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 investigates the quality of numberings measured in three different ways: (a) the complexity of finding witnesses of Kleene's Recursion Theorem in the numbering; (b) for which learning notions from inductive inference the numbering is an optimal hypothesis space; (c) the complexity needed to translate the indices of other numberings to those of the given one. In all three cases, one assumes that the corresponding witnesses or correct hypotheses are found in the limit and one measures the complexity with respect to the best criterion of convergence which can be achieved. The convergence criteria considered are those of finite, explanatory, vacillatory and behaviourally correct convergence. The main finding is that the complexity of finding witnesses for Kleene's Recursion Theorem and the optimality for learning are independent of each other. Furthermore, if the numbering is optimal for explanatory learning and also allows to solve Kleene's Recursion Theorem with respect to explanatory convergence, then it also allows to translate indices of other numberings with respect to explanatory convergence.

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

1 Introduction

In constructive mathematics, one not only asks whether a certain object exists but also how complex it is to find or construct this object. Hayashi [Hay06] studied this question of finding or constructing objects as a limiting process where the solution (or programs to compute the solution) are approximated in the limit in the same way as inductive inference approximates the program of a concept to be learnt. In other words, Hayashi [Hay06] carried over the convergence

⋆ 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.

criteria from inductive inference and applied them to finding solutions to specific mathematical tasks which cannot be carried out directly but need a limiting process for finding the solution.

Case and Moelius [CM09,Moe09] applied this idea to the task of finding fixed points in numberings. From the beginnings of recursion theory, it is known that acceptable numberings of all partial recursive functions admit fixed-points, that is, for every recursive function f there is an e with $\varphi_{f(e)} = \varphi_e$ (where φ_e denotes the e-th function in the acceptable numbering φ). This means in particular that one cannot modify the behaviour of programs systematically; all methods to modify the program of functions fail to have an effect on some program e [Rog67]. In their work on this topic, Case and Moelius [CM09] noted that Kleene's original version [Kle52] of the fixed point theorem – known as Kleene's Recursion Theorem (KRT) – is in many applications much more useful than Rogers' variant. Informally, Kleene's Recursion Theorem says that for every partial recursive function η with two inputs there is an index e such that, for all x, $\varphi_e(x) = \eta(e, x)$. Moelius [Moe09] studied in detail numberings ψ which satisfy KRT and, if they do, how complex it is to obtain the fixed point e as above. The above mentioned work of Moelius motivated choosing KRT rather than the popular fixed point theorem as the basic notion for the present paper.

A numbering is a uniformly partially recursive listing of all partial recursive functions with one input. It depends on the numbering ψ whether it satisfies KRT and, in the case that it does, how complex is the process to find, for any given f, the fixed point e with $\psi_e = f(e, \cdot)$. The complexity of this process is measured in terms of the convergence criterion (finite, explanatory, vacillatory or behaviourally correct, see formal definitions in Section 2) used to find such a fixed point for all recursive functions f. These properties permit to classify numberings as those which obey FinKrt, ExKrt, VacKrt and BcKrt, respectively. Similarly one can measure the numberings ψ with respect to the question whether they are optimal for the corresponding learning criteria I; here ψ is optimal for I-learning iff every I-learnable class can be learnt using the numbering ψ as hypothesis space. A third way to classify numberings is to say that a numbering ψ is I-acceptable iff for every further numbering φ, one can translate each index e of φ into a sequence of ψ-indices which converge to ψ-indices for φ_e in the way prescribed by I.

Building on prior work of Case and Moelius [CM09,Moe09], the present work relates three ways to categorise quality of numberings (degree of effectiveness of KRT, degree of optimality of the numbering for learning, degree of acceptability of the numbering) with each other. The results show that in general, one can find arbitrary combinations of the degrees of the effectiveness of KRT and the degrees of optimality of the numbering. Furthermore, for I being finite, explanatory or vacillatory convergence, numberings which are optimal for I learning and permit I-effective KRT are also I-acceptable. This means, that KRT and optimality are complementary independent properties which together give that the numbering is acceptable. This is now described more formally and precisely in the remaining part of the paper.

2 Technical Definitions and Overview

Recursion-theoretic concepts not covered below are treated as by Rogers [Rog67]; the interested reader might also consult the books of Calude [Cal02], Downey and Hirschfeldt [DH10], Nies

[Nie09], Odifreddi [Odi89,Odi99], Li and Vitányi [LV08] and Soare [Soa87] for further background on recursion theory and Kolmogorov complexity. Furthermore, the interested reader is referred to the two editions of "Systems that Learn" [OSW86,JORS99] which both emphasise on quite different aspects of inductive inference.

Let $\mathbb{N}$ be the set of natural numbers, $\{0, 1, 2, \ldots\}$. Lowercase italicised letters range either over elements of $\mathbb{N}$ (such as a, b, c, d, e, i, j, k, m, n, x, y, z) or over functions (such as f, g, h); it is always clear from the context which of the two cases applies. Uppercase italicised letters (such as A, B, D, E, K) range over subsets of $\mathbb{N}$, unless stated otherwise. Lowercase Greek letters (such as α, β, γ, ψ, φ, ϕ, θ) range over partial functions from $\mathbb{N}$ to $\mathbb{N}$, unless stated otherwise.

For each non-empty $X \subseteq \mathbb{N}$, $\min X$ denotes the minimum element of X, where $\min \emptyset \stackrel{\text{def}}{=} \infty$. For each non-empty, finite $X \subseteq \mathbb{N}$, $\max X$ denotes the maximum element of X; furthermore, $\max \emptyset \stackrel{\text{def}}{=} -1$. The finite sets can be indexed by canonical indices where the e-th finite set D_e is chosen such that $D_0 = \emptyset$ and, for $e > 0$, D_e is the unique set satisfying $\sum_{d \in D_e} 2^d = e$.

Let $\langle \cdot, \cdot \rangle$ be Cantor's pairing function: $\langle x, y \rangle = (x+y)(x+y+1)/2 + y$. The pairing function is a recursive, order preserving bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ [Rog67, page 64]; here order preserving means that $x \leq x' \wedge y \leq y' \Rightarrow \langle x, y \rangle \leq \langle x', y' \rangle$. 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. So, for example, $\lambda x \centerdot {\uparrow}$ denotes the everywhere divergent partial function.

For each partial function f, $\mathrm{rng}(f)$ denotes the range of f. A *text* is a total (not necessarily recursive) function from $\mathbb{N}$ to $\mathbb{N} \cup \{\#\}$. For each text T and $i \in \mathbb{N}$, $T[i]$ denotes the initial segment of T of length i: $T(0)T(1)\ldots T(i-1)$. SEQ 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 $\mathrm{rng}(T) - \{\#\}$ is the graph of α as coded using the pairing function $\langle \cdot, \cdot \rangle$:

$$\mathrm{rng}(T) - \{\#\} = \{\langle x, y \rangle : \alpha(x) = y \ \wedge \ x, y \in \mathbb{N}\}. \tag{1}$$

For a total function f, one often identifies 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.

Let $\mathcal{S}_{const} = \{f : (\exists c)(\forall x)[f(x) = c]\}$ denote the class of all constant functions.

Let $\mathcal{P}$ be the collection of all partial recursive functions from $\mathbb{N}$ to $\mathbb{N}$. 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

$$(\forall \alpha \in \mathcal{P})(\exists p \in \mathbb{N})[\psi_p = \alpha]. \tag{2}$$

The present work only deals with effective numberings of partial recursive functions; hence, for notational simplicity, the requirements "$(\forall \alpha \in \mathcal{P})(\exists p \in \mathbb{N})[\psi_p = \alpha]$" and that the "numbering is effective" are usually assumed without being explicitly postulated. For all ψ and p, s, if $x < s$ and $\psi_p(x)$ halts within s steps with output y, then let $\psi_{p,s}(x) = y$; else let $\psi_{p,s}(x) = \uparrow$. Note that while it is in general undecidable whether $\psi_p(x)$ is defined, it is decidable whether $\psi_{p,s}(x)$ is defined and what value in $\mathbb{N} \cup \{\uparrow\}$ it takes. So there is an asymmetry between the usage of $\uparrow$ in the case of the approximation $\psi_{p,s}$ and in the case of the partial recursive function ψ_p. Let INF_ψ

denote the set $\{i : \psi_i \text{ has infinite domain}\}$ of indices of functions which are defined for infinitely many inputs. Furthermore, given a numbering ψ of partial-recursive functions, one defines W_p^ψ to be the domain of the function ψ_p in the numbering.

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 $\varphi_0 = \lambda x.\uparrow$. For each p, $W_p \stackrel{\text{def}}{=} \{x : \varphi_p(x)\downarrow\}$. K denotes the diagonal halting problem with respect to φ, that is, $\{x : x \in W_x\}$. Let $\text{pad} : \mathbb{N}^2 \to \mathbb{N}$ be a recursive one-one function such that, for each e and y, $\text{pad}(0, 0) = 0$, $\text{pad}(e, y) < \text{pad}(e, y + 1)$, $\text{pad}(e, y) < \text{pad}(e{+}1, y)$ and $\varphi_{\text{pad}(e,y)} = \varphi_e$. Sometimes computations using oracles are considered. Let φ_e^A denote the partial function computed by the e-th program using oracle A in some fixed A-acceptable numbering φ^A.

