Effectivity Questions for Kleene's Recursion Theorem

John Case[1], Sanjay Jain[⋆2] and Frank Stephan[⋆⋆3]

[1] Department of Computer and Information Sciences
University of Delaware, Newark, DE 19716-2586, USA
case@cis.udel.edu

[2] Department of Computer Science
National University of Singapore, Singapore 117417
sanjay@comp.nus.edu.sg

[3] Department of Mathematics and Department of Computer Science
National University of Singapore, Singapore 119076
fstephan@comp.nus.edu.sg

Abstract. The present paper explores the interaction between two recursion-theoretic notions: program self-reference and learning partial recursive functions in the limit. Kleene's Recursion Theorem formalises the notion of program self-reference: It says that given a partial-recursive function ψ_p there is an index e such that the e-th function ψ_e is equal to the e-th slice of ψ_p. The paper studies constructive forms of Kleene's recursion theorem which are inspired by learning criteria from inductive inference and also relates these constructive forms to notions of learnability. For example, it is shown that a numbering can fail to satisfy Kleene's Recursion Theorem, yet that numbering can still be used as a hypothesis space when learning explanatorily an arbitrary learnable class. The paper provides a detailed picture of numberings separating various versions of Kleene's Recursion Theorem and learnability.

Keywords: inductive inference, Kleene's Recursion Theorem, Kolmogorov complexity, optimal numberings.

1 Introduction

Program self-reference is the ability of a program to make use of its own source code in its computations. This notion is formalized by Kleene's Recursion Theorem.[4] Intuitively, this theorem asserts that, for each preassigned algorithmic

⋆ Supported in part by NUS grant numbers R252-000-420-112 and C252-000-087-001.

⋆⋆ Supported in part by NUS grant number R252-000-420-112.

[4] Other "recursion theorems" do not so well capture the notion of program self-reference. For example, consider the (quasi-fixed-point) recursion theorem as formalised by Rogers [Rog67, Theorem 11-I]. In contrast to Kleene's Recursion Theorem, Rogers' recursion theorem is not strong enough to guarantee that a numbering of partial-recursive functions satisfying it has a self-reproducing program which outputs its own index [CM09].

task, there exists a program e that computes exactly the e-th slice of this algorithmic task. The theorem is stated below, following some necessary definitions.

Let $\mathbb{N}$ be the set of natural numbers, $\{0, 1, 2, \ldots\}$. Let $\mathcal{P}$ be the collection of all partial recursive functions from $\mathbb{N}$ to $\mathbb{N}$. Let $\langle \cdot, \cdot \rangle$ be Cantor's pairing function [Rog67, page 64]: $\langle x, y \rangle = (x + y)(x + y + 1)/2 + y$, which is a recursive, order preserving bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ [Rog67, page 64]; here order preserving means that $x \leq x' \wedge y \leq y' \Rightarrow \langle x, y \rangle \leq \langle x', y' \rangle$. For each $\psi \in \mathcal{P}$ and $p \in \mathbb{N}$, let ψ_p be shorthand for $\psi(\langle p, \cdot \rangle)$. An *effective numbering* of $\mathcal{P}$ is a $\psi \in \mathcal{P}$ such that

$$(\forall \alpha \in \mathcal{P})(\exists p \in \mathbb{N})[\psi_p = \alpha]. \tag{1}$$

For this paper, we shall be concerned only with numberings that are effective, and that number the elements of $\mathcal{P}$. Hence, we shall generally omit the phrases "effective" and "of $\mathcal{P}$".

The following is the formal statement of Kleene's Recursion Theorem.

Definition 1 (Kleene [Kle38]). For each numbering ψ, *Kleene's Recursion Theorem holds in* $\psi \Leftrightarrow$

$$(\forall p \in \mathbb{N})(\exists e \in \mathbb{N})[\psi_e = \psi_p(\langle e, \cdot \rangle)]. \tag{2}$$

Equation (2) can be interpreted as follows: Suppose the ψ-program p represents an arbitrary, algorithmic task to perform; then the equation says that there is a ψ-program e such that ψ_e is equal to the e-th slice of this algorithmic task. This is often used in diagonalizations by defining ψ_e in a way implicitly employing parameter e (in effect, a self-copy of e) in some algorithmic task ψ_p.

The following constructive form of Kleene's Recursion Theorem has been well-studied. For reasons that will become apparent shortly, we call this form of the theorem FinKrt.

Definition 2 (Kleene, see [Ric80,Ric81,Roy87]). A numbering ψ is called a FinKrt*-numbering* $\Leftrightarrow$

$$(\exists \text{ recursive } r : \mathbb{N} \to \mathbb{N})(\forall p)[\psi_{r(p)} = \psi_p(\langle r(p), \cdot \rangle)]. \tag{3}$$

In (3), ψ-program $r(p)$ plays the role played by e in (2). In this sense, the function r *finds* a ψ-program $r(p)$ such that $\psi_{r(p)}$ is equal to the $r(p)$-th slice of ψ_p.

In this paper, additional constructive forms of the theorem are considered. Each is inspired by a Gold-style criterion for learning partial recursive functions in the limit. The Gold-style criteria differ in when a learning device is considered to have learned a target partial recursive function. However, the following is common to all. The learning device is fed the elements of the graph of a partial recursive function α.[5] After being fed each such element, the device outputs either '?' or a hypothesis, i.e., a program, possibly corresponding to the partial

[5] The device may also be fed one or more instances of the pause symbol '#'. This allows that graph of the target partial recursive function to be empty, i.e., in such a case, the device is fed nothing but #.

recursive function α. In the present paper, the device is expected to be algorithmic, that is, definable by a computer program.

For the finite (Fin) learning criterion, the device is considered to have learned the target partial recursive function α iff the device outputs finitely many '?' immediately followed by a hypothesis corresponding to α. The constructive form of Kleene's Recursion Theorem, given in Definition 2, may be viewed in a similar way. A device is fed a program p for a preassigned task. After finitely many steps, that device outputs a program e that uses its own source code in the manner prescribed by p.

A numbering ψ is said to be optimal for Fin-learning iff every Fin-learnable class of partial recursive functions can be Fin-learned using ψ as the hypothesis space. A numbering ψ is said to be effectively optimal for Fin-learning iff one can effectively translate every Fin-learning device into a Fin-learning device that uses ψ as its hypothesis space [JS10,Jai11].

Not every numbering is optimal for Fin-learning [JS10], let alone effectively optimal. Similarly, not every numbering is a FinKrt-numbering [Ric80,Ric81]. Hence, one might ask: is every FinKrt-numbering optimal for Fin-learning? Conversely, if a numbering is optimal for Fin-learning, then is it necessarily a FinKrt numbering?

Additional Gold-style learning criteria are introduced in Section 2 below and will be familiar to most readers familiar with inductive inference. These criteria, which are successively less stringent in when a learning device is considered to have learned a target partial recursive function, are: single mind-change explanatory (Ex_1), explanatory (Ex), vacillatory (Vac) and behaviorally correct (Bc). Section 2 also introduces additional constructive forms of Kleene's Recursion Theorem (ExKrt, VacKrt and BcKrt). Each is inspired by one of the just mentioned learning criteria. Our results include the following.

- There is a numbering which does not satisfy Kleene's Recursion Theorem, but which is optimal for Fin-learning and effectively optimal for Ex, Vac and Bc-learning (Theorem 7).
- There is a FinKrt-numbering which is not optimal for any of the learning criteria Fin, Ex, Vac, Bc (Theorem 8).
- There is an ExKrt-numbering which is not a FinKrt-numbering and which is effectively optimal for Ex-learning, but not optimal for Fin or Bc-learning (Theorem 10).
- There is a VacKrt-numbering which is not an ExKrt-numbering and which is effectively optimal for Vac-learning but not optimal for Fin, Ex or Bc-learning (Theorem 11).
- There is a BcKrt-numbering which is not a VacKrt-numbering and which is effectively optimal for Bc-learning, but not optimal for Fin, Ex or Vac-learning (Theorem 13).
- There is a numbering satisfying Kleene's Recursion Theorem which is not a BcKrt-numbering and which is not optimal for any of the learning criteria Fin, Ex, Vac, Bc (Theorem 15).

