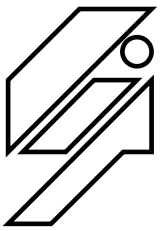


Laboratoire de l'Informatique du Parallélisme

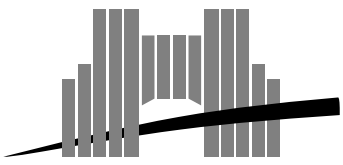
Ecole Normale Supérieure de Lyon
Unité de recherche associée au CNRS n°1398

Do most strong definitions of randomness exist?

Bruno Durand
Vladimir Kanovei
Vladimir A. Uspensky
Nikolai Vereshagin

April 1998

Research Report N^{o} 98-22



Ecole Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : (+33) (0)4.72.72.80.00 Télécopieur : (+33) (0)4.72.72.80.80
Adresse électronique : lip@lip.ens−lyon.fr

Do most strong definitions of randomness exist?

Bruno Durand
Vladimir Kanovei
Vladimir A. Uspensky
Nikolai Vereshagin

April 1998

Abstract

The goal of our paper is to propose a way to obtain more refined definitions of randomness than the notions known so far (e.g. Martin-Löf randomness). We show that a "perfect" definition of randomness based on provability does not exist. We then weaken our requirements on the definition by replacing provability by consistency and obtain a formula that defines a set of random sequences that fulfills rather strong conditions.

Keywords: Randomness, logics, Solovay model

Résumé

Nous proposons ici de raffiner les définitions classiques du caractère aléatoire des suites infinies (en particulier la très classique définition de Martin-Löf). Nous prouvons qu'il n'existe pas de définition parfaite fondée sur la notion de prouvabilité. En remplaçant la prouvabilité par la consistence, nous obtenons une définition des suites aléatoires très générale qui remplit des conditions raisonablement fortes.

Mots-clés: Suites aléatoire, logique, modèle de Solovay

Do most strong definitions of randomness exist?

Bruno Durand [*] Vladimir Kanovei [†]
Vladimir A. Uspensky [‡§] Nikolai Vereshagin[‡¶]

April 21, 1998

Abstract

The goal of our paper is to propose a way to obtain more refined definitions of randomness than the notions known so far (e.g. Martin-Löf randomness). We show that a "perfect" definition of randomness based on provability does not exist. We then weaken our requirements on the definition by replacing provability by consistency and obtain a formula that defines a set of random sequences that fulfills rather strong conditions.

Introduction

If somebody tells us that he have tossed a coin sixty times getting a string u_1

$$101010110010011010010010101000101100011010000101011110100101$$

(0 stands for head, 1 for tail) we are not surprised. However, the string u_2

$$010101010101010101010101010101010101010101010101010101010101$$

looks suspicious and we are ready to reject the assumption that the coin is fair. Why? Is not the probability of both sequences the same— 2^{-60} ?

There are four explanations why the former sequence, u_1, looks more random than the latter, u_2:

[*] L.I.P., Ecole Normale Supérieure de Lyon - CNRS, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. E-mail: Bruno.Durand@ens-lyon.fr.

[†] Moscow Transport Engineering Institute, kanovei@mech.math.msu.su. Supported by DFG and University of Wuppertal.

[‡] Dept. of Mathematical Logic and Theory of Algorithms, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Vorobjyovih Gorih, Moscow 119899, Russia.

[§] Email: uspensky@lpcs.math.msu.ru

[¶] Work done at L.I.P., Ecole Normale Supérieure de Lyon. E-mail: ver@mech.math.msu.su.

a) u_2 belongs to a simply described set of small measure, namely, to the set of strings with alternating digits having the measure 2^{-59}. In contrast, we do not see any simply described set of small measure containing u_1;

b) u_2 has much regularities, it may be described very easily as "01 thirty times", in contrast the first string seems to have no shorter descriptions than the displayed one;

c) u_2 is predictable: if you give its 30 first bits to a person he will surely predict the rest;

d) the subsequence of u_2 consisting of all odd terms has much more zeros than ones (actually no ones at all), and we expect that for any rule of choice of a subsequence (which does not use the information of the value of the chosen term) in the resulted subsequence, the frequency of zeros is about one half.

It is pretty clear that it is impossible to divide finite strings in random and non-random. One may hope only to measure the amount of randomness, which should reflect our belief in that the sequence was obtained by fair coin tossing.

