Available online at www.sciencedirect.com
SCIENCE DIRECT®
Annals of Pure and Applied Logic 138 (2006) 20–30
ANNALS OF PURE AND APPLIED LOGIC
www.elsevier.com/locate/apal

On partial randomness

Cristian S. Calude[a,*], Ludwig Staiger[b], Sebastiaan A. Terwijn[c]

[a]*Department of Computer Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand*
[b]*Martin-Luther-Universität Halle-Wittenberg, Institut für Informatik, D - 06099 Halle, Germany*
[c]*Institute for Discrete Mathematics and Geometry, Technical University of Vienna, Wiedner Hauptstraße 8-10/E104, A-1040 Vienna, Austria*

Available online 7 July 2005

Abstract

If $\mathbf{x} = x_1x_2 \cdots x_n \cdots$ is a random sequence, then the sequence $\mathbf{y} = 0x_10x_2 \cdots 0x_n \cdots$ is clearly not random; however, $\mathbf{y}$ seems to be "about half random". L. Staiger [Kolmogorov complexity and Hausdorff dimension, Inform. and Comput. 103 (1993) 159–194 and A tight upper bound on Kolmogorov complexity and uniformly optimal prediction, Theory Comput. Syst. 31 (1998) 215–229] and K. Tadaki [A generalisation of Chaitin's halting probability Ω and halting self-similar sets, Hokkaido Math. J. 31 (2002) 219–253] have studied the degree of randomness of sequences or reals by measuring their "degree of compression". This line of study leads to various definitions of partial randomness. In this paper we explore some relations between these definitions. Among other results we obtain a characterisation of Σ_1-dimension (as defined by Schnorr and Lutz in terms of martingales) in terms of strong Martin-Löf ε-tests (a variant of Martin-Löf tests), and we show that ε-randomness for $\varepsilon \in (0, 1)$ is different (and more difficult to study) than the classical 1-randomness.

Keywords: Program-size complexity; Hausdorff dimension; Martingales; (Partial) randomness

1. Introduction

The program-size complexity $H(w)$ of a binary string w is the size, in bits, of the shortest program for a universal self-delimiting Turing machine U to calculate w. This

* Corresponding author. Tel.: +64 9 3737 599x85751; fax: +64 9 3737 453.
E-mail addresses: cristian@cs.auckland.ac.nz (C.S. Calude), staiger@informatik.uni-halle.de (L. Staiger), terwijn@logic.at (S.A. Terwijn).

0168-0072/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.apal.2005.06.004

complexity measure plays an important role in characterising the (algorithmic) randomness of infinite sequences and provides an elegant tool for proving information-theoretical forms of Gödel incompleteness (see [2,3,5,17]).

Although the class of random sequences is large (it has constructive measure one), there are many interesting examples of sequences which are not random, but "nearly random". For example, assume $\mathbf{x} = x_1 x_2 \cdots x_n \cdots$ is a random sequence; although the sequence $\mathbf{y} = 0x_1 0x_2 \cdots 0x_n \cdots$ is not random, $\mathbf{y}$ seems to be "about half random". Can we model this intuition?

Staiger [14,15] and Tadaki [16] have studied the degree of randomness of sequences (or reals) by measuring their "degree of compression" with a computable real number ε in the unit interval $[0, 1]$ as a parameter indicating the degree of compression. As ε becomes larger, the degree of randomness increases, so that in the case where $\varepsilon = 1$ one obtains the classical randomness. This line of study leads to various definitions of ε-randomness, probably not all equivalent. It is the aim of this paper to study various definitions for ε-randomness. Some natural results true for the case $\varepsilon = 1$ are false for $\varepsilon < 1$. For example, the analogue of the theorem stating that "a real $\alpha = 0.\mathbf{x}$ is classically 1-random iff there exist a constant $c \geq 0$ and an infinite computable set $M \subseteq \mathbb{N}$ such that $H(\mathbf{x}(n)) \geq n - c$, for each $n \in M$" is false for ε-randomness with $0 < \varepsilon < 1$, that is, the statement "a real $\alpha = 0.\mathbf{x}$ is ε-random iff there exist a constant $c \geq 0$ and an infinite computable set $M \subseteq \mathbb{N}$ such that $H(\mathbf{x}(n)) \geq \varepsilon \cdot n - c$, for each $n \in M$" is false. The study of ε-randomness with $\varepsilon < 1$ is more difficult than the study of classical 1-randomness; one of the reasons is that, as we shall see below, the ε-analogue of the Lebesgue measure, an essential tool for the study of randomness, is the ε-dimensional Hausdorff measure $\mathbb{L}_\varepsilon$, which is, unfortunately, infinite on every non-empty open set. This difficulty can be circumvented by using measures μ^ε for sets of finite strings and relating them to $\mathbb{L}_\varepsilon$.

