# Probabilistic Limit Identification up to "Small" Sets *

Juris Vīksna

Institute of Mathematics and Computer Science
The University of Latvia
Rainis boulevard 29, Rīga, LV – 1459, Latvia
email: jviksna@cclu.lv

**Abstract.** In this paper we study limit identification of total recursive functions in the case when "small" sets of errors are allowed. Here the notion of "small" sets we formalize in a very general way, i.e. we define a notion of measure for subsets of natural numbers, and we consider as being small those sets, which are subsets of sets with zero measure.
We study relations between classes of functions identifiable up to "small" sets for different choices of measure. In particular, we focus our attention on properties of probabilistic limit identification. We show that regardless of particular measure we always can identify a strictly larger class of functions with probability $1/(n+1)$ than with probability $1/n$. Besides that, for computable measures we show that, if there do not exist sets with an arbitrary small non zero measure, then identifiability of a set of functions with probability larger than $1/(n+1)$ implies also identifiability of the same set with probability $1/n$. Otherwise (in the case when there exist sets with an arbitrary small non zero measure), we always can identify a strictly larger class of functions with probability $(n+1)/(2n+1)$ than with probability $n/(2n-1)$, and identifiability with probability larger than $(n+1)/(2n+1)$ implies also identifiability with probability $n/(2n-1)$.

## Introduction

Inductive inference in the case when hypotheses produced by inductive inference machine are allowed to have errors have already been widely studied. Clearly, we are usually interested in situations, when number of allowed errors is in some sense "small". There have been several different attempts to describe what exactly we understand with the notion "small set".

The simplest approach is to define "small" sets as sets that are finite (see [2, 5]). However situation becomes less clear, if we want to include in the class of "small" sets also some infinite sets, which by our intuition we are considering as being small. One possibility is to define small sets as sets with small (or even zero) density. Such approach is independently studied by J. Royer (see [11]) and C. Smith, M. Velauthapillai (see [12, 13]). Other possibilities are considered in

* This work has been supported by Grant No.93-593 from Latvian Council of Science.

[7], where small sets are defined as in some sense "easy computable", and in [6], where, in quite the opposite way, small sets are those, which are quite hard to compute, namely, immune, hyperimmune etc. sets.

However, such definitions of "small" sets in general are dependent from a particular encoding of total recursive functions. That is, if we consider some practical algorithm, designed for learning of some class of objects which can be described by recursive functions, we hardly can expect that our intuitive understanding of "small" sets of errors will coincide with one of the definitions given above. Therefore, it seems that some more general notion of "smallness" could be quite useful.

In this paper we are trying to describe "small" sets in terms of measure. Namely, we define a notion of measure on the set of natural numbers and study limit identification up to sets which either have zero measure, or are subsets of sets with zero measure. Such approach generalizes a large part of notions of "small" sets studied so far, in particular, finite, zero density, immune, hyperimmune, etc. sets. Besides that, it only requires that sets, which we regard as being small, satisfy some simple and quite natural properties, i.e. we can hope that in most cases it will be adequate for our intuitive notion of "small sets".

We show that most of results about identification up to sets with zero measure hold regardless of the particular measure we use to describe which sets are "small". Besides that, in the case when considered measure can be in some sense effectively computed (we call such measures *computable*), we show that probabilistic behaviour of identification up to sets with zero measure can be of one of two different types, and belonging to any of these types can be easily characterized by the properties of measure. In particular, we show that for computable measures the power of $EX$ identification either grows in discrete steps at probabilities $1, 1/2, 1/3, 1/4, \ldots$, or grows in discrete steps at probabilities $1, 2/3, 3/5, 4/7, \ldots$.

## Main Definitions and Notation

In general we use more or less standard mathematical notation, specific notation from the theory of inductive inference is similar as, for example, in [2].

By $A - B$, where $A$ and $B$ are sets, we denote the set $\{x \in A \mid x \notin B\}$. Intersection and union of two sets are denoted in the usual way – by $A \cap B$ and $A \cup B$ respectively. By $A \subseteq B$ we understand that $A$ is a subset of $B$, $A \subset B$ means that $A \subseteq B$ and $A \neq B$. The set of all natural numbers we denote by $\mathbf{N}$, the set of all positive natural numbers by $\mathbf{N}_+$. Similarly, the set of all real numbers is denoted by $\mathbf{R}$ and the set of all non negative real numbers by $\mathbf{R}^+$. The complement $\bar{A}$ of a set $A \subseteq \mathbf{N}$ is defined with respect to the set $\mathbf{N}$, i.e. $\bar{A} = \{x \in \mathbf{N} \mid x \notin A\}$.

By $\mathcal{R}$ and $\mathcal{P}$ we correspondingly denote the set of all total recursive and the set of all partial recursive functions with one argument. Similarly, the sets of $n$ argument functions will be denoted by $\mathcal{R}_n$ and $\mathcal{P}_n$. For $f \in \mathcal{P}$ and natural number $x$ by $f(x)\downarrow$ we understand that $f(x)$ is defined, and by $f(x)\uparrow$ we understand that $f(x)$ is undefined.