– There is a numbering satisfying Kleene's Recursion Theorem which is not a BcKrt-numbering, but which is effectively optimal for Ex, Vac and Bc-learning (Theorem 16).

The remainder of this paper is organized as follows. Section 2 covers preliminaries. Section 3 presents our results concerning numberings that do not satisfy Kleene's Recursion Theorem. Section 4 presents our results concerning numberings that satisfy Kleene's Recursion Theorem in an effective way. Section 5 presents our results concerning numberings that satisfy Kleene's Recursion Theorem, but not in an effective way.

2 Preliminaries

Recursion-theoretic concepts not covered below are treated as by Rogers [Rog67].

Lowercase math-italic letters (e.g., a, b, c) range over elements of $\mathbb{N}$, unless stated otherwise. Uppercase math-italic italicized letters (e.g., A, B, C) range over subsets of $\mathbb{N}$, unless stated otherwise. Lowercase Greek letters (e.g., α, β, γ) range over partial functions from $\mathbb{N}$ to $\mathbb{N}$, unless stated otherwise.

For each non-empty X, $\min X$ denotes the minimum element of X. We let $\min \emptyset \stackrel{\text{def}}{=} \infty$. For each non-empty, finite X, $\max X$ denotes the maximum element of X. We let $\max \emptyset \stackrel{\text{def}}{=} -1$. $D_0, D_1, D_2, \ldots$ denotes a recursive canonical enumeration of all finite subsets of $\mathbb{N}$.

The pairing function $\langle \cdot, \cdot \rangle$ was introduced in Section 1. Note that $\langle 0, 0 \rangle = 0$ and, for each x and y, $\max\{x, y\} \leq \langle x, y \rangle$.

For each one-argument partial function α and $x \in \mathbb{N}$, $\alpha(x){\downarrow}$ denotes that $\alpha(x)$ converges; $\alpha(x){\uparrow}$ denotes that $\alpha(x)$ diverges. We use $\uparrow$ to denote the value of a divergent computation. So, for example, $\lambda x \centerdot {\uparrow}$ denotes the everywhere divergent partial function.

$\mathbb{N}_{\#} \stackrel{\text{def}}{=} \mathbb{N} \cup \{\#\}$ and $\mathbb{N}_{?} \stackrel{\text{def}}{=} \mathbb{N} \cup \{?\}$. For each partial function f (of arbitrary type), $\text{rng}(f)$ denotes the range of f. A *text* is a total (not necessarily recursive) 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. Init denotes the set of all finite initial segments of all texts. For each text T and partial function α, T is a text *for* α iff $\text{rng}(T) - \{\#\}$ is the graph of α as coded by $\langle \cdot, \cdot \rangle$, i.e.,

$$\text{rng}(T) - \{\#\} = \{\langle x, y \rangle \mid \alpha(x) = y \ \wedge \ x, y \in \mathbb{N}\}. \tag{4}$$

For a total function f, we often identify f with its canonical text, that is, the text T with $T(i) = \langle i, f(i) \rangle$. Thus, $f[n]$ represents the initial segment of length n of this canonical text.

A numbering φ is *acceptable* iff for each numbering ψ, there exists a recursive function $t : \mathbb{N} \to \mathbb{N}$ such that, for each p, $\varphi_{t(p)} = \psi_p$ [Rog67,Ric80,Ric81,Roy87]. Let φ be any fixed acceptable numbering satisfying $\varphi_0 = \lambda x \centerdot {\uparrow}$. For each p, $W_p \stackrel{\text{def}}{=} \{x \mid \varphi_p(x){\downarrow}\}$. K denotes the diagonal halting problem with respect to φ, i.e., $\{x \mid x \in W_x\}$. Let $\text{pad} : \mathbb{N}^2 \to \mathbb{N}$ be a recursive function such that, for each e and y, $\varphi_{\text{pad}(e,y)} = \varphi_e$ and $\text{pad}(e, y) < \text{pad}(e, y + 1)$, where we assume

$\mathrm{pad}(0,0) = 0$.

The following are the Gold-style learning criteria considered in this paper.

Definition 3. Let α be any partial recursive function. For each recursive function $M : \mathrm{Init} \to \mathbb{N}_?$ and each numbering ψ, (a)–(e) below.

(a) [Gol67] M Fin_ψ*-learns* α iff for each text T for α, there exist i_0 and e such that

$$(\forall i < i_0)\big[M(T[i]) = ?\big] \ \wedge\ (\forall i \geq i_0)\big[M(T[i]) = e\big] \ \wedge\ \psi_e = \alpha. \tag{5}$$

(b) [CS83] M $\mathsf{Ex}_{\psi,1}$*-learns* α iff for each text T for α, there exist i_0, i_1, e_0 and e_1 such that

$$\begin{aligned}(\forall i < i_0)\big[M(T[i]) = ?\big] \ \wedge\ (\forall i \in \{i_0, \ldots, i_1 - 1\})\big[M(T[i]) = e_0\big] \\ \wedge\ (\forall i \geq i_1)\big[M(T[i]) = e_1\big] \ \wedge\ \psi_{e_1} = \alpha.\end{aligned} \tag{6}$$

(c) [Gol67] M Ex_ψ*-learns* α iff for each text T for α, there exist i_0 and e such that

$$(\forall i \geq i_0)\big[M(T[i]) = e\big] \ \wedge\ \psi_e = \alpha. \tag{7}$$

(d) [Cas99] M Vac_ψ*-learns* α iff for each text T for α, there exist i_0 and a finite set E such that

$$(\forall i \geq i_0)\big[M(T[i]) \in E\big] \ \wedge\ (\forall e \in E)[\psi_e = \alpha]. \tag{8}$$

(e) [Bar74,OW82] M Bc_ψ*-learns* α iff for each text T for α, there exists an i_0 such that

$$M(T[i_0]) \neq ? \text{ and } (\forall i \geq i_0)(\forall e)\big[M(T[i]) = e \ \Rightarrow\ \psi_e = \alpha\big]. \tag{9}$$

Let $I \in \{\mathsf{Fin}, \mathsf{Ex}_1, \mathsf{Ex}, \mathsf{Vac}, \mathsf{Bc}\}$ and let $\mathcal{S}$ be a class of partial recursive functions. M I_ψ*-learns* $\mathcal{S}$ iff M I_ψ-learns each partial recursive function in $\mathcal{S}$. We say that $\mathcal{S}$ is I_ψ-learnable if some M I_ψ-learns $\mathcal{S}$. In above definitions, we omit the subscript ψ when ψ is the fixed acceptable numbering φ.

Definition 4 (Jain & Stephan [JS10]). Let φ be an acceptable numbering. For each $I \in \{\mathsf{Fin}, \mathsf{Ex}_1, \mathsf{Ex}, \mathsf{Vac}, \mathsf{Bc}\}$ and each numbering ψ, (a) and (b) below.

(a) ψ is *optimal for I-learning* iff each I_φ-learnable class is I_ψ-learnable.

(b) ψ is *effectively optimal for I-learning* iff there exists a recursive function $t : \mathbb{N} \to \mathbb{N}$ such that, for each p and each class of partial recursive functions $\mathcal{S}$, if φ_p I_φ-learns $\mathcal{S}$, then $\varphi_{t(p)}$ I_ψ-learns $\mathcal{S}$.

Note that while for learning criteria and the below constructive versions of KRT, the implications $\mathsf{Fin} \to \mathsf{Ex}_1 \to \mathsf{Ex} \to \mathsf{Vac} \to \mathsf{Bc}$ hold, the corresponding implications do not always hold with respect to numberings being optimal or effectively optimal for I-learning. For example, there are numberings which are optimal for Vac-learning but not optimal for Bc-learning [JS10]. However, if a numbering is effectively optimal for Fin-learning, then it is effectively optimal for Ex, Vac and Bc-learning. Furthermore, if a numbering is effectively optimal for Ex-learning then it is effectively optimal for Vac-learning [JS10].

The following are the constructive forms of Kleene's Recursion Theorem considered in this paper. The reader will note the similarity to Definition 3.

