Theoretical Computer Science 110 (1993) 131–144
Elsevier

On the power of inductive inference from good examples

R. Freivalds and E.B. Kinber

Institute of Mathematics and Computer Science, University of Latvia, 226250 Riga, Latvia

R. Wiehagen

Department of Informatics, Humboldt-University, O-1086 Berlin, Germany

Communicated by A.R. Meyer
Received July 1990
Revised August 1991

Abstract

Freivalds, R., E.B. Kinber and R. Wiehagen, On the power of inductive inference from good examples, Theoretical Computer Science 110 (1993) 131–144.

The usual information in inductive inference available for the purposes of identifying an unknown recursive function f is the set of *all* input/output examples $(x, f(x))$, $n \in N$. In contrast to this approach we show that it is considerably more powerful to work with *finite* sets of "good" examples even when these good examples are required to be effectively computable. The influence of the underlying numberings, with respect to which the identification has to be realized, to the capabilities of inference from good examples is also investigated. It turns out that nonstandard numberings can be much more powerful than Gödel numberings.

1. Introduction

The main problem in recursion-theoretic inductive inference is the following. Let f be any recursive function, and let ψ be any numbering of some class of partial recursive functions containing f. Then the task is to synthesize an index (a "program") of f with respect to ψ solely from the sequence $((n, f(n))_{n \in N}$. Thus, an inductive inference strategy can use the sequence of *all* input/output examples of the unknown function. In the sequel we refer to this approach as inductive inference from all examples (abbreviated: aex-inference). For an overview of aex-inference the reader is referred to the surveys [1, 14] and to the monographs [16, 3].

Correspondence to: R. Wiehagen, Department of Informatics, Humboldt-Universität, O-1086 Berlin, Germany.

In the following we investigate inductive inference from (a *finite* number of) "good" examples (abbreviated: gex-inference). Our results show that gex-inference is considerably more powerful than aex-inference.

We will need the following definitions.

Let P, R, P^2, R^2 denote the set of all partial recursive and recursive functions of one and two arguments, respectively. Let E denote the set of all functions from N into N with finite domain.

A function $\psi \in P^2$ is called a numbering. We write ψ_i instead of $\lambda x[\psi(i, x)]$. Let $P_\psi = \{\psi_i \mid i \in N\}$ and $R_\psi = P_\psi \cap R$. $\varphi \in P^2$ is called a Gödel numbering (cf. [18]) iff, for any $\psi \in P^2$, there is $c \in R$ such that, for any i, $\psi_i = \varphi_{c(i)}$. Note that this definition implies $P_\varphi = P$. Let G denote the set of all Gödel numberings.

For a function $f \in R$ and $n \in N$, let $f^n = \text{cod}(f(0), f(1), \ldots, f(n))$, where cod denotes an effective and bijective mapping from the set of all finite sequences of natural numbers onto N.

For functions $f, g \in P$ and $n \in N$, let $f =_n g$ iff $\{(x, f(x)) \mid x \leqslant n$ and $f(x)$ is defined$\} = \{(x, g(x)) \mid x \leqslant n$ and $g(x)$ is defined$\}$; otherwise $f \neq_n g$. Sometimes we identify a function $f \in R$ with the sequence of its values; so $0^m 1^\infty$ denotes the function f such that $f(x) = 0$ if $x < m$, and $f(x) = 1$ if $x \geqslant m$.

For $\delta \in E$ and $f \in R$, we write $\delta \subset f$ instead of "δ is a proper subfunction of f". By init(δ) we denote the subfunction $\{(x, \delta(x)) \mid x \leqslant m\}$ of δ, where m is the maximal argument such that $\delta(x)$ is defined for all $x \leqslant m$. By max(δ) we denote the maximal argument of the domain of δ.

For any set A, by pA we denote the set of all subsets of A.

Definition 1.1. Let $U \subseteq R$ and let ψ be any numbering. U is called finitely identifiable with respect to ψ iff there is a strategy $S \in P$ such that, for any function $f \in U$, there is $n \in N$ such that

(1) for all $x < n$, $S(f^x) = ?$,

(2) $\psi_{S(f^n)} = f$.

Here ? is a special symbol, the output of which can be interpreted as saying by the strategy "I don't know yet.". It is required that the first "real" hypothesis be a correct ψ-program for the function f.

Finite identification was introduced in [12]. The reader is also referred to [8, 13, 21]. Let us define

$$\text{FIN}_\psi = \{U \mid U \text{ is finitely identifiable with respect to } \psi\},$$

$$\text{FIN} = \bigcup_{\psi \in P^2} \text{FIN}_\psi.$$

Definition 1.2 (Barzdin [2], Blum and Blum [4], Gold [12]). Let $U \subseteq R$ and let ψ be any numbering. U is called limit identifiable with respect to ψ iff there is a strategy

$S \in P$ such that, for any function $f \in U$, there is $i \in N$ such that

(1) $\psi_i = f$,

(2) $S(f^n) = i$ for almost all n.

