# On Some Open Problems in Reflective Inductive Inference

Sanjay Jain [a,1]

[a] *School of Computing, National University of Singapore, Singapore 117590. Email: sanjay@comp.nus.edu.sg*

**Abstract**

In this paper we show that there exist classes of functions which can be learnt by a finite learner which reflects on its capability, but not learnable by a consistent learner which optimistically reflects on its capability. This solves the two mentioned open problems from [Gri08].

*Key words:* Theory of Computation, Inductive Inference, Reflection, Consistency.

## 1 Introduction

A function learning criterion can be described as follows. A learning machine $M$ (a computable device) receives growing segments $\sigma_0, \sigma_1, \ldots$ of the graph of a function $f$. During this process, the learner $M$ outputs a sequence of conjectures $p_0, p_1, \ldots$. The learner is said to identify $f$ if the sequence of conjectures, as above, converges to a program for $f$. This is essentially the notion of learning in the limit, also called **Ex** learning, as introduced by Gold [Gol67]. Note that in the above process, the learner does not know if and when it has converged to its final hypothesis. If we impose this additional requirement, then the model is equivalent to requiring that the learner in the above process output only one conjecture (which is a correct program for the input function). This model of learning is known as finite learning (**Fin**-learning) [Gol67].

The intermediate programs output by the learner in the **Ex**-learning model may not be consistent with the input data. Consistency is the requirement that each hypothesis conjectured by the learner, after seeing input $\sigma$, must contain the data in $\sigma$ [Bār74,BB75,WL76,JB81]. It is known that consistency

[1] Supported in part by NUS grant number R252-000-308-112.

is a restriction, that is, some **Ex**-learnable class of functions cannot be learnt by a consistent learner [Bār74].

The above models of learning, and various modifications, have been extensively studied in the literature (see for example, [AS83,JORS99,Wie78,WZ95,ZZ08]). In this paper we will restrict ourselves to the above mentioned criteria. Formal definitions of the above learning models, as well as of the models of reflection considered below, are given in Section 2.

The criteria considered above specify how the learner behaves on the functions in the class to be learnt, but specify nothing about learner's ability to indicate potential non-learnability of functions outside the class being learnt. This may be considered as a weakness from a user's point of view: the learner may lead the user into believing that learning is taking place, even though the conjectures of the learner may be completely wrong.

There have been several proposals in the literature about how a learner may recognize its own limitations. For the following fix a criterion of learning. Reliable learning [BB75,Min76] requires that the learner does not converge to a final conjecture on sequences for functions not learnt by the learner (that is, the learner signals incorrectness of its hypothesis by making a mind change). Mukouchi and Arikawa [MA93] consider the model in which a learner is required to *refute* sequences for functions which are not learnt by the learner.

Jantke [Jan95] and later Grieser [Gri08] considered learners that can reflect on their own capabilities. Such learners are equipped with a *reflection function* that indicates, in the limit, whether the input data represents a function in the class being learnt. Let $R$ be a reflection function belonging to some learner. For each finite piece of input data $\sigma$, the learner is said to *accept* $\sigma$ if $R(\sigma) = 1$; the learner is said to *reject* $\sigma$ if $R(\sigma) = 0$. Grieser [Gri08] considered three different models of reflection, constraining the behaviour of $R$ on finite pieces $\sigma$ of input data.

(1) *Exact reflection*: If the learner accepts $\sigma$, then $\sigma$ belongs to some function in the class being learnt; if the learner rejects $\sigma$, then $\sigma$ does not belong to any function in the class being learnt.
(2) *Pessimistic Reflection*: If the learner accepts $\sigma$, then $\sigma$ belongs to some function in the class being learnt; however, nothing can be said about $\sigma$ if $\sigma$ is rejected.
(3) *Optimistic Reflection*: If the learner rejects $\sigma$, then $\sigma$ does not belong to any function in the class being learnt; however, nothing can be said about $\sigma$ if $\sigma$ is accepted.

