# Identification Criteria in Uniform Inductive Inference

Sandra Zilles
Fachbereich Informatik
Universität Kaiserslautern
Postfach 3049
D - 67653 Kaiserslautern

zilles@informatik.uni-kl.de

**Abstract**

Uniform Inductive Inference is concerned with the existence and the learning behaviour of strategies identifying infinitely many classes of recursive functions. The success of such strategies depends on the hypothesis spaces they use, as well as on the chosen identification criteria resulting from additional demands in the basic learning model. These identification criteria correspond to different hierarchies of learning power – depending on the choice of hypothesis spaces. In most cases finite classes of recursive functions are sufficient to expose an increase in the learning power given by the uniform learning models corresponding to a pair of identification criteria.

# Contents

# 1 Introduction

The scope of Inductive Inference is concerned with theoretical models simulating learning processes. Such models do not only include a learner and a set of objects to be identified, but also a "hypothesis space" which allows us to associate an object with the output of the learner, i.e. the answers (hypotheses) of the learner are interpreted as "names" of certain objects. A model of quite simple mathematical description is for example identification of classes of recursive functions. This concept in general includes three main components:

- a partial-recursive function $S$ – also called strategy – simulating the learner,
- a class $U$ of total recursive functions which have to be identified by $S$,
- a partial-recursive numbering $\psi$ – called hypothesis space – which enumerates at least all functions in $U$.

In each step of the identification process $S$ is presented a finite subgraph of some unknown arbitrary function $f$ contained in $U$; the strategy $S$ returns a hypothesis which is interpreted as an index of a function in the given numbering $\psi$. It is the learner's job to eventually return a single correct hypothesis, i.e. the sequence of outputs ought to converge to a $\psi$-number of $f$. This model – called identification in the limit – has first been analysed by Gold in [Go67] and gave rise to the investigation and comparison of several new learning models ("inference criteria") basing on that principle. The common idea was to restrict the definition of identifiability by means of additional – and in some way natural – demands concerning the properties of the hypotheses. The corresponding models have been compared with respect to the resulting identification power; for some more background the reader is referred to [Ba74a], [CS83], [FKW95], [Fu88], [JB81] and [Wie78]. Some definitions and results in this context will be summarized in Section 2.

This paper studies Inductive Inference models on a meta-level. Considering collections of infinitely many classes $U$ of recursive functions we are looking for meta-learners synthesizing an appropriate strategy for each class $U$ to be learned. For that purpose we agree on a method to describe a class $U$, because for the synthesis of a learner our meta-strategy should be given some description of $U$. That means we do not only try to solve a learning problem by an expert learner but to design a higher-level learner which constructs a method for solving a learning problem from a given description. Thus the meta-learner is able to simulate all the expert learners.

Uniform identification of classes of total recursive functions has already been studied by Jantke in [Ja79]. Unfortunately, his results are rather negative; he proves that there is no strategy which – given any description of an arbitrary class $U$ consisting of just a single recursive function – synthesizes a learner which identifies $U$ with respect to a fixed hypothesis space. Even if we allow different hypothesis spaces for the different classes of recursive functions, no meta-learner is successful for all descriptions of finite classes (cf. [Zi00]). Since in the non-uniform case finite classes can be identified easily with respect to any common inference criterion, these results might suggest that the model of uniform learning yields a concept the investigation of which is not worthwile. As we will see, the results in this paper allow a more optimistic point of view. Of course it is quite natural to consider the same inference criteria known from the non-uniform model also in our meta-level. The aim of this paper is to investigate whether the comparison of these criteria concerning the resulting identification power yields hierarchies analogous to those approved in the classical context. In most cases we will see, that the classical separation results can be transferred to uniform learning. And we can prove even more. If we consider uniform learning with respect to fixed hypothesis spaces, all separations of inference criteria can be achieved by collections of *finite* classes of recursive functions; interestingly some results seem to indicate that even classes consisting of just one function are often sufficient for these separations in the context of uniform learning with fixed hypothesis spaces. The resulting hierarchies correspond to the non-uniform case. If we drop the restrictions concerning the hypothesis spaces, we obtain slightly different results, although many of the criteria can still be separated by finite classes. Here classes consisting of just one recursive function are not suitable for the corresponding proofs, but many separations are witnessed by classes of no more than two functions. So whereas finite classes are very simple regarding their identifiability in the classical model of Inductive Inference, they are in most cases sufficient for the separation of inference criteria in uniform learning. Furthermore we conclude that the hierarchies obtained are very much influenced by the choice of the hypothesis spaces. Now, since the hierarchies of inference criteria do not collapse in our meta-level – even by restricting ourselves to the choice of very simple identification problems – we conclude that the concept of uniform learning is neither trivial nor fruitless. Furthermore this paper corroborates the interpretation that our different inference criteria possess some really substantial specific properties, which yield separations of such a strong nature that they still hold for uniform learning of finite classes.

In [Zi00] the reader may also find positive results encouraging further research. It is shown that the choice of descriptions for the classes $U$ has more influence on the uniform identifiability than the classes themselves, i.e. many meta-strategies fail because of a bad description of the learning problem rather than because of the complexity of the problem. So it might be interesting to find out what kinds of descriptions are suitable for uniform learnability and whether they can be characterized by any specific properties. This question also arises in the context of separating identification criteria in uniform learning. Perhaps there are certain characteristic features of the description sets witnessing our separation results. Though this paper does not provide a solution to that problem, it gives many examples of appropriate description sets, which may be helpful on the way to an answer.

Further research on uniform identification has also been made in the context of language learning, see for example [KB92], [OSW88] and [BCJ96]. Because of its numerous positive results, in particular the work of Baliga, Case and Jain [BCJ96] motivates the investigation of meta-strategies. In [OSW88] Osherson, Stob and Weinstein especially consider several techniques of describing classes of objects to be learned and thus also prove that the way the learning problems are described influences their uniform identifiability.

My special thanks are due to Jochen Nessel and Martin Memmel for fruitful discussions during my work on this paper, as well as to Prof. Rolf Wiehagen for his advice and motivating support.

# 2 Preliminaries

We will first agree on some notations used in this paper. The second subsection is then concerned with a short introduction into Inductive Inference. Several inference criteria are introduced and compared with respect to their learning power.

## 2.1 Notation

Recursion theoretic terms used here without explicit definition can be found in [Ro87].

By $\mathbb{N}$ we denote the set of all nonnegative integers, $\mathbb{N}^*$ is the set of all finite tuples over $\mathbb{N}$; the variable $n$ always ranges over $\mathbb{N}$. For fixed $n$, the notion $\mathbb{N}^n$ is used for the set of all $n$-tuples of integers. By implicit use of a bijective computable function $\mathrm{cod} : \mathbb{N}^* \mapsto \mathbb{N}$ we will identify any $\alpha \in \mathbb{N}^*$ with its coding $\mathrm{cod}(\alpha) \in \mathbb{N}$. If $\alpha \in \mathbb{N}^*$ is any finite tuple, we use $|\alpha|$ to refer to its length, i.e. $|\alpha| = n$ for each $\alpha \in \mathbb{N}^n,\ n \in \mathbb{N}$. For each $n \in \mathbb{N}$ we define the value $\overline{n}$ by

$$\overline{n} := \begin{cases} 0 & \text{if } n > 0 \\ 1 & \text{if } n = 0 \end{cases}$$

A statement is quantified with $\forall^\infty n$ in order to indicate that the statement is fulfilled for all but finitely many $n$; quantifiers $\forall$ and $\exists$ are used in the common way.

For any set $X$ the expression $\mathrm{card}\, X$ denotes the cardinality of $X$, where $\mathrm{card}\, X = \infty$ indicates that $X$ is an infinite set; $\wp X$ denotes the set of all subsets of $X$. The notion $X^*$ is used by analogy with $\mathbb{N}^*$. $X^+$ denotes the set of all non-empty finite tuples over $X$. As a symbol for set inclusion we use $\subseteq$, proper inclusion is indicated by $\subset$. Incomparability of sets is expressed by $\#$.

The set of all partial-recursive functions is denoted by $\mathcal{P}$, the set of total recursive functions by $\mathcal{R}$. If we want to refer to functions of a fixed number $n$ of input variables, we sometimes add the superscript $n$ to these symbols. $\mathcal{R}_{01}$ denotes the set of all recursive functions, the range of which is contained in $\{0, 1\}$. For any $f \in \mathcal{P}$ and any $x \in \mathbb{N}$ we write $f(x)\downarrow$, if $f$ is defined on input $x$; $f(x)\uparrow$ otherwise. If $f \in \mathcal{P}$ and $n$ are given such that $f(0)\downarrow, \ldots, f(n)\downarrow$ we set $f[n] := \mathrm{cod}(f(0), \ldots, f(n))$, i.e. $f[n]$ corresponds to the initial segment of length $n + 1$ of $f$. We often compare $f, g \in \mathcal{P}$ and write $f =_n g$, if

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

otherwise $f \neq_n g$. If the functions $f$ and $g$ differ only for finitely many arguments, that means 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 $f =^* g$. By the notion $f \subseteq g$ we indicate that

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

and use proper inclusion by analogy. But $f \in \mathcal{P}$ may also be identified with the sequence $(f(n))_{n\in\mathbb{N}}$, so we sometimes write $f = 0^n 1\uparrow^\infty$ for the function defined for $x \in \mathbb{N}$ by

$$f(x) = \begin{cases} 0 & \text{if } x < n \\ 1 & \text{if } x = n \\ \uparrow & \text{if } x > n \end{cases}$$

and the like. We often identify a tuple $\alpha \in \mathbb{N}^*$ with the function $\alpha\uparrow^\infty$ implicitly. Thus we may for example write $\alpha \subseteq f$ for some function $f \in \mathcal{P}$, if $\alpha\uparrow^\infty \subseteq f$, that means if $\alpha = (f(0), \ldots, f(|\alpha| - 1))$; furthermore we may denote the $i$-th component ($i \in \mathbb{N},\ i < |\alpha|$) of $\alpha$ by $\alpha(i)$. So $\alpha = (\alpha(0), \ldots, \alpha(|\alpha|-1))$. By $\mathrm{rng}(f)$ we refer to the range $\{f(x) \mid x \in \mathbb{N},\ f(x)\downarrow\}$ of a function $f \in \mathcal{P}$. Analogously – for every $\alpha \in \mathbb{N}^*$ – we use $\mathrm{rng}(\alpha)$ to refer to $\mathrm{rng}(\alpha\uparrow^\infty) = \{\alpha(0), \ldots, \alpha(|\alpha| - 1)\}$.

A function $\psi \in \mathcal{P}^{n+1}$ is used as a numbering for the set $\mathcal{P}_\psi := \{\psi_i \mid i \in \mathbb{N}\}$, where $\psi_i(x) := \psi(i, x)$ for all $i \in \mathbb{N},\ x \in \mathbb{N}^n$ as usual. $i$ is called $\psi$-number of the function $\psi_i$. In order to refer to the set of all total functions in $\mathcal{P}_\psi$, we use the notion $\mathcal{R}_\psi$, i.e. $\mathcal{R}_\psi := \mathcal{P}_\psi \cap \mathcal{R}$. $\mathcal{R}_\psi$ is called the recursive core or "$\mathcal{R}$-core" of $\mathcal{P}_\psi$. If $\psi \in \mathcal{P}^{n+2}$, every $b \in \mathbb{N}$ corresponds to a numbering $\psi^b \in \mathcal{P}^{n+1}$, if we define $\psi^b(i, x) := \psi(b, i, x)$ for all $i \in \mathbb{N},\ x \in \mathbb{N}^n$. Again $i$ is a $\psi^b$-number for the function $\psi^b_i$ defined in the common way.

Any acceptable numbering $\varphi$ corresponds to a Blum complexity measure $\Phi$, as can be found in [Bl67]. Intuitively, $\Phi_i(x)$ returns the number of steps needed for the computation of $\varphi_i(x)$ whenever $\varphi_i(x)\downarrow$; if $\varphi_i(x)\uparrow$, then also $\Phi_i(x)\uparrow$. If $i, x, n \in \mathbb{N}$, we use the notation $\varphi_i(x)\downarrow_{\leq n}$ instead of $\Phi_i(x) \leq n$ and $\varphi_i(x)\uparrow_{\leq n}$ instead of $[\Phi_i(x)\uparrow$ or $\Phi_i(x) > n]$.

## 2.2 Inductive Inference Criteria

Now we introduce our basic Inductive Inference criterion called identification in the limit, which was first defined in [Go67]. It may be regarded as a fundamental learning model from which we define further restrictive inference criteria.

**Definition 1** *Let $U \subseteq \mathcal{R},\ \psi \in \mathcal{P}^2$. The class $U$ is called identifiable in the limit with respect to the hypothesis space $\psi$ if and only if there is a function $S \in \mathcal{P}$ (called strategy) such that for any $f \in U$ the following conditions are fulfilled:*

1. *$S(f[n])$ is defined for all $n \in \mathbb{N}$ ($S(f[n])$ is called hypothesis on $f[n]$),*
2. *there is some $j \in \mathbb{N}$ such that $\psi_j = f$ and $S(f[n]) = j$ for all but finitely many $n \in \mathbb{N}$.*

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

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

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

On any function $f \in U$ the strategy $S$ must generate a sequence of hypotheses converging to a $\psi$-number of $f$. But a user reading the hypotheses generated by $S$ up to a certain time will never know whether the actual hypothesis is correct or not, because he cannot decide whether the time of convergence is already reached. If there was a bound on the number of mind changes, he could at least rely on the actual hypothesis whenever the bound is reached. Learning with such bounds has first been studied in [CS83].

**Definition 2** *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 with respect to* $\psi$, *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 return the sign "?"),*
2. *for all* $f \in U$ *there is an* $n_f \in \mathbb{N}$ *satisfying*
   - $\forall x < n_f \; [S(f[x]) = ?]$,
   - $\forall x \geq n_f \; [S(f[x]) \in \mathbb{N}]$,
3. $\text{card}\,\{n \in \mathbb{N} \mid\, ? \neq S(f[n]) \neq S(f[n+1])\} \leq m$ *for all* $f \in U$.

*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 with respect to* $\psi\}$.

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

*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 output "?" allows our strategy to indicate that its hypothesis is left open for the actual time being, in order not to waste a mind change in the beginning of the learning process. In [CS83] the reader may find a proof of $\text{EX}_m \subset \text{EX}_{m+1} \subset \text{EX}$ for all $m \in \mathbb{N}$.

Instead of restricting our learning model by bounding the number of mind changes we might also try to mitigate the constraints in the definition of identification in the limit – for example by foregoing the demand for convergence of the sequence of hypotheses. Behaviourally correct identification – as defined in [Ba74a] – allows the learner to switch several correct hypotheses infinitely often.

**Definition 3** *Let $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. U is called behaviourally correctly identifiable (BC-identifiable) with respect to $\psi$ if and only if there exists an $S \in \mathcal{P}$, such that for all $f \in U$ the following conditions are fulfilled:*

1. *$S(f[n])$ is defined for all $n \in \mathbb{N}$,*
2. *$\psi_{S(f[n])} = f$ for all but finitely many $n \in \mathbb{N}$.*

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

Although BC contains some classes of functions not learnable under the EX-criterion (see [Ba74a]), a further increase of learning power can be achieved by allowing "slightly incorrect" hypotheses. BC-identification with anomalies has been studied in [CS83].

**Definition 4** *Let $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. U is called BC-identifiable with respect to $\psi$ 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. *$S(f[n])$ is defined for all $n \in \mathbb{N}$,*
2. *$\psi_{S(f[n])} =^* f$ for all but finitely many $n \in \mathbb{N}$.*

*We also write $U \in BC^*_\psi(S)$ and use the notations $BC^*_\psi$ and $BC^*$ by analogy with the previous definitions.*

In [CS83] Case and Smith verify $\mathrm{BC} \subset \mathrm{BC}^* = \wp\mathcal{R}$, where the proof of $\mathrm{BC}^* = \wp\mathcal{R}$ is based on a private communication to Leo Harrington (1978).

In Definition 2 the criterion EX has been changed by strengthening the demands concerning the convergence of the sequence of hypotheses. It is also a quite natural thought to strengthen the demands concerning the intermediate hypotheses themselves. A successful learning behaviour might be to generate intermediate hypotheses agreeing with the information received up to the actual time of the learning process ("consistent" hypotheses, cf. [Go67] and [Ba74b]).

**Definition 5** *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. *$\psi_{S(f[n])} =_n f$ for all $f \in U$ and $n \in \mathbb{N}$ (we say that $S(f[n])$ is a consistent hypothesis for $f[n]$ with respect to $\psi$).*

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

*$CONS_\psi := \{U \mid U$ is identifiable consistently with respect to $\psi\}$.*

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

A very natural example of learning with consistent intermediate hypotheses is "identification by enumeration" – a method introduced by Gold in [Go67]. The idea is to search within the hypothesis space for the first hypothesis agreeing with the information received so far – that means, the learner looks for the minimal consistent index in the given numbering. In general consistency is not decidable, so this method does not work for arbitrary hypothesis spaces. It is typically used, if the given numbering $\psi$ is recursive itself, because in this case consistency with respect to $\psi$ can be checked and the sequence of hypotheses will converge to the minimal $\psi$-number of the function to be learned (if this function has got any $\psi$-number).

**Definition 6** *Let $\psi \in \mathcal{P}^2$. A class $U \subseteq \mathcal{R}$ of recursive functions is identifiable by enumeration, if and only if $U \in CONS_\psi(Enum_\psi)$, where the partial-recursive function $Enum_\psi$ is defined by*

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

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

Note that $\mathcal{R}_\psi \in \mathrm{CONS}_\psi(\mathrm{Enum}_\psi)$, if $\psi \in \mathcal{R}^2$.

In order to be less demanding than in Definition 5, one could also ask for hypotheses which do not disagree convergently (i.e. in their *defined* values) with the actual information ("conform" hypotheses, see [Wie78] and [Fu88]).

**Definition 7** *Assume $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ is called conformly identifiable with respect to $\psi$ if and only if there exists an $S \in \mathcal{P}$ satisfying*

1. *$U \in EX_\psi(S)$,*
2. *$\forall f \in U \ \forall n \in \mathbb{N} \ \forall x \leq n \ [\psi_{S(f[n])}(x) = f(x) \text{ or } \psi_{S(f[n])}(x)\uparrow]$ (we say that $S(f[n])$ is a conform hypothesis for $f[n]$ with respect to $\psi$).*

*We also write $U \in CONF_\psi(S)$ and use the notions $CONF_\psi$ and $CONF$ by analogy with our former definitions.*

Obviously consistent identification is a special case of conform identification. The proper inclusions CONS $\subset$ CONF $\subset$ EX are verified in [Wie78].

Since any hypothesis representing a function not contained in $\mathcal{R}$ must be wrong, another natural demand would be to allow only $\psi$-numbers of total recursive functions ("total" hypotheses, cf. [JB81]) as outputs of $S$.

**Definition 8** *Assume $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ is called identifiable with respect to $\psi$ with total intermediate hypotheses if and only if there exists an $S \in \mathcal{P}$ satisfying*

*1.* $U \in EX_\psi(S)$,

*2.* $\psi_{S(f[n])} \in \mathcal{R}$ *for all* $f \in U$ *and* $n \in \mathbb{N}$.

*We also write:* $U \in TOTAL_\psi(S)$.
$TOTAL_\psi$ *and* $TOTAL$ *are defined by analogy with Definition 5.*

A tightening of this idea is to forbid all hypotheses corresponding to a function not contained in the class to be learned. Such hypotheses must also be wrong for the relevant target functions, so we might wish to exclude them. The remaining hypotheses are called "class-preserving" hypotheses, because they all correspond to functions in the class to be learned.

**Definition 9** *Assume* $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ *is called identifiable with class-preserving intermediate hypotheses with respect to* $\psi$ *if and only if there exists an* $S \in \mathcal{P}$ *satisfying*

*1.* $U \in EX_\psi(S)$,

*2.* $\psi_{S(f[n])} \in U$ *for all* $f \in U$ *and* $n \in \mathbb{N}$.

*We also write* $U \in CP_\psi(S)$ *and define* $CP_\psi$ *and* $CP$ *as usual.*

For a proof of CP $\subset$ TOTAL $\subset$ CONS see [JB81]. Now let $m \geq 1$ be an arbitrary positive integer. In [Zi99] the reader may find a proof for $\mathrm{EX}_m$ # CONF and $\mathrm{EX}_m$ # CONS. Thus $\mathrm{EX}_m \not\subseteq$ TOTAL and $\mathrm{EX}_m \not\subseteq$ CP. A class in CP $\setminus \mathrm{EX}_m$ can also be found easily: the set of all recursive functions of finite support can be identified by enumeration with class-preserving intermediate hypotheses. This set is not an element of $\mathrm{EX}_m$, as can be verified easily (see for example [CS83]). This implies $\mathrm{EX}_m$ # TOTAL and $\mathrm{EX}_m$ # CP. $\mathrm{EX}_0 \subset$ CP then follows by definition (the output "?" may be replaced by any fixed class-preserving hypothesis).

Since in general the halting problem in $\psi$ is not decidable, it might be hard for our strategy to detect the incorrectness of a hypothesis, if the corresponding function differs from the function to be learned only by being undefined for some arguments. For learning with "convergently incorrect" hypotheses (cf. [FKW95]) such outputs are forbidden.

**Definition 10** *Assume* $U \subseteq \mathcal{R}$, $\psi \in \mathcal{P}^2$. $U$ *is called identifiable with respect to* $\psi$ *with convergently incorrect intermediate hypotheses if and only if there exists an* $S \in \mathcal{P}$ *satisfying*

*1.* $U \in EX_\psi(S)$,

*2.* $\psi_{S(f[n])} \not\subset f$ *for all* $f \in U$ *and* $n \in \mathbb{N}$.

*We also write:* $U \in CEX_\psi(S)$.
$CEX_\psi$ *and* $CEX$ *are defined by analogy with Definition 5.*

Note that for every recursive numbering $\psi \in \mathcal{R}^2$ the class $\mathcal{R}_\psi$ is identifiable according to the inference criteria CP, TOTAL and CEX, for example by the strategy $\mathrm{Enum}_\psi$.

Freivalds, Kinber and Wiehagen have proved $\mathrm{EX}_1 \not\subseteq \mathrm{CEX} \subset \mathrm{EX}$ as well as CEX # CONS in [FKW95]. Since $\mathrm{CP} \not\subseteq \mathrm{EX}_m \not\subseteq \mathrm{CEX}$ (for $m \geq 1$) and $\mathrm{CP} \subseteq \mathrm{CEX}$, we know CEX # $\mathrm{EX}_m$ for all $m \geq 1$. By definition TOTAL is a subset of CEX. That this inclusion is proper, follows from $\mathrm{TOTAL} \subset \mathrm{CONS}$ and CEX # CONS. With similar methods as in [FKW95] we can also verify CEX # CONF.

**Theorem 1** *CEX # CONF.*

*Proof.* As CONS # CEX and CONS is a subclass of CONF, we already know that CONF is not contained in CEX. It remains to prove $\mathrm{CEX} \setminus \mathrm{CONF} \neq \emptyset$. For that purpose we use the class $U$ of recursive functions defined in [FKW95] in the proof of $\mathrm{CEX} \setminus \mathrm{CONS} \neq \emptyset$. Let $\tau \in \mathcal{P}^2$ be a fixed acceptable numbering. $U_0 \subseteq \mathcal{R}$ and $U_1 \subseteq \mathcal{R}$ are defined as follows:

$$\begin{aligned} U_0 &:= \{jp \mid j \in \mathbb{N} \text{ and } p \in \mathcal{R}_{01} \text{ and } \tau_j = jp\} \ , \\ U_1 &:= \{j\alpha 120^\infty \mid j \in \mathbb{N} \text{ and } \alpha \in \{0,1\}^+ \text{ and } \tau_j(|\alpha|+1) = 0 \ (\neq 1)\} \ . \end{aligned}$$

If $j\alpha 120^\infty \in U_1$, then $\tau_j \not\subseteq j\alpha 120^\infty$, so $j$ is a convergently incorrect hypothesis for $j\alpha 120^\infty$ with respect to $\tau$.

Now let $U := U_0 \cup U_1$. Similar ideas as in [FKW95] are used to verify $U \in \mathrm{CEX} \setminus \mathrm{CONF}$.

*Proof of "$U \in$ CEX".* On any function $f$ our strategy first returns the value $f(0)$. This hypothesis is maintained, until the value 2 is found in the following input. Thus all functions in $U_0$ are identified with respect to $\tau$. As soon as the value 2 is found, the strategy returns a $\tau$-number of the function $\alpha 0^\infty$, where $\alpha \in \mathbb{N}^*$ is the initial segment of $f$ read so far (ending with the value 2). This strategy also CEX-identifies the functions $f$ in $U_1$ with respect to $\tau$, because $f(0)$ is convergently incorrect for $f$ in this case. So $U \in \mathrm{CEX}_\tau$.

*Proof of "$U \notin$ CONF".* Suppose by way of contradiction that $U \in$ CONF. Without loss of generality this yields the existence of a strategy $S \in \mathcal{P}$ such that $U \in \mathrm{CONF}_\tau(S)$. We will now deduce a contradiction by proving that there is a function $f \in U$ which is not learned conformly by $S$. For that purpose we define a function $g = \tau_j$ by implicit use of the recursion theorem in the following way:

*Construction of $g = \tau_j$.*

The function $g$ is defined in stages $k$, $k \in \mathbb{N}$. Let $n_0 := 0$, $g(0) := j$ and go to stage 0.

*Stage 0.* We set $g(2) := 0$. The value $g(1)$ is defined as follows:

$$g(1) := \begin{cases} 0 & \text{if } S(j)\downarrow \ \wedge S(j0)\downarrow \ \wedge S(j1)\downarrow \ \wedge S(j0) \neq S(j) \\ 1 & \text{if } S(j)\downarrow \ \wedge S(j0)\downarrow \ \wedge S(j1)\downarrow \ \wedge S(j0) = S(j) \ \wedge \ S(j1) \neq S(j) \\ 1 & \text{if } S(j)\downarrow \ \wedge S(j0)\downarrow \ \wedge S(j1)\downarrow \ \wedge S(j0) = S(j1) = S(j) \\ \uparrow & \text{otherwise} \end{cases}$$

If $g(1)$ is undefined, all further values of $g$ (except $g(2)$) shall also be undefined.

If $g(1)$ is defined and $[S(j0) \neq S(j)$ or $S(j1) \neq S(j)]$, then let $n_1 := 2$ (the maximal argument for which $g$ has been defined up to now) and go to stage 1.

If $g(1)$ is defined and $S(j0) = S(j1) = S(j)$, then test by parallel computation, whether one of the following two properties are fulfilled, and – if yes – which one is fulfilled first:

*(i).* $\tau_{S(j)}(1)$ is defined.
*(ii).* There is an integer $y \in \mathbb{N}$ such that $S(g[2]0^y)$ is defined and $S(g[2]0^y) \neq S(j)$.

With each testing step $y$ define $g(y+2) := 0$.

If property *(i)* is fulfilled first, then let $g(x+2)$ be undefined for all $x$ greater than the actual testing step number $y$.

If property *(ii)* is fulfilled first (for some fixed $y \in \mathbb{N}$), then we have already defined $g(x) = 0$ for all $x \in \{3, \ldots, y+2\}$. Furthermore, let $n_1 := y+2$ (the maximal argument for which $g$ has been defined up to now) and go to stage 1.

If neither property *(i)* nor property *(ii)* is fulfilled, we obtain $g = g[2]0^\infty$.

*End stage 0.*

*Stage $k$ for $k \in \mathbb{N}$.* We set $g(n_k+2) := 0$. The value $g(n_k+1)$ is defined as follows:

$$g(n_k+1) := \begin{cases} 0 & \text{if } S(g[n_k])\downarrow \ \wedge S(g[n_k]0)\downarrow \ \wedge S(g[n_k]1)\downarrow \\ & \quad \wedge S(g[n_k]0) \neq S(g[n_k]) \\ 1 & \text{if } S(g[n_k])\downarrow \ \wedge S(g[n_k]0)\downarrow \ \wedge S(g[n_k]1)\downarrow \\ & \quad \wedge S(g[n_k]0) = S(g[n_k]) \ \wedge \ S(g[n_k]1) \neq S(g[n_k]) \\ 1 & \text{if } S(g[n_k])\downarrow \ \wedge S(g[n_k]0)\downarrow \ \wedge S(g[n_k]1)\downarrow \\ & \quad \wedge S(g[n_k]0) = S(g[n_k]1) = S(g[n_k]) \\ \uparrow & \text{otherwise} \end{cases}$$

If $g(n_k+1)$ is undefined, all further values of $g$ (except $g(n_k+2)$) shall also be undefined.

If $g(n_k+1)$ is defined and $[S(g[n_k]0) \neq S(g[n_k])$ or $S(g[n_k]1) \neq S(g[n_k])]$, then let $n_{k+1} := n_k+2$ (the maximal argument for which $g$ has been defined up to now) and go to stage $k+1$.

