# Category, Measure, Inductive Inference: A Triality Theorem and its Applications *

Rūsiņš Freivalds†
Institute of Math and Computer Science
University of Latvia
Raiņa bulvāris 29
LV-1459, Riga, Latvia
rusins@cclu.lv

Carl H. Smith ‡
Institute for Logic, Language and Computation
University of Amsterdam
Plantage Muidergracht 24
NL-1018 TV Amsterdam
The Netherlands
csmith@fwi.uva.nl

**Abstract**

The famous Sierpinski-Erdös Duality Theorem [Sie34b, Erd43] states, informally, that any theorem about effective measure 0 and/or first category sets is also true when all occurrences of "effective measure 0" are replaced by "first category" and vice versa. This powerful and nice result shows that "measure" and "category" are equally useful notions neither of which can be preferred to the other one when making formal the intuitive notion "almost all sets." Effective versions of measure and category are used in recursive function theory and related areas, and resource-bounded versions of the same notions are used in Theory of Computation. Again they are dual in the same sense.

We show that in the world of recursive functions there is a third equipotent notion dual to both measure and category. This new notion is related to learnability (also known as inductive inference or identifiability). We use the term "triality" to describe this three-party duality.

## 1 Introduction

Mathematicians have invented many ways to determine the relative sizes of infinite sets. Cardinality is the most natural technique, but it considers infinite sets as either denumerable or undenumerable. Measure is a generalization of volume and area. Hence, it is the most popular method employed to show that one set of infinite cardinality is larger than another set of the same cardinality. Hence, it is widely used to discuss the relative size of of infinite sets. However there are other ways how to say that one set is smaller than the other one. Category has become a popular notion to discuss smallness of sets [Oxt71]. W. Sierpinski [Sie34b] and P. Erdös [Erd43] proved a fundamental theorem with the following important consequence.

**Theorem 1 (Duality Principle)** *Let $P$ be an assertion containing only notions of "set of measure 0", "set of the first category" and the notions from set theory. Let $P^*$ be an assertion obtained by interchange of the notions "set of measure 0" and "set of the first category." Then each of the assertions $P$ and $P^*$ is implied by the other one, provided the Continuum Hypothesis holds.*

*This project was supported by an International Agreement under NSF Grant 9421640.

†The first author was supported by Latvian Science Council Grant No. 93.599

‡On leave from the Department of Computer Science, University of Maryland, College Park, MD 20742 USA, The second author was supported in part by NSF Grant 9301339 and a grant from the Netherlands Organization for Scientific Research.

Hence, by the Duality Principle, "measure zero" and "first category" are dual notions (like disjunction and conjuction, or existence and universality) neither of which can be preferred to the other one.

The set of all the (0-1)-valued total recursive functions is denumerable. Hence, from the viewpoint of the classical mathematics this is a small set, e.g. a set of measure zero and of the first category. However two decades ago counterparts of the notions "measure zero" and "first category" were defined for the set of all the (0-1)-valued total recursive functions (see [Lis81]) and they also were found to be dual in the same sense, without any reference to the Continuum Hypothesis. In recent papers resource-bounded versions of the traditional measure and category notions have been developed [AS95, ASMWZ96, BM95a, BM95b, BL96, Fen91, Fen95, Kum96, Lut93, Lut96]. For instance, K. Melhhorn [Meh73] and J. Lutz [Lut93] used constructive notions of category and measure to study subrecursive degree structures. The Sierpinski and Erdös Duality Principle is known to hold for the resource-bounded versions of the notions of category and measure as well.

It is a well-known general observation that it is difficult to invent natural notions on a high abstract level but it is much easier to invent such natural notions on a lower level where there is a considerable amount of specific information. Hence. it is no wonder that on the level of classical mathematics there have been no attempts to invent a third notion dual to both measure and category while at the level of recursive functions this is achieved in this paper, opening even more possibilities on the resource-bounded level.

For recursive functions there is a third possible approach to make precise the notion "the set of functions is small," namely "the set of functions is identifiable." Gold[Gol67] considered the reconstruction of computer programs from sample computations. Nowadays this process is called "learning," "identification" or "inductive inference." The values $f(0), f(1), f(2), \ldots$ of a function are given to the identifying algorithm, and a valid program which computes this function, is to be constructed. Various types of identification are considered (finite identification, identification in the limit, -see [AS83, Fre91]). It turns out that the class of all recursive functions is not identifiable (according to any of these identification types). Identifiable sets of recursive functions might be considered "small," however not in terms of measure or category.

In [FFG$^+$] an attempt was made to discuss the relative sizes of learnable sets using effective notions of measure and category. It was shown that, for all known identification types, the complements of all identifiable sets of functions are huge both in measure and in category. This seemed to eliminate any possibility of proving an analogue of the Duality Principle. In [AFS96] another, more general inference type, was introduced with the intent of trying to find an identification type which might be useful to construct the third notion dual simultaneously to measure and category. In this paper we have succeeded in finding such an identification type. The new type is very closely related to one considered in [BFS96].

