Robust Learning — Rich and Poor

John Case [a] Sanjay Jain [b,1] Frank Stephan [c,2] Rolf Wiehagen [d]

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

[b] *School of Computing, National University of Singapore, Singapore 119260, sanjay@comp.nus.edu.sg*

[c] *National ICT Australia LTD, Sydney Node, The University of New South Wales, Sydney NSW 2052, Australia, fstephan@cse.unsw.edu.au*

[d] *Department of Computer Science, Technical University of Kaiserslautern, D-67653 Kaiserslautern, Germany, wiehagen@informatik.uni-kl.de*

Abstract

A class $\mathcal{C}$ of recursive functions is called *robustly learnable* in the sense **I** (where **I** is any success criterion of learning) if not only $\mathcal{C}$ itself but even all transformed classes $\Theta(\mathcal{C})$ where Θ is any general recursive operator, are learnable in the sense **I**. It was already shown before, see [Ful90,JSW01], that for $\mathbf{I} = \mathbf{Ex}$ (learning in the limit) robust learning is rich in that there are classes being both not contained in any recursively enumerable class of recursive functions and, nevertheless, robustly learnable. For several criteria **I**, the present paper makes much more precise *where* we *can* hope for robustly learnable classes and where we *cannot*. This is achieved in two ways. First, for $\mathbf{I} = \mathbf{Ex}$, it is shown that only *consistently* learnable classes can be uniformly robustly learnable. Second, some other learning types **I** are classified as to whether or not they contain rich robustly learnable classes. Moreover, the first results on separating robust learning from uniformly robust learning are derived.

1 Introduction

Robust learning has attracted much attention recently. Intuitively, a class of objects is *robustly learnable* if not only this class itself is learnable but all of its

[1] Supported in part by NUS grant number R252–000–127–112.

[2] Supported by the Deutsche Forschungsgemeinschaft (DFG) under Heisenberg grant Ste 967/1–1 while F. Stephan was working at the Mathematical Institute, University of Heidelberg.

effective transformations remain learnable as well. In this sense, being learnable robustly seems to be a desirable property in all fields of learning. In inductive inference, i.e., informally, learning of recursive functions in the limit, a large collection of function classes was already known to be robustly learnable. Actually, in [Gol67] any *recursively enumerable* class of recursive functions was shown to be learnable. This was achieved even by one and the same learning algorithm, the so-called identification by enumeration, see [Gol67]. Moreover, any reasonable model of effective transformations maps any recursively enumerable class again to a recursively enumerable and, hence, learnable class. Consequently, all these classes are *robustly* learnable. Clearly, the same is true for all *sub*classes of recursively enumerable classes. Thus, the challenging remaining question was if robust learning is even possible *outside* the world of the recursively enumerable classes. This question remained open for about 20 years, until it has been answered positively! [Ful90,JSW01] showed that there are classes of recursive functions which are both "algorithmically rich" and robustly learnable, where *algorithmically rich* means being not contained in any recursively enumerable class of recursive functions. Earliest examples of (large) algorithmically rich classes featured direct self-referential coding. Though ensuring the learnability of these classes themselves, these direct codings could be destroyed already by simple effective transformations, thus proving that these classes are *not* robustly learnable. An early motivation from Bārzdiņš for studying robustness was just to examine what happens to learnability when at least the then known direct codings are destroyed (by the effective transformations). Later examples of algorithmically rich classes, including some indeed *robustly* learnable examples, featured more indirect, "topological" coding. [Ful90,JSW01] mainly had focussed on the *existence* of rich *and* robustly learnable classes; however, in the present paper we want to make much more precise *where* we *can* hope for robustly learnable classes and where we can*not*. In order to reach this goal we will follow two lines.

The first line, outlined in Section 3, consists in exhibiting a "borderline" that separates the region where robustly learnable classes do exist from the region where robustly learnable classes provably cannot exist. More exactly, for the basic type **Ex** of learning in the limit, such a borderline is given just by the type **Cons** of learning in the limit *consistently* (i.e., each hypothesis correctly and completely reflects all the data seen so far). Actually, in Theorem 23 we show that all the uniformly robustly **Ex**-learnable classes must be already contained in **Cons**, and hence the complementary region $\mathbf{Ex} - \mathbf{Cons}$ is free of any such classes. Notice that $\mathbf{Ex} - \mathbf{Cons}$ is far from being empty, since it was shown before that **Cons** is a *proper* subset of **Ex**, see [Bar74a,BB75,Wie76]; the latter is also known as *inconsistency phenomenon*, see [WZ94,Ste98,CF99] where it has been shown that this phenomenon is present in polynomial-time learning as well. We were surprised to find the "robustness phenomenon" and the inconsistency phenomenon so closely related this way. There is another interpretation suggested by Theorem 23 which in a sense nicely contrasts the results on robust learning from [JSW01]. All the robustly learnable classes exhibited in that paper were of some non-trivial topological complexity, see [JSW01] for details. On the other hand, Theorem 23 intuitively says that uniformly robustly learnable classes may not be "too complex", as they all are located in the "lower part" **Cons** of the type **Ex**. Finally, this location in **Cons** in turn seems useful in that just consistent learning plays an important role not only in inductive inference, see [JB81,Ful88,Lan90,WZ95], but also in various other fields of learning such as PAC learning, machine learning and statistical learning, see the books [AB92,Mit97,Vap00], respectively.

In Section 4, we follow another line to solve the problem where rich robustly learnable classes can be found and where they cannot. Therefore let us call any type **I** of learning such as **I** = **Ex**, **Cons** etc. *robustly rich* if **I** contains rich classes being robustly learnable in the sense **I** (where "rich" is understood as above, i.e., a class is rich if it is not contained in any recursively enumerable class); otherwise, the type **I** is said to be *robustly poor*. Then, for a few types, it was already known if they are robustly rich or robustly poor. The first results in this direction are due to Zeugmann [Zeu86] where the types $\mathbf{Ex}_0$ (**Ex**-learning without any mind change) and **Reliable** (see Definition 8) were proved to be robustly poor. In [Ful90,JSW01] the type **Ex** was shown robustly rich. Below we classify several other types as to whether they are robustly rich or poor, respectively. We exhibit types of both categories, rich ones (hence the first rich ones after **Ex**) and poor ones, thus making the whole picture noticeably more complete. This might even serve as an appropriate starting point for solving the currently open problem to derive conditions (necessary, sufficient, both) for when a type is of which category. Notice that in proving types robustly rich below, in general, we show some stronger results, namely, we *robustly separate* the corresponding types from some other "close" types, thereby strengthening known separations in a robust way. From these separations, the corresponding richness results follow easily.

In Section 5, we deal with a problem which was competely open up to now, namely, separating robust learning from uniformly robust learning. While in robust learning any transformed class is required to be learnable in the mere sense that there *exists* a learning machine for it, uniformly robust learning intuitively requires to get such a learning machine for the transformed class *effectively at hand*, see Definition 14. As it turns out by our results, this additional requirement can really lead to a difference. Actually, for a number of learning types, we show that uniformly robust learning is stronger than robust learning. Notice that fixing this difference is also interesting for the following reason. As said above, in Theorem 23 all the uniformly robustly **Ex**-learnable classes are shown to be learnable consistently. However, at present, it is open if this result remains valid when uniform robustness will be replaced by robustness only. Some results of Section 5 can be considered as first steps to attack this apparently difficult problem.

Recently, several papers were published that deal with robustness in inductive inference, see [Zeu86,KS89a,KS89b,Ful90,JSW01,Jai99,OS02,CJO$^+$00]. Each of them has contributed interesting points to a better understanding of the challenging phenomenon of robust learning. [Zeu86,Ful90,JSW01] are already quoted above. In addition, notice that in [JSW01] the mind change hierarchy for **Ex**-type learning is proved to stand robustly. This contrasts the result from [Ful90] that the anomaly hierarchy for **Ex**-type learning does *not* hold robustly. In [KS89a,KS89b] the authors were dealing with so-called Bārzdiņš' Conjecture which, intuitively, stated that the type **Ex** is robustly poor. In [CJO$^+$00] robust learning has been studied for another specific learning scenario, namely learning aided by context. The intuition behind this model is to present the functions to be learned not in a pure fashion to the learner, but together with some "context" which is intended to help in learning. It is shown that within this scenario several results hold robustly as well. In [OS02] the notion of hyperrobust learning is introduced. A class of recursive functions is called *hyperrobustly learnable* if there is one and the same learner which learns not only this class itself but also all of its images under all primitive recursive operators. Hence this learner must be capable to learn the *union* of all these images. This definition is then justified by the following results. First, it is shown that the power of hyperrobust learning does not change if the class of primitive recursive operators is replaced by any larger, still recursively enumerable class of general recursive operators. Second, based on this stronger definition, Bārzdiņš' Conjecture is proved by showing that a class of recursive functions is hyperrobustly **Ex**-learnable iff this class is contained in a recursively enumerable class of recursive functions. Note that, by this equivalence and by Corollary 34 below, hyperrobust **Ex**-learning is stronger than uniformly robust **Ex**-learning. From [OS02] hyperrobustness destroys both direct and topological coding tricks. In [CJO$^+$00] it is noted that hyperrobustness destroys any advantage of context, but, since, empirically, context does help, this provides evidence that the real world, in a sense, has codes for some things buried inside others. In [Jai99] another basic type of inductive inference, namely **Bc**, has been robustly separated from **Ex**, thus solving an open problem from [Ful90]. While in the present paper general recursive operators are taken in order to realize the transformations of the classes under consideration (the reason for this choice is mainly a technical one, namely, that these operators "automatically" map any class of recursive functions to a class of recursive functions again), in some of the papers above other types of operators are used such as effective, recursive, primitive recursive operators, respectively. At this moment, we do not see any choice to this end that seems to be superior to the others. Indeed, each approach appears justified if it yields interesting results.

For references surveying the theory of learning recursive functions, the reader is referred to [AS83,BB75,CS83,Fre91,KW80,OSW86,JORS99].

2 Notation and Preliminaries

Recursion-theoretic concepts not explained below are treated in [Rog67]. N denotes the set of natural numbers. $*$ denotes a non-member of N and is assumed to satisfy $(\forall n)[n < * < \infty]$. Let $\in, \subseteq, \subset, \supseteq, \supset$, respectively denote the membership, subset, proper subset, superset and proper superset relations for sets. The empty set is denoted by $\emptyset$. We let $\mathrm{card}(S)$ denote the cardinality of the set S. So "$\mathrm{card}(S) \leq *$" means that $\mathrm{card}(S)$ is finite. The minimum and maximum of a set S are denoted by $\min(S)$ and $\max(S)$, respectively. We take $\max(\emptyset)$ to be 0 and $\min(\emptyset)$ to be ∞. χ_A denotes the characteristic function of A, that is, $\chi_A(x) = 1$, if $x \in A$, and 0 otherwise.

$\langle \cdot, \cdot \rangle$ denotes a 1-1 computable mapping from pairs of natural numbers onto natural numbers. π_1, π_2 are the corresponding projection functions. $\langle \cdot, \cdot \rangle$ is extended to n-tuples of natural numbers in a natural way. Λ denotes the empty function. η, with or without subscripts, superscripts, primes and the like, ranges over partial functions. If η_1 and η_2 are both undefined on input x, then, we take $\eta_1(x) = \eta_2(x)$. We say that $\eta_1 \subseteq \eta_2$ iff for all x in domain of η_1, $\eta_1(x) = \eta_2(x)$. We let $\mathrm{domain}(\eta)$ and $\mathrm{range}(\eta)$ respectively denote the domain and range of the partial function η. $\eta(x)\downarrow$ denotes that $\eta(x)$ is defined. $\eta(x)\uparrow$ denotes that $\eta(x)$ is undefined.

We say that a partial function η is *conforming* with η' iff for all $x \in \mathrm{domain}(\eta) \cap \mathrm{domain}(\eta')$, $\eta(x) = \eta(x')$. $\eta \sim \eta'$ denotes that η is conforming with η'. η is *non-conforming* with η' iff there exists an x such that $\eta(x)\downarrow \neq \eta'(x)\downarrow$. $\eta \not\sim \eta'$ denotes that η is non-conforming with η'.

For $r \in N$, r-extension of η denotes the function f defined as follows:

$$f(x) = \begin{cases} \eta(x), & \text{if } x \in \mathrm{domain}(\eta); \\ r, & \text{otherwise.} \end{cases}$$

For a finite set S of programs, we let $\mathrm{Union}(S)$ denote the partial recursive function: $\mathrm{Union}(S)(x) = \varphi_p(x)$, for the first $p \in S$ found such that $\varphi_p(x)\downarrow$ using some standard dovetailing mechanism for computing φ_p's. When programs $q_1, q_2, \ldots, q_n$ for partial recursive functions $\eta_1, \eta_2, \ldots, \eta_n$ are implicit, we sometimes abuse notation and use $\mathrm{Union}(\{\eta_1, \eta_2, \ldots, \eta_n\})$, to denote $\mathrm{Union}(\{q_1, q_2, \ldots, q_n\})$.

f, g, h, F and H, with or without subscripts, superscripts, primes and the like, range over total functions. $\mathcal{R}$ denotes the class of all *recursive* functions, i.e., total computable functions with arguments and values from N. $\mathcal{T}$ denotes the class of all *total* functions. $\mathcal{R}_{0,1}$ ($\mathcal{T}_{0,1}$) denotes the class of all *recursive* functions (total functions) with range contained in $\{0, 1\}$. $\mathcal{C}$ and $\mathcal{S}$, with or without subscripts, superscripts, primes and the like, range over subsets of $\mathcal{R}$. $\mathcal{P}$ denotes the class of all *partial recursive* functions over N. φ denotes a *fixed* acceptable programming system. φ_i denotes the partial recursive function computed by program i in the φ-system. Note that in this paper all programs are interpreted with respect to the φ-system. We let Φ be an arbitrary Blum complexity measure [Blu67] associated with the acceptable programming system φ; many such measures exist for any acceptable programming system [Blu67]. We assume without loss of generality that $\Phi_i(x) \geq x$, for all i, x. $\varphi_{i,s}$ is defined as follows:

$$\varphi_{i,s}(x) = \begin{cases} \varphi_i(x), & \text{if } x < s \text{ and } \Phi_i(x) < s; \\ \uparrow, & \text{otherwise.} \end{cases}$$

For a given partial computable function η, we define $\mathrm{MinProg}(\eta) = \min(\{i \mid \varphi_i = \eta\})$.

A class $\mathcal{C} \subseteq \mathcal{R}$ is said to be recursively enumerable iff there exists an r.e. set X such that $\mathcal{C} = \{\varphi_i \mid i \in X\}$. For any non-empty recursively enumerable class $\mathcal{C}$, there exists a recursive function f such that $\mathcal{C} = \{\varphi_{f(i)} \mid i \in N\}$.

A class $\mathcal{C}$ is said to be *h-bounded* iff for all $f \in \mathcal{C}$, for all but finitely many x, $f(x) \leq h(x)$. A class $\mathcal{C}$ is *bounded* iff it is h-bounded for some recursive h. A class $\mathcal{C}$ is *unbounded* iff it is not h-bounded for any recursive h.

The following functions and classes are commonly considered below. Zero is the everywhere 0 function, i.e., $\mathrm{Zero}(x) = 0$, for all $x \in N$. $\mathrm{CONST} = \{f \mid (\forall x)[f(x) = f(0)]\}$ denotes the class of constant functions. $\mathrm{FINSUP} = \{f \mid (\forall^\infty x)[f(x) = 0]\}$ denotes the class of all recursive functions of finite support.

2.1 Function Identification

We first describe inductive inference machines. We assume that the graph of a function is fed to a machine in canonical order.

For $f \in \mathcal{R}$ and $n \in N$, we let $f[n]$ denote $f(0)f(1)\ldots f(n-1)$, the finite initial segment of f of length n. Clearly, $f[0]$ denotes the empty segment. SEG denotes the set of all finite strings, $\{f[n] \mid f \in \mathcal{R} \wedge n \in N\}$. $\mathrm{SEG}_{0,1} = \{f[n] \mid f \in \mathcal{R}_{0,1} \wedge n \in N\}$. We let σ, τ and γ, with or without subscripts, superscripts, primes and the like, range over SEG. Λ denotes the empty sequence. We assume some computable ordering of elements of SEG. $\sigma < \tau$, if σ appears before τ in this ordering. Similary one can talk about least element of a subset of SEG.

For a finite string σ and finite or infinite string β, we let $\sigma \cdot \beta$ denote the concatenation of σ and β. We identify $\sigma = a_0 a_1 \ldots a_{n-1}$ with the partial function

$$\sigma(x) = \begin{cases} a_x, & \text{if } x < n; \\ \uparrow, & \text{otherwise.} \end{cases}$$

Similarly a total function g is identified with the infinite sequence $g(0)g(1)g(2)\ldots$. Thus, for example, $0^\infty = \text{Zero}$.

Let $|\sigma|$ denote the length of σ. If $|\sigma| \geq n$, then we let $\sigma[n]$ denote the prefix of σ of length n. $\sigma \subseteq \tau$ denotes that σ is a prefix of τ. An *inductive inference machine* (IIM) [Gol67] is an algorithmic device that computes a (possibly partial) mapping from SEG into N. Since the set of all finite strings, SEG, can be coded onto N, we can view these machines as taking natural numbers as input and emitting natural numbers as output. We say that $\mathbf{M}(f)$ converges to i (written: $\mathbf{M}(f)\downarrow = i$) iff $(\forall^\infty n)[\mathbf{M}(f[n]) = i]$; $\mathbf{M}(f)$ is undefined if no such i exists. $\mathbf{M}_0, \mathbf{M}_1, \ldots$ denotes a recursive enumeration of all the IIMs. The next definitions describe several criteria of function identification.

Definition 1 [Gol67] Let $f \in \mathcal{R}$ and $\mathcal{S} \subseteq \mathcal{R}$.

(a) $\mathbf{M}$ $\mathbf{Ex}$-*identifies* f (written: $f \in \mathbf{Ex}(\mathbf{M})$) just in case there exists a program i for f such that $\mathbf{M}(f)\downarrow = i$.
(b) $\mathbf{M}$ $\mathbf{Ex}$-*identifies* $\mathcal{S}$ iff $\mathbf{M}$ $\mathbf{Ex}$-identifies each $f \in \mathcal{S}$.
(c) $\mathbf{Ex} = \{\mathcal{S} \subseteq \mathcal{R} \mid (\exists \mathbf{M})[\mathcal{S} \subseteq \mathbf{Ex}(\mathbf{M})]\}$.

By the definition of convergence, only finitely many data points from a function f have been observed by an IIM $\mathbf{M}$ at the (unknown) point of convergence. Hence, some form of learning must take place in order for $\mathbf{M}$ to learn f. For this reason, hereafter the terms *identify*, *learn* and *infer* are used interchangeably.

Definition 2 [Bar74b,CS83] Let $f \in \mathcal{R}$.

(a) $\mathbf{M}$ $\mathbf{Bc}$-*identifies* f (written: $f \in \mathbf{Bc}(\mathbf{M})$) iff, for all but finitely many $n \in N$, $\mathbf{M}(f[n])$ is a program for f.
(b) $\mathbf{M}$ $\mathbf{Bc}$-*identifies* $\mathcal{S}$ iff $\mathbf{M}$ $\mathbf{Bc}$-identifies each $f \in \mathcal{S}$.
(c) $\mathbf{Bc} = \{\mathcal{S} \subseteq \mathcal{R} \mid (\exists \mathbf{M})[\mathcal{S} \subseteq \mathbf{Bc}(\mathbf{M})]\}$.

Definition 3 [Bar74a] $\mathbf{M}$ is said to be *consistent* on f iff, for all n, $\mathbf{M}(f[n])\downarrow$ and $f[n] \subseteq \varphi_{\mathbf{M}(f[n])}$.

Definition 4 [Wie78] $\mathbf{M}$ is said to be *conforming* on f iff, for all n, $\mathbf{M}(f[n])\downarrow$ and $f[n] \sim \varphi_{\mathbf{M}(f[n])}$.

Definition 5 (a) [Bar74a] **M** **Cons**-*identifies* f iff **M** is consistent on f, and **M** **Ex**-identifies f.

