# On Uniform Learning of Classes of 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 learnable uniformly without restricting the set of possible descriptions. Furthermore the influence of the hypothesis spaces on uniform learnability is analysed.

Some more technical results are concerned with transferring well-known characterizations of inference criteria from the classical scenario to the context of uniform learning.

# Contents

# 1 Introduction

Consider a scenario consisting of a learner and an object unknown to the learner; for example the learner might be you and the object might be an infinite sequence of nonnegative integers. In your daily newspaper you have found a page with some brainteasers containing the following exercise:

```
Exercise:  Find out the next three numbers in the row
           1,2,3,3,1,6,4,7,3,9,10,9,15,...
```

Now if you can solve this exercise and additionally state a rule for computing further numbers, you have found a description of the unknown object (i.e. of the whole infinite sequence of numbers), although the information you had about the object was incomplete. Of course, the more information you get about the sequence of numbers, the better are your chances for solving the given problem. If you only saw the first three numbers 1,2,3, perhaps you would rather expect that the sequence consists exactly of all positive integers, listed in ascending order, but the information of the fourth number in the row tells you that such a guess must be wrong. As soon as your guesses have become correct and stay correct, you have "learned" the full unknown object from just a finite amount of information.

Inductive Inference is concerned with similar learning processes considered in a more recursion-theoretic environment. Here the objects to be learned are total recursive functions and the learner is a partial-recursive function. The growing sequence of information about an object $f$ corresponds to the sequence $f(0), f(1), f(2), \ldots$ of its output values. In the $n$-th step of the learning process the learner reads the values $f(0), \ldots, f(n)$ and generates a hypothesis. That hypothesis is a nonnegative integer and shall be interpreted as an index of a function enumerated by a partial-recursive numbering. Thus the semantics of the hypotheses depend on the choice of the hypothesis space, i.e. on the choice of a partial-recursive numbering. In the initial approach of identification in the limit introduced by Gold in [Go67] the learner is supposed to generate a sequence of hypotheses converging to a correct index of the unknown function. Several further identification criteria have been introduced and analysed in [Ba74a], [Ba74b] and [CS83]. 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 single learner identifying *each* function in the class $U$ with respect to the hypothesis space $\psi$ in a manner satisfying the conditions of the criterion $I$.

Now imagine you have accumulated lists 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 your 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 [Ja79] 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. Results on uniform identification of classes of languages can be found in [OSW88], [KB92] and [BCJ96]. 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[1] 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 with respect to other inference criteria. It is proved that the collection of all descriptions of classes consisting of just two recursive functions is not suitable for uniform identification in the limit, 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. In Section 6 Jantke's result is strengthened 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

[1] for the definitions of inference criteria mentioned here see Section 2.2

hypothesis spaces.

But on the other hand this paper also contains some quite positive results. For example, if the descriptions of 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[2]. These positive results at least seem to justify some further research on uniform identifiability. Therefore Section 5 additionally provides some results concerning the suitability of unions of description sets; simple results on the learning power of special "natural" uniform strategies are presented in Section 7. Furthermore Wiehagen's characterizations of several inference criteria in [Wie78] – based on numbering theory – are transferred to the context of uniform identification. In general, Section 2 provides preliminary notions and definitions and Section 3 deals with the notion of uniform identification. General results are presented in Section 4, followed by results on particular description sets in Section 5. Section 6 is concerned with the influence of the choice of the hypothesis spaces and finally Section 7 mentions some typical identification techniques.

I thank Prof. Wiehagen, Jochen Nessel and Martin Memmel for their advice.

[2]see Section 2.2 for definitions

# 2 Preliminaries

First of all we will fix important notions to be used in this paper. All definitions and theorems in the context of recursion theory not explicitly introduced here can be found in the common literature such as for example [Ro87].

## 2.1 Notation

We denote the set of nonnegative integers by $\mathbb{N}$ and write $\mathbb{N}^*$ for the set of all finite tuples of elements of $\mathbb{N}$. Instead of "nonnegative integer" we simply write "integer" or "number". If $n$ is any integer, we refer to the set of all $n$-tuples of integers by $\mathbb{N}^n$. The length function $\lambda\alpha.|\alpha| : \mathbb{N}^* \to \mathbb{N}$ assigns the number of components to any finite tuple of integers (i.e. $|\alpha| = n$ for all $\alpha \in \mathbb{N}^n$ with $n \in \mathbb{N}$). Here – and throughout this paper – we use the common $\lambda$-notation for functions. For any tuples $\alpha, \beta \in \mathbb{N}^*$ we use the notion $\alpha \sqsubseteq \beta$, whenever there exists a tuple $\gamma \in \mathbb{N}^*$ such that the concatenation $\alpha\gamma$ equals $\beta$. If in addition $\gamma$ is not the empty tuple, we may also write $\alpha \sqsubset \beta$. The notion $\alpha \sqsubseteq f$, where $f : \mathbb{N} \to \mathbb{N}$ is a function, means that $\alpha = (f(0), \ldots, f(|\alpha| - 1))$. By means of a bijective and computable mapping $\lambda x_1. \ldots \lambda x_n. \langle x_1, \ldots, x_n \rangle$ from $\mathbb{N}^n$ onto $\mathbb{N}$ we can 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 $\mathrm{cod} : \mathbb{N}^* \to \mathbb{N}$. For convenience we sometimes use the simple notion $\alpha$ to refer to $\mathrm{cod}(\alpha)$, where $\alpha \in \mathbb{N}^*$. Let $\emptyset$ be a symbol for the empty set. 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 $n \in \mathbb{N}$.

For any pair of sets $X$ and $Y$ the notation $X \subset Y$ expresses a proper inclusion of $X$ in $Y$. As long as we do not want to exclude the absolute equality of the given sets, we denote this by $X \subseteq Y$. As a symbol for the incomparability of sets we use $\#$. $\mathrm{card}\, X$ serves as a notation for the cardinality of a set $X$, and we write $\mathrm{card}\, X = \infty$, whenever $X$ is an infinite set. The set of all subsets of $X$ is referred to by $\wp X$. 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 $X \subseteq Y$.

$\mathcal{P}^n$ denotes the class of all partial-recursive functions of $n$ variables. Its subclass of total functions is denoted by $\mathcal{R}^n$. Whenever $n = 1$ or the number of arguments is of no special interest, we omit the note for the number of arguments and simply write $\mathcal{P}$ or $\mathcal{R}$. We use the phrase "recursive function" to mean total partial-recursive function. For any function $f \in \mathcal{P}$ and any integer $n$ the notation $f[n]$ refers to the coding $\mathrm{cod}(f(0), \ldots, f(n))$ of the *initial segment* of length $n + 1$ of $f$, as long as the values $f(0), \ldots, f(n)$ are all defined. If $f$ is a partial-recursive function and $x$ is any integer, we write

- $f(x)\!\downarrow$, if $f$ is defined for the argument $x$,

- $f(x)\uparrow$, if $f$ is not defined for the argument $x$.

For the comparison of two functions $f, g \in \mathcal{P}$ with respect to their initial segments of length $n+1$ (with $n \in \mathbb{N}$) we also agree on a special notation. In the case

$$\{(x, f(x)) \mid x \leq n \text{ and } f(x)\downarrow\} = \{(x, g(x)) \mid x \leq n \text{ and } g(x)\downarrow\}$$

we write $f =_n g$, 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

$$\forall^\infty n \; [[f(n)\uparrow \wedge g(n)\uparrow] \text{ or } [f(n)\downarrow \wedge g(n)\downarrow \wedge f(n) = g(n)]],$$

we write $f =^* g$.

If we identify two functions $f, g$ with the corresponding sets of input-output tuples, the notion $f \subseteq g$ means that

$$\{(x, f(x)) \mid x \in \mathbb{N},\ f(x)\downarrow\} \subseteq \{(x, g(x)) \mid x \in \mathbb{N},\ g(x)\downarrow\}.$$

A partial-recursive function $f$ may also be identified with the sequence $(f(n))_{n\in\mathbb{N}}$ of its values. That explains the use of notations like for example $f = 0^{k+1}\uparrow^\infty$ for the function

$$f(x) = \begin{cases} 0 & x \leq k \\ \uparrow & x > k \end{cases} \quad \text{for } x \in \mathbb{N},$$

or $g = 0^{k+1}12^\infty$ for the function

$$g(x) = \begin{cases} 0 & x \leq k \\ 1 & x = k+1 \\ 2 & x > k+1 \end{cases} \quad \text{for } x \in \mathbb{N}.$$

If $\psi \in \mathcal{P}^{n+1}$ for a number $n \geq 1$, we regard $\psi$ as a numbering for the set $\mathcal{P}_\psi := \{\psi_i \mid i \in \mathbb{N}\}$ by means of the definition

$$\psi_i(x) := \psi(i, x) \text{ for all } i \in \mathbb{N},\ x \in \mathbb{N}^n.$$

The index $i \in \mathbb{N}$ is also called $\psi$-number of the function $\psi_i$.

If we choose an acceptable numbering $\varphi$, we can assign a "computing-time function" $\Phi_i$ to any function $\varphi_i$ ($i \in \mathbb{N}$) by a uniform procedure. Informally we identify $\varphi_i$ with a Turing-program executed by a Turing-machine $M_i$. With the value $\Phi_i(x)$ we associate the number of computational steps made by $M_i$ on the input $x$ (if the computation does not stop, then $\Phi_i(x)$ is undefined).[3] For any $i, x, n \in \mathbb{N}$ we introduce the following notations:

[3] compare with a Blum complexity measure in [Bl67]

- $\varphi_i(x)\downarrow_{\leq n}$ means that the computation of $\varphi_i(x)$ terminates within $n$ steps;
- $\varphi_i(x)\uparrow_{\leq n}$ means that the computation of $\varphi_i(x)$ does not terminate within $n$ steps.

If $\psi \in \mathcal{P}^{n+2}$ ($n \in \mathbb{N}$) is a computable function, every integer $b \in \mathbb{N}$ "describes" a numbering of partial-recursive functions 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:

$$\psi_i^b(x) := \psi^b(i,x) \text{ for all } i \in \mathbb{N},\ x \in \mathbb{N}^n.$$

If a function $\psi \in \mathcal{P}^{n+1}$ ($n \geq 1$) is given, we are often particularly interested in the recursive functions in $\mathcal{P}_\psi$. In their entirety they will be called the "recursive core" or "$\mathcal{R}$-core" of $\mathcal{P}_\psi$ (abbreviated by $\mathcal{R}_\psi$). Hence

$$\mathcal{R}_\psi \ = \ \mathcal{R} \cap \mathcal{P}_\psi.$$

## 2.2 Inductive Inference Criteria

### 2.2.1 Identification in the Limit – the Inference Criterion EX

Identification in the limit provides the fundamentals for learning models examined in the field of inductive inference and has first been analysed by Gold in [Go67]. Since identification in the limit is often called "explanatory correct" identification, we use the notation EX to refer to that criterion. Definition 2.1 provides a formal notation.

**Definition 2.1 (Identification in the Limit)** *Let $U \subseteq \mathcal{R},\ \psi \in \mathcal{P}^2$. The class $U$ is called identifiable in the limit with respect to $\psi$ if and only if there is a function $S \in \mathcal{P}$ such that for any $f \in U$:*

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

*We also write: $U \in EX_\psi(S)$.*

$EX_\psi := \{U \mid U$ *is identifiable in the limit with respect to* $\psi\}$.

$EX := \bigcup_{\psi \in \mathcal{P}^2} EX_\psi$.

If $\{f\} \subset \mathcal{R}$ is identified in the limit by a strategy $S$, there must be a minimal "time" $n_0 \in \mathbb{N}$, such that the sequence $(S(f_n))_{n\in\mathbb{N}}$ of hypotheses remains constant from that time on. That time is referred to by $\text{conv}(S,f)$. Formally we define:

$$\text{conv}(S,f) := \min\{n_0 \in \mathbb{N} \mid S(f^n) = S(f^{n_0}) \text{ for all } n \geq n_0\}$$

for any $S \in \mathcal{P},\ f \in \mathcal{R}$, provided $(S(f_n))_{n\in\mathbb{N}}$ converges.

Wiehagen's work [Wie78] supplies a characterization of the classes learnable in the limit (see Theorem 2.1). Since characterizations of inference criteria will be very useful for the proofs in this paper, we present the corresponding results here.

**Theorem 2.1** *Let $U \subseteq \mathcal{R}$. $U$ is identifiable in the limit, if and only if there is a numbering $\psi \in \mathcal{P}^2$ and a function $d \in \mathcal{R}^2$ satisfying*

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

### 2.2.2 Identification with a Bounded Number of Mind Changes – the Inference Criteria $\mathbf{EX}_m$, $m \in \mathbb{N}$

The sequence of hypotheses produced by an EX-strategy on any of the functions it identifies will converge. Imagine a user reading the strategy's output up to a certain time. Though he knows that the sequence will converge, he cannot decide whether the time of convergence is already reached. Thus he cannot rely on the latest hypothesis produced by the strategy, i.e. even if the strategy has reached its final correct hypothesis, the user will not know that the actual hypothesis is correct. He simply cannot be sure that the strategy will never change its "mind" again. If on the other hand we introduce a bound $m \in \mathbb{N}$ and our strategy identifies functions without changing its hypothesis for more than $m$ times, the user can rely on the correctness of the output at least when the strategy has already changed its mind exactly $m$ times. Case and Smith have first studied the learning power of strategies working with a bounded number of mind changes in [CS83].

**Definition 2.2 (Identification in the Limit with a Bounded Number of Mind Changes)**
*Assume $U \subseteq \mathcal{R},\ \psi \in \mathcal{P}^2$ and $m \in \mathbb{N}$. $U$ is called identifiable in the limit with no more than $m$ mind changes*[4] *wrt $\psi$, if and only if there exists a function $S \in \mathcal{P}$ satisfying*

1. *$U \in EX_\psi(S)$ (where $S$ is additionally permitted to put out the sign "?"),*
2. *for all $f \in U$ there is an $n_f \in \mathbb{N}$ satisfying*
   - $\forall x < n_f\ [S(f[x]) = ?]$,
   - $\forall x \geq n_f\ [S(f[x]) \in \mathbb{N}]$,

[4] we often abbreviate this formulation with "identifiable with no more than $m$ mind changes"

*3.* $\forall f \in U$ [ $card\{n \in \mathbb{N} \mid ? \neq S(f[n]) \neq S(f[n+1])\} \leq m$ ].

*We also write:* $U \in (EX_m)_\psi(S)$.

$(EX_m)_\psi := \{U \mid U$ *is identifiable in the limit with no more than* $m$ *mind changes wrt* $\psi\}$.

$EX_m := \bigcup_{\psi \in \mathcal{P}^2} (EX_m)_\psi$.

*A class* $U \subseteq \mathcal{R}$ *is identifiable with a bounded number of mind changes if and only if there exists a number* $m \in \mathbb{N}$ *such that* $U \in EX_m$.

