

# On the Synthesis of Strategies Identifying Recursive Functions

Sandra Zilles

Fachbereich Informatik  
Universität Kaiserslautern  
Postfach 3049  
D - 67653 Kaiserslautern  
zilles@informatik.uni-kl.de

**Abstract.** A classical learning problem in Inductive Inference consists of identifying each function of a given class of recursive functions from a finite number of its output values. Uniform learning is concerned with the design of single programs solving infinitely many classical learning problems. For that purpose the program reads a description of an identification problem and is supposed to construct a technique for solving the particular problem.

As can be proved, uniform solvability of collections of solvable identification problems is rather influenced by the description of the problems than by the particular problems themselves. When prescribing a specific inference criterion (for example learning in the limit), a clever choice of descriptions allows uniform solvability of all solvable problems, whereas even the most simple classes of recursive functions are not uniformly learnable without restricting the set of possible descriptions. Furthermore the influence of the hypothesis spaces on uniform learnability is analysed.

## 1 Introduction

Inductive Inference is concerned with methods of identifying objects in a target class from incomplete information. The learning model is based on a recursion-theoretic background, i.e. target objects as well as learners are represented by computable functions. From an input sequence consisting of finite subgraphs of the graph of a target function the learner produces a sequence of hypotheses interpreted as indices of functions enumerated by a partial-recursive numbering. In the initial approach of identification in the limit introduced by Gold in [6] that sequence of hypotheses is supposed to converge to a correct index of the target function. Several further identification criteria have been introduced and analysed in [2], [3] and [5]. In general, a learning problem is given by

- a class  $U$  of recursive functions,
- a hypothesis space  $\psi$  and
- an identification criterion  $I$ .

The aim is to find a learner identifying *each* function in the class  $U$  with respect to  $\psi$  within the scope of the criterion  $I$ .

Now imagine a collection of infinitely many learning problems solvable according to a given criterion. Uniform Inductive Inference is concerned with the question, whether there exists a single program which – given a description of a special learning problem of our collection – synthesizes an appropriate learner solving the actual problem. Such a program may be interpreted as a very “intelligent” learner able to simulate infinitely many learners of the classical type. Instead of tackling each problem in a specific way we want to use a kind of uniform strategy coping with the whole accumulation of problems.

Jantke’s work [7] is concerned with the uniform identification of classes of recursive functions in the limit, particularly for the case that in each learning step the intermediate hypothesis generated by the learner is consistent with the information received up to the actual time of the learning process. Jantke proved that his model of uniform identification does not allow the synthesis of a program learning a class consisting of just a single recursive function, as long as the synthesizer is supposed to cope with any possible description of such a class. His negative result indicates that the concept of uniform learning might be rather fruitless. But this suggestion is mitigated by the results on uniform identification of classes of languages – a concept which is studied for example in [10], [9] and [4]. Especially [4] contains lots of positive results allowing a more optimistic point of view concerning the fruitfulness of the uniform identification model. The work of Osherson, Stob and Weinstein additionally deals with several possibilities for the description of learning problems.

The present paper provides its own definition of uniform identifiability with the special feature that any of the learning problems described may be solved with respect to any appropriate hypothesis space without requiring the synthesis of the particular hypothesis spaces. The first result in Section 4 shows the existence of a special set of descriptions accumulating all learning problems solvable according to a given criterion  $I$ , such that synthesizing learners successful with respect to  $I$  is possible. The trick is to encode programs for the learners within the descriptions. Of course in general such tricks should be avoided, for example by fixing the set of descriptions in advance. But then it is still possible to use tricks by a clever choice of the hypothesis spaces. The results in Section 5 show that such tricks provide a uniform strategy for behaviourally correct identification<sup>1</sup> of any class learnable according to that criterion, even coping with any description of such a class. Nevertheless the free choice of the hypothesis spaces does not trivialize uniform learning in the limit. For example the collection of all descriptions of classes consisting of just two recursive functions is not suitable in that sense, i.e. there is no uniform strategy constructing a successful program for learning in the limit from such a description. Unfortunately, those results are rather negative: either uniform learnability is achieved by tricks and thus becomes trivial or it cannot be achieved at all. When fixing the hypothesis spaces in advance, our situation gets even worse. Jantke’s result is strengthened

---

<sup>1</sup> For the definitions of inference criteria mentioned here see Section 2.

by proving that there is no uniform learner for behaviourally correct identification with respect to an acceptable numbering coping with all descriptions of sets of just one recursive function. The same collection of learning problems becomes unsolvable even for behaviourally correct identification with anomalies, if we further tighten our demands concerning the hypothesis spaces.

On the other hand we also present some quite positive results, which at least seem to justify some further research on uniform learning. For example, if the descriptions of the learning problems fulfill some special topological conditions, one can uniformly construct strategies learning the corresponding classes in the limit – even with total and consistent intermediate hypotheses<sup>2</sup>. Results of this kind strongly substantiate the suggestion that uniform identification is indeed a model of rich learning power. The negative results mentioned above just have to be interpreted carefully. The reason for the failure of uniform learners is most often not a substantial lack of power, but lies in the choice of unsuitable descriptions. So our model really seems worthy of investigation. In general, Section 2 provides preliminaries and Section 3 deals with the notion of uniform identification as well as basic results. General results are presented in Section 4, followed by results on particular description sets in Section 5. Finally Section 6 is concerned with the influence of the choice of the hypothesis spaces. We also transfer Wiehagen’s characterizations of classes identifiable in the limit (see [13]) to the case of uniform learning.

## 2 Preliminaries

First we fix some notions that will be used in this paper. All conceptions in the context of recursion theory not explicitly introduced here can be found in [12].

We denote the set of (nonnegative) integers by  $\mathbb{N}$  and write  $\mathbb{N}^*$  for the set of all finite tuples of elements of  $\mathbb{N}$ . If  $n$  is any integer, we refer to the set of all  $n$ -tuples of integers by  $\mathbb{N}^n$ . By means of a bijective and computable mapping from  $\mathbb{N}^n$  onto  $\mathbb{N}$  we identify  $n$ -tuples of integers with elements in  $\mathbb{N}$ . Between  $\mathbb{N}^*$  and  $\mathbb{N}$  we also choose a bijective, computable mapping and denote it by  $\text{cod} : \mathbb{N}^* \rightarrow \mathbb{N}$ . Thus we may use  $\alpha$  to refer to  $\text{cod}(\alpha)$ , where  $\alpha \in \mathbb{N}^*$ . The quantifiers  $\forall$  and  $\exists$  are used in the common way. Quantifying an expression with  $\forall^\infty n$  indicates that the expression is true for all but finitely many  $n \in \mathbb{N}$ .

Inclusion of sets is expressed by the symbol  $\subseteq$ , proper inclusion by  $\subset$ .  $\text{card } X$  serves as a notation for the cardinality of a set  $X$ , and we write  $\text{card } X = \infty$ , whenever  $X$  is an infinite set. The set of all subsets of  $X$  is referred to by  $\wp X$ .

$\mathcal{P}^n$  denotes the class of all partial-recursive functions of  $n$  variables. Its subclass of total functions (called recursive functions) is denoted by  $\mathcal{R}^n$ . Whenever the number of arguments is of no special interest, we omit the superscripts. For any function  $f \in \mathcal{P}$  and any integer  $n$  the notation  $f[n]$  refers to the coding  $\text{cod}(f(0), \dots, f(n))$  of the initial segment of length  $n + 1$  of  $f$ , as long as the values  $f(0), \dots, f(n)$  are all defined. For  $f \in \mathcal{P}$  and  $x \in \mathbb{N}$  we write  $f(x)\downarrow$ , if  $f$  is defined for the argument  $x$ ;  $f(x)\uparrow$ , if  $f$  is not defined for the argument  $x$ .

<sup>2</sup> See Section 6 for definitions.

For piecewise comparison of two functions  $f, g \in \mathcal{P}$  we agree on the notation  $f =_n g$ , if  $\{(x, f(x)) \mid x \leq n \text{ and } f(x) \downarrow\} = \{(x, g(x)) \mid x \leq n \text{ and } g(x) \downarrow\}$ ; otherwise  $f \neq_n g$ . If the set of arguments on which the functions  $f, g \in \mathcal{P}$  disagree is finite, i.e. if  $[[f(n) \uparrow \wedge g(n) \uparrow] \text{ or } [f(n) \downarrow \wedge g(n) \downarrow \wedge f(n) = g(n)]]$  for all but finitely many  $n \in \mathbb{N}$ , we write  $f =^* g$ .

