On Reliability and Refutability in Nonconstructive Identification

Iļja Kucevalovs

Institute of Mathematics and Computer Science,
University of Latvia, Raina bulv. 29, Riga, Latvia
`ilja.kucevalovs@lais.lv`

Abstract

Identification in the limit, originally due to Gold [10], is a widely used computation model for inductive inference and human language acquisition. We consider a nonconstructive extension to Gold's model. Our current topic is the problem of applying the notions of reliability and refutability to nonconstructive identification. Four general identification situations are defined and two of them are studied. Thus some questions left open in [13] are now closed.

Keywords and phrases inductive inference, identification, reliability, refutability, nonconstructive computation

Digital Object Identifier 10.4230/OASIcs.MEMICS.2010.62

1 Introduction

The computational model of inductive inference known as identification (also: identification in the limit, algorithmic learning, etc.), introduced by Gold [10], and its many variations have been widely studied. The reader is encouraged to refer to [20] and [14] for detailed surveys on applying Gold's model to learning recursive functions and recursive languages from positive data, respectively.

Gold's original model deals with some abstract computational device which makes guesses about some object it has as an input. That device, also referred to as inductive inference machine (IIM), is usually said to identify a class of objects if it identifies (i.e. correctly guesses) any object within that class. However, in the general case the IIM behaviour on objects not from the class in question is not specified.

This issue is dealt with in the "reliable identification" model, which is due to Blum and Blum [2], Minicozzi [16], and later Sakurai [19]; "refutable identification", first considered by Mukouchi and Arikawa [17], is a strengthening of the reliable model. Both these models prohibit the IIM to output any sequence which could be a correct sequence of guesses for some class member (if the given object is not such a member), and the refutable model specifies the IIM to explicitly refute any non-member.

In the thesis [13], Gold's model was extended with nonconstructive computational methods, which allow the IIM to utilize some additional information. There is, however, a certain difference between the nonconstructive identification and the traditional "identification with additional information" (see e.g. [4, 7, 11, 12]). We briefly introduce the two main distinctions. First, all the three nonconstructive identification models (K, S and F) are given on a general level, instead of defining a separate criterion for each particular situation (as it is usually being done). Second, these models are specially constructed so that a trivial help (i.e. supplying the desired answer) would not be possible.

Allowing additional information for identification introduces the following dilemma: Should we always assume that the help the IIM gets is correct (and if not, should we refute

© Iļja Kucevalovs;
licensed under Creative Commons License NC-ND
Sixth Doctoral Workshop on Math. and Eng. Methods in Computer Science (MEMICS'10)—Selected Papers.
Editors: L. Matyska, M. Kozubek, T. Vojnar, P. Zemčík, D. Antoš; pp. 62–68
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

this help)? Note the obvious similarity between this problem and the one stated above, solved by introducing reliability for identification. That motivated the author to consider two levels of reliability and refutability and, correspondingly, four different identification situations in nonconstructive identification.

In the current paper, we show a class of non-recursive functions that is not identifiable without any additional information, but is nonconstructively identifiable with the following properties. On the one hand, the help information it is identifiable with can be biased by a function that grows to infinity; on the other hand, it is refutably identifiable (utilizing, however, another kind of help).

To our knowledge, the problems of such type were neither previously solved nor considered at all.

2 Preliminaries

In this section, we briefly introduce the notions used in this paper. Notions from recursive function theory not explained here are treated in e.g. Rogers' textbook[18]; a more brief introduction into recursion theoretic notions can be found in Gallier and Hicks' online books [8, 9]. The notion of Kolmogorov complexity is explained in detail in [15].

N denotes the set of nonnegative integers $\{0, 1, 2, ...\}$. N^+ denotes the set of positive integers $\{1, 2, 3, ...\}$. R denotes the set of all total recursive functions.

$\langle x_0, x_1, ..., x_k\rangle$ is an ordered tuple of elements x_0, x_1, ..., x_k (in that order). If X is such a tuple, we refer to each its kth element as X_k. Given some universal set U, we write $\langle U^k\rangle$ to denote a set of all tuples of some length k over U. $\langle U^*\rangle \equiv \cup_{k\in U}\langle U^k\rangle$, i.e. the set of all finite length tuples over U. $\langle U^\infty\rangle$ denotes the set of all infinite length tuples over U. $\langle U\rangle \equiv \langle U^*\rangle \cup \langle U^\infty\rangle$, i.e. the set of all possible tuples over U, of both finite and infinite length.

