Enlarging Learnable Classes

Sanjay Jain[1,⋆], Timo Kötzing[2,⋆⋆] and Frank Stephan[3,⋆⋆⋆]

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

[2] Max-Planck-Institut für Informatik, Campus E 1 4, 66123 Saarbrücken, Germany
koetzing@mpi-inf.mpg.de

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

Abstract. An early result in inductive inference shows that the class of **Ex**-learnable sets is *not* closed under unions. In this paper we are interested in the following question: For what classes of functions is the union with an arbitrary **Ex**-learnable class again **Ex**-learnable, either effectively (in an index for a learner of an **Ex**-learnable class) or non-effectively? We show that the effective case and the non-effective case separate, and we give a sufficient criterion for the effective case. Furthermore, we extend our notions to considering unions with classes of single functions, as well as to other learning criteria, such as finite learning and behaviorally correct learning.

Furthermore, we consider the possibility of (effectively) extending learners to learn (infinitely) more functions. It is known that all **Ex**-learners learning a *dense* set of functions can be effectively extended to learn infinitely more. It was open whether the learners learning a *non-dense* set of functions can be similarly extended. We show that this is *not* possible, but we give an alternative split of all possible learners into two sets such that, for each of the sets, all learners from that set can be effectively extended. We analyze similar concepts also for other learning criteria.

1 Introduction

One branch of inductive inference investigates the learnability of functions; the basic scenario given in the seminal paper by Gold [7] is as follows. Let $\mathcal{S}$ be a class of recursive functions; we say that $\mathcal{S}$ is *explanatorily learnable* iff there is a learner M which issues conjectures $e_0, e_1, \ldots$ with e_n being based on the data $f(0)f(1)\ldots f(n-1)$ such that, for all $f \in \mathcal{S}$, almost all of these conjectures are the same index e explaining f, that is, satisfying $\varphi_e = f$ with respect to an underlying numbering $\varphi_0, \varphi_1, \ldots$ of all partial recursive functions. In this paper,

⋆ Supported by NUS grants C252-000-087-001 and R252-000-420-112.

⋆⋆ Major parts of this paper were written when Timo Kötzing was visiting the Department of Computer Science at the National University of Singapore.

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

we consider learnability by partial recursive learners; with M_e we refer to the learner derived from the e-th partial recursive function.

During the course of time, several variants of this basic notion of explanatory learning (**Ex**) have been considered; most notably, *behaviorally correct learning* (**BC**) [1], in which the learner has to almost always output a correct index for the input function (these indices though are not constrained to be the same).

Another variant considered is *finite learning* (**Fin**) where the learner outputs a special symbol (?) until it makes one conjecture e which is never abandoned; this conjecture must of course be correct for a function to be learnt. Osherson, Stob and Weinstein [10] introduced a generalization of this notion, namely *confident learning* (**Conf**), where the learner can revise the hypothesis finitely often; it must, however, on each function f, even if it is not in the class to be learnt, eventually stabilize on one conjecture e. In inductive inference, one often only needs the weak version of this property where the convergence criterion only applies to recursive functions while the convergence behavior on non-recursive ones is not constrained (**WConf**, [14]).

Minicozzi [9] called a learner *reliable* iff the learner, on every function, either converges to a correct index or signals infinitely often that it does not find the index (by doing a mind change or outputting a question mark). One can combine the notion of reliability and confidence: A learner is *weakly reliable and confident* (**WConfRel**) iff the learner, for every recursive function f, either converges to an index e with $\varphi_e = f$ or almost always outputs ? (in order to signal non-convergence).

The above crtieria and the relations between them have been extensively studied, giving the following inclusion relations [2, 5–7, 9, 10, 14]:

- **Fin** $\subset$ **Conf** $\subset$ **WConf** $\subset$ **Ex** $\subset$ **BC**;
- **ConfRel** $\subset$ **WConfRel** $\subset$ **Rel** $\subset$ **Ex** $\subset$ **BC**;
- **Fin** $\not\subset$ **Rel** and **Rel** $\not\subset$ **WConf**.

Besides inclusion (learnability with respect to which criterion implies learnability with respect to another criterion), structural questions have also been studied: Is the union of two learnable classes learnable? Can one extend each learnable class?

Blum and Blum's *Non-Union Theorem* [2] (see also [1]) gave a quite strong answer to the first question: There are two classes $\mathcal{S}$ and $\mathcal{S}'$ of recursive functions such that each of them is learnable under the criterion **Ex** but their union is not learnable even under the more general criterion **BC**. Indeed, one can even learn the class $\mathcal{S}$ confidently and the class $\mathcal{S}'$ reliably. Thus, the Non-Union Theorem gives an interesting contrast to the fact that both confident learning and reliable learning are effectively closed under union.

Furthermore, it is interesting to ask how effective the union is. That is, if the union of two classes is learnable, can one effectively construct a learner for the union, given programs for the learners of the two given classes? The answer is "No" in general as can be seen directly by the proof of the Non-Union Theorem.

The confidently learnable class $\mathcal{S}$ above consists of all the functions f such that $f(0)$ is an index for f, and the class $\mathcal{S}'$ consists of all the functions f which are almost everywhere 0 (Blum and Blum [2] used slightly different classes $\mathcal{S}$ and $\mathcal{S}'$ which were $\{0,1\}$-valued; our $\mathcal{S}$ and $\mathcal{S}'$ makes the presentation simpler). Now consider the union of $\mathcal{S}'$ with a class $\mathcal{S}_e$, where $\mathcal{S}_e$ contains φ_e in the case that φ_e is total and $\varphi_e(0) = e$; otherwise $\mathcal{S}_e$ is empty. It is easy to show that, for each e, the class $\mathcal{S}_e \cup \mathcal{S}'$ is explanatory (**Ex**) learnable. If this union would be effective, giving rise to a learner $M_{h(e)}$ for the class $\mathcal{S}_e \cup \mathcal{S}'$, then one could make a learner N for $\mathcal{S} \cup \mathcal{S}'$ as follows: for non-empty sequences σ, $N(\sigma) = M_{h(\sigma(0))}(\sigma)$; a contradiction to the non-union theorem.