## 2 Technical Preliminaries

The natural numbers are denoted by $\mathbb{N}$. For functions $f$ and $g$, $f \sqsubseteq g$ means that $f(x) = g(x)$ for all $x$ in the domain of $f$. This work is concerned primarily with the recursive (computable) functions as formalized in [MY78, Smi94a] and many other places. $\mathcal{R}$ is the set of $\{0, 1\}$ valued recursive functions. This paper is also concerned with the ideas of measure and category that are defined over point sets in real interval $[0, 1]$. In order to compare the learning of recursive functions with with either measure or category, it is necessary to represent functions as points in the interval $[0, 1]$. To do so, it will be convenient to represent a function by a sequence of values from its range. Such a representation is called a *string* representation. So, for example, the sequence $01^2 0^4 1^\infty$ represents the (total) function:

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \text{ or } 3 \leq x \leq 6, \\ 1 & \text{if } 1 \leq x \leq 2 \\ 1 & \text{otherwise.} \end{cases}$$

We tersely give review the notions of effective measure and category, omitting many of the details. The idea of *measure* comes from [Bor05, Bor14], see [Oxt71]. In the recursion-theoretic setting it is more convenient to use the martingale functions [Lut93]. "Small" sets in terms of measure are those with effective measure 0.

Next we review the the notion of category [Kur58, Oxt71]. A set $A$ is *nowhere dense* if every interval has a subinterval contained in the complement of $A$. A set is of the *first category* if it can be represented as a countable union of nowhere dense sets. "Small" sets in terms of category are the sets of the first category. Sets of the first category can also be characterized in terms of the Banach-Mazur game [Oxt71]. The recursion-theoretic counterpart of the "set of the first category" is "effectively meager set" as defined by K. Mehlhorn [Meh73]. L. Lisagor showed that Mehlhorn's definition was equivalent to a Banach-Mazur game [Lis81].

**Definition 1** *(Due to Mehlhorn [Meh73]) A set $C \subseteq \mathcal{R}$ is effectively meager iff there is a sequence $\{h_k\}_{k\in\omega}$ of uniformly recursive functions $h_k : \{0,1\}^\star \to \{0,1\}^\star$ such that $\sigma \sqsubseteq h_k(\sigma)$ for all $k, \sigma$, and for every $f \in C$ there is $k$ such that $h_k(\sigma) \not\sqsubseteq f$ for all $\sigma$. (This formalizes that $C$ is contained in an effective union of effectively nowhere dense sets.)*

Inductive inference emerged when the philosophers became interested in modeling the scientific method [Put75]. An *inductive inference machine* (IIM) is an algorithmic device that inputs the graph of a recursive function and, while doing so, outputs a sequence of computer programs. Typically, an IIM $M$ learns a recursive function $f$, if $M$, when given the graph of $f$ as input, outputs a sequence a of programs that converges to a program that computes $f$. Each IIM $M$ learns some set of recursive functions denoted by $EX(M)$. $EX$ is used to denote the collection of all such sets as $M$ varies across IIMs. Several variations of the Gold model of learning in the limit have been discussed in the literature. Each such variation results in what is called an *identification type*. Our results will apply to some, but far from all, of the investigated identification types.

# 3 The Triality Theorem

The proof of our main result makes extensive use of the following lemma.

**Lemma 2** *Let $\mathcal{X}$ be a subset of $\mathcal{R}$. Let $G$ be a family of sets $G \subset 2^{\mathcal{X}}$ such that*

1. *$G$ is an ideal in $2^{\mathcal{X}}$, i.e.*
   *(a) $[c \in G$ and $d \in G] \Rightarrow c \cup d \in G$; and*
   *(b) $[a \in 2^{\mathcal{X}}$ and $c \in G] \Rightarrow a \cap c \in G$;*
2. *the union of all the sets in $G$ is $\mathcal{X}$;*
3. *$G$ has a countable subfamily $K$ such that an arbitrary element of $G$ is contained in some element of $K$; and*
4. *the complement of every set in $G$ contains an infinite set from $G$.*

*Then $\mathcal{X}$ can be divided into an infinite sequence of disjoint infinite sets $\{\mathcal{X}_n\}$ such that an arbitrary subset $E \subset \mathcal{X}$ belongs to $G$ iff $E$ is a subset of a finite union of sets in $\{\mathcal{X}_n\}$.*

Proof: Since the subfamily $K$ is countable, let $K_0$, $K_1$, $\cdots$ be a (non effective) sequence of the sets in $K$ such that $K_0$ is infinite. Let $H_n$ denote $K_0 \cup K_1 \cup \cdots \cup K_n$. Define a sequence of sets by:

$$L_n = \begin{cases} H_0 & \text{if } n = 0; \\ \left(H_n - \bigcup_{i=0}^{n-1} L_i\right) & \text{if } n \neq 0 \text{ and the set } (H_n - \bigcup_{i=0}^{n-1} L_i) \text{ is infinite;} \\ \emptyset & \text{otherwise} \end{cases}$$

It follows from (4.) that there are infinitely many sets in $\{K_n\}$, and there are infinitely many infinite sets in $\{L_n\}$. Hence the union $\cup_{n=1}^{\infty} L_n = \mathcal{X}$. By construction, the sets $L_n$ are either empty or infinite. Now, we remove the empty sets from the list $\{L_n\}$ but do not change the relative ordering of the remaining sets. The new list is denoted by $\{\mathcal{X}_n\}$. These sets are disjoint and infinite. Furthermore, their union is $\mathcal{X}$. If $E \subset G_\alpha \in G$, then it follows from (3.) that there is a $K_n$ such that $E \subset K_n$ and *either* $E \subset L_n$ (if $L_n$ is nonempty) *or* $E \subset L_m$ (if $L_n$, $L_{n+1}$, $\cdots$, $L_{m-1}$ are all empty and $L_m$ is nonempty). Hence, $E$ is in the finite union $\cup_{i=0}^{n} \mathcal{X}_i$ or $\cup_{i=0}^{m} \mathcal{X}_i$, depending on the emptyness of $L_n$. By (1a), any finite union of sets in $\{\mathcal{X}_n\}$ is in $G$, and any subset of $E$ of such a finite union is in $G$ by (1b). ☒

Lemma 2 has three corollaries, one for each of measure, category and identifiability.