The paper is organised as follows. In Section 2 we describe the notation; in Section 3 we review the main definitions, notions and results used later in the paper; in Sections 4 and 5 we study new types of ε-randomness and relate them to classical 1-randomness; in the last section we state some open problems.

2. Notation

We will follow the notation in [2]. By $\mathbb{N}_+ = \{1, 2, \ldots\}$ we denote the set of positive natural numbers; $\mathbb{Q}, \mathbb{R}, \mathbb{R}_+$ are the sets of rationals, reals, positive reals, respectively. The cardinality of the set A is denoted by $\mathrm{card}\,(A)$. Let us fix $X = \{0, 1\}$; by X^* we denote the set of finite strings (words) on X, including the *empty* string λ; occasionally we write $u \cdot v = uv$ to denote the concatenation of the strings u and v. The length of the string w is denoted by $|w|$ and $X^i = \{w \in X^* : |w| = i\}$. If v is a prefix of w we write $v \sqsubseteq w$. If $W, W' \subseteq X^*$, then $WW' = \{wv : w \in W, v \in W'\}$. A set $W \subseteq X^*$ is called prefix-free if for each $u, v \in W$ with $u \sqsubseteq v$, we have $u = v$.

A *self-delimiting Turing machine* (for short, a *machine*) is a Turing machine T processing binary strings such that its program set (domain) $PROG_T = \{x \in X^* : T \text{ halts on } x\}$ is a prefix-free set of strings. The *program-size complexity* of the string $x \in X^*$ (induced by T) is $H_T(x) = \min\{|y| \ : y \in X^*,\ T(y) = x\}$, where $\min \emptyset = \infty$.

We can effectively construct a machine U (called *universal*) such that for every machine T, $H_U(x) \le H_T(x) + O(1)$. In what follows we will fix U and put $H = H_U$.

A real s is left (right) computable if there is a computable sequence of rationals (s_t) such that $s_t \le s_{t+1}$ ($s_t \ge s_{t+1}$) and $s = \lim_{t\to\infty} s_t$. Left computable reals are also called c.e. reals.

We consider the Cantor space X^ω of infinite sequences (ω-words) over X. As we will focus mainly on irrational numbers, we will identify reals α in the unit interval with sequences $\mathbf{x} \in X^\omega$ via $\alpha \leftrightarrow 0.\mathbf{x}$. If $\mathbf{x} = x_1x_2 \cdots x_n \cdots \in X^\omega$, then $\mathbf{x}(n) = x_1x_2 \cdots x_n$ is the prefix of length n of $\mathbf{x}$. Strings and sequences will be denoted respectively by $u, v, w, \ldots$ and $\mathbf{x}, \mathbf{y}, \ldots$.

For $W \subseteq X^*$, WX^ω denotes the set $\{w\mathbf{x} : w \in W \wedge \mathbf{x} \in X^\omega\}$ of sequences having a prefix in W. The sets WX^ω are the open sets in the natural topology on X^ω. Computably enumerable (c.e.) open sets are sets of the form WX^ω, where $W \subseteq X^*$ is c.e. Let μ denote the usual product (Lebesgue) measure on X^ω given by $\mu(\{w\}X^\omega) = 2^{-|w|}$, for $w \in X^*$. For a measurable set $\mathcal{R}$ of infinite sequences, $\mu(\mathcal{R})$ is the probability that $\mathbf{x} \in \mathcal{R}$ when $\mathbf{x}$ is chosen by a random experiment in which an independent toss of a fair coin is used to decide whether $x_n = 1$. If W is prefix-free, then $\mu(WX^\omega) = \sum_{w\in W} 2^{-|w|}$.

Fix $\varepsilon > 0$. For any (not necessarily prefix-free) set $W \subseteq X^*$ we will write

$$\mu^\varepsilon(W) = \sum_{w\in W} 2^{-\varepsilon|w|}.$$

Note that $\mu^\varepsilon(w) < \sum_{x\in X} \mu^\varepsilon(wx)$ when $\varepsilon < 1$.

The ith section of a set $V \subseteq X^* \times \mathbb{N}_+$ is $V_i = \{w \in X^* : (w, i) \in V\}$.

Finally we will need the Kraft–Chaitin Theorem (see [2], p. 53): Let $n_1, n_2, \ldots$ be a computable sequence of non-negative integers such that $\sum_{i=1}^{\infty} 2^{-n_i} \le 1$. Then, we can effectively construct a prefix-free sequence of strings $w_1, w_2, \ldots$ such that for each $i \ge 1$, $|w_i| = n_i$.

3. Martingales, supermartingales and Hausdorff dimension

In this section we review the main definitions and results we need from the theories of martingales and Hausdorff dimension.

