Kolmogorov Numberings and Minimal Identification

Rusins Freivalds*
Institute of Mathematics and Computer Science
University of Latvia
Raiņa bulvāris 29, LV-1459, Riga, Latvia
Email: rusins@mii.lu.lv

Sanjay Jain
Department of Information Systems and Computer Science
National University of Singapore
Singapore 0511
Republic of Singapore
Email: sanjay@iscs.nus.sg

Abstract

Identification of programs for computable functions from their graphs by algorithmic devices is a well studied problem in learning theory. Freivalds and Chen consider identification of 'minimal' and 'nearly minimal' programs for functions

*Supported in part by Latvian Council of Science Grant 93.599 and NSF Grant 9119540. Work done in part while the author was visiting the Institute of Systems Science, at the National University of Singapore.

from their graphs. To address certain problems in minimal identification for Gödel numberings, Freivalds later considered minimal identification in Kolmogorov Numberings. Kolmogorov numberings are in some sense optimal numberings and have some nice properties. We prove certain hierarchy results for minimal identification in every Kolmogorov numbering. In addition we also compare minimal identification in Gödel numbering versus minimal identification in Kolmogorov numberings.

1 Introduction

Let $N = \{0, 1, 2, \ldots\}$, the set of natural numbers. Let f be a function : $N \to N$. For any $n \in N$, we let $f[n]$ denote $\{(x, f(x)) \mid x < n\}$, the finite initial segment of f consisting of the first n data points in the graph of f. The quantifier '$\stackrel{\infty}{\forall}$' means 'for all but finitely many natural numbers'. Criteria of inference informally described below are formally defined in Section 3.

A *function learning machine* $\mathbf{M}$ is an algorithmic device which, on any input segment $f[n]$, returns either ? or a program, where, if it returns a program on a segment, it returns a program on all extensions of that segment. The output of $\mathbf{M}$ on $f[n]$ is denoted by $\mathbf{M}(f[n])$. If $\mathbf{M}(f[n])$ is a program, we think of that program as $\mathbf{M}$'s conjecture, based on the data $f[n]$, about how to compute all of f; $\mathbf{M}(f[n]) =?$ then represents the situation where $\mathbf{M}$ does not conjecture a program based on the data $f[n]$. The restriction that $\mathbf{M}$ must continue to conjecture programs once it has done so is essentially without loss of generality since a machine which hasn't had enough time to think of a new conjecture can be thought of re-outputting its previous conjecture.

As is by now well known, there are various senses in which $\mathbf{M}$ can be thought of as *successfully* learning or inferring a program for f. For $n \in N$, let $p_n = \mathbf{M}(f[n])$. The criterion of success known as **Ex**-*identification* [Gol67, BB75, CS83] requires that the sequence $p_0, p_1, p_2, \ldots$ contains a program p, *which computes* f , such that $(\stackrel{\infty}{\forall} i)[p_i = p]$. In this case one speaks of p as being the *final* program output by $\mathbf{M}$ on f.

Freivalds [Fre75] and Chen [Che81, Che82] studied the effect of requiring that the final hypothesis held by the learner in the above model be of (nearly) minimal size. Minimal identification of classes of functions depends on the acceptable programming system (acceptable numbering) chosen to interpret programs output by learning machines. Suppose ψ is a numbering (programming system). In $\mathbf{Min}_\psi\mathbf{Ex}$-identification criterion one requires, in addition to $\mathbf{Ex}$-identification of function (in the programming system ψ), that the final programs be of minimal size (see formal definitions in Section 3). We direct the reader to [Fre75, Che81, Che82] for results dealing with minimal identification and its relationship with $\mathbf{Ex}$-identification.

Freivalds [Fre75] showed that there are Gödel numberings in which no function class of infinite cardinality can be learned by minimal programs. This led Freivalds to consider minimal identification in Kolmogorov numberings. Kolmogorov numberings are numberings to which every other numbering is reducible by a linearly bounded function. Freivalds showed that for every Kolmogorov numbering ψ, there exists an infinite class of functions in $\mathbf{Min}_\psi\mathbf{Ex}$. In this paper we consider the relationship between identification criteria such as $\mathbf{Min}_\psi\mathbf{Ex}$, $\mathbf{Min}_\psi\mathbf{FIN}$ and $\mathbf{Min}_\psi\mathbf{CoLearn}$ in Kolmogorov numberings ψ (see formal definitions in Section 3). We also show the existence of classes of functions which can be minimally identified in some Gödel numbering but cannot be minimally identified in any Kolmogorov numbering.

We now proceed formally.

2 Notation

Recursion-theoretic concepts not explained below are treated in [Rog67]. N denotes the set of natural numbers, $\{0, 1, 2, \ldots\}$; The symbols c, e, i, j, k, l, m, n, p, r, s, t, u, v, x, y and z, with or without decorations (decorations are subscripts, superscripts and the like), range over natural numbers unless otherwise specified. $\subseteq, \subset, \supseteq, \supset, \in$, denote subset, proper subset, superset, proper superset and element of relationship respectively. $\emptyset$ denotes the empty set. A, C, L, S, with or without decorations, range over subsets of

N. $\overline{L}$ denotes the complement of L, i.e. $\overline{L} = N - L$. We denote the cardinality of a set S by card(S). max( ), min( ) denote the maximum and minimum of a set, respectively. By convention $\max(\emptyset) = 0$ and $\min(\emptyset) = \infty$.

$\mathcal{R}$ denotes the set of all total recursive functions. h, f, g, q, with or without decorations, range over $\mathcal{R}$. $\mathcal{C}$ and $\mathcal{S}$, with or without decorations, range over subsets of $\mathcal{R}$. $\downarrow$ denotes defined. $\uparrow$ denotes undefined.