A function  $f$  may be identified with the sequence  $(f(n))_{n \in \mathbb{N}}$  of its values, which yields notations like e.g.  $f = 0^k \uparrow^\infty$  or  $g = 0^k 12^\infty$ . A finite tuple  $\alpha \in \mathbb{N}^*$  is often regarded as the function  $\alpha \uparrow^\infty$  implicitly. For two functions  $f, g$  the notation  $f \sqsubseteq g$  means that  $\{(x, f(x)) \mid x \in \mathbb{N}, f(x) \downarrow\} \subseteq \{(x, g(x)) \mid x \in \mathbb{N}, g(x) \downarrow\}$ .

Any  $\psi \in \mathcal{P}^{n+1}$  ( $n \in \mathbb{N}$ ) is used as a numbering for the set  $\mathcal{P}_\psi := \{\psi_i \mid i \in \mathbb{N}\}$  by means of the definition  $\psi_i(x) := \psi(i, x)$  for all  $i \in \mathbb{N}$ ,  $x \in \mathbb{N}^n$ . The index  $i \in \mathbb{N}$  is called  $\psi$ -number of the function  $\psi_i$ .

Given  $\psi \in \mathcal{P}^{n+2}$  ( $n \in \mathbb{N}$ ), every integer  $b \in \mathbb{N}$  “describes” a partial-recursive numbering, which we will denote by  $\psi^b$ . We set  $\psi^b(i, x) := \psi(b, i, x)$  for all  $i \in \mathbb{N}$ ,  $x \in \mathbb{N}^n$  and thus write by analogy with the notations above:  $\psi_i^b(x) := \psi^b(i, x)$  for all  $i \in \mathbb{N}$ ,  $x \in \mathbb{N}^n$ .

For any  $\psi \in \mathcal{P}^{n+1}$ ,  $n \in \mathbb{N}$  we will often refer to the entirety of total functions in  $\mathcal{P}_\psi$ , which will be called the “recursive core” or “ $\mathcal{R}$ -core” of  $\mathcal{P}_\psi$  (abbreviated by  $\mathcal{R}_\psi$ ). Hence  $\mathcal{R}_\psi = \mathcal{R} \cap \mathcal{P}_\psi$ .

Identification in the limit<sup>3</sup> provides the fundamentals for learning models examined in Inductive Inference and has first been analysed by Gold in [6].

**Definition 1.** Let  $U \subseteq \mathcal{R}$ ,  $\psi \in \mathcal{P}^2$ . The class  $U$  is an element of  $EX_\psi$  and called identifiable in the limit with respect to  $\psi$  iff there is a function  $S \in \mathcal{P}$  (called strategy) such that for any  $f \in U$ :

1.  $\forall n \in \mathbb{N} [S(f[n]) \downarrow] \quad (S(f[n]) \text{ is called hypothesis on } f[n]),$
2.  $\exists j \in \mathbb{N} [\psi_j = f \text{ and } \forall^\infty n [S(f[n]) = j]].$

If  $S$  is given, we also write  $U \in EX_\psi(S)$ . We set  $EX := \bigcup_{\psi \in \mathcal{P}^2} EX_\psi$ .

On every function  $f \in U$  the strategy  $S$  generates a sequence of indices converging to a  $\psi$ -number of  $f$ . [13] supplies the following characterization of the classes learnable in the limit, which will be useful for us later on.

**Theorem 1.** Let  $U \subseteq \mathcal{R}$ .  $U \in EX$  iff there exist  $\psi \in \mathcal{P}^2$  and  $d \in \mathcal{R}^2$  such that

1.  $U \subseteq \mathcal{P}_\psi,$
2.  $\forall i, j \in \mathbb{N} [i \neq j \Rightarrow \psi_i \neq_{d(i,j)} \psi_j].$

If we omit the demand for convergence to a single hypothesis, we talk of “behaviourally correct” identification, defined for example in [2].

**Definition 2.** Let  $U \subseteq \mathcal{R}$ ,  $\psi \in \mathcal{P}^2$ .  $U$  is called BC-identifiable wrt  $\psi$  iff there exists an  $S \in \mathcal{P}$ , such that for all  $f \in U$  the following conditions are fulfilled:

1.  $\forall n \in \mathbb{N} [S(f[n]) \downarrow],$

<sup>3</sup> We also use the term “explanatory identification”, abbreviated by EX-identification.

2.  $\forall^\infty n [\psi_{S(f[n])} = f]$ .

We also write  $U \in BC_\psi(S)$  and define  $BC_\psi$  and  $BC$  as usual.

BC-identifiability has also been characterized in [13]. But for our purpose the following characterization proved in [11] is more useful.

**Theorem 2.** *Let  $U \subseteq \mathcal{R}$ .  $U \in BC$  iff there exist  $\psi \in \mathcal{P}^2$  and  $d \in \mathcal{R}^2$  satisfying*

1.  $U \subseteq \mathcal{P}_\psi$ ,
2.  $\forall i, j \in \mathbb{N} [\psi_i = \psi_j \iff \psi_i =_{d(i,j)} \psi_j]$ .

Though BC-identification provides more learning power than EX-identification – a proof can be found in [2] – there are still classes of recursive functions which are not in BC. In [5] we find a variation of BC-identification, which allows learnability of the whole class  $\mathcal{R}$ .

**Definition 3.** *Let  $U \subseteq \mathcal{R}$ ,  $\psi \in \mathcal{P}^2$ .  $U$  is called BC-identifiable with finitely many anomalies wrt  $\psi$  iff there exists an  $S \in \mathcal{P}$ , such that for all  $f \in U$  the following conditions are fulfilled:*

1.  $\forall n \in \mathbb{N} [S(f[n]) \downarrow]$ ,
2.  $\forall^\infty n [\psi_{S(f[n])} =^* f]$ .

We write  $U \in BC_\psi^*(S)$  and use the notations  $BC_\psi^*$  and  $BC^*$  in the usual way.

From now on let  $\mathcal{I} := \{\text{EX}, \text{BC}, \text{BC}^*\}$  denote the set of all previously declared inference criteria. The following results have been proved (see [2] and [5]):

**Theorem 3.**  $\text{EX} \subset \text{BC} \subset \text{BC}^* = \wp \mathcal{R}$ .

### 3 Uniform Learning – Model and Basic Results

Throughout this paper let  $\varphi \in \mathcal{P}^3$  be a fixed acceptable numbering of  $\mathcal{P}^2$ . If we choose a number  $b \in \mathbb{N}$ , we may interpret it as an index for the partial-recursive numbering  $\varphi^b \in \mathcal{P}^2$ , which assigns the value  $\varphi(b, x, y)$  to each pair  $(x, y)$  of integers. Thus we can regard  $b$  as a description of a class of recursive functions, namely the recursive core of  $\mathcal{P}_{\varphi^b}$ . We will denote this class by  $\mathcal{R}_b$ , i.e.  $\mathcal{R}_b := \mathcal{R}_{\varphi^b} = \mathcal{R} \cap \mathcal{P}_{\varphi^b}$  for  $b \in \mathbb{N}$ . Similarly each set  $B \subseteq \mathbb{N}$  describes a set  $\mathcal{R}_B := \{\mathcal{R}_b \mid b \in B\}$  of classes of recursive functions.

If  $V$  and  $W$  are sets of sets, we will write  $V \preceq W$  if and only if for all  $X \in V$  there exists a set  $Y \in W$  such that  $X \subseteq Y$ .

**Definition 4.** *Let  $I, I'$  be elements of  $\mathcal{I}$  satisfying  $I \subseteq I'$ . A set  $J \subseteq \wp \mathcal{R}$  of classes of recursive functions is said to be uniformly learnable with respect to  $I$  and  $I'$  iff there exists a set  $B \subseteq \mathbb{N}$  such that the following conditions are fulfilled:*

1.  $J \preceq \mathcal{R}_B$ ,
2.  $\mathcal{R}_B \subseteq I$ ,
3.  $\exists S \in \mathcal{P}^2 \forall b \in B \exists \psi \in \mathcal{P}^2 [\mathcal{R}_b \in I'_\psi(\lambda x. S(b, x))]$ .