Definition 3.1 (*Ville [19]*). (a) A *martingale* is a function $d : X^* \to \mathbb{R}_+$ such that $2d(w) = d(w0) + d(w1)$, for every $w \in X^*$. A *supermartingale* is a function $d : X^* \to \mathbb{R}_+$ such that $2d(w) \ge d(w0) + d(w1)$, for every $w \in X^*$.

(b) A (super)martingale d *succeeds* on a sequence $\mathbf{x}$ (real $0.\mathbf{x}$) if $\limsup_{n\to\infty} d(\mathbf{x}(n)) = \infty$. A (super)martingale d *succeeds on*, or *covers a* set of sequences (reals) if it succeeds on each sequence (real) in the set. The *success set* $S[d]$ of d is the class of all sequences (reals) on which d succeeds.

(c) A (super)martingale d is *left computable* if the set $\{(x, r) : x \in X^*, r \in \mathbb{Q}, d(x) > r\}$ is c.e.; d is *right computable* if $\{(x, r) : x \in X^*, r \in \mathbb{Q}, d(x) < r\}$ is c.e.; d is *computable* if it is both left computable and right computable.

The following property of supermartingales with respect to prefix-free sets (see [13]) will be useful in what follows.

Proposition 3.2. *Let $C \subseteq X^*$ be prefix-free, and let $d : X^* \to \mathbb{R}_+$ be a supermartingale. Then $d(\lambda) \geq \sum_{w \in C} 2^{-|w|} \cdot d(w)$.*

We continue with the following results linking Lebesgue measure and supermartingales.

Theorem 3.3 (*Ville [19]*). (a) *A set of reals $\mathcal{R}$ has Lebesgue measure zero iff there is a (super)martingale that succeeds on $\mathcal{R}$.*

(b) *Let d be a (super)martingale and define*

$$S^k[d] = \big\{\mathbf{x} \in X^\omega : (\exists n > 0)\big[d(\mathbf{x}(n)) \geq k\big]\big\}.$$

Then, $\mu(S^k[d]) \leq d(\lambda)k^{-1}$.

Definition 3.4 (*Schnorr [13]*). An *order* is a nondecreasing unbounded function $h : \mathbb{N} \to \mathbb{N}$.[1] For a martingale d and order h we define

$$S_h[d] = \left\{\mathbf{x} \in X^\omega : \limsup_{n \to \infty} \frac{d(\mathbf{x}(n))}{h(n)} = \infty\right\}.$$

Schnorr also used null sets of the form $S_h[d]$ with $h(n) = 2^{\varepsilon n}$, $\varepsilon \in (0, 1]$, of *exponential* order.

Finally, we define the classical Hausdorff dimension [8] (see also Falconer [6]).

Definition 3.5. (a) A set $C \subseteq X^*$ is an *n-cover* if every string $w \in C$ has the length $|w| \geq n$. A set $C \subseteq X^*$ *covers* the set $\mathcal{R} \subseteq X^\omega$ if $\mathcal{R} \subseteq \bigcup_{w \in C} wX^\omega$.

(b) Put

$$\mathbb{L}_{\varepsilon,n}(\mathcal{R}) = \inf\left\{\sum_{w \in C} \left(2^{-|w|}\right)^\varepsilon : C \text{ is an } n\text{-cover of } \mathcal{R}\right\}, \tag{1}$$

and define the *ε-dimensional outer Hausdorff measure* of $\mathcal{R}$ to be

$$\mathbb{L}_\varepsilon(\mathcal{R}) = \lim_{n \to \infty} \mathbb{L}_{\varepsilon,n}(\mathcal{R}).$$

(c) The *Hausdorff dimension* of a set $\mathcal{R} \subseteq X^\omega$ is defined as $\dim \mathcal{R} = \inf\{\varepsilon : \mathbb{L}_\varepsilon(\mathcal{R}) = 0\}$.

It should be remarked that for every $\mathcal{R} \subseteq X^\omega$ there is exactly one "changeover point" α such that $\mathbb{L}_\varepsilon(\mathcal{R}) = \infty$ for $\varepsilon < \alpha$ and $\mathbb{L}_\varepsilon(\mathcal{R}) = 0$ for $\varepsilon < \alpha$. Moreover, Hausdorff dimension is countably stable, that is, $\dim \bigcup_{i \in \mathbb{N}} \mathcal{R}_i = \sup\{\dim \mathcal{R}_i : i \in \mathbb{N}\}$.

Observe further that in Eq. (1) the sum $\sum_{w \in C}\left(2^{-|w|}\right)^\varepsilon$ equals $\mu^\varepsilon(C)$ provided C is prefix-free.

The following theorem links Hausdorff dimension and supermartingales.

Theorem 3.6 (*Lutz [9]*). *For any class $\mathcal{R} \subseteq X^\omega$ the following statements are equivalent:*

(i) *The class $\mathcal{R}$ has Hausdorff dimension α,*

[1] An "Ordnungsfunktion" in Schnorr's terminology is always computable, whereas we prefer to leave the complexity of orders unspecified.

