm{Seq}(L)] \,\}. \tag{12}$$

For each one-argument partial function ψ and $x \in \mathbb{N}$, $\psi(x)\!\downarrow$ denotes that $\psi(x)$ converges; $\psi(x)\!\uparrow$ denotes that $\psi(x)$ diverges. We use $\uparrow$ to denote the value of a divergent computation.

For each $\mathcal{X}$, a *numbering of $\mathcal{X}$* is an onto function of type $\mathbb{N} \to \mathcal{X}$. A numbering $(X_j)_{j \in \mathbb{N}}$ is *effective* $\overset{\text{def}}{\Leftrightarrow}$ the predicate $\lambda j, x \,\textbf{.}\, [x \in X_j]$ is partial computable. $\mathcal{EN}$ denotes the collection of all effective numberings.

$\mathcal{CE}$ denotes the collection of all computably enumerable (ce) subsets of $\mathbb{N}$. For each m and n, $\mathcal{PC}_{m,n}$ denotes the collection of partial computable functions mapping $\mathbb{N}^m \times \mathbb{N}_{\#}^n$ to $\mathbb{N}$. We shall be concerned primarily with $\mathcal{PC}_{1,0}$ and $\mathcal{PC}_{1,1}$. $(\varphi_p)_{p \in \mathbb{N}}$ denotes any fixed, acceptable numbering of $\mathcal{PC}_{1,0}$. For each i, $W_i \overset{\text{def}}{=} \{x \mid \varphi_i(x)\!\downarrow\}$. Thus, $(W_i)_{i \in \mathbb{N}}$ is an effective numbering of $\mathcal{CE}$. K denotes the diagonal halting problem, i.e., $\{x \mid x \in W_x\}$.

For each $M \in \mathcal{PC}_{1,1}$ and p, the partial function M_p^* is such that, for each $\sigma \in \mathrm{Seq}$ and $x \in \mathbb{N}_{\#}$:

$$M_p^*(\lambda) \quad = \quad p; \tag{13}$$

$$M_p^*(\sigma \cdot x) \quad = \quad \begin{cases} M\big(M_p^*(\sigma), x\big), & \text{if } M_p^*(\sigma)\!\downarrow; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{14}$$

For each text t and $j \in \mathbb{N}$, M_p^* *converges on t to j* $\overset{\text{def}}{\Leftrightarrow}$ there exists an i_0 such that $(\forall i \geq i_0)\big[M_p^*(t[i]) = j\big]$; M_p^* *diverges on t* $\overset{\text{def}}{\Leftrightarrow}$ M_p^* does not converge on t (to any j).

2.1. Enumeration Operators

An *enumeration operator* is a mapping of type $\mathcal{P}(\mathbb{N})^m \times \mathbb{N}_{\#}^n \to \mathcal{P}(\mathbb{N})$, for some m and n, and that is algorithmic in the following precise sense. To each

enumeration operator $\Theta : \mathcal{P}(\mathbb{N})^m \times \mathbb{N}_\#^n \to \mathcal{P}(\mathbb{N})$, there corresponds a ce set H, such that, for each $(X_0, ..., X_{m-1}) \in \mathcal{P}(\mathbb{N})^m$ and $(x_0, ..., x_{n-1}) \in \mathbb{N}_\#^n$,

$$
\begin{aligned}
&\Theta(X_0, ..., X_{m-1}, x_0, ..., x_{n-1}) \\
&\quad = \{y \mid \langle j_0, ..., j_{m-1}, \hat{x}_0, ..., \hat{x}_{n-1}, y \rangle \in H \ \wedge \ (\forall i < m)[D_{j_i} \subseteq X_i]\}.
\end{aligned}
\tag{15}
$$

A strategy for enumerating $\Theta(X_0, ..., X_{m-1}, x_0, ..., x_{n-1})$, given $X_0, ..., X_{m-1}$ and $x_0, ..., x_{n-1}$, can easily be generalized from that given for enumeration operators of type $\mathcal{P}(\mathbb{N}) \times \mathbb{N}_\# \to \mathcal{P}(\mathbb{N})$ in Section 1.

For each m and n, $\mathcal{EO}_{m,n}$ denotes the collection of all enumeration operators of type $\mathcal{P}(\mathbb{N})^m \times \mathbb{N}_\#^n \to \mathcal{P}(\mathbb{N})$. We shall be concerned primarily with $\mathcal{EO}_{1,0}$ and $\mathcal{EO}_{1,1}$.

Enumeration operators exhibit monotonicity and continuity properties [Rog67, Theorem 9-XXI], described below for $\mathcal{EO}_{1,1}$.

- *Monotonicity*: for each $\mathcal{M} \in \mathcal{EO}_{1,1}$, $X, Y \subseteq \mathbb{N}$, and $x \in \mathbb{N}_\#$,

$$
X \subseteq Y \ \Rightarrow \ \mathcal{M}(X, x) \subseteq \mathcal{M}(Y, x).
\tag{16}
$$

- *Continuity*: for each $\mathcal{M} \in \mathcal{EO}_{1,1}$, $X \subseteq \mathbb{N}$, $x \in \mathbb{N}_\#$, and $y \in \mathbb{N}$,

$$
y \in \mathcal{M}(X, x) \ \Rightarrow \ (\exists A \in \mathcal{F}in)[A \subseteq X \ \wedge \ y \in \mathcal{M}(A, x)].
\tag{17}
$$

For each $\mathcal{M} \in \mathcal{EO}_{1,1}$ and X, the function $\mathcal{M}_X^* : \mathrm{Seq} \to \mathcal{P}(\mathbb{N})$ is such that, for each $\sigma \in \mathrm{Seq}$ and $x \in \mathbb{N}_\#$:

$$
\begin{aligned}
\mathcal{M}_X^*(\lambda) &= X; \tag{18} \\
\mathcal{M}_X^*(\sigma \cdot x) &= \mathcal{M}\big(\mathcal{M}_X^*(\sigma), x\big). \tag{19}
\end{aligned}
$$

2.2. Iterative and Iterative-like Learning Models

The following are the formal definitions of the learning models described in Section 1. The symbols **Fr**, **Inj**, **ClassInj**, **OrdInd**, **Ext**, and **EO** are mnemonic for *Friedberg*, *injective*, *class injective*, *order independent*, *extensional*, and *enumeration operator*, respectively.

Definition 1. For each $\mathcal{L}$, (a)-(g) below. In parts (a)-(f), $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. In part (g), $(\mathcal{M}, X) \in \mathcal{EO}_{1,1} \times \mathcal{CE}$.

(a) **(Wiehagen [Wie76])** $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-*identifies* $\mathcal{L} \Leftrightarrow$ for each $t \in \mathrm{Txt}(\mathcal{L})$, there exists a j such that M_p^* converges on t to j and $X_j = \mathrm{content}(t)$.

(b) **(Jain & Stephan [JS08])** $(M, p, (X_j)_{j \in \mathbb{N}})$ **FrIt**-*identifies* $\mathcal{L} \Leftrightarrow (M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and $(X_j)_{j \in \mathbb{N}}$ is a Friedberg numbering.

(c) $(M, p, (X_j)_{j \in \mathbb{N}})$ **InjIt**-*identifies* $\mathcal{L} \Leftrightarrow (M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and $(X_j)_{j \in \mathbb{N}}$ is 1-1.

(d) $(M, p, (X_j)_{j \in \mathbb{N}})$ **ClassInjIt**-*identifies* $\mathcal{L} \Leftrightarrow (M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, $(X_j)_{j \in \mathbb{N}}$ is 1-1, and $\{X_j \mid j \in \mathbb{N}\} = \mathcal{L}$.

10

(e) $(M, p, (X_j)_{j \in \mathbb{N}})$ **OrdIndIt**-*identifies* $L \Leftrightarrow (M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies L, and, for each $L \in L$, there exists a j such that, for each $t \in \text{Txt}(L)$, M_p^* converges on t to j.

(f) $(M, p, (X_j)_{j \in \mathbb{N}})$ **ExtIt**-*identifies* $L \Leftrightarrow (M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies L, and, for each $\sigma_0, \sigma_1 \in \text{Seq}(L) - \{\lambda\}$,

$$[X_{M_p^*(\sigma_0^-)} = X_{M_p^*(\sigma_1^-)} \ \wedge \ \text{last}(\sigma_0) = \text{last}(\sigma_1)] \ \Rightarrow \ X_{M_p^*(\sigma_0)} = X_{M_p^*(\sigma_1)}. \tag{20}$$

(g) $(\mathcal{M}, X)$ **EOIt**-*identifies* $L \Leftrightarrow$ for each $t \in \text{Txt}(L)$, there exists an i_0 such that $(\forall i \geq i_0)[\mathcal{M}_X^*(t[i]) = \text{content}(t)]$.

Definition 2. Let **It** be as follows.

$$\textbf{It} = \big\{ L \mid \big(\exists (M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}\big)[(M, p, (X_j)_{j \in \mathbb{N}}) \ \textbf{It-identifies} \ L] \big\}.$$

Let **FrIt**, **InjIt**, **ClassInjIt**, **OrdIndIt**, **ExtIt**, and **EOIt** be defined similarly.

3. Uniform Decidability and Computable Finite Thickness

In this section, we extend a result of Lange & Zeugmann by showing that, for each class L, if there exists a single hypothesis space witnessing that L is both uniformly decidable and computably finitely thick, then L can be **FrIt**-identified (Theorem 6). We also show that there exists a class of languages that *is* uniformly decidable and computably finitely thick, but that is *not* in **It**, let alone **FrIt** (Theorem 9). Thus, one could not arrive at the conclusion of the just mentioned Theorem 6 if one were to merely require that: there exists a uniformly decidable effective numbering of L, and a possibly *distinct* computably finitely thick effective numbering of L.

The following are the formal definitions of the terms *uniformly decidable* and *computably finitely thick*. For additional background, see [LZZ08].

Definition 3.

(a) An effective numbering $(X_j)_{j \in \mathbb{N}}$ is *uniformly decidable* $\Leftrightarrow$ the predicate $\lambda j, x \,.\, [x \in X_j]$ is decidable.

(b) A class of languages L is *uniformly decidable* $\Leftrightarrow$ there exists a uniformly decidable effective numbering of L.

(c) An effective numbering $(X_j)_{j \in \mathbb{N}}$ is *computably finitely thick* $\Leftrightarrow$ there exists a computable function $f : \mathbb{N} \to \mathbb{N}$ such that, for each x,

$$\{X_j \mid j \in D_{f(x)}\} = \{X_j \mid x \in X_j\}. \tag{21}$$

(d) **(Lange & Zeugmann [LZ96, Definition 9])** A class of languages L is *computably finitely thick* $\Leftrightarrow$ there exists a computably finitely thick effective numbering of L.

N.B. In part (c) just above, the function f need *not* satisfy $D_{f(x)} = \{j \mid x \in X_j\}$. However, see Lemma 7 below.

Example 4.

(a) *Fin is* uniformly decidable, but is *not* computably finitely thick.

(b) *CE* is *neither* uniformly decidable *nor* computably finitely thick.

(c) The class $\big\{\{e\}, \{e\} \cup (W_e + e + 1) \mid e \in \mathbb{N}\big\}$ is *not* uniformly decidable, but *is* computably finitely thick.[9]

(d) The class

$$\mathit{Coinit} \overset{\text{def}}{=} \{\mathbb{N} + e \mid e \in \mathbb{N}\} \tag{22}$$

is both uniformly decidable and computably finitely thick. Moreover, there exists a single effective numbering witnessing both properties simultaneously.

(e) Let L be as follows.

$$L = \big\{\{e\} \mid e \in \mathbb{N}\big\} \cup \big\{\{e, \varphi_e(0) + e + 1\} \mid e \in \mathbb{N} \ \wedge \ \varphi_e(0)\!\downarrow\big\}. \tag{23}$$

Then, L is both uniformly decidable and computably finitely thick,[10] but there is *no* effective numbering of L witnessing both properties simultaneously. In fact, no such numbering exists for any class containing L.

The following result, due to Lange & Zeugmann, gives a sufficient condition for a class of languages to be **It**-identifiable.

Theorem 5 (Lange & Zeugmann [LZ96, Theorem 12]). For each L, if there exists an effective numbering of L that is both uniformly decidable and computably finitely thick, then $L \in$ **It**.[11]

The following result strengthens Theorem 5 (Lange & Zeugmann) just above.

Theorem 6. For each L, if there exists an effective numbering of L that is both uniformly decidable and computably finitely thick, then $L \in$ **FrIt**.

The proof of Theorem 6 relies on Lemma 7 just below.

Lemma 7. Suppose that L satisfies the conditions of Theorem 6, and that L is infinite. Then, there exists an effective numbering $(X'_j)_{j \in \mathbb{N}}$ of L satisfying (i) and (ii) below.

---

[9]The classes given in parts (c) and (e) of Example 4 can be shown to be computably finitely thick using a technique similar to that used in the proof of Theorem 9 below (see Figure 4(b), specifically).

[10]See footnote 9.

[11]In [LZ96], Theorem 12 is stated as follows.

Let $\mathcal{C}$ be any [uniformly decidable] class. If $\mathcal{C}$ has [computable] finite thickness, then $\mathcal{C} \in$ **It**.

$$(*)$$

Note that $(*)$ differs slightly from Theorem 5 above, in that $(*)$ does not require the existence a single effective numbering witnessing both properties simultaneously. However, based on the proof of this result, it is clear that such a requirement was intended. Furthermore, in light of Theorem 9 below, $(*)$ cannot be established in its literal form.

```
 0. if ∅ ∈ L then
 1.     X'₀ ← ∅;
 2.     j_max ← 0;
 3. else
 4.     j_max ← −1;
 5. end if;
 6. for x = 0, 1, ... do
 7.     let {j₁, ..., jₙ} = {j ∈ D_{f(x)} | x = min X_j};
 8.     (X'_{j_max+1}, ..., X'_{j_max+n}) ← (X_{j₁}, ..., X_{jₙ});
 9.     j_max ← j_max + n;
10.     D_{f'(x)} ← {j ≤ j_max | x ∈ X'_j};
11. end for.
```

Figure 2: The construction of $(X'_j)_{j \in \mathbb{N}}$ and $f' : \mathbb{N} \to \mathbb{N}$ in the proof of Lemma 7.

(i) $(X'_j)_{j \in \mathbb{N}}$ is uniformly decidable.

(ii) $(X'_j)_{j \in \mathbb{N}}$ satisfies the following *strong* form of computable finite thickness. There exists a computable function $f' : \mathbb{N} \to \mathbb{N}$ such that, for each x,

$$D_{f'(x)} = \{j \mid x \in X'_j\}. \tag{24}$$

Proof. Suppose that L satisfies the conditions of Theorem 6, as witnessed by $(X_j)_{j \in \mathbb{N}}$ and $f : \mathbb{N} \to \mathbb{N}$, and that L is infinite. Intuitively, the effective numbering $(X'_j)_{j \in \mathbb{N}}$ is constructed as follows. First, $\emptyset$ is listed into $(X'_j)_{j \in \mathbb{N}}$, if necessary. Then, each set in L whose smallest element is 0 is listed, followed by each set in L whose smallest element is 1, and so on. The construction maintains a variable j_{max}, which records the largest index used to list any such set. This variable is also used in the construction of $f' : \mathbb{N} \to \mathbb{N}$. Formally, $(X'_j)_{j \in \mathbb{N}}$ and f' are constructed as in Figure 2.