Definition 5. [Moe09] For each numbering ψ, (a)–(d) below.

(a) ψ is a FinKrt-*numbering* iff there exists a recursive function $r : \mathbb{N} \to \mathbb{N}$ such that, for each p,
$$\psi_{r(p)} = \psi_p\big(\langle r(p), \cdot\rangle\big). \tag{10}$$

(b) ψ is an ExKrt-*numbering* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exist i_0 and e such that
$$(\forall i \geq i_0)[f(p,i) = e] \ \wedge\ \psi_e = \psi_p(\langle e, \cdot\rangle). \tag{11}$$

(c) ψ is a VacKrt-*numbering* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exist i_0 and a finite set E such that
$$(\forall i \geq i_0)[f(p,i) \in E] \ \wedge\ (\forall e \in E)[\psi_e = \psi_p(\langle e, \cdot\rangle)]. \tag{12}$$

(d) ψ is a BcKrt-*numbering* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exists an i_0 such that
$$(\forall i \geq i_0)(\forall e)[f(p,i) = e \ \Rightarrow\ \psi_e = \psi_p(\langle e, \cdot\rangle)]. \tag{13}$$

Definition 6. For each numbering ψ, (a) and (b) below.

(a) ψ is Ex_1-*acceptable* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exist i_0, e_0 and e_1 such that
$$(\forall i < i_0)[f(p,i) = e_0] \ \wedge\ (\forall i \geq i_0)[f(p,i) = e_1] \ \wedge\ \psi_{e_1} = \varphi_p. \tag{14}$$

(b) **(Case, Jain and Suraj [CJS02])** ψ is Ex-*acceptable* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exist i_0 and e such that
$$(\forall i \geq i_0)[f(p,i) = e] \ \wedge\ \psi_e = \varphi_p. \tag{15}$$

We use the convention that, for each y, $\log(y) \stackrel{\text{def}}{=} \min\{x \mid 2^x \geq y\}$. So, for example, $\log(0) = 0$ and $\log(3) = 2$. For each e, $C(e)$ denotes the plain Kolmogorov complexity of e [LV08,Nie09]. Note that there exists an approximation $\lambda s, e \centerdot C_s(e)$ such that, for each e, $C(e) = \lim_s C_s(e)$. Further note that, for each 1–1 recursive sequence $e_0, e_1, e_2, \ldots$, there exists a constant c such that, for each i, $C(e_{i+1}) < C(e_i) + c$.

3 When Kleene's Recursion Theorem is Absent

This section presents our results concerning numberings that do not satisfy Kleene's Recursion Theorem. Note that every acceptable numbering is a FinKrt-numbering [Kle38]. However, as the next result shows, this does not generalize to other criteria of acceptability. In particular, there is an Ex_1-acceptable numbering that does not satisfy Kleene's Recursion Theorem.

Theorem 7. *There exists a numbering ψ satisfying (a)–(d) below.*

(a) ψ does not satisfy Kleene's Recursion Theorem.
(b) ψ is an Ex_1*-acceptable numbering.*
(c) ψ effectively optimal for Ex*,* Vac *and* Bc*-learning.*
(d) ψ is optimal for Fin*-learning.*

Proof. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } \mathrm{rng}(\varphi_e) \not\subseteq \{e, e+1, e+2, \ldots\}; \\ \uparrow, & \text{if } \mathrm{rng}(\varphi_e) \subseteq \{e, e+1, e+2, \ldots\}. \end{cases} \tag{16}$$

The numbering ψ does not satisfy Kleene's Recursion Theorem: There is no e such that $\psi_e = \lambda x \centerdot e$; hence, there is no e such that $\psi_e = \alpha(\langle e, \cdot\rangle)$ when $\alpha = \lambda\langle e, x\rangle \centerdot e$.

To show that ψ is Ex_1-acceptable, there exists a translator which behaves as follows. On input e, the translator first conjectures 0 for $\psi_0 = \varphi_0 = \lambda x \centerdot \uparrow$. Then, in the case that $\varphi_e(x) = y$, for some x and y, the translator outputs $\mathrm{pad}(e, y)$. Note that $y < \mathrm{pad}(e, y)$. Hence, it follows from the definition of ψ that $\psi_{\mathrm{pad}(e,y)} = \varphi_e$.

To show that ψ is effectively optimal for Ex, Vac, and Bc-learning, given a Bc-learner M, the new learner N first conjectures 0 for $\lambda x \centerdot \uparrow$. If, however, a datum (x, y) is ever seen, then, from that point onward, N simulates M and translates each conjecture e of M into $\mathrm{pad}(e, y)$.

To show that ψ is optimal for Fin-learning, suppose that M is a Fin-learner for a class not containing $\lambda x \centerdot \uparrow$. Then, the new learner N waits for the first pair (x, y); from that point onward, N simulates M and translates each conjecture e of M into $\mathrm{pad}(e, y)$. On the other hand, suppose that M is a Fin-learner for a class containing $\lambda x \centerdot \uparrow$. Then, this class contains no other partial functions. Hence, N can just ignore all input and output 0 as $\psi_0 = \lambda x \centerdot \uparrow$. Hence, ψ is optimal for Fin-learning.[6] □ (**Theorem 7**)

4 When Kleene's Recursion Theorem is Effective

This section presents our results concerning numberings that satisfy Kleene's Recursion Theorem in an effective way. These results include the following. First, a numbering can be a FinKrt-numbering, yet not be optimal for learning (Theorem 8). Second, there exists an ExKrt-numbering that is not a FinKrt-numbering (Theorem 10). Third, there exists a VacKrt-numbering that is not an ExKrt-numbering (Theorem 11). Finally, there exists a BcKrt-numbering that is not a VacKrt-numbering (Theorem 13).

Theorem 8. *There exists a numbering ψ satisfying (a) and (b) below.*

(a) ψ is a FinKrt*-numbering.*
(b) ψ is not optimal for any of the learning criteria Fin*,* Ex*,* Vac*,* Bc*.*

[6] Note that, as ψ is not acceptable, ψ cannot be effectively optimal for Fin-learning [JS10]. Hence, the non-uniform case distinction in this proof is unavoidable.

Proof. The construction of ψ is in two parts. First, we construct a numbering ϑ such that the set $\{e : \vartheta_e \text{ has finite domain}\}$ is dense simple relative to K, i.e., the function that maps n to the n-th index of a function with infinite domain dominates every K-recursive function.[7] From ϑ, we construct ψ.

For each e, s, let $F_{e,s}(\cdot)$ be a uniformly recursive sequence of recursive functions such that, for $F_e(x) = \lim_{s\to\infty} F_{e,s}(x)$,

(a) for all e, x, $F_{e,s}(x) \leq F_{e,s+1}(x)$;
(b) if $\varphi_e^K(x)\downarrow$, then $F_e(x)\downarrow \geq \varphi_e^K(x)$.
(c) if $\varphi_e^K(x)\uparrow$, then $F_e(x)\uparrow$.

Note that such $F_{e,s}$ exist and are uniformly recursive from e, s. Furthermore, each of $F_0, F_1, F_2, \ldots$ is a partial K-recursive function. It should also be noted that, for each e, φ_e^K is majorized by F_e. Hence, the function $n \mapsto \max\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\}$ dominates every K-recursive function.

Let ϑ be such that, for each n, m and x,

$$\vartheta_{\langle n,m\rangle}(x) = \begin{cases} \varphi_n(x), & \text{if } (\exists s > x)\big[\quad (\exists e \leq n)[F_{e,s}(n) = m] \\ & \qquad\qquad \wedge\ (\forall d \leq n)\big[\quad F_{d,s}(n) \leq m \\ & \qquad\qquad\qquad\qquad \vee\ F_{d,s}(n) > F_{d,x}(n)\big]\big]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{17}$$

We show that, for each n and m, $\vartheta_{\langle n,m\rangle}$ has infinite domain iff φ_n has infinite domain and $m = \max\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\}$. To see this, let n and m be given and consider the following four cases.

Case 1: φ_n has finite domain. Clearly, for each n and m, φ_n extends $\vartheta_{\langle n,m\rangle}$. Hence, if φ_n has finite domain, then so does $\vartheta_{\langle n,m\rangle}$.

