Enlarging Learnable Classes

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

[1] Department of Computer Science, National University of Singapore, Singapore 117417, Republic of Singapore
sanjay@comp.nus.edu.sg
[2] Hasso Plattner Institute, 14482 Potsdam, Germany
timo.koetzing@hpi.de
[3] Department of Mathematics, National University of Singapore, Singapore 119076, Republic of Singapore
fstephan@comp.nus.edu.sg

Abstract. We study which classes of recursive functions satisfy that their union with any other explanatorily learnable class of recursive functions is again explanatorily learnable. We provide sufficient criteria for classes of recursive functions to satisfy this property and also investigate its effective variants. Furthermore, we study the question which learners can be effectively extended to learn a larger class of functions. We solve an open problem by showing that there is no effective procedure which does this task on all learners which do not learn a dense class of recursive functions. However, we show that there are two effective extension procedures such that each learner is extended by one of them.

1 Introduction

One branch of inductive inference investigates the learnability of recursive functions; the basic scenario given in the seminal paper by Gold [10] 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 $f(0), f(1), \ldots, f(n-1)$ (e_0 being based on no data) 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, we consider learnability by partial recursive learners; with M_e we refer to the learner derived from the e-th partial recursive function. This setting of learning is also called *learning in the limit* (see also [20], which surveys recursive function learning).

⋆ Supported by NUS grants C252-000-087-001, R252-000-420-112, R146-000-181-112 and R252-000-534-112.

⋆⋆ Major parts of this paper were written when Timo Kötzing was visiting the Department of Computer Science at the National University of Singapore and was at Max-Planck-Institut für Informatik, Saarbrücken.

⋆⋆⋆ Supported in part by NUS grant R252-000-420-112, R146-000-181-112 and R252-000-534-112.

During the course of time, several variants of this basic notion of explanatory learning (**Ex**) have been considered; most notably, *behaviourally correct learning* (**BC**) [3], 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 [16] introduced a generalisation of this notion, namely *confident learning* (**Conf**), where the learner can revise the hypothesis finitely often; it must, however, on each total function f, even if it is not in the class to be learnt or not even recursive, eventually stabilise 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 behaviour on non-recursive ones is not constrained (**WConf**, Sharma, Stephan and Ventsov [19]).

Minicozzi [15] called a learner *reliable* (**Rel**) iff the learner, on every function (including total but non-recursive functions), 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); Osherson, Stob and Weinstein [16, Exercise 4.6.1A] generalised this notion to *weakly reliable* (**WRel**) where the learner is weakly reliable iff the learner is reliable on every recursive function and there are no constraints on its behaviour on non-recursive functions. One can combine the notion of reliability and confidence: A learner is *weakly confident and weakly reliable* (**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 to any conjecture).

Formal definitions of the above criteria are given in Section 2. The relations between those criteria have been extensively studied, giving the following inclusion relations [4, 6, 8, 10, 12, 15, 16, 19]:

– $\mathbf{Fin} \subset \mathbf{Conf} \subset \mathbf{WConf} \subset \mathbf{Ex} \subset \mathbf{BC}$;
– $\mathbf{ConfRel} \subset \mathbf{WConfRel} \subset \mathbf{WRel} \subset \mathbf{Ex} \subset \mathbf{BC}$;
– $\mathbf{Rel} \subset \mathbf{WRel} \subset \mathbf{Ex} \subset \mathbf{BC}$;
– $\mathbf{Fin} \not\subseteq \mathbf{Rel}$ and $\mathbf{Rel} \not\subseteq \mathbf{WConf}$.

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?

Bārzdiņš [3] and Blum and Blum [4] gave with the *non-union theorem* 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, that is, given two confident (reliable) learners, one can effectively find a confident (reliable) learner which explanatorily learns the union of functions explanatorily learnt by the two learners. In fact Minicozzi [15] proved a stronger result that, given an index for a recursively enumerable set A of reliable learners, one can effectively find a reliable learner which learns the union of the classes of functions learnt by the individual learners in the set A. Apsītis, Freivalds, Simanovskis and Smotrovs [2] also considered closedness properties of **Ex**-identification.

The non-union theorem has been extended in various ways; for example, if one considers learning with oracles, there is a choice of a confidently learnable $\mathcal{S}$ and a consistently learnable $\mathcal{S}'$ such that their union is not **Ex**-learnable, even with any non-high oracle A [9, 13] — "non-high" is the best that one can expect in this context as a high oracle permits to learn the whole class of recursive functions [1].

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 [4] used slightly different classes $\mathcal{S}$ and $\mathcal{S}'$ which were $\{0,1\}$-valued; our $\mathcal{S}$ and $\mathcal{S}'$, which are taken from [3], make the presentation simpler; a proof of the non-union theorem using these classes can also be found in [20]). 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. However, learnability of these unions is not effective. If, given e, one can effectively find a **Ex**-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, $N(\sigma) = M_{h(\sigma(0))}(\sigma)$. This learner N, **Ex**-learns $\mathcal{S} \cup \mathcal{S}'$, as $\sigma(0)$ constrains the functions in $\mathcal{S}$ to be only from $\mathcal{S}_{\sigma(0)}$. However, this is in contradiction to the non-union theorem, and thus one cannot effectively find, given e, an **Ex**-learner for $\mathcal{S}_e \cup \mathcal{S}'$.

The above 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 recursive 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 behaviourally 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.

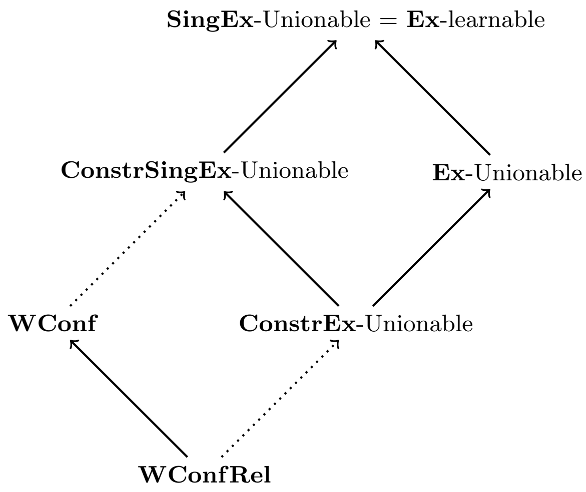

Fig. 1. The inclusion relations for the various unionability notions. It is unknown whether the dotted arrows might also go in the converse direction. All inclusions are given by arrows (and possibly reversed dotted arrows) and the concatenations of these.

2. If a class $\mathcal{S}$ has a weakly confident and weakly 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 fail for all non-empty classes; only non-constructive unions with singletons are possible in the case that every pointwise limit of functions in the class is either in the class or not recursive.