Clearly, $(X'_j)_{j \in \mathbb{N}}$ is a numbering of L. Furthermore, it is straightforward to show that $(X'_j)_{j \in \mathbb{N}}$ satisfies (i) in the statement of the lemma. To show that $(X'_j)_{j \in \mathbb{N}}$ satisfies (ii): let x be fixed, and consider the loop in lines 6-11 of Figure 2 for this value of x. Note that any set listed into $(X'_j)_{j \in \mathbb{N}}$ subsequent to this iteration of the loop will have a minimum element larger than x. Thus, immediately following line 9, it will be the case that

$$(\forall j)[x \in X'_j \;\Rightarrow\; j \le j_{max}]. \tag{25}$$

Clearly, then, f' satisfies (24). $\qquad\qquad\qquad\qquad$ □ (**Lemma 7**)

Proof of Theorem 6. Suppose that L satisfies the conditions of the theorem. The proof is straightforward for the case when L is finite. So, suppose that L is infinite. Let $(X_j)_{j \in \mathbb{N}}$ and $f : \mathbb{N} \to \mathbb{N}$ be as asserted to exist by Lemma 7 for L, e.g., witnessing the strong form of computable finite thickness of Lemma 7(ii). Let $F : \mathcal{Fin} \to \mathcal{P}(\mathbb{N})$ be such that, for each $A \in \mathcal{Fin}$,

$$F(A) = \bigcap \{D_{f(x)} \mid x \in A\}. \tag{26}$$

13

It is straightforward to show that, for each $A \in \mathit{Fin}$,

$$F(A) = \{j \mid A \subseteq X_j\}. \tag{27}$$

For each J, let X_J be such that

$$X_J = \bigcap \{X_j \mid j \in J\}. \tag{28}$$

For each $A \in \mathit{Fin}$, say that A is *exemplary* $\Leftrightarrow$

$$F(A) \neq \emptyset \ \wedge \ A = X_{F(A)} \cap \{0, ..., \max A\}. \tag{29}$$

Note that $F(\emptyset) = \mathbb{N}$, and, moreover, that $\emptyset$ is exemplary.

Claim 6.1. Suppose that A is exemplary, and that $x \leq \max A$. Let A' be such that

$$A' = A \cap \{0, ..., x\}. \tag{30}$$

Then, A' is exemplary.

Proof of Claim. Suppose that A, x, and A' are as stated. Since $A' \subseteq A$,

$$F(A') \supseteq F(A). \tag{31}$$

Thus, since $F(A) \neq \emptyset$, $F(A') \neq \emptyset$. To show that $A' = X_{F(A')} \cap \{0, ..., \max A'\}$:

$$
\begin{array}{rll}
A' & \subseteq \ X_{F(A')} \cap \{0, ..., \max A'\} & \{\text{immediate}\} \\
& \subseteq \ X_{F(A)} \cap \{0, ..., \max A'\} & \{\text{by (31)}\} \\
& \subseteq \ X_{F(A)} \cap \{0, ..., x\} & \{\text{because } \max A' \leq x\} \\
& \subseteq \ (X_{F(A)} \cap \{0, ..., x\}) \cap \{0, ..., x\} & \{\text{immediate}\} \\
& \subseteq \ (X_{F(A)} \cap \{0, ..., \max A\}) \cap \{0, ..., x\} & \{\text{because } x \leq \max A\} \\
& \subseteq \ A \cap \{0, ..., x\} & \{\text{because } A \text{ is exemplary}\} \\
& = \ A' & \{\text{by (30)}\}.
\end{array}
$$

$$\square \ (\textbf{Claim 6.1})$$

Claim 6.2. For each $A \in \mathit{Fin}$ and x, $X_{F(A)} \cap \{0, ..., x\}$ is exemplary.

Proof of Claim. Let $A \in \mathit{Fin}$ and x be fixed, and let $A' = X_{F(A)} \cap \{0, ..., x\}$. Note that, for each $j \in F(A)$,

$$A' \subseteq X_{F(A)} \subseteq X_j. \tag{32}$$

It follows that

$$F(A') \supseteq F(A). \tag{33}$$

Thus, since $F(A) \neq \emptyset$, $F(A') \neq \emptyset$. To show that $A' = X_{F(A')} \cap \{0, ..., \max A'\}$:

$$
\begin{array}{rll}
A' & \subseteq \ X_{F(A')} \cap \{0, ..., \max A'\} & \{\text{immediate}\} \\
& \subseteq \ X_{F(A)} \cap \{0, ..., \max A'\} & \{\text{by (33)}\} \\
& \subseteq \ X_{F(A)} \cap \{0, ..., x\} & \{\text{because } \max A' \leq x\} \\
& = \ A' & \{\text{by the choice of } A'\}.
\end{array}
$$

14

List $\emptyset$ into $(Z_\ell)_{\ell \in \mathbb{N}}$. Then, for each $A \in \mathcal{Fin} - \{\emptyset\}$, act according to the following conditions.

- COND. (a) [A is *not* exemplary $\wedge$ A^- is *not* exemplary]. Do nothing for A.

- COND. (b) [A is *not* exemplary $\wedge$ A^- is exemplary]. For each k, list $A \cup (Y_k + \max A + 1)$ into $(Z_\ell)_{\ell \in \mathbb{N}}$.

- COND. (c) [A is exemplary $\wedge$ $F(A) \subset F(A^-)$]. List $X_{F(A)}$ into $(Z_\ell)_{\ell \in \mathbb{N}}$, and, by letting $J = F(A)$, set $\zeta(J)$ to the index used to list this set.

- COND. (d) [A is exemplary $\wedge$ $F(A) = F(A^-)$]. List A^- into $(Z_\ell)_{\ell \in \mathbb{N}}$.

Figure 3: The construction of $(Z_\ell)_{\ell \in \mathbb{N}}$ in the proof of Theorem 6.

$\square$ (**Claim 6.2**)

Let $(Y_k)_{k \in \mathbb{N}}$ be any Friedberg numbering. An effective numbering $(Z_\ell)_{\ell \in \mathbb{N}}$ is constructed in Figure 3. Claims 6.3 and 6.5 below establish that $(Z_\ell)_{\ell \in \mathbb{N}}$ is a Friedberg numbering. In conjunction with $(Z_\ell)_{\ell \in \mathbb{N}}$, a partial computable function ζ from $\mathcal{Fin} - \{\emptyset\}$ to $\mathbb{N}$ is constructed. Claim 6.6 below establishes that, for each $J \in \mathcal{Fin} - \{\emptyset\}$, $\zeta(J)$ is set at most once. It is clear from the construction of ζ that it is 1-1, i.e., $(\forall J, J' \in \mathcal{Fin} - \{\emptyset\})[\zeta(J)\downarrow = \zeta(J') \Rightarrow J = J']$.

Claim 6.3. $(Z_\ell)_{\ell \in \mathbb{N}}$ is a numbering of $\mathcal{CE}$.

Proof of Claim. Let $S \in \mathcal{CE}$ be fixed. If $S = \emptyset$, then S is listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ at the beginning of the construction of $(Z_\ell)_{\ell \in \mathbb{N}}$. So, suppose that $S \neq \emptyset$. Let $\{s_0 < s_1 < \cdots\} = S$. If there exists an i such that $\{s_0, ..., s_i\}$ is *not* exemplary, then, clearly, for the *least* such i, S is listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ by cond. (b) when $A = \{s_0, ..., s_i\}$. So, suppose that

$$(\forall i)[\{s_0, ..., s_i\} \text{ is exemplary}]. \tag{34}$$

Clearly, $F(\{s_0\})$ is finite, and, by (34), $F(\{s_0\}) \neq \emptyset$. Thus, there exists a *greatest* m such that $F(\{s_0, ..., s_m\}) \subset F(\{s_0, ..., s_{m-1}\})$. If $X_{F(\{s_0,...,s_m\})} = S$, then S is listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ by cond. (c) when $A = \{s_0, ..., s_m\}$. So, suppose that $X_{F(\{s_0,...,s_m\})} \neq S$. If there exists an i such that $s_i \notin X_{F(\{s_0,...,s_m\})}$, then $\{s_0, ..., s_i\}$ is *not* exemplary, contradicting (34). So, there must exist a *least* $x \in X_{F(\{s_0,...,s_m\})} - S$. If there exists an i such that $s_i > x$, then $\{s_0, ..., s_i\}$ is *not* exemplary, again contradicting (34). So, it must be the case that, for each i, $s_i < x$. Let n be such that $\{s_0, ..., s_n\} = S$. Clearly, $\{s_0, ..., s_n, x\}$ is exemplary and $F(\{s_0, ..., s_n, x\}) = F(\{s_0, ..., s_n\})$. Thus, S is listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ by cond. (d) when $A = \{s_0, ..., s_n, x\}$. $\square$ (**Claim 6.3**)

Claim 6.4. For each $A_0, A_1 \in \mathcal{Fin} - \{\emptyset\}$, if $A_0 \neq A_1$, then the sets listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ for A_0 are distinct from those listed into $(Z_\ell)_{\ell \in \mathbb{N}}$ for A_1.

15

Proof of Claim. By way of contradiction, suppose otherwise, as witnessed by A_0 and A_1. Without loss of generality, suppose that the condition that applies for A_0 in Figure 3 is alphabetically no larger than that which applies for A_1, e.g., if cond. (c) applies for A_0, then either cond. (c) or cond. (d) applies for A_1. Consider the following cases.

CASE [cond. (a) applies for A_0]. Since cond. (a) does not list any sets into $(Z_\ell)_{\ell \in \mathbb{N}}$, this case immediately leads to a contradiction.

CASE [cond. (b) applies for both A_0 and A_1]. Then, there exist k_0 and k_1 such that

$$A_0 \cup (Y_{k_0} + \max A_0 + 1) = A_1 \cup (Y_{k_1} + \max A_1 + 1). \tag{35}$$

Without loss of generality, suppose that $\max A_0 \le \max A_1$. Then, by (35),

$$A_0 = A_1 \cap \{0, ..., \max A_0\}. \tag{36}$$

Since $A_0 \ne A_1$, it must be the case that $\max A_0 < \max A_1$. Thus,

$$\max A_0 \le \max A_1^- \ \wedge \ A_0 = A_1^- \cap \{0, ..., \max A_0\}. \tag{37}$$

Since cond. (b) applies for A_1, A_1^- is exemplary. Thus, Claim 6.1 and (37) imply that A_0 is exemplary. But then cond. (b) cannot apply for A_0 — a contradiction.

CASE [cond. (b) applies for A_0 and cond. (c) applies for A_1]. Then, there exists a k such that
$$A_0 \cup (Y_k + \max A_0 + 1) = X_{F(A_1)}. \tag{38}$$

Note that

$$\big(A_0 \cup (Y_k + \max A_0 + 1)\big) \cap \{0, ..., \max A_0\} = A_0 \text{ is } not \text{ exemplary.} \tag{39}$$

On the other hand, by Claim 6.2,

$$X_{F(A_1)} \cap \{0, ..., \max A_0\} \text{ is exemplary.} \tag{40}$$

Formulae (38)-(40) are contradictory.

CASE [cond. (b) applies for A_0 and cond. (d) applies for A_1]. Similar to the previous case.

CASE [cond. (c) applies for both A_0 and A_1]. Then,

$$X_{F(A_0)} = X_{F(A_1)}. \tag{41}$$

Since A_0 and A_1 are exemplary and $A_0 \ne A_1$, it must be the case that either $A_0 \subseteq A_1^-$ or $A_1 \subseteq A_0^-$. Without loss of generality, suppose that $A_0 \subseteq A_1^-$. Since $F(A_1) \subset F(A_1^-)$, there exists a $j \in F(A_1^-)$ such that $\max A_1 \in X_{F(A_1)} - X_j$. Since $A_0 \subseteq A_1^-$, $F(A_0) \supseteq F(A_1^-)$, and, thus, $X_{F(A_0)} \subseteq X_j$. But then $\max A_1 \in X_{F(A_1)} - X_{F(A_0)}$, contradicting (41).

CASE [cond. (c) applies for A_0 and cond. (d) applies for A_1]. Then,

$$X_{F(A_0)} = A_1^-. \tag{42}$$

Since $A_1^- \subseteq X_{F(A_0)}$,

$$F(A_1^-) \supseteq F(A_0). \tag{43}$$

Thus,

$$
\begin{aligned}
A_1 \quad &\subseteq \quad X_{F(A_1)} \quad &\{\text{immediate}\} \\
&= \quad X_{F(A_1^-)} \quad &\{\text{because } F(A_1) = F(A_1^-)\} \\
&\subseteq \quad X_{F(A_0)} \quad &\{\text{by (43)}\} \\
&= \quad A_1^- \quad &\{\text{by (42)}\}
\end{aligned}
$$

— a contradiction.

CASE [cond. (d) applies for both A_0 and A_1]. Then, $A_0^- = A_1^-$. Without loss of generality, suppose that $\max A_0 \leq \max A_1$. Then, since $A_0 \neq A_1$,

$$\max A_0 < \max A_1 \ \wedge \ \max A_0 \notin A_1. \tag{44}$$

On the other hand,

$$
\begin{aligned}
\max A_0 \quad &\in \quad X_{F(A_0)} \quad &\{\text{immediate}\} \\
&= \quad X_{F(A_0^-)} \quad &\{\text{because } F(A_0) = F(A_0^-)\} \\
&= \quad X_{F(A_1^-)} \quad &\{\text{because } A_0^- = A_1^-\} \\
&= \quad X_{F(A_1)} \quad &\{\text{because } F(A_1) = F(A_1^-)\}.
\end{aligned}
\tag{45}
$$

But (44) and (45) contradict the fact that A_1 is exemplary. $\quad\square$ (**Claim 6.4**)

Claim 6.5. $(Z_\ell)_{\ell \in \mathbb{N}}$ is 1-1.

Proof of Claim. Follows from Claim 6.4. $\quad\square$ (**Claim 6.5**)

Claim 6.6. For each $J \in \mathcal{F}in - \{\emptyset\}$, $\zeta(J)$ is set at most once in the construction of $(Z_\ell)_{\ell \in \mathbb{N}}$.