$\forall^\infty, \exists^\infty, \exists!$ denote "for all but finitely many", "there exist infinitely many", "there exists a unique", respectively.

For an object x, $l(x)$ denotes the length of the binary presentation of x. For a set X, $d(X)$ is its cardinality and $\bar{X}$ is its complement. For a sequence X, $l(X)$ is its length. For a sequence X, $X[n]$ denotes its initial segment of length n. For a total function f, $f[n]$ denotes the initial segment of its graph. We write $f(x_0)\downarrow$ and $f(x_0)\uparrow$ to denote that f is defined (undefined) on x_0.

Given a class U, we call any partial recursive function $\varphi : N \to U$ a numbering for U. We say that U is *defined* in φ (written: $\varphi(U)\downarrow$), iff $\forall u \in U : (\exists n \in N)[\varphi_n = u]$. Talking about identification of some U in some φ, we always assume $\varphi(U)\downarrow$.

Having fixed some universal Turing machine M_{uni}, by (plain) Kolmogorov complexity of some object u we call the value

$$C(u) = \min\{l(p) : M_{uni}(p) = u\} \quad (1)$$

In this paper, only the so-called "plain" Kolmogorov complexity will be used (for more details see [15]), and our results will rely on one fixed M_{uni}.

A tuple $J \in \langle N^\infty\rangle$ is a BC-sequence for some object u in some numbering $\varphi = \varphi_0, \varphi_1, \varphi_2, \ldots$ (written: $J \in BC(u, \varphi)$), iff $\forall^\infty n \in N : \varphi_{J_n} = u$.

A tuple $J \in \langle N^\infty\rangle$ is an EX-sequence for some u in some φ (written: $J \in EX(u, \varphi)$), iff $J \in BC(u, \varphi) \wedge (\forall^\infty n \in N : J_n = J_{n+1})$.

A tuple $J \in \langle N^\infty\rangle$ is a FIN-sequence for some u in some φ (written: $J \in FIN(u, \varphi)$), iff $J \in EX(u, \varphi) \wedge (\forall n \in N : J_n = J_{n+1})$.

We say that FIN, EX, BC are identification criteria. We say that $I : U \to \{\langle N \rangle\}$ is an (information) presentation of a class U. We say that I is *injective*, iff $\forall u, v \in U : (u \neq v) \Rightarrow (I(u) \cap I(v) = \varnothing)$. We assume that any presentation I we deal with is injective. We say that I is *unambiguous*, iff $\forall u \in U : d(I(u)) = 1$.

A presentation of a recursive function f is its graph, i.e. pairs $\langle x, f(x) \rangle$, which can be reduced to simply a sequence of values $f(0), f(1), f(2), \ldots$ in case we deal with total functions only (like in this paper). We assume, without a loss of generality [10], that a graph of any function f is always input in its natural order, i.e. $f(0), f(1), f(2)$, and so on. Clearly, a function graph is an injective presentation.

An *inductive inference machine* (IIM) is an abstract device that receives positive integers from time to time and generates positive integers from time to time. If M is some IIM, by $M(X) = Y$ we denote "Y is the output M writes having received some input X".

Given some identification criterion $X : X \in \{FIN, EX, BC\}$, some IIM M, some object u and its presentation I and some numbering φ, we say M X-identifies u from I in φ iff $M(I(u)) \in X(u, \varphi)$. Obviously, if M EX-identifies some u, it BC-identifies it; if M FIN-identifies u, it EX-identifies it.

Given some identification criterion X, some IIM M, some class of objects U, some presentation I and some numbering φ, we say that M X-identifies U from I in φ iff M X-identifies every $u \in U$ from I in φ.

Given some identification criterion X, some IIM M, some class of objects U, some presentation I and some numbering φ, we say that M reliably X-identifies U from I in φ iff the following holds:

1. $\forall u \in U : M(I(u)) \in X(u, \varphi)$;
2. $\forall u \notin U : M(I(u)) \notin X(u, \varphi)$.