The meaningless hypothesis "?" put out in the beginning of the learning process allows our strategy to read some more values of the function to be learned before putting out its first hypothesis. Thus it may "save" one of its precious mind changes.
A special case of identification with a bounded number of mind changes often analysed is finite identification. Its usual definition is equivalent to the definition of identification with (no more than) 0 mind changes. Therefore we often refer to the $EX_0$-criterion with "finite identifiability"[5].

Again we are interested in characterizing the classes identifiable with a bounded number of mind changes. The proof of Theorem 2.2 proceeds similarly to the proof of Theorem 2.1. For a complete proof see [Zi99].

**Theorem 2.2** *Let* $U \subseteq \mathcal{R}$ *and* $m \in \mathbb{N}$. $U$ *is identifiable with no more than* $m$ *mind changes if and only if there exist a numbering* $\psi \in \mathcal{P}^2$ *and a function* $d \in \mathcal{R}$ *satisfying*

*1.* $U \subseteq \mathcal{P}_\psi$,

*2.* $\forall j \in \mathbb{N}$ $[card\ \{i \neq j \mid \psi_i =_{d(i)} \psi_j\} \leq m]$,

*3.* $\forall i, j \in \mathbb{N}$ $[i \neq j \ \Rightarrow \ \psi_i \neq_{\max\{d(i),d(j)\}} \psi_j]$.

As a special case of Theorem 2.2 we can formulate a characterization of finite identifiability, which can also be found in [Wie78].

**Corollary 2.3** *Let* $U \subseteq \mathcal{R}$. $U$ *is finitely identifiable if and only if there exist a numbering* $\psi \in \mathcal{P}^2$ *and a function* $d \in \mathcal{R}$ *satisfying*

*1.* $U \subseteq \mathcal{P}_\psi$,

*2.* $\forall i, j \in \mathbb{N}$ $[i \neq j \ \Rightarrow \ \psi_i \neq_{d(i)} \psi_j]$.

[5] some authors use the notation FIN instead of $EX_0$

#### 2.2.3 Identification in the Limit with Special Intermediate Hypotheses – the Inference Criteria CONS, TOTAL and CONS-TOTAL

A very natural learning behaviour is to construct only consistent intermediate hypotheses, i.e. hypotheses agreeing with the information received so far. The resulting inference criterion was first studied in [Ba74b].

**Definition 2.3 (Consistent Identification)** *Assume $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ is called identifiable consistently with respect to $\psi$ if and only if there exists an $S \in \mathcal{P}$ satisfying*

1. $U \in EX_\psi(S)$,
2. $\forall f \in U\ \forall n \in \mathbb{N}\ [\psi_{S(f[n])} =_n f]$
   *(we say that $S(f[n])$ is a consistent hypothesis for $f[n]$ wrt $\psi$).*

*We also write: $U \in CONS_\psi(S)$.*

$CONS_\psi := \{U \mid U$ *is identifiable consistently wrt* $\psi\}$.

$CONS := \bigcup_{\psi \in \mathcal{P}^2} CONS_\psi$.

The following characterization of consistent identifiability is again taken from [Wie78].

**Theorem 2.4** *Let $U \subseteq \mathcal{R}$. $U$ is identifiable consistently if and only if there exist a numbering $\psi \in \mathcal{P}^2$ and a predicate $d \in \mathcal{R}^3$ satisfying*

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

Whenever a hypothesis put out by a learning strategy is a number of a non-total function in the given hypothesis space, our strategy must be wrong, because the functions to be learned are all taken from $\mathcal{R}$. Thus it is quite reasonable to demand just total functions to be described by the intermediate hypotheses.

**Definition 2.4 (Identification with Total Intermediate Hypotheses)** *Assume $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ is called identifiable wrt $\psi$ with total intermediate hypotheses if and only if there exists an $S \in \mathcal{P}$ satisfying*

1. $U \in EX_\psi(S)$,
2. $\forall f \in U\ \forall n \in \mathbb{N}\ [\psi_{S(f[n])} \in \mathcal{R}]$.

*We also write: $U \in TOTAL_\psi(S)$.*
*$TOTAL_\psi$ and $TOTAL$ are defined by analogy with Definition 2.3.*

More information about the TOTAL-criterion can be found in [JB81]. For the proof of the corresponding characterization theorem see [Wie78].

**Theorem 2.5** *Let $U \subseteq \mathcal{R}$. $U$ is identifiable with total intermediate hypotheses if and only if there exist a numbering $\psi \in \mathcal{P}^2$ and a function $d \in \mathcal{R}$ satisfying*

1. $U \subseteq \mathcal{P}_\psi$,
2. $\forall f \in U \; \forall i \in \mathbb{N} \; [\psi_i =_{d(i)} f \Rightarrow \psi_i \in \mathcal{R}]$.

Since both properties – consistency *and* totality – restricting the set of permitted intermediate hypotheses seem quite natural, one might also combine these criteria. The resulting inference criterion is defined below.

**Definition 2.5 (Identification with Consistent and Total Hypotheses)** *Let $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ is called identifiable wrt $\psi$ with consistent and total intermediate hypotheses if and only if there exists an $S \in \mathcal{P}$ satisfying*

1. $U \in CONS_\psi(S)$ *and*
2. $U \in TOTAL_\psi(S)$.

*We also write: $U \in CONS\text{-}TOTAL_\psi(S)$.*
*$CONS\text{-}TOTAL_\psi$ and $CONS\text{-}TOTAL$ are defined as usual.*

It is not hard to prove that in the case of learning with respect to acceptable numberings consistency can easily be fulfilled whenever the condition of totality is satisfied. The reason is the effective decidability of consistency of total intermediate hypotheses. If the consistency test for a hypothesis put out by a suitable TOTAL-strategy turns out negative, one may put out any consistent and total intermediate hypothesis instead. For more details see [JB81].

**Proposition 2.6** *$TOTAL = CONS\text{-}TOTAL$.*

Though TOTAL = CONS-TOTAL, we will study learning problems for which consistency might not be achieved from TOTAL-strategies in such an easy way (or perhaps might not be achieved at all).

### 2.2.4 Behaviourally Correct Identification – the Inference Criteria BC and BC*

For an arbitrary learning process the mere existence of a certain time, after which all hypotheses are correct, is not sufficient for explanatory correct identification, because the EX-criterion also demands convergence to *a single* hypothesis. If that additional demand is omitted, we talk of "behaviourally correct" identification.

**Definition 2.6 (Behaviourally Correct Identification)** *Let* $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. *U is called BC-identifiable wrt* $\psi$ *if and only if 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 also write* $U \in BC_\psi(S)$ *and define* $BC_\psi$ *and BC as usual.*

Behaviourally correct identifiability has also been characterized by Wiehagen in [Wie78]. But for our purpose the following characterization proved in [Oy98] is more useful.

**Theorem 2.7** *Let* $U \subseteq \mathcal{R}$. *U is BC-identifiable if and only if there exist* $\psi \in \mathcal{P}^2$ *and* $d \in \mathcal{R}^2$ *satisfying the following properties:*

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 behaviourally correct identification provides more learning power than explanatory correct identification – a proof can be found in [Ba74a] – there are still classes of recursive functions which are not BC-identifiable. In [CS83] we find a variation of BC-identification, which allows learnability of the whole class $\mathcal{R}$.

**Definition 2.7 (BC-Identification with Finitely Many Anomalies)** *Let* $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. *U is called BC-identifiable wrt* $\psi$ *with finitely many anomalies if and only if 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 also write* $U \in BC^*_\psi(S)$ *and use the notations* $BC^*_\psi$ *and* $BC^*$ *by analogy with the previous definitions.*

A proof of Theorem 2.8 can be found in [CS83].

**Theorem 2.8** $\mathcal{R} \in BC^*$.

Thus we do not need a characterization of BC$^*$ by means of partial-recursive numberings $\psi \in \mathcal{P}^2$.

### 2.2.5 Comparison of Identification Criteria

For the inference criteria introduced in the previous sections the following comparison results have been proved:

**Theorem 2.9**

1. $EX_0 \subset TOTAL = CONS\text{-}TOTAL \subset CONS \subset EX \subset BC \subset BC^* = \wp\mathcal{R}$.
2. $\forall m \geq 1$ $[EX_m \# CONS \wedge EX_m \# TOTAL]$ *(see [Zi99]).*
3. $\forall m \in \mathbb{N}$ $[EX_m \subset EX_{m+1} \subset EX]$ *(see [CS83]).*

The proof of "$\mathrm{EX}_0 \subset$ TOTAL = CONS-TOTAL" is straightforward and a verification of TOTAL $\subset$ CONS can be found in [JB81]. For a proof of CONS $\subset$ EX see [Ba74b]. Citations for the other results have already been given above.

# 3 Definition of Uniform Learnability

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 $\mathcal{P}_{\varphi^b}$. For convenience we will denote this class by $\mathcal{R}_b$, i.e.

$$\mathcal{R}_b := \mathcal{R}_{\varphi^b} = \mathcal{R} \cap \mathcal{P}_{\varphi^b} \text{ for } b \in \mathbb{N}.$$

Note that the definition of $\mathcal{R}_b$ always depends on our fixed acceptable numbering $\varphi$.
Similarly each subset $B \subseteq \mathbb{N}$ describes a set $\mathcal{R}_B := \{\mathcal{R}_b \mid b \in B\}$ of classes of recursive functions.
From now on let $\mathcal{I}$ denote the set of all previously declared inference criteria:

$$\mathcal{I} := \{\mathrm{EX}, \mathrm{CONS}, \mathrm{TOTAL}, \mathrm{CONS\text{-}TOTAL}, \mathrm{BC}, \mathrm{BC}^*\} \cup \{\mathrm{EX}_m \mid m \in \mathbb{N}\}$$

## 3.1 The Learning Model

**Definition 3.1 (Uniform Learnability)** *Let $I, I'$ be elements of $\mathcal{I}$ satisfying $I \subseteq I'$. A set $J \subseteq \wp\mathcal{R}$ of sets of recursive functions is said to be uniformly learnable with respect to $I$ and $I'$ if and only if 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 $J \in uni(I, I')$. The partial-recursive function $S$ in our third condition will be called a uniform strategy for $J$ wrt $I$ and $I'$; $B$ is called description set for $J, I$ and $I'$.*

Thus $\mathrm{uni}(I, I')$ represents the set of all subsets of $\wp\mathcal{R}$ which are uniformly learnable with respect to $I$ and $I'$.
In order to prove the uniform learnability of a subset $J \subseteq \wp\mathcal{R}$ wrt $I, I' \in \mathcal{I}$ we have to specify the following parameters:

- the set $B \subseteq \mathbb{N}$ describing the classes to be learned,
- the (possibly distinct) numberings $\psi \in \mathcal{P}^2$ serving as hypothesis spaces for the particular classes $\mathcal{R}_b$ $(b \in B)$,
- 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 inference criteria $I$ and $I'$ at all?

2. Which classes $J \subseteq \wp\mathcal{R}$ remain learnable in the sense of $\mathrm{uni}(I, I')$, 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. One part of the work presented here deals with the first question, which turns out to be less difficult than the second. Furthermore we regard lots of special cases of our second problem, where we concentrate our interest on the parameters "description set $B$" and "hypothesis spaces $\psi$". For the purpose of clear arrangement in our propositions and argumentations we have to agree on some more notations first.

### 3.1.1 The Description Set $B$

**Definition 3.2 (Uniform Learning from Special Description Sets)** *Let $I, I' \in \mathcal{I}$, $I \subseteq I'$ and a fixed set $B \subseteq \mathbb{N}$ be given. A set $J \subseteq \wp\mathcal{R}$ will be called uniformly learnable wrt $I$ and $I'$ from description set $B$ if and only if $J \in uni(I, I')$ and the description set for $J, I, I'$ used according to Definition 3.1 equals $B$. We will write $J \in uni_B(I, I')$ for short.*

$\mathrm{uni}_B(I, I')$ thus contains exactly those subsets of $\wp\mathcal{R}$, which can be learned uniformly with respect to $I$ and $I'$ provided that our parameter "description set" is specified by the set $B \subseteq \mathbb{N}$. On condition that $J \in \mathrm{uni}_B(I, I')$ we can obviously conclude $\mathcal{R}_B \in \mathrm{uni}_B(I, I')$. Considering the fact that all classes in $\mathcal{I}$ are closed under inclusion furthermore allows a reverse argumentation: $\mathcal{R}_B \in \mathrm{uni}_B(I, I')$ implies $J \in \mathrm{uni}_B(I, I')$ for all $J \preceq \mathcal{R}_B$. Therefore the sets $J \in \mathrm{uni}(I, I')$ are characterized by those sets $B \subseteq \mathbb{N}$ which are suitable description sets for uniform learning of *some* set $J' \subseteq \wp\mathcal{R}$ in the sense of Definition 3.2:

**Lemma 3.1** *Assume $I, I' \in \mathcal{I}$, $J \subseteq \wp\mathcal{R}$. Then $J \in uni(I, I')$ if and only if there exists a set $B \subseteq \mathbb{N}$ satisfying*

1. *$\mathcal{R}_B \in uni_B(I, I')$ and*

2. *$J \preceq \mathcal{R}_B$.*

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 $\mathcal{R}_B \in \mathrm{uni}(I, I')$ 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 $J \subseteq \wp\mathcal{R}$ were uniformly learnable wrt $I, I'$ from description set $B$ at all. Under these circumstances we may consider the description set $B$ to be unsuitable for uniform learning with respect to $I$ and $I'$. Let us agree on some notation to express the suitability of a description set $B$:

**Definition 3.3 (Suitable Description Sets)** *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 $\mathcal{R}_B \in uni_B(I, I')$. The class of all description sets suitable in that sense will be denoted by $suit(I, I')$.*

These considerations raise the question, whether there are certain specific properties characterizing our appropriate description sets $B \in \text{suit}(I, I')$. What kinds of "natural properties" of the sets $B \subseteq \mathbb{N}$ are necessary or sufficient for their suitability? Which properties seem to be unsuitable and for what reason? Interesting special cases might be:

- $B$ is finite.
- $B$ describes numberings the recursive cores of which consist of just one element, exclusively, i.e. $\forall b \in B\ [\text{card}\ \mathcal{R}_b = 1]$.
- $B$ describes numberings with finite recursive cores, exclusively.
- $B$ describes recursive numberings, exclusively, i.e. $\forall b \in B\ [\varphi^b \in \mathcal{R}^2]$.
- $B$ represents the whole class $I$, i.e. $\forall U \in I\ \exists b \in B\ [U \subseteq \mathcal{R}_b]$.

### 3.1.2 The Hypothesis Spaces $\psi$

**Definition 3.4 (Uniform Learning in Special Hypothesis Spaces)**
*Let $I, I' \in \mathcal{I}$ be given. Let $h : \mathbb{N} \to \mathbb{N}$ denote any function. A set $J \subseteq \wp\mathcal{R}$ is called uniformly learnable with respect to $I$ and $I'$ by interpretation function $h$, if and only if there exists a set $B \subseteq \mathbb{N}$ which fulfills the following conditions:*

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

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

Note that the interpretation function $h$ in our definition is not necessarily computable or total. Of course it might happen, that we want 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 notations $J \in \text{uni}_{B,[h]}(I, I')$ as well as $J \in \text{uni}_{B,\tau}(I, I')$ by analogy.