The argument b) can be formalized by means of the Kolmogorov complexity $\mathrm{K}(u)$ defined, for a finite binary sequence u, as the bit size of minimal program that prints u. The less is $l(u) - \mathrm{K}(u)$ ($l(\cdot)$ stands for the length of strings) the more random looks the string.

The reason a) is more or less equivalent to b). Indeed, if $\mathrm{K}(u)$ is small then u belongs to the set $\{u\}$ of small measure having small Kolmogorov entropy (the Kolmogorov entropy of a set is the bit length of the shortest program printing, in some order, the list of elements in the set). Conversely, if u belongs to a small set A having small Kolmogorov entropy, then u can be identified by the pair (program printing the list of elements of A, the number of u in that list), which has short size since both its components are short.

Reasons c) and d) can be also reduced to a). Thus, in the case of finite binary strings, we have a quite adequate definition of the amount of non-randomness in a string. This is the value $l(u) - \mathrm{K}(u)$.

The things seem to become easier, in a sense, when we turn to the case of infinite binary strings (=binary sequences). One may hope to divide them into random and non-random. That is, to give a rigorous definition of a sequence which may be an outcome of infinite series of coin tosses. For instance, everyone will agree that the infinite coin tossing may not result in the series

$$0101010101010101010101010101010101.....$$

The aim of the definition of randomness is to clear our intuition in this respect.

There are four known approaches (according to four above explanations) to define randomness of infinite sequences. Let us sketch these four approaches (the detailed survey may be found in [7]; and three of the four approaches are exposed in [8] and [9]). We denote by Ω the set of all infinite binary sequences x.

The first approach corresponds to a). One fixes a countable class L of subsets of Ω of measure 1 and then defines a sequence to be random if it belongs to all sets in L. (In this paper, the measure means the Lebesgue, or uniform, measure in Ω, it is denoted by $\mathtt{mes}$.) The set R of random strings has, of course, measure one. The larger class L we take the more refined notion of randomness we obtain and the stronger is our belief that any random sequence may be obtained by fair coin tossing. Equivalently, one can choose a class S of subsets of Ω of measure zero and define a sequence to be random if it avoids (does not belong to) all the sets in S. The common name of obtained notions of randomness is *typicalness*.

The most famous definition of typicalness belongs to Martin-Löf [6] and is as follows. Let Ω_u denote the set of all infinite continuations of a finite string u. Recall the the set $A \subseteq \Omega$ has measure 0 (=is a null set) if for any n there exists a set B_n of finite strings such that 1) $A \subseteq \bigcup_{u \in B_n} \Omega_u$ and 2) $\sum_{u \in B_n} \mathtt{mes}(\Omega_u) = \sum_{u \in B_n} 2^{-l(u)} < 1/n$. In other words, A can be covered by an open set (in Cantor's topology) of arbitrarily small measure. Martin-Löf considers the constructive version of this definition: a set A is called a *effectively null set* if there exists a sequence B_n satisfying 1) and 2) such that the set $\{\langle u, n\rangle : u \in B_n\}$ is enumerable. This means that there exists an algorithm printing all the elements of this set in some order (the order does not matter). As the set may be infinite, the process of printing may last infinitely long.

According to Martin-Löf, the sequence is called random (we will also use the term "typical") if it avoids all effectively null sets. Thus as S one takes the class of sets of the form $\bigcap_n \bigcup_{u \in B_n} \Omega_u$, where $\sum_{u \in B_n} 2^{-l(u)} < 1/n$ and the set $\{\langle u, n\rangle : u \in B_n\}$ is enumerable. The family S is countable, as any its element is identified by an algorithm and the number of algorithms is countable.

The complements of effectively null sets are called *effectively full sets*.

It turned out that every law of probability theory among the laws studied so far includes an effectively full set, and hence is satisfied by any Martin-Löf random sequence. We call a law of probability theory (LPT) an assertion Ψ about an infinite binary sequence such that the set $\{x : \Psi(x)\}$ has measure one. The examples of LPTs are the law of large numbers (the frequency of zeros among first n bits tends to 1/2) or the law of the iterated logarithm. Thus, for any such particular Ψ studied by probability theory there exists an effectively full set included in $\{x : \Psi(x)\}$. For one of them, namely for ergodic

theorem, this has been unknown for a decade. In other words, Martin-Löf random sequences satisfy all known LPTs. Yet one cannot be sure that this will be so for ever: we can by now construct *ad hoc* LPTs that are not satisfied by Martin-Löf random sequences.