Given some additional refutation symbol #, we say that M refutably X-identifies U from I in φ iff the following holds:

1. $\forall u \in U : M(I(u)) \in X(u, \varphi)$;
2. $\forall u \notin U : M(I(u))_{l(M(I(u)))} =$"#".

We say that U is (reliably, refutably) X-identifiable from I in φ iff there exists an IIM that (reliably, refutably) X-identifies U from I in φ.

3 Nonconstructive identification

There are several definitions of nonconstructive identification [13]. Here we consider one of them, the so-called F-nonconstructivity, which is very similar to identification given the upper bound on the program size (see e.g. [4, 12]). However, the definition given here does not limit identification to function or language learning, or to any other particular type of learning.

▶ **Definition 1.** Given some identification criterion X, some IIM M, some class U, some presentation I and some numbering $\varphi : \varphi(U) \downarrow$, we say that M F-nonconstructively X-identifies U from I in φ with amount of nonconstructivity $p(n)$, iff $\forall n \in N : \exists m \in N$ s.t. the following holds:

1. $m \leq p(n)$;
2. $\forall u \in U \cap \{\varphi_i | i \leq n\} : M(\langle I(u), m \rangle) \in X(u, \varphi)$.

Further in the text “nonconstructive”, “nonconstructively” etc. is also referred to as “NK”.[1] Any identification model without additional information is called “constructive”.

4 Application of reliability and refutability

One can consider two levels of applying reliability and/or refutability to nonconstructive identification:

1. Reliable / refutable identification;
2. Reliable / refutable nonconstructivity.

On the first level, one deals with the usual problem of reliability and refutability: Which classes in which numberings can be identified so that input of a non-member presentation would not result in misleading “identification” of it. Nonconstructive methods are used to assist reliability or refutability.

On the second level, the problem is the following: Which is that nonconstructive identification model (if it exists), such that not only correct, but also incorrect nonconstructive information helps identifying some class in some numbering in compliance with some criterion. By saying “helps” we mean that no constructive identification would be possible in that case.

Thus, in accordance with [13], the following four situations are defined:

1. NK-X (the usual nonconstructive identification model);
2. NK-R-X (reliable / refutable models utilizing nonconstructivity);
3. R-NK-X (nonconstructive models which are required to work correctly with incorrect help);
4. R-NK-R-X (reliable / refutable nonconstructive models which are required to work correctly with incorrect help).

We will also use the situation names to denote the corresponding inferring power classes. That is, for a situation Z, the class Z is the set of classes which are Z-identifiable.

In [13], only the NK-X situation was studied. It was noted (and some examples were given) that even emptiness (non-emptiness) of the classes NK-R-X and R-NK-X can be tedious tasks to solve (while proving the R-NK-R-X case would obviously close both these questions). Below we prove non-emptiness of NK-R-X and R-NK-X, leaving out R-NK-R-X.

The above definition of F-nonconstructivity is modified in order to properly define the R-situations.

▶ **Definition 2.** Given some total $E : N \to N$ s.t.

$$\lim_{n\to\infty} E(n) = \infty \tag{2}$$

as well as some identification criterion X, some IIM M, some class U, some presentation I and some numbering $\varphi : \varphi(U)\downarrow$, we say that M E-reliably F-nonconstructively X-identifies U, iff $\forall n \in N : \exists m \in \langle N\rangle$ s.t. the following holds:

1. $\forall u \in U \cap \{\varphi_i | i \le n\}, \forall j \in N : M(\langle I(u), m_j\rangle) \in X(u, \varphi)$;
2. $\forall u \in U \cap \{\varphi_i | i \le n\} : \lim_{j\to\infty} M(\langle I(u), m_j \pm E(j)\rangle) \in X(u, \varphi)$.

[1] From Latvian “nekonstruktīvs” (nonconstructive). This is being done in accordance with the original thesis [13], which was written in Latvian.