**Corollary 3** *Let $\mathcal{X}$ be a subset of $\mathcal{R}$. Let $G$ be the family of sets of $\{0,1\}$ valued functions belonging to $\mathcal{X}$ with effective measure 0. Then $\mathcal{X}$ can be divided into an infinite sequence of disjoint infinite sets $\{\mathcal{X}_n\}$ such that an arbitrary subset $E$ of $\mathcal{X}$ is of effective measure 0 iff $E$ is a subset of a finite union of the sets in $\{\mathcal{X}_n\}$.*

Proof: The family $G$ of the sets with effective measure 0 satisfy (1) through (4) of Lemma 2 with the subfamily $K$ consisting of all the sets of $\{0,1\}$ valued functions $f$ such that a certain partial recursive martingale succeeds on $f$. The family $G$ is of the cardinality $2^{\aleph_0}$ while the family $K$ is countable. ☒

**Corollary 4** *Let $\mathcal{X}$ be a subset of $\mathcal{R}$. Let $G$ be the family of all the effectively meager sets of $\{0,1\}$ valued functions belonging to $\mathcal{X}$. Then $\mathcal{X}$ can be divided into an infinite sequence of disjoint infinite sets $\{\mathcal{X}_n\}$ such that an arbitrary subset $E$ of $\mathcal{X}$ is effectively meager iff $E$ is a subset of a finite union of the sets in $\{\mathcal{X}_n\}$.*

Proof: The family $G$ of all the effectively meager sets satisfy (1) through (4) of Lemma 2 with the subfamily $K$ consisting of all the sets of $\{0,1\}$ valued functions defined by a partial recursive strategy of one player in the Banach-Mazur game (against all possible partial recursive strategies of the other player). ☒

To tie these notions in with inductive inference, we must first give names to the collection of identification types that our results apply to.

**Definition 2** *An identification type $I$ is* smooth *if:*

1. *All the finite classes of recursive $\{0,1\}$ valued functions are $I$-identifiable,*

2. *$\mathcal{R}$ is not $I$-identifiable,*

3. *any subclass of any I-identifiable class is I-identifiable, and*

4. *the union of any two I-identifiable classes is I-identifiable.*

**Corollary 5** *Let $\mathcal{X}$ be a subset of $\mathcal{R}$. Let $G$ be the family of the I-identifiable sets of $\{0,1\}$ valued functions belonging to $\mathcal{X}$. Then $\mathcal{X}$ can be divided into an infinite sequence of disjoint infinite sets $\{\mathcal{X}_n\}$ such that an arbitrary subset $E$ of $\mathcal{X}$ is I-identifiable iff $E$ is a subset of a finite union of the sets in $\{\mathcal{X}_n\}$.*

**Definition 3** *An identification type $I$ is* powerful *if there is a set $U$ of recursive $\{0,1\}$ valued functions such that*

1. *$U$ is I-identifiable,*
2. *$U$ is not a set of effective measure 0, and*
3. *$U$ is not an effectively meager set.*

Now we come to the triality theorem. The proof of this result makes use of the following theorem due to Lisagor:

**Theorem 6** [Lis81] *$\mathcal{R}$ can be divided into two disjoint sets $A'$ and $B'$ such that*

1. *$A'$ is a set of effective measure 0, and*
2. *$B'$ is an effectively meager set.*

**Theorem 7** *Let $I$ be a smooth and powerful identification type. Then there are three one-to-one mappings $\phi$, $\psi$ and $\eta$ taking $\mathcal{R}$ to $\mathcal{R}$ such that for an arbitrary $f \in \mathcal{R}$ and and arbitrary $E \subset \mathcal{R}$:*

1. *$\eta(\psi(\phi(f))) = \psi(\phi(\eta(f))) = \phi(\eta(\psi(f))) = f$,*
2. *$E$ is I-identifiable iff $\phi(E)$ is effectively meager,*
3. *$E$ is effectively meager iff $\psi(E)$ is of effective measure 0, and*
4. *$E$ is of effective measure 0 iff $\eta(E)$ is I-identifiable.*

Proof: Since $I$ is a powerful identification type, there is a set $U \subset \mathcal{R}$ such that $U$ is $I$-identifiable but $U$ is not of partial recursive effective measure 0 and $U$ is not effectively meager. Choose $A'$ and $B'$ from Theorem 6. Let $A$ denote $A' - U$ and $B$ denote $B' - U$. As subsets, $A$ is a set of effective measure 0 and $B$ is a set of the first category. Hence, we have divided $\mathcal{R}$ into three disjoint subsets $A$, $U$ and $B$ situated as in the following table.

| | $A$ | $U$ | $B$ |
|---|---|---|---|
| measure | 0 | 1 | |
| $I$-indentifiability | no | yes | no |
| category | second | | first |

Now we define the (non effective) mappings $\phi$, $\psi$ and $\eta$ such that

1. if $f \in A$, then $\phi(f) \in U$, $\psi(f) \in U$ and $\eta(f) \in U$, and

2. if $f \in U$, then $\phi(f) \in B$, $\psi(f) \in B$ and $\eta(f) \in B$, and

3. if $f \in B$, then $\phi(f) \in A$, $\psi(f) \in A$ and $\eta(f) \in A$.