Via the function $h$ each description $b \in B$ obtains an associated hypothesis space $\varphi^{h(b)}$, by means of which we can interprete the hypotheses produced by the strategy $\lambda x.S(b, x)$. We may come up with several quite obvious possibilities concerning the choice of $h$:

- $h$ is computable; i.e. the hypothesis spaces required or appropriate for the learnability of $\mathcal{R}_b$ ($b \in B \subseteq \mathbb{N}$) in the sense of $I'$ can be computed uniformly from $b$.
- $h(b) = b$ for all $b \in \mathbb{N}$; i.e. the hypotheses produced by $\lambda x.S(b, x)$ on initial segments of functions in $\mathcal{R}_b$ ($b \in \mathbb{N}$) should be interpreted as numbers in the hypothesis space $\varphi^b$ given by $b$.
- $h(b) = j$ for a fixed $j \in \mathbb{N}$ and all $b \in \mathbb{N}$; i.e. the hypotheses produced by $S$ are always interpreted with respect to the same chosen hypothesis space $\varphi^j$. We are particularly interested in the case that $\varphi^j$ is an acceptable numbering.

Regarding the aspect of applications of our theoretical model it seems quite practical to choose one of those possibilities for the specification of our function $h$: in the Definition 3.1 of uniform learning we require for each $b$ in the description set $B$ the mere *existence* of a hypothesis space appropriate for learning $\mathcal{R}_b$ in the sense of $I'$. But whenever there is no computable function assigning an index for that hypothesis space to $b$, the user cannot associate particular functions with the hypotheses generated by $\lambda x.S(b, x)$, because he doesn't know, *which* hypothesis space he should interpret them in. Thus in contrast to learning with respect to a known hypothesis space there is no straightforward way to assign a specific function to a given index. On the other hand, if the user knows a computable interpretation function $h$, he can determine the value $h(b)$ from $b$ as an index for the hypothesis space $\varphi^{h(b)}$. Afterwards a hypothesis $S(b, x) = i \in \mathbb{N}$ can be interpreted unambiguously as an index for the function $\lambda y.\varphi^{h(b)}(i, y) = \varphi_i^{h(b)}$.
For our purpose it is interesting, to what extent the class of all subsets $J \subseteq \wp\mathcal{R}$ uniformly learnable with respect to given criteria $I, I' \in \mathcal{I}$ is restricted by a fixed mapping chosen according to the possibilities listed above. In particular we want to know, *which* sets $J$ remain learnable and whether they can be classified by means of specific properties.

### 3.1.3 The Uniform Strategy $S$

For the purpose of solving inductive learning problems we often find certain procedures, which we intuitively regard as quite promising. A straightforward example is a strategy searching for objects in the hypothesis space, that "match" (i.e. are consistent with) the information read about the target object so far. As these procedures in some ways appear "natural" or "reasonable in content" to us, we often try to model them in the working method of our computable learning strategies. A typical product of this kind is Gold's "Identification by Enumeration" introduced in [Go67]. Obviously, similar techniques can also be used as approaches for solving several learning problems of the classical type by a uniform strategy. For our further studies it may be useful to model a strategy with certain properties in order to analyse its learning power, i.e. to examine what classes $J \subseteq \wp\mathcal{R}$ are uniformly learnable with the given strategy. Therefore we introduce the following notation for uniform identification of a set $J \subseteq \wp\mathcal{R}$ by a fixed strategy $S \in \mathcal{P}^2$.

**Definition 3.5 (Uniform Learning with Special Strategies)**
*Let $I, I' \in \mathcal{I}$, $S \in \mathcal{P}^2$. Then $uni(I, I')(S)$ will denote the set of all classes $J \subseteq \wp\mathcal{R}$, for which $S$ is a uniform strategy wrt $I$ and $I'$ in the sense of Definition 3.1.*

Thus $\text{uni}(I, I')(S)$ contains exactly those subsets of $\wp\mathcal{R}$, which can be learned uniformly with respect to $I$ and $I'$ providing a specification of the parameter "learning strategy" by the function $S$. By analogy with the definitions above we use the notations $\text{uni}_B(I, I')(S)$, $\text{uni}_{[h]}(I, I')(S)$ as well as $\text{uni}_{B,[h]}(I, I')(S)$ (respectively with $\tau$ instead of $[h]$ in the case of a constant interpretation function $h$).

## 3.2 Basic Results

Our agreement on notations and definitions provided so far now allows us to formulate some basic results on uniform learning. Although the corresponding proofs are quite simple, those 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 3.2 note that all classes $I \in \mathcal{I}$ are closed with respect to the inclusion of sets.

**Proposition 3.2** *Let $I, I' \in \mathcal{I}$, $I \subseteq I'$, $J \subseteq \wp\mathcal{R}$. From $J \in uni(I, I')$ we can conclude $J \subseteq I$.*

**Proof:** Let $J \in \text{uni}(I, I')$. Then there exists a set $B \subseteq \mathbb{N}$ which fulfills $J \in \text{uni}_B(I, I')$.
$\Rightarrow$ $\mathcal{R}_B \subseteq I$ and $J \preceq \mathcal{R}_B$.

$\Rightarrow\ \forall U \in J\ \exists b \in B\ [U \subseteq \mathcal{R}_b \in I]$.
$\Rightarrow\ \forall U \in J\ [U \in I]$, because $I$ is closed with respect to the inclusion of sets.
$\Rightarrow\ J \subseteq I$.

*qed Proposition 3.2*

As a direct conclusion from Proposition 3.2 and the definition of uniform learning we state:

**Corollary 3.3** *Let $I, I' \in \mathcal{I},\ I \subseteq I'$. Then $uni(I,I) \subseteq uni(I,I') \subseteq \wp I$.*

Except for the case of consistent learning, any strategy identifying a given class $U \subseteq \mathcal{R}$ with respect to some inference 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. This bound is increased whenever a new input example is exposed to the new strategy. 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". Whereas in the case of $I \in \{\mathrm{EX}, \mathrm{BC}, \mathrm{BC}^*\} \cup \{\mathrm{EX}_m \mid m \in \mathbb{N}\}$ the temporary hypothesis can be fixed deliberately from the hypothesis space, we have to find a temporary hypothesis representing a *total* function if the given criterion is TOTAL. Now we transfer these observations to the level of uniform learning.

**Proposition 3.4** *Let $I, I' \in \mathcal{I}\backslash\{CONS, CONS\text{-}TOTAL\},\ B \subseteq \mathbb{N}$ and let $h : \mathbb{N} \to \mathbb{N}$ be any function. Assume $\mathcal{R}_B \in uni_{B,[h]}(I,I')$. Then there exist a total recursive function $S \in \mathcal{R}^2$ and a function $g$ recursive in $h$ satisfying $\mathcal{R}_B \in uni_{B,[g]}(I,I')(S)$.*

**Remark** If we assume $I' \in \{\mathrm{EX}, \mathrm{BC}, \mathrm{BC}^*\} \cup \{\mathrm{EX}_m \mid m \in \mathbb{N}\}$, then Proposition 3.4 remains valid with $g = h$.

The proof of Proposition 3.4 is arranged exactly as explained above. The reason why in general we cannot keep the numberings $\varphi^{h(b)}$ ($b \in B$) as hypothesis spaces is that the temporary hypotheses needed in the case of learning with respect to TOTAL cannot be computed uniformly from $b$, but just from $h(b)$.

**Proposition 3.5** *Let $I \in \mathcal{I},\ h : \mathbb{N} \to \mathbb{N}$. If $B \subseteq \mathbb{N}$ is a finite set with $\mathcal{R}_b \in I_{\varphi^{h(b)}}$ for all $b \in B$, then $\mathcal{R}_B \in uni_{B,[h]}(I,I')$.*

The proof is obvious: a finite number of strategies – each learning one of the given recursive cores with respect to $I'$ – 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 3.1 without specifying the description set $B$, the hypothesis spaces $\psi$ or the learning strategy $S$ in advance. We just choose two inference criteria $I, I' \in \mathcal{I}$ satisfying $I \subseteq I'$ and try to characterize the subsets $J \subseteq \wp\mathcal{R}$ contained in $\text{uni}(I, I')$. By $\text{uni}(I, I') \subseteq \wp I$ we have already found a necessary condition for the uniform learnability of a set $J \subseteq \wp\mathcal{R}$. In our next theorem the necessary condition supplied by Corollary 3.3 additionally turns out to be sufficient.

**Theorem 4.1** *Assume $I, I' \in \mathcal{I},\ I \subseteq I'$. Then $uni(I, I') = \wp I$.*
*Especially we have $uni(I, I) = \wp I$.*

For the proof of Theorem 4.1 we need the following lemma, which tells us that any class of functions learnable with respect to a given criterion is contained in some recursive core learnable with respect to the same criterion.

**Lemma 4.2** *Let $I \in \mathcal{I}$. For each $U \in I$ there exists a numbering $\psi \in \mathcal{P}^2$ such that*

1. *$U \subseteq \mathcal{P}_\psi$,*
2. *$\mathcal{R}_\psi \in I$.*

**Proof of Lemma 4.2:** The idea is to make use of the numberings supplied by the characterization theorems 2.1 up to 2.7.

*Case 1: $I = EX$.*
Assume $U \in \text{EX}$. Theorem 2.1 provides a numbering $\psi \in \mathcal{P}^2$ and a function $d \in \mathcal{R}$ satisfying

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

This trivially implies

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

According to Theorem 2.1 this statement is again sufficient for the learnability of $\mathcal{R}_\psi$ with respect to the inference criterion EX; thus we conclude $\mathcal{R}_\psi \in \text{EX}$. This proves the claim of Lemma 4.2 for the case $I = \text{EX}$.

*Case 2: $I \in \{BC, CONS\} \cup \{EX_m \mid m \in \mathbb{N}\}$.*

The argumentation in our first case can be applied by analogy.

*Case 3: $I$ = TOTAL.*

Assume $U \in$ TOTAL. Let $\tau$ be an acceptable numbering and $S \in \mathcal{P}^2$ a strategy such that $U \in \text{TOTAL}_\tau(S)$. The construction of the demanded numbering $\psi$ proceeds in a way similar to the proofs in [Wie78]:

We define a set $M$ by

$$M := \{(z,n) \in \mathbb{N}^2 \mid \text{for all } x \leq n \ [\tau_z(x) \text{ is defined}] \text{ and } S(\tau_z[n]) = z\}.$$

The set $M$ describes all initial segments of functions on which the strategy $S$ produces a correct hypothesis. Obviously $M$ is recursively enumerable, so that we can make use of a function $e \in \mathcal{R}$, which enumerates $M$. If the pair $(z,n)$ is the $i$-th element in the enumeration of $e$ (i.e. $e(i) = (z,n)$), we define the $i$-th function of our numbering $\psi$ by the following instructions:

$$\psi_i(x) := \begin{cases} \tau_z(x) & x \leq n \text{ or} \\ & x > n \text{ and } \forall y, m \leq x\ [\tau_{S(\tau_z[y])}(m)\downarrow] \\ & \quad \text{and } \forall y \in \{n, \ldots, x\}\ [\tau_z(y)\downarrow \text{ and } S(\tau_z[y]) = z] \\ \uparrow & \text{otherwise} \end{cases}$$

for all $x \in \mathbb{N}$.

It is easy to verify that $\psi$ is a computable numbering of partial-recursive functions. It remains to prove the following claims:

**Claim 1** $U \subseteq \mathcal{P}_\psi$,

**Claim 2** $\mathcal{R}_\psi \in$ TOTAL.

*Proof of Claim 1:*

Let $f \in U$. As $f$ is identified by $S$ with total intermediate hypotheses with respect to $\tau$, there exist numbers $z, n \in \mathbb{N}$ such that $\tau_z = f$, $S(\tau_z[x]) = z$ for all $x \geq n$ as well as $\tau_{S(\tau_z[y])} \in \mathcal{R}$ for all $y \in \mathbb{N}$. Since $\tau_z = f$ and $S(\tau_z[n]) = z$ we know that $(z,n)$ is an element of $M$. Thus choose $i \in \mathbb{N}$ with $e(i) = (z,n)$. Since furthermore $S(\tau_z[x]) = z$ for all $x \geq n$ and $\tau_{S(\tau_z[y])} \in \mathcal{R}$ for all $y \in \mathbb{N}$, the values of $\psi_i$ must be defined for all arguments. We conclude $\psi_i = \tau_z = f$, and therefore $f \in \mathcal{P}_\psi$.

*qed Claim 1*

*Proof of Claim 2:*

Let $i \in \mathbb{N}$ be an index with $\psi_i \in \mathcal{R}$, $(z,n) := e(i)$.

Claim 2': $\{\psi_i\} \in \text{TOTAL}_\tau(S)$.

Proof: We know that $\psi_i = \tau_z$ and $S(\tau_z[x]) = z$ for all $x \geq n$ as well as $\tau_{S(\tau_z[y])} \in \mathcal{R}$ for all $y \in \mathbb{N}$, because the computation of $\psi_i$ terminates for all arguments.
$\Rightarrow\ \tau_z \in \text{TOTAL}_\tau(S)$.
$\Rightarrow\ \psi_i \in \text{TOTAL}_\tau(S)$.
In summary we obtain $\mathcal{R}_\psi \in \text{TOTAL}_\tau(S)$.

*qed Claim 2*

This proves the claim of Lemma 4.2 for the case $I = \text{TOTAL}$.

*Case 4: $I$ = CONS-TOTAL.*
Here we can use Case 3 and the fact TOTAL = CONS-TOTAL.

*Case 5: $I = BC^*$.*
Since $\mathcal{R} \in \text{BC}^*$, the proof is straightforward in this last case.

Summarizing our five cases we see that we have verified the statement of Lemma 4.2.

*qed Lemma 4.2*

These preparations now enable us to prove Theorem 4.1.

**Proof of Theorem 4.1:** Assume $I, I' \in \mathcal{I},\ I \subseteq I'$. The claim is $\text{uni}(I, I') = \wp I$. Applying Corollary 3.3 we only have to show

$$\wp I \subseteq \text{uni}(I, I').$$

Assume $J \in \wp I$ and choose a fixed acceptable numbering $\tau \in \mathcal{P}^2$ of $\mathcal{P}$. We have to verify $J \in \text{uni}(I, I')$.
For each $U \in J$ Lemma 4.2 supplies a numbering $\psi \in \mathcal{P}^2$ such that $U \subseteq \mathcal{P}_\psi$ and $\mathcal{R}_\psi \in I$. Since $\varphi$ is an acceptable numbering, for each of these numberings $\psi$ there is an index $c \in \mathbb{N}$ with $\varphi^c = \psi$. We obtain:

$$\exists C \subseteq \mathbb{N}\ [\mathcal{R}_C \subseteq I \text{ and } \forall U \in J\ \exists c \in C\ [U \subseteq \mathcal{R}_{\varphi^c}]].$$

Of course for each $c \in C$ there is a strategy $S_c \in \mathcal{P}$ identifying $\mathcal{R}_{\varphi^c}$ in the sense of $I$. On the other hand, each strategy $S_c$ possesses an index $k_c$ in $\tau$ (i.e. $\tau_{k_c} = S_c$ for all $c \in C$). We can conclude

$$\forall c \in C\ [\mathcal{R}_{\varphi^c} \in I(\tau_{k_c})].$$

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:

- 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 $k_c\uparrow^\infty$).
- They contain encodings of $\tau$-indices for strategies identifying their recursive cores according to $I$.