In the last item, we say that a total function $f : \mathbb{N} \to \mathbb{N}$ is the *pointwise limit* of a sequence of total function $(g_n)_{n \in \mathbb{N}}$ iff, for all $i \in \mathbb{N}$, $\lim_{n \to \infty} g_n(i) = f(i)$. All our results for the cases of purely **Ex**-learning are summarised 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 [5] 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 [5] 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?

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 31 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 33 we show that there is a recursive function h such that $M_{h(e,b)}$ **Ex**-learns a proper superclass of what M_e **Ex**-learns 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 (including 0). The symbols $\subseteq, \subset, \supseteq, \supset$ respectively denote subset, proper subset, superset and proper superset. We let $\langle x, y\rangle$ denote $(x + y) \cdot (x + y + 1)/2 + y$; this is Cantor's bijection from $\mathbb{N} \times \mathbb{N}$ onto $\mathbb{N}$, which is increasing in both its arguments and satisfies $\langle 0, 0\rangle = 0$. The pairing function can easily be extended to multiple arguments. Let $\forall^\infty$ denote "for all but finitely many".

Let $\mathcal{R}$ denote the set of all total recursive functions $f : \mathbb{N} \to \mathbb{N}$ and $\mathcal{P}$ denote the set of all partial recursive functions $f : \mathbb{N} \to \mathbb{N}$. Note that [17] calls total recursive functions general recursive functions; following [17], for brevity of notation, in this paper we will sometimes refer to total recursive functions as just recursive functions. Let $\mathcal{R}_{0,1}$ denote the set of all total recursive functions f with range$(f) \subseteq \{0, 1\}$. Let φ denote a fixed acceptable programming system [17] for $\mathcal{P}$. 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 K denote the diagonal halting set $\{x \mid \varphi_x(x)\downarrow\}$. 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, $\varphi_{\mathrm{pad}(i,j)} = \varphi_i$. We let $\mathcal{S}$ range over sets of recursive functions. Please find unexplained recursion theoretic notions in Rogers' book [17].

Let σ, τ range over finite sequences of natural numbers. 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, $\sigma^- \cdot \sigma(|\sigma| - 1) = \sigma$. Let $[\mathcal{S}] = \{f[n] \mid f \in \mathcal{S}\}$. Thus, $[\mathcal{R}] = \mathbb{S}$eq. For notational simplification, $[f] = [\{f\}]$. A class $\mathcal{S}$ is said to be *dense* if $[\mathcal{S}] = [\mathcal{R}]$. A set of sequences S is said to be *dense* if $\{\gamma \mid (\exists \alpha \in S)[\gamma \preceq \alpha]\} = [\mathcal{R}]$. A class $\mathcal{S}$ is *everywhere sparse* iff for all $\tau \in \mathbb{S}\mathrm{eq}$, 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 *recursive operator* [17] Θ is a recursive mapping from $\mathbb{S}\mathrm{eq}$ to $\mathbb{S}\mathrm{eq}$ such that for $\sigma \preceq \tau$, $\Theta(\sigma) \preceq \Theta(\tau)$.

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_i denote the learner derived from φ_i, that is $M_i(\sigma) = \varphi_i(\mathit{code}(\sigma))$, where *code* is a recursive one-one coding of all finite sequences onto $\mathbb{N}$. For ease of notation, we consider M_i itself as a partial-recursive function.

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 $f \in \mathcal{R}$.

(a) [10] 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) [3, 8] 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) [3, 8] 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) [8] We say that M $\mathbf{Ex}_n$-*learns* f (written: $f \in \mathbf{Ex}_n(M)$) iff (i) M **Ex**-learns f and (ii) $\mathrm{card}(\{m \mid ? \neq M(f[m]) \neq M(f[m+1])\}) \leq n$.

Note that the notions of **Fin**-learning and $\mathbf{Ex}_0$-learning are identical. Intuitively, $? \neq M(f[m]) \neq M(f[m{+}1])\})$ denotes a mind change by M on f. We say that M makes a *mind change* at $f[m+1]$ iff $? \neq M(f[m]) \neq M(f[m+1])$ (in particular, changing from ? to a number does *not* count as a mind change).

Definition 2. Let I be **Fin**, **Ex** or **BC**; let $\mathcal{S} \subseteq \mathcal{R}$.

(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) [16] We say that M is *confident* iff (i) M is total recursive 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 recursive and (ii) for all $f \in \mathcal{R}$, $M(f)\!\downarrow$ or for all but finitely many n, $M(f[n]) = ?$.
(c) [4, 15] We say that M is *reliable* iff (i) M is total recursive and (ii) for all total f, $M(f)\!\downarrow$ implies M **Ex**-learns f.

(d) [16] We say that M is *weakly reliable* iff (i) M is total recursive 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 recursive and, for all total f, either M **Ex**-learns f or $M(f[n]) = ?$ for all but finitely many n.

(f) We say that M is *weakly confident and weakly reliable* iff M is total recursive 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.

If M is confident then $I(M) = \mathbf{Ex}(M)$ else $I(M)$ is undefined.

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 weakly reliable, respectively.

Note that the learners from Definition 3 are total recursive; furthermore, for the criteria **Ex** and **BC**, one can assume without loss of generality that the learners are total recursive. In particular, from any learner M, one can effectively construct a total recursive learner M' such that, for $I \in \{\mathbf{Ex}, \mathbf{BC}\}$, $I(M) \subseteq I(M')$, see the proof of Osherson, Stob and Weinstein [16] for the case of $I = \mathbf{Ex}$. We often implicitly assume such conversion of learners into total recursive learners.

Proposition 5 (Lindner [14]). Suppose $f \in \mathcal{R}$ is an accumulation point for $\mathcal{S} \subseteq \mathcal{R}$. Then $\mathcal{S} \cup \{f\} \notin \mathbf{Fin}$.

Proof. For ease of reference, we include the proof. Suppose by way of contradiction that $\mathcal{S} \cup \{f\}$ 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 $\mathcal{S} \cup \{f\}$. □

Proposition 6 (Blum and Blum [4], Minicozzi [15], Osherson, Stob and Weinstein [16]). There exists a recursive function $h_{\mathbf{Rel}}$ such that, if M_i and M_j are reliable then $M_{h_{\mathbf{Rel}}(i,j)}$ is reliable and $\mathbf{Ex}(M_i) \cup \mathbf{Ex}(M_j) \subseteq \mathbf{Ex}(M_{h_{\mathbf{Rel}}(i,j)})$. Similar results hold when one considers confident, weakly confident, weakly reliable or weakly confident and weakly reliable learners, respectively using recursive functions $h_{\mathbf{Conf}}$, $h_{\mathbf{WConf}}$, $h_{\mathbf{WRel}}$, $h_{\mathbf{WConfRel}}$.