Thus, the sequence of hypotheses produced by the strategy S on the function f converges to a correct ψ-program of f. We note that no restriction is made that we should be able to algorithmically determine whether the sequence of hypotheses has already stabilized. It is easy to see that such a restriction would lead to the concept of finite identification.

We define

$$\mathrm{LIM}_\psi = \{U \mid U \text{ is limit identifiable with respect to } \psi\},$$

$$\mathrm{LIM} = \bigcup_{\psi \in P^2} \mathrm{LIM}_\psi.$$

Definition 1.3 (Barzdin [2], Case and Smith [5], Feldman [7]). Let $U \subseteq R$ and let ψ be any numbering. U is called behaviorally correct identifiable with respect to ψ iff there is a strategy $S \in P$ such that, for any function $f \in U$ and for almost all $n \in N$,

$$\psi_{S(f^n)} = f.$$

Thus, on the function f the strategy S produces a sequence of ψ-programs almost all of which compute f.

Define

$$\mathrm{BC}_\psi = \{U \mid U \text{ is behaviorally correct identifiable with respect to } \psi\},$$

$$\mathrm{BC} = \bigcup_{\psi \in P^2} \mathrm{BC}_\psi.$$

The following theorem gives an insight into the possibilities of finite, limit and behaviorally correct aex-inference.

Theorem 1.4 (Barzdin [2], Case and Smith [5], Podnieks [17]). $\mathrm{FIN} \subset \mathrm{LIM} \subset \mathrm{BC} \subset pR$.

Another remarkable fact about aex-inference is pointed out by the following obvious lemma.

Lemma 1.5. *Let $I \in \{\mathrm{FIN}, \mathrm{LIM}, \mathrm{BC}\}$. Then, for any $\varphi \in G$, $I = I_\varphi$.*

Hence, if a function class is aex-inferable at all, then it is always aex-inferable with respect to an arbitrary Gödel numbering. In this sense Gödel numberings are the most "powerful" numberings for aex-inference. Moreover, this seems to imply that the only kind of numberings which are "interesting" for inductive inference are just the Gödel numberings. However, we are convinced that in this strong sense this implication is not justified even for aex-inference. In [9] it is shown that one–one numberings, a special kind of non-Gödel numberings, are very helpful in characterizing the identifiable function classes from FIN and LIM in terms of pure numbering theory. In [10] it is shown that probabilistic inference strategies may have extreme advantages over deterministic ones for FIN-, LIM-, BC-identification just with respect to suitable non-Gödel numberings. Below we show that also for gex-inference, which we will define now, non-Gödel numberings can be more "powerful" than Gödel numberings.

2. Inductive inference from good examples – the approach

The idea of inductive inference from good examples is to use *finite* sets of "well-chosen" examples instead of the infinite sets of *all* examples to identify the unknown functions.

Definition 2.1. Let $U \subseteq R$ and let ψ be any numbering. U is called finitely identifiable from good examples with respect to ψ iff there is a numbering ex, a strategy $S \in P$, and a function $z \in P$ such that $U \subseteq P_\psi$ and, for any $i \in N$ with $\psi_i \in U$,

(1) ex_i is a finite subfunction of ψ_i, $z(i)$ is defined, and $z(i) = \text{card } ex_i$,

(2) for any finite subfunction ε of ψ_i, $\psi_{S(ex_i \cup \varepsilon)} = \psi_i$.

Let us neglect the ε for a moment, i.e. take the special case $\varepsilon = \emptyset$. Then it follows from condition (2) above that, for any function ψ_i from the class U, the strategy S "finitely" produces a correct ψ-program of ψ_i (which may be different from i) solely from ex_i – the finite set of good examples.

Furthermore, it follows from condition (1) that, for any i such that $\psi_i \in U$, ex_i is effectively computable from i.

The ε we need in order to avoid "unfair coding tricks" such as $ex_i = \{(i, \psi_i(i))\}$ which would lead to trivial identification of the whole class R. On the other hand, in "real life" it seems to be seldom to get such a pure set ex_i of good examples. Often one gets additional correct, but nonnecessary information (just the ε) and then one has to deal with the union of all the information, yielding another interpretation of the ε in the definition above.

A possible scenario of inference from good examples is the relationship between teacher and pupil. As a rule the teacher will not tell the pupil only the correct and final answer – say i –, she/he will not present the pupil all she/he knows about the phenomenon to be learned – say $\psi_i(0)$, $\psi_i(1)$, $\psi_i(2)$, … . Actually, she/he will offer some typical information, just "good examples", in order to enable the pupil to learn the unknown phenomenon by processing the good examples.