The following notions of immune and simple sets can be found in most recursion theory books such as [Rog67,Soa87,Odi89]. A set A is *immune* if it is infinite and there exists no infinite r.e. subset of A. A set A is *hyperimmune* if it is infinite and there does not exist a recursive function f such that (a) for all i, $D_{f(i)} \cap A \neq \emptyset$ and (b) $D_{f(i)}$ are pairwise disjoint for different i. A set A is *simple* if A is r.e. and $\overline{A}$ is immune. A set A is *hypersimple* if A is r.e. and $\overline{A}$ is hyperimmune. A function f *dominates* g if $f(x) \geq g(x)$ for all but finitely many x.

Consider log base 2, rounded up to the nearest natural number; more precisely, for each y, $\log(y) \stackrel{\text{def}}{=} \min\{x \in \mathbb{N} : 2^x \geq y\}$. So, for example, $\log(0) = 0$ and $\log(3) = 2$. For any Turing machine M and any e, let $C_M(e)$ denote the length of the least y such that $M(y) = e$. Intuitively, $C_M(e)$ is the Kolmogorov complexity of e with respect to machine M. There is a Turing Machine U such that, for any machine M, for some constant c, for all e, $C_U(e) \leq C_M(e) + c$; such a Turing machine U is called a *Universal Turing Machine*. As the definition $C_U(e)$ depends on the Universal Turing Machine only up to a constant, one can fix such a Universal Turing Machine U and let $C(e)$ denote $C_U(e)$ [LV08,Nie09]. $C(e)$ is also referred to as the plain Kolmogorov complexity of e. For ease of writing some proofs, one assumes without loss of generality, that the Universal Turing Machine diverges on inputs of length 0 or 1, and thus for all $x \geq 1$, there are distinct numbers y and z such that $\log(y) = \log(z) = x$ and $C(y) \geq x$ and $C(z) \geq x$. $C^K(e)$ denotes the notion of plain Kolmogorov complexity of e relative to oracle K. Note that there is a recursive approximation $\lambda s, e. C_s(e)$ 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, $C(e_{i+1}) < C(e_i) + c$.

Now some learning criteria are considered which are based on the learning model considered by Gold [Gol67]. In general, the learning criteria are defined by prescribing how the sequence of hypotheses has to converge to a target concept. Fin learning requires that the learner outputs only one conjecture and that must be correct (before this conjecture the learner is allowed to conjecture ? denoting "no conjecture"). Ex learning allows the learner to revise the hypothesis finitely often and that the last of these hypotheses must be correct. Bc learning allows the learner to make an infinite sequence of conjectures, all of which are correct from some point onwards. Vac is the same as Bc with the additional constraint that there are only finitely many distinct conjectures made, that is, the conjectures are bounded by some bound b. These convergence notions are applied to learning as well as to other processes such as the process of finding a fixed point in the effective versions of KRT. Intuitively, the power of the learning criteria increases from Fin to Ex to Vac to Bc; so Fin_ψ-learning can only deal with simple classes while Bc_ψ-learning can handle quite complex classes.

Definition 1. A *learner* is a recursive function $M : \mathrm{SEQ} \to \mathbb{N} \cup \{?\}$.

Let α be any partial recursive function, M be a learner and ψ be a numbering (hypothesis space).

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

$$(\forall i < i_0)[M(T[i]) = ?] \ \wedge \ (\forall i \geq i_0)[M(T[i]) = e] \ \wedge \ \psi_e = \alpha. \tag{3}$$

(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

$$\begin{aligned}(\forall i < i_0)[M(T[i]) = ?] \ \wedge \ (\forall i \in \{i_0, \ldots, i_1 - 1\})[M(T[i]) = e_0] \\ \wedge \ (\forall i \geq i_1)[M(T[i]) = e_1] \ \wedge \ \psi_{e_1} = \alpha.\end{aligned} \tag{4}$$

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

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

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

$$(\forall i \geq i_0)[M(T[i]) \leq b \ \wedge \ \psi_{M(T[i])} = \alpha]. \tag{6}$$

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

$$M(T[i_0]) \neq ? \text{ and } (\forall i \geq i_0)[\psi_{M(T[i])} = \alpha]. \tag{7}$$

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 $\mathcal{S}$. A class $\mathcal{S}$ is I_ψ-learnable if some M I_ψ-learns $\mathcal{S}$. In above definitions, the subscript ψ is omitted when ψ is the fixed acceptable numbering φ.

Note that in part (d) of the above definition, M Vac_ψ-learning α is equivalent to saying that for all texts T for α, there exists a finite set S such that (i) for all $p \in S$, $\psi_p = \alpha$ and (ii) for all but finitely many i, $M(T[i]) \in S$.

In some cases, one only considers learning from restricted classes of texts rather than all texts as in the definition above. The following folklore result holds as one can effectively convert any text for a total function into its canonical text.

Lemma 2. *Suppose ψ is a numbering and $I \in \{\mathsf{Fin}_\psi, \mathsf{Ex}_\psi, \mathsf{Vac}_\psi, \mathsf{Bc}_\psi\}$. Suppose $\mathcal{S}$ is a class of total functions. If $\mathcal{S}$ is I-learnable from canonical texts then $\mathcal{S}$ is I-learnable from all texts.*

Thus, for learning classes of total recursive functions, one often only considers learnability from canonical texts.

Definition 3 (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, one defines items (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 [JS10]. Furthermore, if a numbering is effectively optimal for Ex-learning then it is effectively optimal for Vac-learning [JS10].

The following is the formal definition of Kleene's Recursion Theorem (KRT).

Definition 4 (Kleene [Kle38,Kle52]). For each numbering ψ, *Kleene's Recursion Theorem holds in* $\psi \Leftrightarrow$

$$(\forall p \in \mathbb{N})(\exists e \in \mathbb{N})[\psi_e = \psi_p(\langle e, \cdot \rangle)]. \tag{8}$$

Equation (8) 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 diagonalisations by defining ψ_e in a way implicitly employing parameter e (in effect, a self-copy of e) in some algorithmic task ψ_p. Equation (8) is both profound and an elegant mathematical tool for proofs in recursion theory and inductive inference.

The following are the constructive forms of Kleene's Recursion Theorem considered in this paper. Intuitively, $f(p)$ in Definition 5 (a) plays the role of e in Definition 4. In the remaining parts of the definition, weaker requirements are placed on how the self reference program e is obtained. The reader will note the similarity to Definition 1. Again, the idea is that when a numbering is a FinKrt-numbering then the complexity to find the self-reference program is small, while, in general, finding a self-reference program in a BcKrt-numbering is a very complex task.

Definition 5 (Moelius [Moe09]). Let ψ be a numbering.

(a) ψ is a FinKrt-*numbering* iff there exists a recursive function $f : \mathbb{N} \to \mathbb{N}$ such that, for each p,

$$\psi_{f(p)} = \psi_p(\langle f(p), \cdot \rangle). \tag{9}$$

(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

$$(\forall i \geq i_0)[f(p, i) = e] \ \wedge \ \psi_e = \psi_p(\langle e, \cdot \rangle). \tag{10}$$

(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 bound b such that

$$(\forall i \geq i_0)[f(p,i) \leq b \ \wedge \ \psi_{f(p,i)} = \psi_p(\langle f(p,i), \cdot \rangle)]. \tag{11}$$

(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

$$(\forall i \geq i_0)[\psi_{f(p,i)} = \psi_p(\langle f(p,i), \cdot \rangle)]. \tag{12}$$

The next definition generalises the notion of an acceptable numbering. Recall that ψ is acceptable iff indices of every other numbering can be translated into ψ (for the same function) via a recursive function.

Definition 6. For each numbering ψ, one defines items (a), (b), (c) and (d) 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

$$(\forall i < i_0)[f(p,i) = e_0] \ \wedge \ (\forall i \geq i_0)[f(p,i) = e_1] \ \wedge \ \psi_{e_1} = \varphi_p. \tag{13}$$

(b) [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

$$(\forall i \geq i_0)[f(p,i) = e] \ \wedge \ \psi_e = \varphi_p. \tag{14}$$

(c) ψ is Vac-*acceptable* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exist i_0 and a bound b such that

$$(\forall i \geq i_0)[f(p,i) \leq b \ \wedge \ \psi_{f(p,i)} = \varphi_p]. \tag{15}$$

(d) ψ is Bc-*acceptable* iff there exists a recursive function $f : \mathbb{N}^2 \to \mathbb{N}$ such that, for each p, there exists i_0 such that

$$(\forall i \geq i_0)[\psi_{f(p,i)} = \varphi_p]. \tag{16}$$

One could define Fin-acceptable numberings accordingly; they coincide, however, with the usual acceptable numberings. Furthermore, Ex-acceptable numberings are also known as K-acceptable numberings [JS10,JST11,ST12,ST14] as they are acceptable via a witness function computed relative to the halting problem K. Translations of numberings via K-recursive functions are widely considered in inductive inference [ABG11,JS94,JS97] and the theory of numberings in general [BG00,BGS03,Ersh77].

The results shown in this paper include the following ones.