Intuitively, the above proposition says that for (weakly) reliable, (weakly) confident learning and their combinations, one can effectively combine the learning powers of two reliable/confident learners (which can thus be extended to combining finitely many reliable/confident learners).

Classification of functions can be made in two versions, one for classifiers classifying all functions [18] and one for classifiers classifying all recursive functions [7]; we take the version with respect to the class of all recursive functions $\mathcal{R}$.

Definition 7 (Case, Kinber, Sharma and Stephan [7]). 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 characterises **WConfRel** in terms of classification; note that we would get **ConfRel** in the case that we consider two-sided classifiable with respect to all functions in the definition above.

Theorem 8. 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.

Proof. Without loss of generality assume that the class $\mathcal{S}$ is not empty. Suppose $\mathcal{S}$ is **WConfRel**-learnable as witnessed by total recursive M. Then $\mathbf{Ex}(M)$ is an **Ex**-learnable superset of $\mathcal{S}$. Let N be such that, for all σ, $N(\sigma) = 0$, if $M(\sigma) = ?$; $N(\sigma) = 1$ otherwise. It is easy to verify that N two-sided classifies $\mathbf{Ex}(M)$.

For the converse direction, suppose now a superset $\mathcal{S}'$ of $\mathcal{S}$ is **Ex**-learnable, as witnessed by M, and two-sided classifiable, as witnessed by N. Let P be such that, for all σ, if $N(\sigma) = 1$ then $P(\sigma) = M(\sigma)$ else $P(\sigma) = ?$. It is easy to see that P **WConfRel**-learns $\mathcal{S}'$. □

3 Initial Results on Unionability

We start with giving the general definition of unionability. Note that a learner **Rel**-learns $\mathcal{S}$ iff it is a reliable learner and **Ex**-learns $\mathcal{S}$; similarly for **Conf**, **WRel**, **WConf**, **WConfRel**-learning. This is relevant for part (b) in the definition below.

Definition 9. 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 I-learnable classes $\mathcal{S}'$ and for all indices e where M_e I-learns $\mathcal{S}'$, $M_{h(e)}$ I-learns $\mathcal{S} \cup \mathcal{S}'$.
(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 $\mathcal{S} \cup (\{\varphi_e\} \cap \mathcal{R})$.

Though I-unionable implies singleton-I-unionable for all criteria of learning considered in this paper, this is not the case for constructive versions. For finite, confident, explanatory or behaviourally correct learning, contructive I-unionability does imply constructive singleton-I-unionability (Proposition 12), however this is not the case for reliable learning (see Proposition 13 along with Proposition 6).

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.

The non-union theorem [3, 4, 20] gives an example of a **Fin**-learnable class $\mathcal{S}$ and a **Rel**-learnable class $\mathcal{S}'$ such that their union is not **BC**-learnable.

Theorem 10 (Bārzdiņš [3], Blum and Blum [4]). Let $\mathcal{S} = \{f \in \mathcal{R} \mid \varphi_{f(0)} = f\}$ and $\mathcal{S}' = \{f \in \mathcal{R} \mid (\forall^\infty x)[f(x) = 0]\}$.

(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 state the characterisation of **Fin**-unionability; we would like to thank the anonymous referee for pointing out that this characterisation is due to Lindner [14].

Theorem 11 (Lindner [14]).

(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) The empty class $\emptyset$ is clearly constructively **Fin**-unionable. Now consider a non-empty $\mathcal{S} \subseteq \mathcal{R}$ 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, as f is an accumulation point of $\mathcal{S}'$ (see Proposition 5).

(c) The empty class is clearly singleton-**Fin**-unionable. Suppose $\mathcal{S}$ is not empty, and f and f_i are as in the proof of part (a) and (b) above. Suppose by way of contradiction that h is a recursive function such that $M_{h(e)}$ **Fin**-learns $\mathcal{S} \cup (\{\varphi_e\} \cap \mathcal{R})$. Fix a recursive enumeration of K. Now, 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. 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.

If $\mathcal{S}$ has a recursive accumulation point, then, by Proposition 5, $\mathcal{S}$ is not singleton-**Fin**-unionable.

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 $\mathcal{S} \cup \{f\}$ 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 inclusion for the criteria not involving reliability.

Proposition 12. 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 $I \in \{$**Rel**, **WRel**, **ConfRel**, **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. So there might be an obstacle on the way to try to prove a hypothetical implication like "constructively **Rel**-unionable ⇒ constructively singleton-**Rel**-unionable". Indeed this obstacle is a real one and the exact opposite of the hypothetical implication holds. On one hand Minicozzi [15] proved that there are I-learnable classes and all I-learnable classes are constructively I-unionable — she did this for $I = \mathbf{Rel}$ and the proof carries over to the other choices of I considered here. On the other hand no class is constructively singleton-I-unionable, as here the singleton is not given by a learner but by a program which might be partial; this program cannot be transformed into a learner effectively. This is now formally proven in the following result.

Proposition 13. Let $I \in \{\mathbf{WRel}, \mathbf{Rel}, \mathbf{WConfRel}, \mathbf{ConfRel}\}$. There is no class $\mathcal{S}$ such that $\mathcal{S}$ is constructively singleton-I-unionable.

Proof. Assume that such a class $\mathcal{S}$ would exist as witnessed by a recursive function h. Then $M_{h(e)}$ is a weakly reliable learner for all e. Now define a function $\varphi_{g(e)}$ inductively as the limit of the following finite sequences σ_s:

Let σ_0 be any given finite sequence and, for $s = 0, 1, \ldots$, let σ_{s+1} be the first proper extension of σ_s found (in some algorithmic search) such that either (i) $M_{h(e)}(\sigma_{s+1}) \neq M_{h(e)}(\sigma_s)$ or (ii) $\varphi_{M_{h(e)}(\sigma_s)}(x)\downarrow \neq \sigma_{s+1}(x)\downarrow$ for some x or (iii) $M_{h(e)}(\sigma_{s+1}) = ?$.

By reliability of $M_{h(e)}$, the function $\varphi_{g(e)}$ is total. It is clear that $M_{h(e)}$ does not I-learn $\varphi_{g(e)}$. However, by the recursion theorem [17], there is an e with $\varphi_{g(e)} = \varphi_e$. As φ_e is total, $M_{h(e)}$ is supposed to I-learn φ_e though, by construction, this is not the case. This contradiction therefore shows that $\mathcal{S}$ is not singleton-I-unionable. $\square$

The proof of the above proposition can be adjusted to prove the following result. Recall that $\mathcal{S}$ is *everywhere sparse* iff for every sequence σ, there is an extension τ such that no function $f \in \mathcal{S}$ extends τ. Written more formally, $\mathcal{S}$ is everywhere sparse iff $\forall\sigma\exists\tau \succeq \sigma\,[\tau \notin [\mathcal{S}]]$.