We will use the following abbreviations:

$$\text{GEX-FIN}_\psi = \{U \mid U \text{ is finitely identifiable from good examples with respect to } \psi\},$$

$$\text{GEX-FIN} = \bigcup_{\psi \in P^2} \text{GEX-FIN}_\psi.$$

Definition 2.2. Let $U \subseteq R$ and ψ be any numbering. U is called limit identifiable from good examples with respect to ψ iff there is a numbering ex, a strategy $S \in P^2$, and a function $z \in P$ such that $U \subseteq P_\psi$ and, for any $i \in N$ with $\psi_i \in U$,

(1) ex_i is a finite subfunction of ψ_i, $z(i)$ is defined, and $z(i) = \text{card } ex_i$,

(2) for any finite subfunction ε of ψ_i, there is j such that $\psi_j = \psi_i$ and, for almost all $n \in N$, $S(ex_i \cup \varepsilon, n) = j$.

Thus, for any function $\psi_i \in U$, on any finite function δ such that $ex_i \subseteq \delta \subseteq \psi_i$ the strategy S produces an infinite sequence of hypotheses converging to a correct ψ-number of the function ψ_i.

Of course, the question arises whether this additional "degree of freedom", namely allowing the strategy of a finite number of mind changes, can really increase its power, since the information processed is always the same, namely δ. In Section 4 we will answer this question affirmatively. We set

$$\text{GEX-LIM}_\psi = \{U \mid U \text{ is limit identifiable from good examples with respect to } \psi\},$$

$$\text{GEX-LIM} = \bigcup_{\psi \in P^2} \text{GEX-LIM}_\psi.$$

Finally, if ψ and ex are numberings and $U \subseteq R$, then ex is said to be an effective U-subnumbering of ψ iff there is $z \in P$, such that, for any $i \in N$ with $\psi_i \in U$, ex_i is a finite subfunction of ψ_i, $z(i)$ is defined and $z(i) = \text{card } ex_i$.

3. Finite gex-inference

Our first result shows that all enumerable classes of recursive functions are finitely identifiable from good examples with respect to suitably chosen numberings.

Therefore, let

$$\text{NUM} = \{R_\psi \mid \psi \in R^2\}$$

denote the family of all enumerable classes of recursive functions. It is well-known that NUM is not contained in FIN (more exactly, NUM and FIN are set-theoretically incomparable, cf. [14]), whereas it follows from Theorem 3.1 that NUM is contained in GEX-FIN. In order to formulate Theorem 3.1, let $R^2_=$ denote the set of all numberings $\psi \in R^2$ such that $\{(i, j) \mid \psi_i = \psi_j\}$ is decidable.

Theorem 3.1. *For any* $\psi \in R^2_=$, $R_\psi \in$GEX-FIN$_\psi$.

Proof. Let $\psi \in R^2_=$. For any $i, j \in N$ such that $\psi_i \neq \psi_j$, let x_{ij} denote the least argument x such that $\psi_i(x) \neq \psi_j(x)$.

Define

$$ex_0 = \emptyset$$

and, for $i > 0$,

$$ex_i = \begin{cases} \{(x_{ij}, \psi_i(x_{ij})) \mid 0 \leqslant j < i\} & \text{if } \psi_i \neq \psi_j \text{ for all } j < i, \\ ex_j & \text{if } j < i \text{ is the least number} \\ & \text{such that } \psi_i = \psi_j. \end{cases}$$

Obviously, ex is an effective R_ψ-subnumbering of ψ.

Furthermore, for any $\delta \in E$, let

$$S(\delta) = \text{the least } i \in N \text{ such that } \delta \subset \psi_i.$$

Clearly, $R_\psi \in$GEX-FIN$_\psi$ by S, since, for any $i \in N$ and any finite subfunction ε of ψ_i, $S(ex_i \cup \varepsilon)$ is equal to the least j such that $\psi_j = \psi_i$. □

Thus, we learn from Theorem 3.1 that good examples for finitely identifying NUM-classes can be effectively computed with respect to *arbitrary* numberings from $R^2_=$. Moreover, we see that the "goodness" of these examples consists in the strategy's ability to distinguish the function ψ_i (without knowing a priori the index i, of course) from all the different previous ones in the numbering ψ, i.e. from any ψ_j, where $j < i$ and $\psi_j \neq \psi_i$. This leads naturally to identification strategies which work enumeratively. On the other hand, the question whether or not there is "another type" of good examples leading to strategies which construct the hypotheses "directly" from the good examples, i.e. without enumerative search through the "space" ψ of hypotheses, has been answered affirmatively for finite identification of pattern languages recently (cf. [15]).