The second, the third and the forth approaches correspond respectively to arguments b), c), d) above. The obtained notions of randomness are called *chaoticness*, *unpredictability* and *stochasticness*. We will not present the definitions of these notions, the interested reader is referred to [7]. Let us just mention that chaoticness is equivalent to typicalness, and both imply unpredictability; unpredictability implies stochasticness, which is weaker than unpredictability. The notions of *chaoticness* and *stochasticness* are also presented in [8] and in [9] (the reader should be warned that in [9] instead of the terms *chaoticness* and *stochasticness* the terms *Kolmogorov randomness* and *Mises–Kolmogorov–Loveland randomness* are used respectively).

The goal of our paper is to propose a way to obtain more refined definitions of typicalness than the notions known so far. Why we think that the existent notions like the Martin-löf's one is not good enough? That is because one can define (in a quite simple way) a particular sequence, which is Martin-Löf random. For instance, the binary representation of the so called Chaitin's number of wisdom [1]; this real number is the probability of a programm to halt when a programming system is fixed and the set of programs is endowed by some standard probability distribution. Or, one can define a particular Martin-Löf random sequence by means of simple diagonal definition. For our intuition it seems slightly uncomfortable to accept a definable sequence as random.

Another argument agains the notions of randomness known so far is that they use the theory of algorithms. The notion of an algorithm cannot be expressed, as far as we know, in terms of set theory (see [9]). Thus it is not quite natural to see that it interfers with the notion of randomness. We would prefer a definition expressed in a logical framework.

The perfect notion of a random sequence in the framework of typicalness would be a notion satisfying two principles.

- Almost all sequences are random. That is, the set of random sequences has measure 1.
- Any random sequence satisfy any mental law of probability theory.

Let us formulate both principles in the rigorous form. By the notion of randomness we mean a formula $\rho(x)$ in a set theoretical language (that of Zermelo–Fraencel system, **ZFC**). The precise form of the principles is:

(1) $\mathbf{ZFC} \vdash \mathtt{mes}\{x : \rho(x)\} = 1$. That is, it is provable, in Zermelo-Fraencel system, that almost all sequences are random.

(2) For any set theoretic formula $\Psi(x)$, if $\mathbf{ZFC} \vdash \mathtt{mes}\{x : \Psi(x)\} = 1$, then $\mathbf{ZFC} \vdash \forall x\ (\rho(x) \Rightarrow \Phi(x))$. That is, if it is provable, in **ZFC**, that the set $\{x \in \Omega : \Psi(x)\}$ has measure 1, then it is provable, in **ZFC**, that all random sequences satisfy $\Psi(x)$.

It is not hard to see that such a perfect notion of randomness does not exist (Theorem 1). Moreover, there is no notion of randomness satisfying (1) and the following weak form of principle (2):

(2′) For any particular (=definable) sequence $x \in \Omega$ it is provable, in **ZFC**, that x is not random. That is, for any formula $F(x)$ in the language of **ZFC** with the single parameter x such that $\mathbf{ZFC} \vdash (\exists\, !\, x \in \Omega)\ F(x)$ it holds $\mathbf{ZFC} \vdash (\forall x \in \Omega)\ (F(x) \Longrightarrow \neg\, \rho(x))$.

Principle (2′) follows from (2), as for any definable sequence $x_0 \in \Omega$ the assertion $x \neq x_0$ is a law of probability theory and therefore on can prove that any random sequence is different from x_0.

Thus we should moderate our requirements. Our proposal to this end, which seems to be a new one, is as follows. Consider the following weaker form of principle (2):

(2″) For any set theoretic formula $\Psi(x)$ such that it is provable in **ZFC** that the set $\{x \in \Omega : \Psi(x)\}$ has measure 1, it is **not** provable in **ZFC** that there **is** a random sequence satisfying $\Psi(x)$.

Informally, the principle states that no one will ever prove that a particular law of probability theory is not satisfied by some random sequence. In particular, any notion of random sequence satisfying (2″) is resistant to the above critics of Martin-Löf randomness.