This example suggests to study four notions of when the unions of a given class $\mathcal{S}$ with another class is **Ex**-learnable:

1. $\mathcal{S}$ is (non-constructively) **Ex**-*unionable* iff for every **Ex**-learnable class $\mathcal{S}'$, the class $\mathcal{S} \cup \mathcal{S}'$ is **Ex**-learnable;
2. $\mathcal{S}$ is *constructively* **Ex**-*unionable* iff one can effectively convert every **Ex**-learner for a class $\mathcal{S}'$ into an **Ex**-learner for the class $\mathcal{S} \cup \mathcal{S}'$;
3. $\mathcal{S}$ is *singleton*-**Ex**-*unionable* iff for every total computable g, $\mathcal{S} \cup \{g\}$ is **Ex**-learnable.
4. $\mathcal{S}$ is *constructively singleton*-**Ex**-*unionable* iff there is a recursive function which assigns, to every index e, an **Ex**-learner for the class $\mathcal{S} \cup \{\varphi_e\}$ if φ_e is total and for the class $\mathcal{S}$ if φ_e is partial.

The same notions can also be defined for other learning criteria like finite, confident and behaviorally correct learning. We get the following results:

1. If a class $\mathcal{S}$ has a weakly confident learner then it is constructively singleton-**Ex**-unionable.
2. If a class $\mathcal{S}$ has a weakly confident and reliable learner then it is constructively **Ex**-unionable.
3. There is a class which is **Ex**-unionable and **BC**-unionable but does not satisfy any of the constructive unionability properties.
4. For finite learning, we show that unionability with classes and constructive union with singletons fails for all non-empty classes; only non-constructive unions with singletons is possible in the case that every pointwise limit of functions in the class is again in the class.

All our results for the cases of purely **Ex**-learning are summarized in Figure 1.

Forming the union with another class or adding a function are specific methods to enlarge a class. Thus, it is natural to ask when a learnable class of functions can be extended at all, without prescribing how to do this. Case and Fulk [4] addressed this question and showed, for the principal learning criteria **Ex** and **BC**, that one can extend learners to learn infinitely more functions whenever the learner satisfies a certain quality, say learns a dense class of functions. This enlargement can be done constructively (under this precondition). Furthermore, one can non-constructively extend any learnable class for many usual learning criteria like **Fin**, **Conf**, **Rel**, **ConfRel**, **WConf**, **WConfRel**, **Ex** and **BC**. Case and Fulk [4] left open two particular questions:

1. Is there a method to extend constructively every learner M_e which does not **Ex**-learn a dense class of functions?

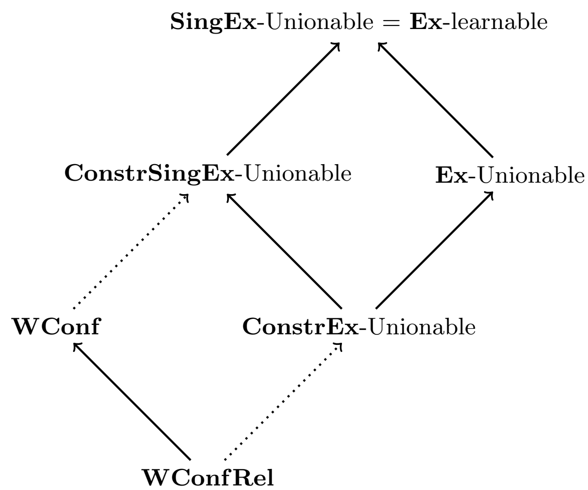

Fig. 1. The inclusion relations for the various unionability notions. It is unknown whether the dotted arrows might also go in the converse direction. All inclusions are given by arrows (and possibly reversed dotted arrows) and the concatenations of these.

2. How much nonconstructive information is needed in order to extend every learner M_e to learn infinitely many more functions? I.e., in how many classes does one have to split the learners so as to have constructive extension for each of the classes?

Theorem 25 answers the first question negatively – such a method does not exist.

On the other hand, the answer to the second question is that only a split into two classes is necessary. This result is not based on the information about whether the class is dense or not; instead it is based on the information about whether there exists a σ such that for no extension τ of σ: $M(\tau)\downarrow \neq M(\sigma)\downarrow$. In Theorem 27 we show that there is a recursive function h such that $\mathbf{Ex}(M_{h(e,b)})$ is a proper superclass of $\mathbf{Ex}(M_e)$ whenever either $b = 1$ and such a σ exists or $b = 0$ and such a σ does not exist.

2 Preliminaries

Let $\mathbb{N}$ denote the set of natural numbers. The symbols $\subseteq, \subset, \supseteq, \supset$ respectively denote subset, proper subset, superset and proper superset. For strings α and β, we let $\alpha \preceq \beta$ denote that α is a prefix of β. We let $\langle \cdot, \cdot \rangle$ denote a fixed computable pairing function from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$, which is increasing in both its arguments. We assume that $\langle 0, 0 \rangle = 0$.

Let φ denote a fixed acceptable programming system [12] for the class of all partial recursive functions. Let φ_i denote the i-th program in this programming system. Then, i is called the index or program for the partial recursive function

φ_i. Let $\mathcal{R}$ denote the set of all total recursive functions and $\mathcal{P}$ denote the set of all partial recursive functions. Let $\mathcal{R}_{0,1}$ denote the set of all total recursive functions f with range$(f) \subseteq \{0,1\}$. Let K denote the diagonal halting set $\{x : \varphi_x(x)\downarrow\}$. For a function η, let $\eta(x)\downarrow$ denote that $\eta(x)$ is defined, and $\eta(x)\uparrow$ denote that $\eta(x)$ is not defined. We let pad be a 1–1 recursive function such that, for all i, j, $\varphi_{\mathrm{pad}(i,j)} = \varphi_i$. Please find unexplained recursion theoretic notions in Rogers' book [12]. We let $\mathcal{S}$ range over sets of recursive functions.

Let σ, τ range over finite sequences. We often identify a total function with its sequence of values, $f(0)f(1)f(2)\ldots$; similarly for finite sequences. Let $f[n] = f(0)f(1)\ldots f(n-1)$. We use the notation $\sigma \preceq \tau$ to denote that σ is a prefix of τ (an initial subfunction of τ). Let Λ denote the empty sequence. Let $|\sigma|$ denote the length of σ. Let $\mathbb{S}$eq denote the set of all finite sequences.

Let $\sigma \cdot \tau$ denote concatenation of sequences, where σ is finite. When it is clear from context, we often drop $\cdot$ and just use $\sigma\tau$ for concatenation. For a finite sequence $\sigma \neq \Lambda$, let σ^- be σ with the last element dropped, that is, $\sigma^- \cdot \sigma(|\sigma|) = \sigma$. Let $[\mathcal{S}] = \{f[n] \mid f \in \mathcal{S}\}$. Thus, $[\mathcal{R}] = \mathbb{S}$eq. For notation simplification, $[f] = [\{f\}]$. A class $\mathcal{S}$ is said to be *dense* if $[\mathcal{S}] = [\mathcal{R}]$. A class $\mathcal{S}$ is *everywhere sparse* iff for all $\tau \in \mathbb{S}$eq, there exists a $\tau' \succeq \tau$ such that $\tau' \notin [\mathcal{S}]$. A total function f is an *accumulation point* of $\mathcal{S}$ iff there exist pairwise distinct functions $g_0, g_1, \ldots$ in $\mathcal{S}$ such that, for all $n \in \mathbb{N}$, $f[n] \preceq g_n$.