Since it is well-known that for any class $U \in$NUM there is a numbering $\psi \in R^2_=$ such that $R_\psi = U$ (cf. [6]), we get immediately the following corollary from Theorem 3.1.

Corollary 3.2. NUM $\subseteq$ GEX-FIN.

This already contrasts to the NUM $\not\subseteq$ FIN result for finite aex-inference.

However, our next result shows that finite gex-inference is even more powerful than finite aex-inference by "two orders of magnitude" corresponding to the inclusions FIN $\subset$ LIM $\subset$ BC.

Theorem 3.3. GEX-FIN = BC.

Proof. GEX-FIN $\subseteq$ BC.

Let $U \in$GEX-FIN. By definition there is a numbering ψ, an effective U-subnumbering ex of ψ, and a strategy S such that, for any i with $\psi_i \in U$ and any $\delta \in E$ with $ex_i \subseteq \delta \subset \psi_i$, $\psi_{S(\delta)} = \psi_i$.

Then define a strategy $T \in P$ as follows:

$$T(f^n) = S(\{(x, f(x)) \mid x \leqslant n\}).$$

Now let $f \in U$ and i be such that $\psi_i \in U$. Then, obviously, for any $n \geqslant \max(ex_i)$, we have

$$ex_i \subseteq \{(x, f(x)) \mid x \leqslant n\} \subset \psi_i.$$

Consequently,

$$\psi_{T(f^n)} = f \quad \text{for any } n \geqslant \max(ex_i).$$

Hence, $U \in \mathrm{BC}_\psi$ by T.

To show BC $\subseteq$ GEX-FIN we need two lemmas.

Lemma 3.4. *For any $U \in$ BC and any $\varphi \in G$, there is a strategy $S \in P$ such that*

(1) *$U \in \mathrm{BC}_\varphi$ by S,*

(2) *for any $f \in U$ and any $n \in N$, if $\varphi_{S(f^n)} = f$, then, for any $m > n$, $\varphi_{S(f^m)} = f$.*

Proof. Let $U \subseteq R$, $\varphi \in G$, and $U \in \mathrm{BC}_\varphi$ by T. Without loss of generality let $T \in R$. Define a strategy S as follows:

$S(f^n) =$ "A φ-number of the function g, where g is defined as follows (let $i := T(f^n)$):

$$g(x) = \begin{cases} \varphi_i(x) & \text{if } x \leqslant n \text{ or, for all } n < y < x \text{ and } 0 \leqslant x' \leqslant x, \\ & \varphi_i(x') \text{ and } \varphi_{T(g^y)}(x') \text{ are both defined and equal,} \\ \text{undefined} & \text{otherwise."} \end{cases}$$

Now let $f \in U$ and let m be the least number such that, for any $n \geqslant m$, $\varphi_{T(f^n)} = f$. Then, for any $n < m$, $\varphi_{S(f^n)} \neq f$. Indeed, let $n < m$ and $\varphi_{S(f^n)} = f$. Then $g = f$. Hence, by the definition of g, $\varphi_{T(f^y)} = f$ for all $y \geqslant n$. But this is a contradiction to the definition of m. On the other hand, for any $n \geqslant m$, obviously $\varphi_{S(f^n)} = f$. Consequently, (1) and (2) hold. □

Lemma 3.5. *If $U \in$ BC, then there is a numbering ψ and a function $r \in R$ such that*

(1) *$U \subseteq P_\psi$,*

(2) *for any $i, n \in N$, if $\psi_i \in U$ and $n \geqslant r(i)$, then there is $j \in N$ such that $r(j) = n$ and $\psi_j =_n \psi_i$,*

(3) *for any $i, j \in N$, if $\psi_i \in U$, $\psi_j =_{r(j)} \psi_i$, and $r(j) \geqslant r(i)$, then $\psi_j = \psi_i$.*

Proof. Let $U \in \mathrm{BC}_\varphi$ by a strategy $S \in R$, where $\varphi \in G$ and S has the property of Lemma 3.4(2).

Let $\{f^n \mid f \in R,\ n \in N, \text{ and } \varphi_{S(f^n)} =_n f\}$ be enumerated by $e \in R$ without repetition. Then, for any $i \in N$, where $e(i) = f^n$, define $\psi_i = \varphi_{S(f^n)}$ and $r(i) = n$.

Let $f \in U$ and let m be the least number such that $\varphi_{S(f^n)} = f$ for any $n \geqslant m$. Let $i \in N$ be such that $e(i) = f^m$. Then $\psi_i = f$; hence, (1) holds.

