Classes with Easily Learnable Subclasses

Sanjay Jain [a,1] Wolfram Menzel [b] Frank Stephan [c,2]

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

[b] Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe, 76128 Karlsruhe, Germany, menzel@ira.uka.de

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

Abstract

In this paper we study the question of whether identifiable classes have subclasses which are identifiable under a more restrictive criterion. The chosen framework is inductive inference, in particular the criterion of explanatory learning (Ex) of recursive functions as introduced by Gold in 1967. Among the more restrictive criteria is finite learning where the learner outputs, on every function to be learned, exactly one hypothesis (which has to be correct). The topic of the present paper are the natural variants (a) and (b) below of the classical question whether a given learning criterion like finite learning is more restrictive than Ex-learning. (a) Does every infinite Ex-identifiable class have an infinite finitely identifiable subclass? (b) If an infinite Ex-identifiable class S has an infinite finitely identifiable subclass, does it necessarily follow that some appropriate learner Ex-identifies S as well as finitely identifies an infinite subclass of S? These questions are also treated in the context of ordinal mind change bounds.

[1] Sanjay Jain was supported in part by NUS grant number R252-000-127-112.

[2] While previously working at the Mathematical Institute of the Universitity of Heidelberg, Frank Stephan was supported by the Deutsche Forschungsgemeinschaft (DFG), Heisenberg grant Ste 967/1–1. National ICT Australia is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence Program.

1 Introduction

Gold [6] introduced a model of learning computable functions, where a learner receives increasing amounts of data about an unknown function and outputs a sequence of hypotheses. The learner has learned or identified the function, if it converges to a single explanation, that is, a program for the function at hand. This concept of explanatory or Ex-learning has been widely studied [3,6,10,14]; see Definition 2.2 below for formal details.

An explanatory learner is often not aware of the fact whether it has already learned the function f or whether the current hypothesis is a preliminary one which must be revised later. It is well-known that, for various restrictive learning criteria, there is a class S which is explanatorily learnable but cannot be learned according to the more restrictive learning type. One might ask, whether there are at least sufficiently large subclasses U of S with better learnability properties. For example, one could impose that the learner on functions from U follows the criterion of finite learning [6], where the learner outputs exactly one hypothesis (which must be correct) on functions from the class, see Definition 2.3 below. In this paper we will consider similar questions for some commonly used criteria of learning, which are at least as restrictive as Ex. Motivation for this comes from various studies in mathematics where one pursues the general theme of when a difficult object can be approximated by a simple object. For example, it is well known that every infinite recursively enumerable set has an infinite recursive subset.

A well-behaved learner satisfies some natural requirements on its behaviour, see Definition 3.1. Such a learner only outputs hypotheses which are extended by total functions from the class to be learned. Furthermore such a learner is consistent whenever it outputs a hypothesis. It turns out that every uniformly recursive class S can be learned by a well-behaved learner (here S is uniformly recursive if there is an enumeration $f_0, f_1, \ldots$ such that $S = \{f_0, f_1, \ldots\}$ and the function $e, x \to f_e(x)$ is recursive in both parameters). Theorem 4.1 shows that the converse is not true: the theorem gives an example of an infinite class, which has a well-behaved learner, while every intersection of the class with a uniformly recursive class is finite.

It is shown that there is an infinite uniformly recursive class without any infinite finitely learnable subclass. This result can be generalized by considering confident learning instead of finite learning. While a finite learner outputs at most one hypothesis on any function, a confident learner may output unbounded, but finitely many, hypothesis on any function (even nonrecursive ones). Theorem 4.2 provides directly this generalized result by giving an example of an infinite uniformly recursive class which does not have an infinite intersection with any confidently learnable class.

Sublearning deals with questions like the following: Is there a learner M

which explanatorily learns a class S and – at the same time – finitely learns an infinite subclass U? In Theorem 5.1 it is shown that there is an explanatorily learnable class S which has an infinite finitely learnable subclass but which does not have a sublearner.

Ordinal counters are used to introduce a hierarchy of mind changes within the concept of confident learning. It turns out that ordinals which are a power of ω, in the way defined in Remark 2.8, play a special role in this theory. Theorem 4.7 states that for a recursive ordinal $\alpha = \omega^\gamma$, with $\gamma \geq 1$, the following holds: There is an infinite class, which has a learner using α mind changes, but no infinite subclass of this class can be learned by a learner using β mind changes, for $\beta < \alpha$. For other recursive ordinals $\alpha \geq 2$ such a class does not exist. Theorem 5.4 is the version of Theorem 4.7 in the context of sublearning.

2 Preliminaries

Notation 2.1. Recursion theoretic notation mainly follows the books of Odifreddi [13,14] and Soare [16]. Let $\mathbb{N} = \{0, 1, \ldots\}$ be the set of natural numbers. For any set $A \subseteq \mathbb{N}$, A^* is the set of finite strings over A and A^∞ the set of total functions from $\mathbb{N}$ to A (viewed as infinite strings). Furthermore, sets are often identified with their characteristic functions, so we may write $A(n) = 1$ for $n \in A$ and $A(n) = 0$ for $n \in \overline{A}$. For a function f, $f[n]$ denotes the string $f(0)f(1)f(2)\ldots f(n-1)$. λ denotes the empty sequence. Strings are viewed upon as partial functions; $\sigma \subseteq \psi$ denotes that ψ extends σ as a partial function. $\sigma \subset \tau$ denotes that τ extends σ properly. $\sigma\tau$ denotes the concatenation of strings σ and τ. σa^m denotes the function coinciding with σ on the domain of σ, taking the value a on the next m inputs and being undefined after that in the case of $m < \infty$; σa^∞ is total. Let φ be a standard acceptable numbering, and φ_e denote the e-th partial recursive function in this numbering.

Definition 2.2 (Explanatory Learning) [6]**.** A *learner* is a total recursive function mapping finite sequences of natural numbers to $\mathbb{N} \cup \{?\}$. An output of M is called *hypothesis* if it is different from ?. Hypotheses are viewed upon as indices for partial recursive functions according to our underlying acceptable numbering φ.

We say that a learner M *converges* on f to a hypothesis e iff for all but finitely many n, $M(f[n]) = e$. A learner M *Ex-learns* (= *Ex-identifies*) a recursive function f if, on input f, it converges to a hypothesis which is a program (or code) for f. We say that M Ex-identifies a class S of recursive functions if and only if M Ex-identifies each function in the class. Ex denotes the family of classes that are learnable by a recursive Ex learner.

The letters "Ex" stand for "explanatory learning".

For learning, we always consider non-empty classes of total and recursive functions. So we can avoid to deal with uninteresting special cases which mess up the statements and proofs of the results but do not give any insight on learning-theory.

Note that the symbol ? stands for the case that the learner cannot make up its mind about what hypothesis to output. The concept of Ex-learning itself does not need this special symbol but additional requirements like bounds on the number of mind changes below will make use of ?, in order to avoid mind changes caused by the lack of data which shows up later.

Definition 2.3 (Mind Change Bounds) [3]**.** We say that a learner M makes a *mind change* on f at n, if there is an $m < n$ such that (I) $M(f[n]) \neq M(f[k])$ for $k = m, m+1, \ldots, n-1$ and (II) $M(f[n]), M(f[m])$ are both different from ?. A class of recursive functions S is in Ex_m, if there is a recursive learner that Ex-learns every $f \in S$ by making at most m mind changes on f. Ex_0-learning without any mind changes is also called *finite learning*.

Definition 2.4 (Consistency) [1,17]**.** A learner M is *consistent* on σ if either (I) $M(\sigma) = ?$ or (II) $M(\sigma)$ outputs an index e such that $\varphi_e(x)\downarrow = \sigma(x)$, for all $x \in \mathrm{domain}(\sigma)$. A learner is *consistent* iff it is consistent on all strings $\sigma \in \mathbb{N}^*$.

Note that the case $M(\sigma) = ?$ was not allowed in the original definition of consistency. Indeed one could remove this case by transforming M to a new learner N which, on input σ, outputs an index for σ if $M(\sigma) = ?$, and outputs the hypothesis $M(\sigma)$ otherwise. However, in order to make it possible that a consistent learner can also be confident or pessimistically reflective (as defined in Definition 2.5) we have explicitly permitted the option that M can output ?.