(ii) $\alpha = \inf\{s \in \mathbb{Q} : \exists d(d \text{ is a supermartingale } \wedge \mathcal{R} \subseteq S_{2^{(1-s)n}}[d])\}$.

Remark 3.7. In fact, Lutz proved Proposition 3.6 using what he called s-gales. He observed that $d' : X^* \to X^*$ is an s-gale iff $d(w) = 2^{(1-s)|w|}d'(w)$ is a martingale. From this it easily follows that the concept of s-gale gives rise to the same concept as was used by Schnorr [13], using martingales with exponential orders $h(|w|) = 2^{-(1-s)|w|}$. (This fact was also observed by Ambos-Spies et al. [1].)

We follow Schnorr's approach, because it seems that the combination of (super)martingales with order functions is more flexible at least in two respects: on the one hand, as in the investigation of Hausdorff dimension, it allows for the use of order functions other than exponential ones, and on the other hand, as the proof of Theorem 11 in [15] shows, computable martingales may achieve non-computable (exponential) order functions, something which is not possible for s-gales, as computable s-gales exist only for computable reals s.

4. Partial randomness

In this section we introduce Tadaki's definition [16] of Martin-Löf ε-randomness and the new notion of "strong Martin-Löf ε-randomness". We derive characterisations of strongly Martin-Löf ε-random sequences in terms of supermartingales and in terms of a priori program-size complexity.

A Martin-Löf test [10] is a uniform sequence $\{V_i\}$ of c.e. subsets of X^* such that the measure $\mu(V_i X^\omega)$ of the i-th set is smaller than 2^{-i}. To adapt this definition to the ε-case Tadaki [16] replaced the condition $\mu(V_i X^\omega) < 2^{-i}$ by $\mu^\varepsilon(V_i) < 2^{-i}$. Thus one obtains the following definition.

Definition 4.1. A *Martin-Löf ε-test* is a c.e. set $V \subseteq X^* \times \mathbb{N}_+$ such that $\mu^\varepsilon(V_i) < 2^{-i}$. A real $\mathbf{x} \in X^\omega$ is *Martin-Löf ε-random* if for every Martin-Löf ε-test V, $\mathbf{x} \notin \bigcap_i V_i X^\omega$.

Since, as was mentioned above, $\mu^\varepsilon(w) < \sum_{x\in X} \mu^\varepsilon(wx)$ whenever $\varepsilon < 1$, the simple procedure for transforming a Martin-Löf test into an equivalent Martin-Löf test having only prefix-free sections V_i (see e.g. [13]) cannot be applied here. Therefore, we introduce the following stronger version of Martin-Löf ε-tests.

Definition 4.2. A *strong Martin-Löf ε-test* is a c.e. set $V \subseteq X^* \times \mathbb{N}_+$ such that for every prefix-free set $C \subseteq V_i$ it holds that $\mu^\varepsilon(C) < 2^{-i}$. A real α is *strongly Martin-Löf ε-random* if for every strong Martin-Löf ε-test V, $\alpha \notin \bigcap_i V_i X^\omega$.

Remark 4.3. (a) Every strong Martin-Löf ε-test is a Martin-Löf ε-test; consequently, every strongly Martin-Löf ε-random real is Martin-Löf ε-random.

(b) If there is a strong Martin-Löf ε-test $V \subseteq X^* \times \mathbb{N}_+$ such that $\mathcal{R} \subseteq \bigcap_n V_n X^\omega$, then $\mathbb{L}_\varepsilon(\mathcal{R}) = 0$ in an effective way.

The last statement needs more explanation. For the case of random reals, that is, when $\varepsilon = 1$, it is well known that every set $\mathcal{R} \subseteq X^\omega$ having non-null Lebesgue measure $\mu(\mathcal{R}) > 0$ contains a random real. This is true also for ε, $0 < \varepsilon \leq 1$, when we replace the Lebesgue measure μ by the ε-dimensional measure $\mathbb{L}_\varepsilon$. Indeed, observe that $\mathbb{L}_\varepsilon(V_i X^\omega) \le \mu^\varepsilon(C) < 2^{-i}$, where $C \subseteq V_i$ is the prefix-free set mentioned in Definition 4.2. Thus $\mathbb{L}_\varepsilon(\bigcap_i V_i X^\omega) = 0$, for every strong Martin-Löf ε-test V, and as there are only countably many strong Martin-Löf ε-tests we have the following.

Proposition 4.4. *Let $\varepsilon \in (0, 1]$ and let $\mathcal{R}_\varepsilon \subseteq X^\omega$ be the set of all strongly Martin-Löf ε-random reals. Then, $\mathbb{L}_\varepsilon(X^\omega \setminus \mathcal{R}_\varepsilon) = 0$.*

The next two lemmata show an intrinsic relationship between left computable supermartingales and strong Martin-Löf ε-tests.