Grieser showed several interesting relationships between the above models of learning with reflection by considering how they affect **Ex**, **Fin** and consistent learning. In his study, Grieser left open several problems, including whether there exists a class of functions which can be finitely identified by a reflecting machine (without any constraints on when the input data is rejected), but which cannot be identified by a consistent learner that optimistically reflects on the input data. In this note we solve the above problem (and others) by constructing such a class of functions.

## 2 Notation and Preliminaries

Let $N$ denote the set of natural numbers, $\{0, 1, 2, \ldots\}$. Let $f$, with or without decorations, range over total functions. Let $\mathcal{R}$ denote the class of all *recursive* functions, i.e., total computable functions with arguments and values from $N$. Let $\mathcal{C}$, with or without decorations, range over subsets of $\mathcal{R}$.

Let $\varphi_0, \varphi_1, \ldots$ be a fixed acceptable numbering for all the partial recursive functions [Rog67]. Let $\Phi$ denote a standard Blum complexity measure [Blu67] for $\varphi$. By $\eta(x)\downarrow$ we denote that the partial function $\eta$ is defined on input $x$.

A sequence is a mapping from $N$ or an initial segment of $N$ to $N \times N$. The content of a sequence $\sigma$, denoted content$(\sigma)$, is the set of pairs appearing in the range of $\sigma$. Let SEQ denote the set of all finite sequences $\sigma$ such that if $(x, y)$ and $(x, z)$ belong to content$(\sigma)$, then $y = z$. (In this paper, we are only interested in sequences whose content represents part of a function. Furthermore we are only interested in the learning of total recursive functions.) If the sequence $\sigma$ contains at least $n$ elements, then we let $\sigma[n]$ denote the first $n$ elements of the sequence $\sigma$. The *canonical sequence* for a function $f$ is the infinite sequence $\sigma$ such that $\sigma(n) = (n, f(n))$.

Concatenation of two finite sequences $\sigma$ and $\tau$ is denoted by $\sigma \diamond \tau$.

A *learning machine* (also called a learner) is a (possibly partial) computable mapping from SEQ to $N \cup \{?\}$. We let $M$, with or without decorations, range over learning machines. We say that a learner $M$ converges on an infinite sequence $\sigma$ to $i$ iff for all but finitely many $n$, $M(\sigma[n]) = i$.

We now formally define the criteria of learning considered in the introduction.

**Definition 1** [Gol67]

(a) $M$ **Ex**-*identifies* a function $f$ (written: $f \in \mathbf{Ex}(M)$) iff for all infinite sequences $\sigma$ such that content$(\sigma) = \{(x, f(x)) : x \in N\}$:

(i) $M(\sigma[n])\downarrow$, for all $n \in N$, and
(ii) there exists an $i$ such that $\varphi_i = f$ and, for all but finitely many $n$, $M(\sigma[n]) = i$.

(b) $M$ **Ex**-identifies a class $\mathcal{C}$ of functions (written: $\mathcal{C} \subseteq \mathbf{Ex}(M)$) iff $M$ **Ex**-identifies each $f \in \mathcal{C}$.

(c) $\mathbf{Ex} = \{\mathcal{C} : (\exists M)[M\ \mathbf{Ex}\text{-identifies } \mathcal{C}]\}$.

**Definition 2** [Gol67] (a) $M$ **Fin**-*identifies* a function $f$ (written: $f \in \mathbf{Fin}(M)$) iff for all infinite sequences $\sigma$ such that content$(\sigma) = \{(x, f(x)) : x \in N\}$:

(i) $M(\sigma[n])\downarrow$, for all $n \in N$, and
(ii) there exist $i, n \in N$ such that $\varphi_i = f$, $M(\sigma[m]) =?$ for $m < n$, and $M(\sigma[m]) = i$ for $m \geq n$.

(b) $M$ **Fin**-identifies a class $\mathcal{C}$ of functions (written: $\mathcal{C} \subseteq \mathbf{Fin}(M)$) iff $M$ **Fin**-identifies each $f \in \mathcal{C}$.