Furthermore, variants of consistency have been considered. For example, a learner M for a class S is *consistent on S* if it is only required that M is consistent on the strings $f[n]$ with $f \in S$. There are classes S which have a learner which is consistent on S but which do not have a consistent learner. An example is the class $\{f : \varphi_{f(0)} = f\}$ of all self-describing functions.

Definition 2.5 (Further Learning-Criteria). A learner M is *prudent* [15] if it Ex-identifies a total extension of φ_e, for each e in its range. A learner M is *pessimistically reflective* [7,8] if M Ex-identifies an extension of σ whenever $M(\sigma) \neq ?$. A learner M is said to be *confident* [15] if it converges on every total function, even the non-recursive functions.

Exact-learning defined below gives a closer connection between the learner and the class to be learned, which goes beyond the fact that the learner identifies the class.

Definition 2.6 (Exact Learning). (Osherson, Stob and Weinstein [15]) For a criterion I which is at least as restrictive as Ex, one says that M *exactly I-identifies* a class S if and only if M I-identifies every function $f \in S$ and does not even Ex-identify any function $f \notin S$.

Note that in the present work, the term *exact learning* is used as in the book "Systems that learn" [10, Definition 4.48]. Therefore this notion differs from the one with the same name used in the field of learning classes represented by indexed families [18,19]. In [10], the following motivation is given for the notion of exact learning in the context of language learning.

> The converse of the dictum that natural languages are learnable by children (via casual exposure) is that nonnatural languages are not learnable. Put differently, the natural languages are generally taken to be the *largest* collection of child-learnable languages. We are thus led to consider paradigms in which learners are required to respond successfully to all languages in a given collection and to respond unsuccessfully to all other languages.

Similar considerations also motivate the notion of exact learning for functions as considered in this paper.

A family $f_0, f_1, \ldots$ of total functions is called *uniformly recursive* if the two-place function $e, x \to f_e(x)$ is recursive. In order to simplify notation, we say that a class S is *uniformly recursive* iff $S = \{f_0, f_1, \ldots\}$ for a uniformly recursive family $f_0, f_1, \ldots$ of functions. The following notion Num captures the subclasses of uniformly recursive classes.

Definition 2.7. A class S of recursive functions is in *Num* if some superclass S' of S is a uniformly recursive class.

Remark 2.8 (Ordinals). Let $<_0, <_1, \ldots$ be an enumeration of all recursively enumerable partial orders. If an ordering $<_e$ is a well-ordering, it is called a *notation for ordinals*. The natural numbers equipped with $<_e$ are isomorphic to an initial segment of the class of all countable ordinals and one can identify every number x with that ordinal α for which $\{y : y <_e x\}$ and $\{\beta : \beta < \alpha\}$ are order-isomorphic sets.

Cantor introduced a *non-commutative* addition $+$ on the ordinals which is invertible: if $\alpha \leq \beta$, there is a unique γ such that $\alpha + \gamma = \beta$. This difference γ is denoted as $\beta - \alpha$. Halmos [9, Section 21] gives an overview on ordinal arithmetic. If $<_e$ is a notation for ordinals having a representative x for α, then there is a notation $<_{e'}$ such that whenever y represents an ordinal $\alpha + \beta$

with respect to $<_e$ then y represents the ordinal β with respect to $<_{e'}$. The ordering $<_{e'}$ is constructed by shifting the part of the ordering strictly below x to the top so that $<_{e'}$ is still a well-ordering and x represents 0:

$$y <_{e'} z \Leftrightarrow (x \leq_e y <_e z) \vee (y <_e z <_e x) \vee (z <_e x \leq_e y)$$

where $x \leq_e y$ stands for $x = y \vee x <_e y$.

Furthermore, Cantor introduced the formal powers of ω, the first infinite ordinal. Cantor showed that one can represent every non-null ordinal by a finite sum

$$\alpha = a_1\omega^{\alpha_1} + a_2\omega^{\alpha_2} + \ldots + a_n\omega^{\alpha_n}$$

where $0 \leq \alpha_n < \ldots < \alpha_2 < \alpha_1$ as ordinals and $a_1, a_2, \ldots, a_n$ are non-null natural numbers [14, page 280].

This representation permits us to view the ordinals as a semimodule over the semiring of the natural numbers with pointwise operations $\oplus, \ominus, \otimes$. Given ordinal α and natural number c, one can define $c \otimes \alpha$ as follows. If $c = 0$ or $\alpha = 0$ then $c \otimes \alpha$ is just 0. Otherwise α has the unique representation $a_1\omega^{\alpha_1} + a_2\omega^{\alpha_2} + \ldots + a_n\omega^{\alpha_n}$ and one defines $c \otimes \alpha = (a_1c)\omega^{\alpha_1} + (a_2c)\omega^{\alpha_2} + \ldots + (a_nc)\omega^{\alpha_n}$. Similarly, one can define the pointwise addition $\alpha \oplus \beta$ which is different from $+$ as it is commutative but has the minimum compatibility $\alpha \oplus 1 = \alpha + 1$. Note that $\alpha \ominus \beta$, the pointwise subtraction, can be undefined even in the case that $\beta < \alpha$: for example, $\omega \ominus 1$ is undefined.

Definition 2.9 [4]**.** A class S is Ex_α-identifiable for a recursive ordinal α iff there is an Ex-learner M, a notation for ordinals $<_e$ having a notation r_α for α, and a total recursive function ord mapping $\mathbb{N}^*$ to $\mathbb{N}$ such that the following hold.

(a) M Ex-identifies every $f \in S$.
(b) $\mathrm{ord}(\lambda) \leq_e r_\alpha$.
(c) For all total f and m, n such that $m < n$, $\mathrm{ord}(f[n]) \leq_e \mathrm{ord}(f[m])$.
(d) For all $f \in S$ and m, n such that $m < n$, $M(f[n]) \neq ?$, $M(f[m]) \neq ?$, and $M(f[n]) \neq M(f[m])$: $\mathrm{ord}(f[n]) <_e \mathrm{ord}(f[m])$.

Remark 2.10. Freivalds and Smith [4] postulated that (d) holds also for all function $f \notin S$. The resulting concept is the same, but in the present paper the restrictions to functions in S will be necessary for studying simultaneous learners. For example, we will consider the case where a learner M simultaneously Ex-identifies R and Ex_α-identifies some $S \subseteq R$. As this class R itself might not be Ex_α-identifiable, the existence of such a simultaneous learner is only possible in a setting where condition (d) is defined as above.

Note that for some $\alpha \geq \omega$ and some classes $S \in \mathrm{Ex}_\alpha$, one must carefully choose the adequate notation for ordinals in order to construct a recursive Ex_α-learner using this notation. If the notation is chosen inadequately, it might happen that the corresponding learner cannot be recursive.

3 Well-Behaved Learners

In this section we introduce the notion of well-behaved learners. Well behaved learners combine the properties of exact, prudent, pessimistically reflective and consistent learners.

Definition 3.1. A learner M is well-behaved for S iff

(a) M exactly Ex-learns S, that is, M Ex-learns f iff $f \in S$;
(b) M is prudent, that is, for all σ with $M(\sigma) \neq ?$, M Ex-learns a function f extending $\varphi_{M(\sigma)}$;
(c) M is consistent, that is, for all σ with $M(\sigma) \neq ?$, $\varphi_{M(\sigma)}$ extends σ.

Every well-behaved learner is pessimistically reflective: If $M(\sigma)$ is an index e, then φ_e extends σ (by consistency) and some f Ex-learned by M extends φ_e (by prudence). Thus, M identifies an extension of σ, whenever $M(\sigma) \neq ?$.

If one would add the property of being pessimistically reflective to the postulated conditions for well-behaved learners, then one could weaken (c) in such a way that M is only required to be consistent on S (since M, being pessimistically reflective, will always output ? on data not belonging to any function in S).

Remark 3.2. Every uniformly recursive class $S = \{f_0, f_1, \ldots\}$ has a well-behaved learner M. This is shown by choosing M as follows: $M(a_0 a_1 \ldots a_n)$ outputs the least $e \leq n$ such that $f_e(m) = a_m$, for $m = 0, 1, \ldots, n$, and outputs ? if such an e is not found.