(b.1) [Bar74a] **M** **Cons**-*identifies* $\mathcal{C}$ iff **M** **Cons**-identifies each $f \in \mathcal{C}$.

(b.2) **Cons** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ **Cons**-identifies $\mathcal{C}]\}$.

(c.1) [JB81] **M** $\mathcal{R}$**Cons**-*identifies* $\mathcal{C}$ iff **M** is total, and **M** **Cons**-identifies $\mathcal{C}$.

(c.2) $\mathcal{R}$**Cons** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ $\mathcal{R}$**Cons**-identifies $\mathcal{C}]\}$.

(d.1) [WL76] **M** $\mathcal{T}$**Cons**-*identifies* $\mathcal{C}$ iff **M** is consistent on each $f \in \mathcal{T}$, and **M** **Cons**-identifies $\mathcal{C}$.

(d.2) $\mathcal{T}$**Cons** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ $\mathcal{T}$**Cons**-identifies $\mathcal{C}]\}$.

Note that for **M** to **Cons**-identify a function f, it must be defined on each initial segment of f. Similarly for **Conf**-identification below.

Definition 6 (a) [Wie78] **M** **Conf**-*identifies* f iff **M** is conforming on f, and **M** **Ex**-identifies f.

(b.1) [Wie78] **M** **Conf**-*identifies* $\mathcal{C}$ iff **M** **Conf**-identifies each $f \in \mathcal{C}$.

(b.2) **Conf** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ **Conf**-identifies $\mathcal{C}]\}$.

(c.1) **M** $\mathcal{R}$**Conf**-*identifies* $\mathcal{C}$ iff **M** is total, and **M** **Conf**-identifies $\mathcal{C}$.

(c.2) $\mathcal{R}$**Conf** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ $\mathcal{R}$**Conf**-identifies $\mathcal{C}]\}$.

(d.1) [Ful88] **M** $\mathcal{T}$**Conf**-*identifies* $\mathcal{C}$ iff **M** is conforming on each $f \in \mathcal{T}$, and **M** **Conf**-identifies $\mathcal{C}$.

(d.2) $\mathcal{T}$**Conf** $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ $\mathcal{T}$**Conf**-identifies $\mathcal{C}]\}$.

Definition 7 [OSW86] **M** is *confident* iff for all total f, $\mathbf{M}(f)\downarrow$.

M **Confident**-*identifies* $\mathcal{C}$ iff **M** is confident and **M** **Ex**-identifies $\mathcal{C}$.

Confident $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ **Confident**-identifies $\mathcal{C}]\}$.

Definition 8 [Min76,BB75] **M** is *reliable* iff **M** is total, and for all total f, $\mathbf{M}(f)\downarrow \Rightarrow$ **M** **Ex**-identifies f.

M **Reliable**-*identifies* $\mathcal{C}$ iff **M** is reliable and **M** **Ex**-identifies $\mathcal{C}$.

Reliable $= \{\mathcal{C} \mid (\exists \mathbf{M})[\mathbf{M}$ **Reliable**-identifies $\mathcal{C}]\}$.

Definition 9 $\mathbf{NUM} = \{\mathcal{C} \mid (\exists \mathcal{C}' \mid \mathcal{C} \subseteq \mathcal{C}' \subseteq \mathcal{R})[\mathcal{C}' \text{ is recursively enumerable}]\}$.

For references on inductive inference within **NUM**, the set of all recursively enumerable classes and their subclasses, the reader is referred to [Gol67,BF74,FBP91].

We let **I** and **J** range over identification criteria defined above.

The following theorem relates the criteria of inference discussed above.

Theorem 10 [WL76,WZ95,Bar74a,Bar74b,BB75,Wie76,Wie78,Ful88,Gra86]

(a) $\mathbf{NUM} \subset \mathcal{T}\mathbf{Cons} = \mathcal{T}\mathbf{Conf} \subset \mathcal{R}\mathbf{Cons} \subset \mathcal{R}\mathbf{Conf} \subset \mathbf{Conf} \subset \mathbf{Ex} \subset \mathbf{Bc}$.

(b) $\mathcal{R}\mathbf{Cons} \subset \mathbf{Cons} \subset \mathbf{Conf}$.

(c) $\mathcal{R}\mathbf{Conf} \not\subseteq \mathbf{Cons}$.

(d) $\mathbf{Cons} \not\subseteq \mathcal{R}\mathbf{Conf}$.

(e) $\mathcal{T}\mathbf{Cons} \subset \mathbf{Reliable} \subset \mathbf{Ex}$.

(f) $\mathbf{Reliable} \not\subseteq \mathbf{Conf}$.

(g) $\mathcal{R}\mathbf{Cons} \not\subseteq \mathbf{Reliable}$.

(h) $\mathbf{NUM} \not\subseteq \mathbf{Confident}$.

(i) $\mathbf{Confident} \not\subseteq \mathbf{Conf}$.

(j) $\mathbf{Confident} \not\subseteq \mathbf{Reliable}$.

Proof. (a) and (b): $\mathbf{NUM} \subset \mathcal{T}\mathbf{Cons} \subset \mathbf{Cons}$ was shown by [WL76]. $\mathcal{T}\mathbf{Cons} \subset \mathcal{R}\mathbf{Cons} \subset \mathbf{Cons}$ was shown by [WZ95]. $\mathbf{Cons} \subset \mathbf{Ex}$ was shown by [Bar74a,BB75,Wie76]. $\mathbf{Cons} \subset \mathbf{Conf} \subset \mathbf{Ex}$ was done by [Wie78]. $\mathcal{T}\mathbf{Cons} = \mathcal{T}\mathbf{Conf}$ was shown by [Ful88]. $\mathbf{Ex} \subset \mathbf{Bc}$ was done by [Bar74b] (see also [CS83]).

$\mathcal{R}\mathbf{Cons} \subset \mathcal{R}\mathbf{Conf}$ and $\mathcal{R}\mathbf{Conf} \subset \mathbf{Conf}$ follow from parts (c) and (d).

Part (c) follows from proof of $\mathbf{Conf} - \mathbf{Cons} \neq \emptyset$ in [Wie78]. We are not sure if anyone has explicitly shown part (d), but it follows as a corollary to Theorem 35.

$\mathcal{T}\mathbf{Cons} \subseteq \mathbf{Reliable}$ follows from definition. $\mathbf{Reliable} - \mathcal{T}\mathbf{Cons} \neq \emptyset$ was shown by [Gra86]; it also follows from part (f). $\mathbf{Reliable} \subset \mathbf{Ex}$ was shown by [BB75]. Thus, part (e) follows.

For part (f), consider the class $\mathcal{C} = \text{FINSUP} \cup \{f \mid \varphi_{f(0)} = f \wedge (\forall x)[\Phi_{f(0)}(x) \leq f(x+1)]\}$. Clearly, $\mathcal{C} \in \mathbf{Reliable}$. Also since FINSUP $\subseteq \mathcal{C}$, if $\mathcal{C} \in \mathbf{Conf}$, then $\mathcal{C} \in \mathcal{T}\mathbf{Conf} = \mathcal{T}\mathbf{Cons}$. Now suppose by way of contradiction that $\mathbf{M}$ $\mathcal{T}\mathbf{Cons}$-identifies $\mathcal{C}$. Then by Kleene recursion theorem [Rog67] there exists an e such that φ_e may be described as follows.

$$\varphi_e(x) = \begin{cases} e, & \text{if } x = 0; \\ \Phi_e(x-1), & \text{if } x > 0, \text{ and} \\ & \mathbf{M}(\varphi_e[x] \cdot \Phi_e(x-1)) \neq \mathbf{M}(\varphi_e[x]); \\ \Phi_e(x-1)+1, & \text{otherwise.} \end{cases}$$

It is easy to verify that φ_e is total. Since $\mathbf{M}$ is consistent on all inputs, for each $f[x]$, there exists at most one y, such that $\mathbf{M}(f[x]) = \mathbf{M}(f[x] \cdot y)$. Thus, by definition of φ_e, for all $x > 0$, $\mathbf{M}(\varphi_e[x]) \neq \mathbf{M}(\varphi_e[x+1])$.

Let $\mathcal{SD} = \{f \in \mathcal{R} \mid \varphi_{f(0)} = f\}$. It was shown by [BB75], that $\mathcal{SD} \notin \mathbf{Reliable}$. Since $\mathcal{SD} \in \mathcal{R}\mathbf{Cons} \cap \mathbf{Confident}$, part (g) and (j) follow.

For (h), consider FINSUP. Clearly, FINSUP $\in \mathbf{NUM}$. Now, suppose FINSUP $\subseteq \mathbf{Ex}(\mathbf{M})$. Then, for all σ, there exists τ such that $\sigma \subseteq \tau$, and $\mathbf{M}(\sigma) \neq \mathbf{M}(\tau)$. It follows that there exists an infinite sequence σ_i, $i \in N$ such that $\sigma_i \subset \sigma_{i+1}$, and $\mathbf{M}(\sigma_i) \neq \mathbf{M}(\sigma_{i+1})$. Thus, $\mathbf{M}$ makes infinite number of mind changes on $\bigcup_i \sigma_i$. Thus $\mathbf{M}$ is not confident. (h) follows.

For part (i), let $\mathcal{C} = \{f \mid [\varphi_{f(0)} = f \ \wedge \ (\forall x > 0)[f(x) \neq 0]]$ or $[(\exists y)[(\forall x > y)[f(x) = 0] \ \wedge \ (\forall x \mid 0 < x \leq y)[f(x) \neq 0]]]\}$.

It is easy to verify that $\mathcal{C} \in \mathbf{Confident}$. Suppose by way of contradiction that $\mathbf{M}$ $\mathbf{Conf}$-identifies $\mathcal{C}$. We consider two cases:

Case 1: There exists an $f[n]$ such that, for x with $0 < x < n$, $f(x) \neq 0$ and $[\mathbf{M}(f[n])\uparrow$ or $f[n] \not\subseteq \varphi_{\mathbf{M}(f[n])}]$.

In this case, let $g = f[n]0^\infty$. Clearly, $g \in \mathcal{C}$, but $\mathbf{M}$ does not $\mathbf{Conf}$-identify g.

Case 2: Not Case 1.

In this case, by Kleene recursion theorem [Rog67], there exists an e such that φ_e may be defined as follows.

$$\varphi_e(x) = \begin{cases} e, & \text{if } x = 0; \\ 1, & \text{if } x > 0 \text{ and} \\ & \mathbf{M}(\varphi_e[x] \cdot 1) \neq \mathbf{M}(\varphi_e[x]); \\ 2, & \text{otherwise.} \end{cases}$$

By hypothesis of the case, φ_e is total and a member of $\mathcal{C}$. Suppose $\mathbf{M}(\varphi_e)\downarrow$. Let n be such that for all $x > n$, $\mathbf{M}(\varphi_e[x]) = \mathbf{M}(\varphi_e[n])$. But then, by definition of φ_e, for all $x > n$, $\varphi_e(x) = 2$, and $\mathbf{M}(\varphi_e[x] \cdot 1) = \mathbf{M}(\varphi_e[x])$. This, by hypothesis of the case, implies that $\mathbf{M}$ does not $\mathbf{Ex}$-identify φ_e. ∎

2.2 Operators

Definition 11 [Rog67] A *recursive operator* is an effective total mapping, Θ, from (possibly partial) functions to (possibly partial) functions, which satisfies the following properties:

(a) Monotonicity: For all functions η, η', if $\eta \subseteq \eta'$ then $\Theta(\eta) \subseteq \Theta(\eta')$.
(b) Compactness: For all η, if $(x, y) \in \Theta(\eta)$, then there exists a finite function $\alpha \subseteq \eta$ such that $(x, y) \in \Theta(\alpha)$.
(c) Recursiveness: For all finite functions α, one can effectively enumerate (in α) all $(x, y) \in \Theta(\alpha)$.

Definition 12 [Rog67] A recursive operator Θ is called *general recursive* iff Θ maps all total functions to total functions.

For each recursive operator Θ, we can effectively (from Θ) find a recursive operator Θ' such that,

(d) for each finite function α, $\Theta'(\alpha)$ is finite, and its canonical index can be effectively determined from α; furthermore if $\alpha \in \mathrm{SEG}$ then $\Theta'(\alpha) \in \mathrm{SEG}$, and
(e) for all total functions f, $\Theta'(f) = \Theta(f)$.

This allows us to get a nice effective sequence of recursive operators.

Proposition 13 [JSW01] *There exists an effective enumeration,* $\Theta_0, \Theta_1, \ldots$, *of recursive operators satisfying condition* (d) *above such that, for all recursive operators* Θ, *there exists an* $i \in N$ *satisfying:*

$$\textit{for all total functions } f,\ \Theta(f) = \Theta_i(f).$$

Since we will be mainly concerned with the properties of operators on total functions, for diagonalization purposes, one can restrict attention to operators in the above enumeration $\Theta_0, \Theta_1, \ldots$.

Definition 14 [Ful90,JSW01]

$\mathbf{RobustEx} = \{\mathcal{C} \mid (\forall \text{ general recursive operators } \Theta)[\Theta(\mathcal{C}) \in \mathbf{Ex}]\}$.

$\mathbf{UniformRobustEx} = \{\mathcal{C} \mid (\exists g \in \mathcal{R})(\forall e \mid \Theta_e \text{ is general recursive})[\Theta_e(\mathcal{C}) \subseteq \mathbf{Ex}(\mathbf{M}_{g(e)})]\}$.

One can similarly define **RobustI** and **UniformRobustI**, for other criteria **I** of learning considered in this paper.

2.3 Some Useful Propositions

Proposition 15 [JSW01,Zeu86] **NUM** = **RobustNUM**

Corollary 16 (a) **NUM** $\subseteq$ **Robust$\mathcal{T}$Cons** $\subseteq$ **RobustReliable**.

(b) **NUM** $\subseteq$ **UniformRobust$\mathcal{T}$Cons**.

Proof. (a) follows immediately from Proposition 15 and Theorem 10(a,e).

In order to show (b) recall that, by Proposition 15, for any recursively enumerable class $\mathcal{C}$ and any general recursive operator Θ_e, $\Theta_e(\mathcal{C})$ is again recursively enumerable and, hence, can be **Ex**-learned using identification by enumeration, see [Gol67]. Moreover, given $\mathcal{C}$, this identification by enumeration strategy can be computed uniformly in e. Recall that identification by enumeration already works consistently, see [Gol67]. Finally, notice that this strategy can easily and uniformly be made to work consistently on all total functions. ∎

We say that $f =^* g$ iff card($\{x \mid f(x) \neq g(x)\}$) is finite. We say that $\mathcal{C} \subseteq \mathcal{R}$ is *closed under finite variations* iff for all $f, g \in \mathcal{R}$ such that $f =^* g$, $f \in \mathcal{C}$ iff $g \in \mathcal{C}$.

Proposition 17 [OS02] *Suppose $\mathcal{C}$ is closed under finite variations. Then $\mathcal{C} \in$ **RobustEx** iff $\mathcal{C} \in$ **NUM**.*

Proposition 18 $\mathcal{T}$**Cons** $\not\subseteq$ **RobustEx**.

Proof. Let $\mathcal{C} = \{f \mid (\exists i)[\Phi_i \in \mathcal{R} \ \wedge \ f =^* \Phi_i]\}$. It is easy to verify that $\mathcal{C} \in \mathcal{T}$**Cons**. However, as $\mathcal{C}$ is closed under finite variations, using Proposition 17, we have $\mathcal{C} \in$ **RobustEx** iff $\mathcal{C} \in$ **NUM**. As $\{\Phi_i \mid \Phi_i \in \mathcal{R}\} \notin$ **NUM**, proposition follows. ∎

The following result was shown in [Gra86].

Lemma 19 [Gra86] *Suppose $\mathcal{S} \in$ **Reliable**, and $\mathcal{S}$ is bounded. Then $\mathcal{S} \in$ **NUM**.*

Corollary 20 *Suppose $\mathcal{S} \in \mathcal{T}$**Cons**, and $\mathcal{S}$ is bounded. Then $\mathcal{S} \in$ **NUM**.*

Proof. This follows immediately from Lemma 19 and Theorem 10(e). ∎

Proposition 21 [FW79] *Let* $\mathcal{S}_e = \{\varphi_i \in \mathcal{R} \mid i \leq e\}$*. Then, there exists a recursive* h *such that, for any* e*,* $\mathcal{S}_e \subseteq \mathbf{Ex}(\mathbf{M}_{h(e)})$*. Moreover, this* $\mathbf{M}_{h(e)}$ *also witnesses that* $\mathcal{S}_e \in \mathbf{Cons}$*.*

Corollary 22 *Let* $\mathcal{S}_e = \{\varphi_i \in \mathcal{R} \mid i \leq e\}$*. There exists a recursive* h *such that, if* Θ_j *is a general recursive operator, then* $\Theta_j(\mathcal{S}_e) \subseteq \mathbf{Ex}(\mathbf{M}_{h(e,j)})$*. Moreover, this* $\mathbf{M}_{h(e,j)}$ *also witnesses that* $\Theta_j(\mathcal{S}_e) \in \mathbf{Cons}$*.*

3 Robust Learning and Consistency

The main result of this section is **UniformRobustEx** $\subseteq$ **Cons**, see Theorem 23. Hence, all the uniformly robustly **Ex**-learnable classes are contained in the "lower part" **Cons** of **Ex**; recall that **Cons** is a *proper* subset of **Ex**, see [Bar74a,BB75,Wie76]. This nicely relates two surprising phenomena of learning, namely the robustness phenomenon and the inconsistency phenomenon. On the other hand, despite the fact that every uniformly robustly **Ex**-learnable class is located in that lower part of **Ex**, **UniformRobustEx** contains "algorithmically rich" classes, since, by Corollary 34, **UniformRobustEx** $\supset$ **NUM**.

Theorem 23 **UniformRobustEx** $\subseteq$ **Cons**.

PROOF. Suppose $\mathcal{C} \in$ **UniformRobustEx**, and g is a recursive function such that, for all i, if Θ_i is general recursive, then $\Theta_i(\mathcal{C}) \subseteq \mathbf{Ex}(\mathbf{M}_{g(i)})$. Without loss of generality we may assume that each $\mathbf{M}_{g(i)}$ is total.

Then, by Kleene recursion theorem [Rog67], there exists an e such that Θ_e may be described as follows. For ease of presentation, $\Theta_e(f[m])$ may be infinite for some f, m (and thus it does not satisfy the invariant (d) we assumed (see discussion around Proposition 13) on the enumeration $\Theta_0, \Theta_1, \ldots$; this can be easily handled by appropriately slowing down Θ_e).

$\Theta_e(\Lambda) = \Lambda$.

$$\Theta_e(f[n{+}1]) = \begin{cases} \Theta_e(f[n]), & \text{if } \Theta_e(f[n]) \text{ is infinite;} \\ \Theta_e(f[n]) \cdot f(n) \cdot 0^\infty, & \text{if for all } m, \\ & \quad \Theta_e(f[n]) \cdot f(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]) \cdot f(n) \cdot 0^m), m}; \\ \Theta_e(f[n]) \cdot f(n) \cdot 0^m, & \text{if } m \text{ is the least number such that} \\ & \quad \Theta_e(f[n]) \cdot f(n) \subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]) \cdot f(n) \cdot 0^m), m}. \end{cases}$$

It is easy to verify that Θ_e is general recursive.

Claim 24 (a) *If* $\Theta_e(f[n])$ *is finite, and* $a \neq b$*, then* $\Theta_e(f[n] \cdot a) \not\sim \Theta_e(f[n] \cdot b)$*.*

(b) *If* $\Theta_e(f[n])$ *and* $\Theta_e(h[m])$ *are both finite, then* $f[n] \sim h[m]$ *iff* $\Theta_e(f[n]) \sim \Theta_e(h[m])$.

(c) *If* $\Theta_e(f[n+1])$ *is finite, then* $\Theta_e(f[n]) \cdot f(n) \subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n+1]))}$.

(d) *For all* p, n, *there exists at most one* $f[n+1]$, *such that* $\Theta_e(f[n]) \cdot f(n) \subseteq \varphi_p$.