This is, however, not all the requirements we find necessary to impose on a notion of randomness. The point is that the principles (1) and (2″) do not imply, that the sequence (say)

000000000000000.....

is not random. Principle (2″) implies, of course, that one cannot prove that it is random. But we expect that such laws as "not to be identically zero" should be proved. This leads us to the third principle:

(3) For any known law $\Psi(x)$ of probability theory it is provable, in **ZFC**, that any random sequence satisfies $\Psi(x)$. More specifically, it is provable, that any random sequence is Martin-Löf random.

Our main result is the notion of randomness denoted by $\rho(x)$ that satisfies (1), (2″) and (3) (Theorem 7). Principle (3) has of course a minor point: the choice of Martin-Löf randomness there is not motivated anyhow (the same minor point is in the Martin-Löf's definition: there is no solid basis to restrict all the LPTs to effectively full sets). However, our construction applies to any previously specified stock of LPTs: for any definable provably countable family of provably measure-one sets there exists a notion of randomness satisfying (1) and (2″), and and such that is is provable that any random sequence belongs to all those sets.

To present the idea let us come back to the Martin-Löf randomness. Recall that countable intersections of open sets are called $\mathbf{G}_\delta$ sets. Let us say that a sequence of sets B_n of finite binary sequences is a code for a $\mathbf{G}_\delta$ set $U \subseteq \Omega$ iff $U = \bigcap_n \bigcup_{u \in B_n} \Omega_u$. By the above definition, a sequence $x \in \Omega$ is Martin-Löf random iff it avoids any $\mathbf{G}_\delta$ set with enumerable code.

Our approach will be to increase the number of full $\mathbf{G}_\delta$ sets to avoid, at least including all those having *arithmetical* codes. This will result in the notion of randomness satisfying (1), (2″) and a much more stronger version of (3) than the above one.

The definition of ρ is as follows. Consider any class of sets A. Let us call a sequence x A-random if it avoids all null $\mathbf{G}_\delta$ sets having a code in A. If A is countable, then the set of A-random sequences has full measure. Let L be the set of all constructible sets (in Gödel's sense). $\rho(x)$ will say that $x \in \Omega$ is L-random whenever the set of all L-random sequences has full measure, and x is arithmetically random (*i. e.* A-random, where A is the class of all arithmetically definable objects, see below) otherwise. It is straightforward that ρ satisfies (1) and (3). Using the Solovay model, one may prove that it satisfies (2″).

The notion ρ satisfies also the common closure properties: it is stable with respect to finite changing, it is stable with respect to choosing a subsequence by means of an algorithm (the algorithm makes decisions which term to choose on the basis of the value of previously chosen terms).

1 "Provable" set theoretic randomness

In what follows, **sequence** will mean: an infinite binary sequence, that is an element of the set $\Omega = 2^\omega$.

One could have the intension to define a sequence r to be random in the case when it avoids any set $X \subseteq \Omega$, definable by a set theoretic formula $\Phi(x)$ such that **ZFC** proves that $\mathtt{mes}\{x \in \Omega : \Phi(x)\} = 0$, where $\mathtt{mes}$ is the Lebesgue measure. However this would not be a good approach.

Indeed in this case one would have got a mixture of mathematical and meta-

mathematical (provability) notions, that hardly can be adequately realized in a mathematically legitimate definition. To see this suppose, towards the contrary, that a set theoretic formula $\rho(x)$ adequately expresses the definition above. Then ρ would satisfy the requirements (1) and (2) above. However:

Theorem 1 *There does not exist any formula ρ satisfying both* (1) *and* (2).

Proof. Suppose that ρ is such a formula.

The argument is based on ideas connected with the Gödel constructibility. Gödel defined in 1938 a class L of sets called *constructible sets* and proved that L is a model of **ZFC**. The statement that all sets are constructible is called *the axiom of constructibility* and formally abbreviated by the equality V = L, where V denotes the universe of all sets. The axiom V = L was proved to be consistent with **ZFC** by Gödel (the key fact is that V = L is true in the class L) and independent from **ZFC** by Cohen in 1961. (We refer to [3, 4] in matters of all general set theoretic facts used below as well as in matters of the history of related set theoretic research.)

The most important here property of L is that there is a well-ordering $<_{\mathrm{L}}$ of L, definable by a concrete set theoretic formula.