We refer to this definition by  $J \in \text{uni}(I, I')$ .

So  $J \in \text{uni}(I, I')$  iff there is a set  $B$  of indices of numberings such that

- every class in  $J$  is contained in some recursive core  $\mathcal{R}_b$  corresponding to an index  $b \in B$ ;
- every recursive core  $\mathcal{R}_b$  described by some  $b \in B$  is learnable under the criterion  $I$ ;
- there is a uniform strategy  $S$  which, given  $b \in B$ , learns  $\mathcal{R}_b$  under the criterion  $I'$  with respect to some appropriate hypothesis space  $\psi$ .

The set  $B$  is called *description set* for  $J, I, I'$ . We also write  $J \in \text{uni}_B(I, I')$ ,  $J \in \text{uni}(I, I')(S)$  or  $J \in \text{uni}_B(I, I')(S)$ , whenever the description set  $B$ , the uniform strategy  $S$  or both of them are fixed.

In order to prove the uniform learnability of a subset  $J \subseteq \wp\mathcal{R}$  wrt  $I, I' \in \mathcal{I}$  we first have to specify the set  $B \subseteq \mathbb{N}$  describing the classes to be learned, secondly the (possibly distinct) numberings  $\psi \in \mathcal{P}^2$  serving as hypothesis spaces for the particular classes  $\mathcal{R}_b$  ( $b \in B$ ) and finally the strategy  $S \in \mathcal{P}^2$  designed to do the actual “learning job”. Starting from this point of view two main questions arise:

1. Which classes  $J \subseteq \wp\mathcal{R}$  are uniformly learnable wrt given criteria  $I, I'$  at all?
2. Which classes  $J \subseteq \wp\mathcal{R}$  remain learnable in the sense of  $\text{uni}(I, I')$ , if we specify in advance one of the parameters mentioned above?

Of course these questions are much too general to be answered exhaustively in this paper. Nevertheless some characterizations and interesting special cases are considered.

On condition that  $J \in \text{uni}_B(I, I')$  we obviously obtain  $\mathcal{R}_B \in \text{uni}_B(I, I')$ . As all classes in  $\mathcal{I}$  are closed under inclusion, we also verify that  $\mathcal{R}_B \in \text{uni}_B(I, I')$  implies  $J \in \text{uni}_B(I, I')$  for all  $J \preceq \mathcal{R}_B$ . Therefore the sets  $J \in \text{uni}(I, I')$  are characterized by those description sets  $B \subseteq \mathbb{N}$  which are suitable for uniform learning of *some* set  $J' \subseteq \wp\mathcal{R}$ :

**Lemma 1.** *Assume  $I, I' \in \mathcal{I}$ ,  $J \subseteq \wp\mathcal{R}$ . Then  $J \in \text{uni}(I, I')$  if and only if there exists a set  $B \subseteq \mathbb{N}$  satisfying  $\mathcal{R}_B \in \text{uni}_B(I, I')$  and  $J \preceq \mathcal{R}_B$ .*

For that reason the appropriate description sets for uniform learning are of particular interest for our further research. Now consider a set  $\mathcal{R}_B$  of recursive cores described by a set  $B \subseteq \mathbb{N}$ . The mere statement that  $\mathcal{R}_B \in \text{uni}(I, I')$  for some  $I, I' \in \mathcal{I}$  does *not* imply the uniform learnability of  $\mathcal{R}_B$  wrt  $I, I'$  from  $B$ . It is quite conceivable that  $\mathcal{R}_B$  might be uniformly learnable from a description set  $B' \subseteq \mathbb{N}$ , but *not* from the description set  $B$ . This would as well involve that *no* set  $J \subseteq \wp\mathcal{R}$  was uniformly learnable wrt  $I, I'$  from description set  $B$  at all; thus we might consider the description set  $B$  to be unsuitable for uniform learning with respect to  $I$  and  $I'$ .

**Definition 5.** *Let  $I, I' \in \mathcal{I}$ ,  $B \subseteq \mathbb{N}$ . The description set  $B$  is said to be suitable for uniform learning with respect to  $I$  and  $I'$  if  $\mathcal{R}_B \in \text{uni}_B(I, I')$ . The class of all description sets suitable in that sense will be denoted by  $\text{suit}(I, I')$ .*

These considerations raise the question whether there are certain specific properties characterizing our appropriate description sets  $B \in \text{suit}(I, I')$ .

**Definition 6.** Fix  $I, I' \in \mathcal{I}$ ,  $h : \mathbb{N} \rightarrow \mathbb{N}$ . A set  $J \subseteq \wp\mathcal{R}$  is called uniformly learnable wrt  $I$  and  $I'$  by the interpretation function  $h$ , iff there exist  $B \subseteq \mathbb{N}$  and  $S \in \mathcal{P}^2$  such that:

1.  $J \in \text{uni}_B(I, I')(S)$ ,
2.  $\forall b \in B [\mathcal{R}_b \in I'_{\varphi^{h(b)}}(\lambda x.S(b, x))]$ .

We abbreviate this formulation by  $J \in \text{uni}_{[h]}(I, I')$ . If additionally there is a numbering  $\tau \in \mathcal{P}^2$  satisfying  $\varphi^{h(b)} = \tau$  for all  $b \in \mathbb{N}$ , we write  $J \in \text{uni}_\tau(I, I')$  instead.

Note that the interpretation function  $h$  in our definition is not necessarily computable or total. Of course we might wish to fix both our hypothesis spaces by means of an interpretation function  $h$  and our description set  $B$  in advance. In that case we use the notions  $\text{uni}_{B, [h]}(I, I')$  as well as  $\text{uni}_{B, \tau}(I, I')$  by analogy. In the usual way we may also refer to fixed uniform strategies in our notations. Via the function  $h$  each description  $b \in B$  obtains an associated hypothesis space  $\varphi^{h(b)}$ , by means of which we can interpret the hypotheses produced by the strategy  $\lambda x.S(b, x)$ . Regarding practical aspects we are interested especially in computable interpretation functions, such as for example the identity function  $id : \mathbb{N} \rightarrow \mathbb{N}$  defined by  $id(x) = (x)$  for all  $x \in \mathbb{N}$ .

We are now able to formulate some basic results on uniform learning. Although the corresponding proofs are quite simple, these results will be useful for our further examinations. First we state a necessary condition for uniform learnability of a subset of  $\wp\mathcal{R}$ . For the proof of Proposition 1 note that all classes  $I \in \mathcal{I}$  are closed with respect to the inclusion of sets.

**Proposition 1.** Let  $I, I' \in \mathcal{I}$ ,  $I \subseteq I'$ ,  $J \subseteq \wp\mathcal{R}$ . If  $J \in \text{uni}(I, I')$ , then  $J \subseteq I$ .

*Proof.* Let  $J \in \text{uni}(I, I')$ . Then there is a set  $B \subseteq \mathbb{N}$  which fulfills  $J \in \text{uni}_B(I, I')$ . This implies  $\mathcal{R}_B \subseteq I$  and  $J \preceq \mathcal{R}_B$ . Thus for all  $U \in J$  there exists  $b \in B$  such that  $U \subseteq \mathcal{R}_b \in I$ . So  $J \subseteq I$ , because  $I$  is closed under inclusion.  $\square$

From Proposition 1 and the definition of uniform learning we conclude:

**Corollary 1.** Let  $I, I' \in \mathcal{I}$ ,  $I \subseteq I'$ . Then  $\text{uni}(I, I) \subseteq \text{uni}(I, I') \subseteq \wp I$ .