Lemma 4.5. *Let $s \in (0, 1]$ be a right computable real number and let $V \subseteq X^* \times \mathbb{N}_+$ be a strong Martin-Löf s-test. Then there is a left computable supermartingale $d : X^* \to \mathbb{R}_+$ such that $\bigcap_n V_n X^\omega \subseteq S_{2^{(1-s)n}}[d]$.*

Proof. Suppose V is a strong Martin-Löf s-test. Since s is right computable there is a computable sequence of rationals (s_t) such that $s_t \ge s_{t+1}$ and $s = \lim_{t\to\infty} s_t$. Let further $V_{n,t}$ be the computable approximation of V_n at stage t, and define

$$d_{n,t}(w) = n \cdot \max\Big\{2^{|w|} \cdot \sum_{v\in C} 2^{-s_t|v|} : C \subseteq V_{n,t} \cap wX^* \text{ prefix-free}\Big\}.$$

Then

$$d_{n,t}(wx) = n \cdot \max\Big\{2^{|wx|} \cdot \sum_{v\in C_x} 2^{-s_t|v|} : C_x \subseteq V_{n,t} \cap wxX^* \text{ prefix-free}\Big\},$$

for $x \in X$. Since the maximum achievable sets C_x, $x \in X$, may be chosen independently from each other such that $C_x \subseteq V_{n,t} \cap wxX^*$, their union $C = \bigcup_{x\in X} C_x$ is a prefix-free subset of $V_{n,t} \cap wX^*$ and

$$\begin{aligned}\sum_{x\in X} d_{n,t}(wx) &= \sum_{x\in X} n \cdot \max\Big\{2^{|wx|} \cdot \sum_{v\in C_x} 2^{-s_t|v|} : C_x \subseteq V_{n,t} \cap wxX^* \text{ prefix-free}\Big\}\\ &\le n \cdot \max\Big\{2 \cdot 2^{|w|} \cdot \sum_{v\in C} 2^{-s_t|v|} : C \subseteq V_{n,t} \cap wX^* \text{ prefix-free}\Big\}\\ &= 2 \cdot d_{n,t}(w).\end{aligned}$$

This proves that for each n and t, $d_{n,t}$ is a supermartingale. Observe that in view of $s_t \ge s_{t+1}$ and $V_{n,t} \subseteq V_{n,t+1}$ we have $d_{n,t}(w) \le d_{n,t+1}(w)$.

Evidently, each $d_{n,t}$ is a computable function. Next we define

$$d_n(w) = \lim_{t\to\infty} d_{n,t}(w), \quad d(w) = \sum_{n=0}^{\infty} d_n(w).$$

Then $d(w) < \infty$ for every $w \in X^*$, since $d(\lambda) \le \sum_n d_n(\lambda) \le \sum_n n \cdot 2^{-n}$ and d is a supermartingale. Furthermore, d is left computable. Finally, if $w \in V_n$, then $d_n(w) \ge n \cdot 2^{(1-s)|w|}$; hence if $\mathbf{x} \in \bigcap_n V_n X^\omega$, then $\mathbf{x} \in S_{2^{(1-s)n}}[d]$. $\square$

Lemma 4.6. *Let $s \in (0, 1]$ be a left computable real number and let $d : X^* \to \mathbb{R}_+$ be a left computable supermartingale. Then there is a strong Martin-Löf s-test $V \subseteq X^* \times \mathbb{N}_+$ such that $S_{2^{(1-s)n}}[d] \subseteq \bigcap_n V_n X^\omega$.*

Proof. Suppose s is a left computable real and d is a left computable supermartingale. We define the strong Martin-Löf s-test V in the following way. We choose $k \geq d(\lambda)$ and define

$$V_n = \left\{w : 2^{-|w|} \cdot d(w) \geq 2^{-s|w|} \cdot 2^n k\right\}.$$

Note that if $\mathbf{x} \in S_{2^{(1-s)n}}[d]$, then we have

$$\limsup_{n\to\infty} \frac{d(\mathbf{x}(n))}{2^{(1-s)n}} = \infty,$$

so $\mathbf{x} \in \bigcap_n V_n$.

We claim that the sets $\{V_n\}$ form a strong Martin-Löf s-test. First observe that V is c.e. since s and d are left computable. Next, let $C \subseteq V_n$ be prefix-free. Then, by construction of V_n, we have

$$2^n k \cdot \mu^s(C) = \sum_{w\in C} 2^{-s|w|} \cdot 2^n k \leq \sum_{w\in C} 2^{-|w|} d(w).$$

Using Proposition 3.2 this yields $2^n k \cdot \mu^s(C) \leq d(\lambda) \leq k$, so $\mu^s(C) \leq 2^{-n}$. □

Now Lemmata 4.5 and 4.6 yield the following.