Now a uniform learner for the target class $J$ just has to read the indices of the particular strategies and afterwards simulate their jobs with the help of the functions associated by $\tau$. More formally:

For each $c \in C$ let the numbering $\eta^{[c]} \in \mathcal{P}^2$ be defined by

$$\eta_0^{[c]} := k_c\uparrow^\infty,\ \eta_{n+1}^{[c]} := \lambda x.\varphi_n^c(x) \text{ for } n \in \mathbb{N}.$$

The numberings $\eta^{[c]} \in \mathcal{P}^2$ again have indices in $\varphi$, which we collect in the set $B \subseteq \mathbb{N}$:

$$B := \{b \in \mathbb{N} \mid \exists c \in C\ [\eta^{[c]} = \varphi^b]\}.$$

In order to make our uniform strategy work according to the informal description above, we set:

$$S(b, f^n) := \tau_{\varphi_0^b(0)}(f^n) \text{ for } b \in \mathbb{N},\ f^n \in \mathbb{N}.$$

Now the proof of $J \in \text{uni}_B(I, I)(S)$ is straightforward. Since $\text{uni}(I, I) \subseteq \text{uni}(I, I')$ by definition, we obtain $J \in \text{uni}(I, I')$ as has been demanded.

*qed Theorem 4.1*

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

**Corollary 4.3** *Let $I, I' \in \mathcal{I},\ I \subset I'$. Then*

$$uni(I, I) = uni(I, I') \subset uni(I', I').$$

**Proof:** By Theorem 4.1 we know $\text{uni}(I, I) = \text{uni}(I, I')$. Since $\text{uni}(I, I') \subseteq \text{uni}(I', I')$, 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 $U \in I' \backslash I$ and obtain a class $J \in \text{uni}(I', I') \backslash \text{uni}(I, I')$ by defining $J := \{U\}$.

*qed Corollary 4.3*

Of course our uniform strategy defined in the proof of Theorem 4.1 does not really *learn* anything. Since the programs for learning the described classes are coded within the described numberings in advance, it just has to look up these programs and simulate them on its second input argument. 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

## 5.1 Characterizations of Suitability

In this section we investigate the suitability of a given description set $B$, i.e. its influence upon the uniform learnability of $\mathcal{R}_B$ from $B$ with respect to some inference criteria $I, I' \in \mathcal{I}$. We start with a simple, but useful observation.

**Proposition 5.1** *Let $I, I' \in \mathcal{I}$ and $B \subseteq \mathbb{N}$ satisfy the following properties:*

1. $\forall b \in B\ [\mathcal{R}_b \in I]$,
2. $\bigcup_{b \in B} \mathcal{R}_b \in I'$.

*Then $B \in suit(I, I')$.*

The proof of Proposition 5.1 is straightforward from the definitions. As a direct consequence we obtain a simple characterization of the description sets suitable for uniform learning with BC$^*$-strategies:

**Theorem 5.2** *$suit(I, BC^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$ for all $I \in \mathcal{I}$. In particular $suit(BC^*, BC^*) = \wp\mathbb{N}$.*

**Proof:** Choose $I \in \mathcal{I}$. Obviously $\text{suit}(I, \text{BC}^*) \subseteq \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$. Now let $B \subseteq \mathbb{N}$ fulfill $\mathcal{R}_B \subseteq I$. Then $B$ satisfies the following properties:

1. $\forall b \in B\ [\mathcal{R}_b \in I]$,
2. $\bigcup_{b \in B} \mathcal{R}_b \subseteq \mathcal{R} \in \text{BC}^*$.

From Proposition 5.1 we conclude $B \in \text{suit}(I, \text{BC}^*)$. Altogether we obtain $\text{suit}(I, \text{BC}^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$. In particular $\text{suit}(\text{BC}^*, \text{BC}^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq \text{BC}^*\} = \wp\mathbb{N}$, because $\mathcal{R} \in \text{BC}^*$.

*qed Theorem 5.2*

For any criterion $I \in \mathcal{I} \backslash \{\text{BC}^*\}$ we obtain a similar characterization. Since BC is not closed with respect to the union of sets, we cannot use Proposition 5.1 for the corresponding proofs. Instead we choose special hypothesis spaces which simplify the work of a uniform strategy. Again – as in the proof of Theorem 4.1 – we make use of special "tricks", such that the resulting strategy does not really have to do any labour. Although our strategy produces its hypotheses in a very "stupid" way, we can always rely on the existence of appropriate hypothesis spaces such that the conditions of uniform learning are fulfilled.

**Theorem 5.3** *Set* $B := \{b \in \mathbb{N} \mid \mathcal{R}_b \in BC\}$*. Then* $B \in suit(BC, BC)$.

**Proof:** Fix an acceptable numbering $\tau \in \mathcal{P}^2$. Since each class of recursive functions learnable in the sense of BC can be identified with respect to *any* acceptable numbering by a *recursive* strategy, we conclude:

$$\forall b \in B\ \exists S_b \in \mathcal{R}\ [\mathcal{R}_b \in \mathrm{BC}_\tau(S_b)].$$

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 a numbering $\psi^{[b]} \in \mathcal{P}^2$ by

$$\psi^{[b]}_i(x) := \tau_{S_b(i)}(x) \text{ for any } i, x \in \mathbb{N}.$$

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 obviously be used by a uniform BC-strategy: for all $b, n \in \mathbb{N}$ and $f \in \mathcal{R}$ we define

$$S(b, f[n]) := f[n]$$

and try to verify $\mathcal{R}_b \in \mathrm{BC}_{\psi^{[b]}}(\lambda x.S(b, x))$ for each $b \in B$. Our argumentation proceeds as follows:

Let $b \in B$.
$\Rightarrow$ $\forall f \in \mathcal{R}_b\ \forall^\infty n\ [\ \tau_{S_b(f[n])} = f\ ]$, because $\mathcal{R}_b \in \mathrm{BC}_\tau(S_b)$.
$\Rightarrow$ $\forall f \in \mathcal{R}_b\ \forall^\infty n\ [\ \psi^{[b]}_{f[n]} = f\ ]$ from the definition of $\psi^{[b]}$.
$\Rightarrow$ $\forall f \in \mathcal{R}_b\ \forall^\infty n\ [\ \psi^{[b]}_{S(b,f[n])} = f\ ]$ because of the definition of $S$.
$\Rightarrow$ $\mathcal{R}_b \in \mathrm{BC}_{\psi^{[b]}}(\lambda x.S(b, x))$.

Thus we know a strategy $S \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in \mathrm{BC}_{\psi^{[b]}}(\lambda x.S(b, x))$ for all $b \in B$. This implies $B \in \mathrm{suit}(\mathrm{BC}, \mathrm{BC})$.

*qed Theorem 5.3*

Hence we obtain the desired characterization of sets suitable for uniform behaviourally correct identification:

**Corollary 5.4** $suit(I, BC) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$ *for all* $I \in \mathcal{I}\backslash\{BC^*\}$.

**Proof:** Assume $I \in \mathcal{I}\backslash\{\mathrm{BC}^*\}$.

- If $B \in \mathrm{suit}(I, \mathrm{BC})$ and $b \in B$ we have $\mathcal{R}_b \in I$. Thus $\mathcal{R}_B \subseteq I$.

- Let $B$ be a set of integers satisfying $\mathcal{R}_B \subseteq I$. Then $\mathcal{R}_B \subseteq \mathrm{BC}$ and furthermore $B \subseteq \{b \in \mathbb{N} \mid \mathcal{R}_b \in \mathrm{BC}\} \in \mathrm{suit}(\mathrm{BC}, \mathrm{BC})$. Since $\mathrm{suit}(\mathrm{BC}, \mathrm{BC})$ is closed under inclusion, we conclude $B \in \mathrm{suit}(\mathrm{BC}, \mathrm{BC})$.

This proves Corollary 5.4.

*qed Corollary 5.4*

As we have seen, the trick of encoding much information within the description sets (see Theorem 4.1) or within the hypothesis spaces (see Theorem 5.3) often supplies quite simple uniform strategies with a huge learning power. If for any given pair of criteria from $\mathcal{I}$ such tricks were successful, uniform learnability with respect to those criteria would be trivial. 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 strategy identifying *any* recursive core consisting of just a single element *finitely* from its description, but on the other hand there is *no* uniform strategy identifying all recursive cores consisting of two elements *in the limit* from their descriptions. In view of classical learning problems any classes consisting of just two elements are not much more complex than classes consisting of one element (both kinds are finitely identifiable), whereas their complexity is very different regarding uniform learning problems. Proposition 5.5 concerns the suitability of the set describing all recursive cores consisting of just one element.

**Proposition 5.5** $\{b \in \mathbb{N} \mid \mathit{card}\ \mathcal{R}_b = 1\} \in \mathit{suit}(EX_0, EX_0)$.

**Proof:** Let $B := \{b \in \mathbb{N} \mid \mathrm{card}\ \mathcal{R}_b = 1\}$. Then of course $\mathcal{R}_b \in \mathrm{EX}_0$ for all $b \in B$, i.e. $\mathcal{R}_B \subseteq \mathrm{EX}_0$. Since for all $f \in \mathcal{R}$ there exists a numbering $\psi \in \mathcal{P}^2$ with $\psi_0 = f$, the strategy $S$ defined by

$$S(b, f[n]) := 0 \text{ for any } f \in \mathcal{R},\ b, n \in \mathbb{N}$$

is appropriate for the purpose of learning $\mathcal{R}_B$ uniformly from $B$ with respect to $\mathrm{EX}_0$ and $\mathrm{EX}_0$.

*qed Theorem 5.5*

Now, in contrast to Proposition 5.5 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 hypotheses spaces.

**Theorem 5.6** *Let* $B := \{b \in \mathbb{N} \mid \mathit{card}\ \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} \leq 2\}$. *Then* $B \notin \mathit{suit}(EX, EX)$.

**Proof:** Assume $B \in \text{suit}(\text{EX}, \text{EX})$. Then there exists an $S \in \mathcal{R}^2$ satisfying $\mathcal{R}_b \in \text{EX}(\lambda x.S(b,x))$ for all $b \in B$. We will prove the existence of an index $b_0 \in B$ describing a recursive core which cannot be identified in the limit by the strategy $\lambda x.S(b_0, x)$. This strategy will fail for at least one function $f \in \mathcal{R}_{b_0}$ by either changing its hypotheses infinitely often or by producing incorrect hypotheses in infinitely many learning steps. That means that the sequence $(S(f[n]))_{n\in\mathbb{N}}$ of hypotheses will either converge to an incorrect hypothesis or not converge at all. To achieve this we define a function $\psi \in \mathcal{P}^3$ according to the following algorithm:

**Algorithm for the definition of $\boldsymbol{\lambda x.\lambda y.\psi(b,x,y)}$ for $\boldsymbol{b} \in \mathbb{N}$:**

First we omit the definition of $\lambda y.\psi(b,0,y)$. In general step $n$, $n \in \mathbb{N}$ is used for the definition of $\lambda y.\psi(b,2n{+}1,y)$ and $\lambda y.\psi(b,2n{+}2,y)$. The function $\lambda y.\psi(b,0,y)$ will be defined afterwards.

Set $\alpha_0 := 0$.

*Step 0:*

$\psi(b,1,0) := 0,\ \psi(b,2,0) := 0,\ \psi(b,2,1) := 1,$

$h_0 := S(b,0),\ y := 1.$

As long as $S(b, 0^{y+1})(= S(b, \alpha_0 0^y)) = h_0$ and $S(b, 010^y)(= S(b, \alpha_0 10^y)) = h_0$, set

$$\{\ \psi(b,1,y) := 0;\ \psi(b,2,y+1) := 0;\ y := y+1;\ \}$$

Afterwards we define $k_0 := y$ and

$$\alpha_1 := \begin{cases} 0^{k_0+1} & S(b,0^{k_0+1}) \neq h_0 \\ 010^{k_0} & S(b,0^{k_0+1}) = h_0 \text{ and } S(b,010^{k_0}) \neq h_0 \end{cases}$$

Altogether we have defined up to now:

$$\lambda y.\psi(b,1,y) = \begin{cases} 0^{k_0+1}\uparrow^\infty & k_0\downarrow \\ 0^\infty & k_0\uparrow \end{cases} \text{ as well as } \lambda y.\psi(b,2,y) = \begin{cases} 010^{k_0}\uparrow^\infty & k_0\downarrow \\ 010^\infty & k_0\uparrow \end{cases}$$

*End Step 0*

*Step $n$ for $n > 0$:*

$(\psi(b,2n+1,0), \ldots, \psi(b,2n+1,|\alpha_n|-1)) := \alpha_n,$

$(\psi(b,2n+2,0), \ldots, \psi(b,2n+2,|\alpha_n|)) := \alpha_n 1.$

$h_n := S(b,\alpha_n),\ y := 1.$

As long as $S(b,\alpha_n 0^y) = h_n$ and $S(b,\alpha_n 10^y) = h_n$, set

$$\{\ \psi(b,2n+1,|\alpha_n|-1+y) := 0;\ \psi(b,2,|\alpha_n|+y) := 0;\ y := y+1;\ \}$$

Afterwards we define $k_n := y$ and

$$\alpha_{n+1} := \begin{cases} \alpha_n 0^{k_n} & S(b,\alpha_n 0^{k_n}) \neq h_n \\ \alpha_n 10^{k_n} & S(b,\alpha_n 0^{k_n}) = h_n \text{ and } S(b,\alpha_n 10^{k_n}) \neq h_n \end{cases}$$

Of course all these values shall only be defined if the analogous values in Step $n-1$ are also defined. We can summarize:

$$\lambda y.\psi(b,2n+1,y) = \begin{cases} \uparrow^\infty & n>0 \wedge [k_0\uparrow \vee \ldots \vee k_{n-1}\uparrow] \\ \alpha_n 0^{k_n}\uparrow^\infty & k_0\downarrow \wedge \ldots \wedge k_n\downarrow \\ \alpha_n 0^\infty & [k_0\downarrow \wedge \ldots \wedge k_{n-1}\downarrow] \wedge k_n\uparrow \end{cases} \quad \text{as well as}$$

$$\lambda y.\psi(b,2n+2,y) = \begin{cases} \uparrow^\infty & n>0 \wedge [k_0\uparrow \vee \ldots \vee k_{n-1}\uparrow] \\ \alpha_n 10^{k_n}\uparrow^\infty & k_0\downarrow \wedge \ldots \wedge k_n\downarrow \\ \alpha_n 10^\infty & [k_0\downarrow \wedge \ldots \wedge k_{n-1}\downarrow] \wedge k_n\uparrow \end{cases}$$

*End Step n*

Finally we define the function $\lambda y.\psi(b,0,y)$ by means of the sequence of the tuples $\alpha_n$, $n \in \mathbb{N}$ of growing length, i.e. $\lambda y.\psi(b,0,y) = \lim_{n\to\infty}(\alpha_n)$. In particular $\lambda y.\psi(b,0,y)$ is initial if and only if $k_n\uparrow$ for an $n \in \mathbb{N}$.