Any strategy identifying a class  $U \subseteq \mathcal{R}$  with respect to some criterion  $I \in \mathcal{I}$  can be replaced by a *total* recursive strategy without loss of learning power. This new strategy is defined by computing the values of the old strategy for a bounded number of steps and a bounded number of input examples with increasing bounds. As long as there is no hypothesis found, some temporary hypothesis is produced. Afterwards the hypotheses of the former strategy are put out “with delay”. Now we transfer these observations to the level of uniform learning and get:

**Proposition 2.** *Let  $I, I' \in \mathcal{I}$ ,  $B \subseteq \mathbb{N}$  and let  $h : \mathbb{N} \rightarrow \mathbb{N}$  be any function. Assume  $\mathcal{R}_B \in \text{uni}_{B, [h]}(I, I')$ . Then there exists a total recursive function  $S$  satisfying  $\mathcal{R}_B \in \text{uni}_{B, [h]}(I, I')(S)$ .*

We will use this result implicitly in the forthcoming proofs.

**Proposition 3.** *Let  $I \in \mathcal{I}$ ,  $h : \mathbb{N} \rightarrow \mathbb{N}$ . If  $B \subseteq \mathbb{N}$  is a finite set with  $\mathcal{R}_b \in I_{\varphi^{h(b)}}$  for all  $b \in B$ , then  $\mathcal{R}_B \in \text{uni}_{B, [h]}(I, I)$ .*

The proof is obvious: a finite number of strategies – each learning one of the given  $\mathcal{R}$ -cores wrt  $I$  – can be merged to a single computable uniform strategy.

## 4 Uniform Learning without Specification of the Model Parameters

First we deal with uniform learning according to Definition 4 without specifying the description set  $B$  or the hypothesis spaces  $\psi$  in advance. We choose two inference criteria  $I, I' \in \mathcal{I}$  satisfying  $I \subseteq I'$  and try to characterize the subsets  $J \subseteq \varphi\mathcal{R}$  contained in  $\text{uni}(I, I')$ . From Corollary 1 we already know that these sets must be subsets of  $I$ . Now we will even prove the sufficiency of that simple condition.

**Theorem 4.** *If  $I, I' \in \mathcal{I}$ ,  $I \subseteq I'$ , then  $\text{uni}(I, I) = \text{uni}(I, I') = \varphi I$ .*

*Proof.* Assume  $I, I' \in \mathcal{I}$ ,  $I \subseteq I'$ . Applying Corollary 1 we only have to show  $\varphi I \subseteq \text{uni}(I, I)$ . Assume  $J \in \varphi I$ . For each  $U \in J$  there is a numbering  $\psi \in \mathcal{P}^2$  such that  $U \subseteq \mathcal{P}_\psi$  and  $\mathcal{R}_\psi \in I$  (a proof can be derived from  $\mathcal{R} \in \text{BC}^*$  and the characterizations in Theorems 1 and 2 quite easily). Let  $C \subseteq \mathbb{N}$  be the set of  $\varphi$ -indices of all these numberings. Fix an acceptable numbering  $\tau \in \mathcal{P}^2$ .

For each  $c \in C$  there is a  $\tau$ -index  $k_c$  of a strategy  $S_c \in \mathcal{P}$ , which identifies  $\mathcal{R}_c$  in the limit with respect to  $\tau$  according to the criterion  $I$ .

These  $\tau$ -indices can now be coded within our hypothesis spaces  $\varphi^c$  by simply integrating the function  $k_c \uparrow^\infty$  into the numberings. Thus we achieve that our new numberings obtain two very useful properties. Firstly, their recursive cores are learnable with respect to the criterion  $I$ , because we do not change the recursive cores by integrating functions of the shape  $k_c \uparrow^\infty$ . Secondly, they contain  $\tau$ -indices for strategies identifying their recursive cores according to  $I$ .

Our suitable description set is the set of all indices of numberings achieved by modification of the numberings described by  $C$ . A uniform  $I$ -learner for the target class  $J$  just has to read the indices of the particular strategies and then simulate their jobs with the help of the functions associated by  $\tau$ . We obtain  $J \in \text{uni}(I, I)$  as claimed.  $\square$

Note that we even obtain  $\text{uni}_\tau(I, I') = \text{uni}(I, I') = \varphi I$  for any acceptable numbering  $\tau$ . Now we can easily compare the power of uniform learning criteria resulting in the choice of particular criteria  $I, I' \in \mathcal{I}$ :



**Corollary 2.** *Let  $I, I' \in \mathcal{I}$ ,  $I \subset I'$ . Then  $\text{uni}(I, I) = \text{uni}(I, I') \subset \text{uni}(I', I')$ .*

*Proof.* By Theorem 4 we know  $\text{uni}(I, I) = \text{uni}(I, I')$ . As  $\text{uni}(I, I') \subseteq \text{uni}(I', I')$ , it remains to prove that  $\text{uni}(I', I')$  is *not* a subset of  $\text{uni}(I, I')$ . For this purpose we simply choose any class  $U \in I' \setminus I$  and obtain a class  $J \in \text{uni}(I', I') \setminus \text{uni}(I, I')$  by defining  $J := \{U\}$ .  $\square$

Intuitively, our uniform strategy defined in the proof of Theorem 4 does not really *learn* anything, because the programs for learning the described classes are coded within the described numberings in advance. In the following sections we will see some more examples for such easy “tricks” simplifying the work of uniform strategies. But as we will see later, there are also non-trivial sets of classes of recursive functions uniformly learnable by really “labouring” strategies.

## 5 Uniform Learning from Special Description Sets

Now we investigate the suitability of given description sets  $B$ , i.e. the uniform learnability of  $\mathcal{R}_B$  from  $B$  wrt some criteria  $I, I' \in \mathcal{I}$ . We start with a simple but useful observation.

**Proposition 4.** *Let  $I, I' \in \mathcal{I}$  and  $B \subseteq \mathbb{N}$  such that  $\mathcal{R}_b \in I$  for all  $b \in B$ . If  $\bigcup_{b \in B} \mathcal{R}_b \in I'$ , then  $B \in \text{suit}(I, I')$ .*

The proof of Proposition 4 is straightforward from the definitions. As a direct consequence we obtain a simple characterization of the description sets suitable for uniform learning with  $\text{BC}^*$ -strategies:

**Theorem 5.**  $\text{suit}(I, \text{BC}^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$  for  $I \in \mathcal{I}$ . In particular  $\text{suit}(\text{BC}^*, \text{BC}^*) = \emptyset \mathbb{N}$ .

Since  $\text{BC}$  is not closed under union of sets, the proof of a corresponding characterization for  $\text{suit}(I, \text{BC})$  cannot be based on Proposition 4. Instead – as in the proof of Theorem 4 – we make use of special “tricks”, such that the resulting strategy does not really have to do any work.

**Theorem 6.** *Set  $B := \{b \in \mathbb{N} \mid \mathcal{R}_b \in \text{BC}\}$ . Then  $B \in \text{suit}(\text{BC}, \text{BC})$  and thus  $\text{suit}(I, \text{BC}) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$  for all  $I \in \{\text{EX}, \text{BC}\}$ .*

*Proof.* Fix an acceptable numbering  $\tau \in \mathcal{P}^2$ . Each class learnable in the sense of  $\text{BC}$  can be identified wrt  $\tau$  by a *total* strategy, i.e. for all  $b \in B$  there is an  $S_b \in \mathcal{R}$  such that  $\mathcal{R}_b \in \text{BC}_\tau(S_b)$ .

Given any element  $b \in B$  we can now list *all* hypotheses produced by  $S_b$  on *all* initial segments of recursive functions in a computable way. If we interpret these hypotheses as  $\tau$ -indices, we obtain a numbering of all candidate functions suggested by  $S_b$ .

More formally: for each  $b \in B$  we define  $\psi^{[b]} \in \mathcal{P}^2$  by  $\psi_i^{[b]}(x) := \tau_{S_b(i)}(x)$  for any  $i, x \in \mathbb{N}$ . If  $f \in \mathcal{R}$ ,  $n \in \mathbb{N}$ , then the index  $f[n]$  via  $\psi^{[b]}$  represents exactly

the function “suggested” by  $S_b$  on input  $f[n]$ . This property can be used by a uniform BC-strategy: for  $b, n \in \mathbb{N}$  and  $f \in \mathcal{R}$  we set  $S(b, f[n]) := f[n]$  and obtain

$$\psi_{S(b, f[n])}^{[b]} = \psi_{f[n]}^{[b]} = \tau_{S_b(f[n])} \text{ for any } b \in B, f \in \mathcal{R}_b, n \in \mathbb{N}.$$

Let  $b \in B$ . Since  $\mathcal{R}_b \in \text{BC}_\tau(S_b)$ , we conclude  $\mathcal{R}_b \in \text{BC}_{\psi^{[b]}}(\lambda x.S(b, x))$ . This implies  $B \in \text{suit}(\text{BC}, \text{BC})$ . The second claim follows immediately.  $\square$

As we have seen, the trick of encoding much information within the description sets or within the hypothesis spaces often supplies quite simple uniform strategies with a huge learning power. But nevertheless, our following results will make sure that uniform learning procedures cannot always be simplified to such a trivial level. On the one hand we can easily find a trick to design a uniform EX-strategy identifying *any* recursive core consisting of just a single element from its description, but on the other hand there is *no* uniform EX-strategy identifying all recursive cores consisting of two elements from their descriptions. In view of classical learning problems any classes consisting of just two elements are not more complex than classes consisting of one element, whereas their complexity is very different regarding uniform learning problems.