(c) $\mathbf{Fin} = \{\mathcal{C} : (\exists M)[M\ \mathbf{Fin}\text{-identifies } \mathcal{C}]\}$.

For **Ex** and **Fin** criteria of learning, one may assume without loss of generality that the learner is total. However, for consistent learning this is not the case.

**Definition 3** [Bār74] (a) $M$ is said to be *consistent* on $f$ iff, for all $\sigma \in$ SEQ such that content$(\sigma) \subseteq f$, $\mathbf{M}(\sigma)\downarrow$ and for all $(x, y) \in$ content$(\sigma)$, $\varphi_{\mathbf{M}(\sigma)}(x) = y$.

(b) $M$ **Cons**-identifies $f$ (written: $f \in \mathbf{Cons}(M)$) iff $M$ **Ex**-identifies $f$ and $M$ is consistent on $f$.

(c) $M$ **Cons**-identifies $\mathcal{C}$ (written: $\mathcal{C} \subseteq \mathbf{Cons}(M)$) iff $M$ **Ex**-identifies $\mathcal{C}$ and $M$ is consistent on each $f \in \mathcal{C}$.

(d) $\mathbf{Cons} = \{\mathcal{C} : (\exists M)[M\ \mathbf{Cons}\text{-identifies } \mathcal{C}]\}$.

There are other notions of consistency considered in the literature (a) [BB75,WL76] $\mathcal{T}$**Cons**, in which the learner is required to be consistent on all the total functions and (b) [JB81] $\mathcal{R}$**Cons**, in which the learner is required to be total (however, here the learner is required to be consistent only on functions in the class being learnt).

**Definition 4** [Gri08] Fix a learner $M$ and a criterion of learning $\mathbf{I}$.

(a) A finite sequence $\sigma$ is *acceptable* (for $M$ with respect to $\mathbf{I}$) if content$(\sigma) \subseteq f$ for some $f \in \mathbf{I}(M)$. A finite sequence $\sigma$ is *unacceptable* (for $M$ with respect to $\mathbf{I}$) if content$(\sigma) \not\subseteq f$ for all $f \in \mathbf{I}(M)$.

(b) An infinite sequence $\sigma$ is *acceptable* (for $M$ with respect to $\mathbf{I}$) if content$(\sigma) = \{(x, f(x)) : x \in N\}$, for some function $f \in \mathbf{I}(M)$. An infinite sequence $\sigma$ is *unacceptable* (for $M$ with respect to $\mathbf{I}$) if it has an initial segment which is unacceptable.

Note that an infinite sequence may be neither acceptable nor unacceptable.

**Definition 5** [Gri08,Jan95,MA93] Fix a learner $M$ and a criterion of learning $\mathbf{I}$.

(a) A *reflection function* $R$ is a total computable function from SEQ to $\{0, 1\}$.

(b) $R$ is said to be a *reflection function of* $M$ *with respect to the learning criterion* $\mathbf{I}$ iff

(i) for all infinite acceptable sequences $\sigma$ (for $M$ with respect to $\mathbf{I}$), $R$ converges on $\sigma$ to 1.
(ii) for all infinite unacceptable sequences $\sigma$ (for $M$ with respect to $\mathbf{I}$), $R$ converges on $\sigma$ to 0.

(c) $R$ is said to be a *pessimistic reflection function* (or $p$-reflection function) of $M$ with respect to the learning criterion $\mathbf{I}$ iff $R$ is a reflection function of $M$ with respect to the learning criterion $\mathbf{I}$ and for all $\sigma \in$ SEQ such that $R(\sigma) = 1$, $\sigma$ is acceptable.

(d) $R$ is said to be an *optimistic reflection function* (or $o$-reflection function) of $M$ with respect to the learning criterion $\mathbf{I}$ iff $R$ is a reflection function of $M$ with respect to the learning criterion $\mathbf{I}$ and for all $\sigma \in$ SEQ such that $R(\sigma) = 0$, $\sigma$ is unacceptable.