Let now $\Psi(x)$ say the following: $x \in \Omega$ satisfies $\rho(x)$ but x is **not** the $<_{\mathrm{L}}$-least element of the set $\{x \in \Omega \cap \mathrm{L} : \rho(x)\}$. (The "but" reservation makes sense only when the intersection $\{x \in \Omega : \rho(x)\} \cap \mathrm{L}$ is non-empty.)

It follows from (1) that **ZFC** proves that $\{x : \Psi(x)\}$ is a set of full measure. Therefore, by the assumption of (2), **ZFC** must prove that $\rho(x)$ implies $\Psi(x)$. However the axiom V = L (which is consistent with **ZFC**) clearly implies that there is a sequence x satisfying $\rho(x)$ but not $\Psi(x)$ — namely, the $<_{\mathrm{L}}$-least element of the set $\{x \in \Omega : \rho(x)\}$, which is equal to $\{x \in \Omega \cap \mathrm{L} : \rho(x)\}$ in the assumption V = L. □

2 "Consistent" set theoretic randomness

Thus there is no formula satisfying both (1) and (2). This setback leads us to the idea to reduce our expectations. For instance one may be interested to find out whether there is a set theoretic formula $\rho(x)$ satisfying (1) and a weaker than (2) assumption, (2″). We will demonstrate that such a formula really exists — and that it is a derivative of an even more interesting formula, that of the Solovay randomness.

2.1 Solovay random sequences

Definition 2 A sequence $x \in \Omega$ is *Solovay random over* L iff it is L-random in the sense above, that is, it avoids any null $\mathbf{G}_\delta$ set with a code in L.

The formula saying that $x \in \Omega$ is Solovay random over L is denoted by $\rho_{\rm L}(x)$. Put $\mathcal{R}_{\rm L} = \{x \in \Omega : \rho_{\rm L}(x)\}$ (all Solovay random over L sequences). $\square$

In fact it will not be different to say: whenever $X \subseteq \Omega$ is a null *Borel* set with a code in L. [1] To see this note, first of all, that any null Borel set $X \subseteq \Omega$ can be covered by a null $\mathbf{G}_\delta$ set $U \subseteq \Omega$, which is a classical fact of measure theory. The construction of the covering set U can be maintained effectively enough to get the following refinement:

- any null Borel set coded in L can be covered by a null $\mathbf{G}_\delta$ set coded in L.

2.2 Solovay random sequences in different set universes

It occurs that basic properties of $\mathcal{R}_{\rm L}$ depend on the structure of the set universe. In other words, there is not much to say about $\mathcal{R}_{\rm L}$ in **ZFC**, but some special provisions can make $\mathcal{R}_{\rm L}$ to be a very useful set.

At the trivial side, $\mathcal{R}_{\rm L}$ is obviously empty if the axiom of constructibility V = L is assumed. Thus $\mathcal{R}_{\rm L}$ can be very small, even empty.

To make $\mathcal{R}_{\rm L}$ large, even a set of full measure, another consistent set theoretic hypothesis can be employed. Recall that $\aleph_1$ is the least uncountable cardinal, or, that is the same, the least cardinal bigger than $\aleph_0 = \mathtt{card}\,\mathbb{N}$, the countable cardinality.

By $\aleph_1^{\rm L}$ they denote "$\aleph_1$ in the sense of L", that is, something which is defined, in L, as the least uncountable cardinal. One easily sees that $\aleph_1^{\rm L}$ is, from the point of view of the whole set universe, an ordinal number (perhaps, not a cardinal [2]), which satisfies either $\aleph_1^{\rm L} < \aleph_1$ or $\aleph_1^{\rm L} = \aleph_1$.

The "or" case follows *e.g.* from V = L, and is not much of interest here.

The "either" case is also consistent with **ZFC**, but it needs to apply the method of *forcing* to get a suitable model. Models of **ZFC** which satisfy $\aleph_1^{\rm L} < \aleph_1$ belong to a wide class of *collapse generic models*: if the inequality holds, they say that $\aleph_1$ *collapses* (in the extension from L to the whole set universe V).