If $g(n_k + 1)$ is defined and $S(g[n_k]0) = S(g[n_k]1) = S(g[n_k])$, then test by parallel computation, whether one of the following two properties are fulfilled, and – if yes – which one is fulfilled first:

*(i).* $\tau_{S(g[n_k])}(n_k + 1)$ is defined.

*(ii).* There is an integer $y \in \mathbb{N}$ such that $S(g[n_k + 2]0^y)$ is defined and $S(g[n_k + 2]0^y) \neq S(g[n_k])$.

With each testing step $y$ define $g(n_k + y + 2) := 0$.

If property *(i)* is fulfilled first, then let $g(n_k + 2 + x)$ be undefined for all $x$ greater than the actual testing step number $y$.

If property *(ii)* is fulfilled first (for some fixed $y \in \mathbb{N}$), then we have already defined $g(x) = 0$ for all $x \in \{n_k + 3, \ldots, n_k + y + 2\}$. Furthermore, let $n_{k+1} := n_k + y + 2$ (the maximal argument for which $g$ has been defined up to now) and go to stage $k + 1$.

If neither property *(i)* nor property *(ii)* is fulfilled, we have $g = g[n_k + 2]0^\infty$.

*End stage $k$.*

*End construction $g$.*

Now consider two cases.

*Case (i).* All stages $k$ ($k \in \mathbb{N}$) are reached in the definition of $g$.

Then $g = jp$ for some $p \in \mathcal{R}_{01}$, so $g \in U_0$. But by construction $S$ changes its mind on $g$ infinitely often, so $S$ cannot identify $g$ in the limit. Hence $U \notin \mathrm{CONF}_\tau(S)$; a contradiction to our assumption.

*Case (ii).* Stage $k$ (for some fixed $k \in \mathbb{N}$) is the last stage reached in the construction of $g$.

Then we will either show that $\{g\} \notin \mathrm{CONF}_\tau(S)$, where $g \in U$, or we can prove the existence of some $y \in \{0, 1\}$ such that the function $f \in \mathcal{R}$, defined by

$$f(x) = \begin{cases} g(x) & \text{if } x \leq n_k \\ y & \text{if } x = n_k + 1 \\ 1 & \text{if } x = n_k + 2 \\ 2 & \text{if } x = n_k + 3 \\ 0 & \text{if } x > n_k + 3 \end{cases}$$

cannot be identified conformly with respect to $\tau$ by our strategy $S$. Note that $f \in U_1$ for any $y \in \{0, 1\}$, because then $f = j\alpha 120^\infty$ for some $\alpha \in \{0, 1\}^+$ and $\tau_j(|\alpha| + 1) = g(|\alpha| + 1) = g(n_k + 2) = 0$. For the choice of $y$ we regard two possibilities.

*Case (ii)a.* $S(g[n_k])\uparrow$ or $S(g[n_k]0)\uparrow$ or $S(g[n_k]1)\uparrow$.

As $f[n_k] = g[n_k]$, we obtain $S(f[n_k])\uparrow$ or $S(f[n_k]y')\uparrow$ for some $y' \in \{0,1\}$. In the first case let $y := 0$, in the latter case $y := y'$. Thus our strategy $S$ is undefined on some initial segment of the function $f \in U_1$ as defined above. We obtain $U \notin \mathrm{CONF}_\tau(S)$, which is again a contradiction.

*Case (ii)b.* $S(g[n_k])\downarrow$ and $\tau_{S(g[n_k])}(n_k+1)\downarrow$ and $S(g[n_k]) = S(g[n_k]0) = S(g[n_k]1)$.

With $f[n_k] = g[n_k]$ we obtain $S(f[n_k])\downarrow$ and $\tau_{S(f[n_k])}(n_k+1)\downarrow$ and $S(f[n_k]) = S(f[n_k]0) = S(f[n_k]1)$. As $\tau_{S(f[n_k])}(n_k+1)\downarrow$, the hypothesis $S(f[n_k])$ cannot be conform for both $f[n_k]0$ and $f[n_k]1$. Then choose $y \in \{0,1\}$, such that $S(f[n_k]y)$ $(= S(f[n_k]))$ is not conform for $f[n_k]y$ and define $f$ as described above. Since $f \in U_1$, this yields the contradiction $U \notin \mathrm{CONF}_\tau(S)$.

*Case (ii)c.* $S(g[n_k])\downarrow$ and $S(g[n_k]) = S(g[n_k]0) = S(g[n_k]1)$, but none of the properties tested by parallel computation in the construction of $g$ is fulfilled.

Then $g = g[n_k+2]0^\infty \in U_0$ and $\tau_{S(g[n_k])}(n_k+1)\uparrow$ and $S(g[n]) = S(g[n_k])$ for all $n \geq n_k$. $\tau_{S(g[n_k])}(n_k+1)\uparrow$ implies $\tau_{S(g[n_k])} \neq g$; therefore the sequence of hypotheses produced by $S$ on $g$ converges to an index incorrect for $g$ with respect to $\tau$. Again we obtain $U \notin \mathrm{CONF}_\tau(S)$.

By construction of $g$ further cases cannot occur, so $U \notin \mathrm{CONF}$. Now we have verified $\mathrm{CEX} \setminus \mathrm{CONF} \neq \emptyset$, which implies $\mathrm{CEX} \,\#\, \mathrm{CONF}$. $\square$

As we have already mentioned above, for the inference criteria introduced in this section the following comparison results have been proved:

**Theorem 2** *[Ba74a, CS83, FKW95, JB81, Wie78, Zi99]*

1. $EX \subset BC \subset BC^* = \wp\mathcal{R}$,
2. $\forall m \in \mathbb{N}\ [EX_m \subset EX_{m+1} \subset EX]$,
3. $EX_0 \subset CP \subset TOTAL \subset CONS \subset CONF \subset EX$,
4. $TOTAL \subset CEX \subset EX$,
5. $CEX \,\#\, CONF$ *and* $CEX \,\#\, CONS$,
6. *if* $I \in \{CP, TOTAL, CONS, CONF, CEX\}$ *and* $m \geq 1$, *then* $EX_m \,\#\, I$.

These results are also summarized in Figure 1. The aim of this paper is to compare these hierarchies of inference criteria to the corresponding hierarchies resulting in uniform learning according to the same inference criteria. The definition of uniform learning is given in Section 3; Sections 4 and 5 are concerned with the comparison of inference criteria in uniform learning.

From now on let $\mathcal{I}$ denote the set of all previously declared inference criteria, i.e.

$$\mathcal{I} = \{\mathrm{EX}, \mathrm{BC}, \mathrm{BC}^*, \mathrm{CP}, \mathrm{TOTAL}, \mathrm{CEX}, \mathrm{CONS}, \mathrm{CONF}\} \cup \{\mathrm{EX}_m \mid m \in \mathbb{N}\} \ .$$

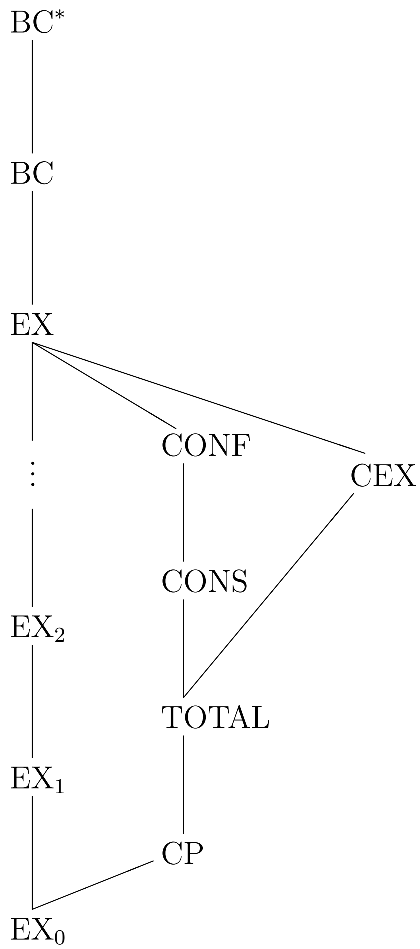


Figure 1: *The hierarchy of inference criteria according to Theorem 2. Any line drawn upwards indicates a proper inclusion. If two classes $I \in \mathcal{I}$ are not connected by a line or a sequence of lines drawn upwards, they are incomparable.*

# 3 Uniform Learning

The scope of this section is to give a formal introduction into uniform learning of classes of recursive functions. We will start with the basic definitions and afterwards collect some simple but useful results.

## 3.1 Definitions

From now on let $\varphi \in \mathcal{P}^3$ be a fixed acceptable numbering of $\mathcal{P}^2$ and $\tau \in \mathcal{P}^2$ an acceptable numbering of $\mathcal{P}^1$. As $\varphi$ is acceptable, it might be regarded as a numbering of all numberings $\psi \in \mathcal{P}^2$: every $b \in \mathbb{N}$ corresponds to the function $\varphi^b$ which is defined by $\varphi^b(i,x) := \varphi(b,i,x)$ for any $i, x \in \mathbb{N}$. Thus $b$ also describes a class $\mathcal{R}_b$ of recursive functions, where $\mathcal{R}_b := \mathcal{R}_{\varphi^b} = \mathcal{P}_{\varphi^b} \cap \mathcal{R}$; i.e. $\mathcal{R}_b$ is the recursive core of $\mathcal{P}_{\varphi^b}$. Therefore any set $B \subseteq \mathbb{N}$ will be called *description set* for the collection $\{\mathcal{R}_b \mid b \in B\}$ of recursive cores corresponding to the indices in $B$. Considering each recursive core as a set of functions to be identified, any description set $B \subseteq \mathbb{N}$ may be associated to a collection of learning problems. Now we are looking for a meta-learner which – given any description $b \in B$ – develops a special learner coping with the learning problem described by $b$, i.e. the special learner must identify each function in $\mathcal{R}_b$.

**Definition 11** *Let $J \subseteq \wp\mathcal{R}$, $I \in \mathcal{I}$, $J \subseteq I$, $B \subseteq \mathbb{N}$. The set $B$ is called suitable for uniform learning with respect to $J$ and $I$ iff the following conditions are fulfilled:*

1. *$\mathcal{R}_b \in J$ for all $b \in B$,*
2. *there is a function $S \in \mathcal{P}^2$ such that for all $b \in B$ there is a numbering $\psi \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in I_\psi(\lambda x.S(b,x))$.*

*We abbreviate this formulation by $B \in suit(J,I)$ and write $B \in suit(J,I)(S)$, if $S$ is given.*

So $B \in \mathrm{suit}(J,I)$ if and only if every recursive core described by some index $b \in B$ belongs to the class $J$ and additionally there is a strategy $S \in \mathcal{P}^2$ which, given $b \in B$, synthesizes an $I$-learner successful for $\mathcal{R}_b$ with respect to some appropriate hypothesis space $\psi$. Note that the synthesis of these appropriate hypothesis spaces is *not* required. This means in particular, that in general the output of a meta-learner cannot be interpreted practically, because we might not know which numbering is actually used as a hypothesis space. Of course we might restrict our definition of suitable description sets by demanding uniform learnability with respect to the acceptable numbering $\tau$ for all classes $\mathcal{R}_b$. Another possibility is to use the numberings $\varphi^b$, $b \in B$, already given by the description set $B$ as hypothesis spaces for $I$-identification of the classes $\mathcal{R}_b$.

**Definition 12** *Let* $J \subseteq \wp\mathcal{R}$, $I \in \mathcal{I}$, $J \subseteq I$, $B \subseteq \mathbb{N}$, $S \in \mathcal{P}^2$. *Assume* $B \in suit(J,I)(S)$. *We write* $B \in suit_\tau(J,I)(S)$ *if* $\mathcal{R}_b \in I_\tau(\lambda x.S(b,x))$ *for all* $b \in B$. *Furthermore the notation* $B \in suit_\varphi(J,I)(S)$ *shall indicate that* $\mathcal{R}_b \in I_{\varphi^b}(\lambda x.S(b,x))$ *for all* $b \in B$. *We also use the notations* $suit_\tau(J,I)$ *and* $suit_\varphi(J,I)$ *in the usual way.*

Note that the definition of $\text{suit}_\varphi$ corresponds to a special case of the definition of $\text{suit}_\tau$, because $\varphi^b$-numbers can be translated into $\tau$-numbers uniformly in $b$, as Proposition 1 states.

**Proposition 1** *There exists a recursive function* $c \in \mathcal{R}$, *such that* $\varphi^b_i = \tau_{c(b,i)}$ *for all* $b, i \in \mathbb{N}$.

*Proof.* Let $p : \mathbb{N}^2 \to \mathbb{N}$ be a bijective recursive function and choose $\pi_1, \pi_2 \in \mathcal{R}$, such that $\pi_1(p(x,y)) = x$ and $\pi_2(p(x,y)) = y$ for all $x, y \in \mathbb{N}$. Then define a function $\psi \in \mathcal{P}^2$ by

$$\psi_j(x) := \varphi(\pi_1(j), \pi_2(j), x) \text{ for all } j, x \in \mathbb{N} .$$

As $\psi \in \mathcal{P}^2$ and $\tau$ is acceptable, there exists some $d \in \mathcal{R}$ satisfying

$$\psi_j = \tau_{d(j)} \text{ for all } j \in \mathbb{N} .$$

By defining $c(b,i) := d(p(b,i))$ for all $b, i \in \mathbb{N}$ we obtain

$$\tau_{c(b,i)} = \tau_{d(p(b,i))} = \psi_{p(b,i)} = \varphi^{\pi_1(p(b,i))}_{\pi_2(p(b,i))} = \varphi^b_i$$

for all $b, i \in \mathbb{N}$. □

As many results in this paper are concerned with the uniform learnability of finite or singleton sets, we introduce the following notions:

**Definition 13**

- $J^* := \{U \subseteq \mathcal{R} \mid \mathit{card}\ U < \infty\}$,
- $J^1 := \{U \subseteq \mathcal{R} \mid \mathit{card}\ U = 1\} = \{\{f\} \mid f \in \mathcal{R}\}$.

Thus $J^*$ denotes the set of all finite subsets of $\mathcal{R}$; $J^1$ is the set of all singleton sets of recursive functions.

## 3.2 Basic Results

Of course it would be nice to find characterizations of the sets suitable for uniform learning with respect to $J, I$, where $I \in \mathcal{I}$ and $J \subseteq I$ are given. This paper compares the uniform identification power of several criteria $I \in \mathcal{I}$ and concentrates on the case $J = J^*$, i.e. all recursive cores to be identified with respect to $I$ are finite. Our first result follows obviously from our definitions and Proposition 1.

**Proposition 2** *Let $I \in \mathcal{I},\ J \subseteq I$. Then $suit_\varphi(J, I) \subseteq suit_\tau(J, I) \subseteq suit(J, I)$.*

Whether these inclusions are proper inclusions or not depends on the choice of $J$ and $I$. If they turned out to be equalities for all $J$ and $I$, then Definition 12 would be superfluous. But in fact, as Theorem 6 will show, we have proper inclusions in the general case.

Any strategy identifying a class $U \subseteq \mathcal{R}$ with respect to some criterion $I \in \mathcal{I} \setminus \{\text{CONS}, \text{CONF}\}$ can be replaced by a *total* recursive strategy without loss of learning power – a result from folklore. This new strategy is defined by computing the values of the former strategy for a bounded number of steps and a bounded number of input examples with increasing bounds. As long as no hypothesis is found, some temporary hypothesis agreeing with the restrictions in the definition of $I$ is produced. Afterwards the hypotheses of the former strategy are put out "with delay". This does not work for CONS and CONF, since in general after the delay the hypotheses are no longer consistent or conform with the information in the actual time of the learning process. An example of a class in CONS which is not consistently identifiable by any total recursive strategy can be found in [WZ95]. Now we transfer these observations to the level of uniform learning and get the following result, which we will use in several proofs:

**Proposition 3** *Let $I \in \mathcal{I} \setminus \{CONS, CONF, CP\},\ J \subseteq I,\ B \subseteq \mathbb{N}$. Assume $B \in suit(J, I)$ ($suit_\tau(J, I)$). Then there is a total recursive function $S$ such that $B \in suit(J, I)(S)$ ($suit_\tau(J, I)(S)$, respectively). Furthermore, if $I \notin \{TOTAL, CEX\}$ and $B \in suit_\varphi(J, I)$, then also $B \in suit_\varphi(J, I)(S)$ for some $S \in \mathcal{R}^2$.*

We also have to exclude the criterion CP here, because in general a class-preserving hypothesis for $\mathcal{R}_b$ cannot be computed uniformly in $b$. A counterclaim to Proposition 3 in the case of learning with respect to $\text{suit}_\varphi$ for the criteria CONS, CONF, CEX, TOTAL and CP is verified by the example given below:

**Example 1** *The description set $B \subseteq \mathbb{N}$ given by*

$$B := \{b \in \mathbb{N} \mid \forall f \in \mathcal{R}_b\ \forall g \in \mathcal{P}_{\varphi^b} \setminus \{f\}\ [g(0)\!\downarrow\ \Rightarrow\ g(0) \neq f(0)]\}$$

*is an element of $suit_\varphi(I, I)$ for all $I \in \{CONS, CONF, CEX, TOTAL, CP\}$, but*

$$B \notin suit_\varphi(I, I)(S)$$

*for all $I \in \{CONS, CONF, CEX, TOTAL, CP\}$ and any total recursive strategy $S \in \mathcal{R}^2$.*

*Proof.* If we set $T(b, f[0]) :=$ *"Return some $i \in \mathbb{N}$ with $\varphi_i^b(0) = f(0)$"* and $T(b, f[n+1]) := T(b, f[0])$ for arbitrary $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$, we observe that $B \in \text{suit}_\varphi(I, I)(T)$ for all $I \in \{\text{CONS}, \text{CONF}, \text{CEX}, \text{TOTAL}, \text{CP}\}$. Otherwise there was an integer $b \in B$ and two functions $f \in \mathcal{R}_b$, $g \in \mathcal{P}_{\varphi^b} \setminus \{f\}$ with $g(0) = f(0)$, which contradicts the definition of $B$.

Now assume that there exists some strategy $S \in \mathcal{R}^2$ satisfying

$$B \in \text{suit}_\varphi(I, I)(S)$$

for some $I \in \{\text{CONS}, \text{CONF}, \text{CEX}, \text{TOTAL}, \text{CP}\}$. We will deduce a contradiction by constructing an integer $b_0 \in \mathbb{N}$ satisfying

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin I_{\varphi^{b_0}}(\lambda x.S(b_0, x))$.

*Construction of $b_0$.*
Define a function $\psi \in \mathcal{P}^3$ for arbitrary $b \in \mathbb{N}$ as follows:

$$\psi_0^b := \begin{cases} 0^\infty & \text{if } S(b,0) \neq 0 \\ \uparrow^\infty & \text{if } S(b,0) = 0 \text{ and } I \neq \text{CONF} \\ 1\uparrow^\infty & \text{if } S(b,0) = 0 \text{ and } I = \text{CONF} \end{cases}$$

$$\psi_1^b := \begin{cases} 0^\infty & \text{if } S(b,0) = 0 \\ \uparrow^\infty & \text{if } S(b,0) \neq 0 \text{ and } I \neq \text{CONF} \\ 1\uparrow^\infty & \text{if } S(b,0) \neq 0 \text{ and } I = \text{CONF} \end{cases}$$

$$\psi_x^b := \begin{cases} \uparrow^\infty & \text{if } I \neq \text{CONF} \\ 1\uparrow^\infty & \text{if } I = \text{CONF} \end{cases} \quad \text{for all } x \geq 2.$$

Since $S \in \mathcal{R}^2$, we know that $S(b, 0)$ is defined for all $b \in \mathbb{N}$, so $\psi$ is a partial-recursive function. Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction $b_0$.*

It remains to prove the following properties:

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin I_{\varphi^{b_0}}(\lambda x.S(b_0, x))$.

*ad 1.* By definition of the function $\psi$ we know that $\mathcal{P}_{\varphi^{b_0}} = \{0^\infty, \uparrow^\infty\}$ (or $\{0^\infty, 1\uparrow^\infty\}$ if $I = \text{CONF}$). Thus for any $f \in \mathcal{R}_{b_0}$ and $g \in \mathcal{P}_{\varphi^{b_0}} \setminus \{f\}$ we obtain $f = 0^\infty$ as well as $g =\uparrow^\infty$ (or $\{g = 1\uparrow^\infty\}$ if $I = \text{CONF}$); in particular $g(0) \neq f(0)$. Thus $b_0 \in B$. *qed 1.*

*ad 2.* Since $S \in \mathcal{R}^2$, it suffices to consider the following two cases.

*Case (i).* $S(b_0, 0) = 0$.
Then $\varphi_1^{b_0} = 0^\infty \in \mathcal{R}_{b_0}$ and $\varphi_0^{b_0} =\uparrow^\infty$ (or $\varphi_0^{b_0} = 1\uparrow^\infty$ if $I = \text{CONF}$). This implies $S(b_0, \varphi_1^{b_0}[0]) = S(b_0, 0) = 0$. But $\varphi_0^{b_0}$ is not an element of the recursive core $\mathcal{R}_{b_0}$ and $\varphi_0^{b_0}(0)\uparrow$ (or $\varphi_0^{b_0}(0) = 1$ if $I = \text{CONF}$). So on $\varphi_1^{b_0}$ the strategy $\lambda x.S(b_0, x)$ returns a hypothesis which is neither consistent nor class-preserving, nor total, nor convergently incorrect (or which is non-conform, if $I = \text{CONF}$). This implies $\mathcal{R}_{b_0} \notin I_{\varphi^{b_0}}(\lambda x.S(b_0, x))$.

*Case (ii).* $S(b_0, 0) \neq 0$.
Then $\varphi_0^{b_0} = 0^\infty \in \mathcal{R}_{b_0}$ and $\varphi_x^{b_0} =\uparrow^\infty$ (or $\varphi_x^{b_0} = 1\uparrow^\infty$ if $I = \text{CONF}$) for all $x \geq 1$. By analogy with our first case we verify $\mathcal{R}_{b_0} \notin I_{\varphi^{b_0}}(\lambda x.S(b_0, x))$.

Further cases do not occur, so $\mathcal{R}_{b_0} \notin I_{\varphi^{b_0}}(\lambda x.S(b_0, x))$. *qed 2.*

These two properties of $b_0$ now contradict our assumption. This implies that there is no recursive strategy $S \in \mathcal{R}^2$ satisfying $B \in \text{suit}_\varphi(I, I)(S)$, where $I \in \{\text{CONS}, \text{CONF}, \text{CEX}, \text{TOTAL}, \text{CP}\}$ was chosen arbitrarily. □

A further example will prove the counterclaim to Proposition 3 concerning suit or $\text{suit}_\tau$ in combination with any of the criteria CONS and CONF.

**Example 2** *The description set $B \subseteq \mathbb{N}$ given by*

$$B := \{b \in \mathbb{N} \mid \varphi^b \in \mathcal{R}^2\}$$

*is an element of $\mathit{suit}_\varphi(\mathit{CONS}, \mathit{CONS})$, but $B \notin \mathit{suit}(\mathit{CONS}, \mathit{CONF})(S)$ for any recursive strategy $S \in \mathcal{R}^2$.*

*Proof.* Obviously $\mathcal{R}_b \in \text{CONS}_{\varphi^b}(\text{Enum}_{\varphi^b})$ for all $b \in \mathbb{N}$. Since a program for the strategy $\text{Enum}_{\varphi^b}$ can be found uniformly in the description $b$, we obtain $B \in \text{suit}_\varphi(\text{CONS}, \text{CONS})$.

It remains to prove that $B \notin \text{suit}(\text{CONS}, \text{CONF})(S)$ for all $S \in \mathcal{R}^2$. Assume to the contrary that there exists some strategy $S \in \mathcal{R}^2$ satisfying

$$B \in \text{suit}(\text{CONS}, \text{CONF})(S) \ .$$

We will deduce a contradiction by constructing an integer $b_0 \in \mathbb{N}$ satisfying

1. $b_0 \in B$,

2. $\mathcal{R}_{b_0} \notin \mathrm{CONF}_\eta(\lambda x.S(b_0,x))$ for all $\eta \in \mathcal{P}^2$.

*Construction of* $b_0$.
Define a function $\psi \in \mathcal{P}^3$ for arbitrary $b \in \mathbb{N}$ as follows:

$$\psi_i^b(0) := 0 \text{ for all } i \in \mathbb{N}\ ,$$

$$\psi_0^b(x+1) := \begin{cases} 0 & \text{if } S(b,\psi_0^b[x]) \neq S(b,\psi_0^b[x]0) \quad\quad\quad\quad\quad \text{(A)} \\ 1 & \text{if } S(b,\psi_0^b[x]) = S(b,\psi_0^b[x]0) \text{ and} \\ & S(b,\psi_0^b[x]) \neq S(b,\psi_0^b[x]1) \quad\quad\quad\quad\quad \text{(B)} \\ 0 & \text{if } S(b,\psi_0^b[x]) = S(b,\psi_0^b[x]0) = S(b,\psi_0^b[x]1) \text{ (C)} \end{cases} \quad \text{for } x \in \mathbb{N}\ ,$$

$$\psi_{i+1}^b(x+1) := \begin{cases} \psi_0^b(x+1) & \text{if case (C) occurs at most } i \text{ times} \\ & \text{in the definition of } \psi_0^b[x+1] \\ 1 & \text{otherwise} \end{cases} \quad \text{for } i,x \in \mathbb{N}\ .$$

Whenever case (A) or case (B) occurs, $\lambda x.S(b,x)$ changes its mind on $\psi_0^b$. As $S \in \mathcal{R}^2$, we have $\psi \in \mathcal{R}^3$. Note that for all $x \in \mathbb{N}$ such that $S(b,\psi_0^b[x]) = S(b,\psi_0^b[x]0) = S(b,\psi_0^b[x]1)$ there is some $i \in \mathbb{N}$ which fulfils

$$\psi_{i+1}^b =_x \psi_0^b \text{ and } \psi_{i+1}^b(x+1) = 1 \neq 0 = \psi_0^b(x+1) \tag{1}$$

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction* $b_0$.

It remains to prove the following properties:

1. $b_0 \in B$,

2. $\mathcal{R}_{b_0} \notin \mathrm{CONF}_\eta(\lambda x.S(b_0,x))$ for all $\eta \in \mathcal{P}^2$.

*ad 1.* We know $\psi \in \mathcal{R}^3$. Since $\varphi^{b_0} = \psi^{b_0}$, this implies $\varphi^{b_0} \in \mathcal{R}^2$. Hence we obtain $b_0 \in B$. *qed 1.*

*ad 2.* Assume that $\mathcal{R}_{b_0} \in \mathrm{CONF}_\eta(\lambda x.S(b_0,x))$ for some fixed $\eta \in \mathcal{P}^2$. As $\varphi_0^{b_0} \in \mathcal{R}$, one of the cases (A), (B), (C) must occur infinitely often in the definition of $\psi_0^{b_0} = \varphi_0^{b_0}$. We consider two possible cases.

*Case (i).* Cases (A) or (B) occur infinitely often.
Then – according to the remark in the definition of $\psi$ – the strategy $\lambda x.S(b_0,x)$ changes its mind on $\varphi_0^{b_0}$ infinitely often. Therefore $\mathcal{R}_{b_0} \notin \mathrm{CONF}_\eta(\lambda x.S(b_0,x))$,

which contradicts our assumption.

*Case (ii).* Case (C) occurs infinitely often.
Then there are infinitely many integers $x \in \mathbb{N}$ such that

$$S(b_0, \varphi_0^{b_0}[x]) = S(b_0, \varphi_0^{b_0}[x]0) = S(b_0, \varphi_0^{b_0}[x]1) \ . \tag{2}$$

Let $z$ be one of these integers. According to (1) there must be an $i \in \mathbb{N}$ such that

$$\begin{aligned}&\varphi_{i+1}^{b_0} =_z \varphi_0^{b_0} \text{ and } \varphi_{i+1}^{b_0}(z+1) = 1 \neq 0 = \varphi_0^{b_0}(z+1)\\ &\text{and } S(b_0, \varphi_{i+1}^{b_0}[z+1]) = S(b_0, \varphi_0^{b_0}[z+1]) \ .\end{aligned}$$

Since $\varphi_{i+1}^{b_0}, \varphi_0^{b_0} \in \mathcal{R}_{b_0}$ and $\mathcal{R}_{b_0} \in \text{CONF}_\eta(\lambda x.S(b_0, x))$, the hypothesis $j := S(b_0, \varphi_0^{b_0}[z+1])$ must be conform with respect to $\eta$ for both $\varphi_0^{b_0}[z+1]$ and $\varphi_{i+1}^{b_0}[z+1]$, therefore $\eta_j(z+1)\uparrow$. In particular $j$ is incorrect for $\varphi_0^{b_0}$ with respect to $\eta$. That means for every $x \in \mathbb{N}$ satisfying condition (2), that the hypothesis $S(b_0, \varphi_0^{b_0}[x+1])$ is incorrect for $\varphi_0^{b_0}$ with respect to $\eta$. Now $\lambda x.S(b_0, x)$ returns incorrect hypotheses for $\varphi_0^{b_0}$ infinitely often, because there are infinitely many integers $x$ satisfying condition (2). We conclude $\mathcal{R}_{b_0} \notin \text{CONF}_\eta(\lambda x.S(b_0, x))$ in contradiction to our assumption.