5 Results

▶ **Theorem 3.** *There exists a class U and a numbering W such that the following properties hold:*

1. *U is not constructively identifiable in W;*
2. *U is F-nonconstructively FIN-identifiable in W (NK-FIN);*
3. *U is F-nonconstructively refutably FIN-identifiable in W (NK-Ref-FIN);*
4. *U is reliably F-nonconstructively FIN-identifiable in W (Rel-NK-FIN).*

Proof. The NK-FIN part is immediate from Rel-NK-FIN.
Consider the total functions $h, m : N \to N$ defined as follows:

$$h(x) = \begin{cases} C(1024), & x = 0 \\ \min\{n \in N \mid (n > h(x-1)) \wedge (C(n) > C(h(x-1)))\}, & x > 0 \end{cases} \tag{3}$$

$$m(x) = \min\{C(n) \mid n \in N \wedge n \geq x\} \tag{4}$$

It is obvious that both such functions do exist; however, neither h [13] nor m [15] can be recursive. Moreover, m, despite being unbounded from above, grows slower than any computable function [15].

Let $\langle p_0, p_1, ...\rangle$ be a growing sequence of all the prime numbers starting with $p_0 = 3$. For every $k \in N$, we define

$$f_k(x) \equiv h((p_k)^x) \tag{5}$$

The numbering W is defined as follows:

$$w_i = \begin{cases} f_k, & \exists k \in N, j \in N \setminus [0; k-1], \\ & \quad n \in N \cap \left[h(j) - \lfloor \frac{m(j)}{2} \rfloor; h(j) + \lfloor \frac{m(j)}{2} \rfloor\right] : \\ & \quad f_k(n) = i \\ h, & otherwise \end{cases} \tag{6}$$

We now briefly explain the idea of the above construction. Each function f_k outputs h values from arguments taken from powers of the k-th prime. That is, $\text{range}(f_i) \cap \text{range}(f_j) = \{1\}$ for every natural $i \neq j$. Moreover, range of every f_k fully contains the set of its indices in W. (That is, every f_k is self-referential in W.) However, indices are contained not in the full range of f_k, but only in the intervals defined by the functions h and m starting from the k-th interval.

The class

$$U = \{f_n \mid n \in N\} \tag{7}$$

is not constructively identifiable in W. First of all, W is not recursive due to non-recursiveness of h and m. That is to say, all the information an IIM can rely on is the self-referential values of f_k. If there existed a value x_0 such that every f_k would output a self-reference given x_0, the problem of constructing an IIM would be trivial; however, every f_k does not output self-references up to the argument value from the k-th interval — that is, no self-referential interval is common for all the f_k. So the only possibility left for identifying U is to possess some "knowledge" about the structure of infinitely many such intervals; this is also not possible due to the incomputable properties of h and m stated above.

Nevertheless, U is reliably F-nonconstructively FIN-identifiable in W. For every f_k, we define the set

$$\pi(f_k) = \{h(i) \mid i \geq k\} \tag{8}$$

It is easy to see that an IIM defined as

$$M(\langle\langle f(0), f(1), ...\rangle, \pi_0 \in \pi(f)\rangle) = \langle f(\pi_0), f(\pi_0), ...\rangle \tag{9}$$

F-nonconstructively FIN-identifies U. Moreover, this identification is E-reliable with $E(x) \equiv \lfloor\frac{m(x)}{2}\rfloor$. Indeed, the condition (2) for E does hold, while any additional information word (8), even having been biased by E, would still help FIN-identify U.

What is left for us is to prove the NK-Ref-FIN part. For any input object u, define the infinite additional information word

$$\langle h(0), h(1), ...\rangle \tag{10}$$

(We assume that the elements are mutually separated using some meta symbols.)

The IIM waits until the element $h(i) = u(1)$. If $h(i) > u(1)$ or i is not prime (or is less than 3), IIM outputs "#" and stops; otherwise it continues running the following algorithm:

1. Set $x \leftarrow 2$;
2. Calculate $(p_i)^x$;
3. Wait until the element $h((p_i)^x)$;
   - a. If $h((p_i)^x) = u(x)$:
     - i. Output $u(h(p_i))$ (if it is already received);
     - ii. Set $x \leftarrow x + 1$;
     - iii. Go to Step 2;
   - b. If $h((p_i)^x) \neq u(x)$:
     - i. Output "#";
     - ii. Stop execution.

It is quite obvious that such an IIM NK-Ref-FIN-identifies U with infinite nonconstructivity. ■

6 Conclusions and future work

We have shown that the classes NK-R-X and R-NK-X are not empty; moreover, the intersection of these classes is not empty. However, the above construction is not strong enough to allow proving (or disproving) the strongest case R-NK-R-$X \neq \varnothing$. Indeed, the refutable part of the proof relies on exact comparison of the additional information and the object in question, — that is, no errors in the given help could be allowed.