On the one hand, there are Ex-learnable classes in Num which are not uniformly recursive and even not prudently learnable by an exact learner. An example is the class $S = \{c^\infty : c \notin K\}$ where K is the halting problem. An exact Ex-learner for S can be constructed as follows. On input c^n, such that $n > 0$ and c is not enumerated into K within n computation-steps, the learner outputs a hypothesis for c^∞; otherwise the learner outputs the symbol ?. For any Ex-learner M for S, the set $\{c : (\exists \sigma)\,(\exists x)\,[M(\sigma)$ is a hypothesis that computes c on argument $x]\}$ is a recursively enumerable superset of $\overline{K}$. Thus, M cannot be an exact prudent Ex-learner for S.

On the other hand, there are classes which have a well-behaved learner but which are not in Num. This result can even be strengthened as shown in Theorem 4.1 below.

We now give some results relating well-behaved learners and exact learners which are in addition prudent or pessimistically reflective.

Proposition 3.3. *If a prudent learner M exactly Ex_0-identifies S, then S is uniformly recursive.*

Proof. Recall that in Definition 2.6 it was defined that the learner M exactly Ex_0-identifies S iff the learner M Ex_0-identifies all functions in S and does not Ex-identify any function outside S. Since M is also prudent, every index output by M is extended by a function in S.

Let $E = \{e : (\exists\sigma)\,[M(\sigma) = e \wedge \varphi_e \text{ extends } \sigma]\}$. The set E is recursively enumerable. If $f \in S$ then there is a prefix $\sigma \subseteq f$ such that $M(\sigma)$ outputs an index e for f and this index e is in E. So $S \subseteq \{\varphi_e : e \in E\}$.

If $e \in E$ and σ witnesses $e \in E$, then a function $f \in S$ extends φ_e and thus σ. Since M Ex_0-learns f and M outputs exactly one hypothesis while reading f, this hypothesis is e and thus $f = \varphi_e$. Hence, $\{\varphi_e : e \in E\} \subseteq S$.

Thus, $S = \{\varphi_e : e \in E\}$ and S is uniformly recursive. ▮

The condition of being prudent is necessary. For example, the class $\{f : \varphi_{f(0)} = f\}$ of self-describing functions has an exact Ex_0-learner which on input $f(0)f(1)\ldots f(n)$ outputs $f(0)$. However, this class is not in Num.

Theorem 3.4. *There is a class R having an exact pessimistically reflective Ex_1-learner but no well-behaved Ex-learner.*

Proof. Consider the class R containing all functions f satisfying one of the following conditions.

- $f = \sigma 0^\infty$ for some $\sigma \in \{1,2\}^*$;
- $f = \varphi_e$ and $f \in \{1^e 2\} \cdot \{1,2\}^\infty$ for some $e \in \mathbb{N}$.

R has no well-behaved Ex-learner. For given well-behaved M and number e, construct the following function f_e:

$$f_e(x) = \begin{cases} 1 & \text{if } x < e \text{ or } (x > e \text{ and } M(f_e[x]) = M(f_e[x]2)); \\ 2 & \text{if } x = e \text{ or } (x > e \text{ and } M(f_e[x]) \neq M(f_e[x]2)). \end{cases}$$

Assume now that $x > e$ and $M(f_e[x])$ is the hypothesis $\tilde{e}$. By condition (c) of the definition of a well-behaved Ex-learner, there is at most one $a \in \{1,2\}$ such that $M(f_e[x]a)$ outputs $\tilde{e}$. If $a = 2$ then $f_e(x) = 1$ else $f_e(x) = 2$. So $M(f_e[x+1]) \neq \tilde{e}$. Thus M_e does not converge to a hypothesis on any of the functions f_e. However, by the Fixed-Point Theorem [13, Theorem II.2.10], there is an e such that $f_e = \varphi_e$. Since $f_e \in \{1^e 2\} \cdot \{1,2\}^\infty$, it follows that $f_e \in R$ and M does not Ex-learn R. So R does not have a well-behaved learner.

There is a pessimistically reflective exact Ex_1-learner N for R. On input σ, N behaves as follows. If $\sigma \in \{1^e 2\} \cdot \{1,2\}^*$ and $\varphi_e(x)\!\downarrow = 1$ for $x < e$

and $\varphi_e(e)\downarrow = 2$ within $|\sigma|$ computation-steps, then $N(\sigma) = e'$ where

$$\varphi_{e'}(x) = \begin{cases} \varphi_e(x) & \text{if } \varphi_e(y)\downarrow \in \{1,2\} \text{ for all } y \le x; \\ \uparrow & \text{otherwise.} \end{cases}$$

If $\sigma = \tau 0^k$ for a $k > 0$ and a $\tau \in \{1,2\}^*$, then $N(\sigma)$ is a canonical index for $\tau 0^\infty$. In all other cases, $N(\sigma) = ?$.

It is easy to verify that N Ex_1-identifies R and that all indices output by N are either for functions in R or for non-total functions. Furthermore, N outputs a hypothesis only on σ of the form $\{1,2\}^*$ or $\{1,2\}^* \cdot \{0\}^*$ all of which are extended by functions in R. So N is an exact pessimistically reflective learner for R. ▮

Theorem 3.5. *Every class having an exact pessimistically reflective Ex-learner has also an exact prudent learner, but the converse does not hold.*

Proof. Implication. Consider a class S having an exact and pessimistically reflective Ex-learner M. The Padding Lemma [13, Proposition II.1.6] states that, for every index e, one can effectively find infinitely many equivalent indices (that is, indices computing the same function φ_e). Thus one can assume without loss of generality that M never returns to an abandoned index e (if M needs to reconsider the function φ_e, it can output an equivalent index not used earlier). Thus, if M outputs on a function f an index e infinitely often, then M converges on f to e.

Now assign to every e the index e' such that $\varphi_{e'}(x) = y$ iff there is a $z > x$ such that $0, 1, \ldots, z \in dom(\varphi_e)$, $\varphi_e(x) = y$ and $M(\varphi_e[z]) = e$; otherwise $\varphi_{e'}(x)$ is undefined.

Now one transforms the pessimistically reflective learner M into a prudent learner N by replacing all hypotheses e of M by the corresponding e'. The new learner has the following properties.

- If M Ex-identifies f by converging to the index e, then $\varphi_e = f$, and M converges on φ_e to e. Thus, by definition of e', $\varphi_{e'} = \varphi_e$. Thus N also Ex-identifies f and is an Ex-learner for S.
- If N outputs e' on some input and $\varphi_{e'}$ is a total function f, then M infinitely often outputs e on f. By the assumption on M, M converges on f to e, that is, M Ex-identifies f. Since M is exact, $f \in S$.
- If N outputs e' on some input and $\varphi_{e'}$ is partial, then there is some σ extending $\varphi_{e'}$ with $M(\sigma) = e$. It follows that there is a function $f \in S$ which extends σ and thus $\varphi_{e'}$.

So N is a prudent Ex-learner for S. Furthermore, all total functions computed by some output of N are in S as shown above. It follows that N is exact.

Separation. The following class R' witnesses that the converse direction fails and the implication is proper. R' is obtained by modifying R from Theorem 3.4, by making the first condition more restrictive. R' contains the functions f satisfying one of the following conditions.

- $f = 1^e 2\sigma 0^\infty$ and $1^e 2\sigma \subseteq \varphi_e$ for some $\sigma \in \{1,2\}^*$;
- $f = \varphi_e$ and $f \in \{1^e 2\} \cdot \{1,2\}^\infty$ for some $e \in \mathbb{N}$.

R$'$ has no exact pessimistically reflective Ex-learner. Consider the set $E = \{e : \varphi_e$ is total and $\{1,2\}$-valued and extends $1^e 2\}$. The set E is Π^0_2 complete and thus not K-recursive. But if there were a pessimistically reflective learner M for R', then M would satisfy the following conditions.

- If $e \in E$ then there is a hypothesis $\tilde{e}$ such that, for almost all s, there is $\sigma \in \{1^e 2\} \cdot \{1,2\}^s$ with $M(\sigma) = \tilde{e}$.
- If $e \notin E$ then, for almost all s and all $\sigma \in \{1^e 2\} \cdot \{1,2\}^s$, $M(\sigma) = ?$.