As at least one of these two cases must occur, we have verified that $\mathcal{R}_{b_0} \notin \text{CONF}_\eta(\lambda x.S(b_0, x))$, where the hypothesis space $\eta \in \mathcal{P}^2$ was chosen arbitrarily. This proves 2. *qed 2.*

These two properties of $b_0$ now contradict our assumption. This implies that there is no recursive strategy $S \in \mathcal{R}^2$ satisfying $B \in \text{suit}(\text{CONS}, \text{CONF})(S)$. □

Let us now collect some simple examples of description sets suitable or not suitable for uniform learning. First we consider the identification of classes consisting of just one recursive function. Any set describing such classes turns out to be suitable for identification under any of our criteria:

**Theorem 3** *Let $I \in \mathcal{I}$. Then suit$(J^1, I) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \in J^1$ for all $b \in B\}$.*

*Proof.* Let $B \subseteq \mathbb{N}$ fulfil $\mathcal{R}_b \in J^1$ for all $b \in B$. Since for all $f \in \mathcal{R}$ there exists a numbering $\psi \in \mathcal{P}^2$ with $\psi_0 = f$, the strategy constantly zero yields $B \in \text{suit}(J^1, I)$. Thus $\{B \subseteq \mathbb{N} \mid \mathcal{R}_b \in J^1 \text{ for all } b \in B\} \subseteq \text{suit}(J^1, I)$. The other inclusion is obvious. □

Unfortunately, we would rather not regard the strategy defined in this proof as an "intelligent" learner, because its output does not depend on the input at all. Its success lies just in the choice of appropriate hypothesis spaces. If such a choice of hypothesis spaces is forbidden, we obtain an absolutely negative result:

**Theorem 4**

1. *$\{b \in \mathbb{N} \mid \mathcal{R}_b \in J^1\} \notin suit_\tau(J^1, BC)$;*
   *in particular even $\{b \in \mathbb{N} \mid card\ \{i \in \mathbb{N} \mid \varphi^b_i \in \mathcal{R}\} = 1\} \notin suit_\tau(J^1, BC)$,*

2. *$\{b \in \mathbb{N} \mid \mathcal{R}_b \in J^1\} \notin suit_\varphi(J^1, BC^*)$;*
   *in particular even $\{b \in \mathbb{N} \mid card\ \{i \in \mathbb{N} \mid \varphi^b_i \in \mathcal{R}\} = 1\} \notin suit_\varphi(J^1, BC^*)$.*

For a proof see [Zi00]. So, if we fix our hypothesis spaces in advance, not even the classes consisting of just one element can be BC-identified uniformly. Regarding the identification of arbitrary finite classes (the learnability of which is trivial in the non-uniform case), the situation does not improve. Even by free choice of the hypothesis spaces we cannot achieve uniform EX-identifiability.

**Theorem 5** *$\{b \in \mathbb{N} \mid \mathcal{R}_b \in J^*\} \notin suit(J^*, EX)$.*

A proof can be found in [Zi00]. How can we interpret these results? Is the concept of uniform learning fruitless and further research on this area not worthwile? Fortunately, many results in [BCJ96] and [Zi00] allow a more optimistic point of view. For example, [Zi00] shows that some constraints on the descriptions $b \in B$ – especially concerning the topological structure of the numberings $\varphi^b$ – yield uniform learnability of huge classes of recursive functions, even with consistent *and* total intermediate hypotheses and also with respect to our acceptable numbering $\tau$. The sticking point seems to be that uniform identifiability is not so much influenced by the classes to be learned, but by the numberings $\varphi^b$ chosen as representations for these classes. So the many negative results should be interpreted carefully. For example the reason that there is no uniform EX-learner for $\{b \in \mathbb{N} \mid \mathcal{R}_b \in J^*\}$ is not so much the complexity of finite classes but rather the need to cope with *any* numbering possessing a finite $\mathcal{R}$-core. Based on these aspects we do not tend to a pessimistic view concerning the fruitfulness of the concept of uniform learning. Our results in the following sections will substantiate this opinion.

Theorems 3 and 4 now enable the proof of the following example of a strict version of Proposition 2.

**Theorem 6**

1. *$suit_\varphi(J^1, I) \subset suit_\tau(J^1, I) \subset suit(J^1, I)$ for all $I \in \mathcal{I} \setminus \{BC^*\}$,*

2. *$suit_\varphi(J^1, BC^*) \subset suit_\tau(J^1, BC^*) = suit(J^1, BC^*)$.*

*Proof.* *ad 1.* Let $I \in \mathcal{I} \setminus \{\mathrm{BC}^*\}$. $\mathrm{suit}_\tau(J^1, I) \subset \mathrm{suit}(J^1, I)$ is obtained as follows: by Theorem 4 we know that $B_1 := \{b \in \mathbb{N} \mid \mathcal{R}_b \in J^1\} \notin \mathrm{suit}_\tau(J^1, I)$ (otherwise $B_1$ was also an element of $\mathrm{suit}_\tau(J^1, \mathrm{BC})$). Thus by Theorem 3 we obtain $B_1 \in \mathrm{suit}(J^1, I) \setminus \mathrm{suit}_\tau(J^1, I)$.

It remains to prove $\mathrm{suit}_\varphi(J^1, I) \subset \mathrm{suit}_\tau(J^1, I)$. Again by Theorem 4 we know that there exists a set $B \subseteq \mathbb{N}$ such that card $\{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\} = 1$ for all $b \in B$ and $B \notin \mathrm{suit}_\varphi(J^1, \mathrm{EX})$. Now let $g \in \mathcal{R}$ be a computable function satisfying

$$\varphi_i^{g(b)}(x) = \begin{cases} 0 & \text{if } \varphi_i^b(y)\!\downarrow \text{ for all } y \leq x \\ \uparrow & \text{otherwise} \end{cases}$$

for any $b, i, x \in \mathbb{N}$. Let $B' := \{g(b) \mid b \in B\}$. Since $\mathcal{R}_{g(b)} = \{0^\infty\}$ for all $b \in B$, we get $B' \in \mathrm{suit}_\tau(J^1, I)$ (via a strategy which constantly returns a $\tau$-index of the function $0^\infty$).

Obviously $\{i \in \mathbb{N} \mid \varphi_i^{g(b)} \in \mathcal{R}\} = \{i \in \mathbb{N} \mid \varphi_i^b \in \mathcal{R}\}$ for all $b \in \mathbb{N}$. If there was a strategy $S \in \mathcal{P}^2$ satisfying $B' \in \mathrm{suit}_\varphi(J^1, I)(S)$, we would achieve $B \in \mathrm{suit}_\varphi(J^1, \mathrm{EX})(T)$ by defining $T(b, f[n]) := S(g(b), 0^n)$ for $f \in \mathcal{R},\ b, n \in \mathbb{N}$. This contradicts the choice of $B$, so $B' \in \mathrm{suit}_\tau(J^1, I) \setminus \mathrm{suit}_\varphi(J^1, I)(S)$. Hence $\mathrm{suit}_\varphi(J^1, I) \subset \mathrm{suit}_\tau(J^1, I)$. *qed 1.*

*ad 2.* By Theorem 2 we know $\mathcal{R} \in \mathrm{BC}^*$. Therefore there exists some strategy $T \in \mathcal{R}$ such that $\mathcal{R} \in \mathrm{BC}^*_\tau(T)$. Defining a uniform learner $S$ by

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

for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$, we obtain $\wp\mathbb{N} \in \mathrm{suit}_\tau(\mathrm{BC}^*, \mathrm{BC}^*)(S)$. So

$$\mathrm{suit}_\tau(\mathrm{BC}^*, \mathrm{BC}^*) = \mathrm{suit}(\mathrm{BC}^*, \mathrm{BC}^*) = \wp\mathbb{N}$$

and in particular $\mathrm{suit}_\tau(J^1, \mathrm{BC}^*) = \mathrm{suit}(J^1, \mathrm{BC}^*)$ is verified. The proper inclusion $\mathrm{suit}_\varphi(J^1, \mathrm{BC}^*) \subset \mathrm{suit}_\tau(J^1, \mathrm{BC}^*)$ is a consequence of Theorem 4.2, because the description set $\{b \in \mathbb{N} \mid \mathcal{R}_b \in J^1\}$ is an element of $\mathrm{suit}_\tau(J^1, \mathrm{BC}^*)$, but does not belong to $\mathrm{suit}_\varphi(J^1, \mathrm{BC}^*)$. *qed 2.*

□

# 4 Separation of Inference Criteria: Special Hypothesis Spaces

From now on we will compare the learning power of our inference criteria for uniform learning of finite classes of recursive functions, i.e. we try to find results in the style of Theorem 2, where the criteria $I \in \mathcal{I}$ are replaced by the sets $\mathrm{suit}(J^*, I)$, $\mathrm{suit}_\tau(J^*, I)$ or $\mathrm{suit}_\varphi(J^*, I)$. Please note that a separation like for example $\mathrm{suit}_\tau(\mathrm{CONS}, \mathrm{CONS}) \subset \mathrm{suit}_\tau(\mathrm{CONS}, \mathrm{EX})$ is not a very astonishing result. The remarkable point is that even collections of *finite* classes of recursive functions suffice for a separation (note that in the non-uniform case finite classes can be identified under *any* criterion $I \in \mathcal{I}$ easily).

In this section we concentrate on uniform learning with respect to fixed hypothesis spaces, i.e. according to Definition 12. Our aim is to show that all comparison results in Theorem 2 hold analogously for these concepts, even if all classes to be learned are finite. Lemma 1 summarizes some very simple observations.

**Lemma 1**

1. *$suit_\varphi(J^*, EX) \subseteq suit_\varphi(J^*, BC) \subseteq suit_\varphi(J^*, BC^*)$,*
2. *$suit_\varphi(J^*, EX_m) \subseteq suit_\varphi(J^*, EX_{m+1}) \subseteq suit_\varphi(J^*, EX)$ for all $m \in \mathbb{N}$,*
3. *$suit_\varphi(J^*, CP) \subseteq suit_\varphi(J^*, TOTAL) \subseteq suit_\varphi(J^*, CONS) \subseteq suit_\varphi(J^*, CONF) \subseteq suit_\varphi(J^*, EX)$,*
4. *$suit_\varphi(J^*, TOTAL) \subseteq suit_\varphi(J^*, CEX) \subseteq suit_\varphi(J^*, EX)$,*
5. *$suit_\varphi(J^*, EX_0) \subseteq suit_\varphi(J^*, CONS)$ and $suit_\varphi(J^*, EX_0) \subseteq suit_\varphi(J^*, CEX)$ [Ne01],*
6. *$suit_\tau(J^*, EX_0) \subseteq suit_\tau(J^*, TOTAL)$.*

*These results hold analogously if we substitute $suit_\varphi$ by $suit_\tau$.*

*Proof.* All inclusions except for $\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL}) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{CONS})$ (or analogously with $\mathrm{suit}_\tau$ instead of $\mathrm{suit}_\varphi$) and $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_0) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{CEX})$ follow immediately from the definitions.

*Proof of "$suit_\varphi(J^*, TOTAL) \subseteq suit_\varphi(J^*, CONS)$" and the corresponding $\tau$-case.* If some set $B \subseteq \mathbb{N}$ fulfils $B \in \mathrm{suit}_\varphi(J^*, \mathrm{TOTAL})(S)$ for a strategy $S \in \mathcal{P}^2$, we can easily define $T \in \mathcal{P}^2$ such that $B \in \mathrm{suit}_\varphi(J^*, \mathrm{CONS})(T)$. On input $(b, f[n])$ the strategy $T$ just has to check the hypothesis $S(b, f[n])$ for consistency with respect to $\varphi^b$. For $b \in B$, $f \in \mathcal{R}_b$ this check is possible, because $\varphi^b_{S(b,f[n])}$ must be a total function. If consistency is verified, $T$ returns the same hypothesis as $S$, otherwise $T$ returns some consistent hypothesis (which can be found, if $f \in \mathcal{R}_b$). Convergence to a correct hypothesis follows from the choice of $S$. The $\tau$-case is proved by analogy.

*Proof of "$suit_\varphi(J^*, EX_0) \subseteq suit_\varphi(J^*, CEX)$".*
This proof is based on a personal communication to Jochen Nessel. Let $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_0)$. Then there exists a strategy $T \in \mathcal{R}^2$, such that

$$\mathcal{R}_b \in (\mathrm{EX}_0)_{\varphi^b}(\lambda x.T(b,x))$$

for all $b \in B$. A strategy $S \in \mathcal{P}^2$ successful for $B$ according to the definition of $\mathrm{suit}_\varphi(J^*, \mathrm{CEX})$ can be described in the following way:

> *On input $(b, f[n]) \in \mathbb{N}^2$ first compute $T(b, f[n])$. If $T(b, f[n]) \neq ?$, then return $T(b, f[n])$. Otherwise look for some $(i, m) \in \mathbb{N}^2$, such that $f[n] \subset \varphi_i^b[m]$ and $T(b, \varphi_i^b[m]) = i$. As soon as such a pair $(i, m)$ is found, return $i$.*

Now assume $b \in B$ and $f \in \mathcal{R}_b$. Since $\mathcal{R}_b \in (\mathrm{EX}_0)_{\varphi^b}(\lambda x.T(b,x))$, there must be some $j, n_0 \in \mathbb{N}$, such that $\varphi_j^b = f$ and $T(b, f[n]) = j$ for all $n \geq n_0$. Then also $S(b, f[n]) = j$ for all $n \geq n_0$, so in the limit our hypotheses are correct with respect to $\varphi^b$.

It remains to prove that $S(b, f[n])$ is defined for all $n \in \mathbb{N}$ and that none of the intermediate hypotheses produced by $\lambda x.S(b,x)$ on $f$ correspond to proper subfunctions of $f$. For that purpose choose some $n \in \mathbb{N}$ such that $T(b, f[n]) = ?$ (if such an $n$ exists). Since $f \in \mathcal{R}_b$ and $\mathcal{R}_b \in (\mathrm{EX}_0)_{\varphi^b}(\lambda x.T(b,x))$, our strategy $S$ must find some pair $(i, m) \in \mathbb{N}^2$, such that $f[n] \subseteq \varphi_i^b[m]$ and $T(b, \varphi_i^b[m]) = i$. So $S(b, f[n]) = i$; in particular, $S(b, f[n])$ is defined.

If $\varphi_i^b[m] \subseteq f$, then also $T(b, f[m]) = i$. As $T$ learns $\mathcal{R}_b$ from $b$ without a mind change, this implies $\varphi_i^b = f$. Thus $S(b, f[n]) = i$ does not correspond to a proper subfunction of $f$.

If $\varphi_i^b[m] \not\subseteq f$, then $\varphi_i^b \not\subseteq f$, i.e. again $S(b, f[n]) = i$ does not correspond to a proper subfunction of $f$.

Hence we obtain $\mathcal{R}_b \in \mathrm{CEX}_{\varphi^b}(\lambda x.S(b,x))$ for all $b \in B$ and therefore $B \in \mathrm{suit}_\varphi(J^*, \mathrm{CEX})$. As $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_0)$ was chosen arbitrarily, we conclude $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_0) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{CEX})$. □

Now we want to prove that nearly all these inclusions are in fact proper inclusions.

## 4.1 The Hierarchy in Lemma 1.1

In this subsection we will show that the hierarchy in Lemma 1.1 is given by strict separations. For that purpose we have to verify $\mathrm{suit}_\varphi(J^*, \mathrm{EX}) \subset \mathrm{suit}_\varphi(J^*, \mathrm{BC})$

and $\mathrm{suit}_\varphi(J^*, \mathrm{BC}) \subset \mathrm{suit}_\varphi(J^*, \mathrm{BC}^*)$ as well as the corresponding results for the $\tau$-case. The separation of $\mathrm{suit}_\varphi(J^*, \mathrm{BC})$ and $\mathrm{suit}_\varphi(J^*, \mathrm{BC}^*)$ (and analogously with $\mathrm{suit}_\tau$ instead of $\mathrm{suit}_\varphi$) is achieved by Theorem 7.

**Theorem 7** *$suit_\varphi(J^1, BC^*) \setminus suit_\tau(J^1, BC) \neq \emptyset$.*

*Proof.* We use a strategy $T \in \mathcal{R}$ to define a description set $B \subseteq \mathbb{N}$ suitable for uniform $\mathrm{BC}^*$-identification by $T$. The set $B$ shall describe only singleton sets of recursive functions and will not be suitable for uniform behaviourally correct identification. The choice of the strategy $T$ may seem rather arbitrary, but it will enable an indirect proof.

*Definition of $B \in suit_\varphi(J^1, BC^*) \setminus suit_\tau(J^1, BC)$.*
Define a strategy $T \in \mathcal{R}$ by

$$T(f[n]) := \mathrm{card}\ \{i \leq n \mid f(i) = 1\} + 1$$

for all $f \in \mathcal{R}$ and $n \in \mathbb{N}$. Then set

$$B := \{b \in \mathbb{N} \mid \mathrm{card}\ \mathcal{R}_b = 1 \text{ and } \mathcal{R}_b \in \mathrm{BC}^*_{\varphi^b}(T)\}\ .$$

*Claim.* $B \in \mathrm{suit}_\varphi(J^1, \mathrm{BC}^*) \setminus \mathrm{suit}_\tau(J^1, \mathrm{BC})$.

*Proof of "$B \in suit_\varphi(J^1, BC^*)$".*
Defining $T' \in \mathcal{R}^2$ by $T'(b, f[n]) := T(f[n])$ for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ we obviously have

- $\forall b \in B\ [\mathcal{R}_b \in J^1]$,
- $\forall b \in B\ [\mathcal{R}_b \in \mathrm{BC}^*_{\varphi^b}(\lambda x.T'(b, x))]$.

By Definition 12 this implies $B \in \mathrm{suit}_\varphi(J^1, \mathrm{BC}^*)$.

*Proof of "$B \notin suit_\tau(J^1, BC)$".*
We will verify this claim by way of contradiction.
*Assumption.* $B \in \mathrm{suit}_\tau(J^1, \mathrm{BC})$.

Then there exists a strategy $S \in \mathcal{R}^2$ such that

$$\forall b \in B\ [\mathcal{R}_b \in \mathrm{BC}_\tau(\lambda x.S(b, x))]\ . \tag{3}$$

*Aim.* Construction of an integer $b_0$, such that

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

in contradiction to statement (3). The strategy $\lambda x.S(b_0,x)$ will fail for the only function $f \in \mathcal{R}_{b_0}$ by returning hypotheses incorrect for $f$ with respect to $\tau$ infinitely often.

*Construction of $b_0$.*
Define a function $\psi \in \mathcal{P}^3$ by the following instructions: let $b \in \mathbb{N}$, $\alpha_0^b := 0$. Go to stage 0.

*Stage 0.* $\psi_0^b(0) := 0$, i.e. $\psi_0^b[0] = \alpha_0^b$. For $x \in \mathbb{N}$ set

$$\psi_1^b(x) := \begin{cases} 0 & \text{if } x = 0 \\ 0 & \text{if } x > 0 \text{ and for all } m \in \{1,\ldots,x\} \\ & \quad [\tau_{S(b,0^m)}(m)\downarrow_{\le x} \Rightarrow\ \tau_{S(b,0^m)}(m) \neq 0] \\ \uparrow & \text{if } x > 0 \text{ and } x \text{ is minimal such that} \\ & \tau_{S(b,0^m)}(m)\downarrow_{\le x} \text{ and } \tau_{S(b,0^m)}(m) = 0 \text{ for some } m \in \{1,\ldots,x\} \\ \psi_0^b(x) & \text{otherwise} \end{cases}$$

If there is some $x \in \mathbb{N}$ and some $m \in \{1,\ldots,x\}$ such that

$$\tau_{S(b,0^m)}(m)\downarrow_{\le x} \text{ and } \tau_{S(b,0^m)}(m) = 0\ ,$$

let $m_0^b$ be the minimal integer such that

$$\tau_{S(b,0^{m_0^b})}(m_0^b)\downarrow_{\le x} \text{ and } \tau_{S(b,0^{m_0^b})}(m_0^b) = 0\ .$$

Then define $\alpha_1^b := 0^{m_0^b}1$ and go to stage 1. *End stage 0.*

*Stage k ($k \in \mathbb{N}, k > 0$).* $\psi_0^b[|\alpha_k^b| - 1] := \alpha_k^b$. For $x \in \mathbb{N}$ set

$$\psi_{k+1}^b(x) := \begin{cases} \alpha_k^b(x) & \text{if } x < |\alpha_k^b| \\ 0 & \text{if } x \ge |\alpha_k^b| \text{ and for all } m \in \{1,\ldots,x\} \\ & \quad [\tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m)\downarrow_{\le x} \Rightarrow\ \tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m) \neq 0] \\ \uparrow & \text{if } x \ge |\alpha_k^b| \text{ and } x \text{ is minimal such that} \\ & \tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m)\downarrow_{\le x} \text{ and } \tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m) = 0 \\ & \text{for some } m \in \{1,\ldots,x\} \\ \psi_0^b(x) & \text{otherwise} \end{cases}$$

If there is some $x \in \mathbb{N}$ with $x \ge |\alpha_k^b|$ and some $m \in \{1,\ldots,x\}$ such that

$$\tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m)\downarrow_{\le x} \text{ and } \tau_{S(b,\alpha_k^b 0^m)}(|\alpha_k^b| + m) = 0\ ,$$

let $m_k^b$ be the minimal integer such that

$$\tau_{S(b,\alpha_k^b 0^{m_k^b})}(|\alpha_k^b| + m_k^b)\downarrow_{\le x} \text{ and } \tau_{S(b,\alpha_k^b 0^{m_k^b})}(|\alpha_k^b| + m_k^b) = 0\ .$$

Then define $\alpha^b_{k+1} := \alpha^b_k 0^{m^b_k} 1$ and go to stage $k+1$. *End stage $k$.*

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction $b_0$.*

The following Fact is verified easily.

**Fact 1**

1. *If $\varphi^{b_0}_0 \in \mathcal{R}$ then*
   (a) $\alpha^{b_0}_k \subset \varphi^{b_0}_0$ *for all* $k \in \mathbb{N}$,
   (b) $\varphi^{b_0}_k \notin \mathcal{R}$ *and* $\varphi^{b_0}_k =^* \varphi^{b_0}_0$ *for all* $k \geq 1$;
2. *if $\varphi^{b_0}_0 \notin \mathcal{R}$ then there exists some $k \in \mathbb{N}$ such that $\mathcal{R}_{b_0} = \{\varphi^{b_0}_{k+1}\} = \{\alpha^{b_0}_k 0^\infty\}$,*
3. *if $\alpha^{b_0}_k$ is defined for some $k \in \mathbb{N}$, then card $\{i \in \mathbb{N} \mid \alpha^{b_0}_k(i) = 1\} = k$.*

In order to contradict phrase 3 we will prove the following statements.

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

*ad 1.* From Fact 1.1(b) and 1.2 we conclude that there is exactly one integer $k$ such that $\mathcal{R}_{b_0} = \{\varphi^{b_0}_k\}$. Thus $\mathcal{R}_{b_0} \in J^1$. For the proof of $\mathcal{R}_{b_0} \in \mathrm{BC}^*_{\varphi^{b_0}}(T)$ consider the following two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi^{b_0}_0\}$.
Since $\varphi^{b_0}_k =^* \varphi^{b_0}_0$ for all $k \geq 1$ by Fact 1.1(b), we obviously have $\varphi^{b_0}_{T(\varphi^{b_0}_0[n])} =^* \varphi^{b_0}_0$ for all $n \in \mathbb{N}$. Hence $\mathcal{R}_{b_0} \in \mathrm{BC}^*_{\varphi^{b_0}}(T)$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi^{b_0}_{k+1}\}$ for some $k \in \mathbb{N}$.
By Fact 1.2 and 1.3 we know that

$$\text{card } \{i \in \mathbb{N} \mid \varphi^{b_0}_{k+1}(i) = 1\} = \text{card } \{i \in \mathbb{N} \mid \alpha^{b_0}_k(i) = 1\} = k\ .$$

Now let $n_0 := \max\{i \in \mathbb{N} \mid \varphi^{b_0}_{k+1}(i) = 1\}$ be the maximal argument for which $\varphi^{b_0}_{k+1}$ returns 1. Then the definition of our strategy $T$ implies $T(\varphi^{b_0}_{k+1}[n]) = k+1$ for all $n \geq n_0$. Therefore $\mathcal{R}_{b_0} \in \mathrm{EX}_{\varphi^{b_0}}(T)$ and in particular $\mathcal{R}_{b_0} \in \mathrm{BC}^*_{\varphi^{b_0}}(T)$.

So we have verified $\mathcal{R}_{b_0} \in J^1$ as well as $\mathcal{R}_{b_0} \in \mathrm{BC}^*_{\varphi^{b_0}}(T)$. By definition this yields $b_0 \in B$. *qed 1.*

*ad 2.* Again we consider two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$.
From Fact 1.1(a) we know $\alpha_k^{b_0} \subset \varphi_0^{b_0}$ for all $k \in \mathbb{N}$. By construction in stage $k$, $k \in \mathbb{N}$, we obtain $\alpha_{k+1}^{b_0} = \alpha_k^{b_0} 0^m 1$ for some $m \in \mathbb{N}$ such that

$$\tau_{S(b_0, \alpha_k^{b_0} 0^m)}(|\alpha_k^{b_0}| + m)\downarrow_{\leq x} \text{ and } \tau_{S(b_0, \alpha_k^{b_0} 0^m)}(|\alpha_k^{b_0}| + m) = 0 \ .$$

So $\varphi_0^{b_0}(|\alpha_k^{b_0}| + m) = 1 \neq 0 = \tau_{S(b_0, \varphi_0^{b_0}[|\alpha_k^{b_0}| + m - 1])}(|\alpha_k^{b_0}| + m)$, i.e. the hypothesis returned by $\lambda x.S(b_0, x)$ on $\varphi_0^{b_0}[|\alpha_k^{b_0}| + m - 1]$ is incorrect with respect to $\tau$ for all $k \in \mathbb{N}$. So our strategy infinitely often returns incorrect hypotheses on the only function in $\mathcal{R}_{b_0}$. Therefore $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi_{k+1}^{b_0}\}$ for some $k \in \mathbb{N}$.
By construction in stage $k$ we observe that $\varphi_{k+1}^{b_0}$ is a total function if and only if $\tau_{S(b_0, \alpha_k^{b_0} 0^m)}(|\alpha_k^{b_0}| + m)\uparrow$ or $\tau_{S(b_0, \alpha_k^{b_0} 0^m)}(|\alpha_k^{b_0}| + m) \neq 0 = \varphi_{k+1}^{b_0}(|\alpha_k^{b_0}| + m)$ for all $m \in \mathbb{N}$. Since $\varphi_{k+1}^{b_0} = \alpha_k^{b_0} 0^\infty$, this implies that all but finitely many hypotheses returned by the strategy $\lambda x.S(b_0, x)$ on $\varphi_{k+1}^{b_0}$ are incorrect with respect to $\tau$. Hence $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$.

In both cases we have verified 2. *qed 2.*

The claims 1 and 2 now imply the existence of some integer $b_0 \in B$ such that $\mathcal{R}_{b_0} \notin \mathrm{BC}_\tau(\lambda x.S(b_0, x))$. This contradicts our assumption and we conclude $B \notin \mathrm{suit}_\tau(J^1, \mathrm{BC})$. Therefore $\mathrm{suit}_\varphi(J^1, \mathrm{BC}^*) \setminus \mathrm{suit}_\tau(J^1, \mathrm{BC}) \neq \emptyset$. □

With similar methods we can also separate explanatory identification from behaviourally correct identification.

**Theorem 8** *$suit_\varphi(J^*, BC) \setminus suit(J^*, EX) \neq \emptyset$.*

*Proof.* We use a strategy $T \in \mathcal{R}$ to define a description set $B \subseteq \mathbb{N}$ suitable for uniform behaviourally correct identification by $T$. The set $B$ shall describe only finite recursive cores and will not be suitable for uniform identification in the limit.