- For $I \in \{\mathsf{Fin}, \mathsf{Ex}, \mathsf{Vac}\}$, numberings effectively optimal for I-learning are also I-acceptable and I-acceptable numberings are effectively optimal for I-learning; furthermore, numberings which are both I-effective for KRT and optimal for I-learning are also I-acceptable (Theorem 9). Here a numbering being Fin-effective for KRT means it is a FinKrt-numbering, and so on for other values of I.

- 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 13).
- There is a FinKrt-numbering which is not optimal for any of the learning criteria Fin, Ex, Vac, Bc (Theorem 15).
- There is an ExKrt-numbering which is not a FinKrt-numbering and which is effectively optimal for Ex-learning and Vac-learning, but not optimal for Fin or Bc-learning (Theorem 19).
- 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 21).
- 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 24).
- 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 25).
- 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 26).

The remainder of this paper is organised as follows. Section 3 presents the results connecting optimality, effective KRT and acceptability. Section 4 presents the results concerning numberings that do not satisfy Kleene's Recursion Theorem. Section 5 presents the results concerning numberings that satisfy Kleene's Recursion Theorem in an effective way. Section 6 presents the results concerning numberings that satisfy Kleene's Recursion Theorem, but not in an effective way.

3 Relating Optimality, KRT and Acceptability

The main finding of this paper is that the various versions of KRT and optimality complement each other. On one hand, for certain criteria, one can combine I-KRT and optimality for I-learning in order to achieve I-acceptability. On the other hand, these two notions are independent from each other and one can, as shown in numerous results below, adjust them independently to certain levels in a numbering.

For the learning criteria $\mathsf{Fin}, \mathsf{Ex}, \mathsf{Vac}, \mathsf{Bc}$, it is already known that every numbering which is effectively optimal for I-learning is also I-acceptable. This is based on a result from [JS10].

Proposition 7. *For $I \in \{\mathsf{Fin}, \mathsf{Ex}, \mathsf{Vac}, \mathsf{Bc}\}$, if a numbering ψ is effectively optimal for I-learning, then ψ is I-acceptable.*

Proof. There is a recursive function g such that the learner $\varphi_{g(e)}$ always outputs e, and thus I-learns $\{\varphi_e\}$. If ψ is effectively optimal for I-learning (say via function f), then $\varphi_{f(g(e))}$ will be an I_ψ-learner for $\{\varphi_e\}$. Let T_e be a text, obtained effectively from e, for φ_e. Then, for $I \in \{\mathsf{Ex}, \mathsf{Vac}, \mathsf{Bc}\}$, using $h(e,n) = \varphi_{f(g(e))}(T_e[n])$, one can obtain a sequence of ψ-indices converging to an index for φ_e in the sense I, giving that ψ is I-acceptable. For $I = \mathsf{Fin}$, one can define $h(e) =$ the first non ? conjecture output by $\varphi_{f(g(e))}$ on T_e, giving that ψ is acceptable.□ (**Proposition 7**)

Furthermore, acceptable numberings are effectively optimal for Fin-learning and Ex-acceptable numberings are effectively optimal for Ex-learning [JS10]. Theorem 19 below shows that Bc-acceptable numberings need not to be optimal for Bc-learning; indeed the result will show that even Ex-acceptable numberings need not to be optimal for Bc-learning. The remaining interesting case is vacillatory learning which is resolved by the following proposition.

Proposition 8. *Every* Vac*-acceptable numbering is effectively optimal for* Vac*-learning.*

Proof. Assume that ψ is Vac-acceptable via recursive function f. Furthermore consider any Vac_φ-learner M for some class $\mathcal{C}$. Now one can effectively from M, f construct the new learner N such that $N(\sigma) = f(M(\sigma), |\sigma|)$.

For the verification that N Vac_ψ-learns $\mathcal{C}$, consider any function in $\mathcal{C}$ and any text T for this function. The learner M outputs on T only indices below a bound a; the translation f outputs on inputs $e \leq a$ and t only indices below a bound b; thus b is also an upper bound for all outputs of the form $f(M(T[n]), n)$. Furthermore, if n is sufficiently large then $M(T[n])$ is a φ-index for the function to be learnt and, for each $e \leq a$, $\psi_{f(e,n)} = \varphi_e$. Thus $f(M(T[n]), n)$ is a ψ-index for the function to be learnt for all sufficiently large n. □ (**Proposition 8**)

For the converse direction from optimality to acceptability, it should be noted that in general the implication does not hold for all criteria in the non-effective case. For example, Jain and Stephan [JS10] showed that there is a numbering which is optimal for Ex-learning but neither effectively optimal for Ex-learning nor Ex-acceptable. However, having sufficiently effective KRT along with optimal learning, allows us to prove the converse direction for the case of explanatory, vacillatory and finite learning.

Theorem 9. (a) *If ψ is a* VacKrt*-numbering and optimal for* Vac*-learning, then ψ is a* Vac*-acceptable numbering.*

(b) *If ψ is an* ExKrt*-numbering and optimal for* Ex*-learning, then ψ is an* Ex*-acceptable numbering.*

(c) *If ψ is a* FinKrt*-numbering and optimal for* Fin*-learning, then ψ is an acceptable numbering.*

Proof. Consider any numbering ψ. Let d be any constant such that there exists an x with either $\psi_0(x)$ being undefined or $\psi_0(x)$ being defined and strictly smaller than d. For each e, let α_e be a partial recursive function such that for all p and all x, $\alpha_e(\langle 0, x\rangle) = d + e$ and $\alpha_e(\langle p+1, x\rangle) = \varphi_e(x)$. Let $\mathcal{S} = \{\alpha_e : e \in \mathbb{N}\}$.

Note that $\mathcal{S}$ is Fin_φ-learnable. Let T_e be a text for α_e obtained effectively from e. Note that such a text can be constructed effectively by carrying out the computations of α_e on different inputs in parallel and then arranging them into a text in the order of how the computations terminate.

(a) Suppose ψ is a VacKrt-numbering which is optimal for Vac-learning. Thus, $\mathcal{S}$ is Vac_ψ-learnable via some learner M. Let g be a recursive function which witnesses that ψ is a VacKrt-numbering.

Let $f(e, t) = g(M(T_e[t]), t)$. Thus, for some finite set S, for all but finitely many t, $M(T_e[t]) \in S$ and each $i \in S$ satisfies: $\psi_i = \alpha_e$. Furthermore, for all $i \in S$, there exists a finite set S_i such that, for all but finitely many t, $g(i,t) \in S_i$ and for all $e' \in S_i$, $\psi_{e'}(x) = \psi_i(\langle e', x\rangle) = \alpha_e(\langle e', x\rangle)$. Now, for all but finitely many t, $f(e,t) \in \bigcup_{i \in S} S_i$. Furthermore, by the choice of d and definition of α_e, it is impossible that $0 \in \bigcup_{i \in S} S_i$. Thus, by the definition of α_e, for all $e' \in \bigcup_{i \in S} S_i$, $\psi_{e'} = \varphi_e$. It follows that f witnesses that ψ is a Vac-acceptable numbering.

(b) Suppose ψ is an ExKrt-numbering which is optimal for Ex-learning. Thus, $\mathcal{S}$ is Ex_ψ-learnable via some learner M. Let g be a recursive function which witnesses that ψ is a ExKrt-numbering.

Let $f(e,t) = g(M(T_e[t]), t)$. Now, proceeding as in the proof of part (a), note that S contains exactly one index, say i, and S_i contains exactly one index. Hence, the rest of the argument as in part (a) shows that f witnesses that ψ is an Ex-acceptable numbering.

(c) Suppose ψ is a FinKrt-numbering which is optimal for Fin-learning. Thus, $\mathcal{S}$ is Fin_ψ-learnable via some learner M. Let g be a recursive function which witnesses that ψ is a FinKrt-numbering.

Let $f(e) = g(M(T_e[t_e])$, where t_e is least such that $M(T_e[t_e]) \neq ?$. Note that f is a recursive function. Now, $M_e(T_e[t_e]) = i$ is a ψ-index for α_e, and thus $\psi_{f(e)}(x) = \psi_i(\langle f(e), x\rangle) = \alpha_e(\langle f(e), x\rangle)$, for all x. As in part (a), it is not possible that $f(e) = 0$. Hence, $\psi_{f(e)} = \varphi_e$, and f witnesses that ψ is an acceptable numbering. □ (**Theorem 9**)

The following corollary characterises acceptable numberings in terms of optimality for Fin-learning.

Corollary 10. *The following conditions are equivalent for a numbering ψ:*

(a) *ψ is a* FinKrt*-numbering and optimal for* Fin*-learning;*
(b) *ψ is effectively optimal for* Fin*-learning;*
(c) *ψ is acceptable.*

Proof. Implication from (a) to (c) follows from Theorem 9. Clearly, (c) implies (a) and (b). The argument for the implication from (b) to (c) is given in the second paragraph of Section 3. □ (**Corollary 10**)

While one could show that Theorem 9 does not carry over to Ex_1, the corresponding question is open for behaviourally correct learning.