A *learner* is a partial-recursive mapping from finite sequences to $\mathbb{N} \cup \{?\}$. We let M, N and P range over learners and let $\mathcal{C}$ range over classes of learners. Let $M_0, M_1, \ldots$ denote an acceptable numbering of all the learners.

We say that M converges on function f to i (written: $M(f)\downarrow = i$) iff for all but finitely many n, $M(f[n]) = i$. If $M(f)\downarrow = i$ for some $i \in \mathbb{N}$, then we say that M converges on f (written: $M(f)\downarrow$). We say that $M(f)$ diverges (written: $M(f)\uparrow$) if $M(f)$ does not converge to any $i \in \mathbb{N}$. We now describe some of the learning criteria.

Definition 1. Suppose M is a learner and f is a total function.

(a) [7] We say that M **Ex**-*learns* f (written: $f \in \mathbf{Ex}(M)$) iff (i) for all s, $M(f[s])$ is defined, and (ii) there exists an i such that $\varphi_i = f$ and, for all but finitely many n, $M(f[n]) = i$.
(b) [1, 6] We say that M **BC**-*learns* f (written: $f \in \mathbf{BC}(M)$) iff, (i) for all s, $M(f[s])$ is defined, and (ii) for all but finitely many n, $\varphi_{M(f[n])} = f$.
(c) [1, 6] We say that M **Fin**-*learns* f (written: $f \in \mathbf{Fin}(M)$) iff (i) for all s, $M(f[s])$ is defined, and (ii) there exist n and i such that $\varphi_i = f$, for all $m < n$, $M(f[n]) =?$, and for all $m \geq n$, $M(f[n]) = i$.
(d) [6] We say that M $\mathbf{Ex}_n$-*learns* f (written: $f \in \mathbf{Ex}_n(M)$) iff (i) M **Ex**-learns f and (ii) card$(\{m \mid ? \neq M(f[m]) \neq M(f[m+1])\}) \leq n$.

We say that M makes a *mind change* at $f[m+1]$ if $? \neq M(f[m]) \neq M(f[m+1])$.

Definition 2. Let I be a learning criterion (defined above or later in this paper):

(a) We say that M I-learns $\mathcal{S}$ (written: $\mathcal{S} \subseteq I(M)$) iff M I-learns each $f \in \mathcal{S}$.

(b) We say that $\mathcal{S}$ is I-learnable iff there exists a learner M which I-learns $\mathcal{S}$.
(c) $I = \{\mathcal{S} \mid \exists M\, [\mathcal{S} \subseteq I(M)]\}$.

Definition 3. (a) [10] We say that M is *confident* iff (i) M is total and (ii) for all total f, $M(f)\downarrow$ or for all but finitely many n, $M(f[n]) =?$.
(b) We say that M is *weakly confident* iff (i) M is total and (ii) for all $f \in \mathcal{R}$, $M(f)\downarrow$ or for all but finitely many n, $M(f[n]) =?$.
(c) [2, 9] We say that M is *reliable* iff (i) M is total and (ii) for all total f, $M(f)\downarrow$ implies M **Ex**-learns f.
(d) We say that M is *weakly reliable* iff (i) M is total and (ii) for all $f \in \mathcal{R}$, $M(f)\downarrow$ implies M **Ex**-learns f.
(e) We say that M is *confident and reliable* iff M is total and, either M **Ex**-learns f or $M(f[n]) =?$ for all but finitely many n.
(f) We say that M is *weakly confident and reliable* iff M is total and, for all $f \in \mathcal{R}$, either M **Ex**-learns f or $M(f[n]) =?$ for all but finitely many n.

Definition 4. We say that M **Conf**-learns $\mathcal{S}$ if M **Ex**-learns $\mathcal{S}$ and M is confident. Similarly, we define **Rel**, **WConf**, **WRel**, **ConfRel** and **WConfRel** learning criteria where we require the learners to be reliable, weakly confident, weakly reliable, confident and reliable, and weakly confident and reliable respectively.

For all the learning criteria considered in this paper, one can assume without loss of generality that the learners are total. In particular, from any learner M, one can effectively construct a total learner M' such that, for all the learning criteria I considered in this paper, $I(M) \subseteq I(M')$ (this can be shown essentially using the same proof as for $I = \mathbf{Ex}$ used by [10]). We often implicitly assume such conversion of learners into total learners. The following proposition shows that learners for unions of confidently learnable classes can be effectively found; similarly for learners of unions of reliably learnable classes.

Proposition 5 (Blum and Blum [2], Minicozzi [9], Osherson, Stob and Weinstein [10]). Each criterion I from **Conf**, **WConf**, **Rel**, **WRel**, **ConfRel**, **WConfRel** is closed effectively under union: there exists a recursive function h_I such that, if M_i I-learns $\mathcal{S}$ and M_j I-learns $\mathcal{S}'$ then $M_{h_I(i,j)}$ I-learns $\mathcal{S} \cup \mathcal{S}'$.

Definition 6. [13] A set $\mathcal{S} \subseteq \mathcal{R}$ is *two-sided classifiable* iff there is a machine M such that, for all $f \in \mathcal{R}$,

(i) if $f \in \mathcal{S}$, then $\forall^\infty x\, [M(f[x]) = 1]$;
(ii) if $f \notin \mathcal{S}$, then $\forall^\infty x\, [M(f[x]) = 0]$.

The next theorem characterizes **WConfRel** in terms of classification.

Theorem 7. Let $\mathcal{S} \subseteq \mathcal{R}$. The following are equivalent:

(a) $\mathcal{S}$ is **WConfRel**-learnable;
(b) A superset of $\mathcal{S}$ is **Ex**-learnable and two-sided classifiable.

3 Initial Results on Unionability

We start with giving the general definition of unionability.

Definition 8. Let I be a learning criterion and $\mathcal{S} \subset \mathcal{R}$.