Case 2: $\{F_e(n) : e \leq n \wedge F_e(n)\downarrow\} = \emptyset$. Let w be so large that, for each $e \leq n$,

$$F_{e,w}(n) > m. \tag{18}$$

Then, for each $x \geq w$, there is no $s > x$ such that $(\exists e \leq n)[F_{e,s}(n) = m]$. Hence, for almost all x, $\vartheta_{\langle n,m\rangle}(x)\uparrow$.

Case 3: $\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\} \neq \emptyset$ and $m \neq \max\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\} < \infty$. Let w be so large that, for each $e \leq n$,

$$F_e(n)\downarrow > m \ \Rightarrow\ F_{e,w}(n) = F_e(n), \tag{19}$$

and

$$F_e(n)\uparrow \ \Rightarrow\ F_{e,w}(n) > m. \tag{20}$$

Then, for each $x \geq w$, there is no $s > x$ such that $(\exists e \leq n)[F_{e,s}(n) = m]$ and $(\forall d \leq n)[F_{d,s}(n) \leq m \ \vee\ F_{d,s}(n) > F_{d,x}(n)]$. Hence, for almost all x, $\vartheta_{\langle n,m\rangle}(x)\uparrow$.

Case 4: φ_n has infinite domain and $m = \max\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\}$. Then, for each x, one can find an $s > x$ such that $(\exists e \leq n)[F_{e,s}(n) = m]$ and, for each $d \leq n$,

$$[F_d(n)\downarrow \ \Rightarrow\ F_{d,s}(n) = F_d(n)] \ \wedge\ [F_d(n)\uparrow \ \Rightarrow\ F_{d,s}(n) > F_{d,x}(n)]. \tag{21}$$

[7] The existence of such numberings is a well-known folklore result.

Hence, for each x, if $\varphi_n(x)\downarrow$, then $\vartheta_{\langle n,m\rangle}(x)\downarrow$.

From Case 4 it also follows that, ϑ is a numbering for $\mathcal{P}$. As the function $\lambda n \centerdot \max\{F_e(n) : e \leq n \ \wedge\ F_e(n)\downarrow\}$ dominates every K-recursive function, the set of all pairs $\langle n,m\rangle$ where $\vartheta_{\langle n,m\rangle}$ has a finite domain is dense simple relative to K.

Now, let ψ be such that, for each p, i and x,

$$\psi_{\langle p,0\rangle}(x) = \vartheta_p(x); \tag{22}$$

$$\psi_{\langle p,i+1\rangle}(x) = \psi_{\langle p,i\rangle}(\langle\langle p,i+1\rangle, x\rangle) \tag{23}$$

Note that ψ is defined such that $\psi_{\langle p,i+1\rangle}$ coincides with the $\langle p,i+1\rangle$-th row of $\psi_{\langle p,i\rangle}$. Hence, ψ is a FinKrt-numbering.

To show that ψ is not optimal for any of the criteria Fin, Ex, Vac, Bc, consider the class $\mathcal{S} = \{f_0, f_1, f_2, \ldots\}$ where, for each n and x, $f_n(x) = n + x$. $\mathcal{S}$ is Fin-learnable and, hence, is also Ex, Vac and Bc-learnable.

Note that, if $\psi_{\langle p,i\rangle} = f_n$, then, by induction over j for all $j > i$, $\mathrm{rng}(\psi_{p,i}) - \mathrm{rng}(\psi_{p,j})$ is infinite and hence $\psi_{p,j} \neq f_m$ for all m. Thus, the following claim holds.

Claim 9. *For each p, there exists at most one i such that $\psi_{\langle p,i\rangle} \in \mathcal{C}$.*

We first show $\mathcal{S}$ is not Vac_ψ-learnable. Now a Vac-learner for $\mathcal{C}$ would, for any n, output only finitely many indices while learning the function f_n. Hence, there exists an index e such that F_e is a K-recursive (i.e., total) function and, for each n, $F_e(n)$ is larger than all the indices output by the learner while learning f_n. It follows that $F_e(n)$ is greater than at least one pair $\langle p,i\rangle$ such that $\psi_{\langle p,i\rangle} = f_n$. Using Claim 9, it follows that ϑ has $n+1$ distinct indices of functions with infinite domain below the value $\max\{F_e(0), F_e(1), \ldots, F_e(n)\}$. But this contradicts the fact that ϑ is a numbering in which the set of indices of partial functions with finite domain is dense simple relative to K. Hence, $\mathcal{C}$ is not Vac_ψ-learnable, and thus neither Fin_ψ nor Ex_ψ-learnable.

Now, assume by way of contradiction that there exists a Bc_ψ-learner M for $\mathcal{C}$. By Claim 9, it follows that, for each p and n, M outputs only finitely many different indices of the form $\langle p,i\rangle$ while learning f_n. Furthermore, by an argument similar to that of the previous paragraph, it can be shown that, for each n, the set $\{p \mid (\exists i \in \mathbb{N})[M \text{ outputs } \langle p,i\rangle \text{ while learning } f_n]\}$ is finite. Hence, the overall number of indices output by the learner while learning an f_n is finite. It follows that M is actually a Vac_ψ-learner for $\mathcal{C}$. But such a learner does not exist as shown in the previous paragraph. □ (**Theorem 8**)

The next result shows, in part, that there exist ExKrt-numberings that are not FinKrt-numberings.

Theorem 10. *There exists a numbering ψ satisfying (a)–(e) below.*

(a) ψ is an Ex-acceptable numbering.
(b) ψ is an ExKrt-numbering.

(c) ψ is not a FinKrt*-numbering.*
(d) ψ is effectively optimal for Ex*-learning.*
(e) ψ is neither optimal for Fin *nor for* Bc*-learning.*

Proof. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } (\exists s > x)(\exists n)[n^2 \leq C_s(e) \leq n^2 + n]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{24}$$

First, we show that the set of indices of partial functions with infinite domain is immune. Let E be any infinite r.e. set of indices and let $e_0, e_1, e_2, \ldots$ be any ascending recursive sequence of elements of E. Then, there exists a constant c such that, for each i, $C(e_{i+1}) < C(e_i) + c$. Let n be so large that $n > C(e_0)$ and $n > c$. As there are only finitely many indices with Kolmogorov complexity below $n^2 + n$, there exists a largest i such that $C(e_i) \leq n^2 + n$. Note that

$$n^2 + n < C(e_{i+1}) < C(e_i) + c < n^2 + 2n < (n+1)^2. \tag{25}$$

It follows that $\psi_{e_{i+1}}$ has a finite domain. Hence, the set of indices of partial functions with infinite domain is immune.

To show that ψ is Ex-acceptable, let e be given. It follows by an argument similar to that of the previous paragraph that there exist n and y such that $n^2 \leq C(\mathrm{pad}(e, y)) \leq n^2 + n$. Furthermore, one can find from e the least such y in the limit. One then has that $\psi_{\mathrm{pad}(e,y)} = \varphi_e$.

To show that ψ is an ExKrt-numbering, let $E_0, E_1, E_2, \ldots$ be a uniformly r.e. family of infinite sets such that, for each p and each $e \in E_p$, $\varphi_e = \psi_p(\langle e, \cdot \rangle)$. One can construct a machine M to witness that ψ is an ExKrt-numbering as follows. Given p, M finds (in the limit) the least $e \in E_p$ for which there exists an n such that $n^2 \leq C(e) \leq n^2 + n$. (The existence of such an e follows by an argument similar to that of the first paragraph.) Then, $\psi_e = \varphi_e = \psi_p(\langle e, \cdot \rangle)$.

To show that ψ is not a FinKrt-numbering, assume by way of contradiction otherwise. Let e_0 be such that $\psi_{e_0} = \lambda x \centerdot x$. Then, enumerate $e_1, e_2, e_3, \ldots$ such that, for each n, $\psi_{e_{n+1}} = \psi_{e_n}(\langle e_{n+1}, \cdot \rangle)$. One can show by induction that, for each n, $\mathrm{rng}(\psi_{e_{n+1}})$ is a proper subset of $\mathrm{rng}(\psi_{e_n})$. Hence, $\{e_0, e_1, e_2, \ldots\}$ is an infinite r.e. set of ψ-indices of total functions. But this would contradict the fact that the set of indices of partial functions with infinite domain is immune.