Now let $i, n \in N$ be such that $f = \psi_i \in U$ and $n \geqslant r(i)$. Then, by Lemma 3.4(2), $\varphi_{S(f^n)} = f$. Hence, there is a $j \in N$ such that $e(j) = f^n$. Consequently, $r(j) = n$ and $\psi_j =_n \psi_i$. Thus, (2) holds, too. Moreover, since $\varphi_{S(f^n)} = f$, we have even $\psi_j = \psi_i$; hence, (3). □

Remark. The converse of Lemma 3.5 is also true. However, there is a more "polished" characterization of BC, namely, Lemma BC.

Lemma BC (Wiehagen [20]). *$U \in \mathrm{BC}$ iff there is a numbering $\psi \in P^2$ and a function $r \in R$ such that*

(1) $U \subseteq P_\psi$,

(2) *for any $f \in U$ and almost all i, $\psi_i =_{r(i)} f$ implies $\psi_i = f$.*

Proof of Theorem 3.3 (*continued*). Now in order to prove $\mathrm{BC} \subseteq \text{GEX-FIN}$ let $U \in \mathrm{BC}$ and let a numbering ψ and a function r be chosen according to Lemma 3.5.

Then, for any i, we define

$$ex_i = \{(x, \psi_i(x)) \mid x \leqslant r(i)\}.$$

Clearly, ex is an effective U-subnumbering of ψ.

Now, for any $\delta \in E$, let a strategy S be defined as follows:

$S(\delta) =$ "Search for a j such that
$r(j) = \max(\mathrm{init}(\delta))$ and $\psi_j =_{r(j)} \mathrm{init}(\delta)$.
Output j."

Let $f \in U$, $\psi_i = f$, and $\delta \in E$ be such that $ex_i \subseteq \delta \subset f$. Then it follows from conditions (1) and (2) of Lemma 3.5 that $S(\delta)$ is defined. Furthermore, $\max(\mathrm{init}(\delta)) \geqslant r(i)$. But then condition (3) of Lemma 3.5 yields $\psi_{S(\delta)} = f$. □

We have proved that even using the "strongest" way to get the good examples, namely to compute them effectively, and the "strongest" way of inference, namely finite identification, it is possible to identify all the classes from the aex-type BC. In order to achieve this result the underlying numberings had to be chosen carefully. Our next results point out that this is necessary in some sense. Theorem 3.6 shows that the goal of Theorem 3.3 cannot be achieved if we confine ourselves to Gödel numberings.

Theorem 3.6. *For any $\varphi \in G$,* $\text{GEX-FIN}_\varphi \subset \text{GEX-FIN}$.

Proof. Let $U=\{0^m1^\infty \mid m\geqslant 1\}\cup\{0^\infty\}$ be a subset of the step functions containing the "accumulation point" 0^∞. (Note that U can be used to prove that $\text{NUM}\setminus\text{FIN}\neq\emptyset$, cf. [14].) Clearly, $U\in\text{NUM}$. Hence, by Corollary 3.2, $U\in\text{GEX-FIN}$.

Now let $\varphi\in G$. Assume that $U\in\text{GEX-FIN}_\varphi$. Then there are an effective U-subnumbering ex of φ and a strategy S such that, for any i with $\varphi_i\in U$ and any $\delta\in E$ with $ex_i\subseteq\delta\subset\varphi_i$, $S(\delta)$ is defined and $\varphi_{S(\delta)}=\varphi_i$.

Hence, there is some partial recursive function e mapping N into E such that, for any i with $\varphi_i\in U$, $e(i)=ex_i$.

By applying an analogue of Smullyan's double recursion theorem (cf. [19]) there are $i,j\in N$ such that

$$\varphi_i=0^\infty,$$

$$\varphi_j(x)=\begin{cases}0 & \text{if } x\leqslant\max\{e(i),e(j),s\}, \text{ where } s \text{ is the number of steps needed to discover that } e(i) \text{ and } e(j) \text{ are both subfunctions of } 0^\infty,\\ 1 & \text{otherwise.}\end{cases}$$

Assume that $e(i)$ or $e(j)$ is not a subfunction of 0^∞. Then there exists no s, and $\varphi_j=0^\infty$. But then $\varphi_i=\varphi_j=0^\infty\in U$; hence, $e(i)$ and $e(j)$ are both subfunctions of 0^∞, a contradiction. Consequently, $e(i)$ and $e(j)$ are both subfunctions of 0^∞ and, therefore, $\varphi_j\neq 0^\infty$.