(e) *If* $\Theta_e(f[n+1])$ *is infinite, then* $\mathbf{M}_{g(e)}$ *does not* **Ex**-*identify* $\Theta_e(f)$.

(f) *For all* $f \in \mathcal{C}$, *for all* n, $\Theta_e(f[n])$ *is finite.*

PROOF. (a), (c) follow from construction. (b) follows from part (a). Part (d) follows from part (b).

For part (e), suppose $\Theta_e(f[n+1])$ is infinite. Without loss of generality assume that n is least such that $\Theta_e(f[n+1])$ is infinite. Then, $\Theta_e(f) = \Theta_e(f[n+1]) = \Theta_e(f[n]) \cdot f(n) \cdot 0^\infty$. But, for all m, $\Theta_e(f[n]) \cdot f(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]) \cdot f(n) \cdot 0^m), m}$. Thus, either $\mathbf{M}_{g(e)}$ diverges on $\Theta_e(f)$, or $\Theta_e(f[n]) \cdot f(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f))}$. Thus, $\mathbf{M}_{g(e)}$ does not **Ex**-identify $\Theta_e(f)$.

(f) follows as corollary to part (e). $\square$

Let *prog* be a recursive function such that $\varphi_{prog(p)}$ is defined as follows.

$\varphi_{prog(p)}(x)$
1. Search for $f[x+1]$ such that $\Theta_e(f[x+1])$ is finite, and $\Theta_e(f[x]) \cdot f(x) \subseteq \varphi_p$.
   (* Note that by Claim 24(d), there exists at most one $f[x+1]$ satisfying above. Moreover, if $p \in \{\mathbf{M}_{g(e)}(\Theta_e(h[x+1])) \mid \Theta_e(h[x+1])$ is finite $\}$, then there exists one such $f[x+1]$, by Claim 24(c). *)
2. If and when such an $f[x+1]$ is found, then output $f(x)$.

End

Claim 25 *If* $\Theta_e(f[n])$ *is finite, then* $f[n] \subseteq \varphi_{prog(\mathbf{M}_{g(e)}(\Theta_e(f[n])))}$.

PROOF. If $n = 0$, then claim trivially holds. If $n > 0$, then by Claim 24(c), $\Theta_e(f[n-1]) \cdot f(n-1) \subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]))}$. Thus, by definition of *prog* and Claim 24(d), $f[n] \subseteq \varphi_{prog(\mathbf{M}_{g(e)}(\Theta_e(f[n])))}$. $\square$

We now define $\mathbf{M}$ as follows:

$$\mathbf{M}(f[m]) = \begin{cases} prog(\mathbf{M}_{g(e)}(\Theta_e(f[m]))), & \text{if } \Theta_e(f[m]) \text{ is finite;} \\ \uparrow, & \text{otherwise.} \end{cases}$$

It follows from Claim 24(f) and Claim 25 that $\mathbf{M}$ is consistent on each $f \in \mathcal{C}$. Since $\mathbf{M}_{g(e)}$ converges on $\Theta_e(f)$ for each $f \in \mathcal{C}$, it follows that $\mathbf{M}$ converges on each $f \in \mathcal{C}$. Thus, it follows by consistency of **M** that **M Ex**-identifies each $f \in \mathcal{C}$. It follows that $\mathcal{C} \in$ **Cons**. ∎

Note that the proof of Theorem 23 shows that given any recursive g, one can effectively construct an **M** such that, for any $\mathcal{C}$, if

for all e such that Θ_e is general recursive, $\Theta_e(\mathcal{C}) \subseteq \mathbf{Ex}(\mathbf{M}_{g(e)})$

then **M Cons**-identifies $\mathcal{C}$. Thus we have that **UniformRobustEx** $\subseteq$ **UniformRobustCons**. Since the reverse inclusion holds by definition, it follows that,

Corollary 26 **UniformRobustEx** $=$ **UniformRobustCons**.

Now we investigate the links between boundedness and uniform robust learnability of classes of functions. As a consequence, we derive that **UniformRobustEx** contains algorithmically rich classes, see Corollary 34 below.

Theorem 27 *Suppose* $\mathcal{S} \in$ **UniformRobustCons**. *If there exists an* x *such that* $\max(\{f(x) \mid f \in \mathcal{S}\}) = \infty$, *then* $\mathcal{S} \in$ **NUM**.

Proof. There is an infinite recursive binary tree without an infinite recursive branch where each node σ is either branching (having successors $\sigma \cdot 0$ and $\sigma \cdot 1$) or a leaf (having no successors). Such a tree can be obtained by starting with full binary tree, and then removing σ from this tree iff there exists a $j < |\sigma|/2$ such that $\sigma[2j] \subseteq \varphi_{j,|\sigma|}$.

Let this tree be named T and let $\tau_0, \tau_1, \ldots$ be a one-one recursive enumeration of its leaves; note that the length of each τ_k is at least 1.

Claim 28 *For any* $g \in \mathcal{R}_{0,1}$, *there exists a unique* f (*which is recursive*) *such that* $g = \tau_{f(0)} \cdot \tau_{f(1)} \cdot \tau_{f(2)} \ldots$. *Furthermore a program for* f *can be computed effectively from* g.

Proof. We inductively define g_i and n_i (g_i will be recursive). Let $g_0 = g$. Suppose g_i has been defined. Then we define n_i and g_{i+1} as follows. Since g_i is recursive, g_i is not an infinite branch of T. Let n_i be such that $\tau_{n_i} \subseteq g_i$ (note that n_i is unique). Let g_{i+1} be such that $g_i = \tau_{n_i} \cdot g_{i+1}$.

It is easy to verify that each g_i is recursive, and n_i can be effectively found from g and i. Let $f(i) = n_i$. Proposition is now easy to verify. □

For any recursive g, let $Imag(g)$ denote the unique f such that $g = \tau_{f(0)} \cdot \tau_{f(1)} \cdot \tau_{f(2)} \ldots$. By above proposition, program for $Imag(g)$ can be found effectively from program for g.

We now continue with the proof of the theorem. Let $\mathcal{S}$ and x be as given in the hypothesis. Suppose g is a recursive function such that, for all e such that Θ_e is general recursive, $\mathbf{M}_{g(e)}$ **Cons**-identifies $\Theta_e(\mathcal{S})$.

Let $U_i = \{\sigma \in \mathrm{SEG}_{0,1} \mid \mathbf{M}_{g(i)}(\sigma)\downarrow \ \wedge \ \sigma \subseteq \varphi_{\mathbf{M}_{g(i)}(\sigma)}\}$.

Note that U_i is r.e. Fix some uniform (in i) enumeration of U_i. Let $\sigma^{i,t}$ denote the least σ not enumerated in U_i within t steps. Note that $\sigma^{i,t}$ can be effectively computed from i, t.

By implicit use of Kleene recursion theorem [Rog67], there exists an e such that

$$\Theta_e(f) = \sigma^{e,f(x)} \cdot \tau_{f(0)} \cdot \tau_{f(1)} \cdots$$

Note that Θ_e is general recursive.

Claim 29 $U_e = \mathit{SEG}_{0,1}$.

Proof. Suppose by way of contradiction that $\mathrm{SEG}_{0,1} - U_e \neq \emptyset$. Let σ be the least element of $\mathrm{SEG}_{0,1} - U_e$. Thus, $\mathbf{M}_{g(e)}(\sigma)\uparrow$ or $\varphi_{\mathbf{M}_{g(e)}(\sigma)}$ does not extend σ. Thus $\mathbf{M}_{g(e)}$ does not **Cons**-identify any extension of σ.

Let t be large enough so that all elements of U_e which are less than σ have been enumerated in U_e within t steps. Then, for any $f \in \mathcal{S}$, such that $f(x) > t$, $\sigma^{e,f(x)} = \sigma$. Thus, $\Theta_e(f)$ extends σ. Therefore, $\mathbf{M}_{g(e)}$ does not **Cons**-identify $\Theta_e(f)$ (since $\mathbf{M}_{g(e)}$ does not **Cons**-identify any extension of σ).

A contradiction to $\mathbf{M}_{g(e)}$ **Cons**-identifying $\Theta_e(\mathcal{S})$. $\square$

It follows that $\mathbf{M}_{g(e)}$ $\mathcal{T}$**Cons**-identifies $\Theta_e(\mathcal{S})$. Since $\Theta_e(\mathcal{S}) \subseteq \mathcal{R}_{0,1}$, it follows by Corollary 20 that $\Theta_e(\mathcal{S}) \in \mathbf{NUM}$.

Thus there exists a recursively enumerable family, $f_0, f_1, \ldots$ of recursive functions in $\mathcal{R}_{0,1}$ such that $\Theta_e(\mathcal{S}) \subseteq \{f_i \mid i \in N\}$.

Let $g_{i,n} = f_i(n) f_i(n+1) f_i(n+2) \ldots$.

Thus, $\{g_{i,n} \mid i, n \in N\}$ is recursively enumerable. Suppose $f \in \mathcal{S}$, and $f_k = \Theta_e(f)$. Then, clearly, $\mathit{Imag}(g_{k,n}) = f$, for $n = |\sigma^{e,f(x)}|$. It follows that $\mathcal{S} \subseteq \{\mathit{Imag}(g_{i,n}) \mid i, n \in N\}$. Hence $\mathcal{S} \in \mathbf{NUM}$. ∎

Corollary 30 *Suppose* $\mathcal{S} \in$ **UniformRobustEx**. *If there exists an x such that* $\max(\{f(x) \mid f \in \mathcal{S}\}) = \infty$, *then* $\mathcal{S} \in \mathbf{NUM}$.

Proof. This follows immediately from Corollary 26 and Theorem 27. ∎

Definition 31 A function F is *1-generic* iff, for every recursively enumerable set $A \subseteq \mathrm{SEG}$, one of following holds:

(a) There exists an x such that $F[x] \in A$ or

(b) There exists an x such that for all $\sigma \in A$, $F[x] \not\subseteq \sigma$.

It can be shown that there exists a limiting recursive 1-generic F.

Theorem 32 *There is $\mathcal{S} \in$ **UniformRobustEx** such that $\mathcal{S}$ is unbounded.*

PROOF. Let F be a limiting recursive 1-generic function.

Claim 33 *For all general recursive operators Θ, one of* (a) *or* (b) *is satisfied.*

(a) *There exists a $\sigma \subseteq F$ such that $(\forall f \supseteq \sigma)[\Theta(F) = \Theta(f)]$, or*

(b) $\Theta(F)$ *is not recursive.*

PROOF. Suppose $\Theta(F)$ is recursive. Let $A = \{\sigma \mid \Theta(\sigma) \not\sim \Theta(F)\}$. Then, A is recursively enumerable. Moreover, for all $\sigma \in A$, $\sigma \not\subseteq F$. Thus, by definition of 1-generic function, there exists an x such that, for all $\sigma \in A$, $F[x] \not\subseteq \sigma$. It follows that for any f extending $F[x]$, $\Theta(f) = \Theta(F)$. Part (a) follows. □

Now define $\mathcal{S}$ as follows. Let

$$\eta_{i,r}(x) = \begin{cases} F(x), & \text{if } x < r; \\ \varphi_i(x), & \text{otherwise.} \end{cases}$$

$\mathcal{S} = \{\eta_{i,r} \mid i < r \text{ and } \eta_{i,r} \in \mathcal{R}\}$.

Clearly, $\mathcal{S}$ is unbounded (since for all recursive h, there exists an i, such that $\varphi_i \in \mathcal{R}$ and $\varphi_i(x) > h(x)$ infinitely often. Thus $\eta_{i,r} \in \mathcal{S}$, for all but finitely many r, but $\eta_{i,r}$ is not h-bounded).

We show that $\mathcal{S} \in$ **UniformRobustEx**. Note that there is a recursively enumerable approximation $G_0, G_1, \ldots$ of total recursive functions to F.

Let h be as in Corollary 22. By parameterized s-m-n theorem, let g be a recursive function such that $\mathbf{M}_{g(e)}$ may be defined as follows.

$\mathbf{M}_{g(e)}(f[n])$

1. If there is a $k \leq n$ such that $\Theta_e(G_k) \supseteq f[n]$, then output a canonical index for $\Theta_e(G_k)$ for the least such k.
2. Otherwise let m be least number such that $\Theta_e(G_n[m]) \not\sim f[n]$. Let $\eta'_{i,r}$

denote the function:

$$\eta'_{i,r}(x) = \begin{cases} G_n(x), & \text{if } x < r; \\ \varphi_i(x), & \text{otherwise.} \end{cases}$$

Let m' be such that, for all $i, r \leq m$, there exists a program $j \leq m'$ for $\eta'_{i,r}$.

Output $\mathbf{M}_{h(m',e)}(f[n])$.

End

We claim that $\mathbf{M}_{g(e)}$ **Ex**-identifies $\Theta_e(\mathcal{S})$. To see this, suppose $f \in \Theta_e(\mathcal{S})$.

Case 1: There exists a k such that $\Theta_e(G_k) = f$.

In this case by step 1 of the construction of $\mathbf{M}_{g(e)}$, we have that $\mathbf{M}_{g(e)}$ **Ex**-identifies f.

Case 2: Not Case 1.

In this case we first claim that $\Theta_e(F) \neq f$. If this were not the case, then since f is recursive, by Claim 33 there would exist an m such that, for all $g \supseteq F[m]$, $\Theta_e(g) = f$. Since for all but finitely many k, $F[m] \subseteq G_k$, we would have Case 1 (that is, $\Theta_e(G_k) = f$).

Thus, for all but finitely many n, $\mathbf{M}_{g(e)}$ executes step 2 on input $f[n]$, and m as computed in step 2 is the least number such that $\Theta_e(F[m]) \not\sim f$, and $F[m] \subseteq G_n$.

But then, by construction of $\mathcal{S}$, there exists an $i < r \leq m$, such that $\eta_{i,r} \in \mathcal{S}$ and $f = \Theta_e(\eta_{i,r})$.

But then, for all but finitely many n, the value m' as defined in $\mathbf{M}_{g(e)}(f[n])$ converges (as n goes to ∞) and is a bound on a program for $\eta_{i,r}$. It follows that $\mathbf{M}_{h(m',e)}$ and thus $\mathbf{M}_{g(e)}$ **Ex**-identifies f.

Above cases prove the theorem. ∎

Corollary 34 **UniformRobustEx** $\supset$ **NUM**.

PROOF. The inclusion follows from Corollary 16(b) and Theorem 10(a). The proper inclusion now follows from Theorem 32, since every class from **NUM** is bounded. ∎

0

Fig. 1. Robust Learning and Consistency/Conformity

4 Robustly Rich and Robustly Poor Learning

In this section, we show for several learning types that they are "robustly rich"; that is, these types contain classes being both robustly learnable and not contained in any recursively enumerable class. Notice that in proving these types robustly rich below, in general, we show some stronger results, namely, we *robustly separate* (or even *uniformly* robustly separate) the corresponding types from some other types, thereby strengthening known separations in a (uniformly) robust way. From these separations, the corresponding richness results follow easily by applying Theorem 10. On the other hand, there are also "robustly poor" learning types; that is, every robustly learnable class of these types is contained in a recursively enumerable class. Some further robustly poor types will be exhibited in Section 5.

A summary of results relating robust consistent and conforming learning is shown in Figure 1. Note that the relationship between **Cons** and **RobustConf** is open at this point (along with whether **RobustCons** is a proper subset of **RobustConf**).

Besides the results given in Figure 1, we also give results involving confident learning in this section.

Theorem 35 **UniformRobustCons** $\not\subseteq$ $\mathcal{R}$**Conf**. *(Moreover the non-inclusion can be witnessed using a class from* $\mathcal{R}_{0,1}$*.)*

Proof. Let F be a limiting recursive function which dominates every recursive function, and $F(x) < F(x+1)$, for all x. Let $F_s(\cdot)$ denote recursive approximation of F.

Let $\mathcal{C} = \{f \in \mathcal{R}_{0,1} \mid (\exists e)[f(e) = 1, (\forall x < e)[f(x) = 0]$, and $(\exists j \leq F(e))[\varphi_j = f]]\}$.

Claim 36 $\mathcal{C} \in$ **UniformRobustEx**.

Proof. Let h be as in Corollary 22. Let g be a recursive function such that $\mathbf{M}_{g(e)}$ may be defined as follows.

$\mathbf{M}_{g(e)}(f[n])$
1. If $\Theta_e(\text{Zero}[n]) \sim f[n]$, then output a standard program for $\Theta_e(\text{Zero})$.
2. Else let m be least number such that $\Theta_e(0^m) \not\sim f[n]$.
   Output $\mathbf{M}_{h(F_n(m),e)}(f[n])$.

End

Suppose Θ_e is general recursive. Then we claim that $\mathbf{M}_{g(e)}$ **Ex**-identifies $\Theta_e(\mathcal{C})$. To see this, suppose $f \in \Theta_e(\mathcal{C})$. If $f = \Theta_e(\text{Zero})$, then clearly, by step 1, $\mathbf{M}_{g(e)}$ **Ex**-identifies f.

If $f \neq \Theta_e(\text{Zero})$, then let m be least number such that $f \not\sim \Theta_e(\text{Zero}[m])$. Thus, if $f = \Theta_e(\eta)$, then $\eta(x) \neq 0$, for some $x < m$. Therefore, by definition of $\mathcal{C}$, there exists a $j \leq F(m)$, such that $f = \Theta_e(\varphi_j)$. Now it follows from definition of h in Corollary 22 that $\mathbf{M}_{h(F(m),e)}$ **Ex**-identifies f, and thus $\mathbf{M}_{g(e)}$ **Ex**-identifies f. $\Box$

Thus $\mathcal{C} \in$ **UniformRobustCons** by Corollary 26.

We now show that $\mathcal{C} \notin \mathcal{R}$**Conf**. Assume by way of contradiction that $\mathcal{C} \in \mathcal{R}$**Conf** as witnessed by $\mathbf{M}$.

Claim 37 *For all but finitely many* j*, for all* σ *extending* 0^j1*, if* $\mathbf{M}(\sigma \cdot 0) = \mathbf{M}(\sigma \cdot 1)$*, then* $\varphi_{\mathbf{M}(\sigma\cdot 0)}(|\sigma|)\uparrow$.

Proof. Let f be such that for $j \in N$, $\varphi_{f(j)}$ is defined as follows:

Search for a $\sigma \in \text{SEG}_{0,1}$ extending 0^j1, such that $\mathbf{M}(\sigma \cdot 0) = \mathbf{M}(\sigma \cdot 1)$, and $\varphi_{\mathbf{M}(\sigma\cdot 0)}(|\sigma|)\downarrow$. If and when such a σ is found, let $\varphi_{f(j)}$ be $(1 \dot- \varphi_{\mathbf{M}(\sigma\cdot 0)}(|\sigma|))$-extension of σ.

Note that for all but finitely many j, $f(j) \leq F(j)$. Let j_0 be such that for all $j \geq j_0$, $f(j) \leq F(j)$.

Thus, for all $j \geq j_0$, if there exists a σ extending 0^j1 such that $\mathbf{M}(\sigma \cdot 0) = \mathbf{M}(\sigma \cdot 1)$, and $\varphi_{\mathbf{M}(\sigma \cdot 0)}(|\sigma|)\downarrow$, then $\varphi_{f(j)}$ is total and belongs to $\mathcal{C}$, and $\mathbf{M}$ is not conforming on $\varphi_{f(j)}$. A contradiction to $\mathbf{M}$ $\mathcal{R}$**Conf**-identifying $\mathcal{C}$.

Thus, for all $j \geq j_0$, for all σ extending 0^j1, if $\mathbf{M}(\sigma \cdot 0) = \mathbf{M}(\sigma \cdot 1)$, then $\varphi_{\mathbf{M}(\sigma \cdot 0)}(|\sigma|)\uparrow$. $\square$

Now for $\sigma \in \mathrm{SEG}_{0,1}$ define f_σ as follows:

$$f_\sigma(x) = \begin{cases} \sigma(x), & \text{if } x < |\sigma|; \\ 1, & \text{if } \mathbf{M}(f_\sigma[x] \cdot 1) = \mathbf{M}(f_\sigma[x]), \text{ and} \\ & \mathbf{M}(f_\sigma[x] \cdot 0) \neq \mathbf{M}(f_\sigma[x]); \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, f_σ is total for each $\sigma \in \mathrm{SEG}_{0,1}$.