(a) $\mathcal{S}$ is *I-unionable* iff, for all I-learnable classes $\mathcal{S}'$, $\mathcal{S} \cup \mathcal{S}'$ is I-learnable.
(b) $\mathcal{S}$ is *constructively I-unionable* iff there is an $h \in \mathcal{R}$ such that, for all e, $\mathcal{S} \cup I(M_e) \subseteq I(M_{h(e)})$.
(c) $\mathcal{S}$ is *singleton-I-unionable* iff, for all $f \in \mathcal{R}$, $\mathcal{S} \cup \{f\}$ is I-learnable.
(d) $\mathcal{S}$ is *constructively singleton-I-unionable* iff there is $h \in \mathcal{R}$ such that, for all e, $M_{h(e)}$ I-learns $\mathcal{S} \cup \{\varphi_e\} \cap \mathcal{R}$.

For the various versions of unionability, in the following sections we will consider in detail which classes are I-unionable for I being **Fin**, **Ex** or **BC**, starting with **Fin**-unionability in this section.

Theorem 9 (Blum and Blum [2]). There are classes $\mathcal{S}$ and $\mathcal{S}'$ such that

(a) $\mathcal{S}$ is **Fin**-learnable (and thus $\mathcal{S} \in$ **Conf** and $\mathcal{S} \in$ **WConf**);
(b) $\mathcal{S}'$ is **Rel**-learnable;
(c) $\mathcal{S} \cup \mathcal{S}' \notin$ **BC**.

Thus, both classes $\mathcal{S}$ and $\mathcal{S}'$ are neither **Ex**-unionable nor **BC**-unionable. In the following, we want to characterise **Fin**-unionability.

Theorem 10. (a) $\mathcal{S}$ is **Fin**-unionable iff $\mathcal{S} = \emptyset$.
(b) $\mathcal{S}$ is constructively **Fin**-unionable iff $\mathcal{S} = \emptyset$.
(c) $\mathcal{S}$ is constructively singleton-**Fin**-unionable iff $\mathcal{S} = \emptyset$.
(d) $\mathcal{S}$ is singleton-**Fin**-unionable iff $\mathcal{S}$ is **Fin**-learnable and $\mathcal{S}$ has no recursive accumulation point.

Proof. (a) and (b) Let $\mathcal{S} \neq \emptyset$ be a set of total computable functions and let $f \in \mathcal{S}$. For all i, let f_i be such that $f_i(i) = f(i)+1$ and, for all $x \neq i$, $f_i(x) = f(x)$. Then the class $\mathcal{S}' = \{f_i \mid i \in \mathbb{N}\}$ is **Fin**-learnable, but $\mathcal{S} \cup \mathcal{S}'$ is not **Fin**-learnable.

(c) We keep $\mathcal{S}$ and f and f_i as in part (a and b) above. Furthermore, we consider a recursive function g such that $\varphi_{g(e)} = f_s$, if e is enumerated into K in exactly s steps; $\varphi_{g(e)} = f$, if e is not enumerated into K. Furthermore, let h be a recursive function such that $M_{h(e)}$ **Fin**-learns $\mathcal{S} \cup \{\varphi_e\} \cap \mathcal{R}$. Let $k(e)$ be the first number found, in some algorithmic search, such that $M_{h(e)}(f[k(e)])\downarrow \neq ?$. The function k is total recursive, as, for all e, $M_{h(e)}$ **Fin**-learns f. If e is enumerated into K in exactly s steps, then $k(g(e)) \geq s$, as otherwise, $\varphi_{g(e)}[k(g(e))] = f_s[k(g(e))] = f[k(g(e))]$, and thus $M_{h(g(e))}$ cannot **Fin**-learn both f and $\varphi_{g(e)}$. Hence e is in K iff e is enumerated within $k(g(e))$ steps into K, a contradiction to K being undecidable.

(d) Clearly $\mathcal{S}$ must be in **Fin** to be singleton-**Fin**-unionable.

We first show that **Fin**-learnable classes with a recursive accumulation point are not singleton-**Fin**-unionable. Let $\mathcal{S}$ be such that there is a recursive accumulation point f of $\mathcal{S}$. Suppose $\mathcal{S} \cup \{f\}$ is **Fin**-learnable, as witnessed by M.

Let x be such that $M(f[x])\downarrow \neq ?$. Furthermore, let $f' \in \mathcal{S}$, $f \neq f'$ be such that $f[x] \preceq f'$. Such an f' exists as f is an accumulation point of $\mathcal{S}$. Now M cannot **Fin**-learn both f and f', as $f[x] \preceq f$ and $f[x] \preceq f'$. This is a contradiction to M **Fin**-learning $\mathcal{S} \cup \{f\}$.

Now suppose $\mathcal{S}$ is **Fin**-learnable as witnessed by M and $\mathcal{S}$ has no recursive accumulation point. Let $f \in \mathcal{R}$. We show that $\mathcal{S}_0 \cup \{f\}$ is **Fin**-learnable. If $f \in \mathcal{S}$, nothing is left to be shown. Suppose $f \notin \mathcal{S}$; thus, there exists an x such that $f[x] \notin [\mathcal{S}]$. Let e be an index for f; we define N such that, for all σ,

$$N(\sigma) = \begin{cases} ?, & \text{if } \sigma \prec f[x]; \\ e, & \text{if } f[x] \preceq \sigma; \\ M(\sigma), & \text{otherwise.} \end{cases}$$

It is easy to verify that N **Fin**-learns $\mathcal{S} \cup \{f\}$. □

It is clear that every constructively I-unionable class is I-unionable and every constructively singleton-I-unionable class is singleton-I-unionable. The next proposition gives the third straight-forward inclusion.

Proposition 11. Let $I \in \{\mathbf{Fin}, \mathbf{Conf}, \mathbf{WConf}, \mathbf{Ex}, \mathbf{BC}\}$. If $\mathcal{S}$ is constructively I-unionable then $\mathcal{S}$ is constructively singleton-I-unionable.

Proof. Given e, consider the I-learner $M_{h(e)}$ which always outputs e; if φ_e is total, then $I(M_{h(e)}) = \{\varphi_e\}$, else $I(M_{h(e)}) = \emptyset$. Now, due to the constructive I-unionability of $\mathcal{S}$, the class is also constructively singleton-I-unionable by forming constructively the union with the class I-learnt by $M_{h(e)}$. □

For the criteria **Rel**, **WRel**, **ConfRel** and **WConfRel**, one cannot translate an index e into a learner for φ_e of the given type, as one is not able to test in the limit whether φ_e is partial or total. This obstacle on the way to prove a hypothetical implication like "constructively **Rel**-unionable ⇒ constructively singleton-**Rel**-unionable" is real and the conjectured implication does not hold: On the one hand, every **Rel**-learnable class is constructively **Rel**-unionable [9]; on the other hand, Theorem 17 as well as Blum and Blum's Non-Union-Theorem exhibit a **Rel**-learnable class which is not constructively singleton-**Rel**-unionable.