**Proposition 5.**  $\{b \in \mathbb{N} \mid \text{card } \mathcal{R}_b = 1\} \in \text{suit}(\text{EX}, \text{EX})$ .

*Proof.* Let  $B := \{b \in \mathbb{N} \mid \text{card } \mathcal{R}_b = 1\}$ . Then of course  $\mathcal{R}_b \in \text{EX}$  for all  $b \in B$ , i.e.  $\mathcal{R}_B \subseteq \text{EX}$ . Since for all  $f \in \mathcal{R}$  there exists a numbering  $\psi \in \mathcal{P}^2$  with  $\psi_0 = f$ , the function constantly zero learns  $\mathcal{R}_B$  uniformly from  $B$  wrt EX and EX.  $\square$

Now, in contrast to Proposition 5 we can prove that no kind of trick can help a strategy to uniformly identify all recursive cores consisting of up to two elements from their descriptions. In particular we observe that there are collections of quite simple identification problems, which even cannot be solved uniformly by encoding information within the hypothesis spaces.

**Theorem 7.**  $\{b \in \mathbb{N} \mid \text{card } \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} \leq 2\} \notin \text{suit}(\text{EX}, \text{EX})$ .

We omit the proof which can be found in [14]. The idea is to proceed indirectly. Let  $B := \{b \in \mathbb{N} \mid \text{card } \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} \leq 2\}$ . Assuming  $B \in \text{suit}(\text{EX}, \text{EX})$  implies the existence of an  $S \in \mathcal{R}^2$  satisfying  $\mathcal{R}_b \in \text{EX}(\lambda x.S(b, x))$  for all  $b \in B$ . But it is possible to construct  $b_0 \in B$  describing an  $\mathcal{R}$ -core which cannot be identified in the limit by  $\lambda x.S(b_0, x)$ . This strategy fails for at least one function  $f \in \mathcal{R}_{b_0}$  by either converging to an incorrect hypothesis or by diverging.

The following corollaries are direct consequences of Theorem 7.

**Corollary 3.**  $\{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite}\} \notin \text{suit}(\text{EX}, \text{EX})$ .

**Corollary 4.**  $\text{suit}(\text{EX}, \text{EX}) \subset \text{suit}(\text{EX}, \text{BC})$ .

*Proof.* Obviously  $\text{suit}(\text{EX}, \text{EX}) \subseteq \text{suit}(\text{EX}, \text{BC})$ . From Theorem 6 and Corollary 3 we conclude  $\{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite}\} \in \text{suit}(\text{EX}, \text{BC}) \setminus \text{suit}(\text{EX}, \text{EX})$ .  $\square$

Theorem 8 is a summary of our main results in this section.

**Theorem 8.** *Fix a criterion  $I \in \mathcal{I}$ . The following conditions are equivalent:*

1.  $\text{suit}(I, I) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$ ,
2.  $\text{suit}(\text{EX}, I) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq \text{EX}\}$ ,
3.  $I \in \{BC, BC^*\}$ .

In order to characterize the sets in  $\text{suit}(\text{EX}, \text{EX})$  we use Theorem 1 with arguments like those presented in [13], which can be transferred to the case of uniform learning quite easily (the proof is left out, but can be found in [14]). Note the similarity of our properties to the conditions for identification of languages in the limit from text, introduced in [1].

**Theorem 9.** *A set  $B \subseteq \mathbb{N}$  belongs to  $\text{suit}(\text{EX}, \text{EX})$  iff  $\exists d \in \mathcal{R}^2 \forall b \in B \exists \psi \in \mathcal{P}^2$*

1.  $\mathcal{R}_b \subseteq \mathcal{P}_\psi$ ,
2.  $\forall i \in \mathbb{N} [d(b, i) \sqsubseteq \psi_i]$ ,
3.  $\forall i, j \in \mathbb{N} [d(b, i) \sqsubseteq d(b, j) \sqsubseteq \psi_i \Rightarrow i = j]$ .

## 6 Uniform Learning with Special Hypothesis Spaces

The trick of encoding information within the hypothesis spaces supplies a strategy uniformly identifying all BC-identifiable recursive cores from their corresponding descriptions. This is a consequence of the freedom of choosing the hypothesis spaces in the definition of uniform learnability. The question arises, to what extent the learning power of uniform strategies is influenced by fixing special hypothesis spaces – for example acceptable numberings – in advance. From Jantke’s work [7] we already know that the set of descriptions of  $\mathcal{R}$ -cores consisting of just a single function is not suitable for uniform EX-identification, if we demand the hypotheses to be correct with respect to an acceptable numbering. Here we tighten Jantke’s result by proving that for the same set of descriptions even BC-identification is not strong enough.

**Theorem 10.** *Assume  $B := \{b \in \mathbb{N} \mid \text{card} \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} = 1\}$  and let  $\tau \in \mathcal{P}^2$  be an acceptable numbering. Then  $\mathcal{R}_B \notin \text{uni}_{B, \tau}(BC, BC)$ .*

An indirect proof is contained in [14]. The assumption  $\mathcal{R}_B \in \text{uni}_{B, \tau}(BC, BC)$  implies the existence of  $S \in \mathcal{R}^2$  satisfying  $\mathcal{R}_b \in \text{BC}_\tau(\lambda x.S(b, x))$  for all  $b \in B$ . A contradiction is obtained by construction of an index  $b_0 \in B$ , such that  $\lambda x.S(b_0, x)$  produces infinitely many hypotheses incorrect wrt  $\tau$  for the only function in  $\mathcal{R}_{b_0}$ .

Let  $B$  be the description set defined in Theorem 10. Since  $\mathcal{R} \in \text{BC}_\tau^*$  for any acceptable numbering  $\tau$ , we obtain  $\mathcal{R}_B \in \text{uni}_{B, \tau}(\text{EX}, \text{BC}^*) \setminus \text{uni}_{B, \tau}(\text{EX}, \text{BC})$ , but we can prove that our set  $B$  is not suitable for uniform  $\text{BC}^*$ -identification with respect to the hypothesis spaces  $\varphi^b$ ,  $b \in B$  given *a priori*:

**Corollary 5.** *Let  $B := \{b \in \mathbb{N} \mid \text{card} \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} = 1\}$ . Then  $\mathcal{R}_B \notin \text{uni}_{B,[id]}(\text{EX}, \text{BC}^*)$ .*

*Proof.* Assume  $\mathcal{R}_B \in \text{uni}_{B,[id]}(\text{EX}, \text{BC}^*)$ . Then there exists  $S \in \mathcal{R}^2$  which provides  $\mathcal{R}_b \in \text{BC}_{\varphi^b}^*(\lambda x. S(b, x))$  for all  $b \in B$ . Let  $\tau \in \mathcal{P}^2$  be acceptable. We will define a strategy  $T \in \mathcal{P}^2$  satisfying  $\mathcal{R}_b \in \text{EX}_\tau(\lambda x. T(b, x))$  for all  $b \in B$ . For that purpose we choose  $g \in \mathcal{R}$ , such that

$$\varphi_i^{g(b)}(j) := \begin{cases} \varphi_i^b(j) & \forall x \leq j \ [\varphi_i^b(x) \downarrow] \\ \uparrow & \text{otherwise} \end{cases} \quad \text{for all } b, i, j \in \mathbb{N}.$$