Theorem 4.7. *For any class $\mathcal{R} \subseteq X^\omega$ the following statements are equivalent:*

(i) *The real α is minimal such that for all $\varepsilon > \alpha$ there is a strong Martin-Löf ε-test $V \subseteq X^* \times \mathbb{N}_+$ with $\mathcal{R} \subseteq \bigcap_n V_n X^\omega$.*
(ii) *The Σ_1-dimension of $\mathcal{R}$ is α, that is, $\alpha = \inf\{s \in \mathbb{Q} : \exists d$ (d is a left computable supermartingale and $\mathcal{R} \subseteq S_{2^{(1-s)n}}[d])\}$.*

The existence of a universal Martin-Löf ε-test, for computable ε was mentioned in [16, Remark 3.1]. In the case of strong Martin-Löf ε-tests the existence of universal left computable supermartingales (see [2, Theorem 4.17] or [9, Theorem 3.6]) gives a simple derivation of the existence of universal strong Martin-Löf ε-tests.

Let $\mathbf{d}$ be Levin's universal left computable supermartingale, that is, for every left computable supermartingale d there is a constant c_d such that $d(w) \leq c_d \cdot \mathbf{d}(w)$ holds for all $w \in X^*$. Then $S_{2^{(1-s)n}}[d] \subseteq S_{2^{(1-s)n}}[\mathbf{d}]$ and Lemmata 4.5 and 4.6 yield the existence of a universal strong Martin-Löf ε-test.

Theorem 4.8. *If $\varepsilon \in (0, 1]$ is a computable real number, then there is a universal strong Martin-Löf ε-test $U \subseteq X^* \times \mathbb{N}_+$, that is, U is a strong Martin-Löf ε-test and $\bigcap_n V_n X^\omega \subseteq \bigcap_n U_n X^\omega$, for every strong ε-test $V \subseteq X^* \times \mathbb{N}_+$.*

For individual sequences $\mathbf{x} \in X^\omega$ we obtain the following:

Theorem 4.9. *Let $\varepsilon \in (0, 1]$ be a computable real number and let $\mathbf{x} \in X^\omega$. Then the following are equivalent:*

1. $\mathbf{x}$ *is strongly Martin-Löf ε-random.*
2. $\mathbf{x} \notin S_{2^{(1-\varepsilon)n}}[\mathbf{d}]$.

The a priori Kolmogorov complexity KA is defined by $\mathrm{KA}(w) = |w| - \log_2 \mathbf{d}(w)$ (see [18]). Thus we obtain the following complexity-theoretic characterisation of strongly Martin-Löf ε-random sequences:

Corollary 4.10. *Let $\varepsilon \in (0, 1]$ be a computable real number. Then, $\mathbf{x}$ is strongly Martin-Löf ε-random iff there is a constant c such that $\mathrm{KA}(\mathbf{x}(n)) \geq \varepsilon \cdot n - c$, for almost all n.*

A similar property relating Martin-Löf ε-randomness to the program-size complexity H was shown by Tadaki [16].

Lemma 4.11. *Let $\varepsilon \in (0, 1]$ be a computable real number. Then, $\mathbf{x}$ is Martin-Löf ε-random iff there is a constant c such that $H(\mathbf{x}(n)) \geq \varepsilon \cdot n - c$, for almost all n.*

Both Lemma 4.11 and Corollary 4.10 show that the two versions of Martin-Löf ε-randomness do not limit the upper complexity of sequences. Thus every (strong) Martin-Löf ε-random $\mathbf{x} \in X^\omega$ is also (strong) Martin-Löf ε'-random for $\varepsilon' < \varepsilon$.

5. Randomness versus ε-randomness

In this section we continue to compare the classical theory of 1-randomness with the theory of ε-randomness with $\varepsilon \in (0, 1]$. First we mention that random reals have the following regular behaviour on a computable set of grid points ([7] and [11]; for a proof see [12]).

Theorem 5.1. *A real $\mathbf{x} \in X^\omega$ is random iff there exist a constant $c \geq 0$ and an infinite computable set $M \subseteq \mathbb{N}$ such that $H(\mathbf{x}(n)) \geq n - c$, for each $n \in M$.*

This result is no longer true for Martin-Löf ε-random reals $0.\mathbf{x}$ $(0 < \varepsilon < 1)$; see also Lemma 4.11. It was shown in [14, Example 3.18] that there are reals $\alpha = 0.\mathbf{x}$ which satisfy $\limsup_{n\to\infty} H(\mathbf{x}(n))/n = 1$ and, simultaneously, $\liminf_{n\to\infty} H(\mathbf{x}(n))/n = 0$. A closer look into this phenomenon yields the following:

Example 5.2. There is an $\mathbf{x} \in X^\omega$ such that for every $1/2 < \varepsilon < 1$, there are infinite computable sets $M_\varepsilon, M'_\varepsilon \subseteq \mathbb{N}$ for which $H(\mathbf{x}(n)) \geq \varepsilon \cdot n$, when $n \in M_\varepsilon$ and $H(\mathbf{x}(n)) \leq (1 - \varepsilon) \cdot n$, when $n \in M'_\varepsilon$.