Open Problem 11. *Is there a* BcKrt*-numbering which is optimal for* Bc*-learning but not* Bc*-acceptable?*

So far, one can only prove a weaker result: if a BcKrt-numbering is optimal for Vac-learning, Ex-learning or Fin-learning then it is Bc-acceptable; the reason is that when learning the function α_e in the proof of Theorem 9, there will be only finitely many correct indices involved among which the learner vacillates and for each of these indices, the corresponding BcKrt-algorithm outputs, from some time on, only ψ-indices of φ_e.

Note that Theorem 13 below will show effectively optimal numberings for Ex-learning, Vac-learning and Bc-learning do not need to satisfy KRT at all. Thus one cannot obtain the equivalence of three conditions, as in Corollary 10, for the case of Ex, Vac or Bc replacing Fin.

Proposition 12. *Suppose $\mathcal{S}$ is* Vac*-learnable and ψ is* Bc*-acceptable. Then, $\mathcal{S} \in \mathsf{Bc}_\psi$.*

Proof. Suppose M witnesses that $\mathcal{S} \in \mathsf{Vac}$ and f witnesses that ψ is Bc-acceptable. Define a learner N as follows. $N(T[n]) = f(M(T[n]), n)$, where $f(?, n)$ is taken as ?. Now, if T is a text for $g \in \mathcal{S}$, then for for some bound b, for all but finitely many n, $M(T[n])$ is a φ-program for g, and $M(T[n]) \leq b$. Furthermore, for all but finitely many n, for all $e \leq b$, $\psi_{f(e,n)} = \varphi_e$. It follows that, for all but finitely many n, $\psi_{N(T[n])} = g$. Thus, N Bc_ψ-learns $\mathcal{S}$. □ (**Proposition 12**)

4 When Kleene's Recursion Theorem is Absent

This section presents the 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 generalise to other criteria of acceptability. In particular, there is an Ex_1-acceptable numbering that does not satisfy Kleene's Recursion Theorem. Note that "Ex_1-acceptable" is quite close to being "acceptable", as one can translate any other numbering into this numbering by making at most one mind change during the translation process. Furthermore, when used as a hypothesis space, the complexity of learning is not increased.

Theorem 13. *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 is 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,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } \mathrm{rng}(\varphi_e) \not\subseteq \{e, e+1, e+2, \ldots\}; \\ \uparrow, & \text{if } \mathrm{rng}(\varphi_e) \subseteq \{e, e+1, e+2, \ldots\}. \end{cases} \tag{17}$$

As every partial recursive function has infinitely many φ-indices, it has a φ-index larger than the minimal element in its range (if any). Thus, every partial recursive function has a ψ-index.

(a) There is no e such that $\psi_e = \lambda x\,.\,e$; hence, there is no e such that $\psi_e = \alpha(\langle e, \cdot \rangle)$ when $\alpha = \lambda\langle e, x\rangle\,.\,e$.

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

(c) To show that ψ is effectively optimal for Ex, Vac and Bc-learning, given a learner M, the new learner N first conjectures 0 for $\lambda x\,.\uparrow$. 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 $\mathrm{pad}(e, y)$.

(d) To show that ψ is optimal for Fin-learning, consider any learner M.

First suppose that M is a Fin-learner for a class not containing $\lambda x\,.\uparrow$. Define a new learner N as follows. If $\mathrm{rng}(\sigma) \subseteq \{\#\}$ or $M(\sigma) =?$, then $N(\sigma) =?$. Otherwise, let (x, y) be the first pair different from # which appears in σ, and let $N(\sigma) = \mathrm{pad}(M(\sigma), y)$. It is easy to verify that N Fin_ψ-learns every partial function Fin-learnt by M.

Now suppose that M Fin-learns $\lambda x.\uparrow$. Thus, M cannot Fin-learn any other function $\eta \supset \lambda x.\uparrow$. Consider learner N which outputs 0 on all inputs. N Fin_ψ-learns $\lambda x.\uparrow$.

It follows that ψ is optimal for Fin-learning. $\square$ (**Theorem 13**)

Note that in the above theorem, as ψ is not acceptable, ψ cannot be effectively optimal for Fin-learning [JS10]; hence, the non-uniform case distinction in the proof cannot be avoided.

5 When Kleene's Recursion Theorem is Effective

This section presents the results concerning numberings that satisfy Kleene's Recursion Theorem in an effective way. These results establish that there is a hierarchy based on how complex it is to find witnesses for Kleene's Recursion Theorem; this hierarchy is inspired by the corresponding complexity hierarchy for learning (in terms of convergence constraints). First, it is shown that a numbering can be a FinKrt-numbering, yet not be optimal for learning (Theorem 15). Second, there exists an ExKrt-numbering that is not a FinKrt-numbering (Theorem 19). Third, there exists a VacKrt-numbering that is not an ExKrt-numbering (Theorem 21). Finally, there exists a BcKrt-numbering that is not a VacKrt-numbering (Theorem 24).

Lemma 14. *There exists a numbering ϑ such that the set $\{i_0, i_1, \ldots\} = \mathrm{INF}_\vartheta$ of indices of functions with infinite domain (in ascending order $i_0 < i_1 < \ldots$) satisfies that $\lambda x.i_x$ dominates every K-recursive function and that no infinite K-r.e. set is contained in INF_ϑ.*

Proof. Let K_s denote K enumerated within s steps. Let $F_{e,s}(x)$ be a recursive approximation to limit-recursive values $F_e(x)$ defined inductively as follows: $F_{e,0}(x) = 0$; if the computations of $\varphi_{e,s}^{K_s}(x)$ and $\varphi_{e,s+1}^{K_{s+1}}(x)$ differ at some oracle query or take longer than s steps or output different results or output a result larger than $F_{e,s}(x)$, then $F_{e,s+1}(x) = F_{e,s}(x)+1$, else $F_{e,s+1}(x) = F_{e,s}(x)$. Now $F_e(x) = \lim_{s\to\infty} F_{e,s}(x)$ is finite iff $\varphi_e^K(x)$ is defined. More precisely, the following statements hold:

(i) for all e, x, $F_{e,s}(x) \le F_{e,s+1}(x)$;
(ii) if $\varphi_e^K(x)\downarrow$, then $F_e(x)\downarrow \ge \varphi_e^K(x)$;
(iii) if $\varphi_e^K(x)\uparrow$, then $F_e(x)\uparrow$.

Note that each F_e is a partial K-recursive function with the same domain as φ_e^K and that the set $\{(e, x, y) : F_e(x) = y\}$ is K-recursive. By (ii), the function $G(n) = \max\{F_e(n) : e \le n \wedge F_e(n)\downarrow\}$ dominates every K-recursive function. Let ϑ be such that, for each n, m and x,

$$\vartheta_{\langle n,m\rangle}(x) = \begin{cases} \varphi_n(x), & \text{if } (\exists s > x)[\quad (\exists e \le n)[F_{e,s}(n) = m] \\ & \qquad\qquad \wedge\ (\forall d \le n)[\quad F_{d,s}(n) \le m \\ & \qquad\qquad\qquad \vee\ F_{d,s}(n) > F_{d,x}(n)]]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{18}$$

Now it is shown that, for each n and m, $\vartheta_{\langle n,m\rangle}$ has infinite domain iff φ_n has infinite domain and $m = G(n)$. 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, $\vartheta_{\langle n,m\rangle} \subseteq \varphi_n$. Thus, $\vartheta_{\langle n,m\rangle}$ has finite domain.

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

$$F_{e,w}(n) > m. \tag{19}$$

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

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

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

and

$$F_e(n)\uparrow \ \Rightarrow \ F_{e,w}(n) > m. \tag{21}$$

Now, 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 all but finitely many x, $\vartheta_{\langle n,m\rangle}(x)\uparrow$.

Case 4: φ_n has infinite domain and $m = G(n)$. Now, for each x, there exists an $s > x$ such that $(\exists e \leq n)[F_{e,s}(n) = m]$ and, for each $d \leq n$,

$$[F_d(n)\downarrow \ \Rightarrow \ F_{d,s}(n) = F_d(n) \leq m] \ \wedge \ [F_d(n)\uparrow \ \Rightarrow \ F_{d,s}(n) > F_{d,x}(n)]. \tag{22}$$

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

It follows from the above analysis that INF_ϑ is contained in the set $\{\langle n, G(n)\rangle : n \in \mathbb{N}\}$. As $G(n)$ dominates every K-recursive function, it immediately follows that if $\{i_0, i_1, \ldots,\} = \mathrm{INF}_\vartheta$ (in ascending order, $i_0 < i_1 < \ldots$), then $\lambda x.i_x$ dominates every K-recursive function. Hence, as every infinite K-r.e. set contains an infinite K-recursive subset, it also follows that no infinite K-r.e. set is contained in INF_ϑ.

Note that for each n, $\vartheta_{\langle n,G(n)\rangle} = \varphi_n$; hence ϑ is a numbering for $\mathcal{P}$. □ (**Lemma 14**)

Theorem 15. *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 *and ψ is not* Bc*-acceptable.*

Proof. Let ϑ be as in Lemma 14. Define ψ as follows. For each p, i and x,

$$\psi_{\langle p,0\rangle}(x) = \vartheta_p(x); \tag{23}$$

$$\psi_{\langle p,i+1\rangle}(x) = \psi_{\langle p,i\rangle}(\langle\langle p, i+1\rangle, x\rangle). \tag{24}$$

As ϑ is a numbering for $\mathcal{P}$, so is ψ. Furthermore, ψ is defined such that $\psi_{\langle p,i+1\rangle}$ coincides with the $\langle p, i+1\rangle$-th slice of $\psi_{\langle p,i\rangle}$. Hence, ψ is a FinKrt-numbering.