By $\varphi$ we denote some previously fixed Gödel numbering of all partial recursive functions. For technical convenience we also assume that $\varphi_0(x)\uparrow$ for all $x \in \mathbf{N}$. By $W_n$ we denote a recursively enumerable set with index $n$, i.e. $W_n = \{x \in \mathbf{N} \mid \varphi_n(x)\downarrow\}$.

For $f \in \mathcal{P}$ by $\#f$ we denote some number of function $f$ in numbering $\varphi$. This notation we mainly will use in the places where we will exploit Recursion theorem, i.e. we will use $\#f$ to define some function $f \in \mathcal{P}$. For brevity in such cases we usually will not refer to Recursion theorem explicitly.

$\langle\cdot,\cdot\rangle : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ will be some computable bijective mapping from the set of pairs of natural numbers to the set of natural numbers. For $z = \langle x, y\rangle$ we will refer to $x$ as $z\langle 1\rangle$ and to $y$ as $z\langle 2\rangle$.

We consider identification of total recursive functions with inductive inference machines (or IIM) – a modification of Turing machines (either deterministic, or probabilistic), which on a special input tape receive the graph of some total recursive function $f$, i.e. the sequence $f(0), f(1), \ldots$, and which, as hypotheses about function $f$, output some sequence of numbers $h_1, h_2, \ldots$. For convenience we will denote by $f[n]$ the sequence of the first $n+1$ elements of the graph of function $f$, and by $M(f[n])$ the hypothesis of machine $M$, produced after receiving the sequence $f[n]$. Without loss of generality we can assume that $M(f[n])$ is always defined. (For probabilistic machines $M(f[n])$ can have different values, each of them produced with some probability. In this case we can assume that the total sum of probabilities is always 1.)

In order to formalize the notion of "small" sets we define the notion of measure of subsets of natural numbers in the following way.

**Definition 1.** A set $S \subseteq \{X \mid X \subseteq \mathbf{N}\}$ is called a collection of *measurable* sets and a function $\mu : S \to [0,1]$ is called a *measure*, if $\emptyset \in S$, $\mathbf{N} \in S$, $\mu(\mathbf{N}) = 1$ and for arbitrary sets $X, Y \in S$ the following holds:

1. $X \cup Y, X \cap Y, X - Y \in S$,
2. $\mu(X) = \mu(\{x + c \mid x \in X\})$ for an arbitrary $c \in \mathbf{N}$, and
3. if $X \cap Y = \emptyset$, then $\mu(X) + \mu(Y) = \mu(X \cup Y)$.$\diamond$

Note that collection of measurable sets not necessarily must contain all subsets of $\mathbf{N}$, i.e. there can exist sets that are not measurable.

If $\mu$ is some given measure and $S$ is a collection of measurable sets (with respect to $\mu$), then we consider as being "small" all sets $A \subseteq \mathbf{N}$, such that $A \subseteq B$ for some $B \in S$ with $\mu(B) = 0$. Such sets we will also call *null* sets. Note that from definition of measure it easily follows that $\mu(\emptyset) = 0$ for an arbitrary measure $\mu$. Thus, the empty set will be "small" regardless of the particular measure.

For convenience, the fact that two functions $f, g \in \mathcal{P}$ are equal up to null set (i.e. $\exists A \in S : \ (\mu(A) = 0 \ \& \ \forall x \notin A : \ f(x) = g(x) \ \& \ f(x)\downarrow)$) we will denote by $f =^{\mu 0} g$.

The given definition of measure is analogous with the one traditionally used in mathematical analysis. However, here we do not require the property of so called $\sigma$-additivity, i.e. we do not require that for an arbitrary enumerable sequence of disjoint sets $X_1, X_2, \ldots \in S$ the equality

$$\mu(\bigcup_{i \in \mathbf{N}} X_i) = \sum_{i \in \mathbf{N}} \mu(X_i)$$

must hold, since the requirement of $\sigma$-additivity for our situation is too restrictive (there exists only trivial $\sigma$-additive measure $\mu$ with collection of measurable sets $S = \{\emptyset, \mathbf{N}\}$ and $\mu(\emptyset) = 0$, $\mu(\mathbf{N}) = 1$).

Besides that, we also assume that the set of all natural numbers have measure 1. The other somewhat different possibility is to consider $\mu : S \to \mathbf{R}^+ \cup \{\infty\}$ and to require $\mu(\mathbf{N}) = \infty$ – in this case measure in general can be defined on different collections $S$ of sets as in the case $\mu(\mathbf{N}) = 1$. As far as we know all our results also hold for this second case, when $\mu(\mathbf{N}) = \infty$. However, because we do not have any interesting practical examples of sets, which are "small" only with respect to some measure $\mu$ with $\mu(\mathbf{N}) = \infty$, this case seems not to be very interesting and for simplicity we do not consider it here.

For us will be useful also the possibility to compare two different measures.

**Definition 2.** Let $\mu_1 : S_1 \to [0,1]$ and $\mu_2 : S_2 \to [0,1]$ be two measures. We say that measure $\mu_1$ is an *extension* of measure $\mu_2$, if $\{A \in \mathbf{N} \mid \exists B \in S_2 : \ A \subseteq B \ \& \ \mu_2(B) = 0\} \subset \{A \in \mathbf{N} \mid \exists B \in S_1 : \ A \subseteq B \ \& \ \mu_1(B) = 0\}$.$\diamond$

The fact that measure $\mu_1$ is an extension of measure $\mu_2$ we denote by $\mu_1 \succ \mu_2$. Note, that for convenience we say that $\mu_1$ is an extension of $\mu_2$ only if the collection of null sets for $\mu_1$ is *strictly* larger than for $\mu_2$.