To do this, we first define, in parallel, $\phi(f)$ for $f \in U$, $\psi(f)$ for $f \in B$ and $\eta(f)$ for $f \in A$. This is the easiest of what will be three cases. Merely list (in a non effective way) the three infinite countable sets of functions $U = \{f_0, f_1, \cdots\}$, $B = \{g_0, g_1, \cdots\}$, $A = \{h_0, h_1, \cdots\}$ and define $\phi(f_i) = g_i$, $\psi(g_i) = h_i$ and $\eta(h_i) = f_i$.

Next, we define in parallel $\phi(f)$ for $f \in B$, $\psi(f)$ for $f \in A$ and $\eta(f)$ for $f \in U$. It follows from Corollary 5 (taking $\mathcal{X} = B$) that $B$ can be divided into an infinite sequence of disjoint infinite sets $\{B_n\}$ such that an arbitrary subset $E$ of $B$ is $I$-identifiable iff $E$ is a subset of a finite union of the sets $\{B_n\}$. It follows from Corollary 4 (taking $\mathcal{X} = A$) that $A$ can be divided into an infinite sequence of disjoint infinite sets $\{A_n\}$ such that an arbitrary subset $E$ of $A$ is effectively meager iff $E$ is a subset of a finite union of the sets $\{A_n\}$. It follows from Corollary 3 (taking $\mathcal{X} = U$) that $U$ can be divided into an infinite sequence of disjoint infinite sets $\{U_n\}$ such that an arbitrary subset $E$ of $U$ is of effective measure 0 iff $E$ is a subset of a finite union of the sets $\{U_n\}$.

In order to define the needed values of $\phi(f)$, $\psi(f)$ and $\eta(f)$, we list (non effectively), for an arbitrary $i \in \mathbb{N}$, $B_i = \{g_0^i, g_1^i, g_2^i, \cdots\}$, $A_i = \{h_0^i, h_1^i, h_2^i, \cdots\}$, $U_i = \{f_0^i, f_1^i, f_2^i, \cdots\}$, and define $\phi(g_j^i) = h_j^i$, $\psi(h_j^i) = f_j^i$ and $\eta(f_j^i) = g_j^i$.

Finally, we define, in parallel, $\phi(f)$ for $f \in A$, $\psi(f)$ for $f \in U$ and $\eta(f)$ for $f \in B$. It follows from Corollary 5 (taking $\mathcal{X} = A$) that $A$ can be divided into an infinite sequence of disjoint infinite sets $\{A'_n\}$ such that an arbitrary subset $E$ of $A$ is $I$-identifiable iff $E$ is a subset of a finite union of the sets $\{A'_n\}$. It follows from Corollary 4 (taking $\mathcal{X} = U$) that $U$ can be divided into an infinite sequence of disjoint infinite sets $\{U'_n\}$ such that an arbitrary subset $E$ of $A$ is effectively meager iff $E$ is a subset of a finite union of the sets $\{U'_n\}$. It follows from Corollary 3 (taking $\mathcal{X} = B$) that $B$ can be divided into an infinite sequence of disjoint infinite sets $\{B'_n\}$ such that an arbitrary subset $E$ of $U$ is of effective measure 0 iff $E$ is a subset of a finite union of the sets $\{B'_n\}$.

In order to define the needed values of $\phi(f)$, $\psi(f)$ and $\eta(f)$, we list (non effectively), for an arbitrary $i \in \mathbb{N}$, $A_i = \{\bar{h}_0^i, \bar{h}_1^i, \bar{h}_2^i, \cdots\}$, $U_i = \{\bar{f}_0^i, \bar{f}_1^i, \bar{f}_2^i, \cdots\}$, $B_i = \{\bar{g}_0^i, \bar{g}_1^i, \bar{g}_2^i, \cdots\}$, and define $\phi(\bar{h}_j^i) = \bar{f}_k^i$, $\psi(\bar{f}_j^i) = \bar{g}_j^i$, and $\eta(\bar{g}_j^i) = \bar{h}_j^i$.

As a direct consequence of the construction, assertion 1.) of the theorem holds. Now we prove assertion 2.) of the theorem. Suppose $E$ is $I$-identifiable. Represent $E = E_A \cup E_U \cup E_B$ where $E_A = E \cap A$, $E_U = E \cap U$ and $E_B = E \cap B$. Since $E_U \subseteq U$, $\phi(E_U) \subseteq B$, and $B$ is effectively meager. Hence, $\phi(E_U)$ is effectively meager as well.

$E_A$ is $I$-identifiable as it is a subset of the $I$-identifiable set $E$. Hence, it follows from Corollary 5 (taking $\mathcal{X} = A$) that $E_A$ is contained in a finite union of the sets $A'_n$. Choose $m$ such that $E_A \subseteq \bigcup_{n=0}^m A'_n$. By the definition of $\phi$, $\phi(E_A) \subseteq \bigcup_{n=0}^m U'_n$, and $\phi(E_A)$ is effectively meager by Corollary 4 (with $\mathcal{X} = U$).