*Definition of $B \in suit_\varphi(J^*, BC) \setminus suit(J^*, EX)$.*
Define a strategy $T \in \mathcal{P}$ by

$$T(f[n]) := \begin{cases} 0 & \text{if } n = 0 \\ 2j+1 & \text{if } n > 0 \ \wedge \ f(n) \neq 2 \ \wedge \ \text{card}\ \{i \leq n \mid f(i) = 2\} = j \\ & \ \wedge \ \forall x \in \{\max\{i \leq n \mid f(i) = 2\}, \ldots, n\} \ [f(x) = 0] \\ 2j+2 & \text{if } n > 0 \ \wedge \ f(n) \neq 2 \ \wedge \ \text{card}\ \{i \leq n \mid f(i) = 2\} = j \\ & \ \wedge \ \exists x \in \{\max\{i \leq n \mid f(i) = 2\}, \ldots, n\} \ [f(x) \neq 0] \\ T(f[n-1]) & \text{if } n > 0 \ \wedge \ f(n) = 2 \end{cases}$$

for arbitrary $f \in \mathcal{R}$ and $n \in \mathbb{N}$. Then set

$$B := \{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite and } \mathcal{R}_b \in \text{BC}_{\varphi^b}(T)\} \ .$$

*Claim.* $B \in \text{suit}_\varphi(J^*, \text{BC}) \setminus \text{suit}(J^*, \text{EX})$.

*Proof of "$B \in suit_\varphi(J^*, BC)$".*
Defining $T' \in \mathcal{P}^2$ by $T'(b, f[n]) := T(f[n])$ for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ we obviously have

- $\forall b \in B \ [\mathcal{R}_b \in J^*]$,
- $\forall b \in B \ [\mathcal{R}_b \in \text{BC}_{\varphi^b}(\lambda x.T'(b, x))]$.

By Definition 12 this implies $B \in \text{suit}_\varphi(J^*, \text{BC})$.

*Proof of "$B \notin suit(J^*, EX)$".*
We will verify this claim by way of contradiction.
*Assumption.* $B \in \text{suit}(J^*, \text{EX})$.
Then there exists a strategy $S \in \mathcal{R}^2$ such that

$$\forall b \in B \ [\mathcal{R}_b \in \text{EX}(\lambda x.S(b, x))] \ . \tag{4}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \text{EX}(\lambda x.S(b_0, x))$,

in contradiction to statement (4). The strategy $\lambda x.S(b_0, x)$ will fail for at least one function $f \in \mathcal{R}_{b_0}$ by either

- changing its hypothesis for $f$ infinitely often or
- converging to an index incorrect for the function $f$.

*Construction of $b_0$.*
Define a function $\psi \in \mathcal{P}^3$ by the following instructions: let $b \in \mathbb{N}$, $\alpha_0^b(0) := 0$. Go to stage 0.

*Stage 0.* By dove-tailed computation look for the minimal integer $m_0^b$ satisfying

$$S(b, 000^{m_0^b}) \neq S(b,0) \text{ or } S(b, 010^{m_0^b}) \neq S(b,0) \ .$$

Then let

$$\alpha_1^b := \begin{cases} 000^{m_0^b}2 & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) \neq S(b,0) \\ 010^{m_0^b}2 & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) = S(b,0) \\ \uparrow & \text{if } m_0^b \text{ is not defined} \end{cases}$$

Furthermore set

$$\begin{aligned} \psi_0^b[m_0^b+2] &:= \alpha_1^b \text{ (i.e. } \psi_0^b = \uparrow^\infty \text{ if } m_0^b \text{ is undefined)} \\ \psi_1^b &:= \begin{cases} 000^\infty & \text{if } m_0^b \text{ is not defined} \\ \psi_0^b & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) \neq S(b,0) \\ 000^{m_0^b}\uparrow^\infty & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) = S(b,0) \end{cases} \\ \psi_2^b &:= \begin{cases} 010^\infty & \text{if } m_0^b \text{ is not defined} \\ \psi_0^b & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) = S(b,0) \\ 010^{m_0^b}\uparrow^\infty & \text{if } m_0^b \text{ is defined and } S(b,000^{m_0^b}) \neq S(b,0) \end{cases} \end{aligned}$$

If $m_0^b$ is defined, then go to stage 1. *End stage 0.*

*Stage k ($k \in \mathbb{N}, k > 0$).* By dove-tailed computation look for the minimal integer $m_k^b$ satisfying

$$S(b, \alpha_k^b 00^{m_k^b}) \neq S(b,\alpha_k^b) \text{ or } S(b, \alpha_k^b 10^{m_k^b}) \neq S(b,\alpha_k^b) \ .$$

Then let

$$\alpha_{k+1}^b := \begin{cases} \alpha_k^b 00^{m_k^b}2 & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_k^b}) \neq S(b,\alpha_k^b) \\ \alpha_k^b 10^{m_k^b}2 & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_k^b}) = S(b,\alpha_k^b) \\ \uparrow & \text{if } m_k^b \text{ is not defined} \end{cases}$$

Furthermore set

$$\begin{aligned} \psi_0^b[|\alpha_{k+1}^b|-1] &:= \alpha_{k+1}^b \text{ if } m_k^b \text{ is defined } (\psi_0^b = \alpha_k^b \uparrow^\infty \text{ if } m_k^b \text{ is undefined)} \\ \psi_{2k+1}^b &:= \begin{cases} \alpha_k^b 00^\infty & \text{if } m_k^b \text{ is not defined} \\ \psi_0^b & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_k^b}) \neq S(b,\alpha_k^b) \\ \alpha_k^b 00^{m_k^b}\uparrow^\infty & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_k^b}) = S(b,\alpha_k^b) \end{cases} \\ \psi_{2k+2}^b &:= \begin{cases} \alpha_k^b 10^\infty & \text{if } m_k^b \text{ is not defined} \\ \psi_0^b & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_0^b}) = S(b,\alpha_k^b) \\ \alpha_k^b 10^{m_k^b}\uparrow^\infty & \text{if } m_k^b \text{ is defined and } S(b,\alpha_k^b 00^{m_k^b}) \neq S(b,\alpha_k^b) \end{cases} \end{aligned}$$

If $m_k^b$ is defined, then go to stage $k+1$. *End stage $k+1$.*

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction $b_0$.*

The following Fact is verified easily.

**Fact 2**

1. *$\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$ iff $m_k^{b_0}$ is defined for all $k \in \mathbb{N}$ iff $\{\varphi_{2k+1}^{b_0}, \varphi_{2k+2}^{b_0}\} = \{\varphi_0^{b_0}, \alpha_k^{b_0}\overline{\varphi_0^{b_0}(|\alpha_k^{b_0}|)}0^{m_k^{b_0}}\uparrow^\infty\}$ for all $k \in \mathbb{N}$;*
2. *$\mathcal{R}_{b_0} = \{\varphi_{2k+1}^{b_0}, \varphi_{2k+2}^{b_0}\} = \{\alpha_k^{b_0}00^\infty, \alpha_k^{b_0}10^\infty\}$ iff $k$ is the minimal integer such that $m_k^{b_0}$ is undefined;*
3. *if $\alpha_k^{b_0}$ is defined for some $k \in \mathbb{N}$, then card $\{i \in \mathbb{N} \mid \alpha_k^{b_0}(i) = 2\} = k$.*

In order to contradict phrase (4) we will prove the following statements.

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0, x))$.

*ad 1.* By Fact 2.1 and Fact 2.2 $\mathcal{R}_{b_0}$ is finite. Thus it remains to prove $\mathcal{R}_{b_0} \in \mathrm{BC}_{\varphi^{b_0}}(T)$. It suffices to consider the following two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$.
We know $T(\varphi_0^{b_0}[0]) = 0$.
If $n \geq 1$ and $\varphi_0^{b_0}(n) \neq 2$, there is some $k \in \mathbb{N}$ maximal with the property $\alpha_k^{b_0} \subseteq \varphi_0^{b_0}[n-1]$. Then $t := \varphi_0^{b_0}(|\alpha_k|) \in \{0,1\}$. Since card $\{i \leq n \mid \varphi_0^{b_0}(i) = 2\} =$ card $\{i \in \mathbb{N} \mid \alpha_k^{b_0}(i) = 2\} = k$ by Fact 2.3, we obtain $T(\varphi_0^{b_0}[n]) = 2k+t+1$. Since $\alpha_k^{b_0}t \subseteq \varphi_{2k+t+1}^{b_0}$ and $\alpha_k^{b_0}t \subseteq \varphi_0^{b_0}$ and $\{\varphi_{2k+1}^{b_0}, \varphi_{2k+2}^{b_0}\} = \{\varphi_0^{b_0}, \alpha_k^{b_0}\bar{t}0^{m_k^{b_0}}\uparrow^\infty\}$ by Fact 2.1, we have $\varphi_{2k+t+1}^{b_0} = \varphi_0^{b_0}$. So for any $n \geq 1$ with $\varphi_0^{b_0}(n) \neq 2$ we obtain

$$\varphi_{T(\varphi_0^{b_0}[n])}^{b_0} = \varphi_0^{b_0} \ . \tag{5}$$

Now if $n \geq 2$ and $\varphi_0^{b_0}(n) = 2$, the definition of $T$ implies

$$\varphi_{T(\varphi_0^{b_0}[n])}^{b_0} = \varphi_{T(\varphi_0^{b_0}[n-1])}^{b_0} \overset{(5)}{=} \varphi_0^{b_0} \ .$$

To verify this, please note that $\varphi_0^{b_0}(n-1) \neq 2$ by construction. Thus $\varphi_{T(\varphi_0^{b_0}[n])}^{b_0} = \varphi_0^{b_0}$ for all but finitely many $n \in \mathbb{N}$ and therefore $\mathcal{R}_{b_0} \in \mathrm{BC}_{\varphi^{b_0}}(T)$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi^{b_0}_{2k+1}, \varphi^{b_0}_{2k+2}\} = \{\alpha^{b_0}_k 00^\infty, \alpha^{b_0}_k 10^\infty\}$ for some $k \in \mathbb{N}$.
Assume $n > |\alpha^{b_0}_k|$. Then $\varphi^{b_0}_{2k+1}(n) \neq 2$ and $\varphi^{b_0}_{2k+2}(n) \neq 2$ by construction. Furthermore

$$\begin{aligned} \text{card}\ \{i \leq n \mid \varphi^{b_0}_{2k+1}(i) = 2\} &= \text{card}\ \{i \leq n \mid \varphi^{b_0}_{2k+2}(i) = 2\} \\ &= \text{card}\ \{i \in \mathbb{N} \mid \alpha^{b_0}_k(i) = 2\} \\ &= k \end{aligned}$$

because of Fact 2.3. Since $\varphi^{b_0}_{2k+1}(|\alpha^{b_0}_k|) = 0$ and $\varphi^{b_0}_{2k+2}(|\alpha^{b_0}_k|) = 1$, this yields $T(\varphi^{b_0}_{2k+1}[n]) = 2k+1$ and $T(\varphi^{b_0}_{2k+2}[n]) = 2k+2$. Therefore $\varphi^{b_0}_{T(\varphi^{b_0}_{2k+1}[n])} = \varphi^{b_0}_{2k+1}$ and $\varphi^{b_0}_{T(\varphi^{b_0}_{2k+2}[n])} = \varphi^{b_0}_{2k+2}$ for all but finitely many $n \in \mathbb{N}$. Hence $\mathcal{R}_{b_0} \in \text{BC}_{\varphi^{b_0}}(T)$.

Altogether we obtain $b_0 \in B$. *qed 1.*

*ad 2.* Again we consider two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi^{b_0}_0\}$.
Then $m^{b_0}_k$ is defined for all $k \in \mathbb{N}$ by Fact 2.1. By construction this implies $\alpha^{b_0}_k \subseteq \varphi^{b_0}_0$ as well as $S(b_0, \alpha^{b_0}_k) \neq S(b_0, \varphi^{b_0}_0[|\alpha^{b_0}_{k+1}| - 2])$ for all $k \in \mathbb{N}$. Thus the sequence of hypotheses produced by the strategy $\lambda x.S(b_0, x)$ on $\varphi^{b_0}_0$ does not converge. This implies $\mathcal{R}_{b_0} \notin \text{EX}(\lambda x.S(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi^{b_0}_{2k+1}, \varphi^{b_0}_{2k+2}\} = \{\alpha^{b_0}_k 00^\infty, \alpha^{b_0}_k 10^\infty\}$ for some $k \in \mathbb{N}$.
Then by Fact 2.2 we know that $m^{b_0}_k$ is not defined. This implies

$$S(b_0, \alpha^{b_0}_k 00^m) = S(b_0, \alpha^{b_0}_k) = S(b_0, \alpha^{b_0}_k 10^m) \text{ for all } k \in \mathbb{N}\ .$$

Thus $S(b_0, \varphi^{b_0}_{2k+1}[n]) = S(b_0, \varphi^{b_0}_{2k+2}[n])$ for all but finitely many $n \in \mathbb{N}$, although $\varphi^{b_0}_{2k+1} \neq \varphi^{b_0}_{2k+2}$. Therefore – no matter what hypothesis space is regarded – the sequence of hypotheses generated by $\lambda x.S(b_0, x)$ must converge to an incorrect hypothesis for at least one of the two functions in $\mathcal{R}_{b_0}$. Hence $\mathcal{R}_{b_0} \notin \text{EX}(\lambda x.S(b_0, x))$.

*qed 2.*

Claims 1 and 2 together now contradict statement 4. Therefore our assumption must have been wrong, i.e. $B \notin \text{suit}(J^*, \text{EX})$. Altogether we have $B \in \text{suit}_\varphi(J^*, \text{BC}) \setminus \text{suit}(J^*, \text{EX})$. This completes the proof. □

Note that this result is even much stronger than required. We just needed to prove $\text{suit}_\varphi(J^*, \text{BC}) \setminus \text{suit}_\varphi(J^*, \text{EX}) \neq \emptyset$ and the corresponding statement for learning with respect to $\tau$. Besides we have not only verified

$$\text{suit}(J^*, \text{BC}) \setminus \text{suit}(J^*, \text{EX}) \neq \emptyset\ ,$$

but we observe a further fact: though we know uniform learning with respect to the hypothesis spaces given by $\varphi$ to be much more restrictive than uniform learning without special demands concerning the hypothesis spaces, we still can find collections of class-descriptions which are

- restrictive enough to describe finite classes of recursive functions only,
- suitable for uniform BC-identification with respect to the hypothesis spaces corresponding to their descriptions,
- but *not* suitable for uniform EX-identification even if the hypothesis spaces can be chosen without restrictions.

Similar strict separations are obtained by the theorems below.

In the following corollary we summarize parts of the results obtained by Theorem 7 and Theorem 8.

**Corollary 1** *$suit_\varphi(J^*, EX) \subset suit_\varphi(J^*, BC) \subset suit_\varphi(J^*, BC^*)$. This result holds also for $suit_\tau$ instead of $suit_\varphi$.*

So we have verified that all inclusions in Lemma 1.1 are in fact proper inclusions.

## 4.2 The Hierarchy in Lemma 1.2

With the help of the techniques used in the proofs of Theorem 7 and Theorem 8 we will now also show that Lemma 1.2 can be written with strict inclusions. For that purpose the following theorem is sufficient.

**Theorem 9** *$suit_\varphi(J^*, EX_{m+1}) \setminus suit(J^*, EX_m) \neq \emptyset$ for any $m \in \mathbb{N}$.*

*Proof.* Fix $m \in \mathbb{N}$. Define a description set $B \subseteq \mathbb{N}$ by

$$B := \{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite and } \varphi^b \in \mathcal{R}^2 \text{ and } \mathcal{R}_b \in (\mathrm{EX}_{m+1})_{\varphi^b}(\mathrm{Enum}_{\varphi^b})\}.$$

We want to prove: $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_{m+1}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m)$.

*Proof of "$B \in suit_\varphi(J^*, EX_{m+1})$".*
For $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ define

$$T(b, f[n]) := \begin{cases} \min\{i \mid \varphi_i^b =_n f\} & \text{if the necessary consistency tests are} \\ & \text{decidable and such a minimum exists} \\ \uparrow & \text{otherwise} \end{cases}$$

Thus $T \in \mathcal{P}^2$. Since $\varphi^b \in \mathcal{R}^2$ for all $b \in B$ we obtain with the definition of $B$:

- $\forall b \in B\ [\mathcal{R}_b \in J^*]$,
- $\forall b \in B\ [\mathcal{R}_b \in (\mathrm{EX}_{m+1})_{\varphi^b}(\lambda x.T(b,x))]$.

By definition of uniform identifiability this implies $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_{m+1})$.

*Proof of "$B \notin suit(J^*, EX_m)$".*

We will verify this claim by way of contradiction.
*Assumption.* $B \in \mathrm{suit}(J^*, \mathrm{EX}_m)$.

Then by Proposition 3 there exists a strategy $S \in \mathcal{R}^2$ such that for any $b \in B$ there is a numbering $\eta \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in (\mathrm{EX}_m)_\eta(\lambda x.S(b,x))$; abbreviated

$$\forall b \in B\ [\mathcal{R}_b \in \mathrm{EX}_m(\lambda x.S(b,x))]\ . \tag{6}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{EX}_m(\lambda x.S(b_0,x))$,

in contradiction to statement (6). The strategy $\lambda x.S(b_0,x)$ will fail for at least one function $f \in \mathcal{R}_{b_0}$ by either

- returning "?" for all initial segments of $f$ or
- generating a hypothesis incorrect for $f$ for infinitely many initial segments of $f$ or
- changing its hypothesis on $f$ at least $m+1$ times.

In order to deduce a contradiction it suffices to prove that one of these three cases must occur. Which of these cases actually occurs for the constructed integer $b_0$, does not have to be decidable.

*Construction of* $b_0$.

Generate a finite list of recursive functions according to the following instructions. Begin in stage 0 with an empty list.

*Stage 0.* Add the function $0^\infty$ to the list. Go to stage 1.

*Stage 1.* Add the function

$$x \mapsto \begin{cases} 0 & \text{if } x = 0 \text{ or } S(b, 0^{x-1}) = ? \\ 1 & \text{otherwise} \end{cases} \quad \text{for } x \in \mathbb{N}$$

(i.e. either the function $0^\infty$ or a function $0^k1^\infty$ for some $k \in \mathbb{N}$) to the list. Go to stage 2.

*Stage* $j+1$ $(1 \leq j \leq m)$. For every function $f$ in the list add the function

$$x \mapsto \begin{cases} f(x) & \text{if } x = 0 \text{ or} \\ & \text{card } \{n < x-1 \mid ? \neq S(b, f[n]) \neq S(b, f[n+1])\} < j \\ j+1 & \text{otherwise} \end{cases}$$

for $x \in \mathbb{N}$ (i.e. either the function $f$ or a function $f[k](j+1)^\infty$ for some $k \in \mathbb{N}$) to the list. Go to stage j+2.

*Stage* $m+2$. Stop.

Let $\psi^b$ be the effective numbering which first enumerates all functions in the list generated above and afterwards continues by listing the function $0^\infty$ forever. That means, $\psi^b_0 = 0^\infty$; $\psi^b_1$ is the function listed in stage 1; $\psi^b_2$, $\psi^b_3$ are the functions listed in stage 2; $\ldots$; $\psi^b_{2^m}, \psi^b_{2^m+1}, \ldots, \psi^b_{2^{m+1}-1}$ are the functions listed in stage $m+1$; furthermore $\psi^b_i = 0^\infty$ for all $i \geq 2^{m+1}$.

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction* $b_0$.

In order to contradict phrase (6) we will prove the following statements.

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{EX}_m(\lambda x.S(b_0, x))$.

*ad 1.* By construction $\mathcal{R}_{b_0}$ is finite and $\varphi^{b_0} \in \mathcal{R}^2$. It remains to prove that $\mathcal{R}_{b_0} \in (\mathrm{EX}_{m+1})_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$. For that purpose choose $f \in \mathcal{R}_{b_0}$ arbitrarily. We prove the following claim:

$$\text{If } f \text{ is listed in stage } t \ (0 \leq t \leq m+1), \text{ then } \{f\} \in (\mathrm{EX}_t)_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}}). \quad (7)$$

We use induction on $t$.

$t = 0$. If $f$ is listed in stage 0, then $f = 0^\infty = \varphi^{b_0}_0$. Obviously $\mathrm{Enum}_{\varphi^{b_0}}(f[n]) = 0$ for all $n \in \mathbb{N}$. Thus $\{f\} \in (\mathrm{EX}_0)_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$.

$t \rightsquigarrow t+1$. If $f$ is listed in stage $t+1$, we consider two possible cases:

*Case (i).* $f$ has already been listed in stage $t'$ for some $t' < t+1$.
Then by induction hypothesis $\{f\} \in (\mathrm{EX}_{t'})_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$ and in particular $\{f\} \in (\mathrm{EX}_{t+1})_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$.

*Case (ii).* $f = f'[k](t+1)^\infty$ for some $f'$ listed in or before stage $t$ and some $k \in \mathbb{N}$.
By induction hypothesis

$$\mathrm{card}\,\{n < k \mid ? \neq \mathrm{Enum}_{\varphi^{b_0}}(f'[n]) \neq \mathrm{Enum}_{\varphi^{b_0}}(f'[n+1])\} \leq t\ .$$

Let $i := \mathrm{Enum}_{\varphi^{b_0}}(f'[k])$. As $f'$ is listed before Stage $t+1$, we have $i < 2^t$. Since $f(k+1) = t+1$, this implies $\varphi_i^{b_0} \neq_{k+1} f$. Therefore $\mathrm{Enum}_{\varphi^{b_0}}(f[k+1]) \in \{2^t, \ldots, 2^{t+1}-1\}$ and

$$\mathrm{card}\,\{n < k+1 \mid ? \neq \mathrm{Enum}_{\varphi^{b_0}}(f[n]) \neq \mathrm{Enum}_{\varphi^{b_0}}(f[n+1])\} \leq t+1\ . \qquad (8)$$

If $j \in \{2^t, \ldots, 2^{t+1}-1\}$ and $f'[k](t+1) \subset \varphi_j^{b_0}$, we obtain $\varphi_j^{b_0} = f'[k](t+1)^\infty = f$ by construction. Thus $\mathrm{Enum}_{\varphi^{b_0}}(f[n]) = \mathrm{Enum}_{\varphi^{b_0}}(f[k+1])$ for all $n \geq k+1$ and $\mathrm{Enum}_{\varphi^{b_0}}(f[k+1])$ is a $\varphi^{b_0}$-number for $f$. With (8) we conclude $\{f\} \in (\mathrm{EX}_{t+1})_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$.
This proves (7).

As our construction stops after $m+1$ stages, we obtain with (7):

$$\mathcal{R}_{b_0} \in (\mathrm{EX}_{m+1})_{\varphi^{b_0}}(\mathrm{Enum}_{\varphi^{b_0}})$$

and therefore $b_0 \in B$. *qed 1.*

*ad 2.* Let stage $t$ ($t \leq m+1$) be the last stage in which at least one *new* function is listed. We consider three cases.

*Case (i).* $t = 0$.
Then $\mathcal{R}_{b_0} = \{0^\infty\}$. By construction in stage 1 we know that $S(b_0, 0^k) =?$ for all $k \in \mathbb{N}$. Thus $\mathcal{R}_{b_0} \notin \mathrm{EX}_m(\lambda x.S(b_0, x))$.

*Case (ii).* $1 \leq t \leq m$.
Then there exists an $f' \in \mathcal{R}_{b_0}$ listed in or before stage $t-1$ and an $f \in \mathcal{R}_{b_0}$ listed in stage $t$ such that

$$f = f'[k]t^\infty \neq f'$$

for some $k \in \mathbb{N}$. Since $f \neq f'$, we have $\varphi^{b_0}_{S(b_0,f'[k])} \neq f'$ or $\varphi^{b_0}_{S(b_0,f[k])} \neq f$. But as there is no new function listed in stage $t+1$, we obtain

$$S(b_0, f'[n]) = S(b_0, f'[k]) = S(b_0, f[k]) = S(b_0, f[n])$$

for all $n \geq k$. Therefore on $f$ or on $f'$ the strategy $\lambda x.S(b_0, x)$ must produce a sequence of hypotheses converging to an incorrect index. This implies $\mathcal{R}_{b_0} \notin \mathrm{EX}_m(\lambda x.S(b_0, x))$.

*Case (iii).* $t = m + 1$.
Then there exists an $f' \in \mathcal{R}_{b_0}$ listed in stage $m$ and an $f \in \mathcal{R}_{b_0}$ listed in stage $m + 1$ such that

$$f = f'[k](m+1)^\infty \neq f'$$

for some $k \in \mathbb{N}$ with card $\{n < k \mid ? \neq S(b_0, f'[n]) \neq S(b_0, f'[n+1])\} = m$. Since $f \neq f'$, we have $\varphi^{b_0}_{S(b_0, f'[k])} \neq f'$ or $\varphi^{b_0}_{S(b_0, f[k])} \neq f$. Therefore on $f$ or on $f'$ the strategy $\lambda x.S(b_0, x)$ must produce a sequence of hypotheses converging to an incorrect index *or* change its hypothesis at least $m + 1$ times. This implies $\mathcal{R}_{b_0} \notin \mathrm{EX}_m(\lambda x.S(b_0, x))$.

Thus we have verified 2. *qed 2.*

Properties 1 and 2 together now imply that there is some $b \in B$ satisfying $\mathcal{R}_b \notin \mathrm{EX}_m(\lambda x.S(b, x))$ in contradiction to (6). Therefore our assumption must have been wrong and we conclude $B \notin \mathrm{suit}(J^*, \mathrm{EX}_m)$.

Altogether we obtain $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_{m+1}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m)$. Since $m$ was chosen arbitrarily, this proves Theorem 9. □

As the set

$$B := \{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite and } \varphi^b \in \mathcal{R}^2 \text{ and } \mathcal{R}_b \in (\mathrm{EX}_{m+1})_{\varphi^b}(\mathrm{Enum}_{\varphi^b})\}$$

defined in the previous proof (for some $m \in \mathbb{N}$) is suitable for uniform class-preserving identification with respect to the numberings $\varphi^b$ ($b \in B$), we obtain

$$B \in \mathrm{suit}_\varphi(J^*, \mathrm{CP}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m) \ .$$

This leads us to the following corollary.

**Corollary 2** *$suit_\varphi(J^*, CP) \setminus suit(J^*, EX_m) \neq \emptyset$ for all $m \in \mathbb{N}$.*

In order to complete a strict version of the hierarchy in Lemma 1.2, note that $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_m) \subset \mathrm{suit}_\varphi(J^*, \mathrm{EX})$ for all $m \in \mathbb{N}$, because $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_m) \subset \mathrm{suit}_\varphi(J^*, \mathrm{EX}_{m+1}) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{EX})$. Of course a similar argumentation implies strict inclusions in the $\tau$-case.

**Corollary 3** *$suit_\varphi(J^*, EX_m) \subset suit_\varphi(J^*, EX_{m+1}) \subset suit_\varphi(J^*, EX)$ for all $m \in \mathbb{N}$ (analogously with $suit_\tau$ instead of $suit_\varphi$).*

## 4.3 The Hierarchy in Lemma 1.3

This subsection is concerned with the investigation of the uniform learning hierarchy given by the criteria CP, TOTAL, CONS, CONF and EX. First we want to prove, that $\mathrm{suit}_\varphi(J^*, \mathrm{CONF})$ is indeed a proper subset of $\mathrm{suit}_\varphi(J^*, \mathrm{EX})$. This will be achieved by proving a much stronger result stated in Theorem 10.

**Theorem 10**

1. $suit_\varphi(J^*, EX_1) \setminus suit(J^*, CONF) \neq \emptyset$,
2. $suit_\varphi(J^*, CEX) \setminus suit(J^*, CONF) \neq \emptyset$.

*Proof.* We use a strategy $T \in \mathcal{R}$ to define a description set $B \subseteq \mathbb{N}$ suitable for uniform identification in the limit by $T$. $T$ will change its mind on the relevant functions at most once and all intermediate hypotheses produced will be correct or convergently incorrect with respect to the given numberings $\varphi^b,\ b \in B$. The set $B$ shall describe only finite recursive cores and will not be suitable for uniform conform identification.

*Definition of* $B \in (suit_\varphi(J^*, EX_1) \cap suit_\varphi(J^*, CEX)) \setminus suit(J^*, CONF)$.
Define a strategy $T \in \mathcal{R}$ by

$$T(f[n]) := \begin{cases} 0 & \text{if } f[n] \in \{0,1\}^* \\ \max\{f(0), \dots, f(n)\} - 1 & \text{otherwise} \end{cases}$$

for arbitrary $f \in \mathcal{R}$ and $n \in \mathbb{N}$. Then set

$$B := \{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite and } \mathcal{R}_b \in (\mathrm{EX}_1)_{\varphi^b}(T) \cap \mathrm{CEX}_{\varphi^b}(T)\}\ .$$

*Claim.* $B \in (\mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \cap \mathrm{suit}_\varphi(J^*, \mathrm{CEX})) \setminus \mathrm{suit}(J^*, \mathrm{CONF})$.