To show that ψ is optimal for Ex-learning, it was shown above that ψ is Ex-acceptable. It is known that Ex-acceptable numberings are effectively optimal for Ex-learning [JS10].

To show that ψ is not optimal for Fin-learning, consider the class of all constant functions. This class of functions is Fin-learnable. However, if this class could be Fin_ψ-learned, then there would be an infinite r.e. set consisting only of indices of total functions. Again, this would contradict the fact that the set of indices of functions with infinite domain is immune.

In order to see that ψ is not optimal for Bc-learning, it can be shown that every Bc_ψ-learner can be transformed into a Vac_ψ-learner. This proof follows more or less the same argument as that in the proof of Theorem 14 below.

However, as there are Bc-learnable classes of partial functions which are not Vac-learnable, the numbering ψ cannot be optimal for Bc-learning. $\square$ (**Theorem 10**)

Theorem 11. *There exists a numbering ψ satisfying (a)–(d) below.*

(a) ψ is a VacKrt*-numbering.*
(b) ψ is not an ExKrt*-numbering.*
(c) ψ is effectively optimal for Vac*.*
(d) ψ is not optimal for any of the learning criteria Fin*,* Ex*,* Bc*.*

Proof. For this proof, let $(C_s)_{s\in\mathbb{N}}$ be a sequence of uniformly recursive approximations to C^K, such that $C^K(d) = \limsup_{s\to\infty} C_s(d)$. Here we assume that the approximation is such that, for any s and any e, there are at most 2^e many d such that $C_s(d) < e$. Let, for all d, e,

$$\psi_{\langle d,e\rangle}(x) = \begin{cases} \varphi_e(\langle\langle d,e\rangle, x\rangle), & \text{if } [\log(d) \leq e+1] \text{ and } (\exists s > x)[C_s(d) \geq e]; \\ \uparrow, & \text{otherwise.} \end{cases}$$

Let g, h be recursive functions such that, for all p, x, $\varphi_{g(e)}(x) = \psi_e(x)$ and $\varphi_{h(e)}(\langle p, x\rangle) = \varphi_e(x)$. Note that there exist such g, h.

Claim 12. *(i) If $C^K(d) < e$ or $\log(d) > e+1$, then $\psi_{\langle d,e\rangle}$ is a finite function.*
(ii) For all e, for all d such that, $\log(d) \leq e+1$ and $C^K(d) \geq e$, the following holds: $(\forall x)[\psi_{\langle d,e\rangle}(x) = \varphi_e(\langle\langle d,e\rangle, x\rangle)]$.
Here, note that for all e, there exists a d such that $\log(d) \leq e+1$ and $C^K(d) \geq e$.
(iii) For all e, for all d such that, $\log(d) \leq h(e)+1$ and $C^K(d) \geq h(e)$, the following holds: $(\forall x)[\psi_{\langle d,h(e)\rangle}(x) = \varphi_{h(e)}(\langle\langle d, h(e)\rangle, x\rangle) = \varphi_e(x)]$.
(iv) For all e, for all d such that $\log(d) \leq g(e)+1$ and $C^K(d) \geq g(e)$, the following holds: $(\forall x)[\psi_{\langle d,g(e)\rangle}(x) = \varphi_{g(e)}(\langle\langle d, g(e)\rangle, x\rangle) = \psi_e(\langle\langle d, g(e)\rangle, x\rangle)]$.

Parts (i) and (ii) follow immediately from the construction. Parts (iii) and (iv) follow using part (ii) and the definitions of g and h.

By part (ii) of Claim 12 it follows that ψ is a numbering of all the partial recursive functions. We now show the different parts of the theorem.

(a) Let $f(e, s) = \langle d, g(e)\rangle$ such that $\log(d) \leq g(e)+1$ and $C_s(d) \geq g(e)$. By part (iv) of Claim 12, we have that f witnesses that ψ satisfies VacKrt.

(b) Suppose $\langle d_0, e_0\rangle$ is such that, for all x, $\psi_{\langle d_0,e_0\rangle}(x) = x$. Suppose by way of contradiction that H witness ExKrt for ψ, that is, for all i, x, $\psi_{H(i)}(x) = \psi_i(\langle H(i), x\rangle)$. Then for each n, let $\langle d_{n+1}, e_{n+1}\rangle = H(\langle d_n, e_n\rangle)$. Thus,

$$(\forall n, x)\,[\psi_{\langle d_{n+1},e_{n+1}\rangle}(x) = \psi_{\langle d_n,e_n\rangle}(\langle\langle d_{n+1}, e_{n+1}\rangle, x\rangle)]. \tag{26}$$

Now, for all n, ψ_{d_n,e_n} is total. Furthermore, $range(\psi_{d_{n+1},e_{n+1}}) \subset range(\psi_{d_n,e_n})$. Thus, $\langle d_n, e_n\rangle$ are pairwise different for different n. Thus, for each $a \in \mathbb{N}$, one can effectively find an n_a with $d_{n_a} \geq a \wedge e_{n_a} \geq a$. For sufficiently large a, $C^K(d_{n_a}) \leq 2\log(a)$ and $e_{n_a} \geq a$. But then, for sufficiently large a, by Claim 12, $\psi_{\langle d_{n_a},e_{n_a}\rangle}$, would be finite function. A contradiction.

(c) To see that the numbering is effectively optimal for vacillatory learning note that, by Claim 12 and the definition of h, for all e, for all but finitely many s, for the least d such that $\log(d) \leq h(e)+1$ and $C_s(d) \geq h(e)$, we have $\psi_{\langle d,h(e)\rangle}(x) = \varphi_e(x)$. Thus, one can just convert a Vac-learner M using φ as the hypothesis space to a Vac-learner M' using ψ as the hypothesis space by having $M'(f[n]) = \langle d, h(M(f[n])\rangle$, where d is least such that $\log(d) \leq h(M(f[n]) + 1$ and $C_n(d) \geq h(M(f[n])$.

(d) Let $\mathcal{S}$ be a class of total functions which is Bc-learnable by some learner M using ψ as the hypothesis space. For any total f, let $E_f = \{M(f[n]) : n \in \mathbb{N}\}$. We claim that E_f is finite for each $f \in \mathcal{S}$. Suppose by way of contradiction that for some $f \in \mathcal{S}$, E_f is infinite. Note that E_f is an r.e. set. Let $\eta(e) = d_e$, for the first pair $\langle d_e, e\rangle$ enumerated in E_f, if any. Now, $\eta(e)$ is defined on infinitely many e, and thus $C^K(d_e) \leq 2\log(e)$ for infinitely many e in the domain of η. But then, by Claim 12, $\psi_{\langle d_e,e\rangle}$ is a finite function for infinitely many e in the domain of η, a contradiction to M Bc-learning f. Thus, M is also a Vac-learner for $\mathcal{S}$. As there are classes of total functions which are Bc-learnable but not Vac-learnable [CS83], ψ is not optimal for Bc-learning.

Now, suppose by way of contradiction that M Ex-learns all constant functions using the numbering ψ. Thus, for each a, there exists a constant c such that, for some d_a, e_a, for all but finitely many n, $M(c^\infty[n]) = \langle d_a, e_a\rangle$, with $\min\{d_a, e_a\} \geq a$. Note that one such pair of values d_a, e_a can be computed using the oracle K. Then, for almost all a, $C^K(d_a) \leq 2\log(a)$ and $e_a \geq a$. Hence, by Claim 12, for all but finitely many a, $\psi_{\langle d_a,e_a\rangle}$ is a finite function. Thus, M does not Ex-learn the class of all constant functions using the numbering ψ. It follows that ψ is not optimal for Fin and Ex-learning. □ (**Theorem 12**)

The final result of this section establishes, in part, that there exist BcKrt-numberings that are not VacKrt-numberings.