$E_B$ is also $I$-identifiable as it is a subset of the $I$-identifiable set $E$. Hence, it follows from Corollary 5 (taking $\mathcal{X} = B$) that $E_B$ is contained in a finite union of the sets $B_n$. Choose $s$ such that $E_B \subseteq \bigcup_{n=0}^s B_n$. Then, by the definition of $\phi$, $\phi(E_B) \subseteq \bigcup_{n=0}^s A_s$ and $\phi(E_B)$ is effectively meager by Corollary 4 (with $\mathcal{X} = A$).

It follows from the fact that $E = E_A \cup E_U \cup E_B$ that $\phi(E) = \phi(E_A) \cup \phi(E_U) \cup \phi(E_B)$, i.e. $\phi(E)$ is effectively meager as the union of three effectively meager sets.

Assertions 3.) and 4.) are proved similarly. ☒

As a consequence of Theorem 7 we have the following "Triality Principle."

**Theorem 8** *Let $P$ be an arbitrary assertion using only the notions of pure set theory and the notion "set of effective measure 0." Let $P^\star$ be the assertion obtained by substituting "I-identifiable set" for 'set of effective measure 0" everywhere in $P$. Let $P^{\star\star}$ be the assertion obtained by substituting "effectively meager set" for 'set of effective measure 0" everywhere in $P$. Then the assertions $P$, $P^\star$ and $P^{\star\star}$ are either all true or all false.*

Proof: (sketch) Theorem 8 is easily implied by Theorem 7 in spite of the algorithmic nature of the notions of measure, category and identifiability on the one hand and the non effectiveness of the mappings $\phi$, $\psi$ and $\eta$ on the other hand. The crucial point in these proofs is the presentation of all the considered sets as finite unions from the "small" sets $\{A_N\}$, $\{B_n\}$, $\{U_n\}$, $\{A'_N\}$, $\{B'_n\}$ and $\{U'_n\}$. ☒

## 4 Suitable Identification Types

A large variety of identification types have been studied in inductive inference [OSW86]. However, only a few of them are smooth. The earliest defined identification type that turned out to be smooth is *reliable* inference defined in [BB75] and investigated in [Min76]. Unfortunately, this type is not powerful. Hence, Theorem 7 does not apply.

**Theorem 9** *Let $U \subset \mathcal{R}$ such that $U$ can be reliably identified. Then $U$ is an effectively meager set with partial recursive effective measure 0.*

An identification type that is suitable is a slight modification to the recently introduced *learning with confidence* [BFS96]. A *Confidence inductive inference machine* (CIIM) works like an IIM, except that it outputs a sequence of hypothesis with belief levels (rational numbers between 0 and 1): $s = p_0/c_0$, $p_1/c_1$, $p_2/c_2$, $\cdots$. We say that a CIIM $M$ learns a recursive function $f$ if on input from the graph of $f$ it outputs a sequence $p_0/c_0$, $p_1/c_1$, $p_2/c_2$, $\cdots$ such that for any $i$, the belief levels associated with $p_i$ are monotone nondecreasing and there exists a unique $p$ with the belief level for $p$ approaching 1 and $\varphi_p = f$. In this way, each CIIM learns a set of recursive functions. The collection of all such sets is denoted by CEX.

The type $CEX$ has most of the needed features, but lacks something essential, namely, $\mathcal{R}$ is $CEX$-identifiable[BFS96]. However one cannot CEX-identify the minimal programs for all the set $\mathcal{R}$.

**Theorem 10** [BFS96] *Suppose $\varphi_0$, $\varphi_1$, $\cdots$ is an acceptable programming system. $\mathcal{R}$ is not $CEX$-identifiable if we demand that minimal indices in $\varphi_0$, $\varphi_1$, $\cdots$ are to be found.*

Since we need smoothness to hold for an identification type, we consider the identification type $RMCEX_\varphi$: "Learning with confidence of the minimal indices in the acceptable programming system $\varphi_0$, $\varphi_1$, $\cdots$ with that additional restriction that if the target function is not from the identifiable set, then no hypothesis gets assigned a confidence level growing to 1." This identification type incorporates all the features of reliable identification and much more. It should be noted that this type depends on the particular acceptable programming system used.

**Theorem 11** *For an arbitrary acceptable programming system $\varphi_0$, $\varphi_1$, $\cdots$, the identification type $RMCEX_\varphi$ is smooth and powerful.*

Proof: Use the set from Theorem 4 of [FFG+95]. Details omitted.

Finally, we mention that $RMCEX_\varphi$ is not a *reasonable* identification type in terms of the definition in [FFG+95]. Theorem 17 of [FFG+95] says that if $U$ is $I$-identifiable with $I$ being a reasonable identification type, then $\mathcal{R} - U$ is both not effectively meager and not of effective measure 0. If one could construct a $RMCEX_\varphi$-identifiable set of functions whose complement would not be huge, it might lead to a strengthening of our Triality Principle for this identification type.

There are four other smooth and powerful identification types. Notice that they all involve taking the union over some parameter that typically gives rise to a hierarchy.

1. $EX$-identification by teams of an arbitrary finite size [Smi82, Smi94b],
2. Probabilistic $EX$-identification with and arbitrary probability [Pit89],
3. $EX$-identification with procrastination in the number of mindchanges [FS93],
4. $EX$-identification with procrastination in the number of anomalies [FS93].

# 5 Applications

In this section we consider several applications of Theorem 8. N. Lusin proved the following theorem about classical, continuous, set theory in 1914 (see [Sie34a]).