Lemma 3 *If* $\aleph_1^{\rm L} < \aleph_1$ *then* $\mathcal{R}_{\rm L}$ *is a* $\mathbf{G}_\delta$ *set of full measure.*

---

[1] It would be difficult to fully present here the involved mechanism of coding Borel subsets of Ω. It is based on the observation that construction of a Borel subset of Ω from sets of the form Ω_u, where u is a finite binary sequence (see Introduction) needs only countably many applications of the operations of countable union and countable intersection. This can be adequately coded *e.g.* by a sequence $c \in \Omega$. Sequences which code Borel sets this way are called *Borel codes*. The set of all Borel codes is a co-analytic subset of Ω.

[2] Cardinals are viewed as initial ordinals, *i.e.* those ordinal numbers κ which are not equinumerous to any $\alpha < \kappa$.

Proof. It suffices to prove that the set $\Omega \cap \mathrm{L}$ of all constructible sequences is countable in the assumption $\aleph_1^{\mathrm{L}} < \aleph_1$. To see this note that, in L, the continuum hypothesis $2^{\aleph_0} = \aleph_1$ holds, hence sequences in Ω can be put in $1-1$ correspondence with finite and countable ordinals. However finite and L-countable ordinals is the same as ordinals smaller than $\aleph_1^{\mathrm{L}}$, so that we have only countably many of them by the assumption $\aleph_1^{\mathrm{L}} < \aleph_1$. □

It occurs that $\aleph_1^{\mathrm{L}} < \aleph_1$ is not necessary for $\mathcal{R}_{\mathrm{L}}$ to be of full measure: in so-called amoeba generic models we have $\aleph_1^{\mathrm{L}} = \aleph_1$ but $\mathcal{R}_{\mathrm{L}}$ is of full measure.

2.3 Solovay random sequences in the Solovay model

The behaviour of the Solovay random sequences becomes especially interesting in *the Solovay model*, which is a kind of a collapse generic model of **ZFC**.

To obtain the Solovay model, one has to fix an inaccessible cardinal ϑ in the constructible universe L. Then one defines a generic extension of L, which is a model of **ZFC** where all L-cardinals $\kappa < \vartheta$, including $\aleph_1^{\mathrm{L}}$, but not the cardinal ϑ itself, become countable. The extension is the Solovay model. It has a lot of applications in set theory, for instance it is true in this model that all projective sets of sequences are Lebesgue measurable. This result is based on the following key fact (we refer to [3, 4] for proof):

Proposition 4 *The following is true in the Solovay model. If $X \subseteq \Omega$ is definable by a set theoretic formula containing only sets in* L *as parameters then there is a Borel set $B \subseteq \Omega$ with a code in* L *such that, for any Solovay random over* L *sequence x, we have $x \in X \iff x \in B$.* □

Note that $\aleph_1^{\mathrm{L}} < \aleph_1$ holds in the Solovay model by the construction. Therefore $\mathcal{R}_{\mathrm{L}}$ has full measure in the Solovay model by Lemma 3, so that, in the Solovay model, every set of sequences, definable by a formula with parameters in L, is a Borel set modulo a null set. It follows that every such a set of sequences is Lebesgue measurable in the Solovay model — and this remains true even if we allow, in addition, arbitrary parameters in Ω in definitions of sets.

Corollary 5 *The following is true in the Solovay model. If $X \subseteq \Omega$ is a set of full measure, definable by a formula containing only sets in* L *as parameters, then $\mathcal{R}_{\mathrm{L}} \subseteq X$.*

Proof. By Proposition 4, we can *w. l. o. g.* assume that $X \subseteq \Omega$ is a *Borel* set of full measure, coded in L. Then, it follows from observation at the end of Section 2.1, that there is a null measure $\mathbf{G}_\delta$ set $U \subseteq \Omega$, coded in L, such that $X' = \Omega \setminus X$ is a subset of U. However $U \cap \mathcal{R}_{\mathrm{L}} = \emptyset$ by definition. It follows that $\mathcal{R}_{\mathrm{L}} \subseteq X$, as required. □

2.4 Arithmetical randomness

Let us fix once and for all a recursive enumeration $\{u_l\}_{l\in\omega}$ of all finite binary sequences. Let $J_l = \Omega_{u_l} = \{x \in \Omega : u \subset x\}$. For any (infinite) sequence $c \in \Omega$, let U_c be the set of all pairs $\langle i, l\rangle$ of natural numbers such that $c(2^i 3^l) = 0$. (Thus U_c can be an arbitrary subset of $\mathbb{N}^2$.) We finally define

$$G_c = \bigcap_{i\in\omega} \bigcup_{\langle i,l\rangle\in U_c} J_l\,,$$

which is clearly an arbitrary $\mathbf{G}_\delta$ subset of Ω.