A programming system (or computable numbering) is a (partial) computable function of two variables. We often drop the word 'computable' from 'computable numbering' in this paper, since we will be dealing with computable numberings only. We let ψ, β, η range over computable numberings (programming systems). Suppose $\psi(\cdot, \cdot)$ is a computable numbering. We often refer to the (partial) function $\lambda x.\psi(i, x)$ as ψ_i. ψ_i thus denotes the (partial) function computed by the i-th program in the numbering ψ.

An acceptable numbering is a computable numbering to which every other computable numbering is reducible via a recursive function. Thus if ψ is an acceptable numbering, then for all computable numberings η, there exists a recursive function h, such that $(\forall i)[\eta_i = \psi_{h(i)}]$. Acceptable numberings are also referred to as Gödel numberings. Kolmogorov numbering is an acceptable numbering to which every other computable numbering can be reduced via a linearly bounded function. Thus is ψ is a Kolmogorov numbering, then for all computable numberings η, there exists a recursive function h and a constant c, such that $(\forall i)[\eta_i = \psi_{h(i)} \ \wedge \ h(i) \leq \max(\{c * i, c\})]$.

For a function f, $\mathrm{MinProg}_\psi(f)$ denotes the minimal program for f, if any, in the ψ programming system, i.e., $\mathrm{MinProg}_\psi(f) = \min(\{i \mid \psi_i = f\})$. Let **ZEROSTAR** $= \{f \mid (\stackrel{\infty}{\forall} x)[f(x) = 0]\}$.

We let φ denote a standard acceptable programming system. φ_i thus denotes the partial recursive function computed by the i^{th} program in the standard acceptable programming system φ. We often refer to the i^{th} program as program i. Φ denotes an arbitrary fixed Blum complexity measure [Blu67, HU79] for the φ-system.

$\langle i, j\rangle$ stands for an arbitrary computable one to one encoding of all pairs of natural numbers onto N [Rog67]. We assume that the pairing function is such that $\langle i, j\rangle \geq \max(i, j)$.

The quantifier '$\stackrel{\infty}{\exists}$' means 'there exist infinitely many'.

3 Learning Paradigms

For any partial function η and any natural number n such that, for each $x < n$, $\eta(x)\downarrow$, we let $\eta[n]$ denote the finite initial segment $\{(x, \eta(x)) \mid x < n\}$. Let INIT $= \{f[n] \mid f \in \mathcal{R} \wedge n \in N\}$. We let σ and τ, with or without decorations, range over INIT. $|\sigma|$ denotes the length of σ. Thus for example $|f[n]| = n$.

For all $\sigma \in$ INIT, zeroext(σ) denotes a function such that

$$\text{zeroext}(\sigma)(x) = \begin{cases} \sigma(x), & \text{if } x \leq |\sigma|; \\ 0, & \text{otherwise.} \end{cases}$$

Definition 1 [Gol67] A *learning machine* is an algorithmic device which computes a mapping from INIT into $N \cup \{?\}$ such that, if $\mathbf{M}(f[n]) \neq ?$, then $\mathbf{M}(f[n+1]) \neq ?$.

We let $\mathbf{M}$, with or without superscripts, range over learning machines. (We reserve $\mathbf{M}$ with subscripts for special type of enumeration of learning machines. See Proposition 1 below.) In Definition 1 above, '?' denotes the situation when $\mathbf{M}$ outputs "no conjecture" on some $\sigma \in$ INIT.

In Definition 2 below we spell out what it means for a learning machine to converge in the limit.

Definition 2 Suppose $\mathbf{M}$ is a learning machine and f is a computable function. $\mathbf{M}(f)\downarrow$ (read: $\mathbf{M}(f)$ *converges*) just in case $(\exists i)(\stackrel{\infty}{\forall} n)\ [\mathbf{M}(f[n]) = i]$. If $\mathbf{M}(f)\downarrow$, then $\mathbf{M}(f)$ is defined $=$ the unique i such that $(\stackrel{\infty}{\forall} n)[\mathbf{M}(f[n]) = i]$, otherwise we say that $\mathbf{M}(f)$ diverges (written: $\mathbf{M}(f)\uparrow$).

3.1 Explanatory Function Identification

We now formally define the criteria of inference considered in this paper.

Definition 3 [Gol67, CS83]

(a) A learning machine $\mathbf{M}$ is said to $\mathbf{Ex}$-*identify* f (written: $f \in \mathbf{Ex}(\mathbf{M})$) just in case $(\exists i \mid \varphi_i = f)(\overset{\infty}{\forall} n)[\mathbf{M}(f[n]) = i]$.

(b) $\mathbf{Ex} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{Ex}(\mathbf{M})]\}$.

3.2 Finite function Inference

Definition 4

(a) A learning machine $\mathbf{M}$ is said to $\mathbf{FIN}$-*identify* f (written: $f \in \mathbf{FIN}(\mathbf{M})$) just in case $(\exists n, p \mid \varphi_p = f)[(\forall m < n)[\mathbf{M}(f[m]) =?] \ \wedge \ (\forall m \geq n)[\mathbf{M}(f[m]) = p]]$.

(b) $\mathbf{FIN} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{FIN}(\mathbf{M})]\}$.

3.3 Colearnability

We say that $\mathbf{M}(f)$ co-converges to p, iff $[N - \{\mathbf{M}(f[m]) \mid m \in N\} = \{p\}]$. If there exists a p such that $\mathbf{M}(f)$ co-converges to p, then we say that $\mathbf{M}$ co-converges on f (to p). Otherwise we say that $\mathbf{M}$ co-diverges on f.