Proposition 14. Let I be any learning criterion such that every I-learner for a class $\mathcal{S}$ is also a **BC**-learner for $\mathcal{S}$. Let $\mathcal{S}$ be any I-learnable class.

(a) If $\mathcal{S}$ is constructively singleton-I-unionable then $\mathcal{S}$ is everywhere sparse.
(b) If **Fin**-learnability implies I-learnability and $\mathcal{S}$ is I-unionable then $\mathcal{S}$ is everywhere sparse.

Proof. Consider an I-learnable class $\mathcal{S}$ and a finite sequence τ such that every $\eta \succeq \tau$ satisfies $\eta \in [\mathcal{S}]$.

(a) Assume by way of contradiction that $\mathcal{S}$ is constructively singleton-I-unionable and this is witnessed by a recursive function h. Now let $\varphi_{g(e)}$ be constructed as in the proof of Proposition 13 with the only differences that:

1. $\sigma_0 = \tau \cdot e$;
2. when forming σ_{s+1} from σ_s, only (ii) and (iii) are used to extend σ_s, as syntactic mind changes like in (i) are allowed for **BC**-learning.

It follows from the denseness of $\mathcal{S}$ above τ that each $\varphi_{g(e)}$ is a total recursive function and that $M_{h(e)}$ does not I-learn $\varphi_{g(e)}$. As in Proposition 13, there is an e with $\varphi_e = \varphi_{g(e)}$ and $M_{h(e)}$ does not I-learn φ_e. Hence h cannot witness that $\mathcal{S}$ is constructively singleton-I-unionable.

(b) This is very similar to (a). One constructs $\varphi_{g(e)}$ as in part (a), except that $\varphi_{g(e)}$ is not diagonalising against $M_{h(e)}$ but against M_e (that is one takes $h(e) = e$). Note that for some e, for which M_e is not an I-learner for $\mathcal{S}$, the function $\varphi_{g(e)}$ can be partial. Now let

$$\mathcal{S}' = \{\varphi_{g(e)} \mid e \in \mathbb{N} \wedge \varphi_{g(e)} \in \mathcal{R}\}.$$

The class $\mathcal{S}'$ is **Fin**-learnable and thus I-learnable; the learner waits until it has seen $\tau \cdot e$ and then conjectures $g(e)$. If M_e I-learns $\mathcal{S}$, then $\varphi_{g(e)}$ is total and not I-learnt by M_e. Hence there is no I-learner for $\mathcal{S} \cup \mathcal{S}'$. Therefore the class $\mathcal{S}$ is not I-unionable. $\square$

Proposition 14 (b) does not cover the learning criteria **Rel**, **WRel**, **ConfRel** and **WConfRel**. Among these, all **ConfRel**-learnable and **WConfRel**-learnable classes are everywhere sparse. However, there are **Rel**-learnable and **WRel**-learnable classes, such as the class $\{f \in \mathcal{R} \mid \forall^\infty x\,[f(x) = 0]\}$, which are dense and these classes are, by Minicozzi's result [15], also **Rel**-unionable and **WRel**-unionable, respectively.

4 Ex- and BC-Unionable Classes

Case and Fulk [5] 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 15 (Case and Fulk [5]). 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 16. Suppose I is either **Ex** or **BC**. Suppose $\mathcal{S} \in \mathbf{WConfRel}$. Then $\mathcal{S}$ is constructively I-unionable.

Proof. The empty class is clearly constructively **Ex**-unionable and constructively **BC**-unionable. Now suppose $\mathcal{S}$ is a non-empty class of recursive functions and M witnesses that $\mathcal{S} \in \mathbf{WConfRel}$. 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 recursive and $I(M'_i) \supseteq I(M_i)$. 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 17. Suppose I is either **Ex** or **BC**. Suppose $\mathcal{S} \in \mathbf{WConf}$. Then $\mathcal{S}$ is constructively singleton-I-unionable.

Proof. The empty class is clearly constructively singleton-**Ex**-unionable and constructively singleton-**BC**-unionable. Now suppose $\mathcal{S}$ is a non-empty class of recursive functions. Let g be a recursive function such that $M_{g(e)}$ always outputs e on any input. Then, $M_{g(e)}$ **WConf**-learns $\{\varphi_e\}$. Let M_i be a weakly confident learner which **Ex**-learns $\mathcal{S}$. Let $h_{\mathbf{WConf}}$ be as from Proposition 6. Then, $h_{\mathbf{WConf}}(g(e), i)$ witnesses the theorem. □

The following corollary now follows from the non-union theorem of Bārzdiņš [3] and Blum and Blum [4], Theorem 10 and Theorem 17.

Corollary 18. Suppose I is either **Ex** or **BC**. Then the class $\mathcal{S} = \{f \in \mathcal{R} \mid \varphi_{f(0)} = f\}$ is constructively singleton-I-unionable, but not I-unionable.

Theorem 19. 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 [13, Theorem 8.1] constructed a uniformly partial-recursive family $\varphi_{g(0)}, \varphi_{g(1)}, \ldots$ of functions with the following properties, where $\mathcal{S} = \{f \mid f \in \mathcal{R}$ and $(\exists n)[\varphi_{g(n)} \subseteq f$ and $\varphi_{g(n)}$ is not total$]\}$ and $\mathcal{S}' = \{\varphi_{g(n)} \mid n \in \mathbb{N}$ and $\varphi_{g(n)} \in \mathcal{R}\}$:

(I) for all n, $\varphi_{g(n)}$ is undefined on at most one input;
(II) for all n, $1^n 0 \preceq \varphi_{g(n)}$;
(III) $\mathcal{S}, \mathcal{S}' \in \mathbf{Ex}$;
(IV) $\mathcal{S} \cup \mathcal{S}' \notin \mathbf{Ex}$;
(V) $\mathcal{S} \cup \mathcal{S}' \in \mathbf{BC}$.

The class $\mathcal{S} \cup \mathcal{S}'$ and every subclass of it is constructively **BC**-unionable. To see this, let *patch* be a recursive function satisfying the following equation:

$$\varphi_{patch(i,\sigma)}(x) = \begin{cases} \sigma(x), & \text{if } x < |\sigma|; \\ \varphi_i(x), & \text{otherwise.} \end{cases}$$

Let any total recursive **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 $1^n 0 \preceq \sigma$ 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 $N(\sigma) = M(\sigma)$.