Claim 38 *For all but finitely many j, for all $g \in \mathcal{C}$ such that $0^j1 \subseteq g$, there exists a τ such that $g = f_\tau$.*

Proof. Consider j such that,

for all $\sigma \in \mathrm{SEG}_{0,1}$ extending 0^j1, if $\mathbf{M}(\sigma \cdot 0) = \mathbf{M}(\sigma \cdot 1)$, then $\varphi_{\mathbf{M}(\sigma \cdot 0)}(|\sigma|)\uparrow$.

Note that all but finitely many j satisfy the above condition (by Claim 37). Now suppose $0^j1 \subseteq g \in \mathcal{C}$. Let $\tau \supseteq 0^j1$ be such that $\tau \subseteq g$, and for all τ', $\tau \subseteq \tau' \subseteq g$, $\mathbf{M}(\tau) = \mathbf{M}(\tau')$. Then, it follows that $f_\tau = g$. $\square$

From the above claim it follows that $\mathcal{C}$ is in **NUM** (since, for any j, there are only finitely many functions in $\mathcal{C}$ with prefix 0^j1).

However, $\mathcal{C} \notin$ **NUM**, since $\{f \in \mathcal{R} \mid \varphi_j = f \text{ and } 0^j1 \subseteq f\} \subseteq \mathcal{C}$, but $\{f \in \mathcal{R} \mid \varphi_j = f \text{ and } 0^j1 \subseteq f\} \notin$ **NUM**. ▮

Corollary 39 **UniformRobustCons** $\supset$ **NUM**.

Proof. The inclusion follows from Corollary 16(b) and Theorem 10(a,b). The proper inclusion now follows from Theorem 35 and Theorem 10(a). ▮

Note that the class $\mathcal{S}$ used in proof of Theorem 32, is also not in $\mathcal{R}$**Conf**. So we can have an alternative proof for above theorem using the $\mathcal{S}$ in proof of Theorem 32. However, our current proof is useful for some other results in the paper.

We now consider robust separation of $\mathcal{R}$**Cons** and $\mathcal{T}$**Cons**, see Theorem 44 below. We first show the following proposition and lemma.

Proposition 40 *Suppose Θ_k is general recursive. Then for all n, there exists an m such that, if $\tau \in SEG_{0,1}$, and $|\tau| \geq m$, then* $\mathrm{domain}(\Theta_k(\tau)) \supseteq \{0, 1, \ldots, n-1\}$.

PROOF. This follows from König's Lemma and general recursiveness of Θ_k. ∎

Lemma 41 *There exists a recursive p such that* (a) *to* (e) *are satisfied.*

(a) $\varphi_{p(i)}(i) = 1$ *and* $\varphi_{p(i)}(x) = 0$, *for* $x < i$.

(b) $\varphi_{p(i)} \in \mathcal{R}_{0,1}$ *or* $\varphi_{p(i)} \in SEG_{0,1}$.

(c) *For all i, for all $k \leq i$, either:*

(i) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero})$ *Or*
(ii) *for all $f \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$,* $[\Theta_k(f) \sim \Theta_k(\mathrm{Zero})]$.

(d) *For all k such that Θ_k is general recursive, for all $i, j \geq k$ such that $i \neq j$, Either,*

(i) $\Theta_k(\varphi_{p(i)}) \sim \Theta_k(\mathrm{Zero})$ *Or*
(ii) $\Theta_k(\varphi_{p(j)}) \sim \Theta_k(\mathrm{Zero})$ *Or*
(iii) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\varphi_{p(j)})$ *Or*
(iv) *for all $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$ and $\varphi_{p(j)} \subseteq g$,* $[\Theta_k(f) \sim \Theta_k(g)]$.

(e) $\mathcal{C} = \{\varphi_{p(i)} \mid i \in N, \varphi_{p(i)} \in \mathcal{R}\} \notin$ **NUM**.

PROOF. By operator recursion theorem [Cas74], there exists a recursive p such that $\varphi_{p(i)}$ may be described as follows.

Initially, we let $\varphi_{p(i)}(x) = 0$, for $x < i$, and $\varphi_{p(i)}(i) = 1$. Let $\varphi^s_{p(i)}$ denote the portion of $\varphi_{p(i)}$ defined before stage s. It will always be the case that domain of $\varphi^s_{p(i)}$ is an initial segment of N.

For all i, for all $k \leq i$, let $I(i, k, s)$ denote the predicate $\Theta_k(\varphi^s_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[s])$. Intuitively $I(i, k, s)$ being true denotes that $\Theta_k(\varphi_{p(i)})$ has been made nonconforming with $\Theta_k(\mathrm{Zero})$. Clearly, if $I(i, k, s)$ is true, then $I(i, k, s+1)$ is also true.

Initially, let $\mathrm{Diag}(i) = \emptyset$, for each $i \in N$. Intuitively, $\mathrm{Diag}(i)$ denotes a collection of j such that in the construction we have made $\varphi_{p(i)} \neq \varphi_{\varphi_i(j)}$. Clearly, $\mathrm{Diag}(i)$ is monotonically non-decreasing.

Intuitively, in stages below, if $s \bmod 3 = 0$, then we try to work on satisfying (c), if $s \bmod 3 = 1$, then we work on satisfying (d) and if $s \bmod 3 = 2$, then we work on satisfying (e). Let x_i^s denote the least x not in the domain of $\varphi_{p(i)}^s$. Go to stage 0.

Stage s

If $s \bmod 3 = 0$, then

1. For $i \leq s$ Do,
    2. If there exists a $k \leq i$, and $\tau \in \mathrm{SEG}_{0,1}$, such that following are satisfied
        2.1. $\Theta_k(\varphi_{p(i)}^s) \sim \Theta_k(\mathrm{Zero}[s])$, and
        2.2. $\varphi_{p(i)}^s \subseteq \tau$ and $|\tau| \leq s$, and
        2.3. $\Theta_k(\tau) \not\sim \Theta_k(\mathrm{Zero}[s])$,

        Then pick least such k and a corresponding τ and make $\varphi_{p(i)}^{s+1} = \tau$.

    EndFor
3. Go to stage $s+1$.

If $s \bmod 3 = 1$, then

4. If there exist $k, i, j \in N$ and $\tau, \gamma \in \mathrm{SEG}_{0,1}$, such that the following are satisfied
    4.1. $k \leq i < j \leq s$, and
    4.2. $I(i, k, s)$ and $I(j, k, s)$ are both true, and
    4.3. $\Theta_k(\varphi_{p(i)}^s) \sim \Theta_k(\varphi_{p(j)}^s)$, and
    4.4. $|\tau| \leq s$ and $|\gamma| \leq s$, and
    4.5. $\varphi_{p(i)}^s \subseteq \tau$, $\varphi_{p(j)}^s \subseteq \gamma$, and
    4.6. $\Theta_k(\tau) \not\sim \Theta_k(\gamma)$,

    Then for least such triple (i, j, k), (in some ordering of all triples), pick one such corresponding τ and γ, and make $\varphi_{p(i)}^{s+1} = \tau$ and $\varphi_{p(j)}^{s+1} = \gamma$.
5. Go to stage $s+1$.

If $s \bmod 3 = 2$ then

6. For $i \leq s$ Do
    7. Let j be the least number not in $\mathrm{Diag}(i)$.
    8. If $\Phi_i(j) \leq s$ and $\Phi_{\varphi_i(j)}(x_i^s) \leq s$,

        Then let $\varphi_{p(i)}(x_i^s) = 1 \dot{-} \varphi_{\varphi_i(j)}(x_i^s)$, and $\mathrm{Diag}(i) = \mathrm{Diag}(i) \cup \{j\}$.

    Endfor
9. Go to stage $s+1$.

End stage s

Clearly, (a) and (b) of claim hold.

We now consider (c). Consider any $i = i_0$, and $k = k_0 \leq i_0$. If $\varphi_{p(i_0)}$ is total, then (c) clearly holds. So suppose $\varphi_{p(i_0)}$ is in $\mathrm{SEG}_{0,1}$. Suppose by way of contradiction that $\Theta_{k_0}(\varphi_{p(i_0)}) \sim \Theta_{k_0}(\text{Zero})$, and there exists an extension $f \in \mathcal{R}_{0,1}$ of $\varphi_{p(i_0)}$ such that $\Theta_{k_0}(f) \not\sim \Theta_{k_0}(\text{Zero})$. Then there exists a finite τ, $\varphi_{p(0)} \subseteq \tau \subseteq f$ and an n such that $\Theta_{k_0}(\tau) \not\sim \Theta_{k_0}(\text{Zero}[n])$. Thus for some stage $s > \max(|\tau|, n, i_0)$ and $s \bmod 3 = 0$, in step 2 of stage s, $\varphi_{p(i_0)}$ would be extended so as to make $\Theta_{k_0}(\varphi_{p(i_0)}) \not\sim \Theta_{k_0}(\text{Zero})$. A contradiction.

Thus (c) is satisfied.

We now consider (d). Consider $i = i_0, j = j_0$, and $k = k_0 \leq \min(i_0, j_0)$. Without loss of generality assume $i_0 < j_0$. Suppose $\lim_{s\to\infty} I(i_0, k_0, s)$ and $\lim_{s\to\infty} I(j_0, k_0, s)$ are both true (otherwise (d) trivially holds). Suppose by way of contradiction that $\Theta_{k_0}(\varphi_{p(i_0)}) \sim \Theta_{k_0}(\varphi_{p(j_0)})$ and there exist total $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i_0)} \subseteq f$, $\varphi_{p(j_0)} \subseteq g$, and $\Theta_{k_0}(f) \not\sim \Theta_{k_0}(g)$. Let n be such that $\Theta_{k_0}(f[n]) \not\sim \Theta_{k_0}(g[n])$. Let s be large enough such that $I(i_0, k, s)$ and $I(j_0, k, s)$ are both true, $[\varphi^s_{p(i_0)} = \varphi_{p(i_0)}$ or $\varphi^s_{p(i_0)} \supseteq \varphi_{p(i_0)}[n]]$ and $[\varphi^s_{p(j_0)} = \varphi_{p(j_0)}$ or $\varphi^s_{p(j_0)} \supseteq \varphi_{p(j_0)}[n]]$. Thus, for any $s' > s$ such that $s' \bmod 3 = 1$, (i_0, j_0, k_0) would be possible candidate in step 4, and for large enough s' it will be the smallest triple satisfying 4.1–4.6. Thus, it would be ensured that $\Theta_{k_0}(\varphi_{p(i_0)}) \not\sim \Theta_{k_0}(\varphi_{p(j_0)})$.

Thus (d) holds.

For (e), first note that $\varphi_{p(i)}$ may be modified in stages $s \bmod 3 = 0$ only finitely often, since there are only finitely many $k \leq i$. Similarly, for each $k \leq i$, $\varphi_{p(i)}$ may be modified in stages $s \bmod 3 = 1$ only finitely often, since if $I(i, k, s)$, then there exists a finite n such that $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$, and thus, only $j < n$ can satisfy (4.3). It follows that for all i, for all but finitely many stages, if $s \bmod 3 \neq 2$, then $\varphi^s_{p(i)} = \varphi^{s+1}_{p(i)}$.

Now suppose by way of contradiction that $\mathcal{C} \in \mathbf{NUM}$. Let i be such that φ_i is total, and $\mathcal{C} \subseteq \{\varphi_{\varphi_i(j)} \mid j \in N\} \subseteq \mathcal{R}$. Then, consider $\varphi_{p(i)}$. It is clear that $\varphi_{p(i)}$ is total (due to step 6–8). Also, Diag(i) contains every j eventually. Thus, for all j, $\varphi_{p(i)} \neq \varphi_{\varphi_i(j)}$, and hence $\varphi_{p(i)} \notin \{\varphi_{\varphi_i(j)} \mid j \in N\}$. A contradiction.

Thus (e) holds. ∎

Corollary 42 *Suppose p is as in Lemma 41. Suppose Θ_k is general recursive operator. Then given any i and n, exactly one of* (a) *and* (b) *below holds, and one can effectively determine which of* (a) *and* (b) *holds.*

(a) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$.

(b) *For all $f \in \mathcal{R}_{0,1}$ extending $\varphi_{p(i)}$, $\Theta_k(f) \sim \Theta_k(\text{Zero}[n])$.*

Proof. By Lemma 41(c), we have:

(i) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero})$ or

(ii) For all $f \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$, $[\Theta_k(f) \sim \Theta_k(\text{Zero})]$.

If (ii) holds, then clearly, (b) of corollary holds. So suppose (i) holds. It follows that either $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$ or $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\text{Zero}[n])$. If $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$ then (a) holds, and if $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\text{Zero}[n])$ then (b) holds.

We now show how to effectively determine which of (a) and (b) holds.

Note that if (a) holds, then there exists a t such that $\Theta_k(\varphi_{p(i),t}) \not\sim \Theta_k(\text{Zero}[n])$. If (b) holds, then either

(iii) $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\text{Zero}[n])$ or

(iv) $\varphi_{p(i)}$ is finite and for all $\tau \in \text{SEG}_{0,1}$ such that $\varphi_{p(i)} \subseteq \tau$, $\Theta_k(\tau) \sim \Theta_k(\text{Zero}[n])$.

Thus, if (b) holds, then there exists a t such that

(iii') $\Theta_k(\varphi_{p(i),t}) \supseteq \Theta_k(\text{Zero}[n])$ or

(iv') For all $\tau \in \text{SEG}_{0,1}$ such that $|\tau| = t$ and $\varphi_{p(i),t} \sim \tau$, $\Theta_k(\tau) \supseteq \Theta_k(\text{Zero}[n])$

(where (iv') follows from (iv) using Proposition 40).

Thus, if (b) holds, we can effectively find a witness for it.

It follows that one can determine effectively which of (a) and (b) holds. ∎

Corollary 43 *Suppose p is as in Lemma 41. Suppose Θ_k is general recursive operator. Then given any i, j and n, exactly one of* (a), (b) *and* (c) *holds, and one can effectively determine which of* (a), (b) *and* (c) *holds.*

(a) [*For all $f \in \mathcal{R}_{0,1}$, such that $\varphi_{p(i)} \subseteq f$, $\Theta_k(f) \sim \Theta_k(\text{Zero}[n])$*] *or* [*For all $g \in \mathcal{R}_{0,1}$, such that $\varphi_{p(j)} \subseteq g$, $\Theta_k(g) \sim \Theta_k(\text{Zero}[n])$*];

(b) $[\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$ *and* $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\text{Zero}[n])]$ *and there exists an $x < n$, such that $\Theta_k(\varphi_{p(i)})(x) \neq \Theta_k(\varphi_{p(j)})(x)$,*

(c) $[\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$ *and* $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\text{Zero}[n])]$ *and for all $f \in \mathcal{R}_{0,1}$ extending $\varphi_{p(i)}$, for all $g \in \mathcal{R}_{0,1}$ extending $\varphi_{p(j)}$, $\Theta_k(f)[n] = \Theta_k(g)[n]$.*

Proof. Note that using Corollary 42 one can effectively determine whether (a) holds or $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])$ and $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\text{Zero}[n])$. We now show that if (a) does not hold then one of (b) or (c) holds, and one can effectively determine which one.

By Lemma 41(d), if (a) does not hold then we must have

(iii) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\varphi_{p(j)})$ or

(iv) For all $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$ and $\varphi_{p(j)} \subseteq g$, $[\Theta_k(f) \sim \Theta_k(g)]$.

If (iii) holds, then clearly, either (b) of corollary holds or both $\Theta_k(\varphi_{p(i)})$ and $\Theta_k(\varphi_{p(j)})$ have domain a superset of $\{0, 1, \ldots, n-1\}$ and $\Theta_k(\varphi_{p(i)})[n] = \Theta_k(\varphi_{p(j)})[n]$, and thus (c) is satisfied. Moreover witness for both of above cases can be effectively found (since in first case there exist $x < n$ and t such that $\Theta_k(\varphi_{p(i),t})(x) \neq \Theta_k(\varphi_{p(j),t})(x)$, and in second case there exists a t such that domains of $\Theta_k(\varphi_{p(i),t})$ and $\Theta_k(\varphi_{p(j),t})$, are supersets of $\{0, 1, \ldots, n-1\}$ and $\Theta_k(\varphi_{p(i),t})[n] = \Theta_k(\varphi_{p(j),t})[n]$).

So suppose (iv) holds. Then clearly, (c) of corollary holds. Thus, by Proposition 40, there exists a t such that

(iv') For all τ such that $|\tau| = t$ and $\varphi_{p(i),t} \sim \tau$, for all γ such that $|\gamma| = t$ and $\varphi_{p(j),t} \sim \gamma$,

$\Theta_k(\tau) \sim \Theta_k(\gamma)$, and domain of both $\Theta_k(\tau)$ and $\Theta_k(\gamma)$ is a superset of $\{0, \ldots, n-1\}$.

Thus, we have a witness for (c) holding.

It follows that one can determine effectively which of (a), (b) or (c) holds. ∎

Theorem 44 **Robust$\mathcal{R}$Cons** $\not\subseteq$ $\mathcal{T}$**Cons**. (*Moreover the non-inclusion can be witnessed using a class from* $\mathcal{R}_{0,1}$.)

Proof. Let p be as in Lemma 41.

Let $\mathcal{C} = \{\varphi_{p(i)} \mid i \in N \wedge \varphi_{p(i)} \in \mathcal{R}\}$. Then, by part (b) of Lemma 41, $\mathcal{C} \subseteq \mathcal{R}_{0,1}$, and, by part (e), $\mathcal{C} \notin$ **NUM**. Thus, by Corollary 20, $\mathcal{C} \notin \mathcal{T}$**Cons**.

We now show that $\mathcal{C} \in$ **Robust$\mathcal{R}$Cons**. To show this, suppose Θ_k is a general recursive operator. We need to show that $\Theta_k(\mathcal{C}) \in \mathcal{R}$**Cons**.

Let $\mathcal{S} = \{\varphi_{p(i)} \in \mathcal{C} \mid i \leq k\}$.

Define **M** as follows:

M$(g[m])$
1. If $g[m] \sim \Theta_k(\text{Zero})$, then output a standard program for $\Theta(\text{Zero})$.
2. Else If $g[m] \sim \Theta_k(f)$, for some $f \in \mathcal{S}$, then output a standard program for one such $\Theta_k(f)$.

3. Else let n be such that $\Theta_k(\text{Zero}[n]) \not\sim g[m]$.
   (* Note that if $g \in \Theta_k(\mathcal{C})$, and we are at this point of the construction, then g must be $\Theta_k(\varphi_{p(i)})$, for some i such that $k \leq i < n$. *)
4. Using Corollary 42 determine $A = \{i \mid k \leq i < n \ \wedge \ \Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])\}$.
   (* Note that by Corollary 42, for $k \leq i < n$, and $i \notin A$ — either $\varphi_{p(i)}$ is not total or $\Theta_k(\varphi_{p(i)}) \sim \Theta_k(\text{Zero}[n]) \not\sim g[m]$ — thus, $g \neq \Theta_k(\varphi_{p(i)})$. *)
5. Let B be a subset of A such that:
   for $i \in A - B$, there exists an $x < m$ such that $\Theta_k(\varphi_{p(i)})(x) \neq g(x)$;
   for distinct $i, j \in B$, for all $f, h \in \mathcal{R}_{0,1}$ such that $f \supseteq \varphi_{p(i)}$ and $h \supseteq \varphi_{p(j)}$, $\Theta_k(f)[m] = \Theta_k(h)[m]$.
   (* Note that, by Corollary 43, there exists such a B, and one such B can be effectively found. *)
   (* Note that elements in $A - B$ can be safely ignored. *)
6. Output (standard) program for $\text{Union}(\{\Theta_k(\varphi_{p(i)}) \mid i \in B\})$.

End

We claim that above $\mathbf{M}$ $\mathcal{R}\mathbf{Cons}$-identifies $\Theta_k(\mathcal{C})$.