Let $\mathcal{S} = \{f_0, f_1, f_2, \ldots\}$ 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. To show that ψ is not optimal for any of the criteria Fin, Ex, Vac, Bc, it is shown that $\mathcal{S}$ is not $\mathsf{Fin}_\psi, \mathsf{Ex}_\psi, \mathsf{Vac}_\psi, \mathsf{Bc}_\psi$-learnable. It follows also that ψ is not Bc-acceptable – otherwise, $\mathcal{S}$ would be Bc_ψ-learnable by Proposition 12.

Claim 16. *For each p, there exists at most one i such that $\psi_{\langle p,i\rangle} \in \mathcal{S}$.*

To see the claim, first note that, for all n, $\mathrm{rng}(f_n)$ is cofinite. Suppose $\psi_{\langle p,i\rangle} = f_n$ for some i. Now, for each $j > i$, $\mathrm{rng}(\psi_{\langle p,i\rangle}) - \mathrm{rng}(\psi_{\langle p,j\rangle})$ is infinite as f_n is 1–1 total function. Hence for all m, for all $j > i$, $\psi_{\langle p,j\rangle} \neq f_m$. This proves the claim.

Let T_n be the canonical text for f_n.

Now it is shown that $\mathcal{S}$ is not Vac_ψ-learnable. For this, suppose by way of contradiction that some learner M Vac_ψ-learns $\mathcal{S}$. Let $c_n = \max\{M(T_n[m]) : m \in \mathbb{N}\}$. Now, c_n is defined for all n and can be obtained K-recursively from n. Furthermore, c_n is greater than the minimal ψ-index for f_n and thus ϑ-index for f_n. Thus, for $\mathrm{INF}_\vartheta = \{i_0, i_1, \ldots\}$ where $i_0 < i_1 < \ldots$, for all n, $[1 + \max\{c_i : i \leq n\}] > i_n$ contradicting Lemma 14. Hence, $\mathcal{S}$ is not Vac_ψ-learnable and thus neither Fin_ψ nor Ex_ψ-learnable.

Now suppose by way of contradiction that some learner M Bc_ψ-learns $\mathcal{S}$. By Claim 16 it follows that, for each p and n, only finitely many different indices of the form $\langle p, i\rangle$ belong to $prog_n = \{M(T_n[m]) : m \in \mathbb{N}\}$. For each n, let $prog'_n = \{p : (\exists i \in \mathbb{N})[\langle p, i\rangle \in prog_n]\}$. If for some n, $prog'_n$ is infinite, then by removing the finitely many non-total ϑ-programs from the r.e. set $prog'_n$ one gets an infinite r.e. set of ϑ programs which is contained in INF_ϑ, contradicting Lemma 14. Thus, for each n, the set $prog_n$ is finite. Hence M Vac_ψ-learns $\mathcal{S}$ from canonical texts. It follows that $\mathcal{S}$ is Vac_ψ-learnable. However, this is false as shown in the previous paragraph. □ (**Theorem 15**)

The following theorem gives an abstract condition under which learning or the complexity of finding witnesses for self-reference can be brought down from "behaviourally correct" to "vacillatory".

Theorem 17. *Suppose ψ is a numbering and INF_ψ is immune. Then, the following hold:*

(a) *ψ is BcKrt-numbering implies ψ is a VacKrt-numbering.*
(b) *For any $\mathcal{C}$, $\mathcal{C}$ is Bc_ψ-learnable implies $\mathcal{C}$ is Vac_ψ-learnable.*

Proof. As the proofs for parts (a) and (b) are similar, only part (a) will be shown. Suppose that f witnesses that ψ is a BcKrt-numbering. A new function f' witnessing that ψ is a VacKrt-numbering is defined as follows:

$f'(p, s)$ outputs index $f(p, j)$ for the least j such that

$$(\forall x \leq s)[\psi_{f(p,j),s}(x) = \psi_{p,s}(\langle f(p,j), x\rangle)] \ \vee \ j = s. \tag{25}$$

First consider the case that for some j, $\psi_{f(p,j)} = \psi_p(\langle f(p,j), \cdot\rangle)$ and $\psi_{f(p,j)}$ is a finite function. Let s be such that the following two equations are satisfied:

$$(\forall i \leq j)[\psi_{f(p,i)} \neq \psi_p(\langle f(p,i), \cdot\rangle) \Rightarrow (\exists x \leq s)(\forall s' \geq s)[\psi_{f(p,i),s'}(x) \neq \psi_{p,s'}(\langle f(p,i), x\rangle)]]; \tag{26}$$

$$\psi_{f(p,j),s} = \psi_{f(p,j)}. \tag{27}$$

Note that, for $s' \geq s$, $f'(p, s') \in \{f(p, 0), f(p, 1), \ldots, f(p, j)\}$. Hence, $\{f'(p, s') : s' \in \mathbb{N}\}$ is finite, as required. Furthermore, every $f(p, i)$ such that $f'(p, s') = f(p, i)$ for infinitely many s', satisfies

$\psi_{f(p,i)} = \psi_p(\langle f(p,i), \cdot \rangle)$.

Next consider the case that, for each i, if $\psi_{f(p,i)} = \psi_p(\langle f(p,i), \cdot \rangle)$, then $\psi_{f(p,i)}$ has infinite domain. By INF_ψ being immune, it follows that $\{f(p,0), f(p,1), f(p,2), \ldots\}$ is a finite set. Thus, for some j, $f(p,0), f(p,1), \ldots, f(p,j)$ represents *all* of the indices output by $f(p, \cdot)$. Let s satisfy Equation (26). It follows that, for $s' \geq s$, $f'(p,s') = f(p,i)$ implies $\psi_{f(p,i)} = \psi_p(\langle f(p,i), \cdot \rangle)$. Furthermore, $\{f'(p,s') : s' \in \mathbb{N}\}$ is finite.

It follows from the above analysis that, in both cases, for fixed first argument, f' outputs only finitely many different indices and, for all but finitely many s, $f'(p,s) = e$ satisfying $\psi_e = \psi_p(\langle e, \cdot \rangle)$. Hence f' witnesses that ψ is a VacKrt-numbering. □ (**Theorem 17**)

Theorem 19 shows, in part, that there exist ExKrt-numberings that are not FinKrt-numberings.

Lemma 18. *There exists a numbering ψ such that the following conditions hold:*

(a) INF_ψ *is immune;*

(b) *For any infinite r.e. set E, at least one $m \in E$ satisfies $\psi_m = \varphi_m$ and one such m can be found effectively in the limit from an index for E.*

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

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } (\exists s > x)(\exists n)[n^2 \leq C_s(e) \leq n^2 + n]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{28}$$

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. Now, there exists a constant c such that, for each i, $C(e_{i+1}) < C(e_i) + c$. Let n be sufficiently large so that $n > C(e_0)$ and $n > c$.

For part (a), as there are only finitely many indices with Kolmogorov complexity below n^2+n, there exists a largest i such that $C(e_i) \leq n^2 + n$. Thus,

$$n^2 + n < C(e_{i+1}) < C(e_i) + c < n^2 + 2n < (n+1)^2. \tag{29}$$

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

For part (b), as there are only finitely many indices with Kolmogorov complexity below n^2, there exists a largest i such that $C(e_i) \leq n^2$. Thus,

$$n^2 < C(e_{i+1}) < C(e_i) + c < n^2 + n. \tag{30}$$

Hence, $\psi_{e_{i+1}} = \varphi_{e_{i+1}}$. Note that an index $e \in E$ satisfying $n^2 < C(e) < n^2 + n$, for some n, can be found in the limit from an index for E. Part (b) follows.

As every partial recursive function has an infinite r.e. set of φ-indices, it also follows from part (b) that ψ is a numbering for $\mathcal{P}$. □ (**Lemma 18**)

Theorem 19. *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 and* Vac*-learning;*
(e) *ψ is neither optimal for* Fin *nor for* Bc*-learning.*

Proof. Let ψ be as given by Lemma 18.

(a) To show that ψ is Ex-acceptable, let $S_e = \{\mathrm{pad}(e, y) : y \in \mathbb{N}\}$. Then, by Lemma 18 (b), there exists a $y \in N$ such that $\psi_{\mathrm{pad}(e,y)} = \varphi_{\mathrm{pad}(e,y)} = \varphi_e$ and such a y can be found in the limit from e. It follows that ψ is Ex-acceptable.

(b) Now it is shown that ψ is an ExKrt-numbering. By the Operator Recursion Theorem [Cas74], there exists a 1–1 recursive function q such that $\varphi_{q(p,i)} = \psi_p(\langle q(p, i), \cdot\rangle)$. Let $E_p = \{q(p, i) : i \in \mathbb{N}\}$. Now, $E_0, E_1, E_2, \ldots$ is a uniformly r.e. family of infinite sets such that, for each p and each $e \in E_p$, $\varphi_e = \psi_p(\langle e, \cdot\rangle)$. Therefore, by Lemma 18, effectively from p, one can find in the limit an $e \in E_p$ such that $\psi_e = \varphi_e = \psi_p(\langle e, \cdot\rangle)$. Thus, ψ is an ExKrt-numbering.