This would give that E is recursive in the limit, a contradiction.

There is an exact prudent Ex$_1$-learner N for R$'$. On input σ, N behaves as follows. If $\sigma \in \{1^e 2\} \cdot \{1,2\}^*$ and $\varphi_e(x)\downarrow = 1$ for all $x < e$ and $\varphi_e(e)\downarrow = 2$ within $|\sigma|$ computation-steps, then $N(\sigma) = e'$ where

$$\varphi_{e'}(x) = \begin{cases} \varphi_e(x) & \text{if } \varphi_e(y)\downarrow \in \{1,2\} \text{ for all } y \leq x; \\ \uparrow & \text{otherwise.} \end{cases}$$

If $\sigma = \tau 0^k$ for $k > 0$, $e \geq 0$ and $\tau \in \{1^e 2\} \cdot \{1,2\}^*$ and if it can be verified in k computation steps that φ_e extends τ, then $N(\sigma)$ is a canonical index for $\tau 0^\infty$. In all other cases, $N(\sigma) = ?$.

It is easy to verify that N is an exact Ex-learner for R'. Furthermore, every non-total, partial function conjectured by N is of the form $\varphi_{e'}$ where e' derives from some e as defined above. Then $\varphi_{e'}$ is a finite function such that $\varphi_{e'}$ extends $1^e 2$ and is extended by $\varphi_{e'} 0^\infty$ which is in R'. It follows that N is a prudent Ex-learner for R'. ∎

4 Easier Learning of Infinite Subclasses

Recall the question considered in the introduction: Does every infinite Ex-learnable class have an infinite finitely learnable subclass? In this section, we study this and similar questions for confident and well-behaved learners.

We start by giving an infinite class learnable by a well-behaved learner, which doesn't have an infinite subclass in Num.

Theorem 4.1. *There is an infinite class S, which is Ex-identifiable by a well-behaved learner, such that for every R in Num the intersection $S \cap R$ is a finite class.*

Proof. The basic idea of this proof is to construct a class $S = \{\Psi_0, \Psi_1, \ldots\}$ of total functions with the following properties:

- There is an enumeration Θ of partial-recursive functions containing the functions $\Psi_0, \Psi_1, \ldots$ and some finite functions such that the uniform prefix-closed graph of Θ is recursive. This permits to adapt the technique of learning by enumeration adequately and to guarantee properties (a) and (c) of the definition of well-behaved learners.
- S is dense. Since Θ contains only finite functions and the total functions $\Psi_0, \Psi_1, \ldots$, property (b) of well-behaved learners, that is prudence, will be satisfied.
- Ψ_e dominates all total complexity measures Φ_d with $d \leq e$. Thus every recursive function can only dominate finitely many Ψ_e and therefore every uniformly recursive class can only contain finitely many functions from S.

Now the construction in detail: Let $\sigma_0, \sigma_1, \ldots$ be an enumeration of all strings. Let $\Phi_0, \Phi_1, \ldots$ be the step counting functions associated with $\varphi_0, \varphi_1, \ldots$ such that $\Phi_e(x)$ is the number of steps needed to compute $\varphi_e(x)$, if $\varphi_e(x)$ is defined, and $\Phi_e(x) = \infty$ otherwise. Now define for every e the value a_e as

$$a_e = \min(\{\infty\} \cup \{x : \Phi_e(x) = \infty \vee (\exists y < x)\,[\Phi_e(x) < \Phi_e(y)]\,\}\,).$$

The a_e's can be approximated from below; that is, there is a total recursive mapping $e, s \to a_{e,s}$ such that $a_e = \lim_s a_{e,s}$, and $a_{e,s} \leq a_{e,s+1}$, for all e, s. Note that one can, without loss of generality, have that $a_{e,s} \leq s$ and thus the approximation never takes the value ∞. Now let

$$\Psi_e(x) = \begin{cases} \sigma_e(x) & \text{if } x \in \text{domain}(\sigma_e); \\ \max(\{0\} \cup \{\Phi_d(y) : d \leq e \wedge & \\ \qquad y < \min(\{1+x, a_d\})\,\}\,) & \text{otherwise.} \end{cases}$$

We cannot recursively know the values $a_0, a_1, \ldots$ but can only approximate them in the limit. So we consider the following enumeration of partial functions containing all the Ψ_e. For each tuple $(b_0, b_1, \ldots, b_e) \in (\mathbb{N} \cup \{\infty\})^*$, let

$$\Theta_{(b_0,b_1,\ldots,b_e)}(x) = \begin{cases} \sigma_e(x) & \text{if } x \in \text{domain}(\sigma_e); \\ \max(B \cup \{0\}) & \text{if the following conditions hold:} \\ & \text{(I) } B = \{\Phi_d(y) : d \leq e \wedge \\ & y < \min(\{1+x, b_d\})\,\} \text{ exists and} \\ & \text{can be completely enumerated,} \\ & \text{(II) } a_{d,x} \leq b_d \text{ for all } d \leq e, \\ & \text{(III) } \Phi_d(y) \leq \Phi_d(y+1)\text{, for all} \\ & y < \min(\{1+x, b_d\}) - 1 \text{ and } d \leq e, \\ & \text{(IV) } x \notin \text{domain}(\sigma_e); \\ \uparrow & \text{otherwise.} \end{cases}$$

Note that in (I) above, B exists if $b_d \leq a_d$, for all $d \leq e$.

On the one hand, one can show that the set

$$\{(b_0, b_1, \ldots, b_e, x, y) : x < \infty \wedge y < \infty \wedge \Theta_{(b_0,b_1,\ldots,b_e)}(x) = y\}$$

is recursive. Therefore, there exists a learner M which consistently learns the class of all total $\Theta_{(b_0,b_1,\ldots,b_e)}$, where M outputs only hypotheses for functions of the form $\Theta_{(b_0,b_1,\ldots,b_e)}$. As $\Theta_{(b_0,b_1,\ldots,b_e)}$ is total iff $a_0 = b_0 \wedge a_1 = b_1 \wedge \ldots \wedge a_e = b_e$, it follows that the total functions in this list are exactly the functions Ψ_e and so M is a consistent learner for $S = \{\Psi_0, \Psi_1, \ldots\}$. In particular, M satisfies conditions (a) and (c) in Definition 3.1 of well-behaved learner.

Furthermore, if some $b_k \neq a_k$ for $k \leq e$, then $\Theta_{(b_0,b_1,\ldots,b_e)}$ is equal to a finite string $\sigma_{e'}$ and the function $\Psi_{e'}$ extends $\sigma_{e'}$. As all indices output by M are indices for functions of form $\Theta_{(b_0,b_1,\ldots,b_e)}$, one can conclude that condition (b) in Definition 3.1 of well-behaved learner is also satisfied.

On the other hand, if $f_0, f_1, \ldots$ is a recursive enumeration of total functions, then the function g given by

$$g(x) = f_0(x) + f_1(x) + \ldots + f_x(x) + 1$$

dominates all these functions and there is a total and ascending function Φ_e dominating g. It follows that the functions $\Psi_e, \Psi_{e+1}, \ldots$ are different from all functions $f_0, f_1, \ldots$ and so the intersection of S and any class in Num is finite. ∎

An essential ingredient of the above proof is that one cannot bound the number of mind changes made by the well-behaved learner. In the extreme case that one does not permit any mind changes, Proposition 3.3 gives a different outcome.

Recall from Definition 2.5 that a learner M is *confident* iff M always converges on input function, that is,

$$(\forall f)\,(\forall^\infty n)\,[M(f[n+1]) = M(f[n])].$$

So a confident learner converges on every input function, even if this function is not recursive and therefore cannot be learned at all. Note that any class which can learned with a bound (whether constant bound or ordinal bound) on the number of mind changes can also be learned by a confident learner.

The next result shows that some infinite learnable classes do not have infinite confidently learnable subclasses.