By property (IV) above, $\mathcal{S} \cup \mathcal{S}'$ is not **Ex**-learnable. Thus, $\mathcal{S}$ and $\mathcal{S}'$ are not **Ex**-unionable. As $\mathcal{S}'$ is **Fin**-learnable, by Theorem 17, $\mathcal{S}'$ 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 $1^n 0 \preceq \sigma$ for some n, then $N(\sigma) = M_{h(g(n))}(\sigma)$, else $N(\sigma) = 0$. However, by Kummer and Stephan [13], such a learner does not exist. □

Theorem 20. There is a class $\mathcal{S}$ which is **Rel**-learnable, **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. Clearly $\mathcal{S}$ is **Rel**-learnable.

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 the 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 $f_n(0) = n$ and $\ell_{n,0} = 1$. 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 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, 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 recursive 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 where $f_{n,t}(0) = n$ for all t; the $f_{n,t}$ converge pointwise to f_n. The learner 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 Begin
  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 $f_{n,t}(0) f_{n,t}(1) \ldots f_{n,t}(x) y^{\infty}$
    (* note that in this case, $\sigma(0) = n$ *),
  Else $N(\sigma) = M_e(\sigma)$
  End

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 up to $x+1$ 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. Thus, for all i, $M_{h(i)}$ **BC**-learns $\mathcal{S}$. We will define a learner N below. For the r, k such that $M_r = N$ and $\varphi_k = f_r$, we will then show that $M_{h(k)}$ does not **BC**-learn $\varphi_k = f_r$, thus getting a contradiction.

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.

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

We get the following corollary to Theorem 20 due to the implications among the criteria of unionability and by Theorem 17.

Corollary 21. Let $\mathcal{S}$ be the class from Theorem 20. Then $\mathcal{S}$ fails to be constructively singleton-**Ex**-unionable, constructively **BC**-unionable or constructively **Ex**-unionable. Furthermore, $\mathcal{S}$ is not **WConf**-learnable.

5 Generalising the Non-Union Theorem

In this section we show that we can build the non-union theorem above any learnable class (Theorem 26). That is, for all **Ex**-learnable classes $\mathcal{S}$, there exist classes $\mathcal{S}_1, \mathcal{S}_2$ such that $\mathcal{S} \cup \mathcal{S}_1$ and $\mathcal{S} \cup \mathcal{S}_2$ are **Ex**-learnable, but their union is not.

Definition 22. Let Θ be a recursive operator and I a learning criterion. We say that I is *robust under* Θ iff, for all sets of functions $\mathcal{S}$, $\mathcal{S}$ is I-learnable iff $\Theta(\mathcal{S})$ is I-learnable.

The proof of the following lemma is straightforward.

Lemma 23. If Θ_0 satisfies, for all f, $\Theta_0(f) = 0^{f(0)}10^{f(1)}10^{f(2)}1\ldots$, then **Ex** and **BC** are robust under Θ_0.

Lemma 24. Let $T \subseteq [\mathcal{R}]$ be recursively enumerable and prefix closed set such that any element of T has at least two incomparable extensions in T. Then there are **Ex**-learnable classes $\mathcal{S}_0$ and $\mathcal{S}_1$, with $[\mathcal{S}_0] \subseteq T$ and $[\mathcal{S}_1] \subseteq T$ such that $\mathcal{S}_0 \cup \mathcal{S}_1$ is not **BC**-learnable.

Proof. For each $\tau \in T$, let $r_0(\tau)$ and $r_1(\tau)$ be the first pair of incomparable extensions of τ found in a fixed enumeration of T. Note that r_0 and r_1 are recursive functions. Define $\Theta_1(f)$ for $f \in \mathcal{R}_{0,1}$, as follows. Inductively define a sequence $(\sigma_i)_{i\in\mathbb{N}}$ such that $\sigma_0 = \Lambda$ and $\sigma_{i+1} = r_{f(i)}(\sigma_i)$; then let $\Theta_1(f) = \bigcup_{i\in\mathbb{N}} \sigma_i$. Then **Ex** and **BC** are robust under Θ_1, that is, $\mathcal{S}' \subseteq \mathcal{R}_{0,1}$ is **Ex**-learnable (respectively, **BC**-learnable) iff $\Theta_1(\mathcal{S}')$ is **Ex**-learnable (respectively, **BC**-learnable).

Let $\hat{\mathcal{S}}_0$ and $\hat{\mathcal{S}}_1$ be the two **Ex**-learnable classes with $\hat{\mathcal{S}}_0 \cup \hat{\mathcal{S}}_1 \notin$ **BC** of Blum and Blum [4]. Together with Lemma 23 we now get that $\mathcal{S}_0 = (\Theta_1(\Theta_0(\hat{\mathcal{S}}_0))$ and $\mathcal{S}_1 = (\Theta_1(\Theta_0(\hat{\mathcal{S}}_1))$ are as desired. □

Lemma 25. Let $\mathcal{S}, \mathcal{S}'$ be **Ex**-learnable and T be a set of finite sequences which is closed under prefixes. Suppose that $[\mathcal{S}'] \subseteq T$, T is decidable and, for all $f \in \mathcal{S}$, $[f] \not\subseteq T$. Then $\mathcal{S} \cup \mathcal{S}'$ is **Ex**-learnable.

Proof. Let $M \in \mathcal{R}$ and $M' \in \mathcal{R}$ be **Ex**-learners for $\mathcal{S}$ and $\mathcal{S}'$, respectively. Let N be such that, for all σ,

$$N(\sigma) = \begin{cases} M'(\sigma), & \text{if } \sigma \in T; \\ M(\sigma), & \text{otherwise.} \end{cases}$$

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

Theorem 26. 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

(a) $\mathcal{S} \cup \mathcal{S}_0$ and $\mathcal{S} \cup \mathcal{S}_1$ are **Ex**-learnable;
(b) $\mathcal{S}_0 \cup \mathcal{S}_1$ is not **BC**-learnable.

Proof. Suppose there is a finite sequence $\tau \notin [\mathcal{S}]$. Let $\hat{\mathcal{S}}_0$ and $\hat{\mathcal{S}}_1$ be the two **Ex**-learnable classes with $\hat{\mathcal{S}}_0 \cup \hat{\mathcal{S}}_1 \notin$ **BC** of Blum and Blum [4]. Let

$$\mathcal{S}_0 = \{\tau \cdot f \mid f \in \hat{\mathcal{S}}_0\} \text{ and } \mathcal{S}_1 = \{\tau \cdot f \mid f \in \hat{\mathcal{S}}_1\}.$$

Then, $\mathcal{S} \cup \mathcal{S}_0$ and $\mathcal{S} \cup \mathcal{S}_1$ are each in **Ex**, but $\mathcal{S}_0 \cup \mathcal{S}_1$ is not in **BC**.