(c) To show that ψ is not a FinKrt-numbering, assume otherwise as witnessed by recursive function h. Let e_0 be such that $\psi_{e_0} = \lambda x . x$. Let $e_{i+1} = h(e_i)$. Now, for each n, $\mathrm{rng}(\psi_{e_{n+1}})$ is a proper subset of $\mathrm{rng}(\psi_{e_n})$ and ψ_{e_i} is a total function (and thus has infinite domain). Hence, $\{e_0, e_1, e_2, \ldots\}$ is an infinite r.e. set of ψ-indices of functions with infinite domain. However, this contradicts Lemma 18 (a).

(d) It is known that Ex-acceptable numberings are effectively optimal for Ex-learning and Vac-learning [JS10]. Thus, ψ is effectively optimal for Ex-learning and Vac-learning.

(e) To show that ψ is not optimal for Fin-learning, consider the class $\mathcal{S}_{const}$ of constant functions which is Fin-learnable. Let T^n denote the canonical text for the constant n function. Suppose by way of contradiction that $\mathcal{S}_{const}$ is Fin_ψ-learnable by some learner M. Let $X = \{M(T^n[m]) : m, n \in \mathbb{N}\} - \{?\}$. Then, X is an infinite r.e. set consisting only of ψ-indices of total functions (which thus have infinite domain). Again, this would contradict Lemma 18 (a).

As INF_ψ is immune, by Theorem 17 it follows that any Bc_ψ-learnable class is Vac_ψ-learnable. As there are Bc-learnable classes of partial functions which are not Vac-learnable [CS83], ψ is not optimal for Bc-learning. □ (**Theorem 19**)

Lemma 20. *There exists a numbering ψ and a recursive function f such that, for all e, d:*

(a) If $d < 2^{e+2}$ and $C^K(d) \geq e$, then $\psi_{\langle d,e\rangle} = \varphi_e(\langle\langle d, e\rangle, \cdot\rangle)$, else $\psi_{\langle d,e\rangle}$ is a function with a finite domain on which it coincides with $\varphi_e(\langle d, e\rangle, \cdot)$;

(b) $\{f(e, t) : t \in \mathbb{N}\}$ is finite and, for all but finitely many t, $\psi_{\langle f(e,t),e\rangle} = \varphi_e(\langle\langle f(e, t), e\rangle, \cdot\rangle)$.

(c) for any infinite K-r.e. set $E = \{\langle d_0, e_0\rangle, \langle d_1, e_1\rangle, \ldots\}$, there exist infinitely many $\langle d, e\rangle \in E$ such that $\psi_{\langle d,e\rangle}$ is finite.

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

$$\psi_{\langle d,e\rangle}(x) = \begin{cases} \varphi_e(\langle\langle d,e\rangle,x\rangle), & \text{if } d < 2^{e+2} \text{ and } (\exists s > x)[C_s^*(d) \geq e]; \\ \uparrow, & \text{otherwise.} \end{cases}$$

Now, (a) follows by definition of ψ above.

Note that ψ is a numbering for all partial recursive functions, as for all e, there exists a $d < 2^{e+2}$ such that $C^K(d) \geq e$. Let f be a recursive function such that, $f(e,s) = d$, for the least d such that $d < 2^{e+2}$ and $C_s^*(d) \geq e$. It is easy to verify that part (b) of the lemma is satisfied.

For part (c), note that if there exist infinitely many $\langle d_n, e_n\rangle$ in E such that $d_n \geq 2^{e_n+2}$ or $C^K(d_n) < e_n$, then part (c) holds by part (a). Otherwise, for each a, there exists a $\langle d_{n_a}, e_{n_a}\rangle$ in E such that $e_{n_a} \geq a$ and $d_{n_a} \geq a$, and such $\langle d_{n_a}, e_{n_a}\rangle$ can be found, K-effectively from a. It follows that $C^K(d_{n_a}) \leq 2\log(a) \leq e_{n_a}$ for large enough a. Hence, for large enough a, $\psi_{\langle d_{n_a}, e_{n_a}\rangle}$ is finite. □ (**Lemma 20**)

Theorem 21. *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 a* Vac*-acceptable numbering and effectively optimal for* Vac*;*
(d) *ψ is not optimal for any of the learning criteria* Fin, Ex, Bc.

Proof. Let ψ and f be as given by Lemma 20. Let g, h be recursive functions such that, for all p, x, $\varphi_{g(e)}(x) = \psi_e(x)$ and $\varphi_{h(e)}(\langle p, x\rangle) = \varphi_e(x)$. Now the different parts of the theorem are shown.

(a) By part (b) of Lemma 20, for all but finitely many t, for all x, $\psi_{\langle f(g(e),t),g(e)\rangle}(x) = \varphi_{g(e)}(\langle\langle f(g(e),t), g(e)\rangle, x\rangle) = \psi_e(\langle\langle f(g(e),t), g(e)\rangle, x\rangle)$. Thus, $\lambda e, t\,.\,\langle f(g(e),t), g(e)\rangle$ witnesses that ψ is a VacKrt-numbering.

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

$$(\forall n, x)\,[\psi_{\langle d_{n+1},e_{n+1}\rangle}(x) = \psi_{\langle d_n,e_n\rangle}(\langle\langle d_{n+1}, e_{n+1}\rangle, x\rangle)]. \tag{31}$$

Now, for all n, $\psi_{\langle d_n,e_n\rangle}$ is total. Furthermore, $range(\psi_{\langle d_{n+1},e_{n+1}\rangle}) \subset range(\psi_{\langle d_n,e_n\rangle})$. Thus, $\langle d_n, e_n\rangle$ are pairwise different for different n. Hence, the r.e. set $\{\langle d_n, e_n\rangle : n \in \mathbb{N}\}$ is infinite and consists of ψ-indices of functions with infinite domain. Hence, by Lemma 20(c), there exists a $\langle d_n, e_n\rangle$ such that $\psi_{\langle d_n,e_n\rangle}$ is a finite function, a contradiction.

(c) Part (b) of Lemma 20 and the definition of h show that the mapping given by $e, t \mapsto \langle f(h(e),t), h(e)\rangle$ witnesses that ψ is a Vac-acceptable numbering: for all but finitely many t, for all x, $\psi_{\langle f(h(e),t),h(e)\rangle}(x) = \varphi_{h(e)}(\langle\langle f(h(e),t), h(e)\rangle, x\rangle) = \varphi_e(x)$. Moreover, $\{\langle f(h(e),t), h(e)\rangle : t \in \mathbb{N}\}$ is finite for each e. Furthermore, one can convert a Vac-learner M using φ as the hypothesis space to a Vac-learner M' using ψ as the hypothesis space by having $M'(T[n]) = \langle f(h(M(T[n])), n), h(M(T[n]))\rangle$.

(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 $f \in \mathcal{S}$, let $E_f = \{M(f[n]) : n \in \mathbb{N}\}$. Next it is shown 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. Thus, by Lemma 20(c), there exist infinitely many $\langle d, e\rangle \in E_f$ such that $\psi_{\langle d,e\rangle}$ is a finite function, a contradiction to M Bc-learning f. Thus, E_f is finite for each $f \in \mathcal{S}$ and M Vac-learns $\mathcal{S}$ from canonical texts. Hence $\mathcal{S}$ is Vac-learnable. 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_ψ-learns the class $\mathcal{S}_{const}$. Let T^c be the canonical text for the constant c function. Let $\langle d_c, e_c\rangle$ be the conjecture to which M converges on text T^c. Note that d_c, e_c can be computed using oracle K, and $\{\langle d_c, e_c\rangle : c \in \mathbb{N}\}$ is infinite (as $\langle d_c, e_c\rangle$ is different for different c). Thus, by Lemma 20(c), there exists a c such that $\psi_{\langle d_c,e_c\rangle}$ is a finite function and thus M does not Ex_ψ-learn the constant c function. It follows that M does not Ex_ψ-learn the class $\mathcal{S}_{const}$. Thus, ψ is not optimal for Fin and Ex-learning. $\square$ (**Theorem 21**)

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

Proposition 22. $\{d : C(d) < \log(d)\}$ *is a simple set.*

Proof. Clearly, $\{d : C(d) < \log(d)\}$ is an r.e. set. Furthermore, its complement is infinite as for all x, there are at least 2^x numbers with $\log(y) = x$, but at most $2^x - 1$ numbers with $C(y) < x$. Now, for any infinite r.e. set E, one can effectively find an $e \in E$ such that $C(e) < \log(e)$. Thus, $\{d : C(d) < \log(d)\}$ is simple. $\square$ (**Proposition 22**)

Lemma 23. *There exists a numbering ψ such that the following conditions hold:*

(a) *If $\varphi_e(0)\uparrow$ or $|\mathrm{rng}(\varphi_e)| \geq 2$ or $C(\varphi_e(0)) < \log(\varphi_e(0))$ or $|W_{\log(\varphi_e(0))}| < e$ or $|W_{\log(\varphi_e(0))}| = \infty$, then $\psi_e = \varphi_e$, else ψ_e is a function with a finite domain on which it coincides with φ_e;*

(b) *There is no K-recursive function f such that, for all x, $f(x) \geq \min\{e' : (\exists y)\ [\psi_{e'} = \lambda z.y,$ $C(y) \geq x$ and $\log(y) = x]\}$.*

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

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } [\quad \varphi_{e,x}(0)\uparrow \\ & \quad\ \vee\ |\mathrm{rng}(\varphi_e)| \geq 2 \\ & \quad\ \vee\ [\varphi_e(0)\downarrow\ \wedge\ C(\varphi_e(0)) < \log(\varphi_e(0))] \\ & \quad\ \vee\ [\varphi_e(0)\downarrow\ \wedge\ |W_{\log(\varphi_e(0)),x}| < e] \\ & \quad\ \vee\ [\varphi_e(0)\downarrow\ \wedge\ |W_{\log(\varphi_e(0))}| > x] \qquad]; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{32}$$

Part (a) of the Lemma holds directly by construction. As every partial recursive function has infinitely many indices, it also follows from part (a) that ψ is a numbering for $\mathcal{P}$ (as for any partial recursive function η, either $\eta(0) =\uparrow$, or $W_{\log(\eta(0))}$ is infinite, or for some φ-index e for η, $|W_{\log(\eta(0))}| < e$).