4 Ex- and BC-Unionable Classes

Case and Fulk [4] investigated **Ex**- and **BC**-unionability and obtained the following basic result that one can always add a function to a given class; so in contrast to finite learning, every **Ex**-learnable class is non-constructively singleton-**Ex**-unionable; the same applies to **BC**-learning.

Proposition 12 (Case and Fulk [4]). If I is either **Ex** or **BC**, $f \in \mathcal{R}$ and $\mathcal{S}$ is I-learnable, then $\mathcal{S} \cup \{f\}$ is I-learnable.

Theorem 13. Suppose I is either **Ex** or **BC**. Suppose $\mathcal{S} \in \mathbf{WConfRel}$. Then $\mathcal{S}$ is constructively I-unionable.

Proof. Suppose $\mathcal{S} \in \mathbf{WConfRel}$ as witnessed by $M \in \mathcal{R}$. Let h be a recursive function such that $M_{h(i)}$ behaves as follows.

Let M_i' be obtained effectively from i such that M_i' is total and $I(M_i') = I(M_i)$. If $M(\sigma) = ?$, then $M_{h(i)}(\sigma) = M_i'(\sigma)$. Otherwise, $M_{h(i)}(\sigma) = M(\sigma)$. It is easy to verify that $M_{h(i)}$ I-learns $\mathcal{S} \cup I(M_i)$. □

Theorem 14. Suppose I is either **Ex** or **BC**. Suppose $\mathcal{S} \in \mathbf{WConf}$. Then $\mathcal{S}$ is constructively singleton-I-unionable.

Proof. Let f be a recursive function such that $M_{f(e)}$ always outputs e on any input. Then, $M_{f(e)}$ **WConf**-learns $\{\varphi_e\}$. Let M_i be a **WConf**-learner for $\mathcal{S}$. Let $h_{\mathbf{WConf}}$ be as from Proposition 5. Then, $h_{\mathbf{WConf}}(f(e), i)$ witnesses the theorem. □

Corollary 15. Suppose I is either **Ex** or **BC**. Let $\mathcal{S} = \{f \in \mathcal{R} : \varphi_{f(0)} = f\}$. Then, $\mathcal{S}$ is constructively singleton-I-unionable, but not I-unionable.

Theorem 16. There are classes $\mathcal{S}, \mathcal{S}' \subseteq \mathcal{R}$ such that

(a) $\mathcal{S}$ and $\mathcal{S}'$ are both **Ex**-learnable;
(b) $\mathcal{S}$ and $\mathcal{S}'$ are both constructively **BC**-unionable;
(c) $\mathcal{S} \cup \mathcal{S}'$ is not **Ex**-learnable;
(d) $\mathcal{S}$ is not constructively singleton-**Ex**-unionable;
(e) $\mathcal{S}'$ is constructively singleton-**Ex**-unionable.

Proof. Kummer and Stephan [8, Theorem 8.1] constructed a uniformly partial-recursive family $\varphi_{g(0)}, \varphi_{g(1)}, \ldots$ of functions such that each $\varphi_{g(n)}$ is undefined at most at one place and $1^n0 \preceq \varphi_{g(n)}$ for all n. Let $\mathcal{S}$ be the set of all total extensions of functions $\varphi_{g(n)}$ which are not total. Let $\mathcal{S}'$ be set of all total $\varphi_{g(n)}$. It is easy to verify that $\mathcal{S}$ and $\mathcal{S}'$ are both in **Ex**.

Kummer and Stephan [8] showed that $\mathcal{S}\cup\mathcal{S}'$ is **BC**-learnable. Actually $\mathcal{S}\cup\mathcal{S}'$ and every subclass of it is constructively **BC**-unionable. To see this, let *patch* be a recursive function such that $\varphi_{patch(i,\sigma)}(x) = \sigma(x)$ if $x < |\sigma|$; $\varphi_{patch(i,\sigma)}(x) = \varphi_i(x)$ if $x \geq |\sigma|$.

Now, let any total **BC**-learner M for some class be given. Now, a new **BC**-learner N, obtained effectively from M, learning $\mathbf{BC}(M) \cup \mathcal{S} \cup \mathcal{S}'$ is defined as follows:

If there is an n such that $1^n0 \preceq \sigma$ and no $x < |\sigma|$ satisfies that $\varphi_{g(n)}(x)$ converges within $|\sigma|$ steps to a value different from $\sigma(x)$,
Then $N(\sigma) = patch(g(n), \sigma)$,
Else $N(\sigma) = M(\sigma)$.

Furthermore, Kummer and Stephan [8] showed that $\mathcal{S} \cup \mathcal{S}'$ is not **Ex**-learnable, hence $\mathcal{S}$ and $\mathcal{S}'$ are not **Ex**-unionable. As $\mathcal{S}'$ is **Fin**-learnable, by Theorem 14, $\mathcal{S}'$ is also constructively singleton-**Ex**-unionable.

Furthermore, $\mathcal{S}$ is not constructively singleton-**Ex**-unionable. Suppose by way of contradiction that h witnesses that $\mathcal{S}$ is constructively singleton-**Ex**-unionable.

Then, the following learner N witnesses that $\mathcal{S} \cup \mathcal{S}' \in \mathbf{Ex}$: If $1^n0 \preceq \sigma$ for some n, then $N(\sigma) = M_{h(g(n))}(\sigma)$, else $N(\sigma) = 0$. However, by Kummer and Stephan [8], such a learner does not exist. □

Theorem 17. There is a class $\mathcal{S}$ which is **Ex**-unionable, **BC**-unionable, but is not constructively singleton-**BC**-unionable.

Proof. For each n, we will define function f_n below. The class $\mathcal{S}$ will consist of all functions of the form $f_n(0)f_n(1)\ldots f_n(x)y^\infty$ which start with values of some f_n until a point x and are constant from then onwards.

Without loss of generality assume that learner M_0 **Ex**-learns all eventually constant functions. The functions f_n satisfy the following properties:

(I) $f_n(0) = n$;
(II) Each f_n is recursive;
(III) The mapping $n, x \mapsto f_n(x)$ is limit-recursive;
(IV) For each $m \leq n$,
either for infinitely many s, $(\exists x)\,[\varphi_{M_m(f_n[s])}(x)\downarrow \neq f_n(x)]$,
or there is a $\sigma \preceq f_n$ such that $(\forall \tau)\,[\varphi_{M_m(\sigma\tau)}$ is a subfunction of $\sigma\tau]$.