Clearly, $\mathbf{M}$ is defined on each initial segment. Suppose $g \in \Theta_k(\mathcal{C})$. To see that $\mathbf{M}$ **Cons**-identifies g we argue as follows.

If $g = \Theta_k(\text{Zero})$ or if $g = \Theta_k(\varphi_{p(i)})$, for some $i \leq k$, then clearly by steps 1 and 2, $\mathbf{M}$ is consistent on g and **Ex**-identifies g.

If $g \neq \Theta_k(\text{Zero})$, then let n be least value such that $g \not\sim \Theta_k(\text{Zero}[n])$. Thus, g must be one of $\Theta_k(\varphi_{p(i)})$, for $i < n$. We have already handled the case $i \leq k$, so assume $k < i < n$ for the following.

We first show that $\mathbf{M}$ is consistent on these g. Note that steps 1 and 2 are clearly consistent with the input. So we only consider m such that $\mathbf{M}(g[m])$ in above construction reaches step 3. For these m, by construction of A, if $i \notin A$, then either $\varphi_{p(i)}$ is not total, or $\Theta_k(\varphi_{p(i)})$ is conforming with $\Theta_k(\text{Zero}[n])$ and thus not equal to g. Similarly, if $i \in A - B$, then $\varphi_{p(i)}$ is not total or $\Theta_k(\varphi_{p(i)})$ is nonconforming with g.

Now, since $g \in \Theta_k(\mathcal{C})$, there exists an $i \in B$ such that $\Theta_k(\varphi_{p(i)}) = g$, and since for all $i', j' \in B$, $\Theta_k(\varphi_{p(i')})$ and $\Theta_k(\varphi_{p(j')})$ do not differ on $\{0, 1, \ldots, m-1\}$, we have that $g[m] \subseteq \text{Union}(\{\Theta_k(\varphi_{p(i)}) \mid i \in B\})$. Thus, $\mathbf{M}$ is consistent on g.

To show that $\mathbf{M}$ **Ex**-identifies g, note that B is monotonically non-increasing (with respect to $g[m]$). Also the limiting (with respect to m) value of B satisfies the property:

for any $i, j \in B$, for any $f, h \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$ and $\varphi_{p(j)} \subseteq h$,

$[\Theta_k(f) \sim \Theta_k(h)]$.

Using the fact that $g = \Theta_k(\varphi_{p(i)})$ for some $i \in B$, it immediately follows that **M Ex**-identifies g. ∎

Corollary 45 **Robust$\mathcal{R}$Cons** $\supset$ **NUM**.

Proof. The inclusion follows from Corollary 16(b) and Theorem 10(a). The proper inclusion now follows from Theorem 44 and Theorem 10(a). ∎

We now show that $\mathcal{R}$**Conf** and $\mathcal{R}$**Cons** can be separated robustly, see Theorem 49 below. A portion of the proof is similar to that in Theorem 44. We need the following modification of Lemma 41.

Lemma 46 *There exists a recursive p such that* (a) *to* (e) *are satisfied.*

(a) $\varphi_{p(i)}(i) = 1$ *and* $\varphi_{p(i)}(x) = 0$*, for* $x < i$.

(b) $\varphi_{p(i)} \in \mathcal{R}_{0,1}$ *or* $\varphi_{p(i)} \in SEG_{0,1}$.

(c) *For all i, for all $k \le i$, Either:*

(i) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero})$*, Or*
(ii) *for all $f \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$,* $[\Theta_k(f) \sim \Theta_k(\text{Zero})]$.

(d) *For all k such that Θ_k is general recursive, for all $i, j \ge k$ such that $i \ne j$, Either:*

(i) $\Theta_k(\varphi_{p(i)}) \sim \Theta_k(\text{Zero})$*, Or*
(ii) $\Theta_k(\varphi_{p(j)}) \sim \Theta_k(\text{Zero})$*, Or*
(iii) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\varphi_{p(j)})$*, Or*
(iv) *for all $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$ and $\varphi_{p(j)} \subseteq g$,* $[\Theta_k(f) \sim \Theta_k(g)]$.

(e) *Let* $\mathcal{C} = \{\varphi_{p(i)} \in \mathcal{R} \mid \mathbf{M}_i$ *is total* $\} \cup \{r$*-extension of* $\varphi_{p(i)} \mid \mathbf{M}_i$ *is total and* $\varphi_{p(i)}$ *is not total and* $r \in \{0,1\}\}$.

Then $\mathcal{C} \notin \mathcal{R}$**Cons**, *and there exists a recursive H such that following two conditions are satisfied.*

(e.1) *If* $\mathbf{M}_i$ *is total, then* $\lim_{t\to\infty} H(i,t)\downarrow = 1$ *iff* $\varphi_{p(i)}$ *is total, and* $\lim_{t\to\infty} H(i,t)\downarrow = 0$ *iff* $\varphi_{p(i)}$ *is not total.*

(e.2) *If* $\mathbf{M}_i$ *is not total, then* $\lim_{t\to\infty} H(i,t)$ *is defined. Moreover, if* $\lim_{t\to\infty} H(i,t) = 0$*, then* $\varphi_{p(i)}$ *is not total.*

Proof. Proof of this lemma is a modification of proof of Lemma 41 where we make appropriate changes so that condition (e) is satisfied.

By operator recursion theorem [Cas74], there exists a recursive p such that $\varphi_{p(i)}$ may be described as follows.

Initially, we let $\varphi_{p(i)}(x) = 0$, for $x < i$, and $\varphi_{p(i)}(i) = 1$. Let $\varphi_{p(i)}^s$ denote the portion of $\varphi_{p(i)}$ defined before stage s. It will always be the case that domain of $\varphi_{p(i)}^s$ is an initial segment of N.

For all i, for all $k \leq i$, let $I(i, k, s)$ denote the predicate $\Theta_k(\varphi_{p(i)}^s) \not\sim \Theta_k(\text{Zero}[s])$. Intuitively $I(i, k, s)$ being true denotes that $\Theta_k(\varphi_{p(i)})$ has been made nonconforming with $\Theta_k(\text{Zero})$. Clearly, if $I(i, k, s)$ is true, then $I(i, k, s + 1)$ is also true.

Intuitively, in stages below, if $s \bmod 3 = 0$, then we try to work on satisfying (c), if $s \bmod 3 = 1$, then we work on satisfying (d) and if $s \bmod 3 = 2$, then we work on satisfying (e). Let x_i^s denote the least x not in the domain of $\varphi_{p(i)}^s$. Go to stage 0.

Stage s

If $s \bmod 3 = 0$, then

1. For $i \leq s$ Do,

   2. If there exists a $k \leq i$, and $\tau \in \text{SEG}_{0,1}$, such that following are satisfied

      2.1. $\Theta_k(\varphi_{p(i)}^s) \sim \Theta_k(\text{Zero}[s])$, and

      2.2. $\varphi_{p(i)}^s \subseteq \tau$ and $|\tau| \leq s$, and

      2.3. $\Theta_k(\tau) \not\sim \Theta_k(\text{Zero}[s])$,

      Then pick least such k and a corresponding τ and make $\varphi_{p(i)}^{s+1} = \tau$.

   EndFor

3. Go to stage $s + 1$.

If $s \bmod 3 = 1$, then

4. If there exist $k, i, j \in N$ and $\tau, \gamma \in \text{SEG}_{0,1}$, such that the following are satisfied

   4.1. $k \leq i < j \leq s$, and

   4.2. $I(i, k, s)$ and $I(j, k, s)$ are both true, and

   4.3. $\Theta_k(\varphi_{p(i)}^s) \sim \Theta_k(\varphi_{p(j)}^s)$, and

   4.4. $|\tau| \leq s$ and $|\gamma| \leq s$, and

   4.5. $\varphi_{p(i)}^s \subseteq \tau$, $\varphi_{p(j)}^s \subseteq \gamma$, and

   4.6. $\Theta_k(\tau) \not\sim \Theta_k(\gamma)$,

   Then for least such triple (i, j, k), (in some ordering of all triples), pick one such corresponding τ and γ, and make $\varphi_{p(i)}^{s+1} = \tau$ and $\varphi_{p(j)}^{s+1} = \gamma$.

5. Go to stage $s + 1$.

If $s \bmod 3 = 2$ then

6. If $\mathbf{M}_i(\varphi_{p(i)}^s \cdot 0)$ and $\mathbf{M}_i(\varphi_{p(i)}^s \cdot 1)$ both converge within s steps, then

   7. If $\mathbf{M}_i(\varphi_{p(i)}^s \cdot 0) \neq \mathbf{M}_i(\varphi_{p(i)}^s \cdot 1)$,

7.1 then let $w \in \{0,1\}$ be such that $\mathbf{M}_i(\varphi^s_{p(i)} \cdot w) \neq \mathbf{M}_i(\varphi^s_{p(i)})$, and set $\varphi_{p(i)}(x^s_i) = w$ and go to stage $s+1$.

7.2 else go to stage $s+1$.

Else go to stage $s+1$.

End stage s

Clearly, (a) and (b) of claim hold.

We now consider (c). Consider any $i = i_0$, and $k = k_0 \leq i_0$. If $\varphi_{p(i_0)}$ is total, then (c) clearly holds. So suppose $\varphi_{p(i_0)}$ is in $\mathrm{SEG}_{0,1}$. Suppose by way of contradiction that $\Theta_{k_0}(\varphi_{p(i_0)}) \sim \Theta_{k_0}(\mathrm{Zero})$, and there exists an extension $f \in \mathcal{R}_{0,1}$ of $\varphi_{p(i_0)}$ such that $\Theta_{k_0}(f) \not\sim \Theta_{k_0}(\mathrm{Zero})$. Then there exists a finite τ, $\varphi_{p(0)} \subseteq \tau \subseteq f$ and an n such that $\Theta_{k_0}(\tau) \not\sim \Theta_{k_0}(\mathrm{Zero}[n])$. Thus for some stage $s > \max(|\tau|, n, i_0)$ and $s \bmod 3 = 0$, in step 2 of stage s, $\varphi_{p(i_0)}$ would be extended so as to make $\Theta_{k_0}(\varphi_{p(i_0)}) \not\sim \Theta_{k_0}(\mathrm{Zero})$. A contradiction.

Thus (c) is satisfied.

We now consider (d). Consider $i = i_0, j = j_0$, and $k = k_0 \leq \min(i_0, j_0)$. Without loss of generality assume $i_0 < j_0$. Suppose $\lim_{s\to\infty} I(i_0, k_0, s)$ and $\lim_{s\to\infty} I(j_0, k_0, s)$ are both true (otherwise (d) trivially holds). Suppose by way of contradiction that $\Theta_{k_0}(\varphi_{p(i_0)}) \sim \Theta_{k_0}(\varphi_{p(j_0)})$ and there exist total $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i_0)} \subseteq f$, $\varphi_{p(j_0)} \subseteq g$, and $\Theta_{k_0}(f) \not\sim \Theta_{k_0}(g)$. Let n be such that $\Theta_{k_0}(f[n]) \not\sim \Theta_{k_0}(g[n])$. Let s be large enough such that $I(i_0, k, s)$ and $I(j_0, k, s)$ are both true, $[\varphi^s_{p(i_0)} = \varphi_{p(i_0)}$ or $\varphi^s_{p(i_0)} \supseteq \varphi_{p(i_0)}[n]]$ and $[\varphi^s_{p(j_0)} = \varphi_{p(j_0)}$ or $\varphi^s_{p(j_0)} \supseteq \varphi_{p(j_0)}[n]]$. Thus, for any $s' > s$ such that $s' \bmod 3 = 1$, (i_0, j_0, k_0) would be possible candidate in step 4, and for large enough s' it will be the smallest triple satisfying 4.1–4.6. Thus, it would be ensured that $\Theta_{k_0}(\varphi_{p(i_0)}) \not\sim \Theta_{k_0}(\varphi_{p(j_0)})$.

Thus (d) holds.

For (e), first note that $\varphi_{p(i)}$ may be modified in stages $s \bmod 3 = 0$ only finitely often, since there are only finitely many $k \leq i$. Similarly, for each $k \leq i$, $\varphi_{p(i)}$ may be modified in stages $s \bmod 3 = 1$ only finitely often, since if $I(i, k, s)$, then there exists a finite n such that $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$, and thus, only $j < n$, can satisfy (4.3). It follows that for all i, for all but finitely many stages, if $s \bmod 3 \neq 2$, then $\varphi^s_{p(i)} = \varphi^{s+1}_{p(i)}$.

We now show that $\mathcal{C} \notin \mathcal{R}\mathbf{Cons}$. Suppose $\mathbf{M}_i$ is total. Let t be such that for any stage $s > t$, $s \bmod 3 \neq 2$, $\varphi^s_{p(i)} = \varphi^{s+1}_{p(i)}$. Note that there exists such a stage t. Now note that in step 6, If clause succeeds infinitely often. If 7.1 is executed infinitely often, then $\varphi_{p(i)}$ is total but $\mathbf{M}_i$ makes infinitely many mind changes on $\varphi_{p(i)}$. On the other hand, if 7.2 is executed in any stage $s > t$, $s \bmod 3 = 2$, then step 7.2 would be executed in every stage $s' > s$, $s' \bmod 3 = 2$ (since

$\varphi_{p(i)} = \varphi^s_{p(i)}$ in this case), $\varphi_{p(i)}$ is finite, and $\mathbf{M}_i$ is not consistent on at least one of r-extension of $\varphi_{p(i)}$, for $r \in \{0,1\}$ (since $\mathbf{M}_i(\varphi^s_{p(i)} \cdot 0) = \mathbf{M}_i(\varphi^s_{p(i)} \cdot 1)$). Thus again $\mathbf{M}_i$ does not $\mathcal{R}$**Cons**-identify $\mathcal{C}$. Thus $\mathcal{C} \notin \mathcal{R}\mathbf{Cons}$.

Let $H(i,s) = 0$, if in stage $3s+2$ above construction executes step 7.2; $H(i,s) = 1$ otherwise. It is easy to verify that H satisfies the requirements in (e). ∎

Corollary 47 *Suppose p is as in Lemma 46. Suppose Θ_k is general recursive operator. Then given any i and n, exactly one of* (a) *and* (b) *below holds, and one can effectively determine which of* (a) *and* (b) *holds.*

(a) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$.

(b) *For all $f \in \mathcal{R}_{0,1}$ extending $\varphi_{p(i)}$, $\Theta_k(f) \sim \Theta_k(\mathrm{Zero}[n])$.*

Proof. By Lemma 46(c), we have:

(i) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero})$ or

(ii) For all $f \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$, $[\Theta_k(f) \sim \Theta_k(\mathrm{Zero})]$.

If (ii) holds, then clearly, (b) of corollary holds. So suppose (i) holds. It follows that either $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$ or $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\mathrm{Zero}[n])$. If $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$ then (a) holds, and if $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\mathrm{Zero}[n])$ then (b) holds.

We now show how to effectively determine which of (a) and (b) holds.

Note that if (a) holds, then there exists a t such that $\Theta_k(\varphi_{p(i),t}) \not\sim \Theta_k(\mathrm{Zero}[n])$. If (b) holds, then either

(iii) $\Theta_k(\varphi_{p(i)}) \supseteq \Theta_k(\mathrm{Zero}[n])$ or

(iv) $\varphi_{p(i)}$ is finite and for all $\tau \in \mathrm{SEG}_{0,1}$ such that $\varphi_{p(i)} \subseteq \tau$, $\Theta_k(\tau) \sim \Theta_k(\mathrm{Zero}[n])$.

Thus, if (b) holds, then there exists a t such that

(iii') $\Theta_k(\varphi_{p(i),t}) \supseteq \Theta_k(\mathrm{Zero}[n])$ or

(iv') For all $\tau \in \mathrm{SEG}_{0,1}$ such that $|\tau| = t$ and $\varphi_{p(i),t} \sim \tau$, $\Theta_k(\tau) \supseteq \Theta_k(\mathrm{Zero}[n])$

(where (iv') follows from (iv) using Proposition 40).

Thus if (b) holds we can effectively find a witness for it.

It follows that one can determine effectively which of (a) and (b) holds. ∎

Corollary 48 *Suppose p is as in Lemma 46. Suppose Θ_k is general recursive operator. Then given any i, j and n, exactly one of* (a), (b) *and* (c) *holds, and one can effectively determine which of* (a), (b) *and* (c) *holds.*

(a) [*For all $f \in \mathcal{R}_{0,1}$, such that $\varphi_{p(i)} \subseteq f$, $\Theta_k(f) \sim \Theta_k(\mathrm{Zero}[n])$*] *or* [*For all $g \in \mathcal{R}_{0,1}$, such that $\varphi_{p(j)} \subseteq g$, $\Theta_k(g) \sim \Theta_k(\mathrm{Zero}[n])$*]*;*

(b) [$\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$ *and* $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\mathrm{Zero}[n])$] *and there exists an $x < n$, such that $\Theta_k(\varphi_{p(i)})(x) \neq \Theta_k(\varphi_{p(j)})(x)$,*

(c) [$\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$ *and* $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\mathrm{Zero}[n])$] *and for all $f \in \mathcal{R}_{0,1}$ extending $\varphi_{p(i)}$, for all $g \in \mathcal{R}_{0,1}$ extending $\varphi_{p(j)}$, $\Theta_k(f)[n] = \Theta_k(g)[n]$.*

Proof. Note that using Corollary 47 one can effectively determine whether (a) holds or $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\mathrm{Zero}[n])$ and $\Theta_k(\varphi_{p(j)}) \not\sim \Theta_k(\mathrm{Zero}[n])$. We now show that if (a) does not hold then one of (b) or (c) holds, and one can effectively determine which one.

By Lemma 46(d), if (a) does not hold then we must have

(iii) $\Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\varphi_{p(j)})$ or

(iv) For all $f, g \in \mathcal{R}_{0,1}$ such that $\varphi_{p(i)} \subseteq f$ and $\varphi_{p(j)} \subseteq g$, $[\Theta_k(f) \sim \Theta_k(g)]$.

If (iii) holds, then clearly, either (b) of corollary holds or both $\Theta_k(\varphi_{p(i)})$ and $\Theta_k(\varphi_{p(j)})$ have domain a superset of $\{0, 1, \ldots, n-1\}$ and $\Theta_k(\varphi_{p(i)})[n] = \Theta_k(\varphi_{p(j)})[n]$, and thus (c) is satisfied. Moreover witness for both of above cases can be effectively found (since in first case there exists a t and $x < n$ such that $\Theta_k(\varphi_{p(i),t})(x) \neq \Theta_k(\varphi_{p(j),t})(x)$, and in second case there exists a t such that domains of $\Theta_k(\varphi_{p(i),t})$ and $\Theta_k(\varphi_{p(j),t})$, are supersets of $\{0, 1, \ldots, n-1\}$ and $\Theta_k(\varphi_{p(i),t})[n] = \Theta_k(\varphi_{p(j),t})[n]$).

So suppose (iv) holds. Then clearly, (c) of corollary holds. Thus, by Proposition 40, there exists a t such that

(iv') For all τ such that $|\tau| = t$ and $\varphi_{p(i),t} \sim \tau$, for all γ such that $|\gamma| = t$ and $\varphi_{p(j),t} \sim \gamma$,

$\Theta_k(\tau) \sim \Theta_k(\gamma)$, and domain of both $\Theta_k(\tau)$ and $\Theta_k(\gamma)$ is a superset of $\{0, \ldots, n-1\}$.

Thus, we have a witness for (c) holding.

It follows that one can determine effectively which of (a), (b) or (c) holds. ∎

Theorem 49 Robust$\mathcal{R}$Conf $\not\subseteq$ $\mathcal{R}$Cons. (*Moreover the non-inclusion can* *be witnessed using a class from $\mathcal{R}_{0,1}$.)*

Proof. Let p, H be as in Lemma 46.

Let $\mathcal{C} = \{\varphi_{p(i)} \in \mathcal{R} \mid \mathbf{M}_i \text{ is total }\} \cup \{r\text{-extension of } \varphi_{p(i)} \mid \mathbf{M}_i \text{ is total and } \varphi_{p(i)} \text{ is not total and } r \in \{0,1\}\}$.

Then by part (e) of Lemma 46, $\mathcal{C} \notin \mathcal{R}\mathbf{Cons}$.