Theorem 4.2. *There is an infinite uniformly recursive class GEN such that intersection of GEN with any confidently learnable class is finite.*

Proof. Recall that a 1-generic set G has the following property: for every recursive set U of strings there is a k such that either the string $G(0)G(1)\ldots G(k)$ itself is in U or no extension of $G(0)G(1)\ldots G(k)$ is in U. One can choose G such that G is Turing reducible to K [14, Section XI.2]. Therefore, there is a recursive enumeration $f_0, f_1, \ldots$ of $\{0,1\}$-valued recursive functions pointwise converging to (the characteristic function of) the set G. Let $GEN = \{f_0, f_1, \ldots\}$ for these functions. As G is not recursive and differs from every function f_k, the set GEN is infinite.

Now consider any class S having a confident learner M. By confidence, M converges on G. Thus there exists a $\sigma \subseteq G$ such that $M(\eta) = M(\sigma)$ whenever $\sigma \subseteq \eta \subseteq G$. As G is 1-generic and as G does not contain any string of the recursive set $\{\eta : \eta \supseteq \sigma \wedge M(\eta) \neq M(\sigma)\}$, there is a τ satisfying: $\sigma \subseteq \tau \subseteq G$ and $M(\eta) = M(\sigma)$ for all $\eta \supseteq \tau$. Furthermore, using the nonrecursiveness of G, one may assume that τ is so long that the hypothesis $M(\sigma)$ does not compute an extension of τ.

As the functions f_k approximate the set G and $\tau \subseteq G$, almost all f_k extend τ. Thus the set $\{f_k : \tau \not\subseteq f_k\}$ is finite and also contains all functions in the intersection of S and $\{f_0, f_1, \ldots\}$. The theorem follows. ∎

As all uniformly recursive classes have a well-behaved learner, the following corollary is immediate.

Corollary 4.3. *There is an infinite class R having a well-behaved Ex-learner such that $R \cap S$ is finite for every confidently learnable class S.*

Theorem 4.4. *If an infinite class S has a confident and well-behaved learner, then S has an infinite uniformly recursive subclass U which is Ex_0-identifiable.*

Proof. Let M be a confident and well-behaved learner for S such that $M(\lambda)$ outputs a hypothesis for the everywhere undefined function. Now consider the tree $T \subseteq \mathbb{N}^*$, with root λ, defined as follows. A node σ of T has as successors all the nodes $\tau \supset \sigma$ such that M outputs at τ for the first time a hypothesis different from $M(\sigma)$; that is, (I) $M(\tau) \notin \{M(\sigma), ?\}$ and (II) $M(\eta) \in \{M(\sigma), ?\}$ for all η with $\sigma \subseteq \eta \subset \tau$. An invariant of this construction is that M never outputs ? on the nodes of T. The tree T is well-founded as M converges on all functions, that is, the tree does not have infinite branches. By König's Lemma, T would be finite if T is finitely branching. As S is infinite, T must be infinite. So there is a node $\sigma \in T$ having infinitely many successors and there is a recursive enumeration $\tau_0, \tau_1, \ldots$ producing them. The subclass U is generated from these τ_k as follows.

The function f_k is the limit of strings η_l, where $\eta_0 = \tau_k$ and η_{l+1} is the first string found (in some standard search) such that $\eta_l \subset \eta_{l+1}$ and $M(\eta_{l+1}) \neq ?$.

To see that all f_k are total, assume by way of contradiction that for some f_k, the process terminates at some η_l. Then it would hold that $(\forall \tau \supset \eta_l)\,[M(\tau) = ?]$ and M would not Ex-identify any extension of η_l. However $M(\eta_l)$, by condition (c) in Definition 3.1, computes a partial function extending η_l and, by condition (b), some total extension of $\varphi_{M(\eta_l)}$ (which is also a total extension of η_l) is in S. A contradiction. Thus each f_k is total.

The definition of f_k ensures that M outputs on f_k infinitely often a hypothesis. As M is confident, M converges on f_k to a hypothesis e. The consistency condition (c) from Definition 3.1 implies that φ_e extends infinitely many $\sigma \subseteq f_k$ and so $\varphi_e = f_k$. As φ_e is total, $\varphi_e \in S$ and thus $\{f_0, f_1, \ldots\} \subseteq S$.

An exact Ex_0-learner for $\{f_0, f_1, \ldots\}$ can be built as follows: on input σ, the learner outputs a hypothesis e_k for f_k whenever $\tau_k \subseteq \sigma \subseteq f_k$ for some k. Otherwise the learner outputs ?. ∎

We now consider results that deal with the question when Ex_α-identifiable classes have infinite Ex_β-identifiable subclasses for $\beta < \alpha$. For this, we need the following two results from Freivalds and Smith [4].

Freivalds and Smith [4, Theorem 6] showed that classes of step functions like the ones below separate the various levels of the hierarchy for learning with an ordinal bound on the number of mind changes.

Proposition 4.5 [4]. *For every ordinal α represented by an element r_α with respect to a suitable notation $<_e$ of ordinals, define the class $DEC_{\alpha,e}$ to be the set of all decreasing functions $f : \mathbb{N} \to \mathbb{N}$ with $f(0) \leq_e r_\alpha$ and $(\forall x)\,[f(x+1) \leq_e f(x)]$. Then $DEC_{\alpha,e}$ is Ex_α-identifiable. However, there is no $\beta < \alpha$ such that some, even not necessarily recursive, learner M Ex_β-identifies $DEC_{\alpha,e}$.*

Proof. $DEC_{\alpha,e}$ contains only functions which are decreasing with respect to a well-ordering. So they can properly decrease only finitely often and are thus eventually constant. So the class $DEC_{\alpha,e}$ consists of recursive functions.

$DEC_{\alpha,e}$ has an Ex_α-learner M defined as follows. On input λ, $M(\lambda) = ?$ and the ordinal is initialized as r_α. On input $y_0 y_1 \ldots y_n$ with $r_\alpha \geq y_0 \geq_e y_1 \geq_e \ldots \geq_e y_n$ let m be the minimal number with $y_m = y_n$. Then M outputs the canonical index for $y_0 y_1 \ldots y_m (y_m)^\infty$ and the value of the ordinal counter is y_m. In particular for $m > 1$, the counter is counted down iff $m = n$. On all other inputs, M outputs ? and does not change its ordinal counter.

Now we show that there is no Ex_β-learner for $DEC_{\alpha,e}$ as follows. Suppose by way of contradiction that there exists such a learner N with ordinal counter ord using some notation $<_{e'}$. Define that $y <' z$ if the ordinal represented by y with respect to $<_e$ is below that represented by z with respect to $<_{e'}$, similarly define $y =' z$ and $y \leq' z$.

We construct a counterexample f to N being an Ex_β-learner for $DEC_{\alpha,e}$. In this construction, we use that without loss of generality, N updates its ordinal only if necessary, that is, N outputs a new hypothesis on some $f \in DEC_{\alpha,e}$ and there had already been a previous hypothesis. We now define the diagonalizing f inductively. Let $f(0) = y_0$ for some y_0 with $r_\beta <_e y_0 \leq_e r_\alpha$, where r_β represents the ordinal β. Assume that $f[x]$ is defined and $x > 0$. If there is a b such that

(I) For every y, z such that $y < z \leq x$ and $N(f[y]), N(f[z])$ are neither equal nor ?: $\mathrm{ord}(f[z]) <_{e'} \mathrm{ord}(f[y])$;
(II) $b =' \mathrm{ord}(f[x])$ and $b <_e f(x-1)$;
(III) $\varphi_{N(f[x])}$ extends $f[x]$ but does not extend $(f[x])b$;

then let $f(x) = b$ else let $f(x) = f(x-1)$.

It is easy to see that the resulting function f is total and in $DEC_{\alpha,e}$. Now we look at the behaviour of N on f assuming that N satisfies (I) on f.

Note that the above construction has the following invariant: the ordinal represented by $\mathrm{ord}(f[x])$ (in $<_{e'}$ notation) is not greater than the ordinal represented by $f(x)$ (in $<_e$ notation).

Let y be the least number with $f(z) = f(y)$ for all $z > y$ and x be the least number with $N(f[x])$ being the final hypothesis of N. Let b be the number with $b =' \mathrm{ord}(f[x])$.

If $y = 0$ then $N(f[x])$ is not a hypothesis for the function $(y_0)^\infty$ since otherwise (I), (II) and (III) would be satisfied as $y_0 >' \mathrm{ord}(f[x])$.