(e) $R$ is said to be an *exact reflection function* (or $e$-reflection function) of $M$ with respect to the learning criterion $\mathbf{I}$ iff $R$ is a reflection function of $M$ with respect to the learning criterion $\mathbf{I}$ and for all $\sigma \in$ SEQ, $R(\sigma) = 1$ iff $\sigma$ is acceptable.

(f) $M$, along with reflection function $R$, **I-Refl** (respectively, **I-pRefl**, **I-oRefl**, **I-eRefl**) identifies $\mathcal{C}$ if $M$ **I**-identifies $\mathcal{C}$ and $R$ is a reflection function (respectively, pessimistic reflection function, optimistic reflection function, exact reflection function) of $M$ with respect to the learning criterion $\mathbf{I}$.

(g) **I-Refl** (respectively, **I-pRefl**, **I-oRefl**, **I-eRefl**) denotes the set of classes of functions $\mathcal{C}$ which can be **I-Refl** (respectively, **I-pRefl**, **I-oRefl**, **I-eRefl**) identified by some learner $M$, along with a corresponding reflection function $R$.

Note that the *acceptable* and *reflection* notions are with respect to the class learnt by the learner $M$, and not with respect to the target class $\mathcal{C}$.

Note that the notion of learning considered in the above definitions allow the input to be an arbitrary ordering of the graph of the function being learnt. Researchers have considered the situation in which only canonical sequences are given to the learner. For **Ex** and **Fin** criteria this does not make a difference, though it does make a difference for consistent learning (see for example, [BB75,WZ95,Gri08]). Moreover, for optimistic, pessimistic or exact reflection (for any of the criteria of learning considered in this paper), considering only canonical sequences does make a difference (see [Gri08]). We would not go into details of this, except mentioning that the results obtained hold in the strongest form: for the positive side, the learning can be done for arbitrary input sequences, whereas for the negative side, learning cannot be done even on canonical sequences.

## 3 Main Result

**Theorem 6 Fin-Refl − Cons-oRefl** $\neq \emptyset$.

Proof. Let $\mathcal{C}_1 = \{f : f(0) > 0,\ \varphi_{f(0)} = f \text{ and } (\forall y > 0)(\forall x < y)[\Phi_{f(0)}(x) < f(y)]\}$

and

$\mathcal{C}_2 = \{f : f(0) > 0,\ \varphi_{f(0)} = f \text{ and } (\exists n > 0)[(\forall x < n)[f(x) > 0] \text{ and } f(n) = 0 \text{ and } (\forall y > n)(\forall x < y)[\Phi_{f(0)}(x) < f(y)]]\}$.

Let $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$.

Using Kleene's parameterized recursion theorem [Soa87] we will define a recursive function $p$ (of two variables) below. It will be the case that, for all $i$ and $j$, $\varphi_{p(i,j)}(0) = p(i,j)$, and, whenever $\varphi_{p(i,j)}$ is total, $\varphi_{p(i,j)} \in \mathcal{C}$.

The class witnessing the theorem will be $\mathcal{C}' = \{\varphi_{p(i,j)} : \varphi_{p(i,j)} \text{ is total }\}$. Note that $\mathcal{C}' \subseteq \mathcal{C}$.

It is easy to verify that $\mathcal{C}' \in$ **Fin-Refl**. For any finite input sequence $\sigma \in \mathrm{SEQ}$, the learner outputs $p(i,j)$, if $i+j \leq |\sigma|$ and $(0, p(i,j)) \in \mathrm{content}(\sigma)$; the learner outputs '?' on input $\sigma$ if there exist no such $i, j$. Note that the above learner finitely learns exactly the class $\mathcal{C}'$.

Furthermore, given as input an infinite sequence $\sigma$ such that, for some total $f$, $\mathrm{content}(\sigma) = \{(x, f(x)) : x \in N\}$, one can effectively determine in the limit (a) whether $f$ is in $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ and (b) whether $f(0) = p(i,j)$, for some $i, j$. Clause (b) is easy to verify in the limit. To see (a), note that $f \notin \mathcal{C}_1$ iff $f(0) = 0$, or for some $x, y$: $x < y$ and $\Phi_{f(0)}(x) \geq f(y)$. Thus, one can determine, in the limit, whether $f \in \mathcal{C}_1$. Similar reasoning can be used to show that one can determine, in the limit, whether $f \in \mathcal{C}_2$.