Now suppose that $[\mathcal{S}] = [\mathcal{R}]$. Let M be a total recursive **Ex**-learner for $\mathcal{S}$. We will show that there is an infinite decidable binary tree $T \subseteq [\mathcal{R}]$ such that, on all infinite paths of T, M changes its mind infinitely often. For every sequence τ, we let $r_0(\tau)$ and $r_1(\tau)$ be the first pair of incomparable extensions of τ found such that $M(\tau) \neq M(r_0(\tau))$ and $M(\tau) \neq M(r_1(\tau))$. Note that there must be such incomparable pairs of sequences, as M learns a dense set. We define a sequence $(T_i)_{i\in\mathbb{N}}$ inductively. For all i, let

$$T_0 = \{\Lambda\} \text{ and } T_{i+1} = \bigcup_{\tau\in T_i} \{r_0(\tau), r_1(\tau)\}.$$

It is now easy to see that $T = \{\gamma \mid (\exists\alpha \in \bigcup_{i\in\mathbb{N}} T_i)[\gamma \preceq \alpha]\}$ is an infinite decidable binary tree, and, on all infinite paths of T, M changes its mind infinitely often. Thus, for all $f \in \mathcal{S}$, $[f] \not\subseteq T$. By Lemma 24, there exist $\mathcal{S}_0, \mathcal{S}_1$ such that $\mathcal{S}_0$ and $\mathcal{S}_1$ are **Ex**-learnable, $[\mathcal{S}_0] \subseteq T$, $[\mathcal{S}_1] \subseteq T$ and $\mathcal{S}_0 \cup \mathcal{S}_1$ is not **BC**-learnable. As all $f \in \mathcal{S}$ satisfy $[f] \not\subseteq T$, by Lemma 25, both classes $\mathcal{S} \cup \mathcal{S}_0$ and $\mathcal{S} \cup \mathcal{S}_1$ are **Ex**-learnable. The theorem follows. □

6 Extendability

In the previous sections, the question was whether a learnable class $\mathcal{S}$ can be extended either by adding any learnable class $\mathcal{S}'$ or any function φ_e without losing learnability. In this section we ask whether a learnable class $\mathcal{S}$ can be extended effectively without losing learnability and without prescribing exactly what additions need to be done to $\mathcal{S}$. So on one hand, for a class $\mathcal{S}$, it may be easier to extend $\mathcal{S}$ to a learnable class than it being unionable, as it is not required to learn all relevant extensions of $\mathcal{S}$; on the other hand it may be harder to extend a class $\mathcal{S}$, as one has to find functions not in $\mathcal{S}$ in order to add them (while in unionability, they were given by a learner or an index). Before discussing this in detail, the next definition makes the notion of extending more precise.

Definition 27. 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.

Lemma 28. Suppose $\mathcal{C}$ is a set of learners and $\sigma_0 \in \mathbb{S}\text{eq}$. 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. There is a partial-recursive function g such that, for all e, x, one defines inductively

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

Note that $\varphi^s_{g(e,x)}$ are finite sequences of numbers. Suppose, $M_e \in \mathcal{C}$. It is easy to see that $\{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$ is infinite and can be **Ex**-learnt by a total recursive learner M'_e obtainable effectively from e. Also, there is a two-sided classifier N_e (obtainable effectively from e) for the class $\{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$. Furthermore, each $M_e \in \mathcal{C}$ fails to **Ex**-learn every $\varphi_{g(e,x)}$, $x \in \mathbb{N}$, thus $\mathbf{Ex}(M_e) \cup \{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$ is an infinite extension of $\mathbf{Ex}(M_e)$. Now, $\mathbf{Ex}(M_e) \cup \{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$ is **Ex**-learnable by learner M''_e, obtainable effectively from e, as follows:

$$M''_e(\sigma) = \begin{cases} M_e(\sigma), & \text{if } N_e(\sigma) = 0; \\ M'_e(\sigma), & \text{otherwise.} \end{cases}$$

It is easy to verify that M''_e **Ex**-learns $\mathbf{Ex}(M_e) \cup \{\varphi_{g(e,x)} \mid x \in \mathbb{N}\}$. □

Case and Fulk [5] 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 29 (Case and Fulk [5]). 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 [5] asked whether there is another effective procedure for the complement, that is, for learners that do not learn a dense class.

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. The following proposition is well-known.

Proposition 30. There exists a recursive function g such that, for all i,

(a) either $\varphi_{g(i)}$ is total or it has a finite domain of the form $\{0, 1, \ldots, z\}$ and
(b) M_i does not **BC**-learn any extension of g.

Theorem 31. 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. Suppose, by way of contradiction, that there is a recursive function h such that, for all e with $[\mathbf{Ex}(M_e)] \neq [\mathcal{R}]$, $\mathbf{Ex}(M_{h(e)})$ properly contains $\mathbf{Ex}(M_e)$. Let g be as in Proposition 30.

By Kleene's recursion theorem, there is a program e such that, for all τ,

$$M_e(\tau) = \begin{cases} g(h(e)), & \text{if } [\tau \preceq \varphi_{g(h(e))}]; \\ \mathrm{pad}(M_{h(e)}(\tau), k), & \text{if } \exists x < |\tau|\, [\varphi_{g(h(e))}(x)\downarrow \neq \tau(x)], \\ & \text{where } k \text{ is the number of inputs up to } |\tau| \\ & \text{ on which } \varphi_{g(h(e))} \text{ is defined within } |\tau| \text{ steps;} \\ \uparrow, & \text{otherwise.} \end{cases}$$

Now if M_e does not **Ex**-learn a dense set of functions, then $\mathbf{Ex}(M_{h(e)})$ must properly contain $\mathbf{Ex}(M_e)$.

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

Then M_e **Ex**-learns *only* $\varphi_{g(h(e))}$; note that the padding parameter k used is unbounded as length of τ goes to infinity. Thus, $M_{h(e)}$ **Ex**-learns $\varphi_{g(h(e))}$ by supposition. However, by Proposition 30, $M_{h(e)}$ does not **Ex**-learn $\varphi_{g(h(e))}$, a contradiction.

Case 2: $\varphi_{g(h(e))}$ is finite.