If $y > 0$ and $x \leq y$ then $N(f[x]) = N(f[y])$ and $\varphi_{N(f[x])}$ does not extend $f[y+1]$, so N does not learn f.

If $x > y > 0$ then $\mathrm{ord}(f[x]) <_{e'} \mathrm{ord}(f[y])$. It follows, using invariant stated above, that $b <_e f(x)$. As $f(x) \neq b$, (III) must be violated and whenever $\varphi_{N(f[x])}$ extends $f[x]$, it also extends $f[x]b$ and is thus different from f.

This case-distinction is complete and in all cases, N does not Ex_β-learn f. Thus N is not Ex_β-identifiable. ▮

Freivalds and Smith [4, Theorem 10] showed that $\bigcup_\alpha \mathrm{Ex}_\alpha$ is closed under union, where the number of mind changes needed to show the closure can go up. If one does not require the new learner to be recursive, one can get very tight bounds. Recall the definitions of $\oplus$ and $\otimes$ from Remark 2.8.

Proposition 4.6 [4]**.** *Given classes $S_1, S_2, \ldots, S_n$ such that each S_m is Ex_{β_m}-identifiable and given ordinal $\alpha = \beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_n \oplus (n \ominus 1)$, there is a (not necessarily recursive) Ex_α-learner N for the union $S_1 \cup S_2 \cup \ldots \cup S_n$.*

Proof. Assume that learners $M_1, M_2, \ldots, M_n$ for $S_1, S_2, \ldots, S_n$ with mind change bounds $\beta_1, \beta_2, \ldots, \beta_n$ are given. The new learner N starts with hypothesis ? and mind change counter $\beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_n \oplus (n \ominus 1)$. Furthermore, N has variables $\gamma_1, \ldots, \gamma_n$ such that each γ_m is initialized as $\beta_m \oplus 1$. On input $\sigma = \tau a$ with $\tau \in \mathbb{N}^*$ and $a \in \mathbb{N}$, N checks whether there is an $m \in \{1, 2, \ldots, n\}$ such that the following holds.

- The previous hypothesis $N(\tau)$ is either ? or inconsistent with the data seen so far;
- $e_m = M_m(\sigma)$ computes a total function φ_{e_m} extending σ, and the ordinal counter of M_m (after seeing σ) is strictly below the value of γ_m at τ.

If so, we let $N(\sigma) = e_m$ and γ_m is updated to the value of the ordinal counter of M_m after seeing σ. The other $\gamma_{m'}$ remain unchanged. The ordinal counter of N is set to the updated value of the expression $\gamma_1 \oplus \gamma_2 \oplus \ldots \oplus \gamma_n$.

Otherwise, $N(\sigma) = N(\tau)$ and the ordinal counter of N remains unchanged. The variables $\gamma_1, \ldots, \gamma_n$ also remain unchanged.

The verification is based on the following facts. The ordinal counter is initialized as $\gamma_1 \oplus \gamma_2 \oplus \ldots \oplus \gamma_n \ominus 1$. Whenever N outputs a new hypothesis, the value of $\gamma_1 \oplus \gamma_2 \oplus \ldots \oplus \gamma_n$ strictly decreases and is then copied into the ordinal counter of N. Whenever N makes a mind change, its ordinal counter is counted down. On every input $f \in S_1 \cup S_2 \cup \ldots \cup S_n$, N converges to an index e of a total function. Since N is not required to be recursive, it does not matter how N represents the ordinals.

If N converges on f to an e such that $\varphi_e \neq f$ then it holds for every m that either M_m does not Ex-identify f or m never qualifies in the search condition of N after M_m has converged to an index e_m of f. In this latter case, the ordinal counter of M_m and the variable γ_m must have the same value after M_m has converged to e_m. Since N never took the value e_m and since γ_m was initialized as $\beta_m \oplus 1$ while the counter of M_m was initialized as β_m, this can only happen because M_m did not count down its ordinal at some mind change. That is, M_m does not Ex_{β_m}-identify f. It follows that $f \notin S_m$. Thus, N is a (not necessarily recursive) Ex-learner for $S_1 \cup S_2 \cup \ldots \cup S_n$ with the ordinal bound $\beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_n \oplus (n \ominus 1)$ on the number of mind changes. ∎

We now give the promised result dealing with the question when Ex_α-identifiable classes have infinite Ex_β-identifiable subclasses for $\beta < \alpha$.

Theorem 4.7. *Fix a notation $<_e$ of ordinals used for all ordinal-learners considered below such that $\oplus$ is recursive and $\ominus$ is partial-recursive. Let $\alpha \geq 2$ be a recursive ordinal and consider all recursive learners, including those which are not exact. If $\alpha = \omega^\gamma$ for an ordinal γ*

Then there is an infinite exactly Ex_α-identifiable class S_α such that for every $\beta < \alpha$, S_α does not have an infinite Ex_β-identifiable subclass,
Else there is a $\beta < \alpha$ such that every infinite exactly Ex_α-identifiable class S has an infinite exactly Ex_β-identifiable subclass.

If one does not want to fix a notation of ordinals with the above property, then the same theorem holds, but the learners considered may no longer be exact.

Note that the case $\alpha = 1 = \omega^0$ is omitted as it is too sensitive to the definition of ordinal counters: if one would count hypotheses instead of mind changes and define that exactly the empty class can be learned with 0 hypotheses, then one could omit the condition "$\alpha \geq 2$" in Theorem 4.7.

Proof. Then-Case. Let e be such that $<_e$ is a notation for ordinals having a representative for α. Now one constructs $S_\alpha \subseteq DEC_{\alpha,e}$ as follows.

Let $M_1, M_2, \ldots$ be a list of all partial-recursive learners equipped with an ordinal mind change counter, using the notation given by $<_e$, such that the initial value of the counter, β_k, is strictly below α. Let U_k be the class of functions which at least one of the machines $M_1, M_2, \ldots, M_k$ infers without violating the mind change bound. There is a, not necessarily recursive, learner N_k identifying U_k exactly with mind change bound $\beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_k \oplus (k \ominus 1)$. Note that $\beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_k \oplus k \leq \alpha$. Thus there is a function $f_k = \sigma a^\infty \in DEC_{\alpha,e}$ such that $f_k(0)$ represents the ordinal $\beta_1 \oplus \beta_2 \oplus \ldots \oplus \beta_k \oplus k$ with respect to $<_e$ and f_k is not learned by N_k. In particular, f_k is not in U_k. Since $\oplus$ is recursive, the mapping $k \to f_k(0)$ is recursive, has a recursive range and is one-one.

Furthermore, one can find a program for one such $f_k \notin U_k$, effectively in the limit, from k. To see this, note that such a function $f_k = \sigma a^\infty$ satisfies the following for $l = 1, 2, \ldots, k$: There exist e, h, x (depending on l) such that either (I) $M_l(\sigma a^h)$ is undefined or (II) $M_l(\sigma a^h)$ has already made a mind change without counting down its ordinal or (III) the learner M_l converges to the wrong index e (that is $e = M_l(\sigma a^h)$, M_l does not change its mind on σa^∞ beyond σa^h and, for some x, $\varphi_e(x) \neq (\sigma a^\infty)(x)$). For each l, the above conditions on the k, σ, a, h, e, x are K-recursive. Thus, from k, one can compute in the limit one such (σ, a), and thus a program for one such f_k.

We now show that the class $S_\alpha = \{f_1, f_2, \ldots\}$ can be exactly Ex_α-identified. Given an Ex_α-learner M for $DEC_{\alpha,e}$, one defines an exact identifier N as follows: If $M(\sigma) = ?$ or $\sigma = \lambda$, then $N(\sigma) = ?$. Else N computes the k such that $f_k(0) = \sigma(0)$. If such a k does not exist, then $N(\sigma) = ?$ as well. If the k is found, then N considers a uniform approximation $f_{k,s}$ to f_k and outputs the following modification e' of the index $e = M(\sigma)$: $\varphi_{e'}(x) = \varphi_e(x)$ iff there is $s \geq x$ such that $\varphi_e(y)\downarrow = f_{k,s}(y)$ for all $y \leq x$. If there is no such s, then $\varphi_{e'}(x)$ is undefined. The convergence behaviour of M and N is the same. However, N converges to an index of f iff M also does and $f \in S_\alpha$ — otherwise, N converges to an index of a partial function or to ?.