Fix  $c \in \mathcal{R}^2$  such that  $\tau_{c(b,i)} = \varphi_i^{g(b)}$  for all  $b, i \in \mathbb{N}$ . Provided  $b \in B$  we observe  $g(b) \in B$  and

$$\exists n_b \in \mathbb{N} \forall i \in \mathbb{N} \setminus \{n_b\} \ [\varphi_i^{g(b)} \text{ initial}].$$

Let  $f_b$  denote the function in  $\mathcal{R}_{g(b)}$ ,  $b \in B$ . Since  $\mathcal{R}_{g(b)} \in \text{BC}_{\varphi^{g(b)}}^*(\lambda x. S(g(b), x))$ , we conclude  $S(g(b), f_b[n]) = n_b$  for all but finitely many  $n \in \mathbb{N}$ . This can be explained by the fact that  $\varphi_i^{g(b)}$  is initial and thus  $\varphi_i^{g(b)} \neq^* f_b$  for all  $i \neq n_b$ . Therefore  $\mathcal{R}_{g(b)} \in \text{EX}_{\varphi^{g(b)}}(\lambda x. S(g(b), x))$ . If we define  $T(b, f^n) := c(b, S(g(b), f^n))$  for all  $b, n \in \mathbb{N}$ ,  $f \in \mathcal{R}$ , we obtain  $\mathcal{R}_B \in \text{uni}_{B,\tau}(\text{EX}, \text{EX})$  in contradiction to Theorem 10. Thus the assumption  $\mathcal{R}_B \in \text{uni}_{B,[id]}(\text{EX}, \text{BC}^*)$  has been wrong.  $\square$

One might reason that uniform learning from the description set  $B$  and with respect to the hypothesis spaces  $\varphi^b$  given above is so hard, because in each numbering  $\varphi^b$  the element of the recursive core possesses only one index. But even if we allow infinitely many  $\varphi^b$ -numbers for the functions to be learned, our situation does not improve:

**Corollary 6.** *If  $C := \{b \mid \text{card} \mathcal{R}_b = 1 \wedge \text{card} \{i \mid \varphi_i^b \in \mathcal{R}\} = \infty\}$ , then  $\mathcal{R}_C \notin \text{uni}_{C,[id]}(\text{EX}, \text{BC}^*)$ .*

*Proof.* We use Corollary 5. For this purpose assume  $\mathcal{R}_C \in \text{uni}_{C,[id]}(\text{EX}, \text{BC}^*)(S)$  for some appropriate strategy  $S \in \mathcal{R}^2$ . We will construct a uniform  $\text{BC}^*$ -strategy for our class  $\mathcal{R}_B$ , where  $B$  denotes the description set defined in Corollary 5.

There is a function  $g \in \mathcal{R}$  satisfying

$$\begin{aligned} \varphi_0^{g(b)} &= \varphi_0^b, \\ \varphi_1^{g(b)} &= \varphi_0^b, \quad \varphi_2^{g(b)} = \varphi_1^b, \\ \varphi_3^{g(b)} &= \varphi_0^b, \quad \varphi_4^{g(b)} = \varphi_1^b, \quad \varphi_5^{g(b)} = \varphi_2^b, \dots \end{aligned}$$

for all  $b \in \mathbb{N}$ . The following properties can be verified easily:

1.  $\forall b \in \mathbb{N} \ [\mathcal{R}_{g(b)} = \mathcal{R}_b]$ .
2.  $\forall b \in \mathbb{N} \forall f \in \mathcal{P}_{\varphi^{g(b)}} \ [\text{card} \{i \in \mathbb{N} \mid \varphi_i^{g(b)} = f\} = \infty]$ , i.e. each function in  $\mathcal{P}_{\varphi^{g(b)}}$  possesses infinitely many  $\varphi^{g(b)}$ -indices.

3.  $\exists e \in \mathcal{R} \forall b \in \mathbb{N} \forall i \in \mathbb{N} [\varphi_i^{g(b)} = \varphi_{e(i)}^b]$ , i.e.  $\varphi^{g(b)}$ -indices can be translated effectively into  $\varphi^b$ -indices with a uniform method.

Considering our problem we observe that  $g(b) \in C$  whenever  $b \in B$ . Therefore  $T(b, x) := e(S(g(b), x))$  (for  $b, x \in \mathbb{N}$ ) yields a computable strategy satisfying  $\mathcal{R}_B \in \text{uni}_{B, [id]}(\text{EX}, \text{BC}^*)(T)$ . That contradiction to Corollary 5 now forces us to reject our assumption. This implies  $\mathcal{R}_C \notin \text{uni}_{C, [id]}(\text{EX}, \text{BC}^*)$ .  $\square$

To characterize uniform learning with respect to acceptable numberings, we use Theorems 1 and 2. We omit the proofs which can be easily transferred from the non-uniform case.

**Theorem 11.** *Let  $B \subseteq \mathbb{N}$  fulfill  $\mathcal{R}_b \in \text{EX}$  for all  $b \in B$ . Furthermore, let  $\tau \in \mathcal{P}^2$  be acceptable. Then  $\mathcal{R}_B \in \text{uni}_{B, \tau}(\text{EX}, \text{EX}) \iff \exists \psi \in \mathcal{P}^3 \exists d \in \mathcal{R}^3 \forall b \in B$*

1.  $\mathcal{R}_b \subseteq \mathcal{P}_{\psi^b}$ ,
2.  $\forall i, j \in \mathbb{N} [i \neq j \Rightarrow \psi_i^b \neq_{d(b, i, j)} \psi_j^b]$ .

**Theorem 12.** *Fix  $I \in \{\text{EX}, \text{BC}\}$ . Let  $B \subseteq \mathbb{N}$  fulfill  $\mathcal{R}_b \in I$  for all  $b \in B$ . Furthermore, let  $\tau \in \mathcal{P}^2$  be an acceptable numbering. Then  $\mathcal{R}_B \in \text{uni}_{B, \tau}(I, \text{BC}) \iff \exists \psi \in \mathcal{P}^3 \exists d \in \mathcal{R}^2 \forall b \in B$*

1.  $\mathcal{R}_b \subseteq \mathcal{P}_{\psi^b}$ ,
2.  $\forall i, j \in \mathbb{N} [\psi_i^b =_{\max\{d(b, i), d(b, j)\}} \psi_j^b \iff \psi_i^b = \psi_j^b]$ .

Since  $\mathcal{R} \in \text{BC}_\tau^*$  for any acceptable numbering  $\tau$ , we can use the same reasoning as in Proposition 4 to prove our characterization in Theorem 13.

**Theorem 13.** *Fix  $I \in \mathcal{I}$  and  $B \subseteq \mathbb{N}$ . Furthermore, let  $\tau \in \mathcal{P}^2$  be an acceptable numbering. Then  $\mathcal{R}_B \in \text{uni}_{B, \tau}(I, \text{BC}^*) \iff [\mathcal{R}_b \in I \text{ for all } b \in B]$ .*

A very natural learning behaviour is to construct only consistent intermediate hypotheses, i.e. hypotheses agreeing with the information received so far (cf. [3]).

**Definition 7.** *Assume  $U \in \mathcal{R}$ ,  $\psi \in \mathcal{P}^2$ .  $U$  is called identifiable consistently wrt  $\psi$  iff there is an  $S \in \mathcal{P}$  satisfying  $U \in \text{EX}_\psi(S)$ , such that  $\psi_{S(f[n])} =_n f$  for all  $f \in U$ ,  $n \in \mathbb{N}$ . We use the notions  $\text{CONS}$ ,  $\text{CONS}_\psi$ ,  $\text{CONS}_\psi(S)$  for that criterion in the common way.*

Furthermore it is reasonable to demand just total functions to be described by the intermediate hypotheses (see [8]).