Note that above properties imply that M_m does not **BC**-learn f_n, for any $n \geq m$. Thus, in particular, f_n is not an eventually constant function.

The construction of f_n is done by inductively defining longer and longer initial segments $f_n[\ell_{n,t}]$ of f_n together with the length $\ell_{n,t}$. Let $\ell_{n,0} = 0$. In stage t, $\ell_{n,t+1}$ and $f_n[\ell_{n,t+1}]$ are defined as follows: Let m be the remainder of t divided by $n+1$. Search for τ, η, a hypothesis e and an $x < \ell_{n,t} + |\tau\eta|$ such that $\varphi_{M_m(f_n[\ell_{n,t}]\cdot\tau)}(x)\downarrow \neq (f_n[\ell_{n,t}]\cdot\tau\eta)(x)$. If such τ, η, e, x are found then $\ell_{n,t+1} = \ell_{n,t} + |\tau\eta| + 1$ and $f_n[\ell_{n,t+1}] = f_n[\ell_{n,t}]\cdot\tau\eta\cdot 0$ else $\ell_{n,t+1} = \ell_{n,t} + 1$ and $f_n[\ell_{n,t+1}] = f_n[\ell_{n,t}]\cdot 0$.

Note that if the search does not succeed in stage t then it does not succeed in stage $t+n+1$ either, as that stage also deals with the same m and $f_n[\ell_{n,t+n+1}]$ is an extension of $f_n[\ell_{n,t}]$. Therefore each f_n is recursive. Furthermore, the f_n are uniformly limit-recursive as one can use the oracle for K to decide whether the extension exists in each specific case. It is clear that property (IV) of f_n mentioned above is also met by the way each f_n is constructed.

Now suppose that a total learner M_e **Ex**-learns or **BC**-learns a class $\mathcal{S}'$. Thus the functions $f_e, f_{e+1}, f_{e+2}, \ldots$ are not learnt by M_e and thus not members of $\mathcal{S}'$. Now consider the following new learner N for $\mathcal{S} \cup \mathcal{S}'$. Let $f_{n,t}$ be the t-th approximation (as a recursive function) to f_n; the $f_{n,t}$ converge pointwise to f_n. N, on input σ of length $t > 0$, is defined as follows:

If $\sigma \preceq f_d$ for some $d \in \{0, 1, \ldots, e\}$,
Then $N(\sigma)$ is an index for f_d for the least such d,
Else if $\sigma = f_{n,t}(0)f_{n,t}(1)\ldots f_{n,t}(x)y^{t-x-1}$ for some n, y and $x < t-1$,
Then $N(\sigma)$ outputs a canonical index for $f_{n,t}(0)f_{n,t}(1)\ldots f_{n,t}(x)y^\infty$,
Else $N(\sigma) = M_e(\sigma)$.

One can easily verify that N **Ex**-learns $f_0, f_1, \ldots, f_e$ and also **Ex**-learns every member of $\mathcal{S}$. Furthermore, for each $f \in \mathcal{S}' - \mathcal{S} - \{f_0, f_1, \ldots, f_e\}$, there are $n = f(0)$, a least x with $f(x+1) \neq f_n(x+1)$ and a least $x' > x$ with $f(x'+1) \neq f(x')$. If $\sigma \preceq f$ is long enough, then $f_{n,|\sigma|}$ equals f_n for inputs below $x+1$ and $|\sigma| > x'+1$ and thus the learner N outputs $M_e(\sigma)$. Hence if M_e is an **Ex**-learner for $\mathcal{S}'$ then N is an **Ex**-learner for $\mathcal{S} \cup \mathcal{S}'$ and if M_e is a **BC**-learner for $\mathcal{S}'$ then N is a **BC**-learner for $\mathcal{S} \cup \mathcal{S}'$.

Now assume by way of contradiction that $\mathcal{S}$ is constructively singleton-**BC**-unionable as witnessed by a recursive function h. We will define a learner N below. For ease of notation, we define N as running in stages and think of learners as getting the graph of the whole function as input, and outputting a sequence of conjectures, all but finitely many of which are programs for the input function (for **BC**-learning); for **Ex**-learning, this sequence of programs also converges syntactically.

Let f denote the function to be learnt and let $n = f(0)$. Now define a trigger-event m to be activated iff there is a $t > m$ such that $f[m] \preceq f_{n,t}$ (as defined above). If $f = f_n$ then infinitely many trigger events are eventually activated; otherwise only finitely many trigger events are eventually activated. On any input function f, the learner N starts in stage 0.

Stage $\langle i, j \rangle$:
- In this stage N copies the output of $M_{h(i)}$ until
  - (i) the $(\langle i, j \rangle + 1)$-th trigger event has been activated and
  - (ii) there are x, z such that $x > j$ and $\varphi_{M_{h(i)}(f[x])}(z) \neq f(z)$.
- When both events have occurred, the learner N leaves stage $\langle i, j \rangle$ and goes to the next stage $\langle i, j \rangle + 1$.

End stage $\langle i, j \rangle$

Note that whenever the input function f is from $\mathcal{S}$, then only finitely many trigger-events are activated and therefore the construction leaves only finitely many stages. Hence, the learner N eventually follows the learner $M_{h(i)}$, for some i, and thus **BC**-learns f.

Let n be such that $M_n = N$. Consider the behaviour of N on f_n. As, for each prefix σ of f_n, N **BC**-learns $\sigma 0^\infty$, it follows from property (IV) of f_n that there exist infinitely many x such that, for some z, $\varphi_{N(f_n[x])}(z)\downarrow \neq f_n(z)$. Furthermore, infinitely many trigger events are activated on input function being f_n. Thus, inductively, for each stage $\langle i, j \rangle$, $\varphi_{M_{h(i)}(f_n[x])}(z)\downarrow \neq f_n(z)$, for some $x > j$. Therefore, for all i, $\varphi_{M_{h(i)}(f_n[x])} \neq f_n$, for infinitely many x. Thus, for each i, $M_{h(i)}$ does not **BC**-learn f_n. However, as there exists an i such that $f_n = \varphi_i$, the learner $M_{h(i)}$ must **BC**-learn f_n. A contradiction. Thus, $\mathcal{S}$ is not constructively singleton-**BC**-unionable. □

Corollary 18. Due to the implications among the criteria of unionability, the class $\mathcal{S}$ from Theorem 17 also fails to be constructively singleton-**Ex**-unionable, constructively **BC**-unionable or constructively **Ex**-unionable. Furthermore, $\mathcal{S}$ is not **WConf**-learnable.