**Theorem 12** [N. Lusin, 1914] *Assuming the Continuum Hypothesis, an arbitrary set of reals $E$ of the second category has a subset $M$ of cardinality $2^{\aleph_0}$ such that every non countable subset of $M$ is a set of the second category.*

A counterpart of this result was subsequently shown for Lebesgue measure in 1934.

**Theorem 13** [W. Sierpinski [Sie34a]] *Assuming the Continuum Hypothesis, an arbitrary set of reals $E$ which is not a set of measure 0 has a subset $M$ of cardinality $2^{\aleph_0}$ such that every non countable subset of $M$ is a set of measure 0.*

We formulate and prove three theorems that are the counterparts of Theorems 12 and 13 using the set of $\mathcal{R}$ of recursive $\{0,1\}$ valued functions instead of the real numbers. These results appear to be new, despite the recent interest effective category and measure theory. Our triality principle (Theorem 8) allows us to essentially prove all three at the same time. This reveals that the following three theorems are the same from some higher viewpoint.

**Theorem 14** *Let $I$ be a smooth identification type. An arbitrary $E \subset \mathcal{R}$ which is not $I$-identifiable has an infinite subset $M$ such that every infinite subset of $M$ is not $I$-identifiable.*

**Theorem 15** *An arbitrary $E \subset \mathcal{R}$ which is not a set of effective measure 0 has in infinite subset $M$ such that every infinite subset of $M$ is not a set of effective measure 0.*

**Theorem 16** *An arbitrary $E \subset \mathcal{R}$ which is not effectively meager has an infinite subset $M$ such that every infinite subset of $M$ is not effectively meager.*

We end this section with a final trio of results obtained using our powerful new technique.

**Theorem 17** *Let $I$ be a smooth identification type. There is a one-to-one mapping $F$ from $\mathcal{R}$ to a subset of $\mathcal{R}$ such that $F(E)$ is $I$-identifiable iff $E$ is finite.*

**Theorem 18** *There is a one-to-one mapping $F$ from $\mathcal{R}$ to a subset of $\mathcal{R}$ such that $F(E)$ is a set of effective measure 0 iff $E$ is finite.*

**Theorem 19** *There is a one-to-one mapping $F$ from $\mathcal{R}$ to a subset of $\mathcal{R}$ such that $F(E)$ is effectively meager iff $E$ is finite.*

## 6 Extended Triality Principle

In the proof of theorem 7 the mappings $\phi$, $\psi$, $\eta$ are very similar. They map the sets $A$, $B$ and $U$ into the same $A$, $B$ and $U$. This suggests that stronger forms of Theorems 7 and 8 might be true. Consider the following:

**Assertion 20** *Let $I$ be a smooth and powerful identification type. Then there is a one-to-one mapping $F : \mathcal{R} \to \mathcal{R}$ such that for an arbitrary $f \in \mathcal{R}$ and for an arbitrary $E \subset \mathcal{R}$:*

1. $F(F(F(f))) = f$,
2. *$E$ is $I$-identifiable $\Leftrightarrow$ $F(E)$ is effectively meager,*
3. *$E$ is effectively meager $\Leftrightarrow$ $F(E)$ is of effective measure 0, and*
4. *$E$ is of effective measure 0 $\Leftrightarrow$ $F(E)$ is $I$-identifiable.*

Assertion 20 would imply the following stronger form of Theorem 8:

**Assertion 21** *Let $P$ be an arbitrary statement using only the notions of pure set theory and the notions "set of effective measure 0," "effectively meager set," and "$I$-identifiable set." Let $P^\star$ be the statement obtained from $P$ by simultaneously substituting "effectively meager set" for "set of effective measure 0," "$I$-identifiable set" for "effectively meager set," and "set of effective measure 0" for "$I$-identifiable set" everywhere in $P$. Let $P^{\star\star}$ be the statement obtained from $P$ by simultaneously substituting "$I$-identifiable set" for "set of effective measure 0," "effectively meager set" for "$I$-identifiable set," and "set of effective measure 0" for "effectively meager set" everywhere in $P$. Then $P$, $P^\star$ and $P^{\star\star}$ are either all true or all false.*

Unfortunately, Assertion 21 (and hence Assertion 20 as well) is false for at least some smooth and powerful identification types $I$. Indeed, the identification type "$EX$-identification by teams of an arbitrary finite size" is smooth and powerful [Smi82]. However, for any set $U \subset \mathcal{R}$ that is identifiable by this inference type, the set $(\mathcal{R} - U)$ cannot be effectively meager and it cannot be of effective measure 0. Applying Assertion 21 to Theorem 6 yields the following false statement:

$\mathcal{R}$ can be divided into two disjoint sets $A'$ and $B'$ such that

1. $A'$ is an effectively meager set, and
2. $B'$ is learnable by a finite $EX$ team.

Hence, Assertions 20 and 21 are false at least for this one identification type. However, the following extension does hold.

**Theorem 22** *Let $P$ be an arbitrary statement using only the notions of pure set theory and the notions "sets of effective measure 0," "effectively meager sets," and "$I$-identifiable sets." Let $P^\star$ be the statement obtained from $P$ by interchanging the notions "sets of effective measure 0" and "effectively meager sets." Then the statements $P$ and $P^\star$ are either both true of both false.*

Proof: (sketch) The proof is more complicated than merely referring to Theorem 6. Again a mapping $F : \mathcal{R} \to \mathcal{R}$ is constructed such that:

1. $F(F(f)) = f$,
2. $E$ is effectively meager $\Leftrightarrow$ $F(E)$ is of effective measure 0, and
3. $E$ is of effective measure 0 $\Leftrightarrow$ $F(E)$ is effectively meager.