Proof of Claim. By way of contradiction, let $J \in \mathcal{F}in - \{\emptyset\}$ be such that $\zeta(J)$ is set more than once in the construction of $(Z_\ell)_{\ell \in \mathbb{N}}$. Then, clearly, there exist $A_0, A_1 \in \mathcal{F}in - \{\emptyset\}$ such that $A_0 \neq A_1$, $F(A_0) = J = F(A_1)$, and both $X_{F(A_0)}$ and $X_{F(A_1)}$ are listed into $(Z_\ell)_{\ell \in \mathbb{N}}$. Furthermore, since $F(A_0) = F(A_1)$, $X_{F(A_0)} = X_{F(A_1)}$. But this contradicts Claim 6.4. $\quad\square$ (**Claim 6.6**)

To complete the proof of the theorem, it suffices to show that $\mathcal{L}$ can be **It**-identified using $(Z_\ell)_{\ell \in \mathbb{N}}$ as the hypothesis space. For ease of presentation, suppose that $Z_0 = \emptyset$. Recall that ζ is a 1-1, partial computable function from $\mathcal{F}in - \{\emptyset\}$ to $\mathbb{N}$. Thus, ζ^{-1} is a partial computable function from $\mathbb{N}$ to $\mathcal{F}in - \{\emptyset\}$. For each x, let $G_x : \mathcal{F}in \to \mathcal{F}in$ be such that, for each $J \in \mathcal{F}in$,

$$G_x(J) = \{j \in J \mid x \in X_j\}. \tag{46}$$

17

(a) For each i, execute stage 0 below.

- STAGE 0. Include $\mathbb{N} + i$ and $\{i\}$ in $\mathcal{L}$. Go to stage 1.

- STAGE 1. Let (M, p) be the ith pair in $((M, p)_{i'})_{i' \in \mathbb{N}}$. Search for a $k \geq i$ such that

$$M_p^*\big((i \cdot \ldots \cdot k) \cdot (k+1)\big){\downarrow} \;=\; M_p^*\big((i \cdot \ldots \cdot k) \cdot k\big) \;=\; M_p^*(i \cdot \ldots \cdot k).$$

  If such a k is found, then include $\{i, ..., k\}$ and $\{i, ..., k+1\}$ in $\mathcal{L}$, and terminate the construction (for i). If *no* such k is found, then search indefinitely.

(b) For each i, execute stage 0 below.

- STAGE 0. Set $X_{\text{start}(i)} = \mathbb{N} + i$, and, for each $j \in \{\text{start}(i) + 1, ..., \text{start}(i+1) - 1\}$, set $X_j = \{i\}$. Go to stage 1.

- STAGE 1. In a dovetailing manner, monitor and act according to the following conditions.

  - COND. [in the construction of $\mathcal{L}$ above, a k is found for i]. Set $X_{\text{start}(i)+2} = \{i, ..., k\}$ and $X_{\text{start}(i)+3} = \{i, ..., k+1\}$.
  - COND. [$i \in X_j$, where $j < \text{start}(i)$]. Set $X_{\text{start}(i)+j+4} = X_j$.

Figure 4: **(a)** The construction of $\mathcal{L}$ in the proof of Theorem 9. **(b)** The construction of $(X_j)_{j \in \mathbb{N}}$ in the proof of Theorem 9. The function start is defined in (51).

Let $M \in \mathcal{PC}_{1,1}$ be such that, for each $\ell > 0$ and x:

$$
\begin{align}
M(0, \#) &= 0; & (47)\\
M(0, x) &= (\zeta \circ F)(\{x\}); & (48)\\
M(\ell, \#) &= \ell; & (49)\\
M(\ell, x) &= (\zeta \circ G_x \circ \zeta^{-1})(\ell). & (50)
\end{align}
$$

It is straightforward to show that $(M, 0, (Z_\ell)_{\ell \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$.

$\square$ (**Theorem 6**)

Recall from Section 1 that $\mathcal{P}at^\Sigma$ is the class of pattern languages over Σ, where Σ is an arbitrary alphabet. It is straightforward to show that, for each Σ, there exists an effective numbering of $\mathcal{P}at^\Sigma$ that is both uniformly decidable and computably finitely thick. Thus, one has the following corollary of Theorem 6.

Corollary 8 (of Theorem 6). For each alphabet Σ, $\mathcal{P}at^\Sigma$ is **FrIt**-identifiable.

The proof of Theorem 9 below exhibits a class of languages $\mathcal{L}$ that is uniformly decidable and computably finitely thick, but $\mathcal{L} \notin \textbf{It}$. Thus, one could not

arrive at the conclusion of Theorem 6 if one were to merely require that: there exists a uniformly decidable effective numbering of $\mathcal{L}$, and a possibly *distinct* computably finitely thick effective numbering of $\mathcal{L}$.

Theorem 9. There exists a class of languages $\mathcal{L}$ that is uniformly decidable and computably finitely thick, but $\mathcal{L} \notin \mathbf{It}$.

Proof. Let $((M, p)_{i'})_{i' \in \mathbb{N}}$ be an algorithmic enumeration of all pairs of type $\mathcal{PC}_{1,1} \times \mathbb{N}$. Let $\mathrm{start} : \mathbb{N} \to \mathbb{N}$ be such that, for each i,

$$\mathrm{start}(i) = 2^{i+2} - 4. \tag{51}$$

Note that, for each i,

$$\mathrm{start}(i + 1) - \mathrm{start}(i) = \mathrm{start}(i) + 4. \tag{52}$$

The class $\mathcal{L}$ is constructed in Figure 4(a). An effective numbering $(X_j)_{j \in \mathbb{N}}$, which is easily seen to be of $\mathcal{L}$, is constructed in Figure 4(b). This effective numbering $(X_j)_{j \in \mathbb{N}}$ is used to show that $\mathcal{L}$ is computably finitely thick. It is straightforward to construct an effective numbering witnessing that $\mathcal{L}$ is uniformly decidable.

The following are easily verifiable from the construction of $(X_j)_{j \in \mathbb{N}}$.

- For each $L \in \mathcal{L}$ and $i \in L$, there exists a $j < \mathrm{start}(i) + 4$ such that $X_j = L$.

- For each i and $j < \mathrm{start}(i)$, if $i \in X_j$, then there exists a $j' \in \{\mathrm{start}(i) + 4, ..., \mathrm{start}(i + 1) - 1\}$ such that $X_{j'} = X_j$.

- For each $j \in \{\mathrm{start}(i), ..., \mathrm{start}(i + 1) - 1\}$, $i \in X_j$.

Given these facts, if one lets $f : \mathbb{N} \to \mathbb{N}$ be such that, for each i,

$$D_{f(i)} = \{\mathrm{start}(i), ..., \mathrm{start}(i + 1) - 1\}, \tag{53}$$

then f clearly witnesses that $(X_j)_{j \in \mathbb{N}}$ is computably finitely thick.

It remains to show that $\mathcal{L} \notin \mathbf{It}$. By way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Let i be such that (M, p) is the ith pair in $((M, p)_{i'})_{i' \in \mathbb{N}}$. Since $(\mathbb{N} + i) \in \mathcal{L}$, there exists a $k_0 \geq i$ such that

$$(\forall k > k_0)\big[M_p^*\big((i \cdot \cdots \cdot k_0) \cdot k\big)\!\downarrow\, = M_p^*(i \cdot \cdots \cdot k_0)\big]. \tag{54}$$

Note that if one lets $k_1 = k_0 + 1$, then k_1 satisfies

$$M_p^*\big((i \cdot \cdots \cdot k_1) \cdot (k_1 + 1)\big)\!\downarrow\ =\ M_p^*\big((i \cdot \cdots \cdot k_1) \cdot k_1\big)\ =\ M_p^*(i \cdot \cdots \cdot k_1). \tag{55}$$

Thus, *some* k_1 satisfying (55) is discovered in stage 1 of the construction of $\mathcal{L}$ for i (though, k_1 is not necessarily $k_0 + 1$). Furthermore, $\{i, ..., k_1\}$ and $\{i, ..., k_1 + 1\}$ are in $\mathcal{L}$. Note that

$$t_0 \ =\ (i \cdot \cdots \cdot k_1) \cdot k_1^{\infty}; \tag{56}$$

$$t_1 \ =\ (i \cdot \cdots \cdot k_1) \cdot (k_1 + 1)^{\infty} \tag{57}$$

are, respectively, texts for $\{i, ..., k_1\}$ and $\{i, ..., k_1 + 1\}$. But, by (55), M_p^* converges to the same hypothesis on each of these texts — a contradiction.

$$\square \ (\textbf{Theorem 9})$$

4. Friedberg, Injective, Class Injective, Order Independent, and Extensional Iterative Learning

This section examines the Friedberg, injective, class injective, order independent, and extensional iterative learning models (**FrIt**, **InjIt**, **ClassInjIt**, **OrdIndIt**, and **ExtIt**, respectively). Recall from Definition 1 that, for each $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$ and $\mathcal{L}$:

- $(M, p, (X_j)_{j \in \mathbb{N}})$ **FrIt**-identifies $\mathcal{L}$ $\Leftrightarrow$ $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and $(X_j)_{j \in \mathbb{N}}$ is a Friedberg numbering;

- $(M, p, (X_j)_{j \in \mathbb{N}})$ **InjIt**-identifies $\mathcal{L}$ $\Leftrightarrow$ $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and $(X_j)_{j \in \mathbb{N}}$ is 1-1;

- $(M, p, (X_j)_{j \in \mathbb{N}})$ **ClassInjIt**-identifies $\mathcal{L}$ $\Leftrightarrow$ $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, $(X_j)_{j \in \mathbb{N}}$ is 1-1, and $\{X_j \mid j \in \mathbb{N}\} = \mathcal{L}$;

- $(M, p, (X_j)_{j \in \mathbb{N}})$ **OrdIndIt**-identifies $\mathcal{L}$ $\Leftrightarrow$ $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and, for each $L \in \mathcal{L}$, there exists a j such that, for each $t \in \mathrm{Txt}(L)$, M_p^* converges on t to j;

- $(M, p, (X_j)_{j \in \mathbb{N}})$ **ExtIt**-identifies $\mathcal{L}$ $\Leftrightarrow$ $(M, p, (X_j)_{j \in \mathbb{N}})$ **It**-identifies $\mathcal{L}$, and, for each $\sigma_0, \sigma_1 \in \mathrm{Seq}(\mathcal{L}) - \{\lambda\}$,

$$[X_{M_p^*(\sigma_0^-)} = X_{M_p^*(\sigma_1^-)} \ \wedge \ \mathrm{last}(\sigma_0) = \mathrm{last}(\sigma_1)] \ \Rightarrow \ X_{M_p^*(\sigma_0)} = X_{M_p^*(\sigma_1)}. \tag{58}$$

In terms of the classes of languages identifiable by these models and by **It**, they are clearly related as follows.

$$\mathbf{OrdIndIt}$$
$$\cup|$$
$$\mathbf{ClassInjIt} \subseteq \mathbf{InjIt} \subseteq \mathbf{ExtIt} \subseteq \mathbf{It}$$
$$\cup|$$
$$\mathbf{FrIt}$$

In this section, we establish that **ClassInjIt** $\not\subseteq$ **FrIt** (Proposition 10), that **FrIt** $\not\subseteq$ **ClassInjIt** (Proposition 11 and Theorem 12), that **InjIt** $\not\subseteq$ (**FrIt** $\cup$ **ClassInjIt**) (Proposition 14), that **InjIt** $=$ **OrdIndIt** (Theorem 15), and that **ExtIt** $\not\subseteq$ **InjIt** (Theorem 16).[12] That **It** $\not\subseteq$ **ExtIt** is shown in Section 5 (Theorems 22 and 23).

Proposition 10 just below establishes that **ClassInjIt** $\not\subseteq$ **FrIt**.

Proposition 10. ClassInjIt $\not\subseteq$ **FrIt**.

---

[12]According to an anonymous referee of the conference version of this paper [MZ10], it was already known to Liepe & Wiehagen that **It** $\not\subseteq$ **OrdIndIt**.

Proof. Recall that $\mathcal{F}in$ is the collection of all finite subsets of $\mathbb{N}$. Jain & Stephan observed that $\mathcal{F}in \notin \mathbf{FrIt}$ [JS08, Remark 25]. However, it is easily seen that $\mathcal{F}in \in \mathbf{ClassInjIt}$. $\square$ (**Proposition 10**)

Proposition 11 just below establishes that $\mathbf{FrIt} \not\subseteq \mathbf{ClassInjIt}$.

Proposition 11. $\mathbf{FrIt} \not\subseteq \mathbf{ClassInjIt}$.

Proof. Let $\overline{\mathcal{K}}$ be such that

$$\overline{\mathcal{K}} = \big\{ \{x\} \mid x \in \overline{K} \big\}. \tag{59}$$

It is straightforward to show that $\overline{\mathcal{K}} \in \mathbf{FrIt}$. On the other hand, since no hypothesis space represents $\overline{\mathcal{K}}$ exactly, $\overline{\mathcal{K}} \notin \mathbf{ClassInjIt}$. $\square$ (**Proposition 11**)

Our proof of Proposition 11 might lead one to wonder: does every class of languages $\mathcal{L} \in \mathbf{FrIt} - \mathbf{ClassInjIt}$ have the property that *no* hypothesis space represents $\mathcal{L}$ exactly? Theorem 12 just below establishes that this is, in fact, *not* the case. In particular, the class $\mathcal{C}oinit$ from Example 4(d) satisfies $\mathcal{C}oinit \in \mathbf{FrIt} - \mathbf{ClassInjIt}$. Furthermore, there exists a hypothesis space $(Y_k)_{k \in \mathbb{N}}$ such that $(Y_k)_{k \in \mathbb{N}}$ represents $\mathcal{C}oinit$ exactly.

Theorem 12. Recall from Example 4(d) that

$$\mathcal{C}oinit = \{\mathbb{N} + e \mid e \in \mathbb{N}\}. \tag{60}$$

$\mathcal{C}oinit$ satisfies (a) and (b) below.

(a) There exists an effective numbering $(Y_k)_{k \in \mathbb{N}}$ such that $\{Y_k \mid k \in \mathbb{N}\} = \mathcal{C}oinit$.

(b) $\mathcal{C}oinit \in \mathbf{FrIt} - \mathbf{ClassInjIt}$.

Proof. The proof of part (a) is straightforward. Furthermore, as per Example 4(d), one can arrange that the effective numbering $(Y_k)_{k \in \mathbb{N}}$ is both uniformly decidable and computably finitely thick. Thus, by Theorem 6 above, $\mathcal{C}oinit \in \mathbf{FrIt}$.

It remains to show that $\mathcal{C}oinit \notin \mathbf{ClassInjIt}$. By way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Then, there exists an e such that

$$X_p = \mathbb{N} + e. \tag{61}$$

Furthermore, since p is the *unique* index satisfying (61), it follows that

$$(\forall x \geq e)[M(p, x) = p]. \tag{62}$$

Let t be such that

$$t = (e+1) \cdot (e+2) \cdot \cdots. \tag{63}$$

Note that t is a text for $(\mathbb{N} + e + 1) \in \mathcal{C}oinit$. However, by (61) and (62), $(M, p, (X_j)_{j \in \mathbb{N}})$ does *not* identify $\mathbb{N} + e + 1$ from t — a contradiction. $\square$ (**Theorem 12**)

Note that the **ClassInjIt**-learning model has the following, somewhat unusual property. It is possible for a class $\mathcal{L}$ to be *not* **ClassInjIt**-identifiable, whilst a class $\mathcal{L}' \supseteq \mathcal{L}$ *is* **ClassInjIt**-identifiable. For example, as per the proofs of Propositions 10 and 11 above, $\overline{\mathcal{K}}$ is *not* **ClassInjIt**-identifiable, whilst $\mathcal{Fin} \supseteq \overline{\mathcal{K}}$ *is* **ClassInjIt**-identifiable. Given these facts, one might wonder: does every **InjIt**-identifiable class have a **ClassInjIt**-identifiable superclass? At present, this question is open.