*Proof of "$B \in suit_\varphi(J^*, EX) \cap suit_\varphi(J^*, CEX)$".*
Defining $T' \in \mathcal{R}^2$ by $T'(b, f[n]) := T(f[n])$ for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ we obviously have

- $\forall b \in B\ [\mathcal{R}_b \in J^*]$,
- $\forall b \in B\ [\mathcal{R}_b \in (\mathrm{EX}_1)_{\varphi^b}(\lambda x.T'(b,x))$ and $\mathcal{R}_b \in \mathrm{CEX}_{\varphi^b}(\lambda x.T'(b,x))]$.

By Definition 12 this implies $B \in \mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \cap \mathrm{suit}_\varphi(J^*, \mathrm{CEX})$.

*Proof of "$B \notin suit(J^*, CONF)$".*
We will verify this claim by way of contradiction.
*Assumption.* $B \in \mathrm{suit}(J^*, \mathrm{CONF})$.

Then there exists a strategy $S \in \mathcal{P}^2$ such that for any $b \in B$ there is a hypothesis space $\psi \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in \mathrm{CONF}_\psi(\lambda x.S(b,x))$; abbreviated

$$\forall b \in B\ [\mathcal{R}_b \in \mathrm{CONF}(\lambda x.S(b,x))]\ . \tag{9}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CONF}(\lambda x.S(b_0,x))$,

in contradiction to statement (9). The strategy $\lambda x.S(b_0,x)$ will fail for at least one function $f \in \mathcal{R}_{b_0}$ by either

- changing its hypothesis for $f$ infinitely often or
- not terminating its computation on input of some initial segment of $f$ or
- violating the conformity demand on input of some initial segment of $f$.

*Construction of* $b_0$.
Define a function $\eta \in \mathcal{P}^3$ by the following instructions: for $b \in \mathbb{N}$ set $\eta_0^b(0) := 0$ and go to stage 0.

*Stage 0.*
For the definition of further values of $\eta_0^b$ compute $S(b,0)$, $S(b,00)$, $S(b,01)$. If these values are all equal, we append zeros until we observe that the strategy $\lambda x.S(b,x)$ changes its mind on the initial segment constructed so far. Otherwise we just append one value $t \in \{0,1\}$, such that $S(b,0) \neq S(b,0t)$. Anyway we define $\eta_0^b(2) := 0$, which will prevent the learner $T$ from returning hypotheses corresponding to proper subfunctions of the functions to be learned, that means identification with convergently incorrect intermediate hypotheses is enabled. Formally:

$$\begin{aligned}
\eta_0^b(1) &:= \begin{cases} 0 & \text{if } S(b,0)\downarrow\ \wedge S(b,00)\downarrow\ \wedge S(b,01)\downarrow \text{ and} \\ & [S(b,0) \neq S(b,00) \text{ or } S(b,0) = S(b,00) = S(b,01)] \\ 1 & S(b,0)\downarrow\ \wedge S(b,00)\downarrow\ \wedge S(b,01)\downarrow \text{ and} \\ & S(b,0) = S(b,00) \text{ and } S(b,0) \neq S(b,01) \\ \uparrow & \text{otherwise} \end{cases} \\
\eta_0^b(2) &:= 0
\end{aligned}$$

For any $x > 0$ define

$$\eta_0^b(x+2) := \begin{cases} 0 & \text{if } S(b,0)\downarrow \ \wedge S(b,00)\downarrow \ \wedge S(b,01)\downarrow \ \text{ and} \\ & S(b,0) = S(b,00) = S(b,01) \text{ and} \\ & \forall z \leq x \ [S(b,\eta_0^b[z])\uparrow_{\leq x} \ \vee S(b,\eta_0^b[z]) = S(b,\eta_0^b)] \\ \uparrow & \text{if } S(b,0)\uparrow \ \vee S(b,00)\uparrow \ \vee S(b,01)\uparrow \\ \text{temporarily} & \text{otherwise} \\ \text{ suspended} & \end{cases}$$

The functions $\eta_1^b$ and $\eta_2^b$ are defined as follows: $\eta_1^b[2] := 002$, $\eta_2^b[2] := 013$. Both functions will be extended by zeros until the values $S(b,0)$, $S(b,00)$, $S(b,01)$ are computed and the definition of $\eta_0^b$ is temporarily suspended (if these conditions are never satisfied, we obtain $\eta_1^b = 0020^\infty$ and $\eta_2^b = 0130^\infty$). Formally we define for $x > 0$:

$$\eta_1^b(x+2) := \eta_2^b(x+2) := \begin{cases} 0 & \text{if } [S(b,0)\uparrow_{\leq x} \ \vee S(b,00)\uparrow_{\leq x} \ \vee S(b,01)\uparrow_{\leq x}] \text{ or} \\ & [S(b,0) = S(b,00) = S(b,01) \text{ and} \\ & \forall z \leq x \ [S(b,\eta_0^b[z])\uparrow_{\leq x} \ \vee S(b,\eta_0^b[z]) = S(b,\eta_0^b)]] \\ \uparrow & \text{otherwise} \end{cases}$$

Compute the number

$$y_1 := \begin{cases} \min\{n \in \mathbb{N} \mid \text{the definition of } \eta_0^b(n+1) & \text{if such a minimum exists} \\ \qquad \text{is temporarily suspended in stage } 0\} & \\ \uparrow & \text{otherwise} \end{cases}$$

If $y_1$ is defined, then go to stage 1. *End stage 0.*

From the construction in stage 0 we observe the following Fact.

**Fact 3**

1. *If stage 1 is reached, then*
   - $y_1 > 0$ *and* $S(b,\eta_0^b[n]) \neq S(b,\eta_0^b[0])$ *for some* $n \in \{1,\ldots,y_1\}$ *(i.e.* $\lambda x.S(b,x)$ *is forced into a mind change),*
   - $\eta_1^b = 0020^m\uparrow^\infty$ *and* $\eta_2^b = 0130^m\uparrow^\infty$ *for some* $m \in \mathbb{N}$*;*
2. *if stage 1 is not reached, then*
   - $\eta_1^b = 0020^\infty$ *and* $\eta_2^b = 0130^\infty$*,*
   - $\eta_0^b \in \{0\uparrow 0\uparrow^\infty, 0^\infty\}$*; in particular*

– $\eta_0^b = 0\uparrow 0\uparrow^\infty$ *iff* $S(b,0)\uparrow\ \vee S(b,00)\uparrow\ \vee S(b,01)\uparrow$,
– $\eta_0^b = 0^\infty$ *iff* $S(b,0) = S(b,00) = S(b,01) = S(b,0^m)$ $(\in \mathbb{N})$ *for any* $m \in \mathbb{N}$;

*3.* $\eta_0^b(2)$ *is defined and* $\eta_0^b(2) \neq \eta_1^b(2),\ \eta_0^b(2) \neq \eta_2^b(2)$.

Now we formulate the instructions for stages $k, k \geq 1$.

*Stage* $k$ *for* $k \geq 1$.

For the definition of further values of $\eta_0^b$ compute $S(b,\eta_0^b[y_k]),\ S(b,\eta_0^b[y_k]0)$, $S(b,\eta_0^b[y_k]1)$. If these values are all equal, we append zeros until we observe that the strategy $\lambda x.S(b,x)$ changes its mind on the initial segment constructed so far. Otherwise we just append one value $t \in \{0,1\}$, such that $S(b,\eta_0^b[y_k]) \neq S(b,\eta_0^b[y_k]t)$. Again we set $\eta_0^b(y_k+2) := 0$ in order to enable identification with convergently incorrect intermediate hypotheses for the learner $T$. Formally:

$$\eta_0^b(y_k+1) \ := \begin{cases} 0 & \text{if } S(b,\eta_0^b[y_k]0)\downarrow\ \wedge S(b,\eta_0^b[y_k]1)\downarrow \text{ and} \\ & [S(b,\eta_0^b[y_k]) \neq S(b,\eta_0^b[y_k]0) \text{ or} \\ & \qquad S(b,\eta_0^b[y_k]) = S(b,\eta_0^b[y_k]0) = S(b,\eta_0^b[y_k]1)] \\ 1 & \text{if } S(b,\eta_0^b[y_k]0)\downarrow\ \wedge S(b,\eta_0^b[y_k]1)\downarrow \text{ and} \\ & S(b,\eta_0^b[y_k]) = S(b,\eta_0^b[y_k]0) \text{ and } S(b,\eta_0^b[y_k]) \neq S(b,\eta_0^b[y_k]1) \\ \uparrow & \text{otherwise} \end{cases}$$

$$\eta_0^b(y_k+2) \ := \ 0$$

For any $x > y_k$ define

$$\eta_0^b(x+2) := \begin{cases} 0 & \text{if } S(b,\eta_0^b[y_k]0)\downarrow\ \wedge S(b,\eta_0^b[y_k]1)\downarrow \text{ and} \\ & S(b,\eta_0^b[y_k]) = S(b,\eta_0^b[y_k]0) = S(b,\eta_0^b[y_k]1) \text{ and} \\ & \text{for all } z \in \{y_k,\ldots,x\} \\ & \quad [S(b,\eta_0^b[z])\uparrow_{\leq x}\ \vee S(b,\eta_0^b[z]) = S(b,\eta_0^b[y_k])] \\ \uparrow & \text{if } S(b,\eta_0^b[y_k]0)\uparrow\ \vee S(b,\eta_0^b[y_k]1)\uparrow \\ \text{temporarily suspended} & \text{otherwise} \end{cases}$$

The functions $\eta_{2k+1}^b$ and $\eta_{2k+2}^b$ are defined as follows:
$\eta_{2k+1}^b[y_k+2] := \eta_0^b[y_k]0(2k+2),\ \eta_{2k+2}^b[y_k+2] := \eta_0^b[y_k]1(2k+3)$. Both functions will be extended by zeros until the values $S(b,\eta_0^b[y_k]0)$ and $S(b,\eta_0^b[y_k]1)$ are computed and the definition of $\eta_0^b$ is temporarily suspended (if these conditions are never satisfied, we obtain $\eta_{2k+1}^b = \eta_0^b[y_k]0(2k+2)0^\infty$ and $\eta_{2k+2}^b = \eta_0^b[y_k]1(2k+3)0^\infty$).

Formally we define for $x > y_k$:

$$\eta^b_{2k+1}(x+2) = \eta^b_{2k+2}(x+2) := \begin{cases} 0 & \text{if } [S(b,\eta^b_0[y_k]0)\uparrow_{\leq x} \ \vee \ S(b,\eta^b_0[y_k]1)\uparrow_{\leq x}] \text{ or} \\ & [S(b,\eta^b_0[y_k]) = S(b,\eta^b_0[y_k]0) = S(b,\eta^b_0[y_k]1) \\ & \text{and } \forall z \in \{y_k,\ldots,x\} \\ & [S(b,\eta^b_0[z])\uparrow_{\leq x} \ \vee \ S(b,\eta^b_0[z]){=}S(b,\eta^b_0[y_k])]] \\ \uparrow & \text{otherwise} \end{cases}$$

Compute

$$y_{k+1} := \begin{cases} \min\{n \in \mathbb{N} \mid \text{the definition of } \eta^b_0(n+1) & \text{if such a minimum exists} \\ \qquad \text{is temporarily suspended in stage } k\} & \\ \uparrow & \text{otherwise} \end{cases}$$

If $y_{k+1}$ is defined, go to stage $k+1$. *End stage $k$.*

By analogy with Fact 3 we obtain:

**Fact 4** *Fix $k \geq 1$.*

1. *If stage $k+1$ is reached, then*
   - *$y_{k+1} > y_k$ and $S(b,\eta^b_0[n]) \neq S(b,\eta^b_0[y_k])$ for some $n \in \{y_k{+}1,\ldots,y_{k+1}\}$ ($\lambda x.S(b,x)$ is forced into a mind change),*
   - *$\eta^b_{2k+1} = \eta^b_0[y_k]0(2k+2)0^m\uparrow^\infty$ and $\eta^b_{2k+2} = \eta^b_0[y_k]1(2k+3)0^m\uparrow^\infty$ for some $m \in \mathbb{N}$;*
2. *if stage $k$ is reached and stage $k+1$ is not reached, then*
   - *$\eta^b_{2k+1} = \eta^b_0[y_k]0(2k+2)0^\infty$ and $\eta^b_{2k+2} = \eta^b_0[y_k]1(2k+3)0^\infty$,*
   - *$\eta^b_0 \in \{\eta^b_0[y_k]\uparrow 0\uparrow^\infty, \eta^b_0[y_k]0^\infty\}$; in particular*
     - *$\eta^b_0 = \eta^b_0[y_k]\uparrow 0\uparrow^\infty$ iff $S(b,\eta^b_0[y_k]0)\uparrow \ \vee \ S(b,\eta^b_0[y_k]1)\uparrow$,*
     - *$\eta^b_0 = \eta^b_0[y_k]0^\infty$ iff $S(b,\eta^b_0[y_k]) = S(b,\eta^b_0[y_k]0) = S(b,\eta^b_0[y_k]1) = S(b,\eta^b_0[y_k]0^m)$ $(\in \mathbb{N})$ for any $m \in \mathbb{N}$;*
3. *$\eta^b_0(y_k{+}2)$ is defined and $\eta^b_0(y_k{+}2) \neq \eta^b_{2k+1}(y_k{+}2)$, $\eta^b_0(y_k{+}2) \neq \eta^b_{2k+2}(y_k{+}2)$.*

**Fact 5** *Our construction obviously yields*

1. *$rng(\eta^b_0) \subseteq \{0,1\}$,*
2. *if $x \in \mathbb{N}$, then $\max(rng(\eta^b_{x+1})) = x+2$ with $rng(\eta^b_{x+1}) \subseteq \{0,1,x+2\}$ or $rng(\eta^b_{x+1}) = \emptyset$,*

*3.* $\eta_0^b \not\subseteq \eta_{x+1}^b$ *for all* $x \in \mathbb{N}$.

**Fact 6** *Fact 4 implies*

1. *if in the construction of* $\eta^b$ *all stages* $k$ $(k \in \mathbb{N})$ *are reached, then we have* $\mathcal{R}_{\eta^b} = \{\eta_0^b\}$,

2. *if stage* $k$ *(for some* $k \in \mathbb{N}$*) is the last stage to be reached, we obtain* $\mathcal{R}_{\eta^b} = \{\eta_0^b, \eta_{2k+1}^b, \eta_{2k+2}^b\}$ *or* $\mathcal{R}_{\eta^b} = \{\eta_{2k+1}^b, \eta_{2k+2}^b\}$.

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \eta^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\eta^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \eta^{b_0}$.
*End Construction* $b_0$.

In order to contradict phrase (9) we will prove the following statements.

1. $b_0 \in B$,

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

*ad 1.*
$\mathcal{R}_{b_0}$ *is finite.* This is a consequence of Fact 6, because either all stages in the construction of $\eta_0^{b_0}$ are reached or there must be some number $k \in \mathbb{N}$, such that stage $k$ is the last stage to be reached.

$\mathcal{R}_{b_0} \in (EX_1)_{\varphi^{b_0}}(T) \cap CEX_{\varphi^{b_0}}(T)$*:* Assume $f \in \mathcal{R}_{b_0}$. It suffices to consider the following two cases.

*Case (i).* $f = \varphi_0^{b_0}$.
By Fact 5 we have $\mathrm{rng}(f) \subseteq \{0, 1\}$ and thus $T(f[n]) = 0$ for any $n \in \mathbb{N}$ by definition of $T$. Therefore $\{f\} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(T) \cap \mathrm{CEX}_{\varphi^{b_0}}(T)$.

*Case (ii).* $f = \varphi_{x+1}^{b_0}$ for some $x \in \mathbb{N}$.
By Fact 5 we know that $\max(\mathrm{rng}(f)) = x + 2$. The definition of $T$ then implies $T(f[n]) = x + 1$ for all but finitely many $n \in \mathbb{N}$. Thus $\{f\} \in \mathrm{EX}_{\varphi^{b_0}}(T)$. Since $\mathrm{rng}(f) \subseteq \{0, 1, x + 2\}$ by Fact 5.2, the strategy $T$ changes its mind on $f$ at most once. Furthermore $T(f[n]) \in \{0, x + 1\}$ for all $n \in \mathbb{N}$. As $\varphi_0^{b_0} \not\subseteq \varphi_{x+1}^{b_0}$, all incorrect hypotheses produced by $T$ on $f$ are convergently incorrect. Hence $\{f\} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(T) \cap \mathrm{CEX}_{\varphi^{b_0}}(T)$.

This yields $\mathcal{R}_{b_0} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(T) \cap \mathrm{CEX}_{\varphi^{b_0}}(T)$ and hence $b_0 \in B$. *qed 1.*

*ad 2.* Assume by way of contradiction that $\mathcal{R}_{b_0} \in \mathrm{CONF}(\lambda x.S(b_0, x))$. Then there must be a numbering $\psi \in \mathcal{P}^2$, such that $\mathcal{R}_{b_0} \in \mathrm{CONF}_\psi(\lambda x.S(b_0, x))$. By

Fact 6 it suffices to consider the following three cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$.
Then by Fact 4 all stages are reached in the definition of $\eta^{b_0}$. With Fact 4.1 we observe that in the identification process for $\varphi_0^{b_0}$ the strategy $\lambda x.S(b_0, x)$ changes its hypothesis infinitely often. This contradicts $\mathcal{R}_{b_0} \in \mathrm{CONF}(\lambda x.S(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi_{2k+1}^{b_0}, \varphi_{2k+2}^{b_0}\}$ for some $k \in \mathbb{N}$.
In this case we obtain $S(b_0, \varphi_{2k+1}^{b_0}[y_k + 1])\uparrow$ or $S(b_0, \varphi_{2k+2}^{b_0}[y_k + 1])\uparrow$ from Fact 4.2, although $\lambda x.S(b_0, x)$ ought to be defined on input of any initial segment of $\varphi_{2k+1}^{b_0}$ and $\varphi_{2k+2}^{b_0}$. Again this yields a contradiction to $\mathcal{R}_{b_0} \in \mathrm{CONF}(\lambda x.S(b_0, x))$.

*Case (iii).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}, \varphi_{2k+1}^{b_0}, \varphi_{2k+2}^{b_0}\}$ for some $k \in \mathbb{N}$.
Then stage $k$ is reached and stage $k + 1$ is not reached. Furthermore we have

$$\begin{aligned} i \;&:=\; S(b_0, \varphi_{2k+1}^{b_0}[y_k + 1]) = S(b_0, \eta_0^{b_0}[y_k]0) = S(b_0, \eta_0^{b_0}[y_k]1) \\ &=\; S(b_0, \varphi_{2k+2}^{b_0}[y_k + 1]) \ , \end{aligned}$$

although $\varphi_{2k+1}^{b_0}[y_k + 1] \neq \varphi_{2k+2}^{b_0}[y_k + 1]$. Therefore $i$ cannot be a $\psi$-number for both $\varphi_{2k+1}^{b_0}$ and $\varphi_{2k+2}^{b_0}$. We distinguish between two possibilities.

*Case (iii)a.* $\psi_i(y_k + 1)\uparrow$.
Then $\psi_i \notin \mathcal{R}$ and in particular $\psi_i \neq \varphi_0^{b_0}$. But $S(b_0, \varphi_0^{b_0}[y_k]) = i$ and by Fact 4.2 we know $S(b_0, \varphi_0^{b_0}[n]) = i$ for all but finitely many $n \in \mathbb{N}$. Hence the sequence of hypotheses produced by $\lambda x.S(b_0, x)$ on the function $\varphi_0^{b_0}$ converges to an index incorrect for $\varphi_0^{b_0}$ with respect to $\psi$. This contradicts $\mathcal{R}_{b_0} \in \mathrm{CONF}_\psi(\lambda x.S(b_0, x))$.

*Case (iii)b.* $\psi_i(y_k + 1) \in \mathbb{N}$.
Then $i$ cannot be conform for both $\varphi_{2k+1}^{b_0}[y_k+1]$ and $\varphi_{2k+2}^{b_0}[y_k+1]$ with respect to $\psi$, although on input of initial segments of $\varphi_{2k+1}^{b_0}$ and $\varphi_{2k+2}^{b_0}$ the strategy $\lambda x.S(b_0, x)$ ought to return just hypotheses conform with respect to $\psi$. Again this yields a contradiction to $\mathcal{R}_{b_0} \in \mathrm{CONF}_\psi(\lambda x.S(b_0, x))$.

In each of our cases we reached a contradiction. Since further cases cannot occur, we conclude $\mathcal{R}_{b_0} \notin \mathrm{CONF}(\lambda x.S(b_0, x))$. *qed 2.*

Claims 1 and 2 together now contradict statement (9). Therefore our assumption must have been wrong, i.e. $B \notin \mathrm{suit}(J^*, \mathrm{CONF})$. Altogether we have $B \in (\mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \cap \mathrm{suit}_\varphi(J^*, \mathrm{CEX})) \setminus \mathrm{suit}(J^*, \mathrm{CONF})$. This completes the proof. $\square$

Since $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \subset \mathrm{suit}_\varphi(J^*, \mathrm{EX})$, this yields the following result:

**Corollary 4** $suit_\varphi(J^*, EX) \setminus suit(J^*, CONF) \neq \emptyset$.

So we have also verified, that the set $\text{suit}_\varphi(J^*, \text{CONF})$ is a proper subset of $\text{suit}_\varphi(J^*, \text{EX})$, as well as the corresponding result for uniform learning with respect to $\tau$. By analogy we want to separate the criteria CONS and CONF. For that purpose we use the stronger result in Theorem 11.

**Theorem 11** $suit_\varphi(J^*, CONF) \setminus suit(J^*, CONS) \neq \emptyset$.

*Proof.* Again we use a uniform strategy $T \in \mathcal{P}^2$ to define a description set $B \subseteq \mathbb{N}$ suitable for uniform conform identification in the limit by $T$. The set $B$ shall describe only finite recursive cores and will not be suitable for uniform consistent identification.

*Definition of $B \in suit_\varphi(J^*, CONF) \setminus suit(J^*, CONS)$.*

Define a strategy $T \in \mathcal{R}$ by

$$T(f[n]) := \begin{cases} 0 & \text{if } f[n] \in \{0,1\}^* \\ \max\{f(0), \dots, f(n)\} - 1 & \text{otherwise} \end{cases}$$

for arbitrary $f \in \mathcal{R}$ and $n \in \mathbb{N}$. Then set

$$B := \{b \in \mathbb{N} \mid \mathcal{R}_b \text{ is finite and } \mathcal{R}_b \in \text{CONF}_{\varphi^b}(T)\}.$$

*Claim.* $B \in \text{suit}_\varphi(J^*, \text{CONF}) \setminus \text{suit}(J^*, \text{CONS})$.

*Proof of "$B \in suit_\varphi(J^*, CONF)$".*

Defining $T' \in \mathcal{R}^2$ by $T'(b, f[n]) := T(f[n])$ for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ we obviously have

- $\forall b \in B\ [\mathcal{R}_b \in J^*]$,
- $\forall b \in B\ [\mathcal{R}_b \in \text{CONF}_{\varphi^b}(\lambda x.T'(b, x))]$.

By definition of uniform identifiability this implies $B \in \text{suit}_\varphi(J^*, \text{CONF})$.

*Proof of "$B \notin suit(J^*, CONS)$".*

We will verify this claim by way of contradiction.
*Assumption.* $B \in \text{suit}(J^*, \text{CONS})$.

Then there exists a strategy $S \in \mathcal{P}^2$ such that for any $b \in B$ there is a hypothesis space $\psi \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in \text{CONS}_\psi(\lambda x.S(b, x))$; abbreviated

$$\forall b \in B\ [\mathcal{R}_b \in \text{CONS}(\lambda x.S(b, x))]\ . \tag{10}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0, x))$,

in contradiction to statement (10). The strategy $\lambda x.S(b_0, x)$ will fail for at least one function $f \in \mathcal{R}_{b_0}$ by either

- changing its hypothesis for $f$ infinitely often or
- not terminating its computation on input of some initial segment of $f$ or
- violating the consistency demand on input of some initial segment of $f$.

In order to deduce a contradiction it suffices to prove that one of these three cases must occur. Which of these cases actually occurs for the constructed integer $b_0$, does not have to be decidable.

*Construction of* $b_0$.

For arbitrary $b \in \mathbb{N}$ set $\eta(b, x, 0) := 0$ (for any $x \in \mathbb{N}$) and define $\eta(b, 0, 1)$ by the following instructions:

$\eta(b, 0, 1) :=$"Compute $S(b, 0)$. Look for a $t \in \{0, 1\}$ such that $S(b, 0t)\downarrow$ and $S(b, 0t) \neq S(b, 0)$. Put out the first $t$ you find."

**Fact 7** *The definition of* $\eta(b, 0, 1)$ *implies*

1. *If* $\eta(b, 0, 1)\downarrow$*, then* $S(b, \eta_0^b[1]) \neq S(b, \eta_0^b[0])$*, i.e.* $\lambda x.S(b, x)$ *is forced to change its mind.*
2. $\eta(b, 0, 1)$ *is undefined if and only if* $S(b, 0)\uparrow$ *or neither* $S(b, 00)$ *nor* $S(b, 01)$ *is defined with a value differing from* $S(b, 0)$.

Furthermore define for any $y \geq 1$:

$$\eta(b, 1, y) \quad := \quad \begin{cases} 0 & \text{if } y = 1 \\ 2 & \text{if } y = 2 \\ 0 & \text{if } y > 2 \text{ and either } S(b, 0)\uparrow \text{ or} \\ & [S(b, 00)\uparrow_{\leq y} \ \vee S(b, 00) = S(b, 0)] \\ & \wedge\, [S(b, 01)\uparrow_{\leq y} \ \vee S(b, 01) = S(b, 0)] \\ \uparrow & \text{otherwise} \end{cases}$$

$$\eta(b, 2, y) \quad := \quad \begin{cases} 1 & \text{if } y = 1 \\ 3 & \text{if } y = 2 \\ \eta(b, 1, y) & \text{otherwise} \end{cases}$$

The values $\eta(b, 1, 2) = 2$ and $\eta(b, 2, 2) = 3$ will allow our strategy $\lambda x.T(b, x)$ to identify $\eta_1^b$ and $\eta_2^b$ in the limit, if these functions are recursive.

**Fact 8** *The following statements are equivalent.*

- $\eta_1^b$ *and* $\eta_2^b$ *are recursive functions,*
- $\eta_1^b = 0020^\infty$ *and* $\eta_2^b = 0130^\infty$*,*
- $\eta(b,0,1)$ *is undefined.*

In general, we define for arbitrary $x, y \in \mathbb{N}$:

$$\eta(b,0,y+1) := \text{“Compute } S(b,\eta_0^b[y])\text{. Look for a } t \in \{0,1\} \text{ such that } S(b,\eta_0^b[y]t)\downarrow \text{ and } S(b,\eta_0^b[y]t) \neq S(b,\eta_0^b[y])\text{. Put out the first } t \text{ you find.”}$$

**Fact 9** *By analogy with Fact 7 the definition of* $\eta(b,0,y+1)$ *implies*

1. *If* $\eta(b,0,y+1)\downarrow$*, then* $S(b,\eta_0^b[y+1]) \neq S(b,\eta_0^b[y])$*, i.e.* $\lambda x.S(b,x)$ *is forced to change its mind.*
2. $\eta(b,0,y+1)$ *is undefined if and only if* $S(b,\eta_0^b[y])\uparrow$ *or neither* $S(b,\eta_0^b[y]0)$ *nor* $S(b,\eta_0^b[y]1)$ *is defined with a value differing from* $S(b,\eta_0^b[y])$*.*

Furthermore define:

$$\eta(b,2x+1,y) \;:=\; \begin{cases} \eta_0^b(y) & \text{if } y \leq x \\ 0 & \text{if } y = x+1 \\ 2x+2 & \text{if } y = x+2 \\ 0 & \text{if } y > x+2 \text{ and either } S(b,\eta_0^b[x])\uparrow \text{ or} \\ & [S(b,\eta_0^b[x]0)\uparrow_{\leq y} \;\vee\; S(b,\eta_0^b[x]0) = S(b,\eta_0^b[x])] \\ & \wedge\; [S(b,\eta_0^b[x]1)\uparrow_{\leq y} \;\vee\; S(b,\eta_0^b[x]1) = S(b,\eta_0^b[x])] \\ \uparrow & \text{otherwise} \end{cases}$$

$$\eta(b,2x+2,y) \;:=\; \begin{cases} 1 & \text{if } y = x+1 \\ 2x+3 & \text{if } y = x+2 \\ \eta(b,2x+1,y) & \text{otherwise} \end{cases}$$

The values $\eta(b,2x+1,x+2) = 2x+2$ and $\eta(b,2x+2,x+2) = 2x+3$ will allow our strategy $\lambda x.T(b,x)$ to identify $\eta_{2x+1}^b$ and $\eta_{2x+2}^b$ in the limit, if these functions are recursive.