Definition 5

(a) A learning machine $\mathbf{M}$ is said to $\mathbf{CoLearn}$-*identify* f (written: $f \in \mathbf{CoLearn}(\mathbf{M})$) just in case $(\exists p \mid \varphi_p = f)[\mathbf{M}$ co-converges on f to $p]$.

(b) $\mathbf{CoLearn} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{CoLearn}(\mathbf{M})]\}$.

3.4 Minimal Identification

We next consider identification by nearly minimal programs. Minimal identification usually depends on the numbering system chosen.

Definition 6 [Fre75] Suppose ψ is a numbering.

(a) $\mathbf{M}$ $\mathbf{Min}_\psi\mathbf{Ex}$-identifies f (written $f \in \mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M})$) iff $\mathbf{M}(f)\!\downarrow = \mathrm{MinProg}_\psi(f)$.

(b) $\mathbf{Min}_\psi\mathbf{Ex} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M})]\}$.

(c) $\mathbf{M}$ $\mathbf{Min}_\psi\mathbf{FIN}$-identifies f (written $f \in \mathbf{Min}_\psi\mathbf{FIN}(\mathbf{M})$) iff $(\exists n)[(\forall m < n)[\mathbf{M}(f[n]) = ?] \ \wedge \ (\forall m \geq n)[\mathbf{M}(f[n]) = \mathrm{MinProg}_\psi(f)]]$.

(d) $\mathbf{Min}_\psi\mathbf{FIN} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{Min}_\psi\mathbf{FIN}(\mathbf{M})]\}$.

(e) $\mathbf{M}$ $\mathbf{Min}_\psi\mathbf{CoLearn}$-identifies f (written $f \in \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M})$) iff $\mathbf{M}(f)$ co-converges to $\mathrm{MinProg}_\psi(f)$.

(f) $\mathbf{Min}_\psi\mathbf{CoLearn} = \{\mathcal{C} \mid (\exists \mathbf{M})[\mathcal{C} \subseteq \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M})]\}$.

The following proposition facilitates the proof of some of our results.

Proposition 1 *There exists a recursively enumerable sequence* $\mathbf{M}_0, \mathbf{M}_1, \ldots$ *of learning machines such that, for all machine* $\mathbf{M}$ *and computable numberings* ψ*, there exists an* i *such that*

$\mathbf{Ex}(\mathbf{M}) \subseteq \mathbf{Ex}(\mathbf{M}_i)$,

$\mathbf{FIN}(\mathbf{M}) \subseteq \mathbf{FIN}(\mathbf{M}_i)$,

$\mathbf{CoLearn}(\mathbf{M}) \subseteq \mathbf{CoLearn}(\mathbf{M}_i)$,

$\mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M}) \subseteq \mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M}_i)$,

$\mathbf{Min}_\psi\mathbf{FIN}(\mathbf{M}) \subseteq \mathbf{Min}_\psi\mathbf{FIN}(\mathbf{M}_i)$,

$\mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}) \subseteq \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}_i)$.

For a proof of the above proposition see for example [OSW86]. We let $\mathbf{M}_0, \mathbf{M}_1, \ldots$ be one such enumeration.

4 Results

Theorem 1 $(\forall$ *Kolmogorov Numbering* $\psi)[\mathbf{Min}_\psi\mathbf{FIN} - \mathbf{FINITE} \neq \emptyset]$.

PROOF. This is essentially the proof used by Freivalds [Fre91] to show that for every Kolmogorov numbering ψ, there exists an infinite class of functions in $\mathbf{Min}_\psi\mathbf{Ex}$. We give the proof for completeness. Suppose a Kolmogorov numbering ψ is given. For $i > 0$, let h_i be defined as follows.

$$h_i(x) = \begin{cases} i, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Let c be such that, for all $j > 0$, there exists a $k \leq c * j$, such that $\psi_k = h_j$ (since ψ is Kolmogorov numbering there exists such a c). Let $\mathcal{S} = \{h_j \mid j > 0 \ \wedge \ \text{card}(\{i \leq c * j \mid \psi_i(0) = j\}) \leq 2c\}$. Note that $\mathcal{S}$ is infinite. Let $C_j = \{k \leq c * j \mid \psi_k(0) = j\}$. Now for $1 \leq i \leq 2c$, define $\mathbf{M}^i$ as follows. The only program (if any) output by $\mathbf{M}^i$ on h_j, is the i-th element, if any, in some standard (effective in j) enumeration of C_j. It is easy to see that each function in $\mathcal{S}$ is $\mathbf{Min}_\psi\mathbf{FIN}$ identified by at least one machine in $\mathbf{M}^1, \ldots, \mathbf{M}^{2c}$. Since $\mathcal{S}$ is infinite the theorem follows by Pigeon hole principle. ∎

Theorem 2 $(\forall$ *Kolmogorov Numbering* $\psi)[\mathbf{Min}_\psi\mathbf{CoLearn} - \mathbf{FIN} \neq \emptyset]$.

PROOF. Suppose a Kolmogorov numbering ψ is given. Let h_0 be everywhere 0 function. For $i > 0$, define h_i as follows.

$$h_i(x) = \begin{cases} 1, & \text{if } x = i; \\ 0, & \text{otherwise.} \end{cases}$$

We will construct a class of functions $\mathcal{C} \in \mathbf{Min}_\psi\mathbf{CoLearn}$, such that $h_0 \in \mathcal{C}$ and for infinitely many k, $h_k \in \mathcal{C}$. Since no such $\mathcal{C}$ can be in $\mathbf{FIN}$, this would prove the theorem.