Our definition of an extension of measure reflects the fact that mainly we are interested only in properties of null sets. Note, however, that therefore our notion of extension of measure differs from the one traditionally used in mathematical analysis.

It is clear that we can define a measure over $\mathbf{N}$ in many different ways. Below are given some examples.

1. We define $\mu(A) = 0$, if $A \subseteq \mathbf{N}$ is finite, and $\mu(A) = 1$, if $A \subseteq \mathbf{N}$ is co-finite. In this way defined function $\mu$ is a measure and class of measurable sets is $S = \{X \subseteq \mathbf{N} \mid X \text{ or } \bar{X} \text{ is finite}\}$. From definition, in this case set is "small" if and only if it is finite. $EX$ identification up to such kind of "small" sets is studied, for example, in [2, 5, 9].

2. For an arbitrary $A \subseteq \mathbf{N}$ we define $\mu(A) = \lim_{n\to\infty} card(\{x \in A \mid x < n\})/n$, if $\lim_{n\to\infty} card(\{x \in A \mid x < n\})/n$ exists, otherwise $\mu(A)$ is undefined. Such function $\mu$ is a measure, and in this case "small" will be sets with zero density. Such notion of "smallness" was introduced and studied in [11, 13] and [15].

3. We define $\mu(A) = 0$, if $A \in \mathbf{N}$ is immune (or hyperimmune etc.) or finite set, and $\mu(A) = 1$, if $A \in \mathbf{N}$ is simple or co-finite set. Again, such function $\mu$ is

a measure, and in this case "small" will be subsets of immune sets. This notion of "smallness" was introduced and studied in [6].

Now, we formally define identification up to "small" set.

**Definition 3.** Set $A$ of total recursive functions is $EX$ identifiable up to null sets with respect to measure $\mu$, if there exists an inductive inference machine $M$, such that for an arbitrary function $f \in A$ there exists $k \in \mathbf{N}$, such that $\varphi_{M(f[k])} =^{\mu 0} f$ and $M(f[k]) = M(f[k'])$ for all $k' > k$.$\diamond$

Such type of identification we denote by $EX^{\mu 0}$, the same notation will stand also for the class of all $EX^{\mu 0}$ identifiable sets of functions.

In particular we will be interested in $EX$ identification in the case, when IIM is required to produce a correct answer only with some probability from the interval $(0, 1]$.

**Definition 4.** Set $A$ of total recursive functions is $EX$ identifiable with probability $p \in (0, 1]$ up to null sets with respect to measure $\mu$, if there exists an inductive inference machine $M$, such that for an arbitrary function $f \in A$ with probability $p$ there exists $k \in \mathbf{N}$, such that $\varphi_{M(f[k])} =^{\mu 0} f$ and $M(f[k]) = M(f[k'])$ for all $k' > k$.$\diamond$

The corresponding type of identification we denote by $EX_p^{\mu 0}$, the same notation we also use for the class of all $EX_p^{\mu 0}$ identifiable functions.

It is not surprising that in our situation important is not only the existence of some measure, but also the fact, whether this measure can be in some sense effectively computed.

For an arbitrary set $A \in \mathbf{N}$ we say that $a_0, a_1, \ldots$ is a characteristic sequence of $A$, if $a_i \in \{0, 1\}$ for all $i$, and $a_i = 1$ if and only if $i \in A$.

**Definition 5.** Let $S$ be a collection of measurable sets and let function $\mu : S \to [0, 1]$ be a measure. We say that $\mu$ is *computable*, if there exists a Turing machine $M$, such that

1. for an arbitrary set $X \in S$ machine $M$, when receiving a characteristic sequence $x_0, x_1, \ldots$ of $X$ on a special input tape, outputs a sequence of real numbers $r_0, r_1, \ldots$, such that $\lim_{i \to \infty} r_i$ is defined and $\lim_{i \to \infty} r_i = \mu(X)$;
2. for an arbitrary set $Y \notin S$ machine $M$, when receiving a characteristic sequence $y_0, y_1, \ldots$ of $Y$, outputs a sequence of real numbers $q_0, q_1, \ldots$, such that $\lim_{i \to \infty} q_i$ is undefined.$\diamond$

The measures from Examples 1 and 2 are computable, the measures from Example 3 are not.

## Identification up to Different Types of Small Sets

At first we state without proof some simple and useful results about measures.