   However, this time we also demand:
4. $E$ is $I$-identifiable $\Leftrightarrow$ $F(E)$ is $I$-identifiable.

$\boxtimes$

Theorem 22 and the failure of Assertion 21 imply that not all combinations of the notions measure, category and identifiability are equivalent in spite of the fact that these notions are pairwise dual in the sense of Theorem 8. Consider $EX$ identification by a team of arbitrary finite size. Then, Theorem 8 applies, but the following assertion fails.

**Assertion 23** *(False) Let $P$ be an arbitrary statement using only the notions of pure set theory and the notions "set of effective measure 0," "effectively meager set," and "I-identifiable set." Let $P^\star$ be the statement obtained from $P$ by interchanging the notions "I-identifiable set" and "effectively meager set." Then the statements $P$ and $P^\star$ are either both true or both false.*

The failure of Assertions 21 and 23 does not exclude a possibility that for some identification types Assertion 21 may hold. For instance, we conjecture that Assertion 21 may hold for $RMCEX_\varphi$-identification.

# 7 Directions for Further Research

Because of the failure of Assertion 21, there is no hope to have complete interchangeability of the notions of measure, category and identifiability. However, considering such interchanges is a good source of open problems. Indeed, several prior studies could be viewed as following such a program For example, Bennett and Gill [BG81] proved that $P^A \neq NP^A \neq coNP^A$ for nearly all recursive oracles where "nearly all" was understood as "all but a set of effective measure 0." However, Blum and Impagliazzo [BI87] showed that $P^A = NP^A = coNP^A$ for nearly all oracles where "nearly all" was understood as "all but a set of the first category." It was conjectured that "what is true for nearly all oracles, is true for the empty oracle as well." This conjecture has been disproved for both "all oracles but a set with effective measure 0" and "all oracles except a set of the first category" [Kur83, LFKN92, Fos93, CCG+94]. This conjecture might still be true for a more appropriate notion of "nearly all sets".

Since then, the notions of measure and category have appeared frequently, not only in their effective versions, but also in resource-bounded versions as well. Our Theorem 8 offers a third notion dual to measure and category. It is quite possible that a third approach to the intuitive notion of "nearly all sets" can be useful. In Lebesgue measure, typical reals are called "random," in Baire category, they are called "generic." Generic sets, like random sets, have proved useful in providing a coherent picture of relativized computation. They embody the method of diagonalization in oracle constructions, i.e. requirements which can always be satisfied by a finite extension of the oracle.

Lutz [Lut90] has introduced resource bounded versions of both measure and category. Fenner [Fen91] has studied resource bounded generics. It seems promising to consider the identifiability counterparts of randomness and generics.

## 8 Acknowledgements

Conversations with Sebastiaan Terwijn were most helpful in sorting out the various notions of effective measure.

## References

[AFS96] A. Ambainis, R. Freivalds, and C. Smith. General inductive inference types based on linearly ordered sets. In C. Puech and R. Reischuk, editors, *Proceedings of the* $13^{\text{th}}$ *Symposium on the Theoretical Aspects of Computer Science*, volume LNCS1046, pages 243–253. Springer, 1996.

[AS83] D. Angluin and C. H. Smith. Inductive inference: Theory and methods. *Computing Surveys*, 15:237–269, 1983.

[AS95] K. Ambos-Spies. Resource-bounded genericity. In *Proceedings of the* $10^{\text{th}}$ *Conference on Structure in Complexity Theory*, pages 162–181. IEEE Computer Society, 1995.

[ASMWZ96] K. Ambos-Spies, E. Mayordomo, Y. Wang, and X. Zheng. Resource-bounded genericity, stochasticity and weak randomness. In C. Puech and R. Reischuk, editors, *Proceedings of the* $13^{\text{th}}$ *Symposium on the Theoretical Aspects of Computer Science*, volume LNCS1046, pages 63–74. Springer, 1996.

[BB75] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

[BFS96] J. Bārzdiņš, R. Freivalds, and C. Smith. Learning with confidence. In *Proceedings of STACS'96*, 1996.

[BG81] C.G. Bennet and J. Gill. Relative to a random oracle $A$, $P^a \neq NP^a \neq co - NP^a$ with probability 1. *SIAM Journal of Computer Science*, 10:96–113, 1981.

[BI87] M. Blum and R. Impagliazzo. Generic oracles and oracle classes. In *Proceedings of the* $28^{\text{th}}$ *Symposium on the Foundations of Computer Science*, pages 118–126. IEEE Computer Soceity, 1987.

[BL96] H. Buhrman and L. Longpré. Compressibility and resource bounded measure. In C. Puech and R. Reischuk, editors, *Proceedings of the* $13^{\text{th}}$ *Symposium on the Theoretical Aspects of Computer Science*, volume LNCS1046, pages 13–24. Springer, 1996.

[BM95a] J. Balcázar and E. Mayordomo. A note on genericity and bi-immunity. In *Proceedings of the* $10^{\text{th}}$ *Conference on Structure in Complexity Theory*, pages 193–196. IEEE Computer Society, 1995.

[BM95b] H. Buhrman and E. Mayordomo. An excursion to the Kolmogorov random strings. In *Proceedings of the* $10^{\text{th}}$ *Conference on Structure in Complexity Theory*, pages 197–205. IEEE Computer Society, 1995.