This completes the definition of $\psi$. Now let $b_0$ be an integer satisfying

$$\lambda x.\lambda y.\psi(b_0,x,y) = \varphi^{b_0}.$$

Our contradiction requested to finish the proof will be achieved by verification of the following claims:

**Claim 1** $b_0 \in B$.

**Claim 2** $\mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0,x))$.

*Proof of Claim 1:*

1st case: $k_n\downarrow$ for all $n \in \mathbb{N}$
$\Rightarrow$ $\lambda y.\psi(b_0,0,y) \in \mathcal{R}$ and $\lambda y.\psi(b_0,x,y)$ initial for all $x>0$.
$\Rightarrow$ $\{i \in \mathbb{N} \mid \varphi_i^{b_0} \in \mathcal{R}\} = \{0\}$ $\Rightarrow$ $b_0 \in B$.

2nd case: $n \in \mathbb{N}$ is the least number satisfying $k_n\uparrow$.
$\Rightarrow$ $\lambda y.\psi(b_0,2n+1,y) = \alpha_n 0^\infty$ and $\lambda y.\psi(b_0,2n+2,y) = \alpha_n 10^\infty$.
Obviously $\lambda y.\psi(b_0,m,y)$ is initial for $m \in \mathbb{N}\backslash\{2n+1,2n+2\}$.
$\Rightarrow$ $\{i \in \mathbb{N} \mid \varphi_i^{b_0} \in \mathcal{R}\} = \{2n+1,2n+2\}$ $\Rightarrow$ $b_0 \in B$.

*qed Claim 1*

*Proof of Claim 2:*

1st case: $\varphi_0^{b_0} \in \mathcal{R}$.
$\Rightarrow$ $k_n\downarrow$ for all $n \in \mathbb{N}$.
$\Rightarrow$ $\alpha_{n+1}\downarrow$ and $S(b_0,\alpha_{n+1}) \neq S(b_0,\alpha_{n+2})$ for all $n \in \mathbb{N}$.
Thus, when trying to identify $\varphi_0^{b_0}$, the strategy $\lambda x.S(b_0,x)$ changes its hypotheses infinitely often.
$\Rightarrow$ $\mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0,x))$.

2nd case: There exists a number $n \in \mathbb{N}$ with $\mathcal{R}_{b_0} = \{\varphi^{b_0}_{2n+1}, \varphi^{b_0}_{2n+2}\}$.
$\Rightarrow\ \varphi^{b_0}_{2n+1} = \alpha_n 0^\infty,\ \varphi^{b_0}_{2n+2} = \alpha_n 10^\infty$.
Considering the definition of Step $n$ we observe that
$S(b_0, \alpha_n 0^y) = S(b_0, \alpha_n 10^y) = h_n = S(b_0, \alpha_n)$ for all $y \in \mathbb{N}$.
Thus $\lim_{y\to\infty}(S(b_0, \varphi^{b_0}_{2n+1}[y])) = \lim_{y\to\infty}(S(b_0, \varphi^{b_0}_{2n+2}[y]))$,
though $\varphi^{b_0}_{2n+1} \neq \varphi^{b_0}_{2n+2}$.
$\Rightarrow\ \mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0, x))$.

As we have already observed in the proof of Claim 1, further cases cannot occur.

*qed Claim 2*

So the index $b_0 \in B$ describes a recursive core, which cannot be learned in the sense of the EX-criterion by our uniform strategy $S$, despite its knowledge of the description $b_0$. Hence, contrary to our assumption we conclude $B \notin \mathrm{suit(EX, EX)}$.

*qed Theorem 5.6*

**Corollary 5.7** *$\{b \in \mathbb{N} \mid \mathcal{R}_b$ is finite$\} \notin$ suit(EX, EX).*

**Proof:** This corollary is a direct consequence of Theorem 5.6.

*qed Corollary 5.7*

**Corollary 5.8** *$\forall I \in \mathcal{I}\backslash\{BC, BC^*\}$ [suit($EX_0, I$) $\subset$ suit($EX_0, BC$)].*

**Proof:** Choose $I \in \mathcal{I}\backslash\{\mathrm{BC}, \mathrm{BC}^*\}$. Obviously $\mathrm{suit}(\mathrm{EX}_0, I) \subseteq \mathrm{suit}(\mathrm{EX}_0, \mathrm{BC})$. From Theorem 5.3 and Corollary 5.7 we conclude $\{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite}\} \in \mathrm{suit}(\mathrm{EX}_0, \mathrm{BC})\backslash\mathrm{suit}(\mathrm{EX}_0, I)$.

*qed Corollary 5.8*

Theorem 5.9 is a summary of our main results in this section.

**Theorem 5.9** *Let $I \in \mathcal{I}$ be an inference criterion. Then the following three conditions are equivalent:*

1. *suit($I, I$) $= \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq I\}$,*
2. *suit($EX_0, I$) $= \{B \subseteq \mathbb{N} \mid \mathcal{R}_B \subseteq EX_0\}$,*
3. *$I \in \{BC, BC^*\}$.*

Since we cannot characterize the classes $\mathrm{suit}(I, I')$ for $I, I' \in \mathcal{I}$ and $I \subseteq I' \subset$ BC in such an easy way as for the case $I' \in \{\mathrm{BC}, \mathrm{BC}^*\}$, we use the theory of numberings to achieve results in the style of Theorem 2.1. In our proofs we will use arguments similar to those presented in [Wie78]. As further results are obtained in the same manner, we will only prove our characterization for the case of uniform EX-identification. Note the similarity of our properties to the characteristic condition for identification of languages in the limit from text, introduced in [An80].

**Theorem 5.10** *Assume $I \in \mathcal{I}$, $I \subseteq EX$. Let $B \subseteq \mathbb{N}$ fulfill $\mathcal{R}_B \subseteq I$. Then $B \in suit(I, EX) \iff \exists d \in \mathcal{R}^2\ \forall b \in B\ \exists \psi \in \mathcal{P}^2$*

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

**Proof:** Let $B \subseteq \mathbb{N}$, $I \in \mathcal{I}$ with $\mathcal{R}_B \subseteq I \subseteq \mathrm{EX}$ be given.

*Necessity:*
Assume $B \in \mathrm{suit}(I, \mathrm{EX})$. Choose $S \in \mathcal{R}^2$ such that $\mathcal{R}_B \in \mathrm{uni}_B(I, \mathrm{EX})(S)$.

$$\Rightarrow\ \forall b \in B\ \exists \eta^{[b]} \in \mathcal{P}^2\ [\mathcal{R}_b \in \mathrm{EX}_{\eta^{[b]}}(\lambda x.S(b,x))].$$

For any $b \in \mathbb{N}$ let the set $M_b$ consist of all segments $\alpha \in \mathbb{N}^*$ satisfying

$\exists z, n \in \mathbb{N}$
- $\forall x \le n\ [\varphi^b_z(x)\downarrow]$,
- $(\varphi^b_z(0), \ldots, \varphi^b_z(n)) = \alpha$,
- $n = 0\ \vee\ S(b, \varphi^b_z[n]) \neq S(b, \varphi^b_z[n-1])$.

Thus $M_b$ is the set of all initial segments of functions in $\mathcal{P}_{\varphi^b}$ forcing the strategy $\lambda x.S(b,x)$ to change its mind. Obviously the sets $M_b$ are uniformly r.e. in $b$. Without loss of generality we can assume that the definition of $S$ yields just infinite sets $M_b$, $b \in \mathbb{N}$. Hence define $d \in \mathcal{R}^2$ such that for arbitrary $b \in \mathbb{N}$ the function $\lambda i.d(b,i)$ enumerates the set $M_b$ without repetitions.
Now fix $b \in \mathbb{N}$. We still have to define a function $\psi \in \mathcal{P}^2$ with the desired properties.

Let $i \in \mathbb{N}$. Assume $d(b,i) = \alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1}$. For arbitrary $x \in \mathbb{N}$ we define

$$\psi_i(x) := \begin{cases} \alpha_x & x \le n \\ \eta^{[b]}_{S(b,\alpha)}(x) & x > n \wedge \forall y \in \{n, \ldots, x\} : \\ & \quad [\eta^{[b]}_{S(b,\alpha)}(y)\downarrow\ \wedge\ S(b, \eta^{[b]}_{S(b,\alpha)}[y]) = S(b,\alpha)] \\ \uparrow & \text{otherwise} \end{cases}$$

Note that $\psi$ is computable, but not uniformly computable in $b$, since in general there is no algorithm which – on input $b \in \mathbb{N}$ – produces a program for $\eta^{[b]}$.

Provided $b \in B$, it remains to prove the properties 1, 2 and 3 from Theorem 5.10 for our fixed (but arbitrary) index $b$.

1. $\mathcal{R}_b \subseteq \mathcal{P}_\psi$:
   Choose $f \in \mathcal{R}_b$. Then there exist $z, n \in \mathbb{N}$ such that $\varphi^b_z = f$ and $n = \mathrm{conv}(\lambda x.S(b,x), f)$. Therefore the segment $\alpha := (f(0), \ldots, f(n))$ must be an element of $M_b$ and is thus enumerated by $\lambda k.d(b,k)$. Assume $i \in \mathbb{N}$

fulfills $d(b,i) = \alpha$. Since the hypothesis $S(b,\alpha)$ is correct for $f$ with respect to $\eta^{[b]}$ and will never be changed, we conclude $f = \psi_i$ from the definition of $\psi$. Thus $f \in \mathcal{P}_\psi$.

2. $\forall i \in \mathbb{N}\ [d(b,i) \sqsubseteq \psi_i]$:
   Straightforward from the definition of $\psi$.

3. $\forall i,j \in \mathbb{N}\ [d(b,i) \sqsubseteq d(b,j) \sqsubseteq \psi_i \ \Rightarrow\ i = j]$:
   Assume $i \neq j$ and $d(b,i) \sqsubseteq d(b,j) \sqsubseteq \psi_i$. Since $\lambda k.d(b,k)$ enumerates $M_b$ without repetitions, we obtain $d(b,i) \sqsubset d(b,j)$ and in particular $|d(b,i)| < |d(b,j)|$. From the definition of $M_b$ we know that $d(b,j)$ forces the strategy $\lambda x.S(b,x)$ to change its mind. Therefore $\psi_i(|d(b,j)|-1)$ is undefined and hence $d(b,j) \not\sqsubseteq \psi_i$ in contradiction to our assumption.

*Sufficiency:*
Let the conditions on the right-hand side of Theorem 5.10 be fulfilled. On input $b \in \mathbb{N}$ and gradually growing information about a function $f \in \mathcal{R}$ the strategy $S$ works according to the following instructions:

"Goto Step 0.
Step $i$ ($i \in \mathbb{N}$) : Put out $i$ until either statement $A$ or statement $B$ is proved to be valid.

$A:\ \ d(b,i) \not\sqsubseteq f.$
$B:\ \ \exists j \in \mathbb{N}\ [i \neq j \wedge d(b,i) \sqsubseteq d(b,j) \sqsubseteq f].$

In case of validity of one of the two statements goto Step $i+1$."

Now by the same argumentation as in [Wie78] we observe

$$\mathcal{R}_b \in \mathrm{EX}_\psi(\lambda x.S(b,x))$$

for all $b \in B$. Therefore $B \in \mathrm{suit}(I,\mathrm{EX})$.

*qed Theorem 5.10*

**Theorem 5.11** *Assume $m \in \mathbb{N}$, $I \in \mathcal{I}$, $I \subseteq EX_m$. Let $B \subseteq \mathbb{N}$ fulfill $\mathcal{R}_B \subseteq I$. Then $B \in suit(I, EX_m) \iff \exists d \in \mathcal{R}^2\ \forall b \in B\ \exists \psi \in \mathcal{P}^2$*

*1. $\mathcal{R}_b \subseteq \mathcal{P}_\psi$,*

*2. $\forall i \in \mathbb{N}\ [d(b,i) \sqsubseteq \psi_i]$,*

*3. $\forall i \in \mathbb{N}\ [card\ \{j \neq i \mid d(b,j) \sqsubseteq \psi_i\} \leq m]$,*

*4. $\forall i,j \in \mathbb{N}\ [d(b,i) \sqsubseteq d(b,j) \sqsubseteq \psi_i \ \Rightarrow\ i = j]$.*

**Corollary 5.12** *Assume $B \subseteq \mathbb{N}$ fulfills $\mathcal{R}_B \subseteq EX_0$.
Then $B \in suit(EX_0, EX_0) \iff \exists d \in \mathcal{R}^2\ \forall b \in B\ \exists \psi \in \mathcal{P}^2$*

*1.* $\mathcal{R}_b \subseteq \mathcal{P}_\psi$,

*2.* $\forall i, j \in \mathbb{N}\ [i \neq j \ \Rightarrow\ d(b,j) \not\sqsubseteq \psi_i]$.

Together with the results in Theorem 5.2 and Corollary 5.4 we now have found characterizations for description sets suitable for uniform learning in the sense of the inference criteria $\mathrm{BC}^*$, BC, EX and $\mathrm{EX}_m$ for any $m \in \mathbb{N}$. When trying to apply our argumentation to the case of identification with consistent and/or total intermediate hypotheses, we are confronted with a new problem. In order to achieve sufficiency of the right-hand side statement a condition in the style of

$$\forall i, j \in \mathbb{N}\ [d(b,i) \sqsubseteq d(b,j) \sqsubseteq \psi_i \Rightarrow i = j]$$

seems indispensable. In the proofs of Theorem 5.10 and Theorem 5.11 we construct numberings $\psi$ providing that property without violating the required conditions concerning the convergence of the sequence of hypotheses. Now observe that the inference classes CONS and TOTAL are not characterized by means of specific convergence criteria but by means of specific properties concerning the intermediate hypotheses. Exactly those specific properties are violated as a consequence of constructing the numberings $\psi$ by analogy with the proofs of Theorem 5.10 and Theorem 5.11. Therefore we cannot state any corresponding theorems for the case of CONS- and TOTAL-identification as yet.

## 5.2 Unions of Suitable Description Sets

Assume we are given two inference criteria in $\mathcal{I}$ and two description sets suitable for uniform learning with respect to the given criteria. We are often interested in the question, whether the union of the two description sets is still suitable for uniform learning with respect to the given criteria or not. This subsection does not provide a general answer to that question, but is concerned with its solution for just a few special cases. The first result is a direct consequence of Theorem 5.3 and Theorem 5.2.

**Proposition 5.13**

*1.* *$suit(I, BC)$ is closed under union for any $I \in \mathcal{I} \backslash \{BC^*\}$.*

*2.* *$suit(I, BC^*)$ is closed under union for any $I \in \mathcal{I}$.*

The following result concerns some special description sets suitable for uniform identification in the limit.