Hence, there must be an m such that $\varphi_j(x)=0$ for any $x\leqslant m$, $e(j)\subset 0^\infty$, and $\max(e(j))\leqslant m$. But by the definition of g, we also get $\max(e(i))\leqslant m$. Hence, for $\delta=\{(x,0)\mid x\leqslant m\}$ we have $ex_i\subseteq\delta\subset\varphi_i$ and $ex_j\subseteq\delta\subset\varphi_j$. Consequently, $\varphi_{S(\delta)}=\varphi_i$ and $\varphi_{S(\delta)}=\varphi_j$. But this is a contradiction, since $\varphi_i\neq\varphi_j$. □

Hence, with respect to finite gex-inference nonstandard numberings can be more powerful than Gödel numberings. Moreover, this effect can be realized even on "easy" function classes, namely enumerable ones.

Corollary 3.7. *There is a class* $U\in\text{NUM}$ *such that*
(1) *for any* $\varphi\in G$, $U\notin\text{GEX-FIN}_\varphi$,
(2) *for some* $\psi\in P^2$, $U\in\text{GEX-FIN}_\psi$.

Proof. Let U be the class from the proof of Theorem 3.6. Then (1) follows from the proof of Theorem 3.6 and (2) follows from Corollary 3.2. □

Theorem 3.6 and Corollary 3.7 clearly contrast Lemma 1.5, which characterizes the power of Gödel numberings for aex-inference. For gex-inference a weakened version of Lemma 1.5 holds stating that any two Gödel numberings are equally powerful.

Lemma 3.8. *For any* $\varphi,\varphi'\in G$, $\text{GEX-FIN}_\varphi=\text{GEX-FIN}_{\varphi'}$.

Proof. Let $\varphi, \varphi' \in G$ and $U \in$GEX-FIN$_\varphi$ by $T \in P$. Then it suffices to show that $U \in$GEX-FIN$_{\varphi'}$.

Therefore, let $c, c' \in R$ be such that, for any i, $\varphi_i = \varphi'_{c(i)}$ and $\varphi'_i = \varphi_{c'(i)}$. For any i, define $ex'_i = ex_{c'(i)}$. Then ex' is an effective U-subnumbering of φ'.

For any $\delta \in E$, define

$$S(\delta) = c(T(\delta)).$$

Then, obviously, $U \in$GEX-FIN$_{\varphi'}$ by S. □

In our approach, for any function $\psi_i \in U$, the good examples for finitely identifying ψ_i can be computed effectively from i (by $ex \in P^2$ and $z \in P$). The question naturally arises whether it is possible to enlarge the capabilities of finite gex-inference by weakening the way to get the good examples. It turns out that the answer to this question is negative. We have defined finite gex-inference where the good examples are only computable in the *limit*, or they do simply *exist* but nothing is required concerning the way to compute them. Then we were able to prove that these weaker versions of good examples do not increase the power of finite gex-inference defined above. For the exact definitions and the proofs using similar ideas as in the present paper the reader is referred to [11].

4. Limit gex-inference

Our final result shows that limit gex-inference is also more powerful than limit aex-inference by two orders of magnitude. More exactly, the whole class R of all recursive functions is limit inferable from good examples.

In order to prove this result let us define a "2-dimensional" inference type as follows.

Definition 4.1. Let $U \subseteq R$ and $\varphi \in G$. $U \in$2LIM iff there is a strategy $T \in P^2$ such that, for any $f \in U$ and for almost all n, there is $j \in N$ such that $\varphi_j = f$ and, for almost all k, $T(f^n, k) = j$.

Now we define a strengthened version of this type.

Definition 4.2. Let $U \subseteq R$ and $\varphi \in G$. $U \in S$2LIM iff there is a strategy $T \in R^2$ such that

(1) $U \in$2LIM by T,

(2) for any $f \in U$ and any $n, k \in N$, if $\varphi_{T(f^n, k)} = f$, then, for any $m \geq n$, there is $j \in N$ such that $\varphi_j = f$ and, for almost all l, $T(f^m, l) = j$.

Clearly, the definitions of 2LIM and S2LIM are independent from the particular choice of $\varphi \in G$.

Now we prove that $R \in S$2LIM and S2LIM$=$GEX-LIM. Consequently, $R \in$GEX-LIM.

Proposition 4.3. $R \in S2\text{LIM}$.

Proof. Let $\varphi \in G$. For any $f \in R$ and any $n, k \in N$, let $J_{f^n,k}$ denote the set of all $j \leqslant k$ such that $\varphi_j =_n f$ is discovered within k steps. Now let $i = \min J_{f^n,k}$ if $J_{f^n,k} \neq \emptyset$. Let $g \in P^2$ be such that, for any $f \in R$ and any $n, k, x \in N$,

$$\varphi_{g(f^n,k)}(x) = \begin{cases} \varphi_i(x) & \text{if there is no } j < i \text{ such that } \varphi_j =_n f \\ & \text{is discovered within } x \text{ steps,} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Let $z \in N$ be such that φ_z is the empty function. Then define a strategy S as follows:

$$S(f^n, k) = \begin{cases} g(f^n, k) & \text{if } J_{f^n,k} \neq \emptyset, \\ z & \text{otherwise.} \end{cases}$$

Now let $f \in R$. Let m be the minimal number such that $\varphi_m = f$. Let x_f be the minimal number such that, for any $i < m$, $\varphi_i \neq_{x_f} f$. Now let $n \in N$. Then it is easy to see that there are two cases:

If $n < x_f$, then, for any k, $\varphi_{S(f^n,k)} \neq f$.