**Fact 10** *By analogy with Fact 8 the following statements are equivalent.*

- $\eta_{2x+1}^b$ *and* $\eta_{2x+2}^b$ *are recursive functions,*
- $\eta_{2x+1}^b = \eta_0^b[x]0(2x+2)0^\infty$ *and* $\eta_{2x+2}^b = \eta_0^b[x]1(2x+3)0^\infty$*,*

- *$\eta_0^b(0), \dots \eta_0^b(x)$ are defined and $\eta(b,0,x+1)$ is undefined; $\eta_0^b \subset \eta_{2x+1}^b$, $\eta_0^b \subset \eta_{2x+2}^b$.*

**Fact 11** *From Fact 9 and Fact 10 we can deduce that for any $b \in \mathbb{N}$ exactly one of the following cases occurs.*

1. *$\eta_0^b \in \mathcal{R}$ or*
2. *$\eta_{2x+1}^b \in \mathcal{R}$ and $\eta_{2x+2}^b \in \mathcal{R}$ for exactly one integer $x \in \mathbb{N}$.*

**Fact 12** *Assume $f \in \mathcal{R}_{\eta^b}$. From Fact 11 and the Definition of $\eta$ we conclude*

1. *if $rng(f) \subseteq \{0,1\}$, then $f = \eta_0^b$;*
2. *if $rng(f) \not\subseteq \{0,1\}$, then $f = \eta_{\max(rng(f))-1}^b$ and $\eta_0^b \subset f$.*

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \eta^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\eta^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \eta^{b_0}$.

*End Construction $b_0$.*

In order to contradict phrase (10) we will prove the following statements.

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0,x))$.

*ad 1.* Fact 11 implies that $\mathcal{R}_{b_0}(= \mathcal{R}_{\eta^{b_0}})$ is finite. From Fact 12 and $\varphi^{b_0} = \eta^{b_0}$ we can conclude that $\mathcal{R}_{b_0} \in \mathrm{CONF}_{\varphi^{b_0}}(T)$. Thus $b_0 \in B$. *qed 1.*

*ad 2.* We distinguish between two cases.

*Case (i).* $\varphi_0^{b_0} \in \mathcal{R}$.
From Fact 9 we know that the strategy $\lambda x.S(b_0,x)$ changes its hypothesis infinitely often on the input sequence $(\varphi_0^{b_0}[n])_{n\in\mathbb{N}}$. Thus we even conclude $\mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0,x))$ and in particular $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0,x))$.

*Case (ii).* $\varphi_{2x+1}^{b_0} \in \mathcal{R}$ and $\varphi_{2x+2}^{b_0} \in \mathcal{R}$ for some $x \in \mathbb{N}$.
Again we distinguish between two cases.

*Case (ii)a.* $S(b_0, \eta_0^{b_0}[x]t)$ is undefined for at least one $t \in \{0,1\}$.
Then the definition of $\eta_{2x+1}^{b_0}$ and $\eta_{2x+2}^{b_0}$ implies $S(b_0, \eta_{2x+r}^{b_0}[x+1])\uparrow$ for at least one $r \in \{1,2\}$. Since $\eta_{2x+1}^{b_0} \in \mathcal{R}_{b_0}$ and $\eta_{2x+2}^{b_0} \in \mathcal{R}_{b_0}$, we know $\mathcal{R}_{b_0} \notin \mathrm{EX}(\lambda x.S(b_0,x))$ and in particular $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0,x))$.

*Case (ii)b.* $S(b_0, \eta_0^{b_0}[x]0)\downarrow$ and $S(b_0, \eta_0^{b_0}[x]1)\downarrow$.
From Fact 10 and Fact 9.2 we conclude

$$S(b_0, \eta_0^{b_0}[x]0) = S(b_0, \eta_0^{b_0}[x]1) \ (= S(b_0, \eta_0^{b_0}[x])) \ .$$

But the hypothesis $S(b_0, \eta_0^{b_0}[x])$ $(= S(b_0, \eta_{2x+1}^{b_0}[x+1]) = S(b_0, \eta_{2x+2}^{b_0}[x+1]))$ cannot be consistent for both $\eta_{2x+1}^{b_0}[x+1]$ $(= \eta_0^{b_0}[x]0)$ and $\eta_{2x+2}^{b_0}[x+1]$ $(= \eta_0^{b_0}[x]1)$. Therefore $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0, x))$.

Thus also in Case *(ii)* we have $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0, x))$.

Because of Fact 11 further cases cannot occur. This proves our second statement, namely $\mathcal{R}_{b_0} \notin \mathrm{CONS}(\lambda x.S(b_0, x))$. *qed 2.*

Claims 1 and 2 together now contradict statement (10). Therefore our assumption must have been wrong, i.e. $B \notin \mathrm{suit}(J^*, \mathrm{CONS})$. Altogether we have $B \in \mathrm{suit}_\varphi(J^*, \mathrm{CONF}) \setminus \mathrm{suit}(J^*, \mathrm{CONS})$. This completes the proof. □

So, the results CONS $\subset$ CONF $\subset$ EX can also be transferred to uniform learning with respect to $\tau$ and the numberings given a priori by $\varphi$. Again, finite classes are sufficient for the separations.

In order to prove $\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL}) \subset \mathrm{suit}_\varphi(J^*, \mathrm{CONS})$ (and the same result for $\mathrm{suit}_\tau$) we can use Theorem 12. We even obtain

$$\mathrm{suit}_\varphi(J^*, \mathrm{CONS}) \setminus \mathrm{suit}_\tau(J^*, \mathrm{TOTAL}) \neq \emptyset \ ,$$

because $\mathrm{suit}_\tau(J^*, \mathrm{TOTAL}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{CEX})$. A separation of the inference criteria CP and TOTAL can not be achieved analogously. We will discuss this case later.

**Theorem 12**

1. $\mathit{suit}_\varphi(J^1, \mathit{EX}_1) \setminus \mathit{suit}_\tau(J^1, \mathit{CEX}) \neq \emptyset$,
2. $\mathit{suit}_\varphi(J^1, \mathit{CONS}) \setminus \mathit{suit}_\tau(J^1, \mathit{CEX}) \neq \emptyset$.

*Proof.* Again we use a uniform strategy $T \in \mathcal{P}^2$ to define a description set $B \subseteq \mathbb{N}$. The corresponding classes of functions will by definition be uniformly identifiable by $T$ with consistent intermediate hypotheses and no more than one mind change. The set $B$ shall describe only singleton recursive cores and will not be suitable for uniform CEX-identification with respect to our acceptable numbering $\tau$.

*Definition of* $B \in (suit_\varphi(J^1, EX_1) \cap suit_\varphi(J^1, CONS)) \setminus suit_\tau(J^1, CEX)$.
Define a strategy $T \in \mathcal{P}^2$ by

$$T(b, f[n]) := \begin{cases} 0 & \text{if } 0 \notin \{f(0), \ldots, f(n)\} \\ \min\{i \geq 1 \mid \exists \alpha \in (\mathbb{N} \setminus \{0\})^* & \text{if } 0 \in \{f(0), \ldots, f(n)\} \text{ and} \\ \qquad [\alpha 0 \subseteq \varphi_i^b \text{ and } \alpha 0 \subseteq f]\} & \text{such a minimum is found} \\ \uparrow & \text{otherwise} \end{cases}$$

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

$$B := \{b \in \mathbb{N} \mid \text{card } \mathcal{R}_b = 1 \ \wedge \ \mathcal{R}_b \in (\text{EX}_1)_{\varphi^b}(\lambda x.T(b,x)) \cap \text{CONS}_{\varphi^b}(\lambda x.T(b,x))\} \ .$$

*Claim.* $B \in (\text{suit}_\varphi(J^1, \text{EX}_1) \cap \text{suit}_\varphi(J^1, \text{CONS})) \setminus \text{suit}_\tau(J^1, \text{CEX})$.

*Proof of "$B \in suit_\varphi(J^1, EX_1) \cap suit_\varphi(J^1, CONS)$".*
We obviously have

- $\forall b \in B \ [\mathcal{R}_b \in J^1]$,
- $\exists T \in \mathcal{P}^2 \ \forall b \in B \ [\mathcal{R}_b \in (\text{EX}_1)_{\varphi^b}(\lambda x.T(b,x)) \ \wedge \ \mathcal{R}_b \in \text{CONS}_{\varphi^b}(\lambda x.T(b,x))]$.

By Definition 12 this implies $B \in \text{suit}_\varphi(J^1, \text{EX}_1) \cap \text{suit}_\varphi(J^1, \text{CONS})$.

*Proof of "$B \notin suit_\tau(J^1, CEX)$".*
We will verify this claim by way of contradiction.
*Assumption.* $B \in \text{suit}_\tau(J^1, \text{CEX})$.

Then by Proposition 3 there exists a strategy $S \in \mathcal{R}^2$ such that $\mathcal{R}_b \in \text{CEX}_\tau(\lambda x.S(b,x))$ for any $b \in B$; abbreviated

$$\forall b \in B \ [\mathcal{R}_b \in \text{CEX}_\tau(\lambda x.S(b,x))] \ . \tag{11}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \text{CEX}_\tau(\lambda x.S(b_0,x))$,

in contradiction to (11). The strategy $\lambda x.S(b_0,x)$ will fail for the only function $f \in \mathcal{R}_{b_0}$ by either

- changing its hypothesis for $f$ infinitely often or
- generating a hypothesis incorrect for $f$ with respect to $\tau$ for infinitely many initial segments of $f$ or
- guessing a $\tau$-number of a proper subfunction of $f$ on input of some initial segment of $f$.

*Construction of $b_0$.*
Define a function $\psi \in \mathcal{P}^3$ with the help of initial segments $\alpha_k^b$ $(b, k \in \mathbb{N})$ as follows: for arbitrary $b \in \mathbb{N}$ set $\alpha_0^b := 1$ and begin in stage 0.

*Stage 0.*
Let $\psi_0^b(0) := 1$ and $e := S(b, \alpha_0^b)$. Start a parallel computation until *(i)* or *(ii)* turns out to be true.

*(i).* $\tau_e(0)$ is defined and $\tau_e(0) \neq 1$.
*(ii).* There is an integer $y > 0$ such that $\tau_e(y)$ is defined.

The function $\psi_1^b$ shall have the initial segment $\alpha_0^b 0$ which will be extended by a sequence of 0's, until *(i)* or *(ii)* turns out to be true:

$$\psi_1^b(x) := \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ 0 & \text{if } x > 1 \text{ and neither } \textit{(i)} \text{ nor } \textit{(ii)} \text{ is fulfilled} \\ & \qquad\qquad\quad \text{within } x \text{ steps of computation} \\ \uparrow & \text{otherwise} \end{cases}$$

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

If condition *(i)* turns out to be true first, then $\psi_0^b$ shall have the initial segment $\alpha_0^b$ which will be extended by a sequence of 1's, until $\lambda x.S(b, x)$ is forced to change its mind on $\psi_0^b$; then $\alpha_1^b$ shall be the initial segment of $\psi_0^b$ constructed so far. More formally:

If condition *(i)* turns out to be true first, set for $x \in \mathbb{N}$:

$$\psi_0^b(x) := \begin{cases} 1 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \text{ and } S(b, \psi_0^b[x-1]) = e \\ \text{see stage 1} & \text{otherwise} \end{cases}$$

If stage 1 is reached in this definition, let $j_0^b \in \mathbb{N}$ be the minimal integer such that $\psi_0^b(j_0^b)$ has been defined in stage 0. Let

$$\alpha_1^b := \begin{cases} \psi_0^b[j_0^b] & \text{if } j_0^b \text{ is defined} \\ \uparrow & \text{otherwise} \end{cases}$$

If condition *(ii)* turns out to be true first – with $\tau_e(y_0)\downarrow$ – then $\alpha_1^b := \alpha_0^b 1 \ldots 1(\tau_e(y_0) + 1)$, where the last argument in the domain of $\alpha_1^b$ is $y_0$. Go to stage 1.

Note that $\psi_0^b$ remains initial if neither *(i)* nor *(ii)* is fulfilled. *End stage 0.*

**Fact 13** *The construction in stage 0 implies*

1. *$\psi_1^b \in \mathcal{R}$ if and only if $\tau_{S(b,\psi_1^b[0])} \subseteq 1{\uparrow}^\infty \ \ (\subset \psi_1^b)$,*

2. *if $\psi_1^b \notin \mathcal{R}$ and $\alpha_1^b\uparrow$, then $\psi_0^b = \alpha_0^b 1^\infty \in \mathcal{R}$ and the sequence of hypotheses produced by $\lambda x.S(b,x)$ on $\psi_0^b$ converges to an index incorrect for $\psi_0^b$ with respect to $\tau$,*

3. *if $\psi_1^b \notin \mathcal{R}$ and $\alpha_1^b\downarrow$, then $\alpha_0^b \subseteq \alpha_1^b \subseteq \psi_0^b$; furthermore*

   *(a) $S(b,\alpha_1^b) \neq S(b,\alpha_0^b)$ or*

   *(b) $S(b,f[0])$ is incorrect with respect to $\tau$ for any $f \in \mathcal{R}$ satisfying $\alpha_1^b \subset f$,*

*Proof of Fact 13.*

*ad 1.* We have $\psi_1^b \in \mathcal{R}$ if and only if $\psi_1^b = 10^\infty$ if and only if neither *(i)* nor *(ii)* is fulfilled. This is equivalent to

$$[\tau_e(0)\uparrow \vee \tau_e(0) = 1] \ \wedge \ [\tau_e(y)\uparrow \ \text{ for all } y > 0] \ .$$

Thus $\psi_1^b \in \mathcal{R}$ if and only if $\tau_{S(b,\psi_1^b[0])} = \tau_e \subseteq 1{\uparrow}^\infty \subset 10^\infty = \psi_1^b$.

*ad 2.* If $\psi_1^b \notin \mathcal{R}$ and $\alpha_1^b\uparrow$, then $\psi_0^b = \alpha_0^b 1^\infty \in \mathcal{R}$ follows from the construction for the case that condition *(i)* turns out to be true first. Then we also obtain $S(b,\psi_0^b[x]) = e$ for all $x \in \mathbb{N}$, although $\tau_e(0) \neq 1 = \psi_0^b(0)$ by condition *(i)*. So the sequence of hypotheses returned by $\lambda x.S(b,x)$ on $\psi_0^b$ converges to a wrong $\tau$-number.

*ad 3.* If $\psi_1^b \notin \mathcal{R}$ and $\alpha_1^b\downarrow$, then $\alpha_0^b \subseteq \alpha_1^b \subseteq \psi_0^b$ by definition. Furthermore, as $\psi_1^b \notin \mathcal{R}$, at least one of the conditions *(i)*, *(ii)* must be fulfilled in stage 0. If condition *(i)* is fulfilled first, then $S(b,\alpha_1^b) \neq S(b,\alpha_0^b)$ by the construction of $\psi_0^b$. If condition *(ii)* is fulfilled first – with $\tau_{S(b,1)}(y_0)\downarrow$ – then $\tau_{S(b,\alpha_1^b[0])}(y_0) = \tau_{S(b,1)}(y_0) \neq \tau_{S(b,1)}(y_0)+1 = \alpha_1^b(y_0)$, therefore in this case $S(b,f[0])$ is incorrect with respect to $\tau$ for any $f \in \mathcal{R}$ satisfying $\alpha_1^b \subset f$. *qed Fact 13.*

Now we iterate the construction in stage 0. In general, for $k \in \mathbb{N}$ stage $k$ is used for the definition of $\psi_{k+1}^b$. If stage $k$ is never reached, then $\psi_{k+1}^b$ will be the empty function.

*Stage k.*

We know that $\psi_0^b(x) = \alpha_k^b(x)$ for all $x < |\alpha_k^b|$. Let $e := S(b,\alpha_k^b)$. Start a parallel computation until *(i)* or *(ii)* turns out to be true.

*(i).* There is an integer $y < |\alpha_k^b|$ such that $\tau_e(y)$ is defined and $\tau_e(y) \neq \alpha_k^b(y)$.
*(ii).* There is an integer $y \geq |\alpha_k^b|$ such that $\tau_e(y)$ is defined.

The function $\psi^b_{k+1}$ shall have the initial segment $\alpha^b_k 0$ which will be extended by a sequence of 0's, until *(i)* or *(ii)* turns out to be true:

$$\psi^b_{k+1}(x) := \begin{cases} \alpha^b_k(x) & \text{if } x < |\alpha^b_k| \\ 0 & \text{if } x = |\alpha^b_k| \\ 0 & \text{if } x > |\alpha^b_k| \text{ and neither } \textit{(i)} \text{ nor } \textit{(ii)} \text{ is fulfilled} \\ & \qquad\qquad \text{within } x \text{ steps of computation} \\ \uparrow & \text{otherwise} \end{cases}$$

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

If condition *(i)* turns out to be true first, then $\psi^b_0$ shall have the initial segment $\alpha^b_k$ which will be extended by a sequence of 1's, until $\lambda x.S(b,x)$ is forced to change its mind on $\psi^b_0$; then $\alpha^b_{k+1}$ shall be the initial segment of $\psi^b_0$ constructed so far. More formally:

If condition *(i)* turns out to be true first, set for $x \in \mathbb{N}$:

$$\psi^b_0(x) := \begin{cases} 1 & \text{if } x \geq |\alpha^b_k| \text{ and } S(b, \psi^b_0[x-1]) = e \\ \text{see stage } k+1 & \text{otherwise} \end{cases}$$

If stage $k+1$ is reached in this definition, let $j^b_k \in \mathbb{N}$ be the minimal integer such that $\psi^b_0(j^b_k)$ has been defined in stage $k$. Let

$$\alpha^b_{k+1} := \begin{cases} \psi^b_0[j^b_k] & \text{if } j^b_k \text{ is defined} \\ \uparrow & \text{otherwise} \end{cases}$$

If condition *(ii)* turns out to be true first – with $\tau_e(y_k)\downarrow$ – then $\alpha^b_{k+1} := \alpha^b_k 1 \ldots 1(\tau_e(y_k)+1)$, where the last argument in the domain of $\alpha^b_{k+1}$ is $y_k$. Go to stage $k+1$.

Note that $\psi^b_0$ remains initial if neither *(i)* nor *(ii)* is fulfilled. *End stage $k$.*

**Fact 14** *The construction in stage $k$ implies*

1. $\psi^b_{k+1} \in \mathcal{R}$ *if and only if* $\alpha^b_k\downarrow$ *and* $\tau_{S(b,\alpha^b_k)} \subseteq \alpha^b_k\uparrow^\infty$ $(\subset \psi^b_{k+1})$,

2. *if* $\psi^b_{k+1} \notin \mathcal{R}$ *and* $\alpha^b_{k+1}\uparrow$, *then* $\psi^b_0 = \alpha^b_k 1^\infty \in \mathcal{R}$ *and the sequence of hypotheses produced by* $\lambda x.S(b,x)$ *on* $\psi^b_0$ *converges to an index incorrect for* $\psi^b_0$ *with respect to* $\tau$,

3. *if* $\psi^b_{k+1} \notin \mathcal{R}$ *and* $\alpha^b_{k+1}\downarrow$, *then* $\alpha^b_k \subset \alpha^b_{k+1} \subseteq \psi^b_0$; *furthermore*

   (a) $S(b, \alpha^b_{k+1}) \neq S(b, \alpha^b_k)$ *or*

   (b) $S(b, f[|\alpha^b_k|-1])$ *is incorrect with respect to* $\tau$ *for any* $f \in \mathcal{R}$ *satisfying* $\alpha^b_{k+1} \subset f$, This fact can be verified in a way similar to the proof of Fact 13.

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$. Therefore all facts mentioned above hold analogously, if we substitute $\psi^b$ by $\varphi^{b_0}$ and $b$ by $b_0$ everywhere.

*End Construction* $b_0$.

**Fact 15**

1. $0 \notin rng(\alpha_k^{b_0})$ *for all* $k \in \mathbb{N}$,
2. $0 \notin rng(\varphi_0^{b_0})$,
3. *there is exactly one index* $i \in \mathbb{N}$ *such that* $\varphi_i^{b_0} \in \mathcal{R}$.

*Proof of Fact 15.* Properties 1 and 2 follow immediately from the construction. For the third property argue by distinguishing two cases as follows:

*Case (i).* $\alpha_k^{b_0}\downarrow$ for all $k \in \mathbb{N}$.
Then $\varphi_0^{b_0}$ is a recursive function by construction. Assume $\varphi_{k+1}^{b_0} \in \mathcal{R}$ for some $k \in \mathbb{N}$. Fact 14.1 then tells us that $\tau_{S(b_0,\alpha_k^{b_0})} \subseteq \alpha_k^{b_0}\uparrow^\infty$ and so neither property *(i)* nor property *(ii)* are fulfilled. As we have already mentioned in the description of stage $k$, this implies that $\varphi_0^{b_0}$ is an initial function. This contradicts $\varphi_0^{b_0} \in \mathcal{R}$. Therefore 0 is the only $\varphi^{b_0}$-number of a total function.

*Case (ii).* There is some $k \in \mathbb{N}$ such that $\alpha_k^{b_0}\downarrow$ and $\alpha_{k+1}^{b_0}\uparrow$.

*Case (ii)a.* $\varphi_0^{b_0} \in \mathcal{R}$.
By the same argumentation as in Case (i) we obtain $\varphi_1^{b_0}, \dots, \varphi_{k+1}^{b_0} \notin \mathcal{R}$. Now, since $\alpha_{k+1}^{b_0}\uparrow$, stage $k+1$ is not reached in the construction of $\psi^{b_0}$. Therefore $\varphi_x^{b_0} \notin \mathcal{R}$ for all $x > k+1$. Thus again 0 is the only $\varphi^{b_0}$-number of a total function.

*Case (ii)b.* $\varphi_0^{b_0} \notin \mathcal{R}$.
If *(i)* or *(ii)* were fulfilled in the parallel computation in stage $k$, we would have either $\alpha_{k+1}^{b_0}\downarrow$ or $\varphi_0^{b_0} \in \mathcal{R}$, which is a contradiction to our assumptions in this case. So neither *(i)* nor *(ii)* is fulfilled. This yields $\varphi_{k+1}^{b_0} \in \mathcal{R}$. If for some $x \in \{1, \dots, k\}$ the function $\varphi_x^{b_0}$ was recursive, then $\alpha_k^{b_0}$ would be undefined. So $\varphi_1^{b_0}, \dots, \varphi_k^{b_0} \notin \mathcal{R}$. As $\alpha_{k+1}^{b_0}\uparrow$, we obtain $\varphi_x^{b_0} \notin \mathcal{R}$ for all $x > k+1$ as in Case (ii)a. Hence $k+1$ is the only $\varphi^{b_0}$-number of a total function.

Further cases cannot occur, so there is exactly one index $i \in \mathbb{N}$, such that $\varphi_i^{b_0} \in \mathcal{R}$. This proves Fact 15. *qed Fact 15.*

In order to contradict phrase (11) we will prove the following statements.

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CEX}_\tau(\lambda x.S(b_0, x))$.

*ad 1.* Fact 15.3 implies $\mathcal{R}_{b_0} \in J^1$. Therefore it remains to prove that $\mathcal{R}_{b_0} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(\lambda x.T(b_0, x)) \cap \mathrm{CONS}_{\varphi^{b_0}}(\lambda x.T(b_0, x))$. We consider two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$.
Because of Fact 15.2 we have $0 \notin \mathrm{rng}(\varphi_0^{b_0})$, and hence $T(b_0, \varphi_0^{b_0}[n]) = 0$ for all $n \in \mathbb{N}$. Thus $\mathcal{R}_{b_0} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(\lambda x.T(b_0, x)) \cap \mathrm{CONS}_{\varphi^{b_0}}(\lambda x.T(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi_{k+1}^{b_0}\}$ with $k \in \mathbb{N}$.
Then $\varphi_{k+1}^{b_0} = \alpha_k^{b_0} 0^\infty$ and $0 \notin \mathrm{rng}(\alpha_k^{b_0})$ (see Fact 15.1).

For all $n < |\alpha_k^{b_0}|$ we obtain $T(b_0, \varphi_{k+1}^{b_0}[n]) = 0$ by definition of $T$. Then $\alpha_k^{b_0} \subseteq \varphi_0^{b_0}$ according to stage $k$, which yields $\varphi^{b_0}_{T(b_0, \varphi_{k+1}^{b_0}[n])} =_n \varphi_{k+1}^{b_0}$. So our intermediate hypotheses are consistent in the first steps of the learning process.

Now for all $n \geq |\alpha_k^{b_0}|$ the segment $\alpha_k^{b_0}$ is the only tuple in $(\mathbb{N} \setminus \{0\})^*$ satisfying $\alpha_k^{b_0} 0 \subseteq \varphi_{k+1}^{b_0}$. So for all $n \geq |\alpha_k^{b_0}|$ the hypothesis $T(b_0, \varphi_{k+1}^{b_0}[n])$ will be the first number $i$ to be found such that $\alpha_k^{b_0} 0 \subseteq \varphi_i^{b_0}$.

As $\varphi_{k+2}^{b_0} \notin \mathcal{R}$, we know that $\alpha_{k+1}^{b_0}$ is undefined, so $\varphi_x^{b_0} = \uparrow^\infty$ and in particular $\alpha_k^{b_0} \not\subseteq \varphi_x^{b_0}$ for all $x > k+1$. Hence $T(b_0, \varphi_{k+1}^{b_0}[n]) \leq k+1$ for all $n \in \mathbb{N}$.

By Fact 15.2 we know that $0 \notin \mathrm{rng}(\varphi_0^{b_0})$ and so $\alpha_k^{b_0} 0 \not\subseteq \varphi_0^{b_0}$. Therefore $T(b_0, \varphi_{k+1}^{b_0}[n]) \neq 0$ for all $n \geq |\alpha_k^{b_0}|$.

If $k = 0$, we have already verified $T(b_0, \varphi_{k+1}^{b_0}[n]) = k+1$ for all $n \geq |\alpha_k^{b_0}|$. Now if $k > 0$ and $x \in \{1, \dots, k\}$ we know that $\varphi_x^{b_0} = \alpha_{x-1}^{b_0} 0^m \uparrow^\infty$ for some $m \in \mathbb{N}$. Since $\alpha_{x-1}^{b_0} \subset \alpha_k^{b_0}$ and $0 \notin \mathrm{rng}(\alpha_k^{b_0})$, this yields $\alpha_k^{b_0} 0 \not\subseteq \varphi_x^{b_0}$ for all $x \in \{1, \dots, k\}$. The only hypothesis remaining is $k+1$, i.e.

$$T(b_0, \varphi_{k+1}^{b_0}[n]) = \begin{cases} 0 & \text{if } n < |\alpha_k^{b_0}| \\ k+1 & \text{if } n \geq |\alpha_k^{b_0}| \end{cases}$$

for arbitrary $n \in \mathbb{N}$. As this results in just one mind change and the intermediate hypothesis 0 is consistent in the first steps of the learning process, we conclude $\mathcal{R}_{b_0} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(\lambda x.T(b_0, x)) \cap \mathrm{CONS}_{\varphi^{b_0}}(\lambda x.T(b_0, x))$.

Since $\mathrm{card}\, \{i \in \mathbb{N} \mid \varphi_i^{b_0} \in \mathcal{R}\} = 1$ by Fact 15.3, further cases cannot occur. This implies $\mathcal{R}_{b_0} \in (\mathrm{EX}_1)_{\varphi^{b_0}}(\lambda x.T(b_0, x)) \cap \mathrm{CONS}_{\varphi^{b_0}}(\lambda x.T(b_0, x))$ and we conclude $b_0 \in B$. *qed 1.*

*ad 2.* Because of Fact 15.3 it again suffices to regard two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\}$.
Then on $\varphi_0^{b_0}$ the strategy $\lambda x.S(b_0, x)$ changes its hypothesis infinitely often or returns a hypothesis incorrect with respect to $\tau$ infinitely often (see Fact 14.3). We obtain $\mathcal{R}_{b_0} \notin \mathrm{EX}_\tau(\lambda x.S(b_0, x))$ and in particular $\mathcal{R}_{b_0} \notin \mathrm{CEX}_\tau(\lambda x.S(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi_{k+1}^{b_0}\}$ with $k \in \mathbb{N}$.
With Fact 14.1 we have

$$\tau_{S(b_0,\varphi_{k+1}^{b_0}[n])} \subseteq \alpha_k^{b_0}\uparrow^\infty \subset \alpha_k^{b_0}0^\infty = \varphi_{k+1}^{b_0}$$

for some $n \in \mathbb{N}$. Hence $S(b_0, \varphi_{k+1}^{b_0}[n])$ is a $\tau$-number of a proper subfunction of $\varphi_{k+1}^{b_0}$ for this integer $n$. We conclude $\mathcal{R}_{b_0} \notin \mathrm{CEX}_\tau(\lambda x.S(b_0, x))$.