Thus, M_e is undefined on any proper extension of $\varphi_{g(h(e))}$, and, hence, does not learn a dense set. By Proposition 30, $M_{h(e)}$ also does not **Ex**-learn any total extension of $\varphi_{g(h(e))}$. Let k be the number of elements in the domain of $\varphi_{g(h(e))}$. Suppose $f \in \mathcal{R}$ does not extend $\varphi_{g(h(e))}$. Then, for all large enough j, we now have

$$M_e(f[j]) = \mathrm{pad}(M_{h(e)}(f[j]), k).$$

Thus, for large enough j, $M_e(f[j])$ is semantically equivalent to $M_{h(e)}(f[j])$. Thus, any function that is not an extension of $\varphi_{g(h(e))}$ is **Ex**-learned by $M_{h(e)}$ iff it is **Ex**-learned by M_e. Thus, $M_{h(e)}$ **Ex**-learns exactly the same class as M_e, a contradiction. □

As an immediate corollary, interesting in its own right, we get that we cannot constructively find initial segments where a given learner does not learn any extension, even not relative to the halting problem K.

Corollary 32. There is no partial K-recursive function g 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 [5] 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 31 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 33. Let $\mathcal{C}$ be the set of all total recursive 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 28, 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 finite 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 recursive 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 34. Let $m \in \mathbb{N}$. There are recursive functions h_1, h_2, h_3 such that

(a) If $I \in \{\mathbf{Conf}, \mathbf{WConf}, \mathbf{ConfRel}, \mathbf{WConfRel}\}$ and M_e is an I-learner then $M_{h_1(e)}$ is an I-learner and $I(M_{h_1(e)})$ infinitely extends $I(M_e)$;
(b) For all e, if $\mathbf{Ex}_m(M_e)$ is not empty then $\mathbf{Ex}_m(M_{h_2(e)})$ infinitely extends $\mathbf{Ex}_m(M_e)$;
(c) If $I \in \{\mathbf{Rel}, \mathbf{WRel}, \mathbf{ConfRel}, \mathbf{WConfRel}\}$ and M_e is an I-learner then $M_{h_3(e)}$ is an I-learner and $I(M_{h_3(e)})$ infinitely extends $I(M_e)$.

Proof. For part (a), note that one can, given a learner M_e, find uniformly in e a sequence of finite sequences $\sigma_0, \sigma_1, \ldots$ which converges, in the case that M_e is confident, to a limit σ_n (so $\sigma_r = \sigma_n$ for all $r \geq n$) such that $M_e(\sigma_n \cdot \tau) = M_e(\sigma_n)$ for all τ. It is easy to see that one can make effectively an I-learner N_e for the class $\{\sigma_r a^\infty \mid r, a \in \mathbb{N}\}$: Whenever N_e discovers that the current data stems from some $\sigma_r a^\infty$ it makes a mind change to this hypothesis and it keeps the hypothesis until it becomes inconsistent and then reverts to ?; as there are only finitely many different σ_r, there are also only finitely many mind changes (including those to or from ?). Using Proposition 6, there is a recursive function h_1 such that $M_{h_1(e)}$ I-learns all functions which are I-learnt either by M_e or by N_e.

For part (b), let $p \in \mathcal{R}$ be such that, for all i, σ, $\varphi_{p(i,\sigma)} = \sigma \cdot i^\infty$.

Let N_e be a total recursive learner, obtained effectively from M_e, such that (i) $\mathbf{Ex}_m(M_e) \subseteq \mathbf{Ex}_m(N_e)$, (ii) $N_e(\Lambda) = ?$ and for $\sigma \preceq \tau$, $[N_e(\sigma) \neq ?$ implies $N_e(\tau) \neq ?]$, (iii) the number of mind changes made by N_e on any function f is bounded by the number of mind changes made by M_e on f, and (iv) the number of mind changes made by N_e on any function f is bounded by m. Note that such an N_e exists and can be obtained effectively from M_e.

Let $U_e = \{\tau \mid \tau \neq \Lambda,\ N_e(\tau) \neq N_e(\tau^-)$ and $\varphi_{N_e(\tau)}(|\tau|)\downarrow\}$. Note that U_e is recursively enumerable effectively in e. Let $U_{e,s}$ denote U_e enumerated within s steps.

Let $\tau_{e,\sigma}$ denote the longest element of $U_{e,|\sigma|}$, if any, such that $\tau_{e,\sigma} \prec \sigma$. If there is no such element, then we take $\tau_{e,\sigma} = \Lambda$.

We define $h_2 \in \mathcal{R}$ such that, for all σ,

$$M_{h_2(e)}(\sigma) = \begin{cases} ?, & \text{if } U_{e,|\sigma|} = \emptyset \text{ or } N_e(\tau_{e,\sigma}) = ?; \\ N_e(\tau_{e,\sigma}), & \text{if } U_{e,|\sigma|} \neq \emptyset \text{ and } \varphi_{N_e(\tau_{e,\sigma})}(|\tau_{e,\sigma}|) = \sigma(|\tau_{e,\sigma}|); \\ p(\sigma(|\tau_{e,\sigma}|), \tau_{e,\sigma}), & \text{otherwise.} \end{cases}$$

It is easy to verify that the number of mind changes made by $M_{h_2(e)}$ on any function f is bounded by the number of mind changes made by N_e on f. To see this, let $S_f = \{\tau \in U_e \mid \tau \preceq f\}$. Then, $M_{h_2(e)}$ on f, outputs at most one new conjecture corresponding to each $\tau \in S_f$: either $N_e(\tau)$ or $p(\sigma(|\tau|), \tau)$, based on whether $\varphi_{N_e(\tau)}(|\tau|) = \sigma(|\tau|)$ or not (note that there may not be conjectures corresponding to some such τ, if that τ is not detected to be in U_e before a longer such τ' is found in U_e).

Also, $\mathbf{Ex}_m(M_{h_2(e)}) \supseteq \mathbf{Ex}_m(N_e) \supseteq \mathbf{Ex}_m(M_e)$, as for functions f $\mathbf{Ex}_m$-learnt by N_e, the longest τ such that $\tau \preceq f$ and $N_e(\tau) \neq N_e(\tau^-)$ belongs to U_e and $\varphi_{N_e(\tau)} = f$ for this τ.

Furthermore, let τ be a member of U_e such that no extension of τ is in U_e. Note that there exists such a member τ of U_e as (i) M_e (and thus N_e) $\mathbf{Ex}_m$-learns at least one function and therefore U_e contains at least one element and (ii) U_e does not contain distinct $\tau_1, \tau_2, \ldots, \tau_{m+2}$ such that $\tau_1 \prec \tau_2 \prec \ldots \prec \tau_{m+2}$ (as N_e makes at most m mind changes on any function). Then, $M_{h_2(e)}$ $\mathbf{Ex}_m$-learns τb^∞, for all b with $\varphi_{N_e(\tau)}(|\tau|) \neq b$. However, M_e $\mathbf{Ex}_m$-learns at most one extension of τ. So part (b) follows.