Let $z = \text{MinProg}_\psi(h_0)$. Note that since ψ is Kolmogorov numbering, there exists a c such that for all $k > 0$, $\text{MinProg}_\psi(h_k) \leq c \cdot k$. For $j > 0$, let $C_j = \{k \leq c \cdot j \mid \psi_k(j) = 1 \ \wedge \ (\forall x < j)[\psi_k(x) = 0]\}$. Let p_k^l denote the l-th program enumerated, if any, in some standard (effective in k) enumeration of C_k.

For $l \in N$, consider machine $\mathbf{M}^l$, such that the following conditions are satisfied. Let c' be a constant > 2. (It is easy to construct such $\mathbf{M}^l$.)

(1) $\{\mathbf{M}^l(h_0[x]) \mid x \leq y\} = \{i \leq y/c'\} - \{z\}$.

(note that this implies that $h_0 \in \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}^l)$.)

(2) $\{\mathbf{M}^l(h_k[y+k+1]) \mid y \in N\} = \{z\} \cup \{i \in N\} - \{p_k^l\}$, if $l \leq \mathrm{card}(C_k)$;

$\{\mathbf{M}^l(h_k[k+1+y]) \mid y \in N\} = \{z\}$, if $l > \mathrm{card}(C_k)$.

Now let $S = \{k > 0 \mid \mathrm{card}(\{p \leq c \cdot k \mid h_k[k+1] \subseteq \psi_p\}) \leq 2c\} - \{k > 0 \mid \mathrm{MinProg}_\psi(h_k) \leq k/c'\}$. Note that S is of infinite cardinality (since $c' > 2$). Furthermore for each $k \in S$, there exists $l \leq 2c$, such that $h_k \in \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}^l)$. It follows that there exists an $l \leq 2c$, such that $\mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}^l)$ contains an infinite subset of $\{h_k \mid k > 0\}$. Since $h_0 \in \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}^l)$, it follows that there exists a $\mathcal{C} \in \mathbf{Min}_\psi\mathbf{CoLearn}$, such that $h_0 \in \mathcal{C}$, and for infinitely many k, $h_k \in \mathcal{C}$, as claimed. ∎

As a corollary we have

Corollary 1 $(\forall$ *Kolmogorov Numbering* $\psi)(\exists\mathcal{C})[\mathcal{C} \in \mathbf{Min}_\psi\mathbf{CoLearn} \wedge (\forall$ *Gödel Numbering* $\eta)[\mathcal{C} \notin \mathbf{Min}_\eta\mathbf{FIN}]]$.

Theorem 3 $(\forall$ *Kolmogorov Numbering* $\psi)(\exists\mathcal{C})[\mathcal{C} \in \mathbf{Min}_\psi\mathbf{Ex} \wedge (\forall$ *Kolmogorov Numbering* $\psi')[\mathcal{C} \notin \mathbf{Min}_{\psi'}\mathbf{CoLearn}]]$

PROOF. Let $l_0^0 = 0$, $u_0^0 = 0$. For $i > 0$, let $l_i^0 = u_{i-1}^{i-1} + 1$; For $1 \leq j \leq i$, let $l_i^j = u_i^{j-1} + 1$. For $i > 0, j \leq i$, let $u_i^j = (l_i^j + i) * 2$. We define a numbering η (which is not a Gödel numbering) as follows.

For $l_i^0 \leq k \leq u_i^i$, we define η_k according to the following process (note that for each i, a different such staging construction is executed). Let $\sigma_i^0 = \{(0, i)\}$. Go to stage 0.

Stage s

For $l_i^s \leq k \leq u_i^s$, let

$$\eta_k(x) = \begin{cases} \sigma_i^s(x), & \text{if } x < |\sigma_i^s|; \\ k, & \text{if } x = |\sigma_i^s|; \\ 0, & \text{otherwise.} \end{cases}$$

Search for a τ properly extending σ_i^s, such that, for some k, $l_i^s \leq k \leq u_i^s$, $\tau \subseteq \eta_k$, and

$\text{card}(\{l < i \mid \text{card}(\{x \leq (u_i^i)^2\} - \{\mathbf{M}_l(\tau') \mid \tau' \subseteq \tau\}) \leq 1\}) \geq s+1.$

If and when such a τ is found let $\sigma_i^{s+1} = \tau$, and go to stage $s+1$.

End Stage s

Note that for any i, the last stage executed is $\leq i$. For any i, let s_i, denote the last stage executed in the above construction corresponding to i. Also, note that each η_k is either an empty function or in **ZEROSTAR**. Moreover, if η_k is in **ZEROSTAR**, then it is distinct from every other η_l, $l \neq k$.

By construction, it follows that $\text{card}(\{k \mid l_i^{s_i} \leq k \leq u_i^{s_i} \ \wedge \ (\exists l < i)[\mathbf{M}_l$ on η_k co-converges to a program $\leq (u_i^i)^2\}) \leq i$.

Now consider any Kolmogorov Numbering ψ. Now, clearly there exists a c, such that for all $k > 0$, $\text{MinProg}_\psi(\eta_k) \leq c \cdot k$. Let $\mathcal{S} = \{\eta_k \mid (\exists i)[l_i^{s_i} \leq k \leq u_i^{s_i} \ \wedge \ (\forall l < i)[\eta_k \notin \mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M}_l)] \ \wedge \ \text{card}(\{p \leq c \cdot k \mid \eta_k[|\sigma_i^{s_i}|+1] \subseteq \psi_p\}) \leq 3c]\}$.

It is easy to verify that $\mathcal{S}$ is infinite (for $u_i^i > c$, at least $2u_i^{s_i}/3 - l_i^{s_i} - i$ of the functions in $\{\eta_k \mid l_i^{s_i} \leq k \leq u_i^{s_i}\}$ belong to $\mathcal{S}$).

By definition of $\mathcal{S}$, $\mathbf{Min}_\psi\mathbf{CoLearn}(\mathbf{M})$ contains at most finitely many functions in $\mathcal{S}$.