Theorem 13. *There exists a numbering ψ satisfying (a)–(d) below.*

(a) ψ is a BcKrt*-numbering.*
(b) ψ is not a VacKrt*-numbering.*
(c) ψ is not optimal for any of the learning criteria Fin*,* Ex*,* Vac*.*
(d) ψ is effectively optimal for Bc*-learning.*

Proof. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } [\quad \varphi_{e,x}(0)\uparrow \\ & \quad\vee\ |\mathrm{rng}(\varphi_e)| \geq 2 \\ & \quad\vee\ [\varphi_e(0)\downarrow\ \wedge\ C(\varphi_e(0)) < \log(\varphi_e(0))] \\ & \quad\vee\ [\varphi_e(0)\downarrow\ \wedge\ |W_{\log(\varphi_e(0)),x}| < e] \\ & \quad\vee\ [\varphi_e(0)\downarrow\ \wedge\ |W_{\log(\varphi_e(0))}| > x] \qquad]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{27}$$

To show that ψ is a BcKrt-numbering, let $E_0, E_1, E_2, \ldots$ be a uniformly r.e. family of infinite sets such that, for each p and each $e \in E_p$, $\varphi_e = \psi_p(\langle e, \cdot\rangle)$. One can construct a machine M to witness that ψ is a BcKrt-numbering as follows.

Suppose that M is given p. Then, at stage s, M outputs the first element e in some canonical enumeration of E_p such that

$$\begin{aligned} & \varphi_{e,s}(0)\uparrow \\ \vee\ & |\mathrm{rng}(\varphi_{e,s})| \geq 2 \\ \vee\ & [\varphi_{e,s}(0)\downarrow\ \wedge\ C_s(\varphi_{e,s}(0)) < \log(\varphi_{e,s}(0))] \\ \vee\ & e > s. \end{aligned} \tag{28}$$

Consider the following two cases.

Case 1: There exists an $e \in E_p$ such that $\varphi_e(0)\uparrow$, $|\mathrm{rng}(\varphi_e)| \geq 2$ or $\varphi_e(0) = y$ for some y with $C(y) < \log(y)$. Then, M converges to the first such e in the canonical enumeration of E_p. Furthermore, for this e, it holds that $\psi_e = \varphi_e = \psi_p(\langle e, \cdot\rangle)$.

Case 2: Not Case 1. Then, the set $F = \{\varphi_e(0) : M \text{ outputs } e\}$ has an empty intersection with the simple set $\{d : C(d) < \log(d)\}$ and, hence, is finite. Let $c = \max\{|W_{\log(d)}| : d \in F\ \wedge\ |W_{\log(d)}| < \infty\}$. As F is finite, this maximum c is taken over only finitely many numbers and, hence, $c < \infty$. Furthermore, as Case 1 does not apply, M outputs each index in E_p only finitely often. Hence, M outputs almost always some index $e > c$. If, for such an e, $W_{\log(\varphi_e(0))}$ is finite, then, for each x, $|W_{\log(\varphi_e(0)),x}| \leq c < e$. On the other hand, if $W_{\log(\varphi_e(0))}$ is infinite, then, for each x, $|W_{\log(\varphi_e(0))}| > x$. Either way, M outputs almost always an e such that $\psi_e = \psi_p(\langle e, \cdot\rangle)$.

It follows from the case distinction that M witnesses that ψ is a BcKrt-numbering.

To show that ψ is not a VacKrt-numbering, assume by way of contradiction otherwise, as witnessed by M. We show that, under this assumption, one can decide membership in $\{x \mid W_x \text{ is finite}\}$ using an oracle for K (which is impossible). It is known that, for almost all x, there exist distinct y and z such that $\log(y) = \log(z) = x$, but $C(y) \geq x$ and $C(z) \geq x$.[8] Given x, one can find such y and z using an oracle for K. One can then determine p such that, for each v and w,

$$\varphi_p(\langle v, w\rangle) = \begin{cases} y, & \text{if } v \text{ is even;} \\ z, & \text{if } v \text{ is odd.} \end{cases} \tag{29}$$

Note that $|\mathrm{rng}(\varphi_p)| = 2$ and, hence, $\psi_p = \varphi_p$. One can then run M on input p and, using the oracle for K, determine the largest e among the finitely many indices output by M. Hence, for some $v < e$, ψ_v is either the constantly y function, or the constantly z function. It follows that either $|W_{\log(\varphi_v(0))}| < v$ or $|W_{\log(\varphi_v(0))}|$ is infinite. If the former, then

$$|W_x| = |W_{\log(\varphi_v(0))}| < v < e. \tag{30}$$

[8] Recall from Section 2 that, for each y, $\log(y) \stackrel{\text{def}}{=} \min\{x \mid 2^x \geq y\}$. For each $x \geq 1$, there are 2^{x-1} many numbers y with $\log(y) = x$ and only $2^{x-1} + 1$ many numbers y with $C(y) < x$. Furthermore, for sufficiently large x, there will exist three or more programs less than $2^{x-1} + 1$ that either produce no output, or produce the same output as programs less than themselves. Hence, for sufficiently large x, such y and z exist.

If the latter, then

$$|W_x| = |W_{\log(\varphi_v(0))}| \geq e. \tag{31}$$

Hence, W_x is finite iff $|W_x| < e$. As $|W_x| < e$ can be decided using an oracle for K, this allows one to determine whether W_x is finite. Since this is impossible, it follows that M does not witness that ψ is a VacKrt-numbering and, more generally, that ψ is not a VacKrt-numbering.

To show that ψ is not optimal for any of the learning criteria Fin, Ex, Vac, note that the class of constant functions is Fin-learnable. But it can be shown that this class is neither Fin_ψ, Ex_ψ nor Vac_ψ-learnable using a proof-idea similar to that of the previous paragraph. Assume by way of contradiction otherwise, as witnessed by M. Then, given x, one can use an oracle for K to find a y such that $\log(y) = x$ and $C(y) \geq x$. Then, when M is fed a text for the constantly y function, M outputs finitely many indices whose maximum is some e. Using this e and the oracle for K, one can determine whether W_x is finite as in the previous paragraph (a contradiction). Hence, ψ is not optimal for any of the learning criteria Fin, Ex, Vac.

To show that ψ is effectively optimal for Bc-learning, suppose that M is a Bc-learner that uses φ as its hypothesis space. Further suppose that M is fed a text for a partial recursive function α and that $e_0, e_1, e_2, \ldots$ is the sequence of indices output by M on this text. Without loss of generality, suppose that this sequence is monotonically increasing, e.g., due to padding. We show that, for almost all i, $\psi_{e_i} = \varphi_{e_i}$. Consider the following three cases.

Case 1: $\alpha(0)\uparrow$. Then, for almost all i, $\varphi_{e_i}(0)\uparrow$ and, hence, $\psi_{e_i} = \varphi_{e_i}$.

Case 2: $\alpha(0)\downarrow$ and $|W_{\log(\alpha(0))}|$ is infinite. Then, for almost all i, $|W_{\log(\varphi_{e_i}(0))}|$ is infinite and, hence, $\psi_{e_i} = \varphi_{e_i}$.

Case 3: $\alpha(0)\downarrow$ and $|W_{\log(\alpha(0))}|$ is finite. Then, as $e_0, e_1, e_2, \ldots$ is monotonically increasing, for almost all i, $|W_{\log(\varphi_{e_i}(0))}| < e_i$. Hence, for almost all i, $\psi_{e_i} = \varphi_{e_i}$.

This case distinction shows that Bc_φ-learners that output successively larger indices are also Bc_ψ-learners. Hence, the numbering ψ is effectively optimal for Bc-learning. $\square$ (**Theorem 13**)

5 When Kleene's Recursion Theorem is Ineffective

This section presents our results concerning numberings that satisfy Kleene's Recursion Theorem, but not in an effective way.

Moelius [Moe09, Theorem 4.1] showed that there exist numberings that are not BcKrt-numberings, but in which Kleene's Recursion Theorem holds. Hence, in such numberings, Kleene's Recursion Theorem is extremely ineffective. Theorems 15 and 16 expand on Moelius's result by showing that there exist such numberings that are optimal for learning and such numberings that are not optimal for learning (respectively). Theorems 15 and 16 make use of Theorem 14 just below.