**Theorem 5.14** *Assume $B, C \in suit(I, EX)$ for an inference criterion $I \in \mathcal{I}$, $I \subseteq EX$. If $B$ is recursively enumerable or $B \backslash C$ is recursively enumerable, then $B \cup C \in suit(I, EX)$.*

**Proof:** Choose $I \in \mathcal{I},\ B, C \subseteq \mathbb{N}$ and $S^B, S^C \in \mathcal{P}^2$ such that $I \subseteq \text{EX}$ and

- $\mathcal{R}_B \in \text{uni}_B(I, \text{EX})(S^B)$,
- $\mathcal{R}_C \in \text{uni}_C(I, \text{EX})(S^C)$.

Let $\pi \in \mathcal{P}$ be a function the domain of which is

- $B$, for the case "$B$ r.e.";
- $B \backslash C$, for the case "$B \backslash C$ r.e.".

For any $a, n \in \mathbb{N},\ f \in \mathcal{R}$ define

$$S(a, f^n) := \begin{cases} S^B(a, f^n) & \pi(a)\downarrow_{\leq n} \text{ or the computation of } \pi(a) \\ & \text{stops before the computation of } S^C(a, f^n) \\ S^C(a, f^n) & \text{otherwise} \end{cases}$$

Now the proof of $\mathcal{R}_{B \cup C} \in \text{uni}_{B \cup C}(I, \text{EX})(S)$ is straightforward.

*qed Theorem 5.14*

Next we consider uniform learning with a bounded number of mind changes. Here we again use a "trick" to simplify learning from unions of suitable description sets.

**Theorem 5.15** *Let $k, l, m, n$ be any integers such that $m \geq k$. Assume $B \in suit(EX_k, EX_l)$ as well as $C \in suit(EX_m, EX_n)$. Then $B \cup C \in suit(EX_m, EX_{l+n+1})$.*

**Proof:** From the given conditions we deduce the existence of two *total recursive* strategies $S^B, S^C \in \mathcal{R}$ with the following properties:

1. $\mathcal{R}_B \in \text{uni}_B(\text{EX}_k, \text{EX}_l)(S^B)$ as well as $\mathcal{R}_C \in \text{uni}_C(\text{EX}_m, \text{EX}_n)(S^C)$.
2. $\forall f \in \mathcal{R}\quad [\text{ card } \{x \mid ? \neq S^B(f^x) \neq S^B(f^{x+1})\} \leq l\ ]$ and $[\text{ card } \{x \mid ? \neq S^C(f^x) \neq S^C(f^{x+1})\} \leq n\ ]$.

The first property allows some further conclusions:

1. $\forall b \in B\ \exists \psi^{(b)} \in \mathcal{P}^2\ [\mathcal{R}_b \in (\text{EX}_l)_{\psi^{(b)}}(\lambda x.S^B(b, x))]$.
2. $\forall c \in C\ \exists \psi^{(c)} \in \mathcal{P}^2\ [\mathcal{R}_c \in (\text{EX}_n)_{\psi^{(c)}}(\lambda x.S^C(c, x))]$.
3. $\forall a \in B \cup C\ [\mathcal{R}_a \in \text{EX}_m]$ (because $m \geq k$).

If we now set

$$\eta^{(b)}_{\langle i,j\rangle} := \psi^{(b)}_i \text{ for } b \in B,\ i,j \in \mathbb{N};$$

$$\eta^{(c)}_{\langle i,j\rangle} := \psi^{(c)}_j \text{ for } c \in C\backslash B,\ i,j \in \mathbb{N},$$

we can easily define a uniform strategy for $B \cup C$ in the sense of uni$(\mathrm{EX}_m, \mathrm{EX}_{l+n+1})$. For $a, x \in \mathbb{N}$ and $f \in \mathcal{R}$ set

$$S(a, f^x) := \begin{cases} ? & S^B(a, f^x) = S^C(a, f^x) =? \\ \langle i, 0\rangle & S^B(a, f^x) = i \in \mathbb{N} \wedge S^C(a, f^x) =? \\ \langle 0, j\rangle & S^B(a, f^x) =? \wedge S^C(a, f^x) = j \in \mathbb{N} \\ \langle i, j\rangle & S^B(a, f^x) = i \in \mathbb{N} \wedge S^C(a, f^x) = j \in \mathbb{N} \end{cases}$$

Since $S_B$ changes its hypotheses at most $l$ times on any input sequence and $S_C$ changes its hypotheses at most $n$ times on any input sequence, we observe for any $a \in B \cup C$

$$\mathcal{R}_a \in (\mathrm{EX}_{l+n+1})_{\eta^{(a)}}(\lambda x.S(a, x))$$

and therefore $B \cup C \in \mathrm{suit}(\mathrm{EX}_m, \mathrm{EX}_{l+n+1})$.

*qed Theorem 5.15*

# 6 Uniform Learning with Respect to Special Hypothesis Spaces

## 6.1 Bounds of Uniform Behaviourally Correct Identification

As we have seen in the previous section, the trick of encoding useful information within the particular hypothesis spaces supplies a simple and rather "stupid" strategy uniformly identifying all BC-identifiable recursive cores from their corresponding descriptions (see Theorem 5.3). Such tricks are possible, because the definition of uniform learnability does not demand identification with respect to specific numberings, i.e. the free choice of hypothesis spaces allows special coding tricks. What influence does the choice of hypothesis spaces have on the uniform learnability? Does the learning power decrease if we avoid the encoding of any precious information within the hypothesis spaces, e.g. by demanding learnability with respect to a given acceptable numbering? The following result answers that question for the case of BC-identification. From Jantke's work [Ja79] we already know that the set of descriptions of recursive cores consisting of just a single function is not suitable for uniform learning with respect to the EX-criterion, 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 behaviourally correct identification is not strong enough.

**Theorem 6.1** *Assume* $B := \{b \in \mathbb{N} \mid card\ \{i \in \mathbb{N} | \varphi_i^b \in \mathcal{R}\} = 1\}$ *and let* $\tau \in \mathcal{P}^2$ *be an acceptable numbering. Then* $\mathcal{R}_B \notin uni_{B,\tau}(BC, BC)$.

**Proof:** Let us assume $\mathcal{R}_B \in \text{uni}_{B,\tau}(\text{BC}, \text{BC})$. This implies the existence of a recursive strategy $S \in \mathcal{R}^2$ satisfying $\mathcal{R}_b \in \text{BC}_\tau(\lambda x.S(b,x))$ for all $b \in B$. In contradiction to this statement we will construct an index $b_0 \in B$, such that our strategy $\lambda x.S(b_0, x)$ trying to identify the only function in $\mathcal{R}_{b_0}$ must produce infinitely many hypotheses incorrect with respect to the hypothesis space $\tau$. For this purpose we define a function $\psi \in \mathcal{P}^3$ by means of the following instructions:

Let $b \in \mathbb{N}$ be given. For all integers $n \in \mathbb{N}$ the function $\lambda x.\psi(b, n, x)$ will be initial or recursive. The initial function segments generated intermediately will be denoted by $\alpha_m$ ($m \in \mathbb{N}$).

**Definition of $\boldsymbol{\lambda x.\psi(b, 1, x)}$ and $\boldsymbol{\alpha_1}$:**

Set $\psi(b, 1, 0) := 0$ and $x := 1$. Then define further values of the function $\lambda x.\psi(b, 1, x)$ in the following way:

$M_1$: Compute $\tau_{S(b,0)}(1), \ldots, \tau_{S(b,0^x)}(x)$ for $x$ steps each.

$$\psi(b,1,x) := \begin{cases} 0 & \psi(b,1,x-1)\downarrow \text{ and} \\ & \forall y \in \{1,\ldots,x\}\ [\tau_{S(b,0^y)}(y)\downarrow_{\leq x} \Rightarrow\ \tau_{S(b,0^y)}(y) \neq 0] \\ \uparrow & \text{otherwise} \end{cases}$$

Now if $\psi(b,1,x-1) = 0$ and $\psi(b,1,x)\uparrow$
(this question can be answered effectively), then set
$k_0 := \min\{\ y \in \{1,\ldots,x\} \mid \tau_{S(b,0^y)}(y)\downarrow_{\leq x}$ and $\tau_{S(b,0^y)}(y) = 0\ \}$
and leave $\psi(b,1,x')$ undefined for all $x' > x$.
Otherwise increase $x$ by 1 and go to the mark $M_1$ again.

Finally we define

$$\alpha_1 := \begin{cases} 0^{k_0}1 & k_0\downarrow \\ \uparrow & k_0\uparrow \end{cases}$$

Note that $\lambda x.\psi(b,1,x) = \begin{cases} 0^\infty & \alpha_1\uparrow \\ 0^{k_0+m}\uparrow^\infty \ \ \text{(for some } m \in \mathbb{N}) & \alpha_1\downarrow \end{cases}$

**Definition of $\boldsymbol{\lambda x.\psi(b,n+1,x)}$ and $\boldsymbol{\alpha_{n+1}}$ for $\boldsymbol{n \geq 1}$:**
If $\alpha_n\downarrow$ and $x \leq k_0 + \cdots + k_{n-1} + n - 1$ $(= |\alpha_n| - 1)$, we define

$$\psi(b,n+1,x) := \begin{cases} \psi(b,n,x) & x < k_0 + \cdots + k_{n-1} + n - 1 \\ 1 & x = k_0 + \cdots + k_{n-1} + n - 1 \end{cases}$$

For the definition of further values of the function $\lambda x.\psi(b,n+1,x)$ we set $x := k_0 + \cdots + k_{n-1} + n;\ z := 1$ and proceed in the following way:

$M_{n+1}$: Compute $\tau_{S(b,\alpha_n 0)}(|\alpha_n|+1), \ldots, \tau_{S(b,\alpha_n 0^z)}(|\alpha_n|+z)$
for $x$ steps each.

$$\psi(b,n+1,x) := \begin{cases} 0 & \psi(b,n+1,x-1)\downarrow \text{ and} \\ & \forall y \in \{1,\ldots,z\}\ [\tau_{S(b,\alpha_n 0^y)}(|\alpha_n|+y)\downarrow_{\leq x} \Rightarrow \\ & \qquad \tau_{S(b,\alpha_n 0^y)}(|\alpha_n|+y) \neq 0] \\ \uparrow & \text{otherwise} \end{cases}$$

Now if $\psi(b,n+1,x-1)\downarrow$ and $\psi(b,n+1,x)\uparrow$
(these questions can be answered effectively), then set
$k_n := \min\{\ y \in \{1,\ldots,z\} \mid \tau_{S(b,\alpha_n 0^y)}(|\alpha_n|+y)\downarrow_{\leq x}$ and $\tau_{S(b,\alpha_n 0^y)}(|\alpha_n|+y) = 0\ \}$
and leave $\psi(b,n+1,x')$ undefined for all $x' > x$.
Otherwise increase $x$ and $z$ by 1 and go to the mark $M_{n+1}$ again.

The new initial function segment $\alpha_{n+1}$ is then defined by

$$\alpha_{n+1} := \begin{cases} \alpha_n 0^{k_n}1 & k_n\downarrow \\ \uparrow & k_n\uparrow \end{cases}$$

Note $\lambda x.\psi(b, n+1, x) = \begin{cases} \uparrow^\infty & \alpha_n\uparrow \\ \alpha_n 0^\infty & \alpha_n\downarrow \wedge \alpha_{n+1}\uparrow \\ \alpha_n 0^{k_n+m}\uparrow^\infty \text{ (for some } m \in \mathbb{N}) & \alpha_n\downarrow \wedge \alpha_{n+1}\downarrow \end{cases}$

Thus it remains to fix the values of $\psi$ for the case that the second argument is 0.

**Definition of $\lambda x.\psi(b, 0, x)$:**

We agree on

$$\psi(b, 0, x) := \begin{cases} \psi(b, x+1, x) & k_x\downarrow \\ \uparrow & k_x\uparrow \end{cases}$$

for all $x \in \mathbb{N}$. That means, $\lambda x.\psi(b, 0, x)$ is defined by the sequence of the segments $\alpha_n,\ n \geq 1$ of gradually growing length.

Altogether we hence get a computable function $\psi \in \mathcal{P}^3$. From that function we wish to construct an index in $B$ in order to derive a contradiction to our assumption proposed above. From the *s-m-n-* and recursion theorem we deduce the existence of an index $b_0 \in \mathbb{N}$ which provides

$$\varphi_n^{b_0}(x) = \psi(b_0, n, x) \text{ for all } n, x \in \mathbb{N}.$$

Furthermore we want to prove that $b_0$ is an element of $B$ and describes a recursive core $\mathcal{R}_{b_0}$ which can *not* be identified by $\lambda x.S(b_0, x)$ in a behaviourally correct manner with respect to $\tau$. Thus we propose two claims:

**Claim 1** $b_0 \in B$.

**Claim 2** $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$.

For the proof of these two claims let $k_n$ and $\alpha_n$ be the values defined in the construction of $\lambda y.\lambda x.\psi(b_0, y, x)$ for all $n \in \mathbb{N}$ (corresponding to those defined above).

Proof of Claim 1:

1st case: During the construction of $\lambda x.\lambda y.\psi(b_0, x, y)$ the values $k_n$ are defined for all $n \in \mathbb{N}$.
$\Rightarrow\ \varphi_0^{b_0} \in \mathcal{R};\ \forall n \in \mathbb{N}\ [\varphi_{n+1}^{b_0}$ initial$]$.
$\Rightarrow\ \{i \in \mathbb{N} \mid \varphi_i^{b_0} \in \mathcal{R}\} = \{0\}$.
$\Rightarrow\ b_0 \in B$.

2nd case: During the construction of $\lambda x.\lambda y.\psi(b_0, x, y)$ the number $n$ is the minimal index satisfying $k_n\uparrow$.
$\Rightarrow\ \varphi_{n+1}^{b_0} \in \mathcal{R};\ \forall m \in \mathbb{N}\backslash\{n+1\}\ [\varphi_m^{b_0}$ initial$]$.
$\Rightarrow\ \{i \in \mathbb{N} \mid \varphi_i^{b_0} \in \mathcal{R}\} = \{n+1\}$.
$\Rightarrow\ b_0 \in B$.

*qed Claim 1*

Proof of Claim 2:

1st case: $\varphi_0^{b_0} \in \mathcal{R}$.
$\Rightarrow$ $k_n\downarrow$ for all $n \in \mathbb{N}$.
According to the definition of $k_n$ and $\lambda x.\psi(b_0, 0, x)$ we obtain:
$\tau_{S(b_0,\alpha_n 0^{k_n})}(|\alpha_n 0^{k_n}|) = 0 \neq 1 = \psi(b_0, 0, |\alpha_n 0^{k_n}|)$ for all $n \in \mathbb{N}$.
Regarding $|\alpha_{n+1}| > |\alpha_n|$ for all $n \in \mathbb{N}$ we finally conclude:
$\lambda x.S(b_0, x)$ produces infinitely many incorrect hypotheses (wrt $\tau$) for the function $\varphi_0^{b_0}$.
$\Rightarrow$ $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$.