[Bor05] E. Borel. *Leçons sur les fonctions de variables réeles.* Gauthier-Villars, Paris, 1905.

[Bor14] E. Borel. *Leçons sur la théorie des fonctions.* Gauthier-Villars, Paris, 1914.

[CCG$^+$94] Richard Chang, Benny Chor, Oded Goldreich, Juris Hartmanis, Johan Håstad, Desh Ranjan, and Pankaj Rohatgi. The random oracle hypothesis is false. *Journal of Computer and System Sciences*, 49(1):24–39, August 1994.

[Erd43] P. Erdös. Some remarks on set theory. *Ann. Math*, 44(2):643–646, 1943.

[Fen91] S. Fenner. Notions of resource-bounded category and genericity. In *Proceedings of the* $6^{\text{th}}$ *Conference on Structure in Complexity Theory*, pages 196–212. IEEE Computer Soceity, 1991.

[Fen95] S. Fenner. Resource-bounded Baire category: A stronger approach. In *Proceedings of the* $10^{\text{th}}$ *Conference on Structure in Complexity Theory*, pages 182–192. IEEE Computer Society, 1995.

[FFG$^+$] L. Fortnow, R. Freivalds, W. Gasarch, M. Kummer, S. Kurtz, C. Smith, and F. Stephan. On the relative sizes of learnable sets. *Theoretical Computer Science.*

[FFG$^+$95] L. Fortnow, R. Freivalds, W. Gasarch, M. Kummer, S. Kurtz, C. Smith, and F. Stephan. Measure, category and learning theory. In *Proceedings of ICALP'95, LNCS Vol. 944*, pages 558–569, 1995.

[Fos93] J. Foster. The generic oracle hypothesis is false. *Information Processing Letters*, 45(2):59–62, 1993.

[Fre91] R. Freivalds. Inductive inference of recursive functions: Qualitative theory. In J. Bārzdiņš and D. Bjørner, editors, *Baltic Computer Science*, pages 77–110. Springer Verlag, 1991. Lecture Notes in Computer Science, Vol. 502.

[FS93] R. Freivalds and C. Smith. On the power of procrastination for machine learning. *Information and Computation*, 107:237–271, 1993.

[Gol67] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[Kum96] M. Kummer. On the complexity of random strings. In C. Puech and R. Reischuk, editors, *Proceedings of the* $13^{\text{th}}$ *Symposium on the Theoretical Aspects of Computer Science*, volume LNCS1046, pages 25–38. Springer, 1996.

[Kur58] C. Kuratowski. *Topologie,* $4^{\text{th}}$ *edition*, volume 20-21 of *Monografie Matematyczne.* Panstwowe Wydawnictwo Naukowe, 1958.

[Kur83] Stuart A. Kurtz. On the random oracle hypothesis. *Information and Control*, 57(1):40–47, 1983.

[LFKN92] C. Lund, L. Fortnow, H. Karloff, and N. Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4), 1992.

[Lis81] L. Lisagor. The Banach-Mazur game. Translated Version of *Matematicheskij Sbornik*, 38:201–206, 1981.

[Lut90] J. Lutz. Category and measure in complexity classes. *SIAM Journal of Computing*, 19:1100–1131, 1990.

[Lut93] J. Lutz. The quantitative structure of exponential time. In *Proceedings of the* $8^{\text{th}}$ *Structure in Complexity Theory Conference*, pages 158–175. IEEE Computer Society, 1993.

[Lut96] J. Lutz. Observations on measures and lowness for $\Delta_2^P$. In C. Puech and R. Reischuk, editors, *Proceedings of the* $13^{\text{th}}$ *Symposium on the Theoretical Aspects of Computer Science*, volume LNCS1046, pages 87–98. Springer, 1996.

[Meh73] K. Mehlhorn. On the size of sets of computable functions. In *Proceedings of the* $14^{\text{th}}$ *Symposium on Switching and Automata Theory*, pages 190–196. IEEE Computer Soceity, 1973.

[Min76] E. Minicozzi. Some natural properties of strong-identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.

[MY78] M. Machtey and P. Young. *An Introduction to the General Theory of Algorithms*. North-Holland, New York, 1978.

[OSW86] D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn*. MIT Press, Cambridge, Mass., 1986.

[Oxt71] J. Oxtoby. *Measure and Category*. Springer-Verlag, 1971.

[Pit89] L. Pitt. Probabilistic inductive inference. *Journal of the ACM*, 36(2):383–433, 1989.

[Put75] H. Putnam. Probability and confirmation. In *Mathematics, Matter and Method*, volume 1. Cambridge University Press, 1975.

[Sie34a] W. Sierpinski. *Hypothèse du continu*, volume 4. Monografie Matematyczne, 1934.

[Sie34b] W. Sierpinski. Sur la dualité entre la première catégore et la mesure nulle. *Fund. Math.*, 28:276–280, 1934.

[Smi82] C. H. Smith. The power of pluralism for automatic program synthesis. *Journal of the ACM*, 29(4):1144–1165, 1982.

[Smi94a] C. Smith. *A Recursive Introduction to the Theory of Computation*. Springer-Verlag, 1994.

[Smi94b] C. Smith. Three decades of team learning. In *Proceedings of AII/ALT'94, Lecture Notes in Artificial Intelligence*. Springer Verlag, 1994.