For $l_i^{s_i} \leq k \leq u_i^{s_i}$, let $C_k = \{l \leq c * k \mid \eta_k[|\sigma_i^{s_i}|+1] \subseteq \psi_l\}$. For $1 \leq l \leq 3c$, let $\mathbf{M}^l$ be such that, for $l_i^{s_i} \leq k \leq u_i^{s_i}$, such that η_k is total, $\mathbf{M}^l(\eta_k)$ converges to the l-th element in C_k, in some standard (effective in k) enumeration of C_k. It is easy to observe that each $\eta_k \in \mathcal{S}$, is in $\mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M}_l)$, for some $l \leq 3c$. Thus there exists an infinite subset of $\mathcal{S}$, which is in $\mathbf{Min}_\psi\mathbf{Ex}$. ∎

Theorem 4 *$(\exists$ infinite r.e. $\mathcal{C})(\exists$ Kolmogorov Numbering $\psi)[\mathcal{C} \in \mathbf{Min}_\psi\mathbf{Ex}]$.*

PROOF. It was shown in [Fre91] that there exists a Kolmogorov numbering ψ such that $\{f \mid (\forall x)[f(x) = f(0)]\} \in \mathbf{Min}_\psi\mathbf{Ex}$. In fact it can be shown that for all $\mathcal{C} \in \mathbf{FIN}$, there exists a Kolmogorov numbering ψ, such that $\mathcal{C} \in \mathbf{Min}_\psi\mathbf{Ex}$. ∎

Theorem 5 *($\forall$ infinite r.e. $\mathcal{C}$)($\forall$ Gödel Numbering ψ)*$[\mathcal{C} \notin \mathbf{Min}_{\psi}\mathbf{CoLearn}]$.

PROOF. Suppose an r.e. infinite class $\mathcal{C}$, a Gödel Numbering ψ, $\mathbf{M}$, are given such that $\mathcal{C} \in \mathbf{Min}_{\psi}\mathbf{CoLearn}(\mathbf{M})$. It follows that, for all c, there exists an $f \in \mathcal{C}$ such that $\{\mathbf{M}(f[n]) \mid n \in N\} \supseteq \{x \mid x \leq c\}$. Since one can search for such a function ($\mathcal{C}$ is r.e.), we have, by implicit use Kleene's recursion theorem [Rog67], that there exists e, such that $\varphi_e \in \mathcal{C}$, and $\{\mathbf{M}(\varphi_e[n]) \mid n \in N\} \supseteq \{x \mid x \leq e\}$. But then $\varphi_e \notin \mathbf{Min}_{\psi}\mathbf{CoLearn}(\mathbf{M})$. A contradiction. ∎

Theorem 6 *($\exists$ Kolmogorov Numbering ψ)($\exists$ infinite co-r.e. L)*$[\{\psi_i \mid i \in L\} \in \mathbf{Min}_{\psi}\mathbf{FIN}]$.

PROOF. Let f_k be defined as follows.

$$f_k(x) = \begin{cases} k, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Recall from Section 2 that φ is a Komogorov numbering. Let ψ be defined as follows. $\psi_{3i} = \varphi_i$. For j not divisible by 3, let $\psi_j = f_j$; Let $\overline{L} = \{3i \mid i \in N\} \cup \{j \mid j$ is not divisible by 3 $\wedge$ $(\exists i < j/3)[\varphi_i(0) = j]\}$. Note that $\overline{L}$ is r.e. Consider $\mathbf{M}$, which on f_k, outputs k as its only program.

It is easy to see that $\mathcal{C} = \{\psi_j \mid j \in L\} \subseteq \mathbf{Min}_{\psi}\mathbf{FIN}(\mathbf{M})$. ∎

Theorem 7 *($\exists$ Gödel Numbering ψ)($\exists\mathcal{C}$)*$[\mathcal{C} \in \mathbf{Min}_{\psi}\mathbf{FIN} \wedge$ *($\forall$ Kolmogorov Numbering ψ')*$[\mathcal{C} \notin \mathbf{Min}_{\psi'}\mathbf{CoLearn}]$.

PROOF. Let h_i be a function defined as follows.

$$h_i(x) = \begin{cases} i, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\psi_{l^4} = \varphi_l$. For i not of the form l^4 for some l, let $\psi_i = h_i$.

Consider $\mathbf{M}'$, which on h_i, outputs i as its only program. Let $\mathcal{C} = \{h_i \mid \mathrm{MinProg}_{\psi}(h_i) = i\}$.

Clearly, $\mathcal{C} \in \mathbf{Min}_{\psi}\mathbf{FIN}(\mathbf{M}')$. Suppose $\mathbf{M}$ and a Kolmogorov numbering ψ is given. By Operator recursion theorem [Cas74], there exists a recursive function q, such that (partial) functions $\varphi_{q(\cdot)}$ may be defined as follows.

Let $l_0 = 0$. l_s is used to denote the least l such that $\varphi_{q(l)}$ has not been defined on any input before stage s.

Go to stage 0.

Begin stage s

1. Search for an $r > l_s$, $S \subseteq \{x \mid r^4 < x < (r+1)^4\}$, such that $\text{card}(S) = r+2$ and $\{x \mid x \leq (l_s + r + 2)^2\} \subseteq \{\mathbf{M}(h_i[n]) \mid n \in N\}$.
2. Let S be as found in step 1. Let $\varphi_{q(l_s+x)}$, $x \leq r+1$, be the $r+2$ functions in $\{h_i \mid i \in S\}$.
3. Let $l_{s+1} = l_s + r + 2$.

End stage s

Now we consider the following cases.

Case 1: Stage s starts but does not halt.