**Proposition 6.** *Let $\mu$ be an arbitrary measure. Then $\mu(A) = 0$ for all finite sets $A \subset \mathbf{N}$, and $\mu(B) = 1$ for all co-finite sets $B \subseteq \mathbf{N}$.*$\diamond$

**Proposition 7.** *Let $\mu_1$ and $\mu_2$ be computable measures, such that $\mu_1 \succ \mu_2$. Then there exists a recursive set $T \subseteq \mathbf{N}$, such that $\mu_2(T) \neq 0$ and $\mu_1(T) = 0$.*$\diamond$

**Proposition 8.** *Let $\mu$ be a computable measure, such that for an arbitrary $l \in \mathbf{N}_+$ there exists a set $A \subseteq \mathbf{N}$ with $0 < \mu(A) \leq 1/l$. Then there exists a function $F \in \mathcal{R}_2$, such that for some $t' \in \mathbf{N}$, for all $t > t'$ and for all $l \in \mathbf{N}_+$ set $W_{F(l,t')}$ is recursive, and we have $0 < \mu(W_{F(l,t')}) \leq 1/l$ and $F(l,t) = F(l,t')$.*$\diamond$

Our first theorem shows that regardless of the particular measure the class of all total recursive functions can not be identified in the limit up to sets with zero measure.

**Theorem 9.** *For an arbitrary measure $\mu$ we have $\mathcal{R} \notin EX^{\mu 0}$.*$\diamond$

This result is an easy consequence from Theorem 12.

Further we show that, if $\mu_1$ and $\mu_2$ are two *computable* measures which can be compared, and are such that $\mu_1 \succ \mu_2$, then up to small sets with respect to $\mu_1$ we can identify a strictly larger class of functions, than up to small sets with respect to $\mu_2$.

**Theorem 10.** *Let $\mu_1$ and $\mu_2$ be arbitrary computable measures, such that $\mu_1 \succ \mu_2$. Then $EX^{\mu_2 0} \subset EX^{\mu_1 0}$.*$\diamond$

*Proof.* The fact that $EX^{\mu_2 0} \subseteq EX^{\mu_1 0}$ is trivial. To prove that $EX^{\mu_1 0} - EX^{\mu_2 0} \neq \emptyset$ we consider set of total recursive functions $A = \{f \in \mathcal{R} \mid f =^{\mu_1 0} \varphi_{f(0)}\}$. It is obvious that $A \in EX^{\mu_1 0}$. To show that $EX^{\mu_1 0} - EX^{\mu_2 0} \neq \emptyset$ for each inductive inference machine $M$ we will construct a function $f_M \in A$, such that $M$ will not identify $f_M$ (up to null sets with respect to measure $\mu_2$).

Let $T$ be a recursive set which exists for measures $\mu_1$ and $\mu_2$ due to Proposition 7. We construct two functions $g_M$ and $h_M$ by the following procedure.

<u>*Step* 1.</u>
We define $g_M(0) = \#g_M$, $h_M(0) = \#g_M$, $H = M(g_M[0])$.

<u>*Step* $k, k > 1$.</u>
We define $g_M(k-1) = 0$ and, if $k - 1 \notin T$, we also define $h_M(k-1) = 0$. Then we compute $M(g_M[k-1])$.

If $M(g_M[k-1]) \neq H$ we define $h_M(x) = g_M(x)$ for all $x < k$ for which $h_M(x)$ is undefined and define $H = M(g_M[k-1])$.

<u>*End.*</u>

Let $h'_M$ be a total recursive function, such that for all $x \in \mathbf{N}$

$$h'_M(x) = \begin{cases} h_M(x), & \text{if } h_M(x) \downarrow, \\ 1, & \text{if } h_M(x) \uparrow. \end{cases}$$

We have $g_M, h'_M \in A$, because $g_M(0) = \#g_M$ and $g_M =^{\mu_1 0} h'_M$. If machine $M$ on function $g_M$ changes hypothesis infinitely often, then we can take $f_M = g_M$. Otherwise $M$ will produce an incorrect hypothesis (up to null set with respect to measure $\mu_2$) either on function $g_M$, or $h'_M$. We can take $f_M$ to be equal with the corresponding of these functions.$\diamond$

As it can be seen from our next theorem, the requirement that $\mu_1$ and $\mu_2$ are computable is essential for our previous result to hold.

**Theorem 11.** *There exist measures $\mu_1$ and $\mu_2$, such that $\mu_1 \succ \mu_2$ and $EX^{\mu_1 0} = EX^{\mu_2 0}$.*$\diamond$

*Proof.* Let $U_0, U_1, U_2, \ldots$ be a sequence of all finite unions of recursively enumerable sets with infinite complements. We consider collections of sets $S_1 = \{A \subseteq \mathbf{N} \mid \forall i \in \mathbf{N}\ \exists x \in A \cap U_i\}$ and $S_2 = \{A \subseteq \mathbf{N} \mid A \text{ is finite}\}$. (Meaningfully $S_1$ contains all infinite sets that are sparser than any infinite set with recursively enumerable complement.) Let $\mu_1 : S_1 \cup \bar{S}_1 \to \{0,1\}$ and $\mu_2 : S_2 \cup \bar{S}_2 \to \{0,1\}$ be measures with $\mu_1(A_1) = \mu_2(A_2) = 0$, if $A_1 \in S_1$ and $A_2 \in S_2$. It is not hard to check that $\mu_1$, $\mu_2$ are well defined, and that $\mu_1 \succ \mu_2$. At the same time $EX^{\mu_1 0} = EX^{\mu_2 0}$, because there do not exist functions $f, g \in \mathcal{P}$, such that $f =^{\mu_1 0} g$ and $f \neq^{\mu_2 0} g$.$\diamond$.