In both cases we have verified Claim 2. *qed 2.*

Altogether we see, that $\mathcal{R}_b \notin \mathrm{CEX}_\tau(\lambda x.S(b, x))$ for at least one $b \in B$, which contradicts (11). Thus our assumption is wrong, i.e. $B \notin \mathrm{suit}_\tau(J^1, \mathrm{CEX})$ and so $\mathrm{suit}_\varphi(J^1, \mathrm{EX}_1) \backslash \mathrm{suit}_\tau(J^1, \mathrm{CEX}) \neq \emptyset$ and $\mathrm{suit}_\varphi(J^1, \mathrm{CONS}) \backslash \mathrm{suit}_\tau(J^1, \mathrm{CEX}) \neq \emptyset$. □

As we have already mentioned above, Theorem 12 implies $\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL}) \subset \mathrm{suit}_\varphi(J^*, \mathrm{CONS})$ (and the same result for $\mathrm{suit}_\tau$). For the separation of the criteria CP and TOTAL we have to be more careful, as Proposition 4 shows.

**Proposition 4** *$suit_\varphi(CP, CP) = suit_\varphi(CP, TOTAL)$.*

*Proof.* Assume $B \in \mathrm{suit}_\varphi(\mathrm{CP}, \mathrm{TOTAL})$. Then there exists a strategy $S \in \mathcal{P}^2$ such that $\mathcal{R}_b \in \mathrm{TOTAL}_{\varphi^b}(\lambda x.S(b, x))$ for all $b \in B$. This implies for arbitrary $b \in B$:

- $\mathcal{R}_b \in \mathrm{EX}_{\varphi^b}(\lambda x.S(b, x))$,
- $\forall f \in \mathcal{R}_b \; \forall n \in \mathbb{N} \; [\varphi^b_{S(b,f[n])} \in \mathcal{R}]$.

Since $\varphi^b_{S(b,f[n])} \in \mathcal{R}$ implies $\varphi^b_{S(b,f[n])} \in \mathcal{R}_b$ (by definition of $\mathcal{R}_b$), we obviously have $B \in \mathrm{suit}_\varphi(\mathrm{CP}, \mathrm{CP})(S)$ and thus $\mathrm{suit}_\varphi(\mathrm{CP}, \mathrm{TOTAL}) \subseteq \mathrm{suit}_\varphi(\mathrm{CP}, \mathrm{CP})$. The opposite inclusion follows immediately. □

Still, if we consider uniform learning of finite classes with respect to our acceptable numbering $\tau$, the separation of CP and TOTAL holds by analogy with Theorem 2. For that purpose we first prove the stronger statement in Theorem 13.

**Theorem 13** $suit_\varphi(J^1, EX_0) \setminus suit_\tau(J^1, CP) \neq \emptyset$.

*Proof.* We define a strategy $T \in \mathcal{R}$ and choose $B \subseteq \mathbb{N}$ such that $T$ is suitable for uniform $\mathrm{EX}_0$-identification of the $\mathcal{R}$-cores described by $B$. These $\mathcal{R}$-cores will all consist of just one function and $B$ will not allow uniform CP-identification with respect to $\tau$.

Let $T \in \mathcal{R}$ be defined by

$$T(f[n]) := \begin{cases} ? & \text{if } n = 0 \\ 0 & \text{if } n > 0 \text{ and } f[n] = 0^{n+1} \\ 1 & \text{if } n > 0 \text{ and } f[n] = 010^{n-1} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } f \in \mathcal{R} \text{ and } n \in \mathbb{N} \ .$$

Furthermore, let

$$B := \{b \in \mathbb{N} \mid \text{card } \mathcal{R}_b = 1 \text{ and } \mathcal{R}_b \in (\mathrm{EX}_0)_{\varphi^b}(T)\} \ .$$

*Claim.* $B \in \mathrm{suit}_\varphi(J^1, \mathrm{EX}_0) \setminus \mathrm{suit}_\tau(J^1, \mathrm{CP})$.

*Proof of "$B \in suit_\varphi(J^1, EX_0)$".*
Defining $T' \in \mathcal{R}^2$ by $T'(b, f[n]) := T(f[n])$ for all $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ we obviously have

- $\forall b \in B\ [\mathcal{R}_b \in J^1]$,
- $\forall b \in B\ [\mathcal{R}_b \in (\mathrm{EX}_0)_{\varphi^b}(\lambda x.T'(b, x))]$.

By definition of uniform identifiability this implies $B \in \mathrm{suit}_\varphi(J^1, \mathrm{EX}_0)$.

*Proof of "$B \notin suit_\tau(J^1, CP)$".*
We will verify this claim by way of contradiction.
*Assumption.* $B \in \mathrm{suit}_\tau(J^1, \mathrm{CP})$.

Then there exists a strategy $S \in \mathcal{P}^2$ such that for any $b \in B$ the recursive core $\mathcal{R}_b$ is identified with class-preserving hypotheses with respect to $\tau$ by $\lambda x.S(b, x)$; abbreviated

$$\forall b \in B\ [\mathcal{R}_b \in \mathrm{CP}_\tau(\lambda x.S(b, x))] \ . \tag{12}$$

*Aim.* Construction of an integer $b_0$, such that

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0, x))$,

in contradiction to statement (12). The strategy $\lambda x.S(b_0, x)$ will fail for the only function $f \in \mathcal{R}_{b_0}$ by

- being undefined on some initial segment of $f$ or
- returning a $\tau$-index of a function not in $\mathcal{R}_{b_0}$ on input of some initial segment of $f$ (that means, our strategy makes a non-class-preserving guess on $f$).

*Construction of* $b_0$.
We define a function $\psi \in \mathcal{P}^3$ as follows: Let $b \in \mathbb{N}$. For $i \geq 2$ let $\psi_i^b :=\uparrow^\infty$. Furthermore, let $\psi_0^b$ be defined by

$$\begin{aligned}
\psi_0^b(0) &:= 0 \\
\psi_0^b(x+1) &:= \begin{cases} 0 & \text{if } S(b,0)\uparrow_{\leq x} \text{ or } [S(b,0)\downarrow_{\leq x} \text{ and } \tau_{S(b,0)}(1)\uparrow_{\leq x}] \text{ or} \\ & [S(b,0)\downarrow_{\leq x} \text{ and } \tau_{S(b,0)}(1)\downarrow_{\leq x} \text{ and } \tau_{S(b,0)}(1) \neq 0] \\ \uparrow & \text{if } S(b,0)\downarrow_{\leq x} \text{ and } \tau_{S(b,0)}(1)\downarrow_{\leq x} \text{ and } \tau_{S(b,0)}(1) = 0 \end{cases}
\end{aligned}$$

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

$$\begin{aligned}
\psi_0^b \in \mathcal{R} &\iff \psi_0^b = 0^\infty \\
&\iff S(b,\psi_0^b[0])\uparrow \text{ or } \tau_{S(b,\psi_0^b[0])}(1)\uparrow \text{ or } \tau_{S(b,0)}(1) > 0 = \psi_0^b(1)\ . \quad (13)
\end{aligned}$$

Finally, the function $\psi_1^b$ is defined by

$$\begin{aligned}
\psi_1^b(0) &:= 0 \\
\psi_1^b(x+2) &:= 0 \text{ for all } x \in \mathbb{N} \\
\psi_1^b(1) &:= \begin{cases} \overline{\tau_{S(b,0)}(1)} & \text{if } S(b,0)\downarrow \text{ and } \tau_{S(b,0)}(1)\downarrow \\ \uparrow & \text{if } S(b,0)\uparrow \text{ or } \tau_{S(b,0)}(1)\uparrow \end{cases}
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\psi_1^b \in \mathcal{R} &\iff \psi_1^b = 0(\overline{\tau_{S(b,\psi_1^b[0])}(1)})0^\infty \\
&\iff [S(b,\psi_0^b[0])\downarrow \text{ and } \tau_{S(b,\psi_0^b[0])}(1)\downarrow]\ . \quad (14)
\end{aligned}$$

Now let $g \in \mathcal{R}$ be a compiler function such that $\varphi^{g(b)} = \psi^b$ for all $b \in \mathbb{N}$. Such a function $g$ exists, since $\psi^b$ was defined uniformly in $b$. The recursion theorem then yields an integer $b_0 \in \mathbb{N}$ satisfying $\varphi^{b_0} = \varphi^{g(b_0)}$. Thus we have $\varphi^{b_0} = \psi^{b_0}$.

*End Construction* $b_0$.

In order to contradict phrase (12) we will prove the following statements.

1. $b_0 \in B$,
2. $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0,x))$.

*ad 1.* We observe $\mathcal{R}_{b_0} \subseteq \{\varphi_0^{b_0}, \varphi_1^{b_0}\}$ by construction. From (13) and (14) we conclude that either $\underline{\mathcal{R}_{b_0} = \{\varphi_0^{b_0}, \varphi_1^{b_0}\}} = \{0^\infty\}$ or $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\} = \{0^\infty\} \neq \{\varphi_1^{b_0}\}$ or $\mathcal{R}_{b_0} = \{\varphi_1^{b_0}\} = \{0(\tau_{S(b,\psi_1^b[0])}(1))0^\infty\} = \{010^\infty\}$, so either $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\} = \{0^\infty\}$ or $\mathcal{R}_{b_0} = \{\varphi_1^{b_0}\} = \{010^\infty\}$. So we have $\mathcal{R}_{b_0} \in J^1$. By definition of $T$ we know for all $n > 0$ that $T(0^n) = 0$ and $T(010^n) = 1$. Furthermore $T(0) =?$. Therefore $\mathcal{R}_{b_0} \in (\mathrm{EX}_0)_{\varphi^{b_0}}(T)$ and so $b_0 \in B$. *qed 1.*

*ad 2.* As we have seen in the proof of "$b_0 \in B$", it suffices to consider two cases.

*Case (i).* $\mathcal{R}_{b_0} = \{\varphi_0^{b_0}\} = \{0^\infty\}$.
Then (13) implies $S(b_0, \varphi_0^{b_0}[0])\uparrow$ or $\tau_{S(b_0,\varphi_0^{b_0}[0])}(1)\uparrow$ or $\tau_{S(b_0,\varphi_0^{b_0}[0])}(1) \neq 0 = \varphi_0^{b_0}(1)$.

If $S(b_0, \varphi_0^{b_0}[0])\uparrow$, then of course $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0, x))$.

If $S(b_0, \varphi_0^{b_0}[0])\downarrow$ and $\tau_{S(b_0,\varphi_0^{b_0}[0])}(1)\uparrow$, then in particular $S(b_0, \varphi_0^{b_0}[0])$ is a $\tau$-index of a non-total function. So $\lambda x.S(b_0, x)$ makes a non-class-preserving guess on $\varphi_0^{b_0}[0]$. This implies $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0, x))$.

If $S(b_0, \varphi_0^{b_0}[0])\downarrow$, $\tau_{S(b_0,\varphi_0^{b_0}[0])}(1)\downarrow$ and $\tau_{S(b_0,\varphi_0^{b_0}[0])}(1) \neq 0$, then $\tau_{S(b_0,\varphi_0^{b_0}[0])} \neq \varphi_0^{b_0}$ and thus $\tau_{S(b_0,\varphi_0^{b_0}[0])} \notin \mathcal{R}_{b_0}$. Again $\lambda x.S(b_0, x)$ makes a non-class-preserving guess on $\varphi_0^{b_0}[0]$. We obtain $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0, x))$.

*Case (ii).* $\mathcal{R}_{b_0} = \{\varphi_1^{b_0}\} = \{010^\infty\}$.
Then $S(b_0, \varphi_1^{b_0}[0])\downarrow$ and $\tau_{S(b_0,\varphi_1^{b_0}[0])}(1)\downarrow$, but $\varphi_1^{b_0}(1) \neq \tau_{S(b_0,\varphi_1^{b_0}[0])}(1)$ and therefore $\tau_{S(b_0,\varphi_1^{b_0}[0])} \notin \mathcal{R}_{b_0}$. So $\lambda x.S(b_0, x)$ makes a non-class-preserving guess on $\varphi_1^{b_0}[0]$. This yields $\mathcal{R}_{b_0} \notin \mathrm{CP}_\tau(\lambda x.S(b_0, x))$. *qed 2.*

Claims 1 and 2 together now contradict statement (12). Therefore our assumption must have been wrong, i.e. $B \notin \mathrm{suit}_\tau(J^1, \mathrm{CP})$. Altogether we have $B \in \mathrm{suit}_\varphi(J^1, \mathrm{EX}_0) \setminus \mathrm{suit}_\tau(J^1, \mathrm{CP})$. □

As we have promised before, this yields the proper separation of CP and TOTAL for uniform learning with respect to the hypothesis space $\tau$.

**Corollary 5** *$suit_\tau(J^1, CP) \subset suit_\tau(J^1, TOTAL)$.*

*Proof.* Obviously $\mathrm{suit}_\tau(J^1, \mathrm{CP}) \subseteq \mathrm{suit}_\tau(J^1, \mathrm{TOTAL})$. It remains to prove the existence of some description set $B \subseteq \mathbb{N}$ satisfying

$$B \in \mathrm{suit}_\tau(J^1, \mathrm{TOTAL}) \setminus \mathrm{suit}_\tau(J^1, \mathrm{CP}) \ .$$

From Theorem 13 we know that there is some description set $B$ in $\mathrm{suit}_\varphi(J^1, \mathrm{EX}_0)$ which is not contained in $\mathrm{suit}_\tau(J^1, \mathrm{CP})$. Since we can prove that $\mathrm{suit}_\varphi(J^1, \mathrm{EX}_0)$ is contained in $\mathrm{suit}_\tau(J^1, \mathrm{TOTAL})$, we have already verified our corollary. For the proof of $\mathrm{suit}_\varphi(J^1, \mathrm{EX}_0) \subseteq \mathrm{suit}_\tau(J^1, \mathrm{TOTAL})$ note that according to Proposition 1 all $\varphi^b$-numbers can be compiled to $\tau$-numbers uniformly in $b$. The output "?" just has to be replaced by some $\tau$-number of any total function. $\square$

## 4.4 The Hierarchies in Lemma 1.4, 1.5, 1.6

From Theorem 12 and $\mathrm{suit}_\varphi(J^*, \mathrm{CONS}) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{EX})$ (analogously for $\mathrm{suit}_\tau$) we obtain that the second inclusion $\mathrm{suit}_\varphi(J^*, \mathrm{CEX}) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{EX})$ in Lemma 1.4 and its $\tau$-version are indeed proper.

Together with the fact $\mathrm{suit}_\tau(J^*, \mathrm{TOTAL}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{CONF})$ Theorem 10 yields $\mathrm{suit}_\varphi(J^*, \mathrm{CEX}) \setminus \mathrm{suit}_\tau(J^*, \mathrm{TOTAL}) \neq \emptyset$ and in particular

$$\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL}) \subset \mathrm{suit}_\varphi(J^*, \mathrm{CEX}) \ ,$$

where again $\mathrm{suit}_\varphi$ may be replaced by $\mathrm{suit}_\tau$.

The three inclusions given in Lemma 1.5 and 1.6 are also proper inclusions. We have verified this implicitly in Theorem 9. The set $B$ used to separate $\mathrm{EX}_1$ from $\mathrm{EX}_0$ is suitable for uniform identification by enumeration with respect to total numberings $\varphi^b$. Since identification by enumeration with respect to total recursive numberings always implies identification with total and consistent intermediate hypotheses, we know that the same set $B$ belongs to $\mathrm{suit}_\varphi(J^*, \mathrm{CONS})$ and $\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL})$ (which is a subset of $\mathrm{suit}_\varphi(J^*, \mathrm{CEX})$ and $\mathrm{suit}_\tau(J^*, \mathrm{TOTAL})$), but not to $\mathrm{suit}(J^*, \mathrm{EX}_0)$.

**Corollary 6**

1. *$suit_\varphi(J^*, TOTAL) \subset suit_\varphi(J^*, CEX) \subset suit_\varphi(J^*, EX)$ (analogously with $\tau$ instead of $\varphi$),*
2. *$suit_\varphi(J^*, EX_0) \subset suit_\varphi(J^*, CONS)$ and $suit_\varphi(J^*, EX_0) \subset suit_\varphi(J^*, CEX)$ (analogously with $\tau$ instead of $\varphi$),*
3. *$suit_\tau(J^*, EX_0) \subset suit_\tau(J^*, TOTAL)$.*

## 4.5 Incomparable Classes

The scope of this subsection is to find pairs of inference criteria which do not yield such inclusions as in Lemma 1. That means we want to collect examples of incomparable classes.

In Theorems 10 and 12 we have verified $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \setminus \mathrm{suit}(J^*, \mathrm{CONF}) \neq \emptyset$ and $\mathrm{suit}_\varphi(J^*, \mathrm{EX}_1) \setminus \mathrm{suit}_\tau(J^*, \mathrm{CEX}) \neq \emptyset$. From $\mathrm{suit}_\tau(J^*, \mathrm{CP}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{CEX})$ and $\mathrm{suit}_\tau(J^*, \mathrm{CP}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{TOTAL}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{CONS}) \subseteq \mathrm{suit}_\tau(J^*, \mathrm{CONF})$ we conclude

$$\mathrm{suit}_\varphi(J^*, \mathrm{EX}_m) \setminus \mathrm{suit}_\tau(J^*, I) \neq \emptyset$$

for all $m \geq 1$ and $I \in \{\mathrm{CP}, \mathrm{TOTAL}, \mathrm{CONS}, \mathrm{CONF}, \mathrm{CEX}\}$. By the same reasoning we obtain

$$\mathrm{suit}_\varphi(J^*, I) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m) \neq \emptyset$$

for all $m \in \mathbb{N}$ and $I \in \{\mathrm{CP}, \mathrm{TOTAL}, \mathrm{CONS}, \mathrm{CONF}, \mathrm{CEX}\}$ with Corollary 2. This yields the following corollary.

**Corollary 7** *Let $I \in \{CP, TOTAL, CONS, CONF, CEX\}$. Then*

$$suit_\varphi(J^*, EX_m) \; \# \; suit_\varphi(J^*, I)$$

*for all $m \geq 1$. The same result holds if we replace $suit_\varphi$ by $suit_\tau$.*

Furthermore we can use Theorem 13 to verify a stronger result for learning with class-preserving or total intermediate hypotheses.

**Corollary 8**

1. *$suit_\varphi(J^*, EX_0) \; \# \; suit_\varphi(J^*, CP)$,*
2. *$suit_\tau(J^*, EX_0) \; \# \; suit_\tau(J^*, CP)$,*
3. *$suit_\varphi(J^*, EX_0) \; \# \; suit_\varphi(J^*, TOTAL)$.*

*Proof.* Properties 1 and 2 are direct consequences of Theorem 13. By Proposition 4 we know that $\mathrm{suit}_\varphi(J^*, \mathrm{TOTAL})$ equals $\mathrm{suit}_\varphi(J^*, \mathrm{CP})$, so property 3 is verified with property 1. □

With Theorem 10 and 12 we have also verified the following corollary.

**Corollary 9**

1. *$suit_\varphi(J^*, CEX) \; \# \; suit_\varphi(J^*, CONS)$, $suit_\tau(J^*, CEX) \; \# \; suit_\tau(J^*, CONS)$.*
2. *$suit_\varphi(J^*, CEX) \; \# \; suit_\varphi(J^*, CONF)$, $suit_\tau(J^*, CEX) \; \# \; suit_\tau(J^*, CONF)$.*

## 4.6 Summary

Now we can summarize our separation results for uniform learning of finite classes with respect to fixed hypothesis spaces.

**Summary 1**

1. *$suit_\varphi(J^*, EX) \subset suit_\varphi(J^*, BC) \subset suit_\varphi(J^*, BC^*)$*
   *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*
2. *$suit_\varphi(J^*, EX_m) \subset suit_\varphi(J^*, EX_{m+1}) \subset suit_\varphi(J^*, EX)$ for all $m \in \mathbb{N}$*
   *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

3. $suit_\varphi(J^*, TOTAL) \subset suit_\varphi(J^*, CONS) \subset suit_\varphi(J^*, CONF) \subset suit_\varphi(J^*, EX)$ *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

4. $suit_\varphi(J^*, TOTAL) \subset suit_\varphi(J^*, CEX) \subset suit_\varphi(J^*, EX)$ *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

5. $suit_\varphi(J^*, CP) = suit_\varphi(J^*, TOTAL)$,

6. $suit_\tau(J^*, CP) \subset suit_\tau(J^*, TOTAL)$,

7. $suit_\varphi(J^*, EX_0) \subset suit_\varphi(J^*, CONS)$, $suit_\varphi(J^*, EX_0) \subset suit_\varphi(J^*, CEX)$ *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

8. $suit_\varphi(J^*, CEX) \;\#\; suit_\varphi(J^*, CONS)$, $suit_\varphi(J^*, CEX) \;\#\; suit_\varphi(J^*, CONF)$ *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

9. *if* $I \in \{CP, TOTAL, CONS, CONF, CEX\}$ *and* $m \geq 1$, *then* $suit_\varphi(J^*, EX_m) \;\#\; suit_\varphi(J^*, I)$ *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

10. $suit_\varphi(J^*, EX_0) \;\#\; suit_\varphi(J^*, CP)$, *(holds analogously if we substitute $suit_\varphi$ by $suit_\tau$),*

11. $suit_\varphi(J^*, EX_0) \;\#\; suit_\varphi(J^*, TOTAL)$, $suit_\tau(J^*, EX_0) \subset suit_\tau(J^*, TOTAL)$.

Thus we have transferred the comparison results of Theorem 2 to the concept of meta-learning in fixed hypothesis spaces. Each separation is achieved already by restricting ourselves to the synthesis of strategies for finite classes of recursive functions. Note that none of the results in Summary 1 is formulated as strictly as possible. We never really needed the whole class $J^*$ to prove these separations. Often it was even enough to choose recursive cores from $J^1$, for example

- in Summary 1.1: $\mathrm{suit}_\varphi(J^1, \mathrm{BC}) \subset \mathrm{suit}_\varphi(J^1, \mathrm{BC}^*)$, $\mathrm{suit}_\tau(J^1, \mathrm{BC}) \subset \mathrm{suit}_\tau(J^1, \mathrm{BC}^*)$ (see Theorem 7);

- in Summary 1.3: $\mathrm{suit}_\varphi(J^1, \mathrm{TOTAL}) \subset \mathrm{suit}_\varphi(J^1, \mathrm{CONS})$, $\mathrm{suit}_\tau(J^1, \mathrm{TOTAL}) \subset \mathrm{suit}_\tau(J^1, \mathrm{CONS})$ (see Theorem 12 and remarks above);

- in Summary 1.4: $\mathrm{suit}_\varphi(J^1, \mathrm{CEX}) \subset \mathrm{suit}_\varphi(J^1, \mathrm{EX})$, $\mathrm{suit}_\tau(J^1, \mathrm{CEX}) \subset \mathrm{suit}_\tau(J^1, \mathrm{EX})$ (see Theorem 12);

- in Summary 1.6: $\mathrm{suit}_\tau(J^1, \mathrm{CP}) \subset \mathrm{suit}_\tau(J^1, \mathrm{TOTAL})$ (see Corollary 5);

- in Summary 1.8: $\mathrm{suit}_\varphi(J^1, \mathrm{CONS}) \not\subseteq \mathrm{suit}_\varphi(J^1, \mathrm{CEX})$, $\mathrm{suit}_\tau(J^1, \mathrm{CONS}) \not\subseteq \mathrm{suit}_\tau(J^1, \mathrm{CEX})$ and the same results with CONF instead of CONS (see Theorem 12);

- in Summary 1.10: $\text{suit}_\varphi(J^1, \text{EX}_0) \not\subseteq \text{suit}_\varphi(J^1, \text{CP})$, $\text{suit}_\tau(J^1, \text{EX}_0) \not\subseteq \text{suit}_\tau(J^1, \text{CP})$ (see Theorem 13);
- in Summary 1.11: $\text{suit}_\varphi(J^1, \text{EX}_0) \not\subseteq \text{suit}_\varphi(J^1, \text{TOTAL})$ (see Theorem 13 and Proposition 4).

But still these results are not as strict as possible, because in the corresponding proofs we did not use all descriptions of singleton recursive cores. A further goal of research will be to find out more about the nature of the appropriate description sets for the separation of inference criteria in uniform learning. Perhaps we can characterize a kind of "smallest description set" witnessing the separation of two identification criteria. Our observations above suggest, that for the inference criteria CP, TOTAL, CEX, CONS, CONF and EX each pairwise separation concerning $\text{suit}_\varphi$ or $\text{suit}_\tau$ can already be achieved by restricting ourselves to descriptions of recursive cores in $J^1$. That would imply that a "smallest description set" was somewhere "below" the set of all descriptions of singleton sets of recursive functions. For uniform learning with respect to general hypothesis spaces (i.e. learning according to suit without subscript) in general singleton sets are not sufficient to prove any separations, because $\text{suit}(J^1, I) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \in J^1 \text{ for all } b \in B\}$ for any criterion $I \in \mathcal{I}$ (cf. Theorem 3).