Problem 13. (a) and (b) below.

(a) Is it the case that, for each $\mathcal{L} \in$ **InjIt**, there exists an $\mathcal{L}' \in$ **ClassInjIt** such that $\mathcal{L} \subseteq \mathcal{L}'$?
(b) If the answer to (a) is "no", then is it the case that, for each $\mathcal{L} \in$ **FrIt**, there exists an $\mathcal{L}' \in$ **ClassInjIt** such that $\mathcal{L} \subseteq \mathcal{L}'$?

Proposition 14 just below establishes that **InjIt** $\not\subseteq$ (**FrIt** $\cup$ **ClassInjIt**).

Proposition 14. **InjIt** $\not\subseteq$ (**FrIt** $\cup$ **ClassInjIt**).

Proof. A witnessing class is $\mathcal{Fin} \vee \overline{\mathcal{K}}$. It is straightforward to show that $\mathcal{Fin} \vee \overline{\mathcal{K}} \in$ **InjIt**. The proof that $\mathcal{Fin} \vee \overline{\mathcal{K}} \notin$ **FrIt** is similar to the proof that $\mathcal{Fin} \notin$ **FrIt**. Finally, the existence of a hypothesis space representing $\mathcal{Fin} \vee \overline{\mathcal{K}}$ exactly would imply the existence of an analogous hypothesis space for $\overline{\mathcal{K}}$. Thus, $\mathcal{Fin} \vee \overline{\mathcal{K}} \notin$ **ClassInjIt**. $\qquad \square$ (**Proposition 14**)

Theorem 15 just below establishes that **InjIt** = **OrdIndIt**.

Theorem 15. **InjIt** = **OrdIndIt**.

Proof. Clearly, **InjIt** $\subseteq$ **OrdIndIt**. Thus, it suffices to show that **OrdIndIt** $\subseteq$ **InjIt**. Let $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$ be fixed, and let $\mathcal{L}$ be the *largest* class of languages that $(M, p, (X_j)_{j \in \mathbb{N}})$ **OrdIndIt**-identifies. (It is easily seen from Definition 1(e) that $\mathcal{L}$ is well-defined.) We show that $\mathcal{L}$ is **InjIt**-identifiable. If $\mathcal{L} \subseteq \mathcal{Fin}$, then $\mathcal{L}$ is clearly **InjIt**-identifiable. So, suppose that $\mathcal{L} \not\subseteq \mathcal{Fin}$.

The remainder of the proof is in several parts. First, we construct a hypothesis space $(Y_j)_{j \in \mathbb{N}}$. This hypothesis space $(Y_j)_{j \in \mathbb{N}}$ is *almost* 1-1 in the sense that, for each j and j', if $Y_j = Y_{j'} \neq \emptyset$, then $j = j'$ (see Claim 15.3 below). From $(Y_j)_{j \in \mathbb{N}}$, we construct a second hypothesis space $(Z_\ell)_{\ell \in \mathbb{N}}$. This hypothesis space $(Z_\ell)_{\ell \in \mathbb{N}}$ is truly 1-1. Finally, we construct an $\tilde{M} \in \mathcal{PC}_{1,1}$, and show that $(\tilde{M}, 0, (Z_\ell)_{\ell \in \mathbb{N}})$ **InjIt**-identifies $\mathcal{L}$.

For each j and n, let $X_j^n \in \mathcal{Fin}$ be such that

$$X_j^n = \{x \mid x \text{ is listed in } X_j \text{ in at most } n \text{ steps}\}. \tag{64}$$

For each L, let $t_L : \mathbb{N} \to \mathbb{N}_\#$ be such that, for each i,

$$t_L(i) = \begin{cases} i, & \text{if } i \in L; \\ \#, & \text{otherwise.} \end{cases} \tag{65}$$

Note that the predicate $\lambda A \in \mathcal{Fin}, j \in \mathbb{N}.[M_p^* \text{ converges on } t_A \text{ to } j]$ is partial computable. For each j and n, let $P(j, n) \Leftrightarrow$ by letting $A = X_j^n$, (a)-(c) below are satisfied.

(a) M_p^* converges on t_A to j.

(b) $(\forall x \in A)[M(j, x) = j]$.

(c) For each $\sigma \in \text{Seq}(A)$ such that $|\sigma| < n$, there exists a j' such that (i) and (ii) below.

   (i) $M_p^*(\sigma) = j'$.

   (ii) $j \neq j' \Rightarrow (\exists x \in X_j)[M(j', x)\downarrow \neq j']$.

Note that P is partial computable. Let $(Y_j)_{j \in \mathbb{N}}$ be such that, for each j,

$$Y_j = \bigcup \{X_j^n \mid P(j, n)\}. \tag{66}$$

Claim 15.1. For each j, if Y_j is infinite, then $Y_j = X_j$.

Proof of Claim. Easily verifiable from (66). $\qquad\qquad \square$ (**Claim 15.1**)

Claim 15.2. Suppose that j satisfies (a)-(c) below.

(a) $X_j \neq \emptyset$.

(b) $X_j \in \mathcal{L}$.

(c) M_p^* converges on t_{X_j} to j.

Then, $Y_j = X_j$.

Proof of Claim. Suppose that j satisfies (a)-(c) in the statement of the claim. First, consider the case that X_j is finite. Then, using the fact that $(M, p, (X_j)_{j \in \mathbb{N}})$ **OrdIndIt**-identifies $\mathcal{L}$, it can be shown that there exists an n such that $P(j, n)$ and $Y_j = X_j^n = X_j$. Next, consider the case that X_j is infinite. Then, using the fact that $(M, p, (X_j)_{j \in \mathbb{N}})$ **OrdIndIt**-identifies $\mathcal{L}$, it can similarly be shown that there exist infinitely many n such that $P(j, n)$. It follows that Y_j is infinite. Thus, by Claim 15.1, $Y_j = X_j$. $\quad\square$ (**Claim 15.2**)

Claim 15.3. For each j and j', if $Y_j = Y_{j'} \neq \emptyset$, then $j = j'$.

Proof of Claim. Suppose that j and j' are such that $Y_j = Y_{j'} \neq \emptyset$. Consider the following cases.

CASE $[Y_j$ and $Y_{j'}$ are finite$]$. Let $A = Y_j = Y_{j'}$. Then, by the construction of Y_j and $Y_{j'}$, M_p^* converges on t_A to j and j'. Thus, it must be the case that $j = j'$.

CASE $[Y_j$ and $Y_{j'}$ are infinite$]$. By way of contradiction, suppose that $j \neq j'$. Choose n_0 such that $P(j', n_0)$, and let $A_0 = X_{j'}^{n_0}$. By part (a) of $P(j', n_0)$, M_p^* converges on t_{A_0} to j', i.e., there exists an i_0 such that

$$(\forall i \geq i_0)\left[M_p^*(t_{A_0}[i]) = j'\right]. \tag{67}$$

Choose n_1 such that $A_0 \subseteq X_j^{n_1}$, $i_0 < n_1$, and $P(j, n_1)$. By part (c) of $P(j, n_1)$, there exists a j'' such that (i) and (ii) below.

(i) $M_p^*(t_{A_0}[i_0]) = j''$.

(ii) $j \neq j'' \Rightarrow (\exists x \in X_j)[M(j'', x)\downarrow \neq j'']$.

23

By (67) and (i) just above, $j' = j''$. Thus, by (ii) just above and the assumption that $j \neq j'$ $(= j'')$, there exists an x_0 such that

$$
\begin{aligned}
x_0 \quad \in \quad & X_j \\
= \quad & Y_j \quad \text{\{by Claim 15.1 and the case\}} \\
= \quad & Y_{j'} \quad \text{\{by assumption\}} \\
= \quad & X_{j'} \quad \text{\{by Claim 15.1 and the case\}}
\end{aligned}
$$

and $M(j', x_0){\downarrow} \neq j'$. Choose n_2 such that $x_0 \in X_{j'}^{n_2}$ and $P(j', n_2)$. By part (b) of $P(j', n_2)$, $M(j', x_0) = j'$ — a contradiction. $\qquad \square$ (**Claim 15.3**)

For each j and s, let $Y_j^s \in \mathcal{F}in$ be such that

$$
Y_j^s = \{x \mid x \text{ is listed in } Y_j \text{ in at most } s \text{ steps}\}. \tag{68}
$$

Recall that $L \not\subseteq \mathcal{F}in$. Thus, there exists some infinite language $L_\infty \in L$. Let σ be any locking sequence for $(M, p, (X_j)_{j\in\mathbb{N}})$ on L_∞. Let $f : \mathbb{N}^3 \to \mathbb{N}$ be any 1-1, computable function such that $\mathrm{rng}(f) = L_\infty$. Let $(Z_\ell)_{\ell\in\mathbb{N}}$ be such that, for each j and s:

$$
Z_0 \quad = \quad \emptyset; \tag{69}
$$

$$
Z_{\langle j,s\rangle+1} \quad = \quad \begin{cases} Y_j, & \text{if } s \text{ is } least \text{ such that } Y_j^s \neq \emptyset; \\ \text{content}(\sigma) \cup \{f(j, s, m) \mid m \in \mathbb{N}\}, & \text{otherwise.} \end{cases} \tag{70}
$$

Claim 15.4. Suppose that j satisfies (a)-(c) in the statement of Claim 15.2, and that s is *least* such that $Y_j^s \neq \emptyset$. Then, $Z_{\langle j,s\rangle+1} = X_j$.

Proof of Claim. Immediate by Claim 15.2 and (70). $\qquad \square$ (**Claim 15.4**)

Claim 15.5. For each j, s, j', and s', if $Z_{\langle j,s\rangle+1} = Z_{\langle j',s'\rangle+1}$, then $j = j'$ and $s = s'$.

Proof of Claim. Suppose that j, s, j', and s' are such that $Z_{\langle j,s\rangle+1} = Z_{\langle j',s'\rangle+1}$. Consider the following cases.

CASE [s *is* least such that $Y_j^s \neq \emptyset \wedge s'$ *is* least such that $Y_{j'}^{s'} \neq \emptyset$]. Then,

$$
Y_j = Z_{\langle j,s\rangle+1} = Z_{\langle j',s'\rangle+1} = Y_{j'}. \tag{71}
$$

Thus, by Claim 15.3, $j = j'$. Furthermore, since s and s' are *least* such that $Y_j^s \neq \emptyset$ and $Y_{j'}^{s'} \neq \emptyset$, $s = s'$.

CASE [s *is* least such that $Y_j^s \neq \emptyset \wedge s'$ *is not* least such that $Y_{j'}^{s'} \neq \emptyset$]. Then,

$$
Y_j = Z_{\langle j,s\rangle+1} = Z_{\langle j',s'\rangle+1} = \text{content}(\sigma) \cup \{f(j', s', m) \mid m \in \mathbb{N}\}. \tag{72}
$$

We show that this case leads to a contradiction. Choose n_0 such that $\text{content}(\sigma) \subseteq X_j^{n_0}$, $|\sigma| < n_0$, and $P(j, n_0)$. By part (c) of $P(j, n_0)$, there exists a j'' such that (i) and (ii) below.

(i) $M_p^*(\sigma) = j''$.

(ii) $j \neq j'' \Rightarrow (\exists x \in X_j)[M(j'',x)\!\downarrow \neq j'']$.

Since σ is a locking sequence for $(M, p, (Z_\ell)_{\ell \in \mathbb{N}})$ on L_∞, $X_{j''} = L_\infty$. On the other hand, by (72), $Y_j \subset L_\infty$. Thus, $j \neq j''$. Furthermore, by (ii) just above, there exists an x such that

$$
\begin{aligned}
x \;\in\; & X_j \\
=\; & Y_j \quad \{\text{by Claim 15.1 and (72)}\} \\
\subseteq\; & L_\infty \quad \{\text{by (72)}\}
\end{aligned}
$$

and $M(j'',x) \neq j''$. But this contradicts the choice of σ.

CASE [s is *not* least such that $Y_j^s \neq \emptyset$ $\wedge$ s' *is* least such that $Y_{j'}^{s'} \neq \emptyset$]. Symmetric to the previous case.

CASE [s is *not* least such that $Y_j^s \neq \emptyset$ $\wedge$ s' is *not* least such that $Y_{j'}^{s'} \neq \emptyset$]. Then,

$$
\text{content}(\sigma) \cup \{f(j,s,m) \mid m \in \mathbb{N}\} = Z_{\langle j,s\rangle+1} = Z_{\langle j',s'\rangle+1} = \text{content}(\sigma) \cup \{f(j',s',m) \mid m \in \mathbb{N}\}.
\tag{73}
$$

Note that content(σ) contributes only finitely many elements to each of $Z_{\langle j,s\rangle+1}$ and $Z_{\langle j',s'\rangle+1}$. Thus, for (73) to hold, it must be the case that $j = j'$ and $s = s'$. $\square$ (**Claim 15.5**)

Claim 15.6. For each ℓ and ℓ', if $Z_\ell = Z_{\ell'}$, then $\ell = \ell'$.