2nd case: There exists an $n \in \mathbb{N}$ with $\varphi_{n+1}^{b_0} \in \mathcal{R}$.
$\Rightarrow$ $k_n\uparrow,\ \alpha_m\downarrow$ for all $m \in \{1, \ldots, n\}$.
The definition of $\psi$ implies $\psi_{(b_0,n+1)} = \alpha_n 0^\infty$ and for any $y \in \mathbb{N}$ :
$\tau_{S(b_0,\alpha_n 0^y)}(|\alpha_n 0^y|) \neq 0$, hence $\tau_{S(b_0,\alpha_n 0^y)} \neq \alpha_n 0^\infty = \varphi_{n+1}^{b_0}$.
$\Rightarrow$ $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$.

*qed Claim 2*

Summarizing these results we achieve the desired statement $\mathcal{R}_B \notin \mathrm{uni}_{B,\tau}(\mathrm{BC}, \mathrm{BC})$.

*qed Theorem 6.1*

Let $B$ be the description set defined in Theorem 6.1. Since we know that $\mathcal{R} \in \mathrm{BC}^*_\tau$ for any acceptable numbering $\tau \in \mathcal{P}$, we conclude that

$$\mathcal{R}_B \in \mathrm{uni}_{B,\tau}(\mathrm{EX}_0, \mathrm{BC}^*) \backslash \mathrm{uni}_{B,\tau}(\mathrm{EX}_0, \mathrm{BC}).$$

Hence the question arises, whether we can find a uniform learning problem *not* solvable by means of BC*-identification. By slight changes of the construction in the proof of Theorem 6.1 we obtain that our set $B$ is not suitable for uniform BC*-identification with respect to the hypothesis spaces $\varphi^b,\ b \in B$ given *a priori*.

**Corollary 6.2** *Let* $B := \{b \in \mathbb{N} \mid \mathit{card}\ \{i \in \mathbb{N} | \varphi_i^b \in \mathcal{R}\} = 1\}$.
*Then* $\mathcal{R}_B \notin \mathit{uni}_{B,[id]}(\mathit{EX}_0, \mathit{BC}^*)$.

**Proof:** Assume $\mathcal{R}_B \in \mathrm{uni}_{B,[id]}(\mathrm{EX}_0, \mathrm{BC}^*)$. Hence there exists a strategy $S \in \mathcal{R}^2$ which provides $\mathcal{R}_b \in \mathrm{BC}^*_{\varphi^b}(\lambda x.S(b, x))$ for all $b \in B$. Furthermore, for all acceptable numberings $\tau \in \mathcal{P}^2$ one may define a strategy $T \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in \mathrm{EX}_\tau(\lambda x.T(b, x))$ for all $b \in B$. For that purpose we simply choose a function $g \in \mathcal{R}$, such that

$$\varphi_i^{g(b)}(j) := \begin{cases} \varphi_i^b(j) & \forall x \leq j\ [\varphi_i^b(x)\downarrow] \\ \uparrow & \text{otherwise} \end{cases}$$

for all $b, i, j \in \mathbb{N}$.
Let $\tau \in \mathcal{P}^2$ be an acceptable numbering and choose $c \in \mathcal{R}^2$ such that $\tau_{c(b,i)} = \varphi_i^{g(b)}$ for all $b, i \in \mathbb{N}$.
Provided $b \in B$ we observe the following properties for the index $g(b)$:

1. $g(b) \in B$.

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

Let $f_b$ denote the only function in $\mathcal{R}_{g(b)},\ b \in B$.
Since $\mathcal{R}_{g(b)} \in \mathrm{BC}^*_{\varphi^{g(b)}}(\lambda x.S(g(b),x))$, 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

$$\mathcal{R}_{g(b)} \in \mathrm{EX}_{\varphi^{g(b)}}(\lambda x.S(g(b),x)).$$

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 get for $b \in B$ :

$$\mathcal{R}_b \in \mathrm{EX}_\tau(\lambda x.T(b,x)).$$

Though Theorem 6.1 tells us that $\mathcal{R}_B \notin \mathrm{uni}_{B,\tau}(\mathrm{EX}_0, \mathrm{BC})$, we now obtain $\mathcal{R}_B \in \mathrm{uni}_{B,\tau}(\mathrm{EX}_0, \mathrm{EX})$. Therefore our first assumption in this proof must have been wrong. Hence we conclude $\mathcal{R}_B \notin \mathrm{uni}_{B,[id]}(\mathrm{EX}_0, \mathrm{BC}^*)$, as has been asserted.

*qed Corollary 6.2*

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. Perhaps if the only recursive function in $\mathcal{P}_{\varphi^b}$ were repeated within the numbering more often, it might be easier for our strategy to find a correct index for the element of the recursive core. But even if we allow infinitely many $\varphi^b$-numbers for the functions to be learned, our situation does not improve. The verification of Corollary 6.3 is done by reduction with the help of Corollary 6.2.

**Corollary 6.3** *Let* $C := \{b \in \mathbb{N} \mid$ *card* $\mathcal{R}_b = 1$ *and card* $\{i \mid \varphi_i^b \in \mathcal{R}\} = \infty\}$.
*Then* $\mathcal{R}_C \notin uni_{C,[id]}(EX_0, BC^*)$.

**Proof:** We will use Corollary 6.2.
For this purpose we assume $\mathcal{R}_C \in \mathrm{uni}_{C,[id]}(\mathrm{EX}_0, \mathrm{BC}^*)(S)$ for some appropriate strategy $S \in \mathcal{R}^2$. Then we construct a uniform BC*-strategy for our class $\mathcal{R}_B$, where $B$ denotes the description set defined in Corollary 6.2.
There is a function $g \in \mathcal{R}$ satisfying

$$\varphi_0^{g(b)} = \varphi_0^b,\ \varphi_1^{g(b)} = \varphi_1^b,$$

$$\varphi_2^{g(b)} = \varphi_0^b,\ \varphi_3^{g(b)} = \varphi_1^b,\ \varphi_4^{g(b)} = \varphi_2^b,$$

$$\varphi_5^{g(b)} = \varphi_0^b,\ \varphi_6^{g(b)} = \varphi_1^b,\ \varphi_7^{g(b)} = \varphi_2^b,\ \varphi_8^{g(b)} = \varphi_3^b, \ldots$$

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. $\forall b \in \mathbb{N}\ \forall f \in \mathcal{P}_{\varphi^{g(b)}}$ [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 the following connection:

$$b \in B \;\Rightarrow\; g(b) \in C$$

Therefore $T(b,x) := e(S(g(b),x))$ (for $b, x \in \mathbb{N}$) yields a computable strategy satisfying $\mathcal{R}_B \in \text{uni}_{B,[id]}(\text{EX}_0, \text{BC}^*)(T)$. That contradiction to Corollary 6.2 now makes us reject our assumption. This implies $\mathcal{R}_C \notin \text{uni}_{C,[id]}(\text{EX}_0, \text{BC}^*)$.

*qed Corollary 6.3*

Of course, repetitions of our single function for all but finitely many indices makes the description set suitable for uniform BC*- and even BC-identification with respect to the hypothesis spaces given a priori.

**Proposition 6.4** *If* $D := \{b \in \mathbb{N} \mid \textit{card}\ \mathcal{R}_b = 1 \textit{ and } \forall^\infty i\ [\varphi_i^b \in \mathcal{R}]\}$*, then* $\mathcal{R}_D \in uni_{D,[id]}(EX_0, BC)$*.*

**Proof:** An appropriate uniform strategy $S \in \mathcal{R}^2$ may be defined by

$$S(b, f[n]) := n$$

for all $b, n \in \mathbb{N}$ and $f \in \mathcal{R}$.

*qed Proposition 6.4*

## 6.2 Characterizations of Uniform Identification with Respect to Acceptable Numberings

As for the characterizations of suitable description sets in the previous section we are also interested in transferring the characterization results in the style of Theorem 2.1 to our situation. In the case of demanding learnability with respect to an acceptable numbering we can simply copy the proofs given in [Wie78] for the "uniform dimension". We obtain the results listed below. Since all the proofs proceed in a way analogous to those of the corresponding theorems mentioned in Section 2, we only give an outline of the proof of Theorem 6.5 and omit the verification of the other results listed below.

**Theorem 6.5** *Fix $I \in \mathcal{I}$ with $I \subseteq EX$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, EX) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^3\ \forall b \in B$*

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

**Proof:** Fix $I \in \mathcal{I},\ I \subseteq \text{EX},\ B \subseteq \mathbb{N}$, such that $\mathcal{R}_b \in I$ for all $b \in B$, let $\tau \in \mathcal{P}^2$ be acceptable.

*Necessity:* Assume $\mathcal{R}_B \in \text{uni}_{B,\tau}(I, \text{EX})$. From the definitions and Proposition 3.4 we conclude that there exists an $S \in \mathcal{R}^2$ satisfying

$$\mathcal{R}_b \in \text{EX}_\tau(\lambda x.S(b, x)) \text{ for all } b \in B.$$

For any integer $b \in \mathbb{N}$ define

$$M_b := \{(z, n) \mid n > 0\ \wedge\ \forall x \leq n\ [\tau_z(x)\downarrow]\ \wedge\ S(b, \tau_z[n]) = z \neq S(b, \tau_z[n-1])\}.$$

Without loss of generality we can assume that the definition of $S$ yields just infinite sets $M_b,\ b \in \mathbb{N}$. Since the sets $M_b$ can be enumerated uniformly in $b$, there exists a recursive function $e \in \mathcal{R}^2$ such that for any $b \in \mathbb{N}$:

- $\{e(b, i) \mid i \in \mathbb{N}\} = M_b$,
- $\forall i, j \in \mathbb{N}\ [e(b, i) = e(b, j) \iff i = j]$.

Now fix $b, i, j \in \mathbb{N}$. If $e(b, i) = (z, n)$ and $e(b, j) = (w, m)$, then set

$d(b, i, j) = \max\{n, m\}$ and

$$\psi(b, i, x) := \begin{cases} \tau_z(x) & x \leq n \text{ or} \\ & x > n \text{ and } \forall y \in \{x, \ldots, n\}\ [\tau_z(y)\downarrow \wedge S(b, \tau_z[y]) = z] \\ \uparrow & \text{otherwise} \end{cases}$$

for all $x \in \mathbb{N}$.

The properties

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

can be verified for all $b \in B$ as in the proof of Theorem 2.1 in [Wie78].

*Sufficiency:* Fix $\psi \in \mathcal{P}^3,\ d \in \mathcal{R}^3$ with our two characteristic properties. Let $c \in \mathcal{R}^2$ be a compiler satisfying

$$\tau_{c(b,i)} = \psi_i^b \text{ for all } b, i \in \mathbb{N}.$$

On input sequence $(b, f[n])_{n\in\mathbb{N}}$ the program of a successful uniform strategy can be explained in the following way:

$S$ begins in step 0.

In step $i$ ($i \in \mathbb{N}$) the output of $S$ is $c(b,i)$. Then $S$ looks for an integer $j \neq i$ satisfying $\psi^b_j =_{d(b,i,j)} f$. As soon as such an integer $j$ is found, $S$ goes to step $i+1$.

Further details of this proof are obtained as in [Wie78].

*qed Theorem 6.5*

Below we list the corresponding results for the other inference criteria.

**Theorem 6.6** *Fix $m \in \mathbb{N}$ and $I \in \mathcal{I}$ with $I \subseteq EX_m$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, EX_m) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^2\ \forall b \in B$*

1. $\mathcal{R}_b \subseteq \mathcal{P}_{\psi^b}$.
2. $\forall j \in \mathbb{N}\ [card\ \{i \neq j \mid \psi^b_i =_{d(b,i)} \psi^b_j\} \leq m]$.
3. $\forall i, j \in \mathbb{N}\ [i \neq j \Rightarrow \psi^b_i \neq_{\max\{d(b,i),d(b,j)\}} \psi^b_j]$.

**Corollary 6.7** *Let $B \subseteq \mathbb{N}$ fulfill $\mathcal{R}_b \in EX_0$ for all $b \in B$. Furthermore, let $\tau \in \mathcal{P}^2$ be an acceptable numbering.*
*Then $\mathcal{R}_B \in uni_{B,\tau}(EX_0, EX_0) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^2\ \forall b \in B$*

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

**Theorem 6.8** *Fix $I \in \mathcal{I}$ with $I \subseteq CONS$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, CONS) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^4_{0,1}\ \forall b \in B$*

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

**Theorem 6.9** *Fix $I \in \mathcal{I}$ with $I \subseteq TOTAL$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, TOTAL) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^2\ \forall b \in B$:*

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

Of course, uniform learning with total intermediate hypotheses with respect to an acceptable numbering $\tau$ is equivalent to uniform learning with total *and* consistent intermediate hypotheses with respect to $\tau$. The proof is analogous to the proof of Proposition 2.6.

**Proposition 6.10** *Fix $I \in \mathcal{I}$ with $I \subseteq TOTAL$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, CONS\text{-}TOTAL) \iff \mathcal{R}_B \in uni_{B,\tau}(I, TOTAL)$.*

**Theorem 6.11** *Fix $I \in \mathcal{I}$ with $I \subseteq BC$. 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 $\mathcal{R}_B \in uni_{B,\tau}(I, BC) \iff \exists \psi \in \mathcal{P}^3\ \exists d \in \mathcal{R}^2\ \forall b \in B$:*

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 $\mathcal{R} \in \mathrm{BC}_\tau^*$ for any acceptable numbering $\tau$, we can use the same reasoning as in Proposition 5.1 to prove our characterization in Theorem 6.12.

**Theorem 6.12** *Fix $I \in \mathcal{I}$ and $B \subseteq \mathbb{N}$. Furthermore, let $\tau \in \mathcal{P}^2$ be an acceptable numbering. Then $\mathcal{R}_B \in uni_{B,\tau}(I, BC^*) \iff [\mathcal{R}_b \in I$ for all $b \in B]$.*

## 6.3 Some Topological Structures Enabling Uniform Identification

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. Thus to convince the reader, that there is still 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 6.1 (Discrete Sets of Functions)** *A subset $D \subseteq \mathcal{P}$ will be called discrete if and only if for any $f \in D$ there exists a number $n \in \mathbb{N}$, such that all functions $g \in D$ fulfill the condition*

$$f \neq g \;\Rightarrow\; f \neq_n g.$$

*The integer $n$ will then be called discreteness point for $f$ with respect to $D$.*

**Theorem 6.13** *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 $\mathcal{R}_B \in uni_{B,\tau}(CONS\text{-}TOTAL, CONS\text{-}TOTAL)$.*

**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 *extend* $\in \mathcal{P}$ which helps us to find suitable extensions of initial function segments. For all $b, n, x, k \in \mathbb{N}$ and $f \in \mathcal{R}$ define

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

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

**Definition of $\psi \in \mathcal{P}^3$ with $\mathcal{R}_b \in$ CONS-TOTAL$_{\psi^b}$ for all $b \in B$:**
Let $b, n, x \in \mathbb{N}$, $f \in \mathcal{R}$. We define

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

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

**Claim 1** If $f \in \mathcal{R}_b$ is any element of the $\mathcal{R}$-core $\mathcal{R}_b$ and $n \in \mathbb{N}$ is any integer, then $\psi^b_{f[n]} \in \mathcal{R}$.