Thus, in the limit one can determine if a function $f$ belongs to $\mathcal{C}'$. It follows that one can build the reflection function for the learner.

We now define the function $p$ as follows. By implicit use of the effective version of Kleene's parameterized recursion theorem [Soa87] there exists a recursive function $p$ such that $\varphi_{p(i,j)}$ may be defined as follows. We may assume without loss of generality that $p(i,j) > 0$, for all $i, j$. Intuitively, we will use $\varphi_{p(i,j)}$ to make sure that learner $M = \varphi_i$ (along with corresponding optimistic reflection function $R = \varphi_j$) does not **Cons-oRefl**-identify $\mathcal{C}'$.

Fix $i, j$. For ease of notation below, we will use $M$ and $R$ for $\varphi_i$ and $\varphi_j$ respectively.

Initially let $\varphi_{p(i,j)}(0) = p(i,j)$. Let $x_s$ denote the least $x$ such that $\varphi_{p(i,j)}(x)$ has not been defined before stage $s$. Thus $x_0 = 1$.

Below $\varphi_{p(i,j)}[x]$ denotes the initial segment $(0, \varphi_{p(i,j)}(0)), (1, \varphi_{p(i,j)}(1)), \ldots, (x-1, \varphi_{p(i,j)}(x-1))$. Below $\tau$ ranges over finite initial segments of canonical sequences. Go to stage 0.

Stage $s$

1. Dovetail steps 2 and 3, until, if ever, one of them succeeds. If ever step 2 succeeds before step 3, then go to step 4. If ever step 3 succeeds before step 2, then go to step 5.
2. Search for an extension $\tau$ of $\varphi_{p(i,j)}[x_s]$, such that $R(\tau) = 0$ and $y > 0$ for all $(x, y) \in \text{content}(\tau)$.
3. Search for a number $w > \max(\{\Phi_{p(i,j)}(x) : x < x_s\})$, such that
   $$M(\varphi_{p(i,j)}[x_s] \diamond (x_s, w))\downarrow \neq M(\varphi_{p(i,j)}[x_s])\downarrow.$$
4. (a) Let $\varphi_{p(i,j)}(y) = z$, for $(y, z) \in \text{content}(\tau)$.
   (b) Let $\varphi_{p(i,j)}(y) = 0$, for $y = |\tau|$ (i.e., for the least number $y$, such that $(y, z) \notin \text{content}(\tau)$, for all $z$).
   (c) For $y > |\tau|$, let $\varphi_{p(i,j)}(y) = 1 + \sum_{x<y} \Phi_{p(i,j)}(x)$.
   (* In this case, when step 2 succeeds, step 4 completes the definition of $\varphi_{p(i,j)}$, and we do not need any further stages. Also note that $\varphi_{p(i,j)}$ is total in this case, as $\varphi_{p(i,j)}(y)$ gets defined for all $y < |\tau|$, by step 4(a); for $y = |\tau|$, by step 4(b); and for all $y > |\tau|$ by step 4(c). *)
5. Let $\varphi_{p(i,j)}(x_s) = w$. Let $x_{s+1} = x_s + 1$. Go to stage $s + 1$.

End Stage $s$

If there exist infinitely many stages, then clearly $\varphi_{p(i,j)}$ is total. Also, due to success of step 3 and execution of step 5 in all stages, $\varphi_{p(i,j)} \in \mathcal{C}_1$ and $M$ does not **Ex**-identify $\varphi_{p(i,j)}$ (as $M$ makes infinitely many mind changes on $\varphi_{p(i,j)}$).