Proof of Claim. Suppose that ℓ and ℓ' are such that $Z_\ell = Z_{\ell'}$. Clearly, if $Z_\ell = Z_{\ell'} = \emptyset$, then $\ell = 0 = \ell'$. On the other hand, if $Z_\ell = Z_{\ell'} \neq \emptyset$, then it must be the case that $\ell \neq 0 \neq \ell'$. Thus, by Claim 15.5, $\ell = \ell'$. $\square$ (**Claim 15.6**)

Let $\tilde{M} \in \mathcal{PC}_{1,1}$ be such that, for each $j, s \in \mathbb{N}$ and $x \in \mathbb{N}_\#$:

$$
\tilde{M}(0, \quad x) \;=\; \begin{cases}
0, & \text{if } x = \#; \\
\langle M(p,x),0 \quad \rangle + 1, & \text{if } [x \neq \# \;\wedge\; M(p,x)\!\downarrow]; \\
\uparrow, & \text{otherwise};
\end{cases}
\tag{74}
$$

$$
\tilde{M}(\langle j,s\rangle + 1, x) \;=\; \begin{cases}
\langle j, \quad s \quad \rangle + 1, & \text{if } [M(j,x) \;= j \;\wedge\; Y_j^s \neq \emptyset]; \\
\langle j, \quad s+1 \rangle + 1, & \text{if } [M(j,x) \;= j \;\wedge\; Y_j^s = \emptyset]; \\
\langle M(j,x),0 \quad \rangle + 1, & \text{if } [M(j,x)\!\downarrow \neq j]; \\
\uparrow, & \text{otherwise}.
\end{cases}
\tag{75}
$$

Clearly, $(\tilde{M}, 0, (Z_\ell)_{\ell \in \mathbb{N}})$ identifies $\emptyset$. On the other hand, suppose that $(\tilde{M}, 0, (Z_\ell)_{\ell \in \mathbb{N}})$ is fed a text for a language in $\mathcal{L} - \{\emptyset\}$. Then, intuitively, $\tilde{M}$ simulates M as follows. Suppose that M outputs some hypothesis j. Then, $\tilde{M}$ iterates through hypotheses $\langle j, s\rangle + 1$, with $s = 0, 1, ...$, until either: M switches to some hypothesis other than j, or a *least* s is found such that $Y_j^s \neq \emptyset$. If M never switches to another hypothesis, then, by Claim 15.2, such an s must eventually be found. Furthermore, by Claim 15.4, for this s, $Z_{\langle j,s\rangle+1} = X_j$. Given these facts and Claim 15.6, $(\tilde{M}, 0, (Z_\ell)_{\ell \in \mathbb{N}})$ clearly **InjIt**-identifies $\mathcal{L}$. $\square$ (**Theorem 15**)

Theorem 16 just below establishes that **ExtIt** $\not\subseteq$ **InjIt**.

Theorem 16. **ExtIt** $\not\subseteq$ **InjIt**.

Proof. Let $\mathcal{L}_0$ be as follows.

$$
\begin{aligned}
\mathcal{L}_0 \;=\; & \{2\mathbb{N}\} \\
& \cup\; \Big\{ \{0,2,...,2e\} \cup \{2e+1\} \cup 2W_e, \\
& \qquad \{0,2,...,2e\} \cup \{2e+1\} \cup 2X \quad \mid e \in \mathbb{N} \;\wedge\; \text{(a)-(c) below}\Big\}. \\
& \qquad \text{(a)}\ (\forall e' \in X)[W_{e'} = X]. \\
& \qquad \text{(b)}\ \text{If } W_e = X,\ \text{then } W_e \text{ and } X \text{ are finite.} \\
& \qquad \text{(c)}\ \text{If } W_e \neq X,\ \text{then } 2W_e \subseteq \{0,2,...,2e\} \text{ and } X \text{ is infinite.}
\end{aligned}
\tag{76}
$$

Note that $\mathcal{L}_0$ is a kind of self-describing class. This, in turn, makes it possible to **ExtIt**-identify $\mathcal{L}_0$. The proof that $\mathcal{L}_0$ is not **InjIt**-identifiable can be sketched as follows. By way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j\in\mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Using Case's 1-1 Operator Recursion Theorem [Cas74, Cas94],[13] we construct an infinite computably enumerable sequence $(e_i)_{i\in\mathbb{N}}$ such that one of the following three cases holds.

- The first case is: M_p^* never makes a mind-change on

$$
(0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^m,
\tag{77}
$$

for any m. In this case, it will turn out that $W_{e_0} = \emptyset$. It will then follow that $(M, p, (X_j)_{j\in\mathbb{N}})$ does *not* identify

$$
\{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\},
\tag{78}
$$

which will be in $\mathcal{L}_0$, leading to a contradiction. (See Claim 16.2 below.)

- The second case is: M_p^* makes the just described mind-change, for some m; but, after seeing $(2e_0 + 1)^m$, M_p^* never makes a mind-change on

$$
(0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^m \cdot (2e_1 \cdot 2e_2 \cdot \cdots \cdot 2e_n),
\tag{79}
$$

for any n. In this case, it will turn out that $W_{e_0} = \emptyset$ and $(\forall i \geq 1)[W_{e_i} = \{e_i \mid i \geq 1\}]$. It will then follow that $(M, p, (X_j)_{j\in\mathbb{N}})$ does *not* identify *at least one* of the languages in (80) and (81) just below.

$$
\{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\}.
\tag{80}
$$

$$
\{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\} \cup \{2e_i \mid i \geq 1\}.
\tag{81}
$$

Furthermore, both languages will be in $\mathcal{L}_0$, leading to a contradiction. (See the case "stage 1 is *not* exited" following the proof of Claim 16.2 below.)

---

[13]Intuitively, the 1-1 Operator Recursion Theorem allows one to construct a computably enumerable sequence of pairwise-distinct φ-programs $(e_i)_{i\in\mathbb{N}}$ such that each program e_i *knows* all programs in the sequence and its own index i.

- The third case is: M_p^* makes both of the just described mind-changes. In this case, it will turn out that, for some such n, $(\forall i \leq n)[W_{e_i} = \{e_1, ..., e_n\}]$. It will then follow that either $(M, p, (X_j)_{j \in \mathbb{N}})$ does *not* identify

$$\{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\} \cup \{2e_1, ..., 2e_n\}, \tag{82}$$

which will be in $\mathcal{L}_0$, or $(X_j)_{j \in \mathbb{N}}$ is *not* 1-1. Either way leads to a contradiction. (See the case "stage 1 *is* exited" following the proof of Claim 16.2 below.)

We now show formally that $\mathcal{L}_0 \in \textbf{ExtIt}$. Let $f : \mathbb{N}^2 \to \mathbb{N}$ be a computable, 1-1 function such that, for each e and e':

$$
\begin{align}
W_{f(0,\ 0\ )} &= 2\mathbb{N}; \tag{83}\\
W_{f(e+1,\ 0\ )} &= \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_e; \tag{84}\\
W_{f(e+1,\ e'+1)} &= \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_{e'}. \tag{85}
\end{align}
$$

(It does not matter what $W_{f(\cdot,\cdot)}$ is for the remaining pairs.) For each $e, e' \in \mathbb{N}$ and $x \in \mathbb{N}_\#$, let $\tilde{M}$ be as follows. We use "unchanged" as a synonym for $\tilde{M}$'s first argument.

$$
\begin{align}
\tilde{M}\big(f(0,\ \ 0\ \ ), x\big) &= \begin{cases} f(e+1, 0), & \text{where } x = 2e+1; \\ \text{unchanged}, & \text{otherwise}; \end{cases} \tag{86}\\[2mm]
\tilde{M}\big(f(e+1, 0\ \ ), x\big) &= \begin{cases} f(e+1, e'+1), & \text{where } x = 2e' \\ & \quad \text{and } 2e' > 2e+1; \\ \text{unchanged}, & \text{otherwise}; \end{cases} \tag{87}\\[2mm]
\tilde{M}\big(f(e+1, e'+1), x\big) &= \text{unchanged}. \tag{88}
\end{align}
$$

It is straightforward to show that $(\tilde{M}, f(0,0), (W_i)_{i \in \mathbb{N}})$ **It**-identifies $\mathcal{L}_0$. That $(\tilde{M}, f(0,0), (W_i)_{i \in \mathbb{N}})$ **ExtIt**-identifies $\mathcal{L}_0$ follows from Claim 16.1 just below.

Claim 16.1. $(\tilde{M}, f(0,0), (W_i)_{i \in \mathbb{N}})$ is extensional with respect to $\mathcal{L}_0$, in the sense of Definition 1(f).

Proof of Claim. Suppose that $\sigma_0, \sigma_1 \in \text{Seq}(\mathcal{L}) - \{\lambda\}$ are such that

$$W_{\tilde{M}_p^*(\sigma_0^-)} = W_{\tilde{M}_p^*(\sigma_1^-)} \wedge \text{last}(\sigma_0) = \text{last}(\sigma_1). \tag{89}$$

It must be shown that

$$W_{\tilde{M}_p^*(\sigma_0)} = W_{\tilde{M}_p^*(\sigma_1)}. \tag{90}$$

If $\tilde{M}_p^*(\sigma_0^-) = \tilde{M}_p^*(\sigma_1^-)$, then (90) follows immediately. Similarly, if $W_{\tilde{M}_p^*(\sigma_0)} = W_{\tilde{M}_p^*(\sigma_0^-)}$ and $W_{\tilde{M}_p^*(\sigma_1)} = W_{\tilde{M}_p^*(\sigma_1^-)}$, then (90) follows immediately. So, suppose that

$$\tilde{M}_p^*(\sigma_0^-) \neq \tilde{M}_p^*(\sigma_1^-) \wedge [W_{\tilde{M}_p^*(\sigma_0)} \neq W_{\tilde{M}_p^*(\sigma_0^-)} \vee W_{\tilde{M}_p^*(\sigma_1)} \neq W_{\tilde{M}_p^*(\sigma_1^-)}]. \tag{91}$$

Clearly, the only way that (89) and (91) can occur is if, for some e, e', and e'', with $2e' > 2e + 1$ and $2e'' > 2e + 1$,

$$\tilde{M}_p^*(\sigma_0^-) = f(e+1, e'+1) \ \wedge \ \tilde{M}_p^*(\sigma_1^-) = f(e+1, 0) \ \wedge \ \text{last}(\sigma_0) = \text{last}(\sigma_1) = 2e'', \tag{92}$$

or (92) with σ_0 and σ_1 reversed. If σ_0 and σ_1 are reversed, then the proof is symmetric. So, suppose (92). Note that (92) implies

$$\tilde{M}_p^*(\sigma_0) = f(e+1, e'+1) \ \wedge \ \tilde{M}_p^*(\sigma_1) = f(e+1, e''+1). \tag{93}$$

Let $L \in \mathcal{L}$ be such that $\sigma_0 \in \text{Seq}(L)$. By the first conjunct of (92),

$$\{2e+1, 2e'\} \subseteq \text{content}(\sigma_0^-). \tag{94}$$

Furthermore,

$$
\begin{array}{lll}
& \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_{e'} & \\
= & W_{f(e+1, e'+1)} & \{\text{by (85)}\} \\
= & W_{\tilde{M}_p^*(\sigma_0^-)} & \{\text{by the first conjunct of (92)}\} \\
= & W_{\tilde{M}_p^*(\sigma_1^-)} & \{\text{by the first conjunct of (89)}\} \\
= & W_{f(e+1, 0)} & \{\text{by the second conjunct of (92)}\} \\
= & \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_e & \{\text{by (84)}\}.
\end{array}
\tag{95}
$$

Given (94) and (95), a straightforward analysis of (76) reveals that L must be a language for which $W_e = X$. Furthermore, (i)-(iii) below.

 (i) L is the only language in $\mathcal{L}$ for which $2e + 1 \in L$.
 (ii) $L = \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_e$.
 (iii) $(\forall e'' \in W_e)[W_{e''} = W_e]$.

By the second conjunct of (92),

$$2e + 1 \in \text{content}(\sigma_1^-). \tag{96}$$

Thus: by (i) above, $\sigma_1 \in \text{Seq}(L)$; by (ii) above, $e'' \in W_e$; and by (iii) above,

$$W_{e''} = W_e. \tag{97}$$

Finally,

$$
\begin{array}{lll}
W_{\tilde{M}_p^*(\sigma_0)} & = & W_{f(e+1, e'+1)} \hspace{2cm} \{\text{by the first conjunct of (93)}\} \\
& = & W_{\tilde{M}_p^*(\sigma_0^-)} \hspace{1.9cm} \{\text{by the first conjunct of (92)}\} \\
& = & W_{\tilde{M}_p^*(\sigma_1^-)} \hspace{1.9cm} \{\text{by the first conjunct of (89)}\} \\
& = & W_{f(e+1, 0)} \hspace{2.1cm} \{\text{by the second conjunct of (92)}\} \\
& = & \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_e \quad \{\text{by (84)}\} \\
& = & \{0, 2, ..., 2e\} \cup \{2e+1\} \cup 2W_{e''} \quad \{\text{by (97)}\} \\
& = & W_{f(e+1, e''+1)} \hspace{1.5cm} \{\text{by (85)}\} \\
& = & W_{\tilde{M}_p^*(\sigma_1)} \hspace{1.9cm} \{\text{by the second conjunct of (93)}\}.
\end{array}
$$

- STAGE 0. Search for an $m \geq 1$ such that

$$M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m+1}\big)\!\downarrow = M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^m\big).$$

  If such an m is found, then set $\sigma_1 = (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^m$, and go to stage 1. If *no* such m is found, then search indefinitely.

- STAGE 1. For larger and larger values of n, do (a) and (b) below.
  - (a) Make it the case that $W_{e_1} = \cdots = W_{e_n} = \{e_1, ..., e_n\}$.
  - (b) If there exists an $i \in \{1, ..., n\}$ such that $M_p^*(\sigma_1 \cdot 2e_i)\!\downarrow \neq M_p^*(\sigma_1)$ in at most n steps, then make it the case that $W_{e_0} = \{e_1, ..., e_n\}$, and terminate the construction. If *no* such i exists, then proceed with the next value of n.

  (Note that if *no* i is ever found in part (b) of stage 1 above, then $(\forall i \geq 1)[W_{e_i} = \{e_i \mid i \geq 1\}]$.)

Figure 5: The construction of $(e_i)_{i \in \mathbb{N}}$ in the proof of Theorem 16.

$\square$ (**Claim 16.1**)

It remains to show that $\mathcal{L}_0 \notin \mathbf{InjIt}$. By way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Then, there exists a k_0 such that

$$(\forall e \geq k_0)\big[M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2k_0) \cdot 2e\big)\!\downarrow = M_p^*(0 \cdot 2 \cdot \cdots \cdot 2k_0)\big]. \tag{98}$$

By the 1-1 Operator Recursion Theorem, there exists a computably enumerable sequence of pairwise-distinct φ-programs $(e_i)_{i \in \mathbb{N}}$ such that

$$(\forall i)[e_i \geq k_0], \tag{99}$$

and such that the behavior of $(e_i)_{i \in \mathbb{N}}$ is as in Figure 5.

Claim 16.2. Stage 0 is exited.

Proof of Claim. By way of contradiction, suppose that stage 0 is *not* exited. Then, $W_{e_0} = \emptyset$. Let L be as follows.

$$L = \{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\}. \tag{100}$$

Note that L is a language in $\mathcal{L}_0$. Furthermore,

$$t = (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^\infty \tag{101}$$

is a text for L. But, since stage 0 is *not* exited, M_p^* does *not* converge to a single hypothesis on this text — a contradiction. $\square$ (**Claim 16.2**)

For the remainder of the proof of the theorem, let m_0 be the m discovered in stage 0. By Claim 16.2, such an m_0 exists. Let σ_1 be as it would be set in stage 0, i.e.,

$$\sigma_1 = (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0}. \tag{102}$$

Consider the following cases.