On the other hand, in the current paper we did not distinguish between e.g. Rel-NK-X and Ref-NK-X. Moreover, only the F-type nonconstructivity was studied, while the other nonconstructivity types could be considered for reliable/refutable identification as well. We hope to give a more complete hierarchy of reliably nonconstructive identification classes in the future.

7 Acknowledgment

The authors thanks the anonymous referees of MEMICS 2010 for their valuable comments which improved the presentation of this paper.

References

1 A. Ambainis, K. Apsītis, C. Calude, R. Freivalds, M. Karpinski, T. Larfeldt, I. Sala, J. Smotrovs: "Effects of Kolmogorov Complexity Present in Inductive Inference as Well"; ALT '97: Proceedings of the 8th International Conference on Algorithmic Learning Theory (1997), 244-259.

2 L. Blum, M. Blum: "Toward a mathematical theory of inductive inference"; Information and Control, vol. 28, 1975, 125-155.

3 J. Case, C. Smith: "Comparison of Identification Criteria for Machine Inductive Inference"; Theoretical Computer Science 25 (1983), 193-220.

4 R. Freivald, R. Wiehagen: "Inductive inference with additional information"; Electron. Inform. Kybernetik 15, 1979, 179-195.

5 R. Freivalds, J. Bārzdiņš, K. Podnieks: "Inductive inference of recursive functions: complexity bounds"; Lecture Notes in Computer Science, vol. 502 (Baltic Computer Science) (1991), 111-155.

6 R. Freivalds: "Amount of nonconstructivity in finite automata": Lecture Notes in Computer Science, vol. 5642 (2009), 227-236.

7 M. Fulk: "Inductive Inference with Additional Information" (revised July 13, 1998 by S. Jain, A. Sharma): Journal of Computer and System Sciences 64 (2002), 153-159.

8 J. Gallier. "Formal Languages And Automata; Models of Computation, Computability; Basics of Recursion Function Theory"; `http://www.cis.upenn.edu/~jean/gbooks/tocnotes.html` (2010.04.14)

9 J. Gallier, A. Hicks: "The Theory of Languages and Computation"; `http://www.cis.upenn.edu/~jean/gbooks/tc.html` (2010.04.14)

10 E.M. Gold: "Language identification in the limit"; Information and Control 10 (1967), 447-474.

11 S. Jain, A. Sharma: "Learning in the Presence of Partial Explanations"; Information and Computation, vol. 95-2 (Dec 1991), 162-191.

12 S. Jain, A. Sharma: "Learning with the Knowledge of an Upper Bound on Program Size"; Information and Computation, vol. 102-1 (Jan 1993), 118-166.

13 I. Kucevalovs: "Nekonstruktivitātes daudzums induktīvajā izvedumā" (master thesis); University of Latvia, 2010. (in Latvian)

14 Lange, Zeugmann, Zilles: "Learning indexed families of recursive languages from positive data: a survey"; Theoretical Computer Science, vol. 397 (2008) n.1-3, 194-232.

15 M. Li, P. Vitányi: "An Introduction to Kolmogorov complexity and its applications" (3rd ed.); Springer-Verlag (2008).

16 E. Minicozzi: "Some natural properties of strong identification in inductive inference"; Theoretical Computer Science, vol.2, 1976, 345-360.

17 Y. Mukouchi, S. Arikawa: "Inductive Inference Machines That Can Refute Hypothesis Spaces"; tech. rep. RIFIS-TR-CS-67, Research Institute of Fundamental Information Science, Kyushu University, Jan. 15, 1994.

18 H. Rogers: "Theory of Recursive Functions and Effective Computability"; McGraw-Hill (1967).

19 A. Sakurai: "Inductive Inference of Formal Languages from Positive Data Enumerated Primitive-Recursively"; The Japan Society for Industrial and Applied Mathematics, Vol.1 No.3, 1991, pp.177-193. (in Japanese)

20 Zeugmann, Zilles: "Learning Recursive Functions: A Survey"; Theoretical Computer Science, vol. 397 (2008) n.1-3, 4-56.