If $n \geqslant x_f$, then, for almost all k, $\varphi_{S(f^n,k)} = f$.

Hence, $R \in S2\text{LIM}$ by S. □

Proposition 4.4. $S2\text{LIM} = \text{GEX-LIM}$.

Proof. $S2\text{LIM} \subseteq \text{GEX-LIM}$.

Let $U \in S2\text{LIM}$ by T. Let

$$\{(j, n, k) \mid \varphi_j(x) \text{ is defined for all } x \leqslant n \text{ and } T(\varphi_j^n, k-1) \neq T(\varphi_j^n, k) = j\}$$

be enumerated by $e \in R$ without repetition.

For any $i, x \in N$, where $e(i) = (j, n, k)$, define

$$\psi_i(x) = \begin{cases} \varphi_j(x) & \text{if } x \leqslant n \text{ or } T(\varphi_j^n, k+x) = j, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Furthermore, define

$$ex_i = \{(x, \psi_i(x)) \mid x \leqslant n\}.$$

Then ex is an effective U-subnumbering of ψ.

In order to show that $U \in \text{GEX-LIM}_\psi$, let a strategy S be defined as follows for any $\delta \in E$ and $m \in N$:

$S(\delta, m) =$ "Let $n = \max(\text{init}(\delta))$. Let $k \leqslant m$ be maximal such that
$T(\text{init}(\delta), k-1) \neq T(\text{init}(\delta), k)$. Let $j = T(\text{init}(\delta), k)$.
If $\varphi_j \neq_n \delta$, then $S(\delta, m)$ is undefined.
If $\varphi_j =_n \delta$, then search for the number i such that
$e(i) = (j, n, k)$ and define $S(\delta, m) = i$."

Now let $f \in U$. Let $l \in N$ be such that $\psi_l = f$. Let $\delta \in E$ be such that $ex_l \subseteq \delta \subset f$. We will show that there is an i such that $\psi_i = f$ and, for almost all m, $\psi_{S(\delta, m)} = i$.

Note that condition (2) of the definition of S2LIM guarantees that on init(δ) the strategy T converges to some φ-number of f; formally, there is a j such that $\varphi_j = f$ and, for almost all k, $T(\text{init}(\delta), k) = j$. Let k be maximal such that $T(\text{init}(\delta), k-1) \neq T(\text{init}(\delta), k)$. Let $n = \max(\text{init}(\delta))$ and i be such that $e(i) = (j, n, k)$. Clearly, $\psi_i = f$ and $S(\delta, m) = i$ for almost all m. Hence, $U \in \text{GEX-LIM}_\psi$ by S.

GEX-LIM $\subseteq S$2LIM follows immediately from Proposition 4.3. □

Corollary 4.5. 2LIM = GEX-LIM.

Proof. Immediately from Propositions 4.3 and 4.4, and S2LIM $\subseteq$ 2LIM. □

Now we get the result announced above.

Theorem 4.6. $R \in$ GEX-LIM.

Proof. Immediately from Propositions 4.3 and 4.4. □

5. Conclusions

We have shown that gex-inference is considerably more powerful than the usual approach of aex-inference. The class R of all recursive functions is even gex-identifiable in the limit. In all cases the sets of good examples were effectively computable. In spite of these affirmative results several problems remain open. We want to point out some of them.

First we conjecture that the class R does not remain gex-identifiable in the limit if the number of mind changes will be bounded. More exactly, for $U \subseteq R$ and $m \in N$, let $U \in \text{GEX-LIM}_m$ iff there is a numbering ψ, an effective U-subnumbering ex of ψ, and a strategy $S \in P^2$ such that

(1) $U \in \text{GEX-LIM}_\psi$ by S,

(2) for any $i \in N$ and any $\delta \in E$, if $\psi_i \in U$ and $ex_i \subseteq \delta \subset f$, then $\text{card}\{n \mid S(\delta, n) \neq S(\delta, n+1)\} \leqslant m$.

Then we suggest the following hierarchy result:

For any m, $\text{GEX-LIM}_m \subset \text{GEX-LIM}_{m+1}$.

Clearly, this would imply that, for any m, $R \notin \text{GEX-LIM}_m$.