If one does not require exact learning, and considers the extension mentioned in the theorem: one can use the learner M for the whole class $DEC_{\alpha,e}$ instead of N and can therefore select the functions $f_k \in DEC_{\alpha,e} - U_k$ arbitrarily. This in particular permits to deal with a nonrecursive $\oplus$ and the case that representation for the ordinal counter of the M_k might depend on each k.

Else-Case. The ordinal α can be represented as $c\omega^\gamma + \delta$ for some ordinal γ with $c > 0$ and $\omega^\gamma > \delta$. If $\delta = 0$, then let $\beta = (c-1)\omega^\gamma$; else let $\beta = c\omega^\gamma$. Note that in both possible definitions it holds that $\beta < \alpha \leq \beta + \beta$ (when $\beta = (c-1)\omega^\gamma$, we implicitly have $c > 1$ by the condition that $\alpha \neq \omega^\gamma$). Let M be an Ex_α-learner for a given class S and ord be its ordinal counter. Let U be the set of all $f \in S$ such that $\mathrm{ord}(f[x]) \geq \beta$ for all x. Now consider the following two subcases.

Subcase U finite. We define the following Ex_β-learner N for the whole class S and the associated ordinal counter ord' as follows:

- If $\mathrm{ord}(\sigma) \geq_e \beta$, then $\mathrm{ord}'(\sigma) = \beta$. Furthermore, if exactly one function in U is consistent with the input σ, then N outputs an index for this function; otherwise N outputs ?.
- If $\mathrm{ord}(\sigma) <_e \beta$, then $\mathrm{ord}'(\sigma) = \mathrm{ord}(\sigma)$ and $N(\sigma) = M(\sigma)$.

It is easy to see that N Ex_β-identifies all the functions in U, as well as all the functions in S on which the ordinal counter of M eventually goes below β. Thus N (exactly) Ex_β-identifies the whole class S.

Subcase U infinite. In this case we define the learner N with ordinal counter ord' as follows.

- If $\mathrm{ord}(\sigma) \geq \beta$ then $N(\sigma) = M(\sigma)$ and $\mathrm{ord}'(\sigma) = \mathrm{ord}(\sigma) \ominus \beta$;
- If $\mathrm{ord}(\sigma) < \beta$ then $N(\sigma) = ?$ and $\mathrm{ord}'(\sigma) = 0$.

Note that due to the special form of β, $\delta \ominus \beta$ is defined for all δ with $\beta \leq \delta \leq \alpha$. It is easy to see that N exactly Ex_β-identifies U.

It remains to consider the case where one does not require that the learner is exact and one wants to deal with orders not having recursive operations $\oplus, \ominus$. In this case, one takes the original learner M for S which of course also

Ex-identifies the subclass $U \subseteq S$. But one adjusts the mind change counter to the following ord$'$. Let r_β be the representative of β with respect to $<_e$ and let $<_{e'}$ be such that whenever r' represents $\beta + \epsilon$ with respect to $<_e$, then r' represents ϵ with respect to $<_{e'}$. The ordinal counter ord$'$ is defined as follows.

If $\mathrm{ord}(\sigma) \geq_e r_\beta$ then $\mathrm{ord}(\sigma)$ represents some ordinal $\beta + \epsilon$ with respect to $<_e$. Now $\mathrm{ord}'(\sigma) = \mathrm{ord}(\sigma)$ and represents the ordinal ϵ with respect to $<_{e'}$.

Otherwise $\mathrm{ord}(\sigma) <_e r_\beta$ and the data is from a function not in U. Then let $\mathrm{ord}'(\sigma) = r_\beta$ (note that r_β represents 0 with respect to $<_{e'}$).

As a consequence, M is an Ex_β-learner for the infinite class U using the properties that ord$'$ starts with an ordinal less than or equal to β with respect to the notation $<_{e'}$ and that the Ex-learning capabilities remains the same. Furthermore, as long as the data is from functions in U, each mind change is accompanied by counting down the ordinal.

This completes the proof for the second (Else) part of the theorem. ∎

Note that, in the above Theorem, in Then case, one cannot have that S_α has a well-behaved Ex_α-learner. Otherwise, by Theorem 4.4, S_α would have an infinite Ex_0-identifiable subclass.

5 Sublearners

The main question considered in this section is the following: Given an Ex-identifiable class S satisfying some additional constraints, is there an infinite subclass U and an Ex-learner M for S such that M Ex_β-identifies U? One additional constraint is that S has an infinite Ex_0-identifiable subclass. As confidently identifiable classes are Ex_α-identifiable for some α, Theorem 4.2 has been adapted into this section as follows. There is a class $S = GEN \cup \{g_0, g_1, \ldots\}$, where GEN is from Theorem 4.2, such that $\{g_0, g_1, \ldots\}$ is Ex_0-identifiable, S is Ex-identifiable and no Ex-learner M for S is at the same time an Ex_α-learner for an infinite subclass of S.

Theorem 5.1. *There exists an infinite class S such that*

(a) *S is exactly Ex-identifiable;*
(b) *S contains an infinite exactly Ex_0-identifiable subclass;*
(c) *For any learner M which Ex-identifies S and for any α, M does not Ex_α-identify an infinite subclass of S.*

Proof. Let G and $f_0, f_1, \ldots$ be as in the proof of Theorem 4.2. Furthermore, let $g_k = f_k(0)f_k(1)\ldots f_k(k)2^\infty$, that is, g_k coincides with f_k on $0, 1, \ldots, k$ and takes the constant 2 from then on. Let $S = \{f_0, g_0, f_1, g_1, f_2, g_2, \ldots\}$. The class

S is clearly a uniformly recursive class. Thus S is exactly Ex-identifiable. Furthermore the subclass $\{g_0, g_1, \ldots\}$ is exactly Ex_0-identifiable since the function g_k is the unique one in this enumeration where $k+1$ is the first element to be mapped to 2.

Now consider any Ex-learner M for S equipped with an ordinal counter. As M learns all functions f_k, it follows from the proof of Theorem 4.2 that M makes on the characteristic function of G infinitely many mind changes. Thus there is a number l such that M has made a mind change on the input $G(0)G(1)\ldots G(l)$ without counting down the ordinal. Since almost all functions f_k and g_k extend the string $G(0)G(1)\ldots G(l)$, M can Ex_α-identify only finitely many functions in S. ▮

Theorem 5.2. *For every infinite class S having a confident and well-behaved learner M, there is a class U and a learner N such that*

- *$U \subseteq S$, U is infinite and U is uniformly recursive;*
- *N is an Ex_1-learner for U;*
- *N is a confident and well-behaved learner for S.*

Proof. This is a generalization of the proof of Theorem 4.4. In the proof of Theorem 4.4, we defined strings σ and $\tau_0, \tau_1, \ldots$ and functions $f_0, f_1, \ldots \in S$ with the following properties.

(I) The τ_k's are recursively enumerable and pairwise incomparable.
(II) For any k, $\sigma \subseteq \tau_k$ and $M(\tau_k) \notin \{M(\sigma), ?\}$. Furthermore, for all k and all η with $\sigma \subset \eta \subset \tau_k$, $M(\eta) \in \{M(\sigma), ?\}$.
(III) For all k, f_k extends τ_k and belongs to S. Furthermore, there is a program p_k for f_k which can be obtained effectively from k.
(IV) For all τ, if $\sigma \subseteq \tau$ and $M(\tau) \notin \{M(\sigma), ?\}$, then there exists a k such that $\tau_k \subseteq \tau$.

We now define our learner N as follows.

$$N(\tau) = \begin{cases} ?, & \text{if } \tau \subset \sigma; \\ p_k, & \text{for the unique } k \text{ with } \tau_k \subseteq \tau \subseteq f_k \text{, if there is such a } k; \\ M(\tau), & \text{otherwise.} \end{cases}$$

We argue that the second clause above can be recursively decided. Note that the τ_k are the places after σ where M outputs its first hypothesis not in $\{M(\sigma), ?\}$. Also the τ_k and f_k have both an effective enumeration. Thus, we can determine effectively from τ, whether there exists a k (and find such a k if it exists) such that $\tau_k \subseteq \tau$, and then use this k to check whether the data seen so far is consistent with f_k. It is now easy to verify that N Ex_1-identifies each f_k — N only outputs $M(\sigma)$ and then p_k on f_k; it is easy to assign the corresponding ordinal counter to N. Furthermore, if the input is incomparable to any f_k, then N follows M. Thus, N inherits the property of being a well-behaved and confident learner for S from M. ∎

Note that, in the above theorem, we are not able to achieve Ex_0 instead of Ex_1, as shown by following example.