The next proposition shows that **Ex** and **BC**-unionable classes are everywhere sparse.

Proposition 19. Suppose I is **Ex** or **BC**. Suppose $\mathcal{S}$ is not everywhere sparse. Then $\mathcal{S}$ is neither I-unionable nor constructively singleton-I-unionable.

The following theorem generalises the Non-Union-Theorem.

Theorem 20. Let $\mathcal{S} \subseteq \mathcal{R}$ be **Ex**-learnable. Then there are $\mathcal{S}_0 \subseteq \mathcal{R}$ and $\mathcal{S}_1 \subseteq \mathcal{R}$ such that $\mathcal{S} \cup \mathcal{S}_0$ and $\mathcal{S} \cup \mathcal{S}_1$ are **Ex**-learnable but $\mathcal{S}_0 \cup \mathcal{S}_1$ is not **BC**-learnable.

5 Extendability

In the previous sections, the question was whether a class $\mathcal{S}$ can be extended by either adding a full class $\mathcal{S}'$ or just a function φ_e without losing learnability; in this section we ask whether a class can be extended effectively without prescribing how this should be done. So on one hand, the task becomes easier as it is not prescribed what to add, on the other hand the task might also become more difficult as one has to find functions not yet learnt in order to add them (while previously, they were given by a learner or an index). Before discussing this in detail, the next definition should make the notion of extending more precise.

Definition 21. Let $\mathcal{C}$ be a set of learners and I a learning criterion.

(a) We say that we can *infinitely I-improve learners from $\mathcal{C}$* iff, for all $M \in \mathcal{C}$, there is a learner $N \in \mathcal{P}$ such that $I(M) \subseteq I(N)$ and $I(N) \setminus I(M)$ is infinite.
(b) We say that we can *uniformly infinitely I-improve learners from $\mathcal{C}$* iff there is a recursive function h such that, for all e with $M_e \in \mathcal{C}$, $I(M_e) \subseteq I(M_{h(e)})$ and $I(M_{h(e)}) \setminus I(M_e)$ is infinite.

Proposition 22. Let $\mathcal{C}$ be a set of learners and I be **Ex** or **BC**. Suppose there is a function $g \in \mathcal{R}$ such that, for all e with $M_e \in \mathcal{C}$, $\{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$ is an infinite I-unionable set disjoint from $I(M_e)$. Furthermore, assume that one can determine with a two-sided classifier effectively obtainable from e, for each recursive function f, whether $f \in \{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$. Then we can uniformly infinitely I-improve learners from $\mathcal{C}$.

Lemma 23. Suppose $\mathcal{C}$ is a set of learners and $\sigma_0 \in \mathbb{S}\mathrm{eq}$. Suppose for all e, σ one can effectively find a sequence $\tau_{e,\sigma}$ such that if $M_e \in \mathcal{C}$ and $\sigma_0 \preceq \sigma$, then $\sigma \preceq \tau_{e,\sigma}$ and $M_e(\sigma) \neq M_e(\tau_{e,\sigma})$. Then we can uniformly infinitely **Ex**-improve learners from $\mathcal{C}$.

Proof. By implicit use of the parametric recursion theorem [12], let g be a recursive function such that, for all e, x,

$$\varphi_{g(e,x)} = \bigcup_s \varphi^s_{f(e,x)} \text{ where } \varphi^0_{g(e,x)} = \sigma_0 \cdot e \cdot x \text{ and } \varphi^{s+1}_{g(e,x)} = \tau_{e,\varphi^s_{g(e,x)}}.$$

Now, each $M_e \in \mathcal{C}$ fails to **Ex**-learn every $\varphi_{g(e,x)}$, $x \in \mathbb{N}$. Furthermore, there is a two-sided classifier for each of the classes $\{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$. The theorem now follows from Proposition 22. □

Case and Fulk [4] showed that every **Ex**-learner can be infinitely extended. Furthermore, for the subclass of learners learning a dense set of functions, an effective procedure is implicitly given for turning any such learner into an infinitely more successful one.

Theorem 24 (Case and Fulk [4]). We can infinitely **Ex**-improve every learner. Furthermore, we can *uniformly* infinitely **Ex**-improve all learners M where $\mathbf{Ex}(M)$ is dense.

As an open question, Case and Fulk [4] asked whether there is another effective procedure for the complement, that is, for learners that are not dense.

The next theorem answers this question in the negative by showing that there is no computable function turning any given (index for an) **Ex**-learner which is not successful on a dense set into an (index for a) strictly more successful learner – not even by a single additional function.

Theorem 25. For every recursive function h there is a learner M_e such that $[\mathbf{Ex}(M_e)] \neq [\mathcal{R}]$ and $\mathbf{Ex}(M_{h(e)})$ is not a strict superset of $\mathbf{Ex}(M_e)$.

Proof. It suffices to show that for every recursive h, there is an index e with $[\mathbf{Ex}(M_e)] \neq [\mathcal{R}]$ and either $\mathbf{Ex}(M_{h(e)}) \not\supseteq \mathbf{Ex}(M_e)$ or $\mathbf{Ex}(M_{h(e)}) \setminus \mathbf{Ex}(M_e)$ contains at most one function. (As if, for some recursive function h', for every e, $M_{h'(e)}$ is such that $\mathbf{Ex}(M_{h'(e)})$ exceeds $\mathbf{Ex}(M_e)$ by at least one function, then $\mathbf{Ex}(M_{h'(h'(e))})$ would exceed $\mathbf{Ex}(M_e)$ by at least two functions).

Suppose, by way of contradiction, that there is a recursive function h such that, for all e with $[\mathbf{Ex}(M_e)] \neq [\mathcal{R}]$, $\mathbf{Ex}(M_{h(e)})$ contains $\mathbf{Ex}(M_e)$ and exceeds it by at least two functions.

We define a recursive function g implicitly by inductively defining, for any $e \in \mathbb{N}$, a (possibly finite) $\preceq$-increasing sequence of sequences $(\sigma_i^e)_{i \in \mathbb{N}}$ and a recursive function g by

$$\begin{aligned} \sigma_0^e \quad &= \quad \Lambda; \\ \forall i\, [\sigma_{i+1}^e \text{ is the first } &\sigma \succ \sigma_i^e \text{ found such that } M_{h(e)}(\sigma)\downarrow \neq M_{h(e)}(\sigma_i^e)\downarrow]; \\ \varphi_{g(e)} \quad &= \quad \bigcup_{i \in \mathbb{N}} \sigma_i^e. \end{aligned}$$

We let k be a recursive function such that, for all e, τ, $k(e, \tau)$ is the maximum i such that σ_i^e is defined within $|\tau|$ steps. By Kleene's recursion theorem, there is a program e such that, for all τ,

$$M_e(\tau) = \begin{cases} g(e), & \text{if } \exists i\, [\tau \preceq \sigma_i^e]; \\ \mathrm{pad}(M_{h(e)}(\tau), k(e, \tau)), & \text{if } \exists i\, [\sigma_i^e \cup \tau \text{ is not single-valued}]; \\ \uparrow, & \text{otherwise.} \end{cases}$$

Now if M_e does not learn a dense set of functions, then $\mathbf{Ex}(M_{h(e)})$ must exceed $\mathbf{Ex}(M_e)$ by at least two more functions.