One way to prove the hierarchy above could consist in defining $S2\text{LIM}_m$ as a version of S2LIM with bounded mind changes, then proving that, for any m, $\text{GEX-LIM}_m = S2\text{LIM}_m$, and, finally, establishing that, for any m, $S2\text{LIM}_m \subset S2\text{LIM}_{m+1}$. It is open whether this technique will work. The reader is referred to [5] for other mind change hierarchies.

Furthermore, it appears that the main reason why good examples are good is because of the strategy's ability to distinguish a function to be identified from all different previous ones in the underlying numbering. This leads naturally to identification strategies which work enumeratively. Are there other such reasons for goodness or not? Can they lead to "constructive" gex-identification strategies? Is it, therefore, reasonable to deal with classes of objects to be identified which possess more "structure" such as formal languages, finite automata, Boolean functions? The first of our results in this direction lead to the conjecture that the answer to these questions will be affirmative (cf. [15]).

Finally, we have seen that nonstandard numberings can be more powerful for gex-inference than Gödel numberings. Is this also true for limit gex-inference? Or, is there $\varphi \in G$ such that $R \in \text{GEX-LIM}_\varphi$? Up to now we know only that, for any $\varphi \in G$, $R \in \text{LIMGEX-LIM}_\varphi$ (cf. [11]). In any case: Let there be a non-Gödel numbering which is powerful for gex-inference. To what extent one can modify this numbering without decreasing its power for gex-inference, but to come as close as possible to the "programming convenience" of Gödel numberings?

References

[1] D. Angluin and C.H. Smith, Inductive inference: theory and methods, *Computing Surveys* **15** (1983) 237–269.

[2] Ya.M. Barzdin, Two theorems on the limiting synthesis of functions, in: [3], Vol. 1 (1974), 82–88 (in Russian).

[3] Ya.M. Barzdin, ed., *Theory of Algorithms and Programs*, Vols. 1, 2, 3 (Riga Latvian State Univ., 1974, 1975, 1977) (in Russian).

[4] L. Blum and M. Blum, Towards a mathematical theory of inductive inference, *Inform. and Control* **28** (1975) 125–155.

[5] J. Case and C. Smith, Comparison of identification criteria for machine inductive inference, *Theoret. Comput. Sci.* **25** (1983) 193–220.

[6] Yu.L. Eršov, *Theory of Numberings* (Nauka, Moscow, 1977) (in Russian).

[7] J.A. Feldman, Some decidability results on grammatical inference and complexity, *Inform. and Control* **20** (1972) 244–262.

[8] R. Freivalds, Finite identification of general recursive functions by probabilistic strategies, *Math. Res.* **2** (1979) 138–145.

[9] R. Freivalds, E.B. Kinber and R. Wiehagen, Inductive inference and computable one-one numberings, *Z. Math. Logik Grundlag. Math.* **28** (1982) 463–479.

[10] R. Freivalds, E.B. Kinber and R. Wiehagen, Probabilistic versus deterministic inductive inference in nonstandard numberings, *Z. Math. Logik Grundlag. Math.* **34** (1988) 531–539.

[11] R. Freivalds, E.B. Kinber and R. Wiehagen, Inductive inference from good examples, in: K.P. Jantke, ed., *Analogical and Inductive Inference 1989*, Lecture Notes in Artificial Intelligence, Vol. 397 (Springer, Berlin, 1989) 1–17.

[12] E.M. Gold, Language identification in the limit, *Inform. and Control* **10** (1967) 447–474.

[13] S. Jain and A. Sharma, Finite learning by a team, in: *Proc. Computational Learning Theory* (1990) 163–177.

[14] R. Klette and R. Wiehagen, Research in the theory of inductive inference by GDR mathematicians – a survey, *Inform. Sci.* **22** (1980) 149–169.

[15] S. Lange and R. Wiehagen, Polynomial-time inference of arbitrary pattern languages, *New Generation Computing* **8** (1991) 361–370.

[16] D.N. Osherson, M. Stob and S. Weinstein, *Systems that Learn* (MIT Press, Cambridge, MA, 1986).
[17] K. Podnieks, Comparing various concepts of limiting synthesis and function prediction, in: [3], Vol. 1 (1974) 68–81 (in Russian).
[18] H. Rogers Jr., *Theory of Recursive Functions and Effective Computability* (McGraw-Hill, New York, 1967).
[19] R.M. Smullyan, Theory of formal systems, *Ann. Math. Studies* **47** (1961).
[20] R. Wiehagen, Zur Theorie der algorithmischen Erkennung, Diss. B, Humboldt-Univ. Berlin, 1978.
[21] R. Wiehagen, On the complexity of effective program synthesis, in: K.P. Jantke, ed., *Analogical and Inductive Inference 1986*, Lecture Notes in Computer Science, Vol. 265 (Springer, Berlin, 1986) 206–219.