Example 5.3. Consider the class $S = \{0^\infty\} \cup \{0^h 1 0^\infty : h \in \mathbb{N}\}$.

- There is a well behaved Ex_1-learner for S;
- No learner which Ex-identifies S, can Ex_0-identify an infinite subclass of S.

The existence of the well-behaved Ex_1-learner is easy to verify. On the other hand, any Ex-learner for S has to identify 0^∞ and outputs an index for it on input of the form 0^k for some k. Then it can Ex_0-identify only the finite subclass $\{0^\infty, 10^\infty, 010^\infty, \ldots, 0^{k-1}10^\infty\}$.

We now consider the question: Does there exist a class R_α which is Ex_α-identifiable, R_α contains an infinite finitely learnable subclass, but no learner can simultaneously Ex-identify R_α and Ex_β-identify an infinite subset of Ex_α, for $\beta < \alpha$.

The answer to the above question depends on α.

Theorem 5.4. *Fix a notation $<_e$ of ordinals used for all ordinal-learners considered below such that operation $\oplus$ is recursive and $\ominus$ partial-recursive. Let $\alpha \geq 2$ be a recursive ordinal and consider all recursive learners, including those which are not exact. If $\alpha = \omega^\gamma$ for an ordinal γ*

Then there is an infinite exactly Ex_α-identifiable class R_α such that (I) *R_α contains an infinite Ex_0-identifiable subclass and* (II) *for all $\beta < \alpha$, there does not exist an Ex-learner M for R_α which Ex_β-sublearns an infinite subclass of R_α using the notation $<_e$.*

Else there is a $\beta < \alpha$ such that every infinite exactly Ex_α-identifiable class S has an exact Ex-learner M for S which Ex_β-sublearns an infinite subclass of S.

If one does not want to fix a notation of ordinals with the above property, then the same theorem holds, but the learners considered may no longer be exact.

Proof. Then-Case. Assume that $\alpha = \omega^\gamma$ for some γ. The set R_α is defined as the union of two sets $\{f_1, f_2, \ldots\}$ and $\{g_1, g_2, \ldots\}$ where the functions f_k are exactly as in Theorem 4.7. For each function f_k, there is a number $a_k \geq 2$ such that for all $M \in \{M_1, M_2, \ldots, M_k\}$, whenever M makes a mind change on f_k without counting down the ordinal, then this happens before seeing all the data $f_k[a_k]$. Without loss of generality suppose 0 also represents the ordinal 0. The function g_k is taken to be $f_k[a_k+1]0^\infty$.

R_α is clearly infinite. Furthermore, $R_\alpha \subseteq DEC_{\alpha,e}$ and one can compute the characteristic function of g_k from the one of f_k using the oracle K. Thus one can adapt the Ex_α-learner from Theorem 4.7 to an Ex_α-learner N for R_α.

One can construct an exact Ex_0 learner for $\{g_0, g_1, \ldots\}$ as follows. If the input is not of form $\tau 0^r$ for some $\tau \in (\mathbb{N} - \{0\})^+$ then M outputs ?. Otherwise, M computes the k such that $g_k(0) = \sigma(0)$. If such a k does not exist, then $N(\sigma) = ?$ as well. If the k is found, N considers a uniform approximation $g_{k,s}$ to g_k and outputs the following modification e' of the index e for $\tau 0^\infty$: $\varphi_{e'}(x) = \varphi_e(x)$ iff there is $s \geq x$ such that $\varphi_e(y)\downarrow = g_{k,s}(y)$ for all $y \leq x$. If there is no such s, then $\varphi_{e'}(x)$ is undefined. It is now easy to verify that M is an exact Ex_0-learner for $\{g_0, g_1, \ldots\}$.

If M_k is an Ex-learner for R_α, then M_k is total and converges on all functions f_l to its correct index. By the construction in Theorem 4.7, M_k then fails for all f_l with $l \geq k$ to count down the ordinal at some mind change. Thus M_k does not Ex_{β_k}-learn the functions f_l, g_l with $l \geq k$. Thus no infinite subclass of R_α is Ex_β-sublearned for any $\beta < \alpha$.

Else-Case. This proof differs from the one in Theorem 4.7 only at one place: in the subcase that U is infinite and exact learners are desired, one defines that $N = M$ but changes ord to ord$'$ as done there. The reason for it is that this time N must be an exact Ex-learner for S while in Theorem 4.7 N must be an exact learner for U. All other parts of the proof remain unchanged. ∎

Remark 5.5. The negative results made use of the fact that the subclass has to be infinite. Indeed, dropping this constraint destroys all negative results. Given any finite subclass $U \subseteq S$ and any Ex-learner M for S, one can transform M into an Ex-learner N for S, such that N is also an Ex_0-learner for U: There is a number n such that M has converged on every $f \in U$ to the final index for f by the time it has seen $f[n]$. In particular, $M(f[m+1]) = M(f[m])$ for all $m \geq n$ and $f \in U$. The new learner N given by

$$N(\sigma) = \begin{cases} ? & \text{if } |\sigma| < n; \\ M(\sigma) & \text{if } |\sigma| \geq n; \end{cases}$$

has the desired properties: N Ex-identifies the same functions as M but on the functions $f \in U$, N only outputs the symbol ? before outputting the correct hypothesis $M(f[n])$.

Acknowledgement. The authors thank Gunter Grieser and Steffen Lange for discussions about the connections between well-behaved learners and reflective inductive inference. The authors are also grateful to Thomas Kent for proof-reading and comments. Further thanks go to the anonymous referees of the conference Algorithmic Learning Theory (ALT 2002) as well as this journal for suggestions and comments.

References

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

[2] Lenore Blum and Manuel Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

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

[4] Rūsiņš Freivalds and Carl Smith. On the role of procrastination in machine learning. *Information and Computation*, 107:237–271, 1993.

[5] Mark Fulk. Prudence and other conditions on formal language learning. *Information and Computation*, 85:1–11, 1990.

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

[7] Gunter Grieser. Reflective inductive inference machines and its improvement by therapy. *Algorithmic Learning Theory, 7th International Workshop, ALT 1996, Sydney, Australia, October 1996, Proceedings, Springer Lecture Notes in Artificial Intelligence*, 1160:325–336, 2001.

[8] Gunter Grieser. Reflective inductive inference of recursive functions. *Algorithmic Learning Theory, 13th International Conference, ALT 2002, Lübeck, Germany, November 2002, Proceedings, Springer Lecture Notes in Artificial Intelligence*, 2533:203–217, 2002.

[9] Paul R. Halmos. *Naive Set Theory.* Springer-Verlag, NY, 1994.

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

[11] Klaus-Peter Jantke. Reflecting and self-confident inductive inference machines. *Algorithmic Learning Theory, 6th International Workshop, ALT 1995, Fukuoka, Japan, October 1995, Proceedings, Springer Lecture Notes in Artificial Intelligence*, 997:282–297, 1995.

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

[13] Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland, Amsterdam, 1989.

[14] Piergiorgio Odifreddi. *Classical Recursion Theory*, volume II. Elsevier, Amsterdam, 1999.

[15] Daniel Osherson, Michael Stob and Scott Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists.* MIT Press, 1986.

[16] Robert Soare. *Recursively Enumerable Sets and Degrees.* Springer-Verlag, 1987.

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

[18] Thomas Zeugmann, Steffen Lange and Shyam Kapur. Characterizations of monotonic and dual monotonic language learning. *Information and Computation*, 120:155 - 173, 1995.

[19] Thomas Zeugmann and Steffen Lange. A guided tour across the boundaries of learning recursive languages. *Algorithmic Learning for Knowledge-Based Systems*, *Springer Lecture Notes in Artificial Intelligence*, 961:190–258, 1995.