Theorem 14. *Suppose that ψ is a BcKrt-numbering and that the set of indices of partial functions with infinite domain is immune. Then, ψ is a VacKrt-numbering.*

Proof. Suppose that M witnesses that ψ is a BcKrt-numbering. From M, we construct a machine N witnessing that ψ is a VacKrt-numbering. Suppose that $e_0, e_1, e_2, \ldots$ is the sequence of indices output by M on input p. Then, at stage s, N outputs index e_j for the least j such that

$$(\forall x \leq s)[\psi_{e_j,s}(x) = \psi_{p,s}(\langle e_j, x\rangle)] \ \vee \ j = s. \tag{32}$$

First, consider the case that there exists a j such that $\psi_{e_j} = \psi_p(\langle e_j, \cdot\rangle)$ and ψ_{e_j} is a finite function. Let s be such that

$$\begin{aligned} &(\forall i < j)\big[\psi_{e_i} \neq \psi_p(\langle e_i, \cdot\rangle) \ \Rightarrow \ (\exists x \leq s)[\psi_{e_i,s}(x) \neq \psi_{p,s}(\langle e_i, x\rangle)]\big] \\ &\quad \wedge \ \psi_{e_j,s} = \psi_{e_j}. \end{aligned} \tag{33}$$

Note that, from stage s onward, N will only ever output indices from among $e_0, e_1, \ldots, e_j$. Hence, N will vacillate among only finitely many indices, as required. Furthermore, each such index e_i output by N will satisfy $\psi_{e_i} = \psi_p(\langle e_i, \cdot\rangle)$.

Next, consider the case that, for each i, if $\psi_{e_i} = \psi_p(\langle e_i, \cdot\rangle)$, then ψ_{e_i} has infinite domain. Then, it follows from the immunity assumption that $\{e_0, e_1, e_2, \ldots\}$ is a finite set, i.e., there exists some j such that $e_0, e_1, \ldots, e_j$ represents *all* of the indices output by M on input p. Let s be as in the first conjunct of (33). Then, from stage s onward, each index e_i output by N will satisfy $\psi_{e_i} = \psi_p(\langle e_i, \cdot\rangle)$. Furthermore, N will again vacillate among only finitely many indices, as required.

It follows from the analysis of these two cases that N witnesses that ψ is a VacKrt-numbering. $\square$ (**Theorem 14**)

Theorem 15. *There exists a numbering ψ satisfying (a)–(c) below.*

(a) ψ satisfies Kleene's Recursion Theorem.
(b) ψ is not a BcKrt-numbering.
(c) ψ is not optimal for any of the learning criteria Fin, Ex, Vac, Bc.

Proof. Let ψ' be the numbering called "ψ" in the proof of Theorem 10, and let ψ'' be the numbering called "ψ" in the proof of Theorem 13. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } \psi'(x)\downarrow \text{ and } \psi''(x)\downarrow; \\ \uparrow, & \text{if } \psi'(x)\uparrow \text{ or } \psi''(x)\uparrow. \end{cases} \tag{34}$$

To show that ψ satisfies Kleene's Recursion Theorem, let α be a given partial recursive function, and let E be an infinite r.e. set such that, for each $e \in E$, $\varphi_e = \alpha(\langle e, \cdot\rangle)$. Consider the following three cases.

Case 1: There exists a y such that, for infinitely many $e \in E$, $\varphi_e(0) = y$. Let

$$F = \begin{cases} \{e \in E : \varphi_e(0) = y\}, & \text{if } |W_{\log(y)}| = \infty; \\ \{e \in E : \varphi_e(0) = y \ \wedge \ e > |W_{\log(y)}|\}, & \text{if } |W_{\log(y)}| < \infty. \end{cases} \tag{35}$$

Note that, for each $e \in F$, $\psi''_e = \varphi_e$. Further note that F is r.e. and infinite. Hence, there exists an ascending recursive sequence, $e_0, e_1, e_2, \ldots$ of elements of F. It follows that there exists a constant c such that, for each i,

$C(e_{i+1}) < C(e_i) + c$. Let n be so large that $n > e_0$ and $n > c$. Let i be largest such that $C(e_i) < n^2$. Then, $n^2 \leq C(e_{i+1}) \leq n^2 + n$ and, hence, $\psi'_{e_{i+1}} = \varphi_{e_{i+1}}$. It follows that $\psi_{e_{i+1}} = \varphi_{e_{i+1}} = \alpha(\langle e_{i+1}, \cdot\rangle)$.

Case 2: The set $\{y \mid e \in E \wedge \varphi_e(0) = y\}$ is infinite. Then, there are infinitely many y in this set such that $C(y) < \log(y)$. Hence, there exists an ascending recursive sequence $e_0, e_1, e_2, \ldots$ of elements of E such that, for each i, $\varphi_{e_i}(0)\downarrow$, $\varphi_{e_i}(0) < \varphi_{e_{i+1}}(0)$, and $C(\varphi_{e_i}(0)) < \log(\varphi_{e_i}(0))$. Note that, for each i, $\psi''_{e_i} = \varphi_{e_i}$. Let n be so large that $C(e_0) < n$ and, for each i, $C(e_{i+1}) < C(e_i)+n$. Then, there exists a largest i such that $C(e_i) < n^2$. It follows that $n^2 \leq C(e_{i+1}) \leq n^2 + n$ and $\psi'_{e_{i+1}} = \varphi_{e_{i+1}}$. Hence, $\psi_{e_{i+1}} = \varphi_{e_{i+1}} = \alpha(\langle e_{i+1}, \cdot\rangle)$.

Case 3: Not Cases 1 and 2. Then, the set $\{e \in E \mid \varphi_e(0)\downarrow\}$ is finite. It follows that $F = \{e \in E : \varphi_e(0)\uparrow\}$ is r.e. and infinite. Furthermore, for each $e \in E$, $\psi''_e = \varphi_e$. Now, one can show as in Case 1 that there exists an $e' \in F$ such that $\psi'_{e'} = \varphi_{e'}$ and, hence, $\psi_{e'} = \alpha(\langle e', \cdot\rangle)$.

This completes the case distinction to show that ψ satisfies Kleene's Recursion Theorem.

To show that ψ is not a BcKrt-numbering, assume by way of contradiction otherwise. Then, as the indices of partial functions with infinite domain form an immune set in the numbering ψ', they also form an immune set in the numbering ψ. In particular, ψ is then also a VacKrt-numbering by Theorem 14. So, suppose that M witnesses that ψ is a VacKrt-numbering. Then, one can a arrive at a contradiction in much the same way as in the proof of Theorem 13 — the only difference is that an additional side-condition on the choice of p is needed. In particular, using an oracle for K, one finds a p as in (29), satisfying: there exists an n such that $n^2 \leq C(p) \leq n^2+n$. This condition on p establishes that $\psi'_p = \varphi_p$, and, hence, that $\psi_p = \varphi_p$. Then, as in the proof Theorem 13, one runs M on this p. Finally, using the oracle for K, one determines, for the given x, whether W_x is finite (a contradiction). It follows that M does not witness that ψ is a VacKrt-numbering, and, more generally, that ψ is neither a VacKrt-numbering nor a BcKrt-numbering.