For each r, at least $[(r+1)^4 - r^4 - 1] - [r+1]$ of the functions $\{h_i \mid r^4 < i < (r+1)^4\}$, are in $\mathcal{C}$. This is bigger than $(2r+2)^2 + r + 3$, for large enough r. Thus non-success of step 1 in stage s implies that, $\mathcal{C} \not\subseteq \mathbf{Min}_{\psi'}\mathbf{CoLearn}(\mathbf{M})$ (in fact $\mathbf{M}$ does not even $\mathbf{CoLearn}$, $\mathcal{C}$).

Case 2: All stages halt.

In this case note that for each stage, at least one of the $r+2$ functions found in step 1 is in $\mathcal{C}$ (since at most $r+1$ of the $r+2$ functions can be spoiled by ψ_{x^4}, $x \leq r$, and thus not be in $\mathcal{C}$). It follows, that $(\stackrel{\infty}{\exists} i)[\varphi_{q(i)} \in \mathcal{C} \wedge \mathbf{M}$ on $\varphi_{q(i)}$ does not co-converge to a program $\leq i^2]$. But then, since ψ' is a Kolmogorov Numbering , $\mathcal{C} \not\subseteq \mathbf{Min}_{\psi'}\mathbf{CoLearn}(\mathbf{M})$. ∎

Theorem 8 $(\exists \mathcal{C})(\exists$ *Gödel Numbering* $\psi)[\mathcal{C} \in \mathbf{Min}_{\psi}\mathbf{Ex} \wedge (\forall$ *Kolmogorov Numbering* $\eta)[\mathcal{C} \notin \mathbf{Min}_{\eta}\mathbf{Ex}]]$.

Proof. By parameterized s-m-n theorem [Rog67] there exists a recursive 1-1 function g such that $\varphi_{g(k,i,j)}$, may be defined as follows. Let F be a partial recursive function, such that $F(k,i,x)$, denotes the output of the i-th program in the k-th numbering on input x.

Begin definition of $\varphi_{g(k,i,j)}$

Let Cancel $= \emptyset$.

Let $\varphi_{g(k,i,j)})(0) = \langle k,i,j\rangle$.

Let x_s denote the least x such that $\varphi_{g(k,i,j)}(x)$ has not been defined before stage s.

Go to stage 0.

Begin stage s

0. Dovetail steps 1 and 2, until one of them succeeds. If step 1 succeeds (before step 2, if ever) then go to step 3. If step 2 succeeds (before step 1, if ever) then go to step 4.

1. Search for an $l < j^2$, such that $l \notin$ Cancel and $F(k,l,x_s)\downarrow$.

2. Suppose $f = \text{zeroext}(\varphi_{g(k,i,j)}[x_s])$. Search for $m > x_s$, such that $\mathbf{M}_i(f[m]) \geq j^2$.

3. Let l be as found in step 1. Let $\varphi_{g(k,i,j)}(x_s) = F(k,l,x_s)+1$.

   Let Cancel $=$ Cancel $\cup \{l\}$.

   Go to stage $s+1$.

4. Let m be as found in step 2. For $x_s \leq x \leq m$, let $\varphi_{g(k,i,j)}(x) = 0$.

   Go to stage $s+1$.

End stage s

End definition of $\varphi_{g(k,i,j)}$

Now define $\mathcal{C}_{\langle k,i,j\rangle}$ as follows. $\mathcal{C}_{\langle k,i,j\rangle} = \{\text{zeroext}(\varphi_{g(k,i,j)}[x]) \mid x = 1 \vee \varphi_{g(k,i,j)}(x-1)\downarrow \neq 0\}$. Note that $\text{zeroext}(\varphi_{g(k,i,j)}) \in \mathcal{C}_{\langle k,i,j\rangle}$, and $\text{card}(\mathcal{C}_{\langle k,i,j\rangle}) \leq j^2+1$ (since step 1 can succeed only j^2 many times). Moreover the functions in $\mathcal{C}_{\langle k,i,j\rangle}$ are 1–1 enumerable effectively in $\langle k,i,j\rangle$.

Let $S = \{\langle k,i,j\rangle \mid (\forall p \leq \sqrt{j})[\varphi_p(0) \neq \langle k,i,j\rangle]\}$.

Let $\mathcal{C} = \bigcup_{\langle k,i,j\rangle \in S} \mathcal{C}_{\langle k,i,j\rangle}$.

Note that for each k, i, there exist infinitely many j, such that $\mathcal{C}_{\langle k,i,j\rangle} \subseteq \mathcal{C}$.

We first claim that, for all k, i, such that $F(k,\cdot,\cdot)$ is a Kolmogorov Numbering, for all but finitely many j, zeroext$(\varphi_{g(k,i,j)}) \notin \mathbf{Min}_{F(k,\cdot,\cdot)}\mathbf{Ex}(\mathbf{M}_i)$ (thus $\mathcal{C} \not\subseteq \mathbf{Min}_{F(k,\cdot,\cdot)}\mathbf{Ex}(\mathbf{M}_i)$). To see this, suppose $F(k,\cdot,\cdot)$ is a Kolmogorov numbering. Thus for all but finitely many j, $\mathrm{MinProg}_{F(k,\cdot,\cdot)}(\varphi_{g(k,i,j)}) < j^2$. Thus we consider j such that $\mathrm{MinProg}_{F(k,\cdot,\cdot)}(\varphi_{g(k,i,j)}) < j^2$. Now let $f = \mathrm{zeroext}(\varphi_{g(k,i,j)})$. If $(\overset{\infty}{\exists}\ n)[\mathbf{M}_i(f[n]) \geq j^2]$, then clearly, $f = \varphi_{g(k,i,j)} \notin \mathbf{Min}_{F(k,\cdot,\cdot)}\mathbf{Ex}(\mathbf{M}_i)$. Thus $(\overset{\infty}{\forall}\ n)[\mathbf{M}_i(f[n]) < j^2]$. But then by the construction of $\varphi_{g(k,i,j)}$, it follows that $(\forall l < j^2)[l \in \mathrm{Cancel} \vee F(k,l,\cdot)$ is not total$]$. Thus $(\forall l < j^2)[F(k,l,\cdot) \neq f]$. Thus $f \notin \mathbf{Min}_{F(k,\cdot,\cdot)}\mathbf{Ex}(\mathbf{M}_i)$.