**Definition 8.** *Assume  $U \in \mathcal{R}$ ,  $\psi \in \mathcal{P}^2$ .  $U$  is called identifiable with consistent and total hypotheses wrt  $\psi$  iff there is an  $S \in \mathcal{P}$  satisfying  $U \in \text{CONS}_\psi(S)$ , such that  $\psi_{S(f[n])} \in \mathcal{R}$  for all  $f \in U$ ,  $n \in \mathbb{N}$ . The notions  $\text{CT}$ ,  $\text{CT}_\psi$ ,  $\text{CT}_\psi(S)$  are used for that criterion in the usual manner.*

Definitions 4, 5 and 6 can be reformulated for the inference criteria  $\text{CONS}$  and  $\text{CT}$ . A quite simple result, based on Identification by Enumeration as has been introduced in [6], is presented in [10]:

**Theorem 14.** *If  $B \subseteq \mathbb{N}$ ,  $\varphi^b \in \mathcal{R}^2$  for all  $b \in B$ , then  $\mathcal{R}_B \in \text{uni}_{B,[id]}(CT, CT)$ .*

For uniform learning with respect to “meaningful” hypothesis spaces, i.e. in such a way, that all hypotheses produced by the strategy can be “interpreted” by the user, most of our results have been negative. Even very “simple” classes yield bad results. To show that there still remains a sense in the definition of uniform learning, we present some intuitively more complex description sets suitable for uniform learning in the limit – even with consistent and total intermediate hypotheses – with respect to any acceptable numbering.

**Definition 9.** *A set  $D \subseteq \mathcal{P}$  is called discrete iff for any  $f \in D$  there is an  $n \in \mathbb{N}$ , such that  $f \neq_n g$  for all functions  $g \in D \setminus \{f\}$ . This  $n \in \mathbb{N}$  is then called discreteness point for  $f$  wrt  $D$ .*

**Theorem 15.** *Let  $\tau \in \mathcal{P}^2$  be an acceptable numbering,  $B \subseteq \mathbb{N}$ . Assume that  $\mathcal{P}_{\varphi^b}$  is discrete for all  $b \in B$ . Then  $\mathcal{R}_B \in \text{uni}_{B,\tau}(CT, CT)$ .*

*Proof.* Provided that  $B$  fulfills the conditions requested above we first construct appropriate hypothesis spaces uniformly in  $b \in B$ . Of course their indices may then be transformed to equivalent programs in  $\tau$  effectively. For that purpose we will fix  $b \in B$  and collect all initial segments of functions in  $\mathcal{P}_{\varphi^b}$  in order to use them as initial segments for the functions in our new hypothesis space. We will try to extend these initial segments to computable functions, such that finally all functions of the recursive core  $\mathcal{R}_b$  have indices in our constructed numbering. The uniform strategy defined afterwards works iteratively. It always starts with a consistent hypothesis and in each following inference step it tests whether its previous hypothesis is still consistent with the new information received or not. In the first case the previous hypothesis is maintained, otherwise a new consistent hypothesis is constructed.

For the definition of our new hypothesis spaces  $\psi^b$ ,  $b \in B$  we need a function  $\text{extend} \in \mathcal{P}$ , which indicates suitable extensions of initial segments. For  $b, n, x, k \in \mathbb{N}$  and  $f \in \mathcal{R}$  define

$$\text{extend}(b, f[n], x, k) := \begin{cases} 1 & \varphi_k^b(0) \downarrow, \dots, \varphi_k^b(x) \downarrow \text{ and } \varphi_k^b[n] = f[n] \\ \uparrow & \text{otherwise} \end{cases}$$

Thus  $\text{extend}(b, f[n], x, k)$  is defined if and only if  $\varphi_k^b[x]$  is an “extension” of  $f[n]$ .

*Definition of  $\psi \in \mathcal{P}^3$  with  $\mathcal{R}_b \in CT_{\psi^b}$  for all  $b \in B$ :*

Let  $b, n, x \in \mathbb{N}$ ,  $f \in \mathcal{R}$ . We define

$$\psi(b, f[n], x) := \begin{cases} f(x) & x \leq n \\ \varphi_k^b(x) & x > n \text{ and } k \in \mathbb{N} \text{ may be found,} \\ & \text{such that } \text{extend}(b, f[n], x, k) = 1 \\ \uparrow & \text{otherwise} \end{cases}$$

Obviously  $\psi$  is computable. For any  $b \in B$  we observe the following properties:

*Claim 1.* If  $f \in \mathcal{R}_b$  and  $n \in \mathbb{N}$  is any integer, then  $\psi_{f[n]}^b \in \mathcal{R}$ .

*Claim 2.* If  $f \in \mathcal{R}_b$  and  $n_f$  is a discreteness point of  $f$  wrt  $\mathcal{P}_{\varphi^b}$ , then  $\psi_{f[n_f]}^b = f$ .

*Claim 3.*  $\mathcal{R}_b \subseteq \mathcal{P}_{\psi^b}$ .

*Proof of Claim 1.* Since  $f \in \mathcal{P}_{\varphi^b}$ , we know that for all  $x \in \mathbb{N}$  there is an “extension” of  $f[n]$ , i.e.  $\forall x \in \mathbb{N} \exists k \in \mathbb{N} [\text{extend}(b, f[n], x, k) = 1]$ . As there is an extension, it may also be found within a finite amount of time. The definition of  $\psi$  then implies that for all  $n \in \mathbb{N}$  the function  $\psi_{f[n]}^b$  is total and thus recursive.

*Proof of Claim 2.* For all arguments less than or equal to  $n_f$  the values of  $\psi_{f[n_f]}^b$  and  $f$  must agree, because those arguments match the first case in the definition of  $\psi^b$ . For all arguments greater than  $n_f$  the existence of an “extension” of  $f[n_f]$  is checked. As  $f \in \mathcal{R}$ , we observe that this check will always stop with a positive answer. Since  $n_f$  is a discreteness point for  $f$  wrt  $\mathcal{P}_{\varphi^b}$ , the only function in  $\mathcal{P}_{\varphi^b}$  extending  $f[n_f]$  is  $f$ . Hence  $\psi_{f[n_f]}^b = f$ .

*Proof of Claim 3.* Let  $f \in \mathcal{R}_b$  and let  $n_f$  be a discreteness point of  $f$  wrt  $\mathcal{P}_{\varphi^b}$ . As  $\mathcal{P}_{\varphi^b}$  is discrete,  $n_f$  exists. Claim 2 then implies  $\psi_{f[n_f]}^b = f$ , hence  $f \in \mathcal{P}_{\psi^b}$ .

Now let  $c \in \mathcal{R}^2$  be a recursive function satisfying  $\tau_{c(b,y)} = \psi_y^b$  for all  $b, y \in \mathbb{N}$ .

*Definition of a strategy*  $S \in \mathcal{P}^2$  with  $\mathcal{R}_b \in CT_\tau(\lambda x.S(b, x))$  for all  $b \in B$ :  
Let  $f \in \mathcal{R}$ ,  $b, n \in \mathbb{N}$ . We define  $S(b, f[0]) := c(b, f[0])$  and

$$S(b, f[n+1]) := \begin{cases} \uparrow & \exists x \leq n+1 [\tau_{S(b, f[n])}(x) \uparrow] \\ S(b, f[n]) & \tau_{S(b, f[n])} =_{n+1} f \\ c(b, f[n+1]) & \text{otherwise} \end{cases}$$

Now we prove that for all  $b \in B$  and all initial segments of functions in  $\mathcal{R}_b$  our strategy returns consistent indices of total functions. Furthermore we will show that for any function in  $\mathcal{R}_b$  the sequence of hypotheses converges. From consistency we thus obtain convergence to a *correct* index. We have to verify:

- (i)  $\forall b \in B \forall f \in \mathcal{R}_b \forall n \in \mathbb{N} [S(b, f[n]) \downarrow]$ .
- (ii)  $\forall b \in B \forall f \in \mathcal{R}_b \forall n \in \mathbb{N} [\tau_{(b, S(b, f[n]))} =_n f]$ .
- (iii)  $\forall b \in B \forall f \in \mathcal{R}_b \forall n \in \mathbb{N} [\tau_{(b, S(b, f[n]))} \in \mathcal{R}]$ .
- (iv)  $\forall b \in B \forall f \in \mathcal{R}_b \exists n_0 \in \mathbb{N} \forall n \geq n_0 [S(b, f[n]) = S(b, f[n_0])]$ .

*Proof of (i), (ii) and (iii).* Let  $b \in B$ ,  $f \in \mathcal{R}_b$ . We use induction on  $n$ .