To show that ψ is not optimal for any of the learning criteria Fin, Ex, Vac, Bc, note that, for each p, if ψ_p is a constant function, then so is ψ''_p. Hence, if the class of constant functions were Fin_ψ, Ex_ψ, Vac_ψ-learnable, then it would be $\mathsf{Fin}_{\psi''}$, $\mathsf{Ex}_{\psi''}$, or $\mathsf{Vac}_{\psi''}$-learnable (respectively), as well. However, as shown in the proof of Theorem 13, this class is not learnable in any of these senses. Furthermore, as the set of indices of partial functions with infinite domain is immune, every Bc_ψ-learnable class is also Vac_ψ-learnable. Hence, the class of constant functions is also not Bc_ψ-learnable. This shows that ψ is not optimal for any of the learning criteria Fin, Ex, Vac, Bc. □ (**Theorem 15**)

Theorem 16. *There exists a numbering ψ satisfying (a)–(c) below.*

(a) ψ satisfies Kleene's Recursion Theorem.
(b) ψ is not a BcKrt*-numbering.*
(c) ψ is effectively optimal for Ex, Vac *and* Bc*-learning.*

Proof. Let S be a set which is hypersimple relative to K.[9] Let $\lambda t \centerdot S_t$ be a recursive approximation of S in that, for each n, $n \in S \Leftrightarrow (\forall^\infty t)[n \in S_t]$. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } \Big[\quad \varphi_{e,x}(0)\uparrow \\ & \qquad \vee\, \varphi_e(0)\downarrow < e \\ & \qquad \vee\, \big[\varphi_e(0)\downarrow \wedge (\exists n)(\exists t > x)[\quad n \notin S_t \wedge \\ & \qquad\qquad\qquad n^2 \le C_t(\varphi_e(0)) \le n^2 + n]\big]\Big]; \\ \uparrow, & \text{otherwise.} \end{cases}$$

To show that ψ satisfies Kleene's Recursion Theorem, let α be a given partial recursive function, and let E be an infinite r.e. set such that, for each $e \in E$, $\varphi_e = \alpha(\langle e, \cdot \rangle)$. We show that at least one $e \in E$ satisfies $\psi_e = \varphi_e$. If there exists an $e \in E$ such that $\varphi_e(0)\uparrow$ or $\varphi_e(0)\downarrow < e$, then this is immediate. So, suppose that, for each $e \in E$, $\varphi_e(0)\downarrow \ge e$. Then, there exists an ascending recursive sequence $e_0, e_1, e_2, \ldots$ of elements of E such that, for each i, $\varphi_{e_i}(0) < \varphi_{e_{i+1}}(0)$. It follows that there exists a constant c such that, for each i, $C(\varphi_{e_{i+1}}(0)) < C(\varphi_{e_i}(0)) + c$. Let n be so large that $n > C(\varphi_{e_0}(0))$, $n > c$, and $n \notin S$. Let i be largest such that $C(\varphi_{e_i}(0)) \le n^2$. Then, $n^2 \le C(\varphi_{e_{i+1}}(0)) \le n^2 + n$. Furthermore, as $n \notin S$, it follows that $\psi_{e_{i+1}} = \varphi_{e_{i+1}}$. This completes the proof that ψ satisfies Kleene's Recursion Theorem.

To show that ψ is not a BcKrt-numbering, assume by way of contradiction otherwise, as witnessed by M. Let $p_0, p_1, p_2, \ldots$ be such that, for each k, $p_k > k$, and, for each e and x, $\varphi_{p_k}(\langle e, x\rangle) = e + k$. Note that, for each k, $\psi_{p_k} = \varphi_{p_k}$, as $\varphi_{p_k}(0) = \varphi_{p_k}(\langle 0, 0\rangle) = k < p_k$. Now, let k be fixed, and let $e_0, e_1, e_2, \ldots$ be the sequence of indices output by M on input p_k. Then, for almost all i, $\psi_{e_i} = \lambda x \centerdot e_i + k$. Furthermore, it follows from the definition of ψ that, for almost all i, there exists an $n_i \notin S$ such that $n_i^2 \le C(e_i + k) \le n_i^2 + n_i$. Note that, as S is simple relative to K, $\{e_0, e_1, e_2, \ldots\}$ is a finite set, i.e., there exists some j such that $e_0, e_1, \ldots, e_j$ represents *all* of the indices output by M on input p_k. Let $f : \mathbb{N} \to \mathbb{N}$ be such that, for each k, $D_{f(k)}$ is the collection of all n_i corresponding to $e_0, e_1, \ldots, e_j$, where n_i and $e_0, e_1, \ldots, e_j$ are as just described. Note that f is K-recursive. Further note that, as only finitely many numbers have the same Kolmogorov complexity, for each n, there exists some k such that each element of $D_{f(k)}$ is at least n. Let $g : \mathbb{N} \to \mathbb{N}$ be such that, for each i,

$$g(0) = f(0); \tag{36}$$

$$g(i+1) = f(k), \text{ where } k \text{ is least such that } D_{f(k)} \cap (D_{g(0)} \cup D_{g(1)} \cup \cdots \cup D_{g(i)}) = \emptyset. \tag{37}$$

Note that: g is K-recursive; for each i and j, $D_{g(i)}$ and $D_{g(j)}$ are disjoint; and, for each i, $D_{g(i)}$ is not a subset of S. This contradicts the fact that S is hypersimple

[9] For example, one could take S to be $\{e : \vartheta_e \text{ has finite domain}\}$ in the proof of Theorem 8, as every set that is dense simple relative to K is also hypersimple relative to K.

relative to K. Hence, M does not witness that ψ is BcKrt-numbering, and, more generally, ψ is not a BcKrt-numbering.

To show that ψ is effectively optimal for Ex, Vac and Bc-learning, suppose that M is such a learner, and suppose that α is some target partial recursive function. Then, so long as $\alpha(0)$ has not yet been seen, the new learner N simulates M, and outputs whatever M would output. If, however, $\alpha(0)$ is ever seen, then, from that point onward, N simulates M and translates each conjecture e of M into $\mathrm{pad}(e, \alpha(0))$. Note that, if $\varphi_e = \alpha$ and $\alpha(0)\uparrow$, then $\psi_e = \varphi_e$. On the other hand, if $\varphi_e = \alpha$ and $\alpha(0)\downarrow$, then $\psi_{\mathrm{pad}(e,\alpha(0))} = \varphi_e$, as $\alpha(0) < \mathrm{pad}(e, \alpha(0))$. It is easy to see that this translation preserves Ex, Vac and Bc-convergence. Hence, ψ is effectively optimal for Ex, Vac and Bc-learning. $\square$ (**Theorem 16**)

Acknowledgment. The authors would like to thank Samuel E. Moelius III for discussions and support in writing this paper.

References

[Bar74] Janis Bārzdiņš. Two theorems on the limiting synthesis of functions. *Theory of Algorithms and Programs*, Volume 1, pages 82–88, LSU, Riga, Latvia, 1974.

[Cas99] John Case. The power of vacillation in language learning. *SIAM Journal on Computing*, 28:1941–1969, 1999.

[CJS02] John Case, Sanjay Jain and Mandayam Suraj. Control structures in hypothesis spaces: The influence on learning. *Theoretical Computer Science*, 270(1-2):287–308, 2002.

[CM09] John Case and Samuel E. Moelius III. Program self-reference in constructive Scott subdomains. In *Proceedings of Computability in Europe 2009 (CiE'09)*, volume 5635 of *Lecture Notes in Computer Science*, pages 89–98. Springer, 2009. To appear in Theory of Computing Systems.

[CS83] John Case and Carl H. Smith. Comparison of identification criteria for machine inductive inferenc e. *Theoretical Computer Science*, 25:193–220, 1983.

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

[Jai11] Sanjay Jain. Hypothesis spaces for learning. *Information and Computation*, 209(3):513–527, 2011.

[JS10] Sanjay Jain and Frank Stephan. Numberings optimal for learning. *Journal of Computer and System Sciences*, 76(3-4):233–250, 2010.

[Kle38] Stephen C. Kleene. On notation for ordinal numbers. *Journal of Symbolic Logic*, 3(4):150–155, 1938.

[LV08] Ming Li and Paul M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Texts in Computer Science. Springer, third edition, 2008.

[Moe09] Samuel E. Moelius III. *Program Self-Reference.* PhD thesis, University of Delaware, 2009.

[Nie09] André Nies. *Computability and Randomness.* Oxford Logic Guides. Oxford University Press, 2009.

[OW82] Daniel N. Osherson and Scott Weinstein. Criteria of language learning. *Information and Control*, 52:123–138, 1982.

[Ric80] Gregory A. Riccardi. *The Independence of Control Structures in Abstract Programming Systems.* PhD thesis, SUNY Buffalo, 1980.

[Ric81] Gregory A. Riccardi. The independence of control structures in abstract programming systems. *Journal of Computer and System Sciences*, 22(2):107–143, 1981.

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

[Roy87] James S. Royer. *A Connotational Theory of Program Structure*, volume 273 of *Lecture Notes in Computer Science*. Springer, 1987.