Proof. We use an idea of Daley [4] and the construction of Example 3.18 in [14]. We define $\mathbf{x} = \prod_{i=0}^{\infty} w_i \cdot 0^{(2i+1)!}$, where w_i is a string with $|w_i| = (2i)!$ having $H(w_i) \geq |w_i|$. Further let

$$m_n = \sum_{i=0}^{n}((2i)! + (2i+1)!) + (2n+2)!, \;\; m'_n = \sum_{i=0}^{n}((2i)! + (2i+1)!)\ ,$$

and consider the computable sets

$$M = \{m_n : n \in \mathbb{N}\} \text{ and } M' = \left\{m'_n : n \in \mathbb{N}\right\}.$$

Then $\mathbf{x}(m_n) = \mathbf{x}(m'_n) \cdot w_{n+1}$ and $\mathbf{x}(m'_n) = \mathbf{x}(m_{n-1}) \cdot 0^{(2n+1)!}$, for the finite prefixes $\mathbf{x}(m_n)$ and $\mathbf{x}(m'_n)$ of $\mathbf{x}$. This leads to the inequalities

$$H(\mathbf{x}(m_n)) \geq H(w_{n+1}) - c \geq (2n+2)! - c\ ,$$
$$H(\mathbf{x}(m'_n)) \leq 2 \cdot (m_{n-1} + \log(2n+1)!) + c',$$

for all $n \in \mathbb{N}$ and suitably chosen constants c, c'.

Then, by construction of m_n and m'_n, for $1/2 < \varepsilon < 1$ there are only finitely many n such that

$$\frac{(2n+2)! - c}{m_n} < \varepsilon \text{ and } 1 - \varepsilon < \frac{2(m_{n-1} + \log(2n+1)!) + c'}{m'_n}.$$

Consequently, the sets $M_\varepsilon = \{m_n : m_n \in M \wedge (2n+2)! - c \geq \varepsilon \cdot m_n\}$ and

$$M'_\varepsilon = \{m'_n : m'_n \in M' \wedge 2(m_{n-1} + \log(2n+1)!) + c' \leq (1-\varepsilon) \cdot m'_n\}$$

are infinite computable sets satisfying our requirements. □

We proved above that a set $\mathcal{R} \subseteq X^\omega$ having $\mathbb{L}_\varepsilon(\mathcal{R}) > 0$ contains a strongly Martin-Löf ε-random real, and, consequently, it contains also a Martin-Löf ε-random real.

Next we are now going to show that the same is true for reals which were called strongly Chaitin ε-random in [16]. According to Lemma 4.11 every strongly Chaitin ε-random real is also Martin-Löf ε-random.

Definition 5.3. A real $\alpha = 0.\mathbf{x}$ is *strongly Chaitin ε-random* if $\lim_{n\to\infty}(H(\mathbf{x}(n) - \varepsilon \cdot n) = \infty$.

First we derive an auxiliary result which is essentially Theorem 3.4 of [16]. For the sake of completeness we give its proof. To this end we introduce an extra piece of notation, namely for $W \subseteq X^*$,

$$W^\delta = \big\{\mathbf{x} : \mathbf{x} \in X^\omega \wedge \{n : \mathbf{x}(n) \in W\} \text{ is infinite}\big\}.$$

Proposition 5.4. *Let $\varepsilon \in (0, 1]$ be a computable real. An $\mathbf{x} \in X^\omega$ is not strongly Chaitin ε-random iff there is an c.e. set $W \subseteq X^*$ such that $\sum_{w\in W} 2^{-\varepsilon|w|} < \infty$ and $\mathbf{x} \in W^\delta$.*

Proof. Assume that $H(\mathbf{x}(n)) < \varepsilon \cdot |w| + c$, for infinitely many n, and consider the c.e. set

$$W_{\varepsilon,c} = \{w : w \in X^* \wedge H(w) \leq \varepsilon \cdot |w| + c\}.$$

Then, clearly, $\mathbf{x} \in W^\delta_{\varepsilon,c}$.