## Probabilistic Behaviour of Identification up to Small Sets

At first, we show that for identification up to null sets holds the traditional result for probabilistic $EX$ identification, namely, that with probability $1/(n+1)$ we can identify a strictly larger class of functions, than with probability $1/n$.

**Theorem 12.** *For an arbitrary measure $\mu$ and for an arbitrary number $n \in \mathbf{N}_+$ we have $EX^{\mu 0}_{1/(n+1)} - EX^{\mu 0}_{1/n} \neq \emptyset$.*$\diamond$

*Proof.* For brevity here and also further in this chapter we will show the proof only for the simplest interesting case (here for $n = 1$). As a rule, in all cases it will be sufficient to illustrate the main idea of proof. We will also use the fact that $EX^{\mu 0}_1 = EX^{\mu 0}$, formal proof of which we omit here (this fact actually follows from Theorem 13).

We define set of total recursive functions

$$A = \{f \in \mathcal{R} \mid f = \varphi_{f(0)} \vee \exists m \in \mathbf{N}_+ : f = \varphi_{f(m)\langle 1\rangle} \text{ and } \forall m' > m : f(m')\langle 1\rangle = 0\}.$$

It is clear that for each $f \in A$ with probability $1/2$ we can algorithmically compute $\#f$, since it will be either $f(0)$, or the largest $m$, such that $f(m)\langle 1\rangle \neq 0$. Thus $A \in EX^{\mu 0}_{1/2}$.

It remains to show that $A \notin EX^{\mu 0}$. For each inductive inference machine $M$ we will construct a function $f_M \in A$, such that $M$ will not identify $f_M$.

We construct a function $g_M$ by the following procedure.

*Step 1.*

We define $g_M(0) = \#g_M$, $d = 0$, $H_1 = M(g_M[0])$ and $H_2 = M(g_M[0])$. Meaningfully $d$ is the length of initial fragment on which $g_M$ is already defined, $H_1$ and $H_2$ are the last hypotheses produced by $M$ on some variants of function $g_M$.

*Step $k, k > 1$.*

We compute $M(h_{i,k}[k-1])$ for functions $h_{1,i}$ and $h_{2,i}$, where for all $x \in \mathbf{N}$

$$h_{i,k}(x) = \begin{cases} g_M(x), & \text{if } x < d, \\ \langle \#h_{i,k}, i\rangle, & \text{if } x = d, \\ \langle 0, i\rangle, & \text{if } x > d. \end{cases}$$

If $M(h_{1,k}[k-1]) \neq H_1$, then we define $g_M(x) = h_{1,k}(x)$ for all $x \leq k$ for which $g_M(x)$ is undefined, define $H_1 = M(h_{1,k}[k-1])$, and define $d = k$.

Similarly, if $M(h_{2,k}[k-1]) \neq H_2$, then we define $g_M(x) = h_{2,k}(x)$ for all $x \leq k$ for which $g_M(x)$ is undefined, define $H_2 = M(h_{2,k}[k-1])$, and define $d = k$.

*End.*

If $g_M$ is total, then $g_M \in A$ and, at the same time, machine $M$ does not identify function $g_M$. Thus, we can take $f_M = g_M$.

If $g_M$ is not total, then for some $d' \in \mathbf{N}$ it is undefined for all $x \geq d'$. By Proposition 6 machine $M$ produces an incorrect hypothesis either on function $h_{1,d'}$, or on function $h_{2,d'}$. We can take $f_M$ to be equal with the corresponding of these functions.$\diamond$

However, in general it is not true that there do not exist any intermediate learning powers between $EX^{\mu 0}_{1/n}$ and $EX^{\mu 0}_{1/(n+1)}$ (however, such result holds, for example, for $EX$ identification up to finite sets). For an arbitrary measure $\mu$ we can only show that there do not exist any intermediate learning powers between $EX^{\mu 0}_{n/(2n-1)}$ and $EX^{\mu 0}_{(n+1)/(2n-1)}$. Thus, at least in interval of probabilities $(1/2, 1]$ power of identification up to null sets grows in discrete steps, when we are lowering requirements on probability with which identification must be correct.

**Theorem 13.** *For an arbitrary measure $\mu$ and for arbitrary numbers $n \in \mathbf{N}_+$ and $p \in ((n+1)/(2n+1), n/(2n-1)]$ we have $EX^{\mu 0}_p = EX^{\mu 0}_{n/(2n-1)}$.*$\diamond$

*Proof.* We will show the proof for the case $n = 1$.

Let $p \in (2/3, 1]$ and $A \in EX^{\mu 0}_p$. Let $M_p$ be an inductive inference machine that $EX$ identifies an arbitrary function $f \in A$ up to null set (with respect to $\mu$) with probability $p$. We will construct another inductive inference machine $M$ that does the same deterministically (therefore, with probability 1). It will show that $A \in EX^{\mu 0}_1$, and thus, $EX^{\mu 0}_p = EX^{\mu 0}_1$.