To see part (b), suppose otherwise. For each x, let $e_x = f(x)$ be as in the statement of part (b) and let $\psi_{e_x}(0) = y_x$. Thus, ψ_{e_x} is the constant y_x function, $\log(y_x) = x$, and $C(y_x) \geq x$. Thus, by definition of ψ_{e_x}, it follows that either $|W_x| < e_x$ or $|W_x| > z$ for all z. Thus, $|W_x| < e_x \leq f(x)$ or $|W_x| = \infty > f(x)$. Thus, one can determine whether W_x is finite or not using a limit recursive function, which is known to be false [Soa87]. □ (**Lemma 23**)

Theorem 24. *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 as given by Lemma 23.

(a) Using the Operator Recursion Theorem [Cas74], 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. $M(p, s)$ is the first element e in some canonical enumeration of E_p such that

$$\begin{array}{ll} & \varphi_{e,s}(0)\uparrow \\ \vee & |\mathrm{rng}(\varphi_{e,s})| \geq 2 \\ \vee & [\varphi_{e,s}(0)\downarrow \ \wedge \ C_s(\varphi_{e,s}(0)) < \log(\varphi_{e,s}(0))] \\ \vee & e > s. \end{array} \tag{33}$$

Consider the following two cases.

Case 1: There exists an $e \in E_p$ such that $\varphi_e(0)\uparrow$ or $|\mathrm{rng}(\varphi_e)| \geq 2$ or $\varphi_e(0) = y$ for some y with $C(y) < \log(y)$. In this case, $\lim_{s\to\infty} M(p, s)$ 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. In this case, the set $F = \{\varphi_e(0) : M(p, s) = e \text{ for some } s\}$ 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, $M(p, s)$ outputs each index in E_p only for finitely many s. Hence, for all but finitely many s, $M(p, s)$ outputs 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, for all but finitely many s, $M(p, s)$ outputs an e such that $\psi_e = \varphi_e = \psi_p(\langle e, \cdot\rangle)$.

Thus, M witnesses that ψ is a BcKrt-numbering.

(b) To show that ψ is not a VacKrt-numbering, assume by way of contradiction otherwise, as witnessed by M. Define a K-recursive function f as follows. For any given x, using oracle for K, f first finds two distinct numbers y and z such that $\log(y) = \log(z) = x$ and $C(y) \geq x$ and $C(z) \geq x$. Note that there exist such y and z by the assumption on C given in Section 2. Then, f (implicitly using s-m-n theorem) determines a p such that, for all v and w,

$$\varphi_p(\langle v, w\rangle) = \begin{cases} y, & \text{if } v \text{ is even;} \\ z, & \text{if } v \text{ is odd.} \end{cases} \tag{34}$$

Note that $|\text{rng}(\varphi_p)| = 2$ and, hence, $\psi_p = \varphi_p$. Now, let $f(x) = \max\{M(p, s) : s \in \mathbb{N}\}$. Using limit lemma [Soa87], f is K-recursive. Furthermore, $f(x)$ is a bound on index e_x such that $\psi_{e_x} = \psi_p(\langle e_x, \cdot\rangle) = \varphi_p(\langle e_x, \cdot\rangle)$, which is either the constant y function or the constant z function. This however contradicts part (b) of Lemma 23.

(c) To show that ψ is not optimal for any of the learning criteria Fin, Ex, Vac, note that the class $\mathcal{S}_{const}$ is Fin-learnable. Suppose by way of contradiction that M Vac_ψ-learns $\mathcal{S}_{const}$. Let f be a K-recursive function defined as follows. On input x, using oracle for K, first determine a y such that $\log(y) = x$ and $C(y) \geq x$. Let T^y be the canonical text for the constant y function. Let $f(x) = \max\{M(T^y[m]) : m \in \mathbb{N}\}$. Now, $f(x)$ is a bound on a ψ-index e_x for the constant y function, with $\log(y) = x$ and $C(y) \geq x$. However, this contradicts Lemma 23(b). Hence, ψ is not optimal for any of the learning criteria Fin, Ex, Vac.

(d) 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 (this can be easily ensured by using the padding function pad). Now it is shown that, for all but finitely many i, $\psi_{e_i} = \varphi_{e_i}$. Consider the following three cases.

Case 1: $\alpha(0)\uparrow$. Then, for all but finitely many 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 all but finitely many 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 all but finitely many i, $|W_{\log(\varphi_{e_i}(0))}| < e_i$. Hence, for all but finitely many 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. $\Box$ (**Theorem 24**)

6 When Kleene's Recursion Theorem is Ineffective

This section presents the 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. Therefore, in such numberings, Kleene's Recursion Theorem is extremely ineffective. Theorems 25 and 26 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).

Theorem 25. *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 and ψ is not Bc-acceptable.*

Proof. Let ψ' be the numbering called "ψ" in Lemma 18 and let ψ'' be the numbering called "ψ" in Lemma 23. Let ψ be such that, for each e and x,

$$\psi_e(x) = \begin{cases} \varphi_e(x), & \text{if } \psi'_e(x)\downarrow \text{ and } \psi''_e(x)\downarrow; \\ \uparrow, & \text{otherwise.} \end{cases} \tag{35}$$

(a) 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

$$F = \begin{cases} \{e \in E : \varphi_e(0) = y\}, & \text{if } |W_{\log(y)}| = \infty; \\ \{e \in E : \varphi_e(0) = y \ \wedge \ e > |W_{\log(y)}|\}, & \text{otherwise.} \end{cases} \tag{36}$$

Note that, for each $e \in F$, $\psi''_e = \varphi_e$. Further note that F is r.e. and infinite. Hence, by Lemma 18(b), there exists an $e \in F$ such that $\psi'_e = \varphi_e$, and hence $\psi_e = \varphi_e = \alpha(\langle e, \cdot \rangle)$.

Case 2: The set $\{y : (\exists e \in E)\,[\varphi_e(0) = y]\}$ 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, $\psi''_{e_i} = \varphi_{e_i}$. As $F = \{e_0, e_1, e_2, \ldots\}$ is r.e., by Lemma 18(b), there exists an $e \in F$ such that $\psi'_e = \varphi_e$, and hence $\psi_e = \varphi_e = \alpha(\langle e, \cdot \rangle)$.

Case 3: Not Cases 1 and 2. Then, the set $\{e \in E : \varphi_e(0)\downarrow\}$ is finite. It follows that $F = \{e \in E : \varphi_e(0)\uparrow\}$ is r.e. and infinite. Furthermore, for each $e \in F$, $\psi''_e = \varphi_e$. Also, by Lemma 18(b), there exists an $e \in F$ such that $\psi'_e = \varphi_e$, and hence $\psi_e = \varphi_e = \alpha(\langle e, \cdot \rangle)$.

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

Note that the above analysis also shows that ψ is a numbering for $\mathcal{P}$, as for every partial recursive function η, there exists an α such that $\alpha(e, \cdot) = \eta$.

(b) Assume by way of contradiction that ψ is a BcKrt-numbering. 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 17. So, suppose that M witnesses that ψ is a VacKrt-numbering. The following is a slight modification of the argument used in part (b) of the proof for Theorem 24 to make sure that $\psi'_p = \varphi_p$. Define a K-recursive function f as follows. For any given x, using oracle for K, f first finds two distinct numbers y and z such that $\log(y) = \log(z) = x$ and $C(y) \geq x$ and $C(z) \geq x$. Then, f (implicitly using parameterised s-m-n theorem) and oracle for K determines a p such that, for all v and w,

$$\psi'_p(\langle v, w \rangle) = \varphi_p(\langle v, w \rangle) = \begin{cases} y, & \text{if } v \text{ is even;} \\ z, & \text{if } v \text{ is odd.} \end{cases} \tag{37}$$

Note that this additional constraint of $\psi'_p = \varphi_p$ can be easily satisfied using Lemma 18(b). Note that $|\mathrm{rng}(\varphi_p)| = 2$ and, hence, $\psi''_p = \varphi_p$. It follows that $\psi_p = \varphi_p$. Now, let $f(x) = \max\{M(p, s) : s \in \mathbb{N}\}$. Using limit lemma [Soa87], f is K-recursive. Furthermore, $f(x)$ is a bound on index e_x such that $\psi_{e_x} = \psi_p(\langle e_x, \cdot \rangle) = \varphi_p(\langle e_x, \cdot \rangle)$, which is either the constant y function or the constant z function. This however contradicts part (b) of Lemma 23.