We now show that $\mathcal{C} \in \mathbf{Robust}\mathcal{R}\mathbf{Conf}$. To show this, suppose Θ_k is a general recursive operator. We need to show that $\Theta_k(\mathcal{C}) \in \mathcal{R}\mathbf{Conf}$.

Let $\mathcal{S} = \{\varphi_{p(i)} \in \mathcal{C} \mid i \leq k\} \cup \{r\text{-extension of } \varphi_{p(i)} \mid r \in \{0,1\}, i \leq k, \text{ and } r\text{-extension of } \varphi_{p(i)} \in \mathcal{C}\}$.

Define $\mathbf{M}$ as follows:

$\mathbf{M}(g[m])$
1. If $g[m] \sim \Theta_k(\text{Zero})$, then output a standard program for $\Theta(\text{Zero})$.
2. Else If $g[m] \sim \Theta_k(f)$, for some $f \in \mathcal{S}$, then output a standard program for one such $\Theta_k(f)$.
3. Else let n be such that $\Theta_k(\text{Zero}[n]) \not\sim g[m]$.
   (* Note that if $g \in \Theta_k(\mathcal{C})$, and we are at this point of the construction, then g must be $\Theta_k(\varphi_{p(i)})$, or $\Theta_k(r\text{-extension of } \varphi_{p(i)})$, for some $r \in \{0,1\}$, for some i such that $k \leq i < n$. *)
4. Using Corollary 47 determine $A = \{i \mid k \leq i < n \wedge \Theta_k(\varphi_{p(i)}) \not\sim \Theta_k(\text{Zero}[n])\}$.
   (* Note that by Corollary 42, for $k \leq i < n$, and $i \notin A$ — either $\varphi_{p(i)}$ is not total or $\Theta_k(\varphi_{p(i)}) \sim \Theta_k(\text{Zero}[n]) \not\sim g[m]$ — thus, $g \neq \Theta_k(\varphi_{p(i)})$. *)
5. Let B be a subset of A such that:
   for $i \in A - B$, there exists an $x < m$ such that $\Theta_k(\varphi_{p(i)})(x) \neq g(x)$;
   for distinct $i, j \in B$, for all $f, h \in \mathcal{R}_{0,1}$ such that $f \supseteq \varphi_{p(i)}$ and $h \supseteq \varphi_{p(j)}$, $\Theta_k(f)[m] = \Theta_k(h)[m]$.
   (* Note that, by Corollary 48, there exists such a B, and one such B can be effectively found. *)
6 For each $i \in B$, such that $H(i,m) = 0$, let h_i denote
   0-extension of $\varphi_{p(i),m}$ if $\Theta_k(0\text{-extension of } \varphi_{p(i),m}) \supseteq g[m]$;
   1-extension of $\varphi_{p(i),m}$ if $\Theta_k(1\text{-extension of } \varphi_{p(i),m}) \supseteq g[m]$, and $\Theta_k(0\text{-extension of } \varphi_{p(i),m}) \not\supseteq g[m]$;
   Λ if $\Theta_k(1\text{-extension of } \varphi_{p(i),m}) \not\supseteq g[m]$, and $\Theta_k(0\text{-extension of } \varphi_{p(i),m}) \not\supseteq g[m]$.
7. Output (standard) program for $\text{Union}(\{\Theta_k(\varphi_{p(i)}) \mid i \in B\} \cup \{\Theta_k(h_i) \mid i \in B, H(i,m) = 0\})$.

End

We claim that above $\mathbf{M}$ $\mathcal{R}$**Conf**-identifies $\Theta_k(\mathcal{C})$.

Clearly, $\mathbf{M}$ is defined on each initial segment. Suppose $g \in \Theta_k(\mathcal{C})$. To see that $\mathbf{M}$ **Conf**-identifies g we argue as follows.

If $g = \Theta_k(\text{Zero})$ or if $g = \Theta_k(\varphi_{p(i)})$, for some $i \leq k$, or if $g = \Theta_k(r\text{-extension of } \varphi_{p(i)})$, for $r \in \{0, 1\}$, for some $i \leq k$, then clearly by steps 1 and 2, $\mathbf{M}$ is conforming on g and **Ex**-identifies g.

If $g \neq \Theta_k(\text{Zero})$, then let n be least value such that $g \not\sim \Theta_k(\text{Zero}[n])$. Thus, g must be one of $\Theta_k(\varphi_{p(i)})$, for $i < n$, or g must be $\Theta_k(r\text{-extension of } \varphi_{p(i)})$, $r \in \{0, 1\}$, for some $i < n$. We have already handled the case $i \leq k$, so assume $k < i < n$ for the following.

We first show that $\mathbf{M}$ is conforming on these g. Note that steps 1 and 2 are clearly consistent with the input. So we only consider m such that $\mathbf{M}(g[m])$ in above construction reaches step 3. For these m, by construction of A, if $i \notin A$, then for all total f extending $\varphi_{p(i)}$, $\Theta_k(f)$ is conforming with $\Theta_k(\text{Zero}[n])$ and thus not equal to g. Similarly, if $i \in A - B$, then for all total f extending $\varphi_{p(i)}$, $\Theta_k(f)$ is nonconforming with g.

Now, since $g \in \Theta_k(\mathcal{C})$, there exists an $i \in B$ such that $\Theta_k(\varphi_{p(i)}) = g$, or $\lim_{t\to\infty} H(i,t) = 0$ and $\varphi_{p(i)}$ is not total and $g = \Theta_k(r\text{-extension of } \varphi_{p(i)})$, for some $r \in \{0, 1\}$. Fix one such i.

Now for $j \in B$, $j \neq i$, for any total extension f of $\varphi_{p(i)}$ and any total extension h of $\varphi_{p(j)}$, $\Theta_k(f)$ and $\Theta_k(h)$ do not differ on $\{0, 1, \ldots, m-1\}$; thus for all $j \in B$, $\Theta_k(\varphi_{p(j)})$ is conforming with $g[m]$. Also, by definition of h_j in step 6, we have that $g[m]$ is conforming with $\Theta_k(h_j)$, for all $j \in B$ such that $H(j, m) = 0$. Thus, $\mathbf{M}$ is conforming on g.

To show that $\mathbf{M}$ **Ex**-identifies g, first note that B is monotonically non-increasing (with respect to $g[m]$). Also, for all $j \in B$, $\lim_{t\to\infty} H(j,t)\!\downarrow$ and if $\lim_{t\to\infty} H(j,t) = 0$, then h_j stabilizes as m goes to ∞.

Thus, if $g = \Theta_k(\varphi_{p(i)})$, then clearly, $\{\Theta_k(\varphi_{p(j)}) \mid j \in B\}$ contains g. On the other hand, if $\varphi_{p(i)}$ is not total, and g is $\Theta_k(r\text{-extension of } \varphi_{p(i)})$, for some $r \in \{0, 1\}$, then $g = \Theta_k(h_i)$, for the limiting value of h_i (with respect to m).

It follows that, for large enough m, in the definition of $\mathbf{M}(g[m])$,

$$g \subseteq \text{Union}(\{\Theta_k(\varphi_{p(i)}) \mid i \in B\} \cup \{\Theta_k(h_i) \mid i \in B, H(i,m) = 0\}).$$

Along with conformity of $\mathbf{M}$, above implies that $\mathbf{M}$ **Ex**-identifies g. ∎

Corollary 50 $\mathbf{Robust}\mathcal{R}\mathbf{Conf} \supset \mathbf{NUM}$.

Proof. The inclusion follows from Corollary 16(b) and Theorem 10(a). The proper inclusion now follows from Theorem 49 and Theorem 10(a). ∎

Proposition 51 $\mathbf{UniformRobustConfident} \not\subseteq \mathcal{R}\mathbf{Conf}$.

Proof. $\mathcal{C}$ used in the proof of Theorem 35 is also in **UniformRobustConfident**. ∎

Corollary 52 $\mathbf{UniformRobustConfident} \not\subseteq \mathbf{NUM}$.

Proof. This follows immediately from Proposition 51 and Theorem 10(a). ∎

Note that by Theorem 10(h), $\mathbf{NUM} \not\subseteq \mathbf{Confident}$. Thus using Corollary 52, **NUM** and **Confident** are incomparable. The following proposition strengthens Corollary 34.

Proposition 53 $\mathbf{UniformRobustEx} \not\subseteq \mathbf{Confident} \cup \mathbf{NUM}$.

Proof. Let F, $\mathcal{S}$ be as in the proof of Theorem 32. Thus $\mathcal{S} \in \mathbf{UniformRobustEx}$ and $\mathcal{S} \notin \mathbf{NUM}$, by proof of Corollary 34.

Now we show that $\mathcal{S} \notin \mathbf{Confident}$. Suppose that a total-recursive and confident learner $\mathbf{M}$ **Ex**-identifies $\mathcal{S}$. This learner $\mathbf{M}$ makes on F only finitely many mind changes, in particular there exists an n such that, for all $n' > n$, $\mathbf{M}(F[n']) = \mathbf{M}(F[n])$. Now consider the following set B of strings:

$$B = \{\tau \mid \tau \text{ extends } F[n] \text{ and } \mathbf{M}(\tau) \neq \mathbf{M}(F[n])\}.$$

Since for all $\sigma \in B$, $\sigma \not\subseteq F$, by definition of 1-generic function, there exists an $x > n$ such that for all $\tau \supseteq F[x]$, $\mathbf{M}(\tau) = \mathbf{M}(F[x])$. Thus, $\mathbf{M}$ can learn at most one function extending $F[x]$. However, by definition of $\mathcal{S}$, there are infinitely many functions extending $F[x]$ in $\mathcal{S}$. Thus, $\mathbf{M}$ cannot witness that $\mathcal{S} \in \mathbf{Confident}$. ∎

Zeugmann [Zeu86] proved already the following theorem.

Theorem 54 [Zeu86] $\mathbf{RobustReliable} = \mathbf{NUM}$.

Corollary 55 $\mathbf{Robust}\mathcal{T}\mathbf{Cons} = \mathbf{NUM}$.

Proof. This follows immediately from Theorem 54 and Corollary 16(a). ∎

5 Robust Learning versus Uniformly Robust Learning

While in robust learning any transformed class is required to be learnable in the mere sense that there *exists* a learning machine for it, uniformly learning requires to get such a machine for the transformed class *effectively at hand*. For a number of learning types, we now show that uniformly robust learning is indeed stronger than robust learning.

Definition 56 A function F is *i-generic* iff for all sets $A \subseteq \mathrm{SEG}$ in Σ_i, there exists an x such that either

(a) $F[x] \in A$, or

(b) for all $\sigma \in A$, $F[x] \not\subseteq \sigma$.

Proposition 57 *There exist a recursive binary tree T and a limiting recursive function mapping $\tau \in SEG_{0,1}$ to $m_\tau \in SEG_{0,1}$ such that following four properties are satisfied.*

(a) *For all $\sigma \in SEG_{0,1}$, $m_\sigma \in T$.*

(b) *For all $\sigma, \tau \in SEG_{0,1}$, $\sigma \subset \tau$ iff $m_\sigma \subset m_\tau$.*

(c) *Only infinite branches of T are m_g, where $g \in \mathcal{T}_{0,1}$ and $m_g = \bigcup_{n\in N} m_{g[n]}$.*

(d) *For all $\tau \in SEG_{0,1}$, for all $i, j \leq |\tau|$, Either*

  (d.1) $\Theta_i(m_\tau) \not\sim \varphi_j$*, Or*

  (d.2) *for all $\sigma \supseteq \tau$, $\sigma \in SEG_{0,1}$, $\Theta_i(m_\sigma) \sim \varphi_j$.*

Proof. We first define m_τ, for $\tau \in \mathrm{SEG}_{0,1}$ below in stages. m_τ^s denotes the value of m_τ just before the start of stage s. It will be the case that $m_\tau^s \in \mathrm{SEG}_{0,1}$, and for all $\sigma, \gamma \in \mathrm{SEG}_{0,1}$, for all s, $\sigma \subset \gamma$ iff $m_\sigma^s \subset m_\gamma^s$. Also, $m_\tau = \lim_{s\to\infty} m_\tau^s\downarrow$. For $n \leq |\sigma|$, let $\sigma[m, n]$ denote $\sigma(m)\sigma(m+1)\ldots\sigma(n-1)$.

Initially, let $m_\tau^0 = \tau$. Go to stage 0.

Stage s

1. Search for $\tau, \sigma \in \mathrm{SEG}_{0,1}$, $i, j \in N$ such that
   $|\tau|, |\sigma| \leq s$,
   $i, j \leq |\tau|$,
   $\tau \subseteq \sigma$,
   $\Theta_i(m_\tau^s) \sim \varphi_{j,s}$.
   $\Theta_i(m_\sigma^s) \not\sim \varphi_{j,s}$.
2. If there exist such τ, σ, i, j, then for least such τ, and a corresponding σ let

$m_\tau^{s+1} = m_\sigma^s$,
for all $\gamma \supseteq \tau$, $m_\gamma^{s+1} = m^s_{\sigma\gamma[|\tau|,|\gamma|]}$.

3. For all γ, such that m_γ^{s+1} is not defined in step 2, let $m_\gamma^{s+1} = m_\gamma^s$.
4. Go to stage $s+1$.

End stage s

Since each τ can be chosen in step 2 only finitely often — once for each pair $i, j \leq |\tau|$ — after each m_γ, $\gamma \subset \tau$ are stabilized, it is easy to show by induction on $|\tau|$ that m_τ^{s+1} can be defined via step 2 above only finitely often. Thus, $m_\tau = \lim_{s\to\infty} m_\tau^s\downarrow$.

Let $T = \{\sigma \mid (\exists\tau)(\exists t \geq |\sigma|)[\sigma \subseteq m_\tau^t]\}$.

It is easy to verify that (a) and (b) are satisfied. To see (c), suppose $h \in \mathcal{T}_{0,1}$ is not of form m_g for any $g \in \mathcal{T}_{0,1}$. Let $X_h = \{\tau \in \mathrm{SEG}_{0,1} \mid m_\tau \subseteq h\}$. Let $\eta = \bigcup_{\tau \in X_h} \tau$. By part (b) if X_h is infinite, then $\eta \in \mathcal{T}_{0,1}$, and $h = m_\eta$. Thus, X_h is finite. It follows that, for all $\tau \in \mathrm{SEG}_{0,1}$ such that $|\tau| = |\eta|+1$, $m_\tau \not\subseteq h$. Let $w = \max(\{|m_\tau| \mid \tau \in \mathrm{SEG}_{0,1} \ \wedge\ |\tau| = |\eta|+1\})$. Then, it follows that $h[w]$ is not a prefix of any extension of m_τ, for any τ of length $|\eta|+1$.

Let s be such that, for all $\tau \in \mathrm{SEG}_{0,1}$ such that $|\tau| \leq |\eta|+1$, for all $t \geq s$, $m_\tau = m_\tau^t$. Now, it follows that $h[w] \not\subseteq m_\tau^t$, for any $t \geq s$, and any $\tau \in \mathrm{SEG}_{0,1}$. Thus, it follows that $h[\max(w,s)] \not\subseteq m_\tau^t$, for any $t \geq \max(w,s)$, and any $\tau \in \mathrm{SEG}_{0,1}$. Thus, $h[\max(w,s)] \notin T$. This proves (c).

To see (d), suppose $i, j \leq |\tau|$, $\sigma \supseteq \tau$, and $\Theta_i(m_\tau) \sim \varphi_j$, but $\Theta_i(m_\sigma) \not\sim \varphi_j$. Then, let s be such that $s \geq \max(i, j, |\tau|, |\sigma|)$ and $\Theta_i(m_\tau) \sim \varphi_{j,s}$, $\Theta_i(m_\sigma) \not\sim \varphi_{j,s}$, and for all $t \geq s$, $m_\tau^t = m_\tau$, and $m_\sigma^t = m_\sigma$. But then for all $t \geq s$, τ, σ, i, j will satisfy the conditions in step 1. Thus, for large enough $t > s$, τ will be the least candidate picked in stage t, step 2. Hence, $m_\tau^t \neq m_\tau^{t+1}$, a contradiction. Consequently, (d) holds. ∎

Theorem 58 UniformRobustConfident $\subset$ RobustConfident.

Proof. Clearly, **UniformRobustConfident** $\subseteq$ **RobustConfident**. We thus only need to show that **RobustConfident** $-$ **UniformRobustConfident** $\neq \emptyset$.

Fix $T, m_\tau, \tau \in \mathrm{SEG}_{0,1}$ and $m_g, g \in \mathcal{T}_{0,1}$, as in Proposition 57.

Let F be a $\{0,1\}$-valued 3-generic function.

Let $f' = m_F$.

Let $\mathcal{S} = \{f'[x] \cdot 0^\infty \mid x \in N\}$.

For rest of proof let σ, τ range over $\mathrm{SEG}_{0,1}$.

Claim 59 *For all e such that Θ_e is general recursive, there exist x_e, j_e such that $\varphi_{j_e} \in \mathcal{R}$, and*

(A) $(\forall \tau \supseteq F[x_e])[\Theta_e(m_\tau) \subseteq \varphi_{j_e}]$, *or*

(B) $(\forall \varphi_j \in \mathcal{R})(\forall \tau \supseteq F[x_e] \mid |\tau| \geq \max(e, j))[\Theta_e(m_\tau) \not\sim \varphi_j]$.

Proof. Suppose Θ_e is general recursive. Let $R_e = \{\tau \mid (\exists j)[\varphi_j \in \mathcal{R} \wedge (\forall \sigma \supseteq \tau)[\Theta_e(m_\sigma) \sim \varphi_j]]\}$.

Note that $R_e \in \Sigma_3$. Thus, since F is 3-generic, there exists an x_e such that either

(a) $F[x_e] \in R_e$, or

(b) for all $\tau \in R_e$, $F[x_e] \not\subseteq \tau$.

In other words, either

(a') there exists a j such that $\varphi_j \in \mathcal{R}$, and, for all $\sigma \supseteq F[x_e]$, $\Theta_e(m_\sigma) \subseteq \varphi_j$, or

(b') for all $\varphi_j \in \mathcal{R}$, for all $\tau \supseteq F[x_e]$, there exists a $\sigma \supseteq \tau$ such that $[\Theta_e(m_\sigma) \not\sim \varphi_j]$.

By Proposition 57, (b') above is equivalent to

(b") for all $\varphi_j \in \mathcal{R}$, for all $\tau \supseteq F[x_e]$, such that $|\tau| \geq \max(e, j)$, $[\Theta_e(m_\tau) \not\sim \varphi_j]$.

Now (a') is same as (A) and (b") is same as (B). $\Box$

For Θ_e being general recursive, fix x_e, j_e as in Claim 59.

Let $\mathcal{F}_e = \{\sigma 0^\infty \mid \sigma \subseteq m_{F[x_e]}\}$.

Let $\mathcal{G}_e = \{\sigma 0^\infty \mid \sigma \supseteq m_{F[x_e]}\}$.

Note that $\mathcal{S} \subseteq \mathcal{F}_e \cup \mathcal{G}_e$.

Claim 60 *Suppose Θ_e is general recursive. Then, for all $f \in \Theta_e(\mathcal{S}) - [\Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}]$, there exists an n such that, for all $\sigma \in T$, such that $\sigma \supseteq m_{F[x_e]}$ and $|\sigma| = n$, $f \not\sim \Theta_e(\sigma)$.*

Proof. Suppose $f \in \Theta_e(\mathcal{S}) - [\Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}]$.

We first claim that,

(P1) for some $m > x_e$, for all $\tau \supseteq F[x_e]$ such that $|\tau| \geq m$, $\Theta_e(m_\tau) \not\sim f$.

To see this, first note that by Claim 59,

(A) For all $\tau \supseteq F[x_e]$, $\Theta_e(m_\tau) \subseteq \varphi_{j_e}$, or

(B) $(\forall \varphi_j \in \mathcal{R})(\forall \tau \supseteq F[x_e] \mid |\tau| \geq \max(e, j))[\Theta_e(m_\tau) \not\sim \varphi_j]$.

In case (B), (P1) immediately follows by taking $m = \max(\text{MinProg}(f), e)$. In case (A), since $f \neq \varphi_{j_e}$, by König's Lemma and general recursiveness of Θ_e, (P1) follows.