If, in some stage $s$, step 2 succeeds, and thus step 4 is executed, then clearly,

$\varphi_{p(i,j)} \in \mathcal{C}_2$. Also since $R$ is supposed to be optimistic, $R(\tau) = 0$ and content$(\tau)$ is contained in $\varphi_{p(i,j)}$, we have that $M$ (along with $R$) does not **Cons-oRefl** identify $\varphi_{p(i,j)}$.

If, in some stage $s$, neither step 2 nor step 3 succeed, then $\varphi_{p(i,j)}$ is not total. Furthermore, $R$ is not a reflection function for $M$. Otherwise $M$ would need to identify some extension of $\varphi_{p(i,j)}[x_s] \diamond (x_s, w)$, for all $w > \max(\{\Phi_{p(i,j)}(x) : x < x_s\})$, and thus $M$ needs to be consistent on $\varphi_{p(i,j)}[x_s] \diamond (x_s, w)$, for all $w > \max(\{\Phi_{p(i,j)}(x) : x < x_s\})$. This would imply $M(\varphi_{p(i,j)}[x_s] \diamond (x_s, w)) \neq M(\varphi_{p(i,j)}[x_s])$, for all but one $w > \max(\{\Phi_{p(i,j)}(x) : x < x_s\})$, leading to success of step 3.

Thus, we have that $M$ (along with reflection function $R$) does not **Cons-oRefl**-identify $\mathcal{C}'$. Also note that, as claimed earlier, $\varphi_{p(i,j)} \in \mathcal{C}$, whenever $\varphi_{p(i,j)}$ is total.

The theorem follows. ∎

As **Fin** $\subseteq$ $\mathcal{R}$**Cons** $\subseteq$ **Cons** (this holds also for each of the reflection models) [WZ95,Gri08], we have the following corollary.

**Corollary 7** **Fin**-**Refl** $-$ $\mathcal{R}$**Cons**-**oRefl** $\neq \emptyset$.

$\mathcal{R}$**Cons**-**Refl** $-$ **Cons**-**oRefl** $\neq \emptyset$.

$\mathcal{R}$**Cons**-**Refl** $-$ $\mathcal{R}$**Cons**-**oRefl** $\neq \emptyset$.

The above solves the two mentioned open problems in the paper [Gri08].

**Acknowledgements.** We thank the referees for pointing out an inconsistency in the usage of the notion of reflection in an earlier version of the paper. We also thank the referees for several helpful comments and reformulations which improved the presentation of the paper.

## References

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

[Bār74] J. Bārzdiņš. Inductive inference of automata, functions and programs. In *Proceedings of the 20th International Congress of Mathematicians, Vancouver*, pages 455–560, 1974. In Russian. English translation in American Mathematical Society Translations: Series 2, 109:107-112, 1977.

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

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

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

[Gri08] G. Grieser. Reflective inductive inference of recursive functions. *Theoretical Computer Science A*, 397(1–3):57–69, 2008. Special Issue on Forty Years of Inductive Inference. Dedicated to the 60th Birthday of Rolf Wiehagen.

[Jan95] K. P. Jantke. Reflecting and self-confident inductive inference machines. In K. Jantke, T. Shinohara, and T. Zeugmann, editors, *Algorithmic Learning Theory: 6th International Workshop (ALT '95)*, volume 997 of *Lecture Notes in Artificial Intelligence*, pages 282–297. Springer-Verlag, 1995.

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

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

[MA93] Y. Mukouchi and S. Arikawa. Inductive inference machines that can refute hypothesis spaces. In K.P. Jantke, S. Kobayashi, E. Tomita, and T. Yokomori, editors, *Algorithmic Learning Theory: 4th International Workshop (ALT '93)*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 123–136. Springer-Verlag, 1993.

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

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

[Soa87] R. Soare. *Recursively Enumerable Sets and Degrees.* Springer-Verlag, 1987.

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

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

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

[ZZ08] T. Zeugmann and S. Zilles. Learning recursive functions: A survey. *Theoretical Computer Science A*, 397(1–3):4–56, 2008. Special Issue on Forty Years of Inductive Inference. Dedicated to the 60th Birthday of Rolf Wiehagen.