Let us say that a set $G \subseteq \Omega$ is *an arithmetically coded* $\mathbf{G}_\delta$ *set* iff $G = G_c$ for an arithmetically definable sequence $c \in \Omega$. ($c \in \Omega$ is said to be arithmetically definable iff there exists a formula with addition, multiplication, equality relation, the relation "$x(i) = 0$" and with quantifiers over natural numbers which is true if and only if $x = c$.)

Definition 6 (introduced in [2]) A sequence $x \in \Omega$ is *arithmetically random* iff it avoids any null measure arithmetically coded $\mathbf{G}_\delta$ set.

The formula saying that $x \in \Omega$ is arithmetically random is denoted by $\rho_{\rm A}(x)$. Put $\mathcal{R}_{\rm A} = \{x \in \Omega : \rho_{\rm A}(x)\}$ (all arithmetically random sequences). □

One easily proves, in **ZFC**, that $\mathcal{R}_{\rm L} \subseteq \mathcal{R}_{\rm A}$, or, in other words, $\rho_{\rm L}(x)$ implies $\rho_{\rm A}(x)$. Unlike $\mathcal{R}_{\rm L}$, the set $\mathcal{R}_{\rm A}$ is, provably in **ZFC**, a set of full measure. Clearly any Martin-Löf random sequence $x \in \Omega$ belongs to $\mathcal{R}_{\rm A}$.

2.5 A formula for "consistent" randomness

Let $\rho(x)$ be the formula saying:

- $x \in \mathcal{R}_{\rm A}$, and <u>if</u> $\mathcal{R}_{\rm L}$ is a set of full measure <u>then</u> $x \in \mathcal{R}_{\rm L}$.

Thus ρ defines the set $\mathcal{R}_{\rm L}$ of all Solovay random sequences over L — provided this is a set of full measure, while otherwise it defines simply the set $\mathcal{R}_{\rm A}$ of all arithmetically random sequences. It easily follows that ρ satisfies (3).

Theorem 7 *The formula* $\rho(x)$ *also satisfies requirements* (1) *and* $(2'')$.

Proof. It is clear that ρ provably in **ZFC** defines a set of full measure. Thus it remains to check $(2'')$. Let $\Psi(x)$ be a set theoretic formula such that **ZFC** proves that it defines a set of full measure. To prove the consistency of the statement that $\rho(x) \Longrightarrow \Psi(x)$, we show that the set $\{x \in \Omega : \rho(x)\ \&\ \neg\,\Phi(x)\}$ is empty in the Solovay model.

Indeed, in this model the set $X = \{x : \Psi(x)\}$ is definable by $\Psi(x)$, a parameter-free formula, and $\mathtt{mes}X = 1$ by the choice of Ψ, while we have $\{x : \rho(x)\} = \mathcal{R}_{\rm L}$, see above. It remains to apply Corollary 5. □

References

[1] G.J. Chaitin. "A theory of program size formally identical to information theory", *J. of ACM*, 22:329–340, 1975.

[2] B. Durand, A. Shen, N. Vereshagin. "Arithmetical randomness". (Manuscript), 1998.

[3] T. Jech, Set Theory, Academic Press, 1978.

[4] K. Kunen, *Set Theory, an introduction to independence proofs*, North–Holland, 1980.

[5] M. van Lambalgen, Independence, randomness, and the axiom of choice, *J. of Symbolic Logic*, 57:1274 – 1304 (1992).

[6] P. Martin-Löf. "The definition of random sequences", *Information and Control*, 9:602–619, 1966.

[7] An.A. Muchnik, A.L. Semenov, V.A. Uspensky. "Mathematical metaphysics of randomness," To appear in TCS, vol. 207, no. 2 (October 1998).

[8] V. A. Uspensky, A. L. Semenov, and A. Kh. Shen. "Can an individual sequence of zeros and ones be random?" *Russian Math. Surveys,* 45(no. 1):121–189 (1990).

[9] V. A. Uspensky and A. L. Semenov. Algorithms: Main Ideas and Applications. Kluwer Acad. Publ., Dordrecht, 1993.