First assume  $n = 0$ . Obviously  $S(b, f[0]) = c(b, f[0])$  is defined. Furthermore, from the definitions of  $S$ , *extend*,  $\psi$  and  $c$  we observe that

$$\tau_{S(b, f[0])} = \tau_{c(b, f[0])} = \psi_{f[0]}^b =_0 f.$$

This proves the consistency of the hypothesis  $S(b, f[0])$ . Since  $f \in \mathcal{R}_b$ , we observe from Claim 1 with  $n = 0$ , that  $\tau_{S(b, f[0])} = \tau_{c(b, f[0])} = \psi_{f[0]}^b \in \mathcal{R}$ .

Assume for a fixed  $n \in \mathbb{N}$ , that  $S(b, f[n])$  is defined, consistent for  $f[n]$  and a  $\tau$ -index of a total function. From this situation we want to deduce that also  $S(b, f[n+1])$  is defined, consistent for  $f[n+1]$  and a  $\tau$ -index of a total function. Since  $\tau_{S(b, f[n])}$  is total, we can test effectively whether  $\tau_{S(b, f[n])} =_{n+1} f$  or not. If the first case occurs, the hypothesis is maintained. Then the new hypothesis is still defined, consistent and an index of a total function. Otherwise, if the second case occurs, our previous hypothesis must have been wrong. We obtain

$$\tau_{S(b, f[n+1])} = \tau_{c(b, f[n+1])} = \psi_{f[n+1]}^b =_{n+1} f .$$

So  $S(b, f[n+1])$  is consistent for  $f[n+1]$ . Claim 1 now yields  $\tau_{S(b, f[n+1])} \in \mathcal{R}$ . Anyway the hypothesis produced by  $S$  fulfills the conditions (i), (ii) and (iii).

*Proof of (iv).* Again assume  $b \in B$ ,  $f \in \mathcal{R}_b$ . If there exists an  $n_0 \in \mathbb{N}$ , such that for all  $n \geq n_0$  the first case in the definition of  $S(b, f[n])$  occurs, the hypothesis  $S(b, f[n_0])$  will never be changed and the sequence of hypotheses converges. Provided such an  $n_0$  does *not* exist, we may deduce a contradiction as follows:

As  $\mathcal{P}_{\varphi^b}$  is discrete, there is an  $n_f \in \mathbb{N}$  satisfying  $[\varphi_i^b =_{n_f} f \iff \varphi_i^b = f]$  for all  $i \in \mathbb{N}$ . From (ii) we already know  $\tau_{S(b, f[n_f])} =_{n_f} f$ . Since according to our assumption there exists a number  $n > n_f$ , such that the second case in the definition of  $S(b, f[n])$  occurs, the hypothesis put out by  $S$  on input  $(b, f[n])$  equals  $c(b, f[n])$ . Since  $n > n_f$ , the number  $n$  is a discreteness point of  $f$  wrt  $\mathcal{P}_{\varphi^b}$ . Claim 2 now implies  $f = \psi_{f[n]}^b = \tau_{c(b, f[n])} = \tau_{S(b, f[n])}$ . Thus  $S$  has found a correct hypothesis. But correct hypotheses must be consistent for all further inputs; therefore the first case in the definition of  $S$  will occur for all following input segments. Hence we reach the desired contradiction. This implies (iv).

From conditions (i),(ii) and (iv) we conclude, that the output of our uniform strategy converges to a correct hypothesis for all “interesting” input sequences. Together with condition (iii) we finally obtain  $\mathcal{R}_b \in \text{CT}_{\psi^b}(\lambda x.S(b, x))$  for all  $b \in B$ . This completes the proof.  $\square$

By comparison of Theorem 15 with our previous results we conclude that the way the recursive cores are described has much more influence upon their uniform learnability than the  $\mathcal{R}$ -cores themselves. We know from the proof of Theorem 4 that from appropriate descriptions even for the entirety of all classes in EX uniform learning with respect to acceptable numberings is possible. Because of Theorem 15 the set  $B_{\text{discrete}} := \{b \in \mathbb{N} \mid \mathcal{P}_{\varphi^b} \text{ is discrete}\}$  is suitable for uniform CT-identification with respect to any acceptable numbering. On the other hand, there are sets describing finite – and thus very “simple” – recursive cores which are not suitable for uniform learning with respect to EX at all, even if we allow free choice of the hypothesis spaces. The reason for the failure of all uniform strategies might be the inappropriate topological features of the described numberings. Now if a set  $B$  even describes functions  $\varphi^b$  enumerating discrete sets *without repetitions*, we observe that  $B$  is suitable for uniform learning with respect to the hypothesis spaces  $\varphi^b$  given a priori. Unfortunately, to prove that



result we will abandon our demand for total intermediate hypotheses. Theorems 15 and 16 give rise to a more optimistic view regarding the learning power of uniform identification models. Indeed, our concept is not as poor or trivial as our previous results might have suggested.

**Definition 10.** A numbering  $\psi \in \mathcal{P}^2$  is called *absolutely discrete* iff  $\mathcal{P}_\psi$  is discrete and each function in  $\mathcal{P}_\psi$  has exactly one  $\psi$ -number.

**Theorem 16.** Fix  $B \subseteq \mathbb{N}$ . Assume that  $\varphi^b$  is an absolutely discrete numbering for all  $b \in B$ . Then  $\mathcal{R}_B \in \text{uni}_{B,[id]}(\text{CT}, \text{CONS})$ .

*Proof.* From Theorem 15 we know that  $\mathcal{R}_b \in \text{CT}$  for all  $b \in B$ . For any  $f \in \mathcal{R}$ ,  $n, b \in \mathbb{N}$  a uniform strategy  $S \in \mathcal{P}^2$  may look for a number  $i \in \mathbb{N}$  with  $\varphi_i^b =_n f$  and then return  $i$ . As this strategy  $S$  works consistently and  $B$  describes absolutely discrete numberings only, we obtain  $\mathcal{R}_B \in \text{uni}_{B,[id]}(\text{CT}, \text{CONS})(S)$ .  $\square$

## References

1. Angluin, D.; *Inductive Inference of Formal Languages from Positive Data*, Information and Control 45, 117-135 (1980).
2. Barzdins, J.; *Two Theorems on the Limiting Synthesis of Functions*, Theory of Algorithms and Programs, Latvian State University, Riga 210, 82-88 (1974) (in Russian).
3. Barzdins, J.; *Inductive Inference of Automata, Functions and Programs*, In: Proceedings International Congress of Math., Vancouver, 455-460 (1974).
4. Baliga, G.; Case, J.; Jain, S.; *Synthesizing Enumeration Techniques for Language Learning*, In: Proceedings of the Ninth Annual Conference on Computational Learning Theory, ACM Press, 169-180 (1996).
5. Case, J.; Smith, C.; *Comparison of Identification Criteria for Machine Inductive Inference*, Theoretical Computer Science 25, 193-220 (1983).
6. Gold, E.M.; *Language Identification in the Limit*, Information and Control 10, 447-474 (1967).
7. Jantke, K.; *Natural Properties of Strategies Identifying Recursive Functions*, Elektronische Informationsverarbeitung und Kybernetik 15, 487-496 (1979).
8. Jantke, K.; Beick, H.; *Combining Postulates of Naturalness in Inductive Inference*, Elektronische Informationsverarbeitung und Kybernetik 17, 465-484 (1981).
9. Kapur, S.; Bilardi, G.; *On uniform learnability of language families*, Information Processing Letters 44, 35-38 (1992).
10. Osherson, D.N.; Stob, M.; Weinstein, S.; *Synthesizing Inductive Expertise*, Information and Computation 77, 138-161 (1988).
11. Oymanns, L.; *Lernen von rekursiven Funktionen mit guten Beispielen*, Diploma Thesis, University of Kaiserslautern (1998) (in German).
12. Rogers, H.; *Theory of Recursive Functions and Effective Computability*, MIT Press, Cambridge, Massachusetts (1987).
13. Wiehagen, R.; *Zur Theorie der algorithmischen Erkennung*, Dissertation B, Humboldt-University, Berlin (1978) (in German).
14. Zilles, S.; *On Uniform Learning of Classes of Recursive Functions*, Technical Report LSA-2000-05E, Centre for Learning Systems and Applications, University of Kaiserslautern (2000).