CASE [stage 1 is *not* exited]. Then, $W_{e_0} = \emptyset$ and $(\forall i \geq 1)[W_{e_i} = \{e_i \mid i \geq 1\}]$. Furthermore, for each $i \geq 1$,

$$[M_p^*(\sigma_1 \cdot 2e_i) = M_p^*(\sigma_1)] \ \vee \ [M_p^*(\sigma_1 \cdot 2e_i)\uparrow]. \tag{103}$$

Let L_0 and L_1 be as follows.

$$
\begin{aligned}
L_0 &= \{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\}. &(104)\\
L_1 &= \{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\} \cup \{2e_i \mid i \geq 1\}. &(105)
\end{aligned}
$$

Note that each of L_0 and L_1 is a language in $\mathcal{L}_0$. Furthermore,

$$
\begin{aligned}
t_0 &= (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{\infty}; &(106)\\
t_1 &= (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0} \cdot 2e_1 \cdot 2e_2 \cdot \cdots &(107)
\end{aligned}
$$

are, respectively, texts for L_0 and L_1. But, by (103), either M_p^* converges to the same hypothesis on each of these texts, or M_p^* diverges on t_1. Either way, $(M, p, (X_j)_{j \in \mathbb{N}})$ does *not* **InjIt**-identify $\mathcal{L}_0$ — a contradiction.

CASE [stage 1 *is* exited]. Then, for some n_0, $(\forall i \leq n_0)[W_{e_i} = \{e_1, ..., e_{n_0}\}]$. Furthermore, for some $i_0 \in \{1, ..., n_0\}$,

$$M_p^*(\sigma_1 \cdot 2e_{i_0})\downarrow \neq M_p^*(\sigma_1). \tag{108}$$

Let L be as follows.

$$L = \{0, 2, ..., 2e_0\} \cup \{2e_0 + 1\} \cup \{2e_1, ..., 2e_{n_0}\}. \tag{109}$$

Note that L is a language in $\mathcal{L}_0$. Furthermore,

$$
\begin{aligned}
t_0 &= (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_1 \cdot 2e_2 \cdot \cdots \cdot 2e_{n_0}) \cdot (2e_0 + 1)^{\infty}; &(110)\\
t_1 &= (0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_1 \cdot 2e_2 \cdot \cdots \cdot 2e_{n_0}) \cdot (2e_0 + 1)^{m_0} \cdot (2e_{i_0})^{\infty} &(111)
\end{aligned}
$$

are each texts for L. Let j_0 be such that

$$M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0}\big) = j_0. \tag{112}$$

Note that, for each $m \geq m_0$,

$$
\begin{aligned}
&M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_1 \cdot 2e_2 \cdot \cdots \cdot 2e_{n_0}) \cdot (2e_0 + 1)^{m}\big) \\
=\ & M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m}\big) && \text{\{by (98) and (99)\}}\\
=\ & M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0}\big) && \text{\{by the choice of } m_0\}\\
=\ & j_0 && \text{\{by (112)\}}.
\end{aligned}
$$

30

Thus, on t_0, M_p^* converges to j_0. Let j_1 be the hypothesis to which M_p^* converges on t_1. Note that

$$
\begin{aligned}
& M(j_0, 2e_{i_0}) \\
= \ & M\Big(M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0}\big), 2e_{i_0}\Big) && \{\text{by (112)}\} \\
= \ & M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0} \cdot 2e_{i_0}\big) && \{\text{immediate}\} \\
\neq \ & M_p^*\big((0 \cdot 2 \cdot \cdots \cdot 2e_0) \cdot (2e_0 + 1)^{m_0}\big) && \{\text{by (108)}\} \\
= \ & j_0 && \{\text{by (112)}\}.
\end{aligned}
$$

Thus, it must be the case that $j_0 \neq j_1$. But this contradicts the fact that $(M, p, (X_j)_{j \in \mathbb{N}})$ **InjIt**-identifies L. $\qquad\qquad \square$ (**Theorem 16**)

We conclude this section with the following remark.

Remark 17. The fact that **It** $\nsubseteq$ **InjIt** (as opposed to **It** $\nsubseteq$ **ExtIt** or **ExtIt** $\nsubseteq$ **InjIt**) can be shown directly using either of the next two pre-existing results.

- **It** $\nsubseteq$ **OrdIndIt** (see footnote 12 above).

- There exists a class of languages that *can* be **It**-identified, but that can*not* be so identified *strongly non-U-shapedly* [CK10, Theorem 5.4] (see also [Bei84, Wie91, CM08]).

5. Iterative Learning by Enumeration Operator

This section examines the iterative learning by enumeration operator model (**EOIt**). Recall that **EOIt** is similar to **It**, except that the computational component of the learner is modeled as an enumeration operator, as opposed to a partial computable function. Our main results of this section are the following.

- **EOIt** $\nsubseteq$ **It** (Proposition 20).

- **It** $\nsubseteq$ (**ExtIt** $\cup$ **EOIt**) (Theorem 22).

- (**It** $\cap$ **EOIt**) $\nsubseteq$ **ExtIt** (Theorem 23).

- (**FrIt** $\cap$ **ClassInjIt**) $\nsubseteq$ **EOIt** (Theorem 24).

- **ExtIt** $\nsubseteq$ (**InjIt** $\cup$ **EOIt**) (Proposition 26(a)).

- **InjIt** $\nsubseteq$ (**FrIt** $\cup$ **ClassInjIt** $\cup$ **EOIt**) (Proposition 26(b)).

- **ClassInjIt** $\nsubseteq$ (**FrIt** $\cup$ **EOIt**) (Proposition 26(c)).

- **FrIt** $\nsubseteq$ (**ClassInjIt** $\cup$ **EOIt**) (Proposition 26(d)).

This section also includes two other results: Theorem 18 and Proposition 21. The purpose of Theorem 18 is to correct an error in the conference version of this paper. The purpose of Proposition 21 is to fulfill the diagram in Figure 1.

The conference version of this paper incorrectly claimed that every computably finitely thick class of languages can be **EOIt**-identified [MZ10, Theorem 12]. Theorem 18 just below establishes that this claim was incorrect.

For each i, execute stage 0 below.

- STAGE 0. Include $\mathbb{N} + i$ and $\{i\}$ in $\mathcal{L}$. Go to stage 1.

- STAGE 1. Let $\mathcal{M}$ be the ith element of $(\mathcal{M}_{i'})_{i' \in \mathbb{N}}$. Search for a $k \geq i$ such that
$$\mathcal{M}^*_{\{i\}}(i \cdot \cdots \cdot k) \nsubseteq \{i, ..., k\}.$$
  If such a k is found, then include $\{i, ..., k\}$ in $\mathcal{L}$, and terminate the construction (for i). If *no* such k is found, then search indefinitely.

Figure 6: The construction of $\mathcal{L}$ in the proof of Theorem 18.

Theorem 18. There exists a uniformly decidable, computably finitely thick class of languages $\mathcal{L}$ such that $\mathcal{L} \notin \textbf{EOIt}$.

The proof of Theorem 18 relies on Lemma 19 just below.

Lemma 19. For each $(\mathcal{M}, X) \in \mathcal{EO}_{1,1} \times \mathcal{CE}$, and each $\rho, \sigma, \tau \in \text{Seq}$,

$$\mathcal{M}^*_X(\rho) \subseteq \mathcal{M}^*_X(\sigma) \ \Rightarrow \ \mathcal{M}^*_X(\rho \cdot \tau) \subseteq \mathcal{M}^*_X(\sigma \cdot \tau). \tag{113}$$

Proof. A straightforward induction using essentially the monotonicity of $\mathcal{M}$.
$$\square \ (\textbf{Lemma 19})$$

Proof of Theorem 18. Let $(\mathcal{M}_{i'})_{i' \in \mathbb{N}}$ be an algorithmic enumeration of $\mathcal{EO}_{1,1}$. The class $\mathcal{L}$ is constructed in Figure 6. It is straightforward to construct an effective numbering witnessing that $\mathcal{L}$ is uniformly decidable. That $\mathcal{L}$ is computably finitely thick can be shown using a technique similar to that used in the proof of Theorem 9 above (see Figure 4(b), specifically).

To show that $\mathcal{L} \notin \textbf{EOIt}$: by way of contradiction, suppose otherwise, as witnessed by $(\mathcal{M}, X) \in \mathcal{EO}_{1,1} \times \mathcal{CE}$. Let i be such that $\mathcal{M}$ is the ith element of $(\mathcal{M}_{i'})_{i' \in \mathbb{N}}$. Since $\{i\} \in \mathcal{L}$, there exists an m_0 such

$$\mathcal{M}^*_X(i^{m_0}) = \{i\}. \tag{114}$$

Furthermore, since $(\mathbb{N} + i) \in \mathcal{L}$, there exists some $k_0 \geq i$ that is discovered in stage 1 of the construction of $\mathcal{L}$ for i. Thus,

$$\mathcal{M}^*_{\{i\}}(i \cdot \cdots \cdot k_0) \nsubseteq \{i, ..., k_0\}, \tag{115}$$

and $\{i, ..., k_0\}$ is in $\mathcal{L}$. Let σ be such that

$$\sigma = i \cdot \cdots \cdot k_0. \tag{116}$$

Let t be such that

$$t = i^{m_0} \cdot \sigma^\infty. \tag{117}$$

32

Note that t is a text for $\{i, ..., k_0\}$. It follows that there exists an n_0 such that

$$(\forall n \geq n_0)\big[\mathcal{M}_X^*(t[n]) = \{i, ..., k_0\}\big]. \tag{118}$$

Without loss of generality, suppose that $n_0 - m_0$ is divisible by $|\sigma|$. By (114) and (118),

$$\mathcal{M}_X^*(i^{m_0}) = \{i\} \subseteq \{i, ..., k_0\} = \mathcal{M}_X^*(t[n_0]). \tag{119}$$

Note that

$$
\begin{array}{rcll}
\mathcal{M}_{\{i\}}^*(\sigma) & = & \mathcal{M}_X^*(i^{m_0} \cdot \sigma) & \{\text{by } (114)\} \\
& \subseteq & \mathcal{M}_X^*(t[n_0] \cdot \sigma) & \{\text{by } (119) \text{ and Lemma } 19\} \\
& = & \mathcal{M}_X^*(t[n_0 + |\sigma|]) & \{\text{by } (117)\} \\
& = & \{i, ..., k_0\} & \{\text{by } (118)\}.
\end{array}
\tag{120}
$$

But (115) and (120) are contradictory. $\qquad\square$ (**Theorem 18**)

Proposition 20 just below establishes that **EOIt** $\not\subseteq$ **It**.[14]

Proposition 20. EOIt $\not\subseteq$ **It**.

Proof. Let $\mathcal{L}_1$ be the following class of languages.

$$\mathcal{L}_1 = \big\{K \cup \{x\} \mid x \in \mathbb{N}\big\}. \tag{121}$$

It is well known that $\mathcal{L}_1 \notin$ **It** (see, e.g., [JORS99, Proposition 4.7]). On the other hand, it is easily seen that $\mathcal{L}_1 \in$ **EOIt**. $\qquad\square$ (**Proposition 20**)

Proposition 21 just below establishes that several of the classes appearing earlier in this paper are **EOIt**-identifiable.

Proposition 21. Let $\mathcal{L}_0$ be as in the proof of Theorem 16. Then,

$$\big\{\mathit{Pat}^\Sigma, \mathcal{Fin}, \overline{\mathcal{K}}, \mathcal{Fin} \vee \overline{\mathcal{K}}, \mathit{Coinit}, \mathcal{L}_0\big\} \subseteq \textbf{EOIt}. \tag{122}$$

Proof. By a straightforward adaptation of Lange & Wiehagen's pattern language learning algorithm [LW91], it can be shown that $\mathit{Pat}^\Sigma \in$ **EOIt**. It is similarly straightforward to show that $\{\mathcal{Fin}, \overline{\mathcal{K}}, \mathcal{Fin} \vee \overline{\mathcal{K}}, \mathit{Coinit}\} \subseteq$ **EOIt**.

To show that $\mathcal{L}_0 \in$ **EOIt**: Let $\Psi : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ be such that, for each X,

$$
\begin{array}{rl}
\Psi(X) = X & \cup \quad \bigcup\{2W_e \mid 2e + 1 \in X\} \\
& \cup \quad \bigcup\{2W_{e'} \mid 2e' \in X \ \wedge \ (\exists e)[2e + 1 \in X \ \wedge \ 2e + 1 < 2e']\}.
\end{array}
\tag{123}
$$

---

[14]For the reader familiar with **TxtEx**-learning [Gol67, JORS99]: note that the proof of Proposition 20 also establishes that **EOIt** is not contained in **TxtEx**. On the other hand, it is easily seen that **EOIt** is contained in **TxtBc** [CL82, JORS99].

Clearly, $\Psi \in \mathcal{EO}_{1,0}$. Let $\tilde{\mathcal{M}} : \mathcal{P}(\mathbb{N}) \times \mathbb{N}_{\#} \to \mathcal{P}(\mathbb{N})$ be such that, for each X and e:

$$\tilde{\mathcal{M}}(X, \#\ \ \ ) = X; \tag{124}$$

$$\tilde{\mathcal{M}}(X, 2e\ \ \ ) = \Psi(X \cup \{2e\}); \tag{125}$$

$$\tilde{\mathcal{M}}(X, 2e+1) = \Psi(\{0, 2, ..., 2e\} \cup \{2e+1\}) \cup \begin{cases} X, & \text{if } 2e+1 \in X; \\ \emptyset, & \text{otherwise.} \end{cases} \tag{126}$$

Clearly, $\tilde{\mathcal{M}} \in \mathcal{EO}_{1,1}$. Furthermore, it is straightforward to show that $(\tilde{\mathcal{M}}, 2\mathbb{N})$ **EOIt**-identifies $\mathcal{L}_0$. $\qquad\qquad\qquad\qquad\qquad\square$ (**Proposition 21**)

Theorem 22 just below establishes that $\textbf{It} \not\subseteq (\textbf{ExtIt} \cup \textbf{EOIt})$.

Theorem 22. $\textbf{It} \not\subseteq (\textbf{ExtIt} \cup \textbf{EOIt})$.

Proof. For each $i \geq 1$, let A_i and B_i be as follows.

$$A_i = \{x \mid x < 4i\} \cup \{4i\}. \tag{127}$$

$$B_i = \{x \mid x < 4i\} \cup \{4i+1\}. \tag{128}$$

Let $\mathcal{L}_2$ be such that

$$\mathcal{L}_2 = \{2\mathbb{N}, 2\mathbb{N}+1\} \cup \{A_i, B_i \mid i \in \mathbb{N}\}. \tag{129}$$

To show that $\mathcal{L}_2 \in \textbf{It}$, we describe informally a learner that **It**-identifies $\mathcal{L}_2$. Suppose that the first non-$\#$ element that the learner sees is even. (The case when the first such element is odd is handled similarly.) Then, the learner hypothesizes $2\mathbb{N}$ until, if ever, it sees an odd element. From then on, the learner keeps track of the *largest* odd element seen, whilst ignoring all even elements. Let x be the largest odd element seen, and let i be *least* such that $x \leq 4i + 1$. Note that, for each $i \geq 1$:

$$A_i \cap 2\mathbb{N}+1 = (\{x \mid x < 4i\} \cup \{4i\}\ \ \ \ ) \cap 2\mathbb{N}+1 = \{x \in 2\mathbb{N}+1 \mid x \leq 4i-1\}; \tag{130}$$

$$B_i \cap 2\mathbb{N}+1 = (\{x \mid x < 4i\} \cup \{4i+1\}) \cap 2\mathbb{N}+1 = \{x \in 2\mathbb{N}+1 \mid x \leq 4i+1\}. \tag{131}$$

Thus, if $x < 4i + 1$, then the learner hypothesizes A_i; whereas, if $x = 4i + 1$, then the learner hypothesizes B_i. Clearly, such a learner **It**-identifies $\mathcal{L}_2$.