For part (c), consder e such that M_e is an I-learner for some class $\mathcal{S}$. For each n, define the function $f_{e,n}$ as follows:

$f_{e,n} = \bigcup_s \sigma^s_{e,n}$, where $\sigma^0_{e,n} = n$ (sequence of one element n) and
$\sigma^{s+1}_{e,n}$ is a proper extension of $\sigma^s_{e,n}$ such that $M_e(\sigma^{s+1}_{e,n}) \neq M_e(\sigma^s_{e,n})$ or $M_e(\sigma^{s+1}_{e,n}) = ?$, if such an extension can be found in s steps,
$\sigma^{s+1}_{e,n} = \sigma^s_{e,n}$, if an extension as above cannot be found in s steps.

Note that an index for $f_{e,n}$ can be found effectively from e and n. Note that if M_e is an I-learner then $f_{e,n}$ is total recursive and M_e either outputs ? infinitely often on $f_{e,n}$ or makes infinitely many mind changes on $f_{e,n}$ and thus does not I-learn $f_{e,n}$.

Let h_3 be a recursive function such that $M_{h_3(e)}(\Lambda) = ?$ and, for σ with $|\sigma| > 0$ and $n = \sigma(0)$, $M_{h_3(e)}(\sigma)$ is defined as follows:

If there exists an $x < |\sigma|$ such that $f_{e,n}(x)$ is defined within $|\sigma|$ steps (in the construction above) and $f_{e,n}(x) \neq \sigma(x)$,
Then $M_{h_3(e)}(\sigma) = M_e(\sigma)$,
Else $M_{h_3(e)}(\sigma)$ is the canonical index for $f_{e,n}$.

In the case that M_e is an I-learner, then each $f_{e,n}$ is total recursive and either the input function $f = f_{e,n}$ or there is an x such that $f_{e,n}(x){\downarrow} \neq f(x)$. Thus, $M_{h_3(e)}$ is an I-learner for $\{f_{e,n} \mid n \in \mathbb{N}\} \cup \mathbf{Ex}(M_e)$. □

7 Conclusion

The non-union theorem shows that there are $\mathbf{Ex}$-learnable classes such that their union is not even $\mathbf{BC}$-learnable. Therefore, the question arose which classes $\mathcal{S}$ are not only $\mathbf{Ex}$-learnable but also satisfy the property that the union of $\mathcal{S}$ with any other **Ex**-learnable class is still **Ex**-learnable. We distinguished the non-effective version of this notion from the effective one, where one can compute from any index e of a learner an index e' of a new learner which **Ex**-learns the class $\mathcal{S}$ and all functions learnt by the learner with index e. We separated the effective case and the non-effective case for the above notion, and gave a sufficient criterion for the effective case in terms of weakly confident and weakly reliable learning. However, we do not know whether this sufficient criterion is also necessary.

Furthermore, we also considered learning unions with classes containing at most one function. For the notion of effectively forming the union of a class $\mathcal{S}$ with a single function f (if total) given by its index d, we did not require that the index d is guaranteed to be an index of a total function; therefore additionally learning the function f (if total) requires that the learning procedure be able to deal with the case that the function with index d may be partial. For **Ex**-learning and **BC**-learning this is a nontrivial task. Again, in the case of **Ex**-learning for effectively learning unions with a single function, we showed a sufficient criterion, which is not yet known to be necessary: namely that the class $\mathcal{S}$ is weakly confidently learnable.

Furthermore, we considered the possibility of (effectively) extending learners to learn (infinitely) more functions. It was known that all **Ex**-learners learning a *dense* set of functions can be effectively extended to learn infinitely more. We solved the open problem on whether learners learning a *non-dense* set of functions can be similarly extended by showing that this is *not* possible. However, we gave 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 have analysed similar concepts also for other learning criteria.

Acknowledgements: A preliminary version of this paper appeared at the Conference on Algorithmic Learning Theory 2012 [11]. We thank the referees of ALT 2012 and of this journal for various useful comments.

References

1. Lenny Adleman and Manuel Blum. Inductive inference and unsolvability. *Journal of Symbolic Logic*, 56:891–900, 1991.
2. Kalvis Apsītis, Rūsiņs Freivalds, Raimonds Simanovskis and Juris Smotrovs. Closedness properties in Ex-identification. *Theoretical Computer Science*, 268:367–393, 2001. Special issue for ALT 1998.
3. Janis A. Bārzdiņš. Two theorems on the limiting synthesis of functions. *In Theory of Algorithms and Programs I, Latvian State University, Riga, U.S.S.R*, 210:82–88, 1974. In Russian.
4. Lenore Blum and Manuel Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.
5. John Case and Mark Fulk. Maximal machine learnable classes. *Journal of Computer and System Sciences*, 58:211–214, 1999.
6. John Case, Sanjay Jain and Susan Ngo Manguelle. Refinements of inductive inference by Popperian and reliable machines, *Kybernetika*, 30:23–52, 1994.

7. John Case, Efim Kinber, Arun Sharma and Frank Stephan. On the classification of recursive languages. *Information and Computation*, 192:15–40, 2004.
8. John Case and Carl Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
9. Lance Fortnow, William Gasarch, Sanjay Jain, Efim Kinber, Martin Kummer, Steven Kurtz, Mark Pleszkoch, Theodore Slaman, Robert Solovay and Frank Stephan. Extremes in the degrees of inferability. *Annals of Pure and Applied Logic*, 66:231–276, 1994.
10. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
11. Sanjay Jain, Timo Kötzing and Frank Stephan. Enlarging Learnable Classes. *Algorithmic Learning Theory, 23nd International Conference, 2012 (ALT 2012).* Springer LNAI 7568:66–80, 2012.
12. Sanjay Jain, Daniel Osherson, James Royer and Arun Sharma. *Systems that Learn: An Introduction to Learning Theory.* MIT Press, Cambridge, Massachusetts, second edition, 1999.
13. Martin Kummer and Frank Stephan. On the structure of degrees of inferability. *Journal of Computer and System Sciences*, 52:214–238, 1996.
14. Rolf Lindner. Algorithmische Erkennung, Dissertation B. Sektion Mathematik, Friedrich-Schiller-Universität Jena, Jena 1972.
15. Eliana Minicozzi. Some natural properties of strong-identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.
16. 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.
17. Hartley Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw Hill, New York, 1967. Reprinted in 1987.
18. Frank Stephan. On one-sided versus two-sided classification. *Archive for Mathematical Logic*, 40:489–513, 2001.
19. Arun Sharma, Frank Stephan and Yuri Ventsov. Generalized notions of mind change complexity. *Information and Computation* 189:235–262, 2004.
20. Thomas Zeugmann and Sandra Zilles. Learning recursive functions: A survey. *Theoretical Computer Science* 397:4–56, 2008. Special issue Forty Years of Inductive Inference: Dedicated to the 60th Birthday of Rolf Wiehagen.