For each $k \in \mathbf{N}$ and for each $f \in A$ we define the set $H(k, f)$ as follows:

$$H(k,f) = \{(h,q) \mid h \in \mathbf{N}, q \in [0,1] : M_p(f[k]) = h \text{ with probability } p\}.$$

For an arbitrary set of pairs $X \subseteq \mathbf{N} \times [0,1]$ and arbitrary $i \in \mathbf{N}$ let $Prob(X) = \sum_{(j,q)\in X} q$ and $prob(i,X) = \sum_{(i,q)\in X} q$. For such sets of pairs $X_1, X_2 \subseteq \mathbf{N} \times [0,1]$ we also define

$$X_1 \sqcap X_2 = \{(i,q) \mid (i,q) \in X_1 \cup X_2 \ \& \ \nexists \ q' < q : (i,q') \in X_1 \cup X_2\}.$$

We construct an inductive inference machine $M$ as follows.

For each function $f \in A$ machine $M$ initially defines sets $H_1 = \emptyset$ and $H_2 = \emptyset$. Then for each $k \in \mathbf{N}$ machine $M$ computes $H(k,f)$.

1. If $H_1 = \emptyset$, then $M$ defines $H_1 = H_2 = H(k,f)$ and outputs the number $\#f_k$, where $f_k \in \mathcal{P}$ and is computed by the following procedure $P$.

   For each $(h,q) \in H_2$ and $x \in \mathbf{N}$ procedure $P$ simulates the computation of $\varphi_h(x)$, until it finds $y \in \mathbf{N}$, such that

$$\sum_{i\in\mathbf{N}:\ \varphi_i(x)=y} prob(i,H_2) > Prob(H_2)/2.$$

   Then it outputs $y$ as the value of $f_k(x)$.
2. If $H_1 \neq \emptyset$ and $Prob(H(k,f) \sqcap H_1) \leq 1/3$, then $M$ defines $H_1 = H_2 = H(k,f)$ and outputs the number $\#f_k$, computed as in Case 1.
3. If $H_1 \neq \emptyset$ and $Prob(H(k,f) \sqcap H_1) > 1/3$, then M defines $H_2 = H_2 \sqcap H(k,f)$ and outputs the number $\#f_k$, computed as in Case 1.

Since we know that machine $M_p$ identifies function $f$ with probability larger than $2/3$, then only Case 3 can occur more than a finite number of times. Eventually more than a half (with respect to the sum of probabilities) of functions from $H_2$ will be correct numbers for $f$ (up to null sets), thus starting from some $k' \in \mathbf{N}$ majority voting procedure will compute a function equal with $f$ up to a union of a finite number of null sets. Since such union of null sets is a null set, we will have $f =^{\mu 0} f_k$ for all $k > k'$.$\diamond$

However, there are measures for which identifiability of some class of functions with probability $p > 1/(n+1)$ implies also identifiability of the same class with probability $1/n$. If we restrict our attention to computable measures, then we are able to give a complete characterization of measures for which this is the case. Namely, this result will hold only for measures for which there do not exist sets $A$ with $0 < \mu(A) \leq \varepsilon$ for an arbitrary $\varepsilon \in (0,1]$. Moreover, for all computable measures for which the property mentioned above does not hold, we can identify a strictly larger class of functions with probability $(n+1)/(2n+1)$ than with probability $n/(2n-1)$.

**Theorem 14.** *Let $\mu$ be a computable measure, such that for an arbitrary $l \in \mathbf{N}_+$ there exists a set $A \subseteq \mathbf{N}$ with $0 < \mu(A) \leq 1/l$. Then for an arbitrary $n \in \mathbf{N}_+$ we have $EX^{\mu 0}_{(n+1)/(2n+1)} - EX^{\mu 0}_{n/(2n-1)} \neq \emptyset$.*$\diamond$

*Proof.* We will show the proof for the case $n = 1$.

We define set of total recursive functions

$$A = \{f \in \mathcal{R} \mid \exists m > 1,\ \exists a, b \in \{f(0), f(1), f(m)\} : \\ a \neq b \;\&\; f =^{\mu 0} \varphi_{a\langle 1\rangle} =^{\mu 0} \varphi_{b\langle 1\rangle} \;\&\; \forall m' > m : f(m')\langle 1\rangle = 0\}.$$

Clearly, we have $A \in EX^{\mu 0}_{2/3}$, because for each $f \in A$ with probability $2/3$ we can algorithmically compute $\#f$ – at least two elements from the set $\{f(0)\langle 1\rangle, f(1)\langle 1\rangle, f(m)\langle 1\rangle\}$, where $m$ is the largest number with $f(m)\langle 1\rangle \neq 0$, are correct numbers (up to null sets) for function $f$.

It remains to show that $A \notin EX^{\mu 0}$. For each inductive inference machine $M$ we will construct a function $f_M \in A$, such that $M$ will not identify $f_M$.

We construct two functions $g_M$ and $h_M$ by the following procedure. Here $F$ will be the function, which exist for the measure $\mu$ due to Proposition 8.

*Step* 1.

We define $g_M(0) = h_M(0) = \langle \#g_M, 0\rangle$, $g_M(1) = h_M(1) = \langle \#h_M, 0\rangle$, define $d = 1$, $G = M(g_M[0])$ and $H = M(h_M[0])$, $S = W_{F(d,0)}$. Here again $G$ and $H$ are the last hypotheses produced by $M$ on $g_M$ and $h_M$, $d$ approximately is a number of hypotheses produced by $M$, $S$ is some set on which we will try to define functions $g_M$ and $h_M$ differently.