To show that $\mathcal{L}_2 \notin \textbf{ExtIt}$: by way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Let σ and τ be locking sequences for $(M, p, (X_j)_{j \in \mathbb{N}})$ on $2\mathbb{N}$ and $2\mathbb{N} + 1$, respectively. Let $i \geq 1$ be *least* such that $(\text{content}(\sigma) \cup \text{content}(\tau)) \subseteq A_i$. Let t be such that

$$t = (0 \cdot 1 \cdot \cdots \cdot 4i) \cdot \#^{\infty}. \tag{132}$$

Note that each of $\sigma \cdot t$ and $\tau \cdot t$ is a text for A_i. Thus, there exists an $n > 4i$ such that

$$X_{M_p^*(\sigma \cdot t[n])} = A_i = X_{M_p^*(\tau \cdot t[n])}. \tag{133}$$

34

Let t_0 and t_1 be as follows.

$$t_0 = \sigma \cdot (4i+4) \cdot t[n] \cdot 4i \cdot (4i+1) \cdot (4i+2) \cdot (4i+3) \cdot \#^\infty. \quad (134)$$
$$t_1 = \tau \cdot (4i+5) \cdot t[n] \cdot 4i \cdot (4i+1) \cdot (4i+2) \cdot (4i+3) \cdot \#^\infty. \quad (135)$$

Note that t_0 is a text for A_{i+1}, and that t_1 is a text for B_{i+1}. Furthermore, by the choices of σ and τ:

$$M_p^*\big(\sigma \cdot (4i+4)\big) = M_p^*(\sigma); \quad (136)$$
$$M_p^*\big(\tau \cdot (4i+5)\big) = M_p^*(\tau). \quad (137)$$

It follows that

$$X_{M_p^*(\sigma \cdot (4i+4) \cdot t[n])} = A_i = X_{M_p^*(\tau \cdot (4i+5) \cdot t[n])}. \quad (138)$$

Thus, since $(M, p, (X_j)_{j \in \mathbb{N}})$ **ExtIt**-identifies $\mathcal{L}_2$, $(M, p, (X_j)_{j \in \mathbb{N}})$ must converge to hypotheses for the same language on t_0 and t_1. But this contradicts the fact that t_0 and t_1 are texts for *distinct* languages in $\mathcal{L}_2$.

The proof that $\mathcal{L}_2 \notin$ **EOIt** is similar to the proof that $\mathcal{L}_2 \notin$ **ExtIt**.

$\square$ (**Theorem 22**)

Theorem 23 just below establishes that $(\mathbf{It} \cap \mathbf{EOIt}) \not\subseteq \mathbf{ExtIt}$. Let $\mathcal{L}_2$ be as in the proof of Theorem 22, and let $\mathbb{N} \times \mathcal{L}_2$ be shorthand for $\{\mathbb{N} \times L \mid L \in \mathcal{L}_3\}$. The class $\mathcal{L}_3$ constructed in the proof of Theorem 23 is a proper subclass of $\mathbb{N} \times \mathcal{L}_2$. It is straightforward to verify that, like $\mathcal{L}_2$, $\mathbb{N} \times \mathcal{L}_2 \in \mathbf{It} - (\mathbf{ExtIt} \cup \mathbf{EOIt})$. Intuitively, $\mathcal{L}_3$ includes sufficiently much of $\mathbb{N} \times \mathcal{L}_2$ so that $\mathcal{L}_3 \notin \mathbf{ExtIt}$, while $\mathcal{L}_3$ excludes sufficiently much of $\mathbb{N} \times \mathcal{L}_2$ so that $\mathcal{L}_3 \in \mathbf{EOIt}$.

Theorem 23. $(\mathbf{It} \cap \mathbf{EOIt}) \not\subseteq \mathbf{ExtIt}$.

Proof. Let $((M, p)_{i'})_{i' \in \mathbb{N}}$ be an algorithmic enumeration of all pairs of type $\mathcal{PC}_{1,1} \times \mathbb{N}$. A class $\mathcal{L}_3$ is constructed in Figure 7.

To show that $\mathcal{L}_3 \in \mathbf{It}$: Let $\mathcal{L}_2$ be as in the proof of Theorem 22. Let $r : \mathbb{N}_\# \to \mathbb{N}_\#$ be such that: $r(\#) = \#$, and, for each x, $(\exists i)[x = \langle i, r(x) \rangle]$. (Thus, on $\{\#\}$, r is the identity, and, on $\mathbb{N}$, r is the second projection with respect to $\langle \cdot, \cdot \rangle$.) Note that, for each $L \in \mathcal{L}_3$, $L \neq \emptyset$ and $r(L) \in \mathcal{L}_2$. Thus, an **It**-learner for $\mathcal{L}_3$ can be made to work as follows. Suppose that the learner is fed $t \in \text{Txt}(\mathcal{L}_3)$. Then, upon seeing the first non-$\#$ element of t, the learner extracts the unique i such that $\text{content}(t) \cap \{\langle i, x \rangle \mid x \in \mathbb{N}\} \neq \emptyset$. Thereafter, the learner for $\mathcal{L}_3$ simulates a learner for $\mathcal{L}_2$ on $r \circ t$. Whenever the learner for $\mathcal{L}_2$ outputs a hypothesis for a set X, the learner for $\mathcal{L}_3$ outputs a hypothesis for $\{\langle i, x \rangle \mid x \in X\}$. Clearly, such a leaner **It**-identifies $\mathcal{L}_3$.

To show that $\mathcal{L}_3 \in \mathbf{EOIt}$: Let $(E_i)_{i \in \mathbb{N}}$ and $(O_i)_{i \in \mathbb{N}}$ be such that, for each i:

$$E_i = \begin{cases} \{\langle i, 0 \rangle, \langle i, 2 \rangle, ..., \langle i, 4j_0 - 2 \rangle\}, & \text{where } j_0 \geq 1 \text{ is } \textit{least} \text{ such that } \textit{at least one} \text{ of (a)-(c)} \\ & \quad \text{in Figure 7 hold for } i, \text{ if such a } j_0 \text{ exists;} \quad (139) \\ \{\langle i, x \rangle \mid x \in 2\mathbb{N}\}, & \text{otherwise;} \end{cases}$$

$$O_i = \begin{cases} \{\langle i, 1 \rangle, \langle i, 3 \rangle, ..., \langle i, 4j_0 - 1 \rangle\}, & \text{where } j_0 \geq 1 \text{ is } \textit{least} \text{ such that } \textit{at least one} \text{ of (a)-(c)} \\ & \quad \text{in Figure 7 hold for } i, \text{ if such a } j_0 \text{ exists;} \quad (140) \\ \{\langle i, x \rangle \mid x \in 2\mathbb{N}+1\}, & \text{otherwise.} \end{cases}$$

35

For each i, do the following. Let (M, p) be the ith pair in $((M, p)_{i'})_{i' \in \mathbb{N}}$. Let $j_0 \geq 1$ be *least*, if any, such that *at least one* of (a)-(c) below hold.

(a) $M_p^*(\langle i, 0 \rangle \cdot \langle i, 2 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 4 \rangle) \uparrow$.

(b) $M_p^*(\langle i, 1 \rangle \cdot \langle i, 3 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 5 \rangle) \uparrow$.

(c) $\quad M_p^*(\langle i, 0 \rangle \cdot \langle i, 2 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 4 \rangle) = M_p^*(\langle i, 0 \rangle \cdot \langle i, 2 \rangle \cdot \cdots \cdot \langle i, 4j_0 - 2 \rangle)$
$\quad \wedge \ M_p^*(\langle i, 1 \rangle \cdot \langle i, 3 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 5 \rangle) = M_p^*(\langle i, 1 \rangle \cdot \langle i, 3 \rangle \cdot \cdots \cdot \langle i, 4j_0 - 1 \rangle)$.

If such a j_0 exists, then, for each $j \geq j_0$, include $A_{i,j}$ and $B_{i,j}$ in $\mathcal{L}_3$, where:

$$A_{i,j} = \{\langle i, x \rangle \mid x < 4j\} \cup \{\langle i, 4j \rangle\};$$
$$B_{i,j} = \{\langle i, x \rangle \mid x < 4j\} \cup \{\langle i, 4j + 1 \rangle\}.$$

If *no* such j_0 exists, then include $\{\langle i, x \rangle \mid x \in 2\mathbb{N}\}$ and $\{\langle i, x \rangle \mid x \in 2\mathbb{N} + 1\}$ in $\mathcal{L}_3$.

Figure 7: The construction of $\mathcal{L}_3$ in the proof of Theorem 23.

Note that each of E_i and O_i is computably enumerable, uniformly in i. Furthermore, it is easily verifiable that, for each $L \in \mathcal{L}_3$ and i, (i) and (ii) below.

(i) $L \cap \{\langle i, x \rangle \mid x \in 2\mathbb{N} \quad \} \neq \emptyset \ \Rightarrow \ E_i \subseteq L$.
(ii) $L \cap \{\langle i, x \rangle \mid x \in 2\mathbb{N} + 1\} \neq \emptyset \ \Rightarrow \ O_i \subseteq L$.

Let $\mathcal{M} : \mathcal{P}(\mathbb{N}) \times \mathbb{N}_{\#} \to \mathcal{P}(\mathbb{N})$ be such that, for each X, i, and j:

$$\mathcal{M}(X, \# \qquad \quad) \ = \ X; \tag{141}$$
$$\mathcal{M}(X, \langle i, 2j \rangle \quad \ ) \ = \ X \cup E_i \cup \{\langle i, 2j \rangle\}; \tag{142}$$
$$\mathcal{M}(X, \langle i, 2j + 1 \rangle) \ = \ X \cup O_i \cup \{\langle i, 2j + 1 \rangle\}. \tag{143}$$

Clearly, $\mathcal{M} \in \mathcal{EO}_{1,1}$. Using (i) and (ii) above, it is straightforward to show that $(\mathcal{M}, \emptyset)$ **EOIt**-identifies $\mathcal{L}_3$.

Finally, to show that $\mathcal{L}_3 \notin \textbf{ExtIt}$: by way of contradiction, suppose otherwise, as witnessed by $(M, p, (X_j)_{j \in \mathbb{N}}) \in \mathcal{PC}_{1,1} \times \mathbb{N} \times \mathcal{EN}$. Let i be such that (M, p) is the ith pair in $((M, p)_{i'})_{i' \in \mathbb{N}}$. First, consider the case that *no* j_0 exists as in Figure 7 for i. Then, $\{\langle i, x \rangle \mid x \in 2\mathbb{N}\}$ and $\{\langle i, x \rangle \mid x \in 2\mathbb{N} + 1\}$ are in $\mathcal{L}_3$. However, it is easily seen that M_p^* makes infinitely many mind-changes on either $\langle i, 0 \rangle \cdot \langle i, 2 \rangle \cdot \cdots$ or $\langle i, 1 \rangle \cdot \langle i, 3 \rangle \cdot \cdots$ — a contradiction. So, for the remainder of the proof of the theorem, suppose that such a j_0 exists. Thus, for each $j \geq j_0$, $A_{i,j}$ and $B_{i,j}$ are in $\mathcal{L}_3$, where $A_{i,j}$ and $B_{i,j}$ are as in Figure 7. Consider the following cases, based on (a)-(c) in Figure 7.

CASE $[M_p^*(\langle i, 0 \rangle \cdot \langle i, 2 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 4 \rangle) \uparrow]$. Then, clearly, one can construct a text for $A_{i, j_0 + 1}$ on which M_p^* diverges — a contradiction.

CASE $[M_p^*(\langle i, 1 \rangle \cdot \langle i, 3 \rangle \cdot \cdots \cdot \langle i, 4j_0 + 5 \rangle) \uparrow]$. Similar to the previous case.

CASE [$M_p^*(\langle i,0\rangle \cdot \langle i,2\rangle \cdot \cdots \cdot \langle i,4j_0+4\rangle) = M_p^*(\langle i,0\rangle \cdot \langle i,2\rangle \cdot \cdots \cdot \langle i,4j_0-2\rangle)$
$\wedge\ M_p^*(\langle i,1\rangle \cdot \langle i,3\rangle \cdot \cdots \cdot \langle i,4j_0+5\rangle) = M_p^*(\langle i,1\rangle \cdot \langle i,3\rangle \cdot \cdots \cdot \langle i,4j_0-1\rangle)]$.
Then, it can be shown that $(M,p,(X_j)_{j\in\mathbb{N}})$ does *not* identify either A_{i,j_0+1} or B_{i,j_0+1} in much the same way that the learner in the proof of Theorem 22 was shown to not identify either the content of the text in (134) or the content of the text in (135). Again, this leads to a contradiction. □ (**Theorem 23**)

Theorem 24 just below establishes that $(\mathbf{FrIt}\cap\mathbf{ClassInjIt})\nsubseteq\mathbf{EOIt}$.

Theorem 24. $(\mathbf{FrIt}\cap\mathbf{ClassInjIt})\nsubseteq\mathbf{EOIt}$.

The proof of Theorem 24 relies on Lemma 19 above, and on Lemma 25 just below.

Lemma 25. There exists a Friedberg numbering $(X_j)_{j\in\mathbb{N}}$ satisfying

$$(\forall j)[1\le|D_j|\le 3\ \Rightarrow\ X_j=D_j]. \tag{144}$$

Proof. Let $(Y_k)_{k\in\mathbb{N}}$ be any Friedberg numbering. Note that each of the following sets is computably enumerable.

  (i) $\{k\mid Y_k=\emptyset\ \vee\ |Y_k|>3\}$.
  (ii) $\{j\mid D_j=\emptyset\ \vee\ |D_j|>3\}$.