We now construct a Gödel numbering ψ, such that $\mathcal{C} \in \mathbf{Min}_\psi\mathbf{Ex}$. Let $\mathrm{gap}(l) = l^4 + 2$. Let $h(l) = l + \Sigma_{l' \leq l}\mathrm{gap}(l)$. Let $\psi_{h(l)} = \varphi_l$. For $h(l) < x < h(l+1)$, if $\varphi_{l+1}(0) = \langle k,i,j\rangle$, then let $\psi_x = (x - h(l))$-th function in $\mathcal{C}_{\langle k,i,j\rangle}$, if any. Note that if $\varphi_{l+1}(0) = \langle k,i,j\rangle$ and $l+1 > \sqrt{j}$, then $j^2 + 1 \leq \mathrm{gap}(l+1)$ and thus all functions in $\mathcal{C}_{\langle k,i,j\rangle}$ are placed in the gap between $h(l)$ and $h(l+1)$.

Note that for $\langle k,i,j\rangle \in S$, for any l, such that $\varphi_{l+1}(0) = \langle k,i,j\rangle$ we have $l+1 > \sqrt{j}$, and thus for all such l, all functions in $\mathcal{C}_{\langle k,i,j\rangle}$ are placed by the construction in the gap between $h(l)$ and $h(l+1)$.

Now consider the following machine $\mathbf{M}$. $\mathbf{M}(f)$ converges to the least x, such that $h(l) < x < h(l+1)$, $\psi_x = f$, and $\varphi_{l+1}(0) = f(0)$. Note that by the property of ψ and S, discussed above it follows that $\mathcal{C} \subseteq \mathbf{Min}_\psi\mathbf{Ex}(\mathbf{M})$. ∎

Theorem 9 *For all $\alpha_1, \alpha_2, \alpha_3 \in \{\subseteq, \subset\}$, it is possible to construct a Gödel Numbering η such that* **FINITE** α_1 $\mathbf{Min}_\eta\mathbf{FIN}$ α_2 $\mathbf{Min}_\eta\mathbf{CoLearn}$ α_3 $\mathbf{Min}_\eta\mathbf{Ex}$.

Proof. The idea is essentially to interleave the diagonalizations for the relevant proper subset construction with the Gödel numbering in which no infinite set of functions is **MinEx**-identifiable.

So suppose β is the Gödel numbering in which no infinite class of functions can be $\mathbf{Min}_\beta\mathbf{Ex}$-identified. We will first construct a special numbering ψ such that **FINITE** $\subset$

$\mathbf{Min}_\psi\mathbf{FIN} \subset \mathbf{Min}_\psi\mathbf{CoLearn} \subset \mathbf{Min}_\psi\mathbf{Ex}$. This numbering will in addition have some nice properties which will allow us to construct the required η.

Let $\mathrm{gap}(r) = (r+2)^2 + (2r+3) + 2$. Let $h(r) = r + \sum_{l<r} \mathrm{gap}(l)$.

For $1 \le l \le 2$, let $h_1(r,l) = h(r) + (r+2)^2 + (2r+3) + l$.

For $1 \le l \le 2r+3$, let $h_2(r,l) = h(r) + (r+2)^2 + l$.

For $j < r+2, 1 \le l \le r+2$, let $h_3(r,j,l) = h(r) + j(r+2) + l$.

h, h_1, h_2, h_3 are used to partition N into four groups as follows. Let

$S_0 = \{i \mid (\exists r)[i = h(r)]\}$.

$S_1 = \{i \mid (\exists r, l \mid 1 \le l \le 2)[i = h_1(r,l)]\}$.

$S_2 = \{i \mid (\exists r, l \mid 1 \le l \le 2r+3)[i = h_2(r,l)]\}$.

$S_3 = \{i \mid (\exists r, j, l \mid j < r+2 \ \wedge \ 1 \le l \le r+2)[i = h_3(r,j,l)]\}$.

Note that S_0, S_1, S_2, S_3 is a partition of N into four disjoint subsets. We will use S_1, S_2, S_3 for three separate diagonalizations in the construction of ψ.

Let $\psi_{h(r)} = \beta_r$. (This makes ψ a Gödel Numbering).

For $i \in N$, let g_i^1 denote the function defined as follows.

$$g_i^1(x) = \begin{cases} 1, & \text{if } x = 0; \\ i, & \text{if } x = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Suppose $i \in S_1$ and $i = h_1(r,l)$. Let $\psi_i \subseteq g_i^1$ suchthat $\psi_i = g_i^1$ iff $(\forall j \le r)[g_i^1[2] \not\subseteq \beta_j]$.

Let $\mathcal{C}_1 = \{\psi_i \mid i \in S_1 \ \wedge \ \psi_i \text{ is total}\}$.

Let $\sigma_0, \sigma_1, \ldots$ be an 1-1 enumeration of all elements of INIT. We assume without loss of generality that $|\sigma_i| \le i$. Let X be a recursive function from N^2 to N such that the following properties are satisfied (note that such an X can easily be constructed).

(1) $(\forall j)[\sigma_{X(j,0)} \supseteq \{(0,2),(1,j)\}]$.