*Step* $k, k > 1$.

We compute $M(g_M[k])$, $M(h_M[k])$ and $W_{F(d,k)}$.

If $M(g_M[k]) = G$, $M(h_M[k]) = H$ and $W_{F(d,k)} = S$, we define $g_M(k+1) = \langle 0, 0\rangle$, define $h_M(k+1) = \langle 0, 0\rangle$, if $k+1 \notin S$, and define $h_M(k+1) = \langle 0, 1\rangle$, if $k+1 \in S$. Then we go to the next step.

If $M(g_M[k]) \neq G$ or $M(g_M[k]) \neq H$, we define $G = M(g_M[k])$, $H = M(h_M[k])$, $d = d+1$, $S = W_{F(d,k)}$.

If $M(g_M[k]) = G$, $M(h_M[k]) = H$, but $W_{F(d,k)} \neq S$, we define $S = W_{F(d,k)}$.

In both these cases we further define $g_M(k+1) = \langle \#g', 0\rangle$ and $h_M(k+1) = \langle \#h', 0\rangle$, where for all $x \in \mathbf{N}$

$$g'(x) = \begin{cases} g_M(x), & \text{if } x \leq k, \\ \langle \#g', 0\rangle, & \text{if } x = k+1, \\ \langle 0, 0\rangle, & \text{if } x > k+1, \end{cases}$$

$$h'(x) = \begin{cases} h_M(x), & \text{if } x \leq k, \\ \langle \#h', 0\rangle, & \text{if } x = k+1, \\ \langle 0, 0\rangle, & \text{if } x > k+1 \text{ and } x \notin S, \\ \langle 0, 1\rangle, & \text{if } x > k+1 \text{ and } x \in S. \end{cases}$$

Then we go to the next step.

*End.*

It is not hard to see that $g_M \in A$ and $h_M \in A$, besides that $g_M =^{\mu 0} h_M$ if and only if machine $M$ on one of these functions changes hypothesis infinitely often. In this case we can take $f_M$ to be equal with the corresponding of these functions.

Otherwise we will have $\mu(\{x \in \mathbf{N} \mid g_M(x) \neq h_M(x)\}) \neq 0$, thus $M$ cannot produce a correct hypothesis (up to null sets) for both of these functions. Hence, we can take $f_M$ to be the one of these functions on which $M$ hypothesis is wrong.$\diamond$

**Theorem 15.** *Let $\mu$ be a computable measure, such that for some $l \in \mathbf{N}_+$ there does not exist set $A \subseteq \mathbf{N}$ with $0 < \mu(A) \leq 1/l$. Then for arbitrary numbers $n \in \mathbf{N}_+$ and $p \in (1/(n+1), 1/n]$ we have $EX_p^{\mu 0} = EX_{1/n}^{\mu 0}$.*$\diamond$

*Proof.* We will show the proof for the case $n = 1$.

Let $p \in (1/2, 1]$ and $A \in EX_p^{\mu 0}$. Let $M_p$ be an inductive inference machine that $EX$ identifies an arbitrary function $f \in A$ up to null set (with respect to $\mu$) with probability $p$. We will construct another inductive inference machine $M$ that identifies an arbitrary function $f \in A$ deterministically.

Let $H(k, f)$, $Prob(X)$ and $prob(i, X)$ be defined as in proof of Theorem 13. For function $f \in A$ and numbers $h, k \in \mathbf{N}$ we define $E(f, h, k)$ to be equal with number produced by Turing machine $M_\mu$, which computes measure $\mu$, after receiving the sequence $x_0, x_1, \ldots$, where $x_i = 1$, if computation of $\varphi_h(i)$ converges after $k$ steps of simulation with result $\varphi_h(i) \neq f(i)$, and $x_i = 0$ otherwise.

We construct an inductive inference machine $M$ as follows.

For each function $f \in A$ machine $M$ initially defines set $H = \emptyset$. Then for each $k \in \mathbf{N}$ machine $M$ computes $H(k, f)$.

1. If $Prob(H) \leq 1/2$, then $M$ defines $H(k, f)$ and outputs the number $\#f_k$, where $f_k \in \mathcal{P}$ and is computed by the following procedure $P$.

   For each $(h, q) \in H$ and $x \in \mathbf{N}$ procedure $P$ simulates the computation of $\varphi_h(x)$, until it finds $y \in \mathbf{N}$ and $(i, q) \in H$, such that $q > 0$ and $\varphi_i(x) = y$. Then it outputs $y$ as the value of $f_k(x)$.
2. If $Prob(H) > 1/2$ and $Prob(H(k, f) \sqcap H) \leq 1/2$, then $M$ defines $H(k, f)$ and outputs the number $\#f_k$ computed as in Case 1.
3. If $Prob(H) > 1/2$ and $Prob(H(k, f) \sqcap H) > 1/2$, then M defines $H = \{(i, q) \in H' \mid E(f, i, k) \leq 1/2l\}$ and outputs the number $\#f_k$ computed as in Case 1.