Thus, to construct a numbering $(X_j)_{j\in\mathbb{N}}$ as in the statement of the lemma, one can arrange that $(X_j)_{j\in\mathbb{N}}$ satisfy (144) explicitly, and that, for each k listed in (i), there is some *unique* j listed in (ii) such that $X_j=Y_k$. Such a numbering $(X_j)_{j\in\mathbb{N}}$ is clearly a Friedberg numbering. □ (**Lemma 25**)

Proof of Theorem 24. Let $(X_j)_{j\in\mathbb{N}}$ be a Friedberg numbering as asserted to exist by Lemma 25, and let $(Y_k)_{k\in\mathbb{N}}$ be any Friedberg numbering satisfying: $Y_0=\emptyset$. Let $(Z_\ell)_{\ell\in\mathbb{N}}$ be such that, for each j and k, $Z_{\langle j,k\rangle}=(2X_j)\cup(2Y_k+1)$. It is straightforward to show that $(Z_\ell)_{\ell\in\mathbb{N}}$ is a Friedberg numbering. Let $\mathcal{L}_4$ be the following class of languages.

$$\mathcal{L}_4=\{Z_{\langle j,\max D_j\rangle}\mid 1\le|D_j|\le 3\}. \tag{145}$$

Note that, for each j such that $1\le|D_j|\le 3$,

$$\begin{aligned}Z_{\langle j,\max D_j\rangle}&=(2X_j)\cup(2Y_{\max D_j}+1)\\&=(2D_j)\cup(2Y_{\max D_j}+1).\end{aligned} \tag{146}$$

It is straightforward to show that $\mathcal{L}_4\in\mathbf{FrIt}$ (e.g., using $(Z_\ell)_{\ell\in\mathbb{N}}$ as the hypothesis space), and that $\mathcal{L}_4\in\mathbf{ClassInjIt}$. It remains to show that $\mathcal{L}_4\notin\mathbf{EOIt}$. By way of contradiction, suppose otherwise, as witnessed by $(\mathcal{M},X)\in\mathcal{EO}_{1,1}\times\mathcal{CE}$. Recall that $Y_0=\emptyset$. Let $k_0=0$, and let $L_0=\{2k_0\}$. Note that

$$L_0=\{2k_0\}=\{2k_0\}\cup(2\emptyset+1)=\{2k_0\}\cup(2Y_{k_0}+1). \tag{147}$$

Thus, $L_0\in\mathcal{L}_4$. It follows that there exists an m_0 such that

$$\mathcal{M}_X^*\big((2k_0)^{m_0}\big)=L_0. \tag{148}$$

37

Let k_1 be such that $Y_{k_1} = \mathbb{N}$, and let $L_1 = \{2k_0, 2k_1\} \cup (2\mathbb{N} + 1)$. Note that $L_1 \in \mathcal{L}_4$. It follows that there exists an n_0 such that

$$(\forall n \geq n_0)\left[\mathcal{M}_X^*\left((2k_0)^{m_0} \cdot 2k_1 \cdot 1 \cdot 3 \cdot \cdots \cdot 2n + 1\right) = L_1\right]. \tag{149}$$

Let k_2 and n_1 be such that (a)-(c) below.

(a) $k_1 < k_2$.
(b) $n_0 \leq n_1$.
(c) $Y_{k_2} = \{0, ..., n_1\}$.

Clearly, such k_2 and n_1 exist. Let $L_2 = \{2k_0, 2k_1, 2k_2\} \cup \{1, 3, ..., 2n_1 + 1\}$. Note that $L_2 \in \mathcal{L}_4$, and that $L_1 \nsubseteq L_2$. Let σ be such that

$$\sigma = 2k_1 \cdot 1 \cdot 3 \cdot \cdots \cdot 2n_1 + 1, \tag{150}$$

let t be any text for L_2, and let t' be such that

$$t' = \sigma \cdot t(0) \cdot \sigma \cdot t(1) \cdot \cdots . \tag{151}$$

Note that t' is a text for L_2. It follows that there exists an i_0 such that

$$(\forall i \geq i_0)[\mathcal{M}_X^*(t'[i]) = L_2]. \tag{152}$$

Without loss of generality, suppose that i_0 is divisible by $|\sigma| + 1$. By (148) and (152),

$$\mathcal{M}_X^*\left((2k_0)^{m_0}\right) = L_0 \subseteq L_2 = \mathcal{M}_X^*(t'[i_0]). \tag{153}$$

Furthermore, by (149) and the fact that $n_0 \leq n_1$,

$$L_1 \subseteq \mathcal{M}_X^*\left((2k_0)^{m_0} \cdot \sigma\right). \tag{154}$$

From (153), (154), and Lemma 19, it follows that

$$L_1 \subseteq \mathcal{M}_X^*(t'[i_0] \cdot \sigma) = \mathcal{M}_X^*(t'[i_0 + |\sigma|]). \tag{155}$$

But since $L_1 \nsubseteq L_2$, (152) and (155) are contradictory. $\qquad \square$ (**Theorem 24**)

Proposition 26 just below completes the diagram in Figure 1.

Proposition 26. (a)-(d) below.

(a) $\textbf{ExtIt} \nsubseteq (\textbf{InjIt} \cup \textbf{EOIt})$.
(b) $\textbf{InjIt} \nsubseteq (\textbf{FrIt} \cup \textbf{ClassInjIt} \cup \textbf{EOIt})$.
(c) $\textbf{ClassInjIt} \nsubseteq (\textbf{FrIt} \cup \textbf{EOIt})$.
(d) $\textbf{FrIt} \nsubseteq (\textbf{ClassInjIt} \cup \textbf{EOIt})$.

Proof. Recall from the proof of Theorem 16 that $\mathcal{L}_0 \in \textbf{ExtIt} - \textbf{InjIt}$, and from the proof of Theorem 24 that $\mathcal{L}_4 \in (\textbf{FrIt} \cap \textbf{ClassInjIt}) - \textbf{EOIt}$. Witnessing classes for (a)-(d) in the statement of the proposition are as follows.

(a) $\mathcal{L}_0 \vee \mathcal{L}_4 \in \textbf{ExtIt} - (\textbf{InjIt} \cup \textbf{EOIt})$.

(b) $\mathcal{F}\!in \vee \overline{\mathcal{K}} \vee \mathcal{L}_4 \in \mathbf{InjIt} - (\mathbf{FrIt} \cup \mathbf{ClassInjIt} \cup \mathbf{EOIt})$.

(c) $\mathcal{F}\!in \vee \mathcal{L}_4 \in \mathbf{ClassInjIt} - (\mathbf{FrIt} \cup \mathbf{EOIt})$.

(d) $\overline{\mathcal{K}} \vee \mathcal{L}_4 \in \mathbf{FrIt} - (\mathbf{ClassInjIt} \cup \mathbf{EOIt})$.

In each case, the proof of the positive part is straightforward; whereas, the proof of the negative part is slightly more involved. We give some of the details below.

(a) It is straightforward to show that if $\mathcal{L}_0 \vee \mathcal{L}_4$ were **InjIt**-identifiable, then $\mathcal{L}_0$ would be **OrdIndIt**-identifiable. Thus, by Theorem 15, $\mathcal{L}_0$ would be **InjIt**-identifiable. But since this would contradict Theorem 16, $\mathcal{L}_0 \vee \mathcal{L}_4 \notin \mathbf{InjIt}$. Similarly, it is straightforward to show that if $\mathcal{L}_0 \vee \mathcal{L}_4$ were **EOIt**-identifiable, then $\mathcal{L}_4$ would also be **EOIt**-identifiable, contradicting Theorem 24. Thus, $\mathcal{L}_0 \vee \mathcal{L}_4 \notin \mathbf{EOIt}$.

(b) The proof that $\mathcal{F}\!in \vee \overline{\mathcal{K}} \vee \mathcal{L}_4 \notin \mathbf{FrIt}$ is similar to the proof that $\mathcal{F}\!in \notin \mathbf{FrIt}$. The existence of a hypothesis space representing $\mathcal{F}\!in \vee \overline{\mathcal{K}} \vee \mathcal{L}_4$ exactly would imply the existence of an analogous hypothesis space for $\overline{\mathcal{K}}$. Thus, $\mathcal{F}\!in \vee \overline{\mathcal{K}} \vee \mathcal{L}_4 \notin$ **ClassInjIt**. That $\mathcal{F}\!in \vee \overline{\mathcal{K}} \vee \mathcal{L}_4 \notin$ **EOIt** is argued as in part (a).

(c) That $\mathcal{F}\!in \vee \mathcal{L}_4 \notin \mathbf{FrIt}$ is argued as in part (b). That $\mathcal{F}\!in \vee \mathcal{L}_4 \notin \mathbf{EOIt}$ is argued as in part (a).

(d) That $\overline{\mathcal{K}} \vee \mathcal{L}_4 \notin \mathbf{ClassInjIt}$ is argued as in part (b). That $\overline{\mathcal{K}} \vee \mathcal{L}_4 \notin \mathbf{EOIt}$ is argued as in part (a). $\hfill \square$ (**Proposition 26**)

6. Conclusion

In this paper, we considered ways of preventing iterative learners from employing *coding tricks*, which we defined (informally) as: any use by an iterative learner of the syntax of a hypothesis to determine what elements that learner has or has not yet seen. One means of preventing *some* such coding tricks is to require that the hypothesis space used be a Friedberg numbering. Of interest, in this regard, are those classes of languages that are both uniformly decidable and computably finitely thick, and for which there exists a single effective numbering witnessing both properties simultaneously. As we showed in Section 3, any such class can be iteratively identified using a Friedberg numbering as the hypothesis space (Theorem 6).

In addition to iterative learning (**It**-learning) and Friedberg iterative learning (**FrIt**-learning), we considered several other learning models, namely: injective iterative learning (**InjIt**-learning), class injective learning (**ClassInjIt**-learning), order independent iterative learning (**OrdIndIt**-learning), extensional iterative learning (**ExtIt**-learning), and iterative learning by enumeration operator (**EOIt**-learning). We showed, for example, that injective iterative learning is equivalent to order independent iterative learning (Theorem 15), but that these models are otherwise, pairwise inequivalent (Sections 4 and 5).

There are at least three directions in which research might proceed. First, one could ask: is our definition of *coding trick* the right one? For example, as we showed in Section 5, there exists a class of languages that can*not* be

EOIt-identified, but that *can* be **FrIt** and **ClassInjIt**-identified (Theorem 24). Using our definition, one could argue that any learner for such a class employs coding tricks. However, this would seem to include a great many learners (see the discussion near the end of Section 1). Second, recall that in the **EOIt**-learning model, the computational component of a learner is modeled as an enumeration operator, as opposed to a partial computable function. As far as we are aware, this fact distinguishes **EOIt**-learning from any learning model considered previously. As such, the properties of learners in this model should be further investigated (see also footnote 14 above). Finally, there is the remaining open problem of whether every **InjIt** or **FrIt**-identifiable class of languages is contained in a **ClassInjIt**-identifiable class (Problem 13).

[Ang80] D. Angluin. Finding patterns common to a set of strings. *Journal of Computer and System Sciences*, 21(1):46–62, 1980.

[BB75] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28(2):125–155, 1975.

[BCJS10] L. Becerra-Bonache, J. Case, S. Jain, and F. Stephan. Iterative learning of simple external contextual languages. *Theoretical Computer Science*, 411(29–30):2741–2756, 2010.

[Bei84] H-R. Beick. *Induktive Inferenz mit höchster Konvergenzgeschwindigkeit.* PhD thesis, Sektion Mathematik, Humboldt-Universität Berlin, 1984.

[Cas74] J. Case. Periodicity in generations of automata. *Mathematical Systems Theory*, 8(1):15–32, 1974.

[Cas94] J. Case. Infinitary self-reference in learning theory. *Journal of Experimental and Theoretical Artificial Intelligence*, 6(1):3–16, 1994.

[CCJS07] L. Carlucci, J. Case, S. Jain, and F. Stephan. Results on memory-limited U-shaped learning. *Information and Computation*, 205(10):1551–1573, 2007.

[CK10] J. Case and T. Kötzing. Strongly non-U-shaped learning results by general techniques. In *Proc. of the 23rd Conference on Learning Theory (COLT'2010)*, pages 181–193. Omnipress, 2010.

[CL82] J. Case and C. Lynes. Machine inductive inference and language identification. In *Proc. of the 9th International Colloquium on Automata, Languages and Programming (ICALP'1982)*, volume 140 of *Lecture Notes in Computer Science*, pages 107–115. Springer, 1982.

[CM08]   J. Case and S. Moelius.   Optimal language learning.   In *Proc. of the 19th Annual Conference on Algorithmic Learning Theory (ALT'2008)*, volume 5254 of *Lecture Notes in Artificial Intelligence*, pages 419–433. Springer, 2008.

[dBY10]   M. de Brecht and A. Yamamoto. Topological properties of concept spaces (full version). *Information and Computation*, 208(4):327–340, 2010.

[FKW82]   R. Freivalds, E. Kinber, and R. Wiehagen. Inductive inference and computable one-one numberings. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 28(27):463–479, 1982.

[Fri58]   R. Friedberg. Three theorems on recursive enumeration. I. Decomposition. II. Maximal set. III. Enumeration without duplication. *Journal of Symbolic Logic*, 23(3):309–316, 1958.

[Ful90]   M. Fulk. Prudence and other conditions on formal language learning. *Information and Computation*, 85(1):1–11, 1990.

[Gol67]   E. Mark Gold. Language identification in the limit. *Information and Control*, 10(5):447–474, 1967.

[JLMZ10]   S. Jain, S. Lange, S. Moelius, and S. Zilles. Incremental learning with temporary memory. *Theoretical Computer Science*, 411(29–30):2757–2772, 2010.

[JORS99]   S. Jain, D. Osherson, J. S. Royer, and A. Sharma. *Systems that Learn: An Introduction to Learning Theory*. MIT Press, second edition, 1999.

[JS08]   S. Jain and F. Stephan. Learning in Friedberg numberings. *Information and Computation*, 206(6):776–790, 2008.

[Kum90]   M. Kummer. An easy priority-free proof of a theorem of Friedberg. *Theoretical Computer Science*, 74(2):249–251, 1990.

[LW91]   S. Lange and R. Wiehagen. Polynomial time inference of arbitrary pattern languages. *New Generation Computing*, 8(4):361–370, 1991.

[LZ96]   S. Lange and T. Zeugmann. Incremental learning from positive data. *Journal of Computer and System Sciences*, 53(1):88–103, 1996.

[LZZ08]   S. Lange, T. Zeugmann, and S. Zilles. Learning indexed families of recursive languages from positive data: A survey. *Theoretical Computer Science*, 397(1–3):194–232, 2008.

[MZ10]   S. Moelius and S. Zilles. Learning without coding. In *Proc. of the 21st Annual Conference on Algorithmic Learning Theory (ALT'2010)*, volume 6331 of *Lecture Notes in Artificial Intelligence*, pages 300–314. Springer, 2010.

[Rog67] H. Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw Hill, 1967. Reprinted, MIT Press, 1987.

[SSS⁺94] S. Shimozono, A. Shinohara, T. Shinohara, S. Miyano, S. Kuhara, and S. Arikawa. Knowledge acquisition from amino acid sequences by machine learning system BONSAI. *Transactions of the Information Processing Society of Japan*, 35(10):2009–2018, 1994.

[Wie76] R. Wiehagen. Limes-Erkennung rekursiver Funktionen durch spezielle Strategien. *Elektronische Informationsverarbeitung und Kybernetik*, 12(1/2):93–99, 1976.

[Wie91] R. Wiehagen. A thesis in inductive inference. In *Proc. of the 1st International Workshop on Nonmonotonic and Inductive Logic (1990)*, volume 543 of *Lecture Notes in Artificial Intelligence*, pages 184–207. Springer, 1991.