Now, by Proposition 57(c) and König's Lemma, there exist only finitely many $\sigma \in T$ such that $\sigma \supseteq m_{F[x_e]}$, but σ does not extend any m_τ, with $\tau \supseteq F[x_e]$, $|\tau| \geq m$. Thus, there exists an $n > m$, such that for all $\sigma \in T$ with $\sigma \supseteq m_{F[x_e]}$ and $|\sigma| = n$, $f \not\sim \Theta_e(\sigma)$. □

Claim 61 *For all e, such that Θ_e is general recursive,* $\Theta_e(\mathcal{S}) \in$ **Confident**.

PROOF. Let h be as in Corollary 22. Consider the following $\mathbf{M}$.

$\mathbf{M}(g[n])$
1. If for some $f \in \Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}$, $f[n] = g[n]$, then output a standard program for f.
2. Else If there exists a $\sigma \in T$, such that $|\sigma| = n$, $\sigma \supseteq m_{F[x_e]}$, $\Theta_e(\sigma) \sim g[n]$, then output $\mathbf{M}(g[n-1])$ (if $n = 0$, then $\mathbf{M}(g[n]) = 0$).
3. If for all $\sigma \in T$, such that $|\sigma| = n$, and $\sigma \supseteq m_{F[x_e]}$, $[\Theta_e(\sigma) \not\sim g[n]]$, then
   Let p be a bound on the programs for $\{\sigma 0^\infty \mid \sigma \in \text{SEG}_{0,1}, |\sigma| < m\}$, where $m > x_e$ is the least number such that for all $\sigma \in T$ with $|\sigma| = m$ and $\sigma \supseteq m_{F[x_e]}$, $\Theta_e(\sigma) \not\sim g[n]$.
   Output $\mathbf{M}_{h(p,e)}(g[n])$.

End

We claim that $\mathbf{M}$ **Confident**-identifies $\Theta_e(\mathcal{S})$. To see this, first note that $\mathbf{M}$ **Ex**-identifies each $g \in \Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}$. If $g \notin \Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}$, then we consider two cases.

(Note that, if $\mathbf{M}(g[n])$ executes step 3, then $\mathbf{M}(g[n+1])$ also executes step 3.)

Case 1: For all n, $\mathbf{M}(g[n])$ never reaches step 3.

In this case trivially $\mathbf{M}$ is confident on g. Moreover, by Claim 60, $g \notin \Theta_e(\mathcal{S}) - [\Theta_e(\mathcal{F}_e) \cup \{\varphi_{j_e}\}]$.

Case 2: For all but finitely many n, $\mathbf{M}(g[n])$ executes step 3.

Then, for all n such that $\mathbf{M}(g[n])$ reaches step 3, m as computed in step 3 of the algorithm for $\mathbf{M}$, would be the least number such that, for all $\sigma \in T$ with $|\sigma| = m$ and $\sigma \supseteq m_{F[x_e]}$, $\Theta_e(\sigma) \not\sim g$. It follows that values of m and p as computed by $\mathbf{M}$ converge. Thus $\mathbf{M}$ is confident on g. Moreover, by definition of $\mathcal{S}$, if $g \in \Theta_e(\mathcal{S})$, then g must be from $\{\Theta_e(\sigma 0^\infty) \mid |\sigma| < m\}$. Thus, $\mathbf{M}_{h(p,e)}$ and hence $\mathbf{M}$ $\mathbf{Ex}$-identifies g.

From above cases it follows that $\mathbf{M}$ **Confident**-identifies $\Theta_e(\mathcal{S})$. $\Box$

Claim 62 $\mathcal{C} \notin$ **UniformRobustConfident**.

Proof. Suppose by way of contradiction that there exists a recursive g such that, for all e, if Θ_e is general recursive, then $\mathbf{M}_{g(e)}$ **Confident**-identifies $\Theta_e(\mathcal{S})$. Without loss of generality assume that $\mathbf{M}_{g(e)}$ is total for all e.

Then, by implicit use of Kleene recursion theorem [Rog67], there exists an e such that Θ_e may be defined as follows. For ease of presentation, $\Theta_e(f[m])$ may be infinite for some f, m (and thus it does not satisfy the invariant (d) we assumed (see discussion around Proposition 13) on the enumeration $\Theta_0, \Theta_1, \ldots$; this can be easily handled by appropriately slowing down Θ_e).

$\Theta_e(\Lambda) = \Lambda$.

$$\Theta_e(f[n{+}1]) = \begin{cases} \Theta_e(f[n]), & \text{if } \Theta_e(f[n]) \text{ is infinite;} \\ \Theta_e(f[n]) \cdot f(n) \cdot 0^\infty, & \text{if for all } m, \\ & \Theta_e(f[n]) \cdot f(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]) \cdot f(n) \cdot 0^m), m}; \\ \Theta_e(f[n]) \cdot f(n) \cdot 0^m, & \text{if } m \text{ is the least number such that} \\ & \Theta_e(f[n]) \cdot f(n) \subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f[n]) \cdot f(n) \cdot 0^m), m}. \end{cases}$$

It is easy to verify that Θ_e is general recursive. Recall that $f' = m_F$. We now consider two cases.

Case 1: For some n, $\Theta_e(f'[n+1])$ is infinite.

Let n be maximum such that $\Theta_e(f'[n])$ is finite. Thus, $\Theta_e(f'[n{+}1])$ is infinite, and $\Theta_e(f') = \Theta_e(f'[n]) \cdot f'(n) \cdot 0^\infty = \Theta_e(f'[n{+}1] \cdot 0^\infty) \in \Theta_e(\mathcal{S})$. However, by definition of Θ_e, $\Theta_e(f'[n]) \cdot f'(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f'[n]) \cdot f'(n) \cdot 0^m), m}$, for all m. Thus, $\mathbf{M}_{g(e)}(\Theta_e(f'[n{+}1] \cdot 0^\infty))$ diverges, or $\Theta_e(f'[n]) \cdot f'(n) \not\subseteq \varphi_{\mathbf{M}_{g(e)}(\Theta_e(f'[n] \cdot f'(n) \cdot 0^\infty))}$. It follows that $\mathbf{M}_{g(e)}$ does not $\mathbf{Ex}$-identify $\Theta_e(f'[n+1] \cdot 0^\infty) \in \Theta_e(\mathcal{S})$. A contradiction.

Case 2: For all n, $\Theta_e(f'[n])$ is finite.

In this case first note that for all $h \neq f'$, $\Theta_e(h) \neq \Theta_e(f')$. To see this, let y be the least number such that $h(y) \neq f'(y)$. Now it is easy to verify that $\Theta_e(f'[y] \cdot f'(y)) \subseteq \Theta_e(f')$ and $\Theta_e(f'[y] \cdot h(y)) \subseteq \Theta_e(h)$. Thus, $\Theta_e(f') \neq \Theta_e(h)$.

Thus, by Claim 59, it follows that $\Theta_e(f')$ is non-recursive (otherwise, for all m_h such that h extends $F[x_e]$, we would have $\Theta_e(f') = \Theta_e(m_h)$ – a contradiction to above proof that for all $m_h \neq f'$, $\Theta_e(f') \neq \Theta_e(m_h)$).

Now, for all n, $\Theta_e(f'[n]) \cdot f'(n) \subseteq \mathbf{M}_{g(e)}(\Theta_e(f'[n+1]))$. Thus, either $\mathbf{M}_{g(e)}$ diverges on $\Theta_e(f')$, or $\Theta_e(f')$ is recursive. Since it was shown above that $\Theta_e(f')$ is nonrecursive, it follows that $\mathbf{M}_{g(e)}$ diverges on $\Theta_e(f')$, a contradiction to $\mathbf{M}_{g(e)}$ being confident.

From above cases it follows that $\mathcal{S} \notin$ **UniformRobustConfident**. □

Theorem follows from Claims 61 and 62. ∎

For identification with mind changes, we assume $\mathbf{M}$ to be a mapping from SEG to $N \cup \{?\}$. This is to avoid biasing the number of mindchanges made by the machine [CS83].

Definition 63 [Gol67,BB75,CS83] Let $b \in N \cup \{*\}$. Let $f \in \mathcal{R}$.

(a) $\mathbf{M}$ $\mathbf{Ex}_b$*-identifies* f (written: $f \in \mathbf{Ex}_b(\mathbf{M})$) just in case $\mathbf{M}$ **Ex**-identifies f, and $\text{card}(\{n \mid ? \neq \mathbf{M}(f[n]) \neq \mathbf{M}(f[n+1])\}) \leq b$ (i.e., $\mathbf{M}$ makes no more than b mind changes on f).
(b) $\mathbf{M}$ $\mathbf{Ex}_b$*-identifies* $\mathcal{S}$ iff $\mathbf{M}$ $\mathbf{Ex}_b$-identifies each $f \in \mathcal{S}$.
(c) $\mathbf{Ex}_b = \{\mathcal{S} \subseteq \mathcal{R} \mid (\exists \mathbf{M})[\mathcal{S} \subseteq \mathbf{Ex}_b(\mathbf{M})]\}$.

In the following, we present characterizations for $\mathbf{RobustEx}_0$, see Proposition 64; for $\mathbf{UniformRobustEx}_n$ for any $n \in N$, see Corollary 68; and for $\bigcup_{n\in N} \mathbf{UniformRobustEx}_n$, see Corollary 70. As a consequence, we derive that all the types $\mathbf{Ex}_n$, $n \in N$, are uniformly robustly poor. Furthermore, for all the types $\mathbf{Ex}_n$, $n \in N$, uniformly robust learning turns out to be stronger than robust learning, see Corollary 71.

Proposition 64 [Zeu86,JSW01] *For any* $\mathcal{C} \subseteq \mathcal{R}$*,* $\mathcal{C} \in \mathbf{RobustEx}_0$ *iff* $\mathcal{C}$ *is finite.*

Thus, the type $\mathbf{Ex}_0$ of learning without any mind change is robustly poor, since every finite class is recursively enumerable.

Proposition 65 *For any* $n \in N$ *and* $\mathcal{C} \subseteq \mathcal{R}$ *with* $\text{card}(\mathcal{C}) < 2^{n+1}$*,* $\mathcal{C} \in \mathbf{UniformRobustEx}_n$*.*

Proof. Suppose $\mathcal{C} = \{f_0, f_1, \ldots, f_{m-1}\}$ with $m < 2^{n+1}$. Then, one can show $\mathcal{C} \in \mathbf{UniformRobustEx}_n$ as follows. There exists a recursive g such that $\mathbf{M}_{g(e)}$ may be defined as follows.

For any f, let $S(f[l]) = \{i < m \mid f[l] \sim \Theta_e(f_i[l])\}$.

For $S \subseteq \{i \mid i < m\}$, let Maj($S$) denote a program such that

$$\varphi_{\mathrm{Maj}(S)}(x) = \begin{cases} y, & \text{if } \mathrm{card}(\{i \in S \mid \Theta_e(f_i)(x) = y\}) > \frac{\mathrm{card}(S)}{2}; \\ \uparrow, & \text{otherwise.} \end{cases}$$

Now define $\mathbf{M}_{g(e)}(f[l])$ as follows.

$$\mathbf{M}_{g(e)}(f[l]) = \begin{cases} \mathbf{M}_{g(e)}(f[l-1]), & \text{if } \mathrm{card}(S(f[l])) = 0; \\ \mathrm{Maj}(S(f[s])), & \text{if } 2^k \leq \mathrm{card}(S(f[l])) < 2^{k+1}, \text{ and} \\ & s \text{ is the least number such that} \\ & 2^k \leq \mathrm{card}(S(f[s])) < 2^{k+1}. \end{cases}$$

For any f, for $i \leq n$, let s_i be minimal number, if any, such that $\mathrm{card}(S(f[s_i])) < 2^{i+1}$. (If for all l, $\mathrm{card}(S(f[l])) \geq 2^{i+1}$, then $s_i = \infty$). Let $s_{-1} = \infty$.

Note that for all l such that $s_i \leq l < s_{i-1}$, $\mathbf{M}_{g(e)}(f[l]) = \mathrm{Maj}(S[s_i])$. Thus, $\mathbf{M}_{g(e)}$ makes at most n mind changes. Furthermore, for $f \in \mathcal{C}$, let i be the minimum number such that $s_i \neq \infty$. Then, $\mathrm{card}(\{j \mid j \in S(f[s_i]) \wedge \Theta_e(f_j) \sim f\}) \geq 2^i$. Thus, $\mathrm{Maj}(S(f[s_i]))$ is a program for $\Theta_e(f)$. Thus, $\mathbf{M}_{g(e)}$ **Ex**-identifies f. The proposition follows. ∎

The following proposition is a modification of similar proposition in [JSW01].

Proposition 66 *For any $\mathbf{M}_e$ and $n \in N$, one can effectively define $p_i^{e,n}$, $i < 2^{n+1}$, such that $\mathbf{M}_e$ does not $\mathbf{Ex}_n$-identify $\{\varphi_{p_i^{e,n}} \mid i < 2^{n+1}\}$.*

Proof. We give below a description of total functions $F_i^{e,n}$ effectively in i, e, n. It will be the case that programs $p_i^{e,n}$ for $F_i^{e,n}$ can be obtained effectively from i, e, n.

Without loss of generality assume that $\mathbf{M}_e(\Lambda) =?$.

For all binary strings α of length at most n, we define H_α as follows.

$H_\Lambda = 0^j$, where $j = \min(\{x \mid \mathbf{M}_e(0^x) \neq ?\})$.

For a binary string β of length less than n and $b \in \{0, 1\}$, let $H_{\beta \cdot b}$ be defined as follows:

If H_β is infinite, then $H_{\beta \cdot b} = H_\beta$; otherwise, let $H_{\beta \cdot b} = H_\beta \cdot b \cdot 0^j$, where $j = \min(\{x \mid \mathbf{M}_e(H_\beta) \neq \mathbf{M}_e(H_\beta b 0^x)\})$.

It is easy to verify that $H_\beta \subseteq H_{\beta \cdot b}$, and if H_β is finite, then $\mathbf{M}_e$ on H_β has made at least $|\beta|$ mind changes. Now suppose $(n+1)$-bit binary representation of i is $b_0 b_1 \cdots b_n$. Then define

$$F_i^{e,n} = \begin{cases} H_{b_0 b_1 \ldots b_{n-1}}, & \text{if } H_{b_0, b_1 \ldots b_{n-1}} \text{ is infinite;} \\ H_{b_0 b_1 \ldots b_{n-1}} \cdot b_n \cdot 0^\infty, & \text{otherwise.} \end{cases}$$

It is easy to verify that programs $p_i^{e,n}$ for $F_i^{e,n}$ can be obtained effectively from i, e, n.

We claim that $\mathbf{M}_e$ cannot $\mathbf{Ex}_n$-identify $\{F_i^{e,n} \mid i < 2^{n+1}\}$. To see this, note that if H_Λ is infinite, then $\mathbf{M}_e$ does not identify any element of $\{F_i^{e,n} \mid i < 2^{n+1}\}$. Otherwise, let β be the longest string of length at most n such that H_β is finite. Then, consider the following cases:

Case 1: $|\beta| = n$.

In this case, $\mathbf{M}_e$ on H_β has already made n mind changes. Suppose i and i' respectively have binary representation of $\beta \cdot 0$ and $\beta \cdot 1$. Now, $\mathbf{M}_e$ can $\mathbf{Ex}_n$-identify at most one of $F_i^{e,n}$ and $F_{i'}^{e,n}$, since both start with H_β.

Case 2: $|\beta| < n$.

In this case, $\mathbf{M}_e(H_\beta) = \mathbf{M}_e(H_\beta \cdot 1 \cdot 0^\infty) = \mathbf{M}_e(H_\beta \cdot 0 \cdot 0^\infty)$. Suppose i and i' respectively have $(n+1)$-bit binary representation $\beta \cdot 0 \cdot 0^{n-|\beta|}$ and $\beta \cdot 1 \cdot 0^{n-|\beta|}$. Then, $F_i^{e,n} = H_\beta \cdot 0 \cdot 0^\infty$ and $F_{i'}^{e,n} = H_\beta \cdot 1 \cdot 0^\infty$, and thus, $\mathbf{M}_e$ does not $\mathbf{Ex}$-identify at least one of $F_i^{e,n}, F_{i'}^{e,n}$.

The proposition follows from above cases. ∎

Proposition 67 *For any $n \in N$ and $\mathcal{C} \subseteq \mathcal{R}$ with* $\text{card}(\mathcal{C}) \geq 2^{n+1}$, *$\mathcal{C} \notin$* $\mathbf{UniformRobustEx}_n$.

PROOF. Suppose $f_0, f_1, ..., f_{2^{n+1}-1}$ are 2^{n+1} distinct functions in $\mathcal{C}$. Let m be such that, for all i, j with $0 \leq i < j < 2^{n+1}$, $f_i[m] \neq f_j[m]$.

Let $p_i^{e,n}$ be as in Proposition 66.

Suppose by way of contradiction that g is such that for all e, if Θ_e is general recursive, then $\mathbf{M}_{g(e)}$ $\mathbf{Ex}_n$-identifies $\Theta_e(\mathcal{C})$.

Then by Kleene recursion theorem [Rog67], there exists an e such that Θ_e may be described as follows.

$$\Theta_e(f) = \begin{cases} \varphi_{p_i^{g(e),n}}, & \text{if } f[m] = f_i[m] \text{, for some } i < 2^{n+1}; \\ \text{Zero}, & \text{otherwise.} \end{cases}$$

Clearly, then by Proposition 66, $\mathbf{M}_{g(e)}$ does not $\mathbf{Ex}_n$-identify $\Theta_e(\mathcal{C})$. A contradiction. ∎

Corollary 68 *For any $n \in N$ and $\mathcal{C} \subseteq \mathcal{R}$, $\mathcal{C} \in$ **UniformRobustEx**$_n$ iff* $\mathrm{card}(\mathcal{C}) < 2^{n+1}$.

Clearly, Corollary 68 yields that all the types $\mathbf{Ex}_n$, $n \in N$, are uniformly robustly poor.

Corollary 69 *For any $n \in N$, $\mathbf{RobustEx}_0 \not\subseteq \mathbf{UniformRobustEx}_n$.*

PROOF. This follows immediately from Propositions 64 and 67. ∎

Corollary 70 $\mathbf{RobustEx}_0 = \bigcup_{n \in N} \mathbf{UniformRobustEx}_n$.

PROOF. This follows immediately from Proposition 64 and Corollary 68. ∎

Corollaries 69 and 70 above give a situation which is quite rare in the sense that in most cases if one can diagonalize against each of $\mathbf{I}_n$, one would expect to be able to diagonalize against $\bigcup_{n \in N} \mathbf{I}_n$.

Corollary 71 *For any $n \in N$, $\mathbf{UniformRobustEx}_n \subset \mathbf{RobustEx}_n$.*

PROOF. Since $\mathbf{RobustEx}_0 \subseteq \mathbf{RobustEx}_n$, we have $\mathbf{RobustEx}_n - \mathbf{UniformRobustEx}_n \neq \emptyset$ by Corollary 69. ∎

The following Propositions 72 and 73 are needed in proving Theorem 75 below. From this theorem, we then derive that uniformly robust learning is stronger than robust learning for each of the types $\mathbf{Cons}, \mathbf{Ex}$ and $\mathbf{Bc}$.

Proposition 72 *There exists a K-recursive sequence of initial segments, $\sigma_0, \sigma_1, \ldots \in SEG_{0,1}$, such that for all $e \in N$, the following are satisfied.*

(a) $0^e 1 \subseteq \sigma_e$.

(b) *For all $e' \leq e$, if $\Theta_{e'}$ is general recursive, then either $\Theta_{e'}(\sigma_e) \not\sim \Theta_{e'}(0^{|\sigma_e|})$ or for all $f \in \mathcal{T}_{0,1}$ extending σ_e, $\Theta_{e'}(f) = \Theta_{e'}(0^\infty)$.*

PROOF. We define σ_e (using oracle for K) as follows. Initially, let $\sigma_e^0 = 0^e 1$. For $e' \leq e$, define $\sigma_e^{e'+1}$ as follows: if there exists an extension $\tau \in \mathrm{SEG}_{0,1}$ of $\sigma_e^{e'}$, such that $\Theta_{e'}(\tau) \not\sim \Theta_{e'}(0^{|\tau|})$, then let $\sigma_e^{e'+1} = \tau$; otherwise, let $\sigma_e^{e'+1} = \sigma_e^{e'}$.