Since we know that machine $M_p$ identifies function $f$ with probability larger than $1/2$, then only Case 3 can occur more than a finite number of times. Starting from some time set $H$ will contain only numbers of functions that can converge only on null sets to values different from the values of $f$. Therefore, starting from some $k' \in \mathbf{N}$ we will have $f =^{\mu 0} f_k$ for all $k > k'$.$\diamond$

It is an interesting question, whether Theorem 15 holds also for measures that are not computable. However, we expect that most probably this is not the case.

## Conclusions

We have shown that for computable measures we always can $EX$ identify up to "small" sets a larger class of functions, if we allow for larger part of sets to be "small". Besides that, we have obtained a complete characterization of behaviour of such identification in probabilistic case (at least for interval of probabilities $(1/2, 1]$). We have also obtained several results about identification up to null sets for measures that are not computable. However, several interesting questions remain open and we will try to indicate them below.

First, we have studied in this paper only so called $EX$ identification. To consider $BC$ identification up to null sets will not be very promising, since it is known that $\mathcal{R} \in BC^*$ (i.e. class of all total recursive functions is identifiable up to finite sets), and we know that all finite sets must have zero measure. However, it seems that $FIN$ identification up to "small" sets could be interesting and worth to study.

Other, and probably more interesting, problem concerns the generalization of Theorems 13 and 14 for probabilities from interval $(0, 1/2]$. The probabilistic hierarchy, given by these theorems for probabilities from interval $(1/2, 1]$ first was discovered in [4] for $FIN$ identification. However, probabilistic behaviour of $FIN$ identification for probabilities below $1/2$ turns out to be very complicated and still is known only partially. The same hierarchy $1, 2/3, 3/5, 4/7, \ldots$ of probabilities appears also for some other types of $EX$ identification (see [16]). At the same time, in all such cases nothing is known for $EX$ identification with probabilities smaller than $1/2$. Therefore, any results in this direction would be very interesting.

Finally, while our definition of computable measures appears to be very natural, it seems that it is still too restrictive. In particular, while measures from Example 3 are not computable by our definition, the results of Theorem 15 and some variant of Theorem 10 still hold for them. Thus, it could be interesting either to try somehow to extend the notion of computable measures, or to change the notion of computability with some more general property.

## Acknowledgements

The author of the paper wish to thank professor Rūsiņš Freivalds for the idea to use the notion of measure for characterization of "small" sets.

## References

1. L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.
2. J. Case and C. Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.
3. L. Fortnov, R. Freivalds, W. Gasarch, M. Kummer, S. Kurtz, C. Smith, and F. Stephan. Measure, complexity and learning theory. Technical Report TR-95-01, University of Chicago, 1995.

4. R. Freivalds. Finite identification of general recursive functions by probabilistic strategies. In *Proceedings of the Conference on Algebraic, Arithmetic and Categorial Methods in Computation Theory FCT'79*, pages 138–145, 1979.
5. R. Freivalds, C. Smith, and M. Velauthapillai. Trade-off among parameters affecting inductive inference. *Information and Computation*, 82:323–349, 1989.
6. R. Freivalds and J. Vīksna. Inductive inference up to immune sets. In *Proceedings of the International Workshop on Analogical and Inductive Inference AII'89*, volume 397 of *Lecture Notes in Computer Science*, pages 138–147, 1989.
7. W. Gasarch, R. Sitaraman, C. Smith, and M Velauthapillai. Learning programs with an easy to calculate set of errors. In *Proceedings of the International Workshop on Analogical and Inductive Inference AII'89*, volume 397 of *Lecture Notes in Computer Science*, pages 124–137, 1989.
8. E. Kinber and T. Zeugmann. Inductive inference of almost everywhere correct programs by reliably working strategies. *Journal of Information Processing and Cybernetics*, 21:91–100, 1985.
9. L. Pitt. Probabilistic inductive inference. *Journal of the ACM*, 36:383–433, 1989.
10. H. Rogers. *Theory of Recursive Functions and Effective Computability.* MIT Press, 1987.
11. J. Royer. Inductive inference of approximations. *Information and Computation*, 70:156–178, 1986.
12. C. Smith and M. Velauthapillai. On the inference of programs approximately computing the desired function. In *Proceedings of the International Worksop on Analogical and Inductive Inference AII'86*, volume 265 of *Lecture Notes in Computer Science*, pages 164–176, 1986.
13. C. Smith and M. Velauthapillai. On the inference of approximate programs. *Theoretical Computer Science*, 77:249–266, 1990.
14. R.I. Soare. *Recursively Enumerable Sets and Degrees*. Springer Verlag, 1987.
15. J. Vīksna. Probabilistic inference of approximations. In *Proceedings of the Conference of Nonmonotonic Logic and Inductive Inference*, volume 659 of *Lecture Notes in Computer Science*, pages 323–332, 1991.
16. J. Vīksna. Probabilistic inference of limiting functions with bounded number of mind changes. *International Journal of Foundations of Computer Science*, 1995. Accepted.
17. J. Vīksna. Some general properties of probabilistic EX inductive inference. Technical Report LU-IMCS-TRCS-95-3, Institute of Mathematics and Computer Science, University of Latvia, 1995.