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

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

*Proof of Claim 1:*
Fix $n \in \mathbb{N}$. Since there is an integer $i \in \mathbb{N}$ such that $\varphi_i^b = f$, we know that for all

$x \in \mathbb{N}$ there exists a suitable "extension" of $f[n]$, i.e.

$$\forall x \in \mathbb{N}\ \exists k \in \mathbb{N}\ [extend(b, f[n], x, k) = 1].$$

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 $\psi^b_{f[n]}$ is total and thus recursive.

*qed Claim 1*

*Proof of Claim 2:*
For all arguments less than or equal to $n_f$ the values of $\psi^b_{f[n]}$ 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 that 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 $\varphi^b_i$. Hence $\psi^b_{f[n_f]} = \varphi^b_i$.

*qed Claim 2*

*Proof of Claim 3:*
Let $f \in \mathcal{R}_b$ and let $n_f$ be a discreteness point of $f$ with respect to $\mathcal{P}_{\varphi^b}$. The discreteness of $\mathcal{P}_{\varphi^b}$ guarantees the existence of $n_f$. From Claim 2 we now conclude $\psi^b_{f[n_f]} = f$, hence $f \in \mathcal{P}_{\psi^b}$.

*qed Claim 3*

Now let $c \in \mathcal{R}^2$ be a computable function satisfying

$$\tau_{c(b,y)} = \psi^b_y \text{ for all } b, y \in \mathbb{N}.$$

**Definition of a strategy $S \in \mathcal{P}^2$ with $\mathcal{R}_b \in$ CONS-TOTAL$_\tau(\lambda x.S(b,x))$ for all $b \in B$:**
Let $f \in \mathcal{R},\ b, n \in \mathbb{N}$. We define a uniform strategy in the following way:

$$S(b, f[0]) := c(b, f[0])$$

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

It remains to prove that for all $b \in B$ and all initial segments of functions in $\mathcal{R}_b$ our strategy puts out consistent indices of total functions. Furthermore we will show that the sequence of hypotheses put out for any function in $\mathcal{R}_b$ must converge. Together with the consistency of all hypotheses we thus obtain convergence to a *correct* index. Formally, we have to prove:

(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 $n$.
At 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 $\tau_{S(b,f[0])} = \tau_{c(b,f[0])} = \psi^b_{f[0]} =_0 f$. This proves the consistency of the hypothesis $S(b, f[0])$. Furthermore, since $f \in \mathcal{R}_b$, we observe from Claim 1 with $n = 0$, that $\tau_{S(b,f[0])} = \tau_{c(b,f[0])} = \psi^b_{f[0]} \in \mathcal{R}$.
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. On the other hand, if the second case occurs, our previous hypothesis must have been wrong. We obtain $\tau_{S(b,f[n+1])} = \tau_{c(b,f[n+1])} = \psi^b_{f[n+1]} =_{n+1} f$. Therefore $S(b, f[n+1])$ is consistent for $f[n+1]$. Claim 1 now confirms the statement $\tau_{S(b,f[n+1])} \in \mathcal{R}$. Anyway the hypothesis produced by $S$ fulfills the conditions (i), (ii) and (iii).

*qed (i),(ii),(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 can deduce a contradiction in the following way:
Since $\mathcal{P}_{\varphi^b}$ is discrete, there exists an $n_f \in \mathbb{N}$ satisfying

$$\varphi^b_i =_{n_f} f \iff \varphi^b_i = f$$

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 $f = \psi^b_{f[n]} = \tau_{c(b,f[n])} = \tau_{S(b,f[n])}$. 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).

*qed (iv)*

From the 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 $\mathcal{R}_b \in \text{CONS-TOTAL}_{\psi^b}(\lambda x.S(b, x))$ for all $b \in B$. This completes the proof of Theorem 6.13.

*qed Theorem 6.13*

By comparison of Theorem 6.13 with some of our further negative results we conclude that the way the recursive cores are described have much more influence upon their uniform learnability than the $\mathcal{R}$-cores themselves. We know from Theorem 4.1 that there are even suitable sets describing the entirety of all classes of recursive functions learnable with respect to a given criterion. That result remains valid for uniform learning with respect to any fixed acceptable numbering. With Theorem 6.13 we also have a positive result for uniform learnability with respect to a given numbering. If we set

$$B_{discrete} := \{b \in \mathbb{N} \mid \mathcal{P}_{\varphi^b} \text{ is discrete}\},$$

we know that $B_{discrete}$ is suitable for uniform learning with respect to the criterion CONS-TOTAL and 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 is that the numberings described correspond to a non-discrete set of partial-recursive functions. Not the recursive functions to be identified but the *non-recursive* functions in $\mathcal{P}_{\varphi^b}$ described by $b \in \mathbb{N}$ trouble our strategy.

An example for a set describing only discrete sets $\mathcal{P}_{\varphi^b}$ is the set of all indices of "$\text{EX}_0$-characteristic" numberings like those in Corollary 2.3. Hence, with the definition

$$B_{EX_0} := \{b \in \mathbb{N} \mid \varphi^b \text{ is an EX}_0\text{-characteristic numbering}\}$$

we obtain

$$\mathcal{R}_{B_{EX_0}} \in \text{uni}_{B_{EX_0},\tau}(\text{CONS-TOTAL}, \text{CONS-TOTAL})$$

for any acceptable numbering $\tau \in \mathcal{P}^2$. We observe that $B_{EX_0}$ does not only describe discrete sets $\mathcal{P}_{\varphi^b}$, but also functions $\varphi^b$ enumerating those sets *without repetitions*. As we will see, such description sets are even 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.

**Definition 6.2 (Absolutely Discrete Numberings)** *A function* $\psi \in \mathcal{P}^2$ *is called an absolutely discrete numbering :* $\Longleftrightarrow$

1. $\mathcal{P}_\psi$ *is discrete,*
2. $\forall i, j \in \mathbb{N}\ [i \neq j \ \Rightarrow\ \psi_i \neq \psi_j]$.

**Theorem 6.14** *Fix* $B \subseteq \mathbb{N}$. *Assume that* $\varphi^b$ *is an absolutely discrete numbering for all* $b \in B$. *Then* $\mathcal{R}_B \in uni_{B,[id]}(\textit{CONS-TOTAL}, \textit{CONS})$.

**Proof:** From the previous theorem we know that $\mathcal{R}_b \in$ CONS-TOTAL for all $b \in B$. Thus it remains to prove the existence of a strategy $S \in \mathcal{P}^2$ satisfying

$$\mathcal{R}_b \in \text{CONS}_{\varphi^b}(\lambda x.S(b,x))$$

for all $b \in B$. For arbitrary $f \in \mathcal{R},\ n, b \in \mathbb{N}$ we simply define $S(b, f[n])$ by the following instructions:

"Look for a number $i \in \mathbb{N}$ with $\varphi_i^b =_n f$;
Return $i$;"

Since $\mathcal{R}_b \subseteq \mathcal{P}_{\varphi^b}$, we conclude

- $\forall f \in \mathcal{R}_b\ \forall n \in \mathbb{N}\ [S(b, f[n])\downarrow]$,
- $\forall f \in \mathcal{R}_b\ \forall n \in \mathbb{N}\ [\varphi^b_{S(b,f[n])} =_n f]$

for all $b \in B$. Thus – for arbitrary $b \in B$ – we obtain $\mathcal{R}_b \in \text{CONS}_{\varphi^b}(\lambda x.S(b,x))$, because $\varphi^b$ is an absolutely discrete numbering.

*qed Theorem 6.14*

# 7 Uniform Learning with Special Strategies

A very natural way to learn a set of recursive functions is "Identification by Enumeration" introduced by Gold in [Go67]. The idea is to search for the first element in the hypothesis space consistent with the information received so far. Of course such a strategy will not work in any arbitrarily chosen hypothesis space. If our hypothesis space $\psi$ is a recursive function itself, then all functions in $\mathcal{P}_\psi$ can be identified in the limit by means of Identification by Enumeration.

**Definition 7.1 (Identification by Enumeration)** *A class $U \subseteq \mathcal{R}$ is said to be identifiable by enumeration wrt $\psi \in \mathcal{P}^2$, if and only if*

$$U \in CONS_\psi(Enum_\psi),$$

*where the strategy $Enum_\psi$ is defined by*

$$Enum_\psi(f[n]) := \begin{cases} \min X & X := \{i \in \mathbb{N} \mid \psi_i =_n f\} \neq \emptyset \ and \\ & \forall j < \min X \ [\psi_j(0)\downarrow, \dots, \psi_j(n)\downarrow] \\ \uparrow & otherwise \end{cases}$$

*for all $f \in \mathcal{R},\ n \in \mathbb{N}$.*

**Proposition 7.1** *Fix $\psi \in \mathcal{R}^2$. Then any subclass of the set $\mathcal{P}_\psi$ is identifiable by enumeration with respect to $\psi$. Additionally, the $Enum_\psi$-strategy identifies $\mathcal{P}_\psi$ with total intermediate hypotheses.*

Now we want to investigate the use of Enum-strategies for the purpose of uniform learning. Since the definition of Identification by Enumeration depends on a fixed numbering $\psi \in \mathcal{P}^2$, we first have to define a kind of "uniform Enum-strategy".

**Definition 7.2 (Uniform Enum-Strategy)** *Define a uniform strategy Enum by*

$$Enum(b, f[n]) := Enum_{\varphi^b}(f[n])$$

*for all $f \in \mathcal{R},\ b, n \in \mathbb{N}$.*

It is easy to see that Proposition 7.1 remains valid in a "uniform version", as has also been proved in [OSW88]:

**Theorem 7.2** *Let $B \subseteq \mathbb{N}$. Assume $\varphi^b \in \mathcal{R}^2$ for all $b \in B$. Then*

$$\mathcal{R}_B \in uni_{B,[id]}(CONS\text{-}TOTAL, CONS\text{-}TOTAL)(Enum).$$

Though Identification by Enumeration is based on a very natural learning behaviour, the possibility to use it is most often restricted to numberings $\psi$ providing computability of the predicate $d$ defined by

$$d(i,x,y) := \begin{cases} 1 & \psi_i(x) = y \\ 0 & \psi_i(x) \neq y \end{cases} \quad \text{for } i,x,y \in \mathbb{N}.$$

Therefore we consider a variation of our Enum-strategy. Instead of searching the minimal consistent hypothesis, a "temporarily conform" strategy looks for the minimal index of a function with values not contradicting the information received so far within a certain amount of computing time. For further information see [FKW95].

**Definition 7.3 (Temporarily Conform Identification)** *A class $U \subseteq \mathcal{R}$ is called temporarily conformly identifiable wrt $\psi \in \mathcal{P}^2$, if and only if*

$$U \in EX_\psi(TC_\psi),$$

*where the strategy $TC_\psi$ is defined by*

$$TC_\psi(f[n]) := \begin{cases} \min X & X := \{i \in \mathbb{N} \mid \forall x \leq n \; [\varphi_i(x)\downarrow_{\leq n} \Rightarrow \varphi_i(x) = f(x)]\} \neq \emptyset \\ \uparrow & \textit{otherwise} \end{cases}$$

*for all $f \in \mathcal{R}$, $n \in \mathbb{N}$.*

In [FKW95] Freivalds, Kinber and Wiehagen have proved the following quite easy characterization of temporarily conform identifiability.

**Theorem 7.3** *Fix $\psi \in \mathcal{R}^2$ and $U \subseteq \mathcal{R}$. $U$ is temporarily conformly identifiable with respect to $\psi \iff$*

1. *$U \subseteq \mathcal{P}_\psi$,*

2. *$\forall f \in U \; \forall i < \min_\psi f \; [\psi_i \not\subset f]$.*

By analogy with the definition of a uniform Enum-strategy we also obtain a uniform TC-strategy.

**Definition 7.4 (Uniform TC-Strategy)** *Define a uniform strategy $TC$ by*

$$TC(b, f[n]) := TC_{\varphi^b}(f[n])$$

*for all $f \in \mathcal{R}$, $b,n \in \mathbb{N}$.*

Now the result in [FKW95] can easily be transferred to the case of uniform learning.

**Theorem 7.4** *Let* $B \subseteq \mathbb{N}$*.* $\mathcal{R}_B \in uni_{B,[id]}(EX, EX)(TC) \iff$
$\forall b \in B\ \forall f \in \mathcal{R}_b\ \forall i < \min_{\varphi^b} f\ [\varphi_i^b \not\subseteq f]$.

The results on uniform learning with special strategies presented here are rather trivial. Still they might raise the question, whether there are more collections of learning problems uniformly solvable by a strategy similar to our uniform Enum- and TC-strategies.

# References


[An80] Angluin, D. (1980); *Inductive Inference of Formal Languages from Positive Data*, Information and Control 45, 117-135.

[Ba74a] Barzdin, J. (1974); *Two Theorems on the Limiting Synthesis of Functions*, Theory of Algorithms and Programs, Latvian State University, Riga 210, 82-88 (in Russian).

[Ba74b] Barzdin, J. (1974); *Inductive Inference of Automata, Functions and Programs*, In: Proceedings International Congress of Math., Vancouver, 455-460.

[BCJ96] Baliga, G.; Case, J.; Jain, S. (1996); *Synthesizing Enumeration Techniques for Language Learning*, In: Proceedings of the Ninth Annual Conference on Computational Learning Theory, ACM Press, 169-180.

[Bl67] Blum, M. (1967); *A Machine-Independent Theory of the Complexity of Recursive Functions*, Journal of the ACM 14 (2), 322-336.

[CS83] Case, J.; Smith, C. (1983); *Comparison of Identification Criteria for Machine Inductive Inference*, Theoretical Computer Science 25, 193-220.

[FKW95] Freivalds, R.; Kinber, E.B.; Wiehagen, R. (1995); *How Inductive Inference Strategies Discover Their Errors*, Information and Computation 118, 208-226.

[Go67] Gold, E.M. (1967); *Language Identification in the Limit*, Information and Control 10, 447-474.

[Ja79] Jantke, K.P. (1979); *Natural Properties of Strategies Identifying Recursive Functions*, Elektronische Informationsverarbeitung und Kybernetik 15, 487-496.

[JB81] Jantke, K.P.; Beick, H. (1981); *Combining Postulates of Naturalness in Inductive Inference*, Elektronische Informationsverarbeitung und Kybernetik EIK 17, 465-484.

[KB92] Kapur, S.; Bilardi, G. (1992); *On uniform learnability of language families*, Information Processing Letters 44, 35-38.

[OSW88] Osherson, D.N.; Stob, M.; Weinstein, S. (1988); *Synthesizing Inductive Expertise*, Information and Computation 77, 138-161.

[Oy98] Oymanns, L. (1998); *Lernen von rekursiven Funktionen mit guten Beispielen*, Diploma Thesis, University of Kaiserslautern (in German).

[Ro87] Rogers, H. (1987); *Theory of Recursive Functions and Effective Computability*, MIT Press, Cambridge, Massachusetts.

[Wie78] Wiehagen, R. (1978); *Zur Theorie der algorithmischen Erkennung*, Dissertation B, Humboldt-University, Berlin (in German).

[Zi99] Zilles, S. (1999); *Induktive Inferenz mit beschränkter Zahl von Hypothesenwechseln*, Diploma Thesis, University of Kaiserslautern (in German).