Case 1: $\varphi_{g(e)}$ is total.

Then M_e **Ex**-learns *only* $\varphi_{g(e)}$; thus, $M_{h(e)}$ **Ex**-learns $\varphi_{g(e)}$ by supposition. However, by construction of σ_i^e and $g(e)$, $M_{h(e)}$ on $\varphi_{g(e)}$ makes infinitely many mind changes, a contradiction.

Case 2: σ_i^e is defined only for finitely many i.

Let i be the maximum such that σ_i^e is defined. Thus, M_e is undefined on any extension of σ_i^e, and, hence, does not learn a dense set. Suppose $f \in \mathcal{R}$ does not extend σ_i^e. For all j large enough, we now have $M(f[j]) = \text{pad}(M_{h(e)}(f[j]), i)$. Thus, for large enough j, $M(f[j])$ is semantically equivalent to $M_{h(e)}(f[j])$. Thus, any function that is not an extension of σ_i^e, is **Ex**-learned by $M_{h(e)}$ iff it is **Ex**-learned by M_e. Thus, as $M_{h(e)}$ never changes its mind beyond σ_i^e, on any extension of σ_i^e, it can **Ex**-learn at most *one* more function than M_e, a contradiction. □

As an immediate corollary, we get that we cannot constructively find initial segments where a given learner does not learn any extension.

Corollary 26. There is no function $g \in \mathcal{P}$ such that, for all e with $\mathbf{Ex}(M_e)$ not dense, we have that $g(e)$ is a finite sequence with $g(e) \notin [\mathbf{Ex}(M_e)]$.

Case and Fulk [4] ask whether there is any partitioning of all learners into two (or at least finitely many) sets such that, for each of the sets, all learners from that set can be uniformly extended. From Theorem 25 we know that this partitioning cannot be according to whether the set of learned functions is dense. The following theorem answers the open problem by giving a different split of all possible learners into two different classes.

Theorem 27. Let $\mathcal{C}$ be the set of all total learners M such that M changes its mind on a dense set of sequences. Then we can uniformly infinitely **Ex**-improve learners from $\mathcal{C}$ and from $\mathcal{R} \setminus \mathcal{C}$.

Proof. It follows from Lemma 23, that we can uniformly infinitely **Ex**-improve learners from $\mathcal{C}$. We now consider the case of extending learners from $\mathcal{R} \setminus \mathcal{C}$. For any given e and t, let $\tau_{e,t}$ denote the length-lexicographically first sequence found such that M_e does not change its mind on the first t extensions of $\tau_{e,t}$. For any sequence σ and any b we let $g(\sigma, b)$ denote an index for σb^∞. Let $h \in \mathcal{R}$ be such that, for all e and σ,

$$M_{h(e)}(\sigma) = \begin{cases} g(\tau_{e,|\sigma|}, b), & \text{if there is } b \text{ with } \sigma \preceq \tau_{e,|\sigma|} b^\infty; \\ M_e(\sigma), & \text{otherwise.} \end{cases}$$

For all e with $M_e \in \mathcal{R} \setminus \mathcal{C}$, we have that the sequence $\tau_{e,0}, \tau_{e,1}, \ldots$ converges to a τ_e such that M_e does not make any mind changes on any extension of τ_e. Now, $M_{h(e)}$ learns $\mathbf{Ex}(M_e) \cup \{\tau_e \cdot b^\infty \mid b \in \mathbb{N}\}$. Note that M_e can **Ex**-learn at most one function extending τ_e. The theorem follows. □

As one can effectively convert any partial learner to a total learner with the same (or more) learning capacity, the above result also applies for partial learners.

For **Fin**-learning, extending learners is much easier: any learner that learns anything at all can be infinitely extended.

Theorem 28. Let I be one of $\mathbf{Ex}_m$ or $\mathbf{Conf}$. There is a function h such that, for all e with $I(M_e) \neq \emptyset$, $I(M_{h(e)})$ infinitely extends $I(M_e)$. Here $M_{h(e)}$ is confident, if M_e is confident.

Similarly for reliable learning, one can always extend a learner infinitely.

Theorem 29. There is a recursive function h such that, for e with M_e reliable, $M_{h(e)}$ is reliable and $\mathbf{Ex}(M_{h(e)})$ infinitely extends $\mathbf{Ex}(M_e)$.

References

1. Janis A. Bārzdiņš. Two theorems on the limiting synthesis of functions. *In Theory of Algorithms and Programs, Latvian State University, Riga, USSR*, 210:82–88, 1974.
2. Lenore Blum and Manuel Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.
3. John Case. Periodicity in generations of automata. *Mathematical Systems Theory*, 8:15–32, 1974.
4. John Case and Mark Fulk. Maximal machine learnable classes. *Journal of Computer and System Sciences*, 58:211–214, 1999.
5. John Case, Sanjay Jain and Susan Ngo Manguelle. Refinements of inductive inference by Popperian and reliable machines, *Kybernetika*, 30:23–52, 1994.
6. John Case and Carl Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
7. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
8. Martin Kummer and Frank Stephan. On the structure of degrees of inferability. *Journal of Computer and System Sciences*, 52:214–238, 1996.
9. Eliana Minicozzi. Some natural properties of strong-identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.
10. Daniel Osherson, Michael Stob and Scott Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists*. MIT Press, Cambridge, Mass., 1986.
11. Lenny Pitt. Inductive inference, DFAs, and computational complexity. *Analogical and Inductive Inference*. Proceedings of the Second International Workshop (AII 1989). Springer LNAI 397:18–44, 1989.
12. Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw Hill, New York, 1967. Reprinted in 1987.
13. Frank Stephan. On one-sided versus two-sided classification. *Archive for Mathematical Logic*, 40:489–513, 2001.
14. Arun Sharma, Frank Stephan and Yuri Ventsov. Generalized notions of mind change complexity. *Information and Computation* 189:235–262, 2004.