Now let $\sigma_e = \sigma_e^{e+1}$ as defined above. It is easy to verify that the proposition is satisfied. ∎

Proposition 73 *There exists an infinite increasing sequence $a_0, a_1, \ldots$ of natural numbers such that for $A = \{a_i \mid i \in N\}$, the following properties are satisfied for all $k \in N$.*

(a) *The complement of A is recursively enumerable relative to K.*

(b) *φ_{a_k} is total.*

(c) *For all $e \leq a_k$ such that φ_e is total, $\varphi_e(x) \leq \varphi_{a_{k+1}}(x)$ for all $x \in N$.*

(d) *For σ_e as defined in Proposition 72, $|\sigma_{a_k}| \leq a_{k+1}$.*

PROOF. The construction of a_i's is done using movable markers (using oracle for K). Let a_i^s denote the value of a_i at the beginning of stage s in the construction. It will be the case that, for all s and i, either $a_i^s = a_i^{s+1}$, or $a_i^{s+1} > s$. This allows us to ensure property (a). The construction itself directly implements properties (b) to (d). Let *pad* be a 1–1 padding function [Rog67] such that for all i, j, $\varphi_{pad(i,j)} = \varphi_i$, and $pad(i,j) \geq i + j$.

We assume without loss of generality that φ_0 is total. Initially, let $a_0^0 = 0$, and $a_{i+1}^0 = pad(0, |\sigma_{a_i^0}|)$ (this ensures $a_{i+1}^0 \geq |\sigma_{a_i^0}| > a_i^0$). Go to stage 0.

Stage s
1. If there exist a k, $0 < k \leq s$, and $x \leq s$ such that:
   (i) $\varphi_{a_k^s}(x)\uparrow$ or
   (ii) for some $e \leq a_{k-1}^s$, $[(\forall y \leq s)[\varphi_e(y)\downarrow]$ and $\varphi_e(x) > \varphi_{a_k^s}(x)]$
2. Then pick least such k and go to step 3. If there is no such k, then for all i, let $a_i^{s+1} = a_i^s$, and go to stage $s+1$.
3 For $i < k$, let $a_i^{s+1} = a_i^s$.
4. Let j be the least number such that
   (i) $(\forall y \leq s)[\varphi_j(y)\downarrow]$ and
   (ii) for all $e \leq a_{k-1}^s$, if for all $y \leq s$, $\varphi_e(y)\downarrow$, then for all $y \leq s$, $\varphi_j(y) \geq \varphi_e(y)$.
   Let $a_k^{s+1} = pad(j, |\sigma_{a_{k-1}^s}| + s + 1)$.
5. For $i > k$, let $a_i^{s+1} = pad(0, |\sigma_{a_{i-1}^{s+1}}| + s + 1)$.
6. Go to stage $s+1$.
End stage s

We claim (by induction on k) that $\lim_{s\to\infty} a_k^s\downarrow$ for each k. To see this, note that once all the a_i, $i < k$, have stabilized, step 4 would eventually pick a j such that φ_j is total, and for all $e \leq a_{k-1}$, if φ_e is total then $\varphi_e \leq \varphi_j$. Thereafter a_k would not be changed.

We now show the various properties claimed in the proposition. One can enumerate $\overline{A}$ (using oracle for K) using the following property: $x \in \overline{A}$ iff there exists a stage $s > x$ such that, for all $i \leq x$, $a_i^s \neq x$. Thus (a) holds. (b) and (c) hold due to the check in step 1. (d) trivially holds due to padding used for definition of a_i^s for all s. ∎

Definition 74 Suppose $h \in \mathcal{R}$. Let $\mathcal{B}_h = \{\varphi_e \mid \varphi_e \in \mathcal{R} \ \wedge \ (\forall^\infty x)[\Phi_e(x) \leq h(x)]\}$.

Intuitively, $\mathcal{B}_h$ denotes the class of recursive functions whose complexity is almost everywhere bounded by h. We assume without loss of generality that FINSUP $\subseteq \mathcal{B}_{\varphi_0}$. Thus for a_i as in Proposition 73, FINSUP $\subseteq \mathcal{B}_{\varphi_{a_i}}$, for all i.

Theorem 75 RobustCons $\not\subseteq$ UniformRobustBc.

Proof. Fix $\sigma_0, \sigma_1, \ldots$ as in Proposition 72, and $a_0, a_1, \ldots$ as in Proposition 73.

Let $G_k = \mathcal{B}_{\varphi_{a_k}} \cap \{f \mid \sigma_{a_k} \subseteq f\}$.

The main idea of the construction is to construct the diagonalizing class by taking at most finitely many functions from each G_k.

Let Θ_{b_k} be defined as

$$\Theta_{b_k}(f[n]) = \begin{cases} \Lambda, & \text{if } n < |\sigma_{a_k}|; \\ f(|\sigma_{a_k}|)f(|\sigma_{a_k}|+1)\ldots f(n-1), & \text{if } \sigma_{a_k} \subseteq f[n]; \\ f[n], & \text{otherwise.} \end{cases}$$

Note that Θ_{b_k} is general recursive.

Claim 76 $\bigcup_{i \geq k} \Theta_{b_k}(G_i) \notin$ **Bc**.

Proof. Suppose by way of contradiction that **M Bc**-identifies $\bigcup_{i \geq k} \Theta_{b_k}(G_i)$. Then, clearly **M** must **Bc**-identify FINSUP (since FINSUP $\subseteq \Theta_{b_k}(G_k)$). Thus, for all $\tau \in \mathrm{SEG}_{0,1}$, there exists an n such that, for all $m \geq n$, $\varphi_{\mathbf{M}(\tau 0^m)}(|\tau|+m) = 0$. Thus, there exists a recursive function g such that, for all $\tau \in \mathrm{SEG}_{0,1}$ satisfying $|\tau| \leq n$, $\varphi_{\mathbf{M}(\tau 0^{g(n)})}(|\tau| + g(n)) = 0$. Now for each τ, define f_τ inductively by letting $\eta_0 = \tau$, $\eta_{n+1} = \eta_n \cdot 0^{g(|\eta_n|)} \cdot 1$ and $f_\tau = \bigcup_n \eta_n$. Note that **M** does not **Bc**-identify any f_τ. Also, f_τ is uniformly (in τ) computable and thus $\{f_\tau \mid \tau \in \mathrm{SEG}_{0,1}\} \subseteq \mathcal{B}_h$, for some recursive h. Thus, for sufficiently large j, $\{f_\tau \mid \tau \in \mathrm{SEG}_{0,1}\} \subseteq \mathcal{B}_{\varphi_{a_j}}$. Thus, for almost all $j > k$, $f_{\sigma_{a_j}} \in G_j = \Theta_{b_k}(G_j)$.

Since **M** does not **Bc**-identify $f_{\sigma_{a_j}}$, claim follows. □

Let g' be a function dominating all K'-recursive functions. For each $k \in N$ and $e \leq g'(k)$, let $f_{k,e}$ denote a function in $\bigcup_{i \geq k} G_i$, such that $\mathbf{M}_e$ does not **Bc**-identify $\Theta_{b_k}(f_{k,e})$.

Let $\mathcal{S} = \{f_{k,e} \mid k \in N, e \leq g'(k)\}$. Let $F_k = \mathcal{S} \cap G_k$. It is easy to verify that F_k is finite (since $f_{k,e} \notin \bigcup_{i<k} G_i$).

Claim 77 $\mathcal{S} \notin$ **UniformRobustBc**.

Proof. Suppose by way of contradiction that h is a recursive function such that $\mathbf{M}_{h(e)}$ **Bc**-identifies $\Theta_e(\mathcal{S})$. Note that b_k can be recursively computed using oracle for K. Thus, $h(b_k)$ can be recursively computed using oracle for K. Hence, for all but finitely many k, $h(b_k) \leq g'(k)$. Consequently, $\mathbf{M}_{h(b_k)}$ does not **Bc**-identify $\Theta_{b_k}(f_{k,h(b_k)}) \in \Theta_{b_k}(\mathcal{S})$. A contradiction. $\square$

Claim 78 $\mathcal{S} \in$ **RobustCons**.

Proof. Suppose $\Theta = \Theta_k$ is general recursive. We need to show that $\Theta_k(\mathcal{S}) \in$ **Cons**. Let $A = \{a_i \mid i \in N\}$. Since $\overline{A}$ is r.e. in K, there exists a recursive sequence $c_0, c_1, \ldots$, such that each $a \in A$, $a > a_k$, appears infinitely often in the sequence, and each $a \notin A$ or $a \leq a_k$, appears only finitely often in the sequence. Let $\sigma_{e,t} \in \mathrm{SEG}_{0,1}$ be such that $\sigma_{e,t} \supseteq 0^e1$, and $\sigma_{e,t}$ can be obtained effectively from e, t, and $\lim_{t\to\infty} \sigma_{e,t} = \sigma_e$. Note that there exist such $\sigma_{e,t}$ due to K-recursiveness of the sequence $\sigma_0, \sigma_1, \ldots$.

Note that there exists a recursive h such that, if φ_e is recursive then, $\mathbf{M}_{h(e)}$ **Cons**-identifies $\Theta(\mathcal{B}_{\varphi_e})$. Fix such recursive h.

Let $F = \{0^\infty\} \cup F_0 \cup F_1 \cup \ldots \cup F_k$. F and $\Theta(F)$ are finite sets of total recursive functions.

Define $\mathbf{M}$ as follows.

$\mathbf{M}(f[n])$
1. If for some $g \in \Theta(F)$, $g[n] = f[n]$, then output a canonical program for one such g.
2. Else, let $t \leq n$ be the largest number such that $\Theta(\sigma_{c_t,n}) \sim f[n]$, and $\Theta(\sigma_{c_t,n}) \not\sim \Theta(0^\infty)$.

   Dovetail the following steps until one of them succeeds. If steps 2.1 or 2.2 succeed, then go to step 3. If step 2.3 succeeds, then go to step 4.

   2.1 There exists an $s > n$, such that $c_s \neq c_t$, and $\Theta(\sigma_{c_s,s}) \sim f[n]$, and $\Theta(\sigma_{c_s,s}) \not\sim \Theta(0^\infty)$.

   2.2 There exists an $s > n$, such that $\sigma_{c_t,s} \neq \sigma_{c_t,n}$.

   2.3 $\mathbf{M}_{h(c_t)}(f[n])\downarrow$, and $f[n] \subseteq \varphi_{\mathbf{M}_{h(c_t)}(f[n])}$.
3. Output a program for $f[n]0^\infty$.
4. Output $\mathbf{M}_{h(c_t)}(f[n])$.

End

It is easy to verify that whenever $\mathbf{M}(f[n])$ is defined, $f[n] \subseteq \varphi_{\mathbf{M}(f[n])}$. Also, if $f \in \Theta(F)$, then **M Cons**-identifies f.

Now, consider any $f \in \Theta(\mathcal{S}) - \Theta(F)$. Note that there exists a unique $i > k$ such that $f \sim \Theta(\sigma_{a_i})$ and $\Theta(\sigma_{a_i}) \not\sim \Theta(0^\infty)$ (due to definition of σ_{a_j}'s). Fix such i. Also, since $f \neq \Theta(0^\infty)$, there exist only finitely many e such that $f \sim \Theta(0^e 1)$.

We first claim that $\mathbf{M}(f[n])$ is defined for all n. To see this, note that if $c_t \neq a_i$ or $\sigma_{c_t,n} \neq \sigma_{a_i}$, then step 2.1 or step 2.2 would eventually succeed. Otherwise, since $f \in \Theta(F_i) \subseteq \Theta(\mathcal{B}_{\varphi_{a_i}})$, step 2.3 would eventually succeed (since $\mathbf{M}_{h(a_i)}$ **Cons**-identifies $\Theta(\mathcal{B}_{\varphi_{a_i}})$).

Thus, it suffices to show that **M Ex**-identifies f. Let r be such that $f \not\sim \Theta(0^r)$. Let m and $n > m$ be large enough such that (i) to (iv) hold.

(i) $f[n] \not\sim \Theta(0^r)$.

(ii) $c_m = a_i$, and for all $s \geq m$, $\sigma_{a_i,s} = \sigma_{a_i,m}$.

(iii) For all $e < r$ and $t > m$, if $e \notin A$ or $e \leq a_k$, then $c_t \neq e$.

(iv) For all $e < r$ and $t > m$, if $e \in A - \{a_i\}$ and $e > a_k$, then $\Theta(\sigma_{e,t}) \not\sim f[n]$ or $\Theta(\sigma_{e,t}) \sim \Theta(0^\infty)$.

Note that there exist such m, n. Thus, for all $n' \geq n$, in computation of $\mathbf{M}(f[n'])$, c_t would be a_i, and step 2.1 and step 2.2 would not succeed. Thus step 2.3 would succeed, and **M** would output $\mathbf{M}_{h(a_i)}(f[n'])$. Thus **M Ex**-identifies f, since $\mathbf{M}_{h(a_i)}$ **Ex**-identifies f. $\square$

Theorem follows from above claims. ∎

Corollary 79 **UniformRobustCons** $\subset$ **RobustCons**.

UniformRobustEx $\subset$ **RobustEx**.

UniformRobustBc $\subset$ **RobustBc**.

6 Conclusion

In this paper we investigated, for many learnability criteria, whether there are rich function classes which are robustly or uniformly robustly learnable under these criteria.

Furthermore, we showed that every uniformly robustly **Ex**-learnable class is also consistently learnable. It remains an open problem whether the adverb "uniformly" can be dropped: Is every robustly **Ex**-learnable class also consistently learnable?

In addition, we explored the relationship between robust and uniformly robust variants of several learnability criteria. Specifically, we separated robust **Ex**-learning from its uniformly robust counterpart, solving an open problem from [JSW01].

7 Acknowledgements

We thank the anonymous referees for several helpful comments.

References

[AB92] M. Anthony and N. Biggs. *Computational Learning Theory.* Cambridge University Press, 1992.

[AS83] D. Angluin and C. Smith. Inductive inference: Theory and methods. *Computing Surveys*, 15:237–289, 1983.

[Bar74a] J. Bārzdiņš. Inductive inference of automata, functions and programs. In *Int. Math. Congress, Vancouver*, pages 771–776, 1974.

[Bar74b] J. Bārzdiņš. Two theorems on the limiting synthesis of functions. In *Theory of Algorithms and Programs, vol. 1*, pages 82–88. Latvian State University, 1974. In Russian.

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

[BF74] J. Bārzdiņš and R. Freivalds. Prediction and limiting synthesis of recursively enumerable classes of functions. *Latvijas Valsts Univ. Zimatm. Raksti*, 210:101–111, 1974.

[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.

[CF99] C. Costa Florencio. Consistent identification in the limit of some Penn and Buszkowski's classes is NP-hard. In P. Monachesi, editor, *Computational Linguistics in the Netherlands*, pages 1–12, 1999.

[CJO+00] J. Case, S. Jain, M. Ott, A. Sharma, and F. Stephan. Robust learning aided by context. *Journal of Computer and System Sciences (Special Issue for COLT'98)*, 60:234–257, 2000.

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

[FBP91] R. Freivalds, J. Bārzdiņš, and K. Podnieks. Inductive inference of recursive functions: Complexity bounds. In J. Bārzdiņš and D. Bjørner, editors, *Baltic Computer Science*, volume 502 of *Lecture Notes in Computer Science*, pages 111–155. Springer-Verlag, 1991.

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

[Ful88] M. Fulk. Saving the phenomenon: Requirements that inductive machines not contradict known data. *Information and Computation*, 79:193–209, 1988.

[Ful90] M. Fulk. Robust separations in inductive inference. In *31st Annual IEEE Symposium on Foundations of Computer Science*, pages 405–410. IEEE Computer Society Press, 1990.

[FW79] R. Freivalds and R. Wiehagen. Inductive inference with additional information. *Journal of Information Processing and Cybernetics (EIK)*, 15:179–195, 1979.

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

[Gra86] J. Grabowski. Starke Erkennung. In R. Lindner and H. Thiele, editors, *Strukturerkennung diskreter kybernetischer Systeme, Teil I*, pages 168–184. Seminarbericht Nr.82, Department of Mathematics, Humboldt University of Berlin, 1986. In German.

[Jai99] S. Jain. Robust behaviorally correct learning. *Information and Computation*, 153(2):238–248, September 1999.

[JB81] K. P. Jantke and H.-R. Beick. Combining postulates of naturalness in inductive inference. *Journal of Information Processing and Cybernetics (EIK)*, 17:465–484, 1981.

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

[JSW01] S. Jain, C. Smith, and R. Wiehagen. Robust Learning is Rich. *Journal of Computer and System Sciences*, 62:178–212, 2001.

[KS89a] S. Kurtz and C. Smith. On the role of search for learning. In R. Rivest, D. Haussler, and M. Warmuth, editors, *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 303–311. Morgan Kaufmann, 1989.

[KS89b] S. Kurtz and C. Smith. A refutation of Bārzdiņš' conjecture. In K. P. Jantke, editor, *Analogical and Inductive Inference, Proceedings of the Second International Workshop (AII '89)*, volume 397 of *Lecture Notes in Artificial Intelligence*, pages 171–176. Springer-Verlag, 1989.

[KW80] R. Klette and R. Wiehagen. Research in the theory of inductive inference by GDR mathematicians – A survey. *Information Sciences*, 22:149–169, 1980.

[Lan90] S. Lange. Consistent polynomial-time inference of k-variable pattern languages. In J. Dix, K. P. Jantke, and P. Schmitt, editors, *Nonmonotonic and Inductive Logic, 1st International Workshop, Karlsruhe, Germany*, volume 543 of *Lecture Notes in Computer Science*, pages 178–183. Springer-Verlag, 1990.

[Min76] E. Minicozzi. Some natural properties of strong identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.

[Mit97] T. Mitchell. *Machine Learning*. McGraw Hill, 1997.

[OS02] M. Ott and F. Stephan. Avoiding coding tricks by hyperrobust learning. *Theoretical Computer Science*, 284:161–180, 2002.

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

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

[Ste98] W. Stein. Consistent polynomial identification in the limit. In M. M. Richter, C. H. Smith, R. Wiehagen, and T. Zeugmann, editors, *Algorithmic Learning Theory: Ninth International Conference (ALT' 98)*, volume 1501 of *Lecture Notes in Artificial Intelligence*, pages 424–438. Springer-Verlag, 1998.

[Vap00] V. N. Vapnik. *The Nature of Statistical Learning Theory. Second Edition*. Springer-Verlag, 2000.

[Wie76] R. Wiehagen. Limes-Erkennung rekursiver Funktionen durch spezielle Strategien. *Journal of Information Processing and Cybernetics (EIK)*, 12:93–99, 1976.

[Wie78] R. Wiehagen. *Zur Theorie der Algorithmischen Erkennung*. Dissertation B, Humboldt University of Berlin, 1978.

[WL76] R. Wiehagen and W. Liepe. Charakteristische Eigenschaften von erkennbaren Klassen rekursiver Funktionen. *Journal of Information Processing and Cybernetics (EIK)*, 12:421–438, 1976.

[WZ94] R. Wiehagen and T. Zeugmann. Ignoring data may be the only way to learn efficiently. *Journal of Experimental and Theoretical Artificial Intelligence*, 6:131–144, 1994.

[WZ95] R. Wiehagen and T. Zeugmann. Learning and consistency. In K. P. Jantke and S. Lange, editors, *Algorithmic Learning for Knowledge-Based Systems*, volume 961 of *Lecture Notes in Artificial Intelligence*, pages 1–24. Springer-Verlag, 1995.

[Zeu86] T. Zeugmann. On Bārzdiņš' conjecture. In K. P. Jantke, editor, *Analogical and Inductive Inference, Proceedings of the International Workshop*, volume 265 of *Lecture Notes in Computer Science*, pages 220–227. Springer-Verlag, 1986.