(2) $(\forall j,l)[\sigma_{X(j,l)} \subseteq \sigma_{X(j,l+1)}]$.

(3) $(\forall j)[\lim_{l\to\infty} X(j,l)\downarrow]$.

(4) $(\forall j,l)[X(j,l) \neq X(j,l+1) \Rightarrow X(j,l+1) > l]$.

(5) For any given j, suppose $\tau_j = \sigma_{\lim_{l\to\infty} X(j,l)}$. Then $(\forall k \le j)[(\exists \tau' \supseteq \tau_j)[\mathbf{M}_k(\tau') \neq ?] \Rightarrow \mathbf{M}_k(\tau_j) \neq ?]$.

Suppose $i \in S_2$, and $i = h_2(r, l)$, where $r = \langle j, k\rangle$ (note that by our convention $\langle j, k\rangle \geq \max(j, k)$). Let g_i^2 be defined as follows.

$$g_i^2(x) = \begin{cases} \sigma_k(x), & \text{if } x < |\sigma_k|; \\ i, & \text{if } x = |\sigma_k|; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\psi_i \subseteq g_i^2$, such that $\psi_i = g_i^2$, iff $[\lim_{t\to\infty} X(j, t) = k \ \wedge \ (\forall m \leq j)[\beta_{\mathbf{M}_m(\sigma_k)} \not\supseteq g_i^2[|\sigma_k| + 1]] \ \wedge \ (\forall m \leq r)[\beta_m \not\supseteq g_i^2[|\sigma_k| + 1]]]$.

Let $\mathcal{C}_2 = \{\psi_i \mid i \in S_2 \ \wedge \ \psi_i \text{ is total }\}$.

For $i \in S_3$, and $i = h_3(r, j, l)$, let g_i^3 be defined as follows.

$$g_i^3(x) = \begin{cases} 3, & \text{if } x = 0; \\ r, & \text{if } x = 1; \\ j, & \text{if } x = 2; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\psi_i \subseteq g_i^3$, such that $\psi_i = g_i^3$, iff $[(\forall k \leq r)[\beta_k \not\supseteq g_i^3[3]] \ \wedge \ l = \min(\{l' \mid 1 \leq l' \leq r + 2 \ \wedge \ (\forall k \leq r)[\{\mathbf{M}_k(g_i^3[n]) \mid n \in N\} \cap \{h(r, j, k) \mid 1 \leq k \leq r + 2\} \neq \{h(r, j, k) \mid 1 \leq k \leq r + 2 \ \wedge \ k \neq l'\}]\})]$.

Let $\mathcal{C}_3 = \{\psi_i \mid i \in S_3 \ \wedge \ \psi_i \text{ is total }\}$.

It is easy to observe the following properties (for identification in the numbering ψ).

$\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ are infinite and disjoint.

$\mathcal{C}_1 \in \mathbf{Min}_\psi\mathbf{FIN}$.

$\mathcal{C}_1 \cup \mathcal{C}_2 \in \mathbf{Min}_\psi\mathbf{CoLearn}$.

$\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \in \mathbf{Min}_\psi\mathbf{Ex}$.

No infinite subset of $\mathcal{C}_1$ is in **FINITE**.

No infinite subset of $\mathcal{C}_2$ is in $\mathbf{Min}_\psi\mathbf{FIN}$.

No infinite subset of $\mathcal{C}_3$ is in $\mathbf{Min}_\psi\mathbf{CoLearn}$.

There does not exists an infinite class $\mathcal{C}$ of functions such that a machine can **FIN**-identify $\mathcal{C}$, by converging (for functions in $\mathcal{C}$) to minimal ψ indices in $N - S_1$.

There does not exists an infinite class $\mathcal{C}$ of functions such that a machine can **CoLearn**-identify $\mathcal{C}$, by co-converging (for functions in $\mathcal{C}$) to minimal ψ indices in $N - S_1 - S_2$.

There does not exists an infinite class $\mathcal{C}$ of functions such that a machine can **Ex**-identify $\mathcal{C}$, by converging (for functions in $\mathcal{C}$) to minimal ψ indices in $N - S_1 - S_2 - S_3$.

Now consider the numbering η defined as follows. Let $S = S_0 \cup \bigcup_{1 \leq x \leq 3 \ \wedge \ \alpha_x = \mathsf{C}} S_x$. Let $\eta_i = \psi_i$, if $i \in S$. η_i is empty function otherwise.

Due to the properties of ψ discussed above, it is easy to verify that η satisfies the requirements of the theorem. ∎

References

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

[Blu67] M. Blum. A machine independent theory of the complexity of recursive functions. *Journal of the ACM*, 14:322–336, 1967.

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

[Che81] K. Chen. *Tradeoffs in Machine Inductive Inference*. PhD thesis, SUNY at Buffalo, 1981.

[Che82] K. Chen. Tradeoffs in inductive inference of nearly minimal sized programs. *Information and Control*, 52:68–86, 1982.

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

[Fre75] R. Freivalds. Minimal Gödel numbers and their identification in the limit. *Lecture Notes in Computer Science*, 32:219–225, 1975.

[Fre91] R. Freivalds. Inductive inference of recursive functions: Qualitative theory. In J. Barzdins and D. Bjorner, editors, *Baltic Computer Science. Lecture Notes in Computer Science 502*, pages 77–110. Springer-Verlag, 1991.

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

[HU79] J. Hopcroft and J. Ullman. *Introduction to Automata Theory Languages and Computation*. Addison-Wesley Publishing Company, 1979.

[OSW86] D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn, An Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, Cambridge, Mass., 1986.

[Rog67] H. Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw Hill, New York, 1967. Reprinted by MIT Press, Cambridge, Massachusetts in 1987.