(c) 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 $\mathcal{S}_{const}$ was Fin_ψ, Ex_ψ, or 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 24, this class is not learnable in any of these senses. Furthermore, as the set of ψ-indices of partial functions with infinite domain is immune, by Theorem 17, every Bc_ψ-learnable class is also Vac_ψ-learnable. Hence, $\mathcal{S}_{const}$ is also not Bc_ψ-learnable. This shows that ψ is not optimal for any of the learning criteria Fin, Ex, Vac, Bc. One also obtains that the numbering is not Bc-acceptable (otherwise, by Proposition 12, $\mathcal{S}_{const}$ would be Bc-learnable, contradicting the proof above). □ (**Theorem 25**)

Theorem 26. *There exists a numbering ψ satisfying (a)–(d) 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;*
(d) *ψ is* Ex_1*-acceptable.*

Proof. Let S be a set which is hypersimple relative to K. For example, one could take S to be the set $\{e : \vartheta_e \text{ has finite domain}\}$ used in the proof of Lemma 14 and Theorem 15. Let $\lambda t.S_t$ 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,

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

(a) 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 $e \in E$, $\varphi_e = \alpha(\langle e, \cdot\rangle)$. Now it is shown 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 $e \in E$, $\varphi_e(0)\downarrow \geq e$. 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 $n \notin S$ be sufficiently large so that $n > C(\varphi_{e_0}(0))$ and $n > c$. Let i be largest such that $C(\varphi_{e_i}(0)) \leq n^2$. Then, $n^2 \leq C(\varphi_{e_{i+1}}(0)) \leq n^2 + n$. 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.

Note that the above analysis also shows that ψ is a numbering for $\mathcal{P}$, as for every partial recursive function η, there exists an α such that $\alpha(e, \cdot) = \eta$.

(b) 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 $e_{k,r} = M(p_k, r)$. Then, for all but finitely many i, $\psi_{e_{k,i}} = \lambda x \centerdot e_{k,i} + k$. Furthermore, it follows from the definition of ψ that, for all but finitely many i, there exists an $n_{k,i} \notin S$ such that $(n_{k,i})^2 \leq C(e_{k,i} + k) \leq (n_{k,i})^2 + n_{k,i}$. Note that, as S is simple relative to K, $\{e_{k,0}, e_{k,1}, e_{k,2}, \ldots\}$ is a finite set, that is, there exists some j such that $e_{k,0}, e_{k,1}, \ldots, e_{k,j}$ represents *all* of the indices in $\{M(p_k, s') : s' \in \mathbb{N}\}$. Let $f : \mathbb{N} \to \mathbb{N}$ be such that, for each k, $D_{f(k)}$ is the collection of all $n_{k,i}$, $i \leq j$, corresponding to $e_{k,i}$, which are defined (where $n_{k,i}$ and $e_{k,i}$ are as just described, that is, $n_{k,i}$ satisfies $n_{k,i} \notin S$ and $(n_{k,i})^2 \leq C(e_{k,i} + k) \leq (n_{k,i})^2 + n_{k,i}$; note that some $n_{k,i}$ may not exist). 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,

$$g(0) = f(0); \tag{38}$$

$$g(i+1) = f(k), \text{ where } k \text{ is least such that } D_{f(k)} \cap (D_{g(0)} \cup D_{g(1)} \cup \cdots \cup D_{g(i)}) = \emptyset. \tag{39}$$

Note that g satisfies the following statements: g is K-recursive; for each i and j, $D_{g(i)}$ and $D_{g(j)}$ are disjoint; for each i, $D_{g(i)}$ is not a subset of S. This contradicts the fact that S is hypersimple relative to K. Hence, M does not witness that ψ is BcKrt-numbering, and, more generally, ψ is not a BcKrt-numbering.

(c) 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.

(d) The numbering is also Ex_1-acceptable. Given an index p of a function φ_p, the translation f first keeps outputting $f(p, s) = p$ until a stage s is reached where $\varphi_{p,s}(0)$ is defined and smaller than s. For all $t > s$, the value $f(p, t)$ is $pad(p, s)$. $\square$ (**Theorem 26**)

7 Conclusion

Kleene's Recursion Theorem is a great mathematical tool for proving theorems. Proofs using KRT tend to be elegant, succinct and insightful. Specifically this occurs in learning theory where recursion theorems have often been used to prove hierarchies and separations of learning criteria [JORS99]. Thus, it has been interesting herein to explore how presence or absence of KRT in the hypothesis space (whether possessed of some degree of constructivity or not) affects learnability and its complexity, when measured in the type of convergence which can be achieved. It was shown that, in general, presence or absence of constructive or not-so-constructive KRT is independent of whether the numbering is optimal for various learning criteria. Hence, (constructive or otherwise) KRT is not necessary for optimality for various learning criteria. Thus, at least regarding learnability, the value of KRT is elegance, succinctness and insight rather than raw necessity.

Acknowledgment. The authors would like to thank Samuel E. Moelius III for discussions and support in writing this paper. The authors would like to thank the anonymous referees for various comments which helped to improve the paper. A preliminary version of this paper appeared in the conference LFCS 2013 [CJS13].

References

[ABG11] Klaus Ambos-Spies, Serikzhan Badaev and Sergey Goncharov. Inductive inference and computable numberings. *Theoretical Computer Science*, 412:1652–1668, 2011.

[BG00] Serikzhan Badaev and Sergey Goncharov. The theory of numberings: open problems. *Computability Theory and its Applications. Current Trends and Open Problems. Contemporary Mathematics.* Volume 257, American Mathematical Society, Providence, pages 23–38, 2000.

[BGS03] Serikzhan Badaev, Sergey Goncharov and Andrea Sorbi. Completeness and universality of arithmetical numberings. *Computability and Models*, Kluwer / Plenum Publishers, New York, pages 11–44, 2003.

[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.

[Cal02] Cristian S. Calude. *Information and Randomness: An Algorithmic Perspective.* Second edition, Springer, Heidelberg, 2002.

[Cas74] John Case. Periodicity in generations of automata. *Mathematical Systems Theory*, 8:15–32, 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.

[CJS13] John Case, Sanjay Jain and Frank Stephan. Effectivity questions for Kleene's recursion theorem. In *Symposium on Logical Foundations in Computer Science 2013 (LFCS 2013)*, volume 7734 of *Lecture Notes in Computer Science*, pages 89–103, Springer, 2013.

[CM09] John Case and Samuel E. Moelius III. Program self-reference in constructive Scott subdomains. Theory of Computing Systems (special issue for CiE 2009), 51:22–49, 2012.

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

[DH10] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity.* Theory and Applications of Computability. Springer, New York, 2010.

[Ersh77] Yuri Ershov. *Theory of Numberings.* Nauka, Moscow, 1977.

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

[Hay06] Susumu Hayashi. Mathematics based on incremental learning — Excluded middle and inductive inference. *Theoretical Computer Science*, 350(1):125–139, 2006.

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

[JORS99] Sanjay Jain, Daniel Osherson, James Royer and Arun Sharma. *Systems that Learn: An Introduction to Learning Theory.* MIT Press, Cambridge, Mass., second edition, 1999.

[JS94] Sanjay Jain and Arun Sharma. Characterizing language identification by standardizings operations. *Journal of Computer and System Sciences*, 49(1):96–107, 1994.

[JS97] Sanjay Jain and Arun Sharma. Characterizing language identification in terms of computable numberings. *Annals of Pure and Applied Logic*, 84(1):51–72, 1997.

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

[JST11] Sanjay Jain, Frank Stephan and Jason Teutsch. Index sets and universal numberings. *Journal of Computer and System Sciences* 77:760–773, 2011.

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

[Kle52] Stephen C. Kleene. *Introduction to Metamathematics.* North-Holland, Amsterdam, 1952.

[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.

[Odi89] Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland Publishing Co., Amsterdam, 1989.

[Odi99] Piergiorgio Odifreddi. *Classical Recursion Theory, Vol II.* Elsevier, Amsterdam, 1999.

[OSW86] Daniel N. Osherson, Michael Stob and Scott Weinstein. *Systems That Learn, An Introduction to Learning Theory for Cognitive and Computer Scientists.* Bradford — The MIT Press, Cambridge, Massachusetts, 1986.

[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, Heidelberg, 1987.

[Soa87] Robert I. Soare. *Recursively Enumerable Sets and Degrees.* Springer, Heidelberg, 1987.

[ST12] Frank Stephan and Jason Teutsch. An incomplete set of shortest descriptions. *The Journal of Symbolic Logic*, 77:291-307, 2012.

[ST14] Frank Stephan and Jason Teutsch. Things which can be made into themselves. *Information and Computation*, 237:174–186, 2014.