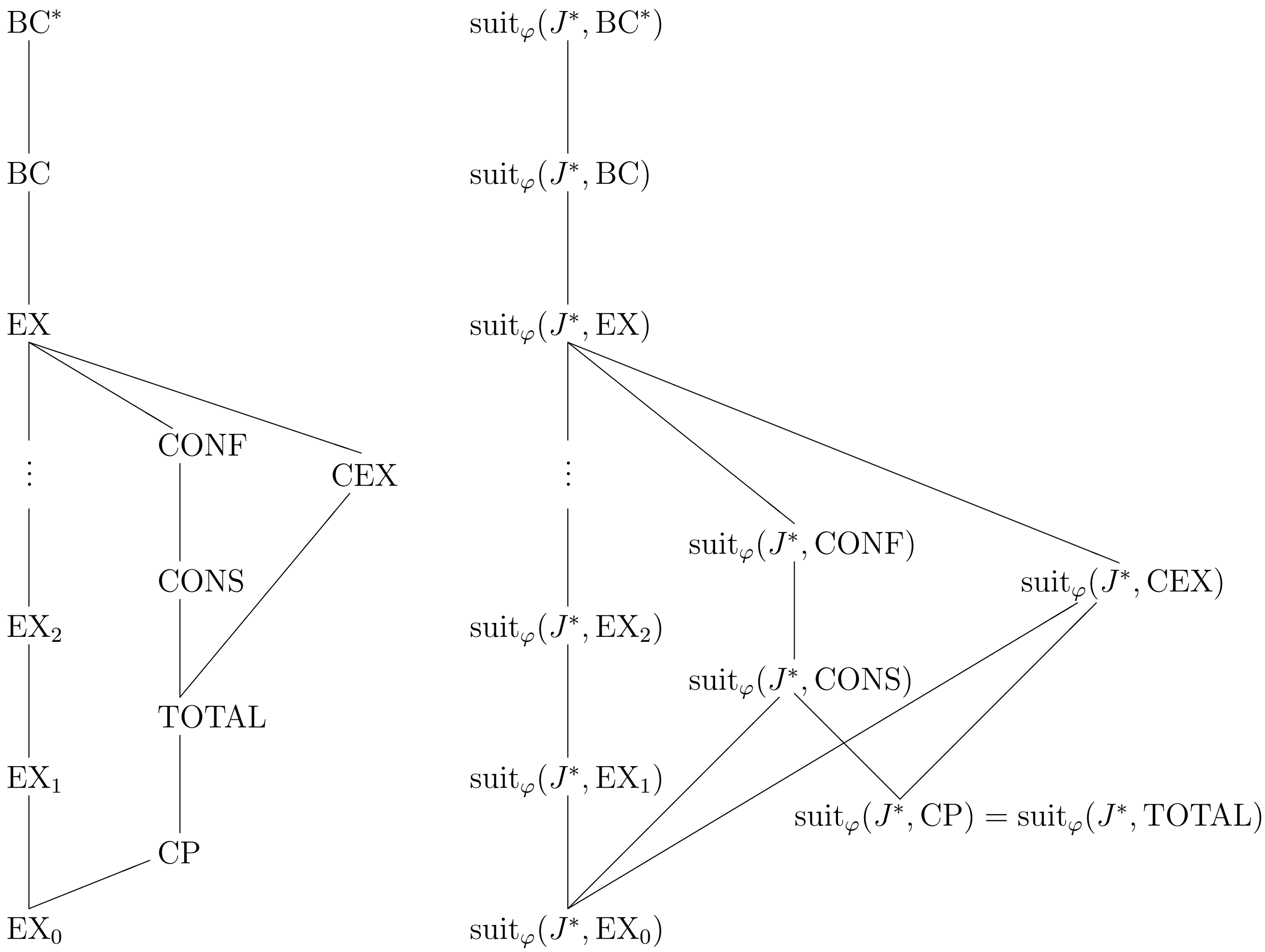


Figure 2: *The hierarchy of inference criteria according to Theorem 2 compared with the corresponding hierarchy for uniform learning with respect to the numberings $\varphi^b$, $b \in B$. Any line drawn upwards indicates a proper inclusion. If two classes are not connected by a line or a sequence of lines drawn upwards, they are incomparable.*

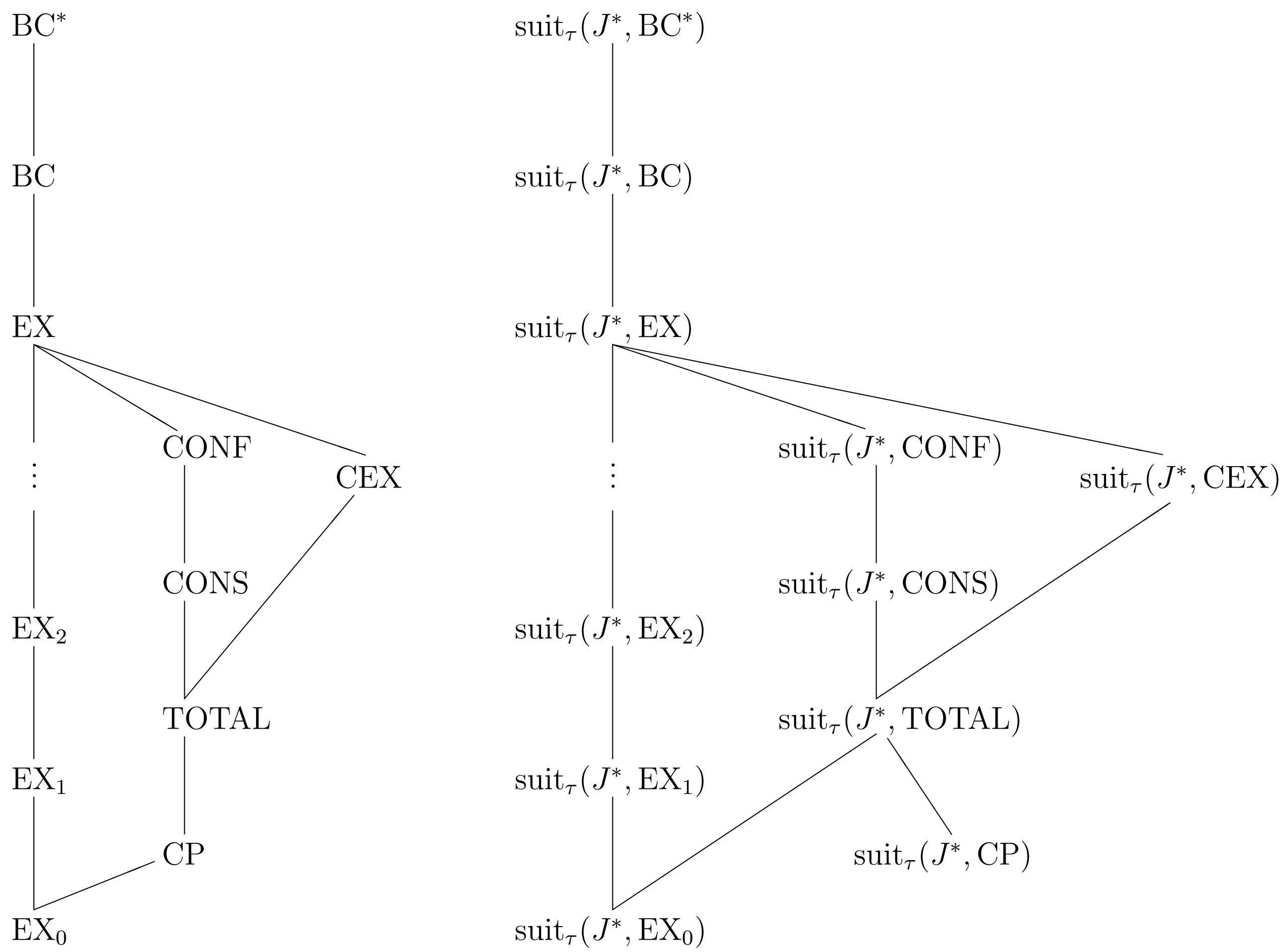


Figure 3: *The hierarchy of inference criteria according to Theorem 2 compared with the corresponding hierarchy for uniform learning with respect to the acceptable numbering $\tau$. Any line drawn upwards indicates a proper inclusion. If two classes are not connected by a line or a sequence of lines drawn upwards, they are incomparable.*

# 5 Separation of Inference Criteria: General Hypothesis Spaces

In this section we investigate the hierarchies of inference criteria for uniform learning without restrictions in the choice of the hypothesis spaces. Again we will concentrate on description sets corresponding to collections of finite classes of recursive functions. Some of the comparison results in Section 4 hold analogously for this concept, but there are differences, too.

**Lemma 2**

1. *suit*$(J^*, \mathit{EX}) \subseteq$ *suit*$(J^*, \mathit{BC}) \subseteq$ *suit*$(J^*, \mathit{BC}^*)$,
2. *suit*$(J^*, \mathit{EX}_m) \subseteq$ *suit*$(J^*, \mathit{EX}_{m+1}) \subseteq$ *suit*$(J^*, \mathit{EX})$ *for all* $m \in \mathbb{N}$,
3. *suit*$(J^*, \mathit{CONS}) \subseteq$ *suit*$(J^*, \mathit{CONF}) \subseteq$ *suit*$(J^*, \mathit{EX})$,
4. *suit*$(J^*, \mathit{CP}) \subseteq$ *suit*$(J^*, \mathit{TOTAL}) \subseteq$ *suit*$(J^*, \mathit{CEX}) \subseteq$ *suit*$(J^*, \mathit{EX})$,
5. *suit*$(J^*, \mathit{EX}_0) \subseteq$ *suit*$(J^*, \mathit{CP})$ *and* *suit*$(J^*, \mathit{EX}_0) \subseteq$ *suit*$(J^*, \mathit{CONS})$.

*Proof.* The results in Lemma 2.1, 2.2, 2.3 and 2.4 are direct consequences of the definitions. In order to verify Lemma 2.5, fix $B \in \mathrm{suit}(J^*, \mathrm{EX}_0)$. Then there is some strategy $S \in \mathcal{R}^2$, such that for every $b \in B$ there exists a numbering $\psi^{[b]}$ satisfying $\mathcal{R}_b \in (\mathrm{EX}_0)_{\psi^{[b]}}(\lambda x.S(b,x))$.

For the proof of $B \in \mathrm{suit}(J^*, \mathrm{CP})$ we change the numberings $\psi^{[b]}$ by copying some elements of the recursive cores into fixed numbers of the hypothesis spaces. Formally, for each $b \in B$ we define a new numbering $\eta^{[b]}$ as follows: let $i_b \in \mathbb{N}$ be a $\psi^{[b]}$-number of any function in the recursive core $\mathcal{R}_b$, i.e. $\psi^{[b]}_{i_b} \in \mathcal{R}_b$ (we can assume without loss of generality that $\mathcal{R}_b \neq \emptyset$).

$$\begin{aligned} \eta^{[b]}_0 &:= \psi^{[b]}_{i_b} , \\ \eta^{[b]}_{x+1} &:= \psi^{[b]}_x \text{ for all } x \in \mathbb{N} . \end{aligned}$$

Now the hypothesis 0 can be used to replace the "?" returned by the strategy, because it is class-preserving with respect to the new hypothesis space. So we define a strategy $S' \in \mathcal{P}^2$ for $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ by

$$S'(b, f[n]) := \begin{cases} 0 & \text{if } S(b, f[n]) = ? \\ S(b, f[n]) + 1 & \text{otherwise} \end{cases}$$

By construction of our numberings $\eta^{[b]}$ this implies $\mathcal{R}_b \in \mathrm{CP}_{\eta^{[b]}}(\lambda x.S'(b,x))$ for all $b \in B$. Therefore $B$ is suitable for uniform class-preserving identification.

For the proof of $B \in \mathrm{suit}(J^*, \mathrm{CONS})$ we also have to change our hypothesis spaces $\psi^{[b]}$. We will enable consistent identification by mixing the old numbering with a new numbering of all recursive functions of finite support. The numbers of these functions will code their own initial segments representing their finite support, so they can be used as consistent intermediate hypotheses while our old strategy returns "?". For each $b \in B$ we define a numbering $\zeta^{[b]}$ as follows: for any $x \in \mathbb{N}$ let

$$\begin{aligned} \zeta^{[b]}_{2x} &:= \mathrm{cod}^{-1}(x)0^\infty \ , \\ \zeta^{[b]}_{2x+1} &:= \psi^{[b]}_x \ . \end{aligned}$$

Now a consistent uniform learner $S''$ can be defined by

$$S''(b, f[n]) := \begin{cases} 2f[n] & \text{if } S(b, f[n]) = ? \\ 2S(b, f[n]) + 1 & \text{otherwise} \end{cases}$$

for $f \in \mathcal{R}$ and $b, n \in \mathbb{N}$ (note that $f[n] = \mathrm{cod}(f(0), \ldots, f(n))$, so $\zeta^{[b]}_{2f[n]} = f(0)f(1)\ldots f(n)0^\infty =_n f$). We obtain $\mathcal{R}_b \in \mathrm{CONS}_{\zeta^{[b]}}(\lambda x.S''(b, x))$ for all $b \in B$. Thus $B \in \mathrm{suit}(J^*, \mathrm{CONS})$. $\square$

Note that we dropped the inclusion for TOTAL-identification in the third line. As in general a uniform strategy $S$ satisfying $B \in \mathrm{suit}(J^*, \mathrm{TOTAL})(S)$ for some $B \subseteq \mathbb{N}$ can *not* synthesize an appropriate hypothesis space for $\mathcal{R}_b$ from $b \in B$, the hypotheses returned by $S$ cannot be checked for consistency. Therefore the proof of Lemma 1 cannot be transferred. In the following subsections we will try to find out, which of the results in Lemma 2 can be written with proper inclusion symbols.

## 5.1 The Hierarchies in Lemma 2.1, 2.2, 2.3

In Section 4 we have seen that $\mathrm{suit}_\varphi(J^*, \mathrm{BC})$ is a proper subset of $\mathrm{suit}_\varphi(J^*, \mathrm{BC}^*)$. The same result is obtained for uniform learning with respect to the acceptable numbering $\tau$. Without these strict demands concerning the hypothesis spaces we observe a difference in our hierarchies.

**Theorem 14** *suit*$(J^*, BC) =$ *suit*$(J^*, BC^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b$ *is finite for all* $b \in B\}$.

*Proof.* As the whole class $\mathcal{R}$ can be behaviourally correctly identified with anomalies (cf. Theorem 2), we obtain $\mathbb{N} \in \mathrm{suit}(\mathrm{BC}^*, \mathrm{BC}^*)$ and in particular

$$\mathrm{suit}(J^*, \mathrm{BC}^*) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \text{ is finite for all } b \in B\} \ .$$

Thus it remains to prove $\mathrm{suit}(J^*, \mathrm{BC}) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b$ is finite for all $b \in B\}$. But this fact follows directly from $\mathrm{suit}(\mathrm{BC}, \mathrm{BC}) = \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \in \mathrm{BC}$ for all $b \in B\}$, which is a result from [Zi00]. $\square$

Although for uniform learning behaviourally correct identification is already sufficient to cope with any description of a finite class (so anomalies cannot increase the learning power), BC and BC$^*$ are not equal uniform learning criteria in the general case. This follows trivially, since one might choose classes not in BC for the recursive cores to be identified.

For explanatory and behaviourally correct identification the hierarchy remains the same, as has already been verified in Theorem 8. Furthermore the hierarchy in Lemma 2.2 – resulting in bounds on the number of mind changes – consists of proper inclusions. This is verified with the strong result in Theorem 9. Because of Corollary 4 and Theorem 11 the inclusions in Lemma 2.3 are also proper. So we can summarize our results in the following corollary.

**Corollary 10**

1. $\mathit{suit}(J^*, \mathit{EX}) \subset \mathit{suit}(J^*, \mathit{BC}) = \mathit{suit}(J^*, \mathit{BC}^*)$
$= \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \textit{ is finite for all } b \in B\}$,
2. $\mathit{suit}(J^*, \mathit{EX}_m) \subset \mathit{suit}(J^*, \mathit{EX}_{m+1}) \subset \mathit{suit}(J^*, \mathit{EX})$ *for all* $m \in \mathbb{N}$,
3. $\mathit{suit}(J^*, \mathit{CONS}) \subset \mathit{suit}(J^*, \mathit{CONF}) \subset \mathit{suit}(J^*, \mathit{EX})$.

## 5.2 The Hierarchy in Lemma 2.4

For uniform learning of finite classes of recursive functions we observe a change in our hierarchy, if we do not fix the hypothesis spaces in advance. As Theorem 15 states, the hierarchy of the criteria CP, TOTAL, CEX and EX collapses in this case.

**Theorem 15** $\mathit{suit}(J^*, \mathit{CP}) = \mathit{suit}(J^*, \mathit{TOTAL}) = \mathit{suit}(J^*, \mathit{CEX}) = \mathit{suit}(J^*, \mathit{EX})$.

*Proof.* Since $\mathrm{suit}(J^*, \mathrm{CP}) \subseteq \mathrm{suit}(J^*, \mathrm{TOTAL}) \subseteq \mathrm{suit}(J^*, \mathrm{CEX}) \subseteq \mathrm{suit}(J^*, \mathrm{EX})$ by definition, it remains to prove

$$\mathrm{suit}(J^*, \mathrm{EX}) \subseteq \mathrm{suit}(J^*, \mathrm{CP}) \ .$$

For that purpose fix a description set $B \in \mathrm{suit}(J^*, \mathrm{EX})$. Then we know

1. $\mathcal{R}_b$ is finite for all $b \in B$,
2. there is a strategy $S \in \mathcal{P}^2$ such that for any $b \in B$ there is a hypothesis space $\psi^{[b]} \in \mathcal{P}^2$ satisfying $\mathcal{R}_b \in \mathrm{EX}_{\psi^{[b]}}(\lambda x.S(b, x))$.

Note that the hypothesis spaces $\psi^{[b]}$ do not have to be computable uniformly in $b$. Now we want to prove that $B \in \mathrm{suit}(J^*, \mathrm{CP})$. We even will see that our given strategy $S$ is already an appropriate strategy for uniform CP-identification from

$B$. This requires a change of the hypothesis spaces $\psi^{[b]}$ for $b \in B$.

*Idea.* Assume $b \in B$ was fixed. Since $\lambda x.S(b,x)$ identifies the finite class $\mathcal{R}_b$ in the limit, there are only finitely many initial segments of functions in $\mathcal{R}_b$ which force the strategy $\lambda x.S(b,x)$ into a "non-class-preserving" guess. If we replace the functions in $\psi^{[b]}$ associated with these non-class-preserving guesses by an element of $\mathcal{R}_b$, we obtain a hypothesis space appropriate for TOTAL-identification of $\mathcal{R}_b$ by $\lambda x.S(b,x)$.

*More formally:* Fix $b \in B$. From statement 2 we obtain

$$\mathrm{card}\ \{n \in \mathbb{N} \mid \psi^{[b]}_{S(b,f[n])} \notin \mathcal{R}_b\} < \infty$$

for all $f \in \mathcal{R}_b$. Defining the set of "forbidden" hypotheses on "relevant" initial segments by

$$H^{[b]} := \{i \in \mathbb{N} \mid \psi^{[b]}_i \notin \mathcal{R}_b \wedge \exists f \in \mathcal{R}_b\ \exists n \in \mathbb{N}\ [S(b,f[n]) = i]\}\ ,$$

we conclude with statement 1, that $H^{[b]}$ is finite. Now we define a new hypothesis space $\eta^{[b]}$ by

$$\eta^{[b]}_i := \begin{cases} \psi^{[b]}_i & \text{if } i \notin H^{[b]} \\ g & \text{if } i \in H^{[b]} \end{cases} \text{ for all } i \in \mathbb{N}\ ,$$

where $g \in \mathcal{R}_b$ is an arbitrary function in the recursive core described by $b$. Since $\psi^{[b]} \in \mathcal{P}^2$ and $H^{[b]}$ is finite, $\eta^{[b]}$ is computable.
Then $\mathcal{R}_b \in \mathrm{CP}_{\eta^{[b]}}(\lambda x.S(b,x))$ by definition of $\eta^{[b]}$. As $b \in B$ was chosen arbitrarily, we conclude $B \in \mathrm{suit}(J^*, \mathrm{CP})$. □

**Corollary 11** *suit*$(J^*$, *CONS*$) \subset$ *suit*$(J^*$, *CONF*$) \subset$ *suit*$(J^*$, *CP*$)$.

*Proof.* This fact follows immediately from Theorem 10 and Theorem 11 and by the result $\mathrm{suit}(J^*, \mathrm{EX}) = \mathrm{suit}(J^*, \mathrm{CP})$ in Theorem 15. □

Obviously, a further change in the hierarchies of inference criteria is witnessed by the fact $\mathrm{suit}(J^*, \mathrm{CONS}) \subset \mathrm{suit}(J^*, \mathrm{CEX})$, which follows by the same argumentation as in the proof of Corollary 11.

## 5.3 The Hierarchy in Lemma 2.5

A strict version of the first inclusion in Lemma 2.5 now follows immediately from our observations in Corollary 10 and Theorem 15. Since $\mathrm{suit}(J^*, \mathrm{EX}_0) \subset \mathrm{suit}(J^*, \mathrm{EX})$ by Corollary 10.2 and $\mathrm{suit}(J^*, \mathrm{EX}) = \mathrm{suit}(J^*, \mathrm{CP})$ by Theorem 15, we conclude

$$\mathrm{suit}(J^*, \mathrm{EX}_0) \subset \mathrm{suit}(J^*, \mathrm{CP})\ .$$

A proof of $\mathrm{suit}(J^*, \mathrm{EX}_0) \subset \mathrm{suit}(J^*, \mathrm{CONS})$ can be deduced from Corollary 2 as follows: as $\mathrm{suit}_\varphi(J^*, \mathrm{CP}) \subseteq \mathrm{suit}_\varphi(J^*, \mathrm{CONS})$ and $\mathrm{suit}_\varphi(J^*, \mathrm{CP}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_0) \neq \emptyset$, we know that $\mathrm{suit}(J^*, \mathrm{CONS}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_0) \neq \emptyset$. Together with Lemma 2.5 we obtain

$$\mathrm{suit}(J^*, \mathrm{EX}_0) \subset \mathrm{suit}(J^*, \mathrm{CONS}) .$$

**Corollary 12**

1. *$suit(J^*, EX_0) \subset suit(J^*, CP)$,*
2. *$suit(J^*, EX_0) \subset suit(J^*, CONS)$.*

## 5.4 Incomparable Classes

There are only a few incomparabilities remaining in uniform learning with general hypothesis spaces: CONF as well as CONS are still incomparable to the criteria resulting in mind change bounds (where at least one mind change is allowed).

If $m \geq 1$ is chosen arbitrarily, then $\mathrm{suit}(J^*, \mathrm{EX}_m) \setminus \mathrm{suit}(J^*, \mathrm{CONF}) \neq \emptyset$ and $\mathrm{suit}(J^*, \mathrm{EX}_m) \setminus \mathrm{suit}(J^*, \mathrm{CONS}) \neq \emptyset$ follow immediately from Theorem 10. For the opposite direction, note that $\mathrm{suit}(J^*, \mathrm{CONS}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m) \neq \emptyset$ is a consequence of Corollary 2. This also yields $\mathrm{suit}(J^*, \mathrm{CONF}) \setminus \mathrm{suit}(J^*, \mathrm{EX}_m) \neq \emptyset$. Altogether we obtain the following corollary.

**Corollary 13**

1. *$suit(J^*, EX_m) \;\#\; suit(J^*, CONF)$ for all $m \geq 1$,*
2. *$suit(J^*, EX_m) \;\#\; suit(J^*, CONS)$ for all $m \geq 1$.*

## 5.5 Summary

Finally we can summarize our separation results for uniform identification of finite classes without any restrictions concerning the hypothesis spaces.

**Summary 2**

1. *$suit(J^*, EX) \subset suit(J^*, BC) = suit(J^*, BC^*)$*
   *$= \{B \subseteq \mathbb{N} \mid \mathcal{R}_b \text{ is finite for all } b \in B\}$,*
2. *$suit(J^*, EX_m) \subset suit(J^*, EX_{m+1}) \subset suit(J^*, EX)$ for all $m \in \mathbb{N}$,*
3. *$suit(J^*, EX_0) \subset suit(J^*, CONS) \subset suit(J^*, CONF) \subset suit(J^*, CP)$*
   *$= suit(J^*, TOTAL) = suit(J^*, CEX) = suit(J^*, EX)$,*
4. *if $I \in \{CONS, CONF\}$ and $m \geq 1$, then $suit(J^*, EX_m) \;\#\; suit(J^*, I)$.*

So in contrast to uniform learning of finite classes with respect to fixed hypothesis spaces just a few of the separations in Theorem 2 can be transferred to the unrestricted concept of uniform learning. Still it is remarkable, how many inference criteria for uniform identification can be separated by collections of finite classes – even with very strong results (cf. the remarks below Theorem 8). But similar considerations as in Section 4 lead us to the observation that again our results are not as strict as possible. For none of these separations the whole class $J^*$ was necessary; in most comparisons a subset of the set of all descriptions of recursive cores consisting of up to two elements was sufficient. To verify this, note that in the associated proofs the specially constructed description $b_0$ corresponds to a set of at most two recursive functions. In these cases the special classes $B$ of descriptions witnessing the separations might have been restricted to classes of descriptions of recursive cores of up to two elements. If we set $J^2 := \{U \subseteq \mathcal{R} \mid \text{card } U \leq 2\}$, we obtain for example

- in Summary 2.1: $\text{suit}(J^2, \text{EX}) \subset \text{suit}(J^2, \text{BC})$, where $\text{suit}(J^2, \text{BC})$ equals the set $\{B \subseteq \mathbb{N} \mid \text{card } \mathcal{R}_b \leq 2 \text{ for all } b \in B\}$ (see the proof of Theorem 8);
- in Summary 2.3: $\text{suit}(J^2, \text{EX}_0) \subset \text{suit}(J^2, \text{CONS}) \subset \text{suit}(J^2, \text{CONF})$ (see the proofs of Theorem 9 and Theorem 11).

For the verification of $\text{suit}(J^*, \text{CONF}) \subset \text{suit}(J^*, \text{EX})$ and $\text{suit}(J^*, \text{EX}_m) \not\subseteq \text{suit}(J^*, \text{CONS})$, as well as $\text{suit}(J^*, \text{EX}_m) \not\subseteq \text{suit}(J^*, \text{CONF})$ for all $m \geq 1$, descriptions of recursive cores of at most three elements were sufficient (see the proof of Theorem 10). But perhaps this number of elements might be reduced to 2, if learning with convergently incorrect intermediate hypotheses was not involved in the proof of Theorem 10. Finally, in order to verify the results $\text{suit}(J^*, \text{EX}_m) \subset \text{suit}(J^*, \text{EX}_{m+1})$ and $\text{suit}(J^*, \text{CONS}) \not\subseteq \text{suit}(J^*, \text{EX}_m)$, $\text{suit}(J^*, \text{CONF}) \not\subseteq \text{suit}(J^*, \text{EX}_m)$ for all $m \in \mathbb{N}$, we might restrict ourselves to recursive cores consisting of no more than $2^{m+1}$ functions (see the construction in the proof of Theorem 9). Of course, in many of our results the description sets might still be reduced further without violating the conditions of separations. As has already been mentioned in Section 4, it might be interesting to learn more about the structure of description sets fit for our separations, and thus perhaps to find something like "smallest description sets" to be used as witnesses for the results in Summary 2. We might conjecture, that for uniform learning according to the definition of suit (without subscript) such smallest descriptions sets are in general "bigger" than for uniform learning with respect to $\text{suit}_\tau$ or $\text{suit}_\varphi$: in many results concerning $\text{suit}_\tau$ and $\text{suit}_\varphi$ descriptions of singleton recursive cores are sufficient (cf. the remarks below Summary 1 in Section 4), whereas – by Theorem 3 – such descriptions can never be enough to separate any of our identification criteria in the context of suit.

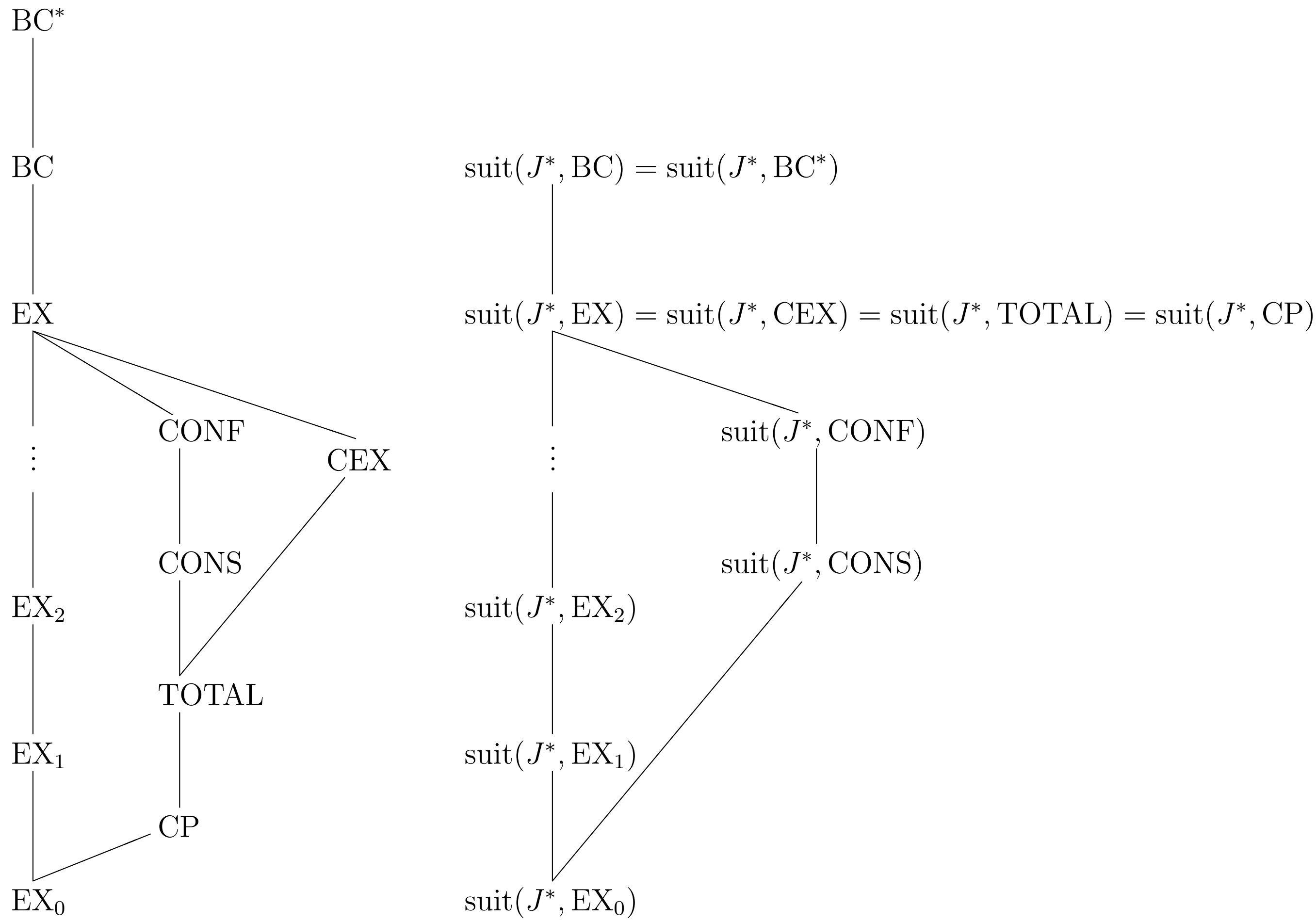


Figure 4: *The hierarchy of inference criteria according to Theorem 2 compared with the corresponding hierarchy for uniform learning without specifying the hypothesis spaces in advance. Any line drawn upwards indicates a proper inclusion. If two classes are not connected by a line or a sequence of lines drawn upwards, they are incomparable.*

## References


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

[Ba74a] Barzdins, 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] Barzdins, 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.

[Fu88] Fulk, M.A. (1988); *Saving the Phenomena: Requirements that Inductive Inference Machines Not Contradict Known Data*, Information and Computation 79, 193-209.

[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 17, 465-484.

[KB92] Kapur, S.; Bilardi, G. (1992); *On Uniform Learnability of Language Families*, Information Processing Letters 44, 35-38.

[Ne01] Nessel, J. (2001); personal communication.

[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).

[WZ95] Wiehagen, R.; Zeugmann, T. (1995); *Learning and Consistency*, In: K.P. Jantke and S. Lange, editors, Algorithmic Learning for Knowledge-Based Systems, Lecture Notes in Artificial Intelligence 961, 1-24.

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

[Zi00] Zilles, S. (2000); *On Uniform Learning of Classes of Recursive Functions*, Technical Report LSA-2000-05E, Centre for Learning Systems and Applications, University of Kaiserslautern.