Next let $W \subseteq X^*$ be c.e. and $\sum_{w\in W} 2^{-\varepsilon|w|} < 2^c$, for some $c \in \mathbb{N}$. Then $\sum_{w\in W} 2^{-(\lceil\varepsilon|w|\rceil+c)} < 1$. Since ε is computable, the set $M_W = \{(\lceil\varepsilon|w|\rceil+c, w) : w \in W\}$ is also c.e. and, because of Kraft–Chaitin Theorem, there is a machine $\phi : X^* \to X^*$ such that $\phi(X^*) = W$, and for every $w \in W$ there is a π such that $\phi(\pi) = w$ and $|\pi| = \lceil\varepsilon|w|\rceil + c$. This shows $H_\phi(w) \leq \lceil\varepsilon|w|\rceil + c$ and every $\mathbf{x} \in W^\delta$ is not strongly Chaitin ε-random. □

Using the fact that $X^\omega \setminus \bigcup_{c\in\mathbb{N}} W^\delta_{\varepsilon,c}$ is the set of all strongly Chaitin ε-random sequences and that $\sum_{w\in W} 2^{-\varepsilon|w|} < \infty$ implies $\mathbb{L}_\varepsilon(W^\delta) = 0$ (see [14, Lemma 3.8]), we obtain the following:

Theorem 5.5. *Let* $\varepsilon \in (0, 1]$ *and let*

$$\mathcal{P}_c = \{\mathbf{x} : \mathbf{x} \in X^\omega \wedge H(\mathbf{x}(n)) \leq \varepsilon \cdot n + c, \text{ for infinitely many } n\}.$$

Then, for all $c \in \mathbb{N}$, $\mathbb{L}_\varepsilon\big(\mathcal{P}_c\big) = 0$.

From this result we get the following analogue of Lemma 3.13 of [14]:

Corollary 5.6. *If* $\mathbb{L}_\varepsilon(\mathcal{R}) > 0$, *then* $\mathcal{R}$ *contains a strongly Chaitin* ε*-random real.*

We conclude this section by showing that the results of Theorem 5.5 and Proposition 4.4 are tight, that is, there are sets $\mathcal{R} \subseteq X^\omega$ having Hausdorff dimension $\dim \mathcal{R} = \varepsilon,\ 0 < \varepsilon \leq 1$, but not containing any Martin-Löf ε-random real.

Example 5.7. For $\varepsilon,\ 0 < \varepsilon \leq 1$ consider the set $Q_\varepsilon = \{(p, q) : p, q \in \mathbb{N}_+ \wedge p/q < \varepsilon\}$ and let

$$\mathcal{R} = \bigcup_{(p,q)\in Q_\varepsilon} (X^p \cdot 0^{q-p})^\omega.$$

The sets $(X^p \cdot 0^{q-p})^\omega$ are definable by finite automata, so the results of [14, Section 4] apply.

Hence we obtain, on the one hand, $\limsup_{n\to\infty} H(\mathbf{x}(n))/n \leq p/q < \varepsilon$, whenever $\mathbf{x} \in (X^p \cdot 0^{q-p})^\omega$, and, on the other hand, $\dim (X^p \cdot 0^{q-p})^\omega = p/q$ and, consequently, $\dim \mathcal{R} = \sup\{\dim (X^p \cdot 0^{q-p})^\omega : (p, q) \in Q_\varepsilon\} = \varepsilon$.

6. Conclusion and open questions

Tadaki had invented in [16] two versions of a concept of ε-randomness. He derived also complexity-theoretic characterisations of them (see Lemma 4.11 and Proposition 5.4 above). Up to now, it is open whether these concepts coincide or not.

As random sequences can be also characterised using left-computable supermartingales, we pursued this route and obtained a third concept of ε-randomness, which gives a close connection to supermartingales as well as a complexity-theoretic characterisation. However, it is also not known whether it coincides with one of Tadaki's concepts.

We conjecture that Martin-Löf ε-randomness does not imply strong Chaitin ε-randomness even for computable ε. It is also open which relations hold between strong Martin-Löf ε-randomness and strong Chaitin ε-randomness.

Acknowledgements

The authors wish to thank an anonymous referee for constructive criticism and useful comments. Part of the work of the first author was done in Halle with support from the Martin-Luther-Universität Halle-Wittenberg. The second author was in part supported by the University of Auckland Research Fund (Project Number 3602449/9343). The third author was supported by the Austrian Research Fund (Lise Meitner grant M699-N05).

References

[1] K. Ambos-Spies, W. Merkle, J. Reimann, F. Stephan, Hausdorff dimension in exponential time, in: Proc. 16th Conference on Computational Complexity, IEEE Computer Society, 2001, pp. 210–217.

[2] C.S. Calude, Information and Randomness. An Algorithmic Perspective, 2nd ed., Revised and Extended, Springer Verlag, Berlin, 2002.

[3] G.J. Chaitin, Algorithmic Information Theory, Cambridge University Press, Cambridge, 1987 (3rd printing 1990).

[4] R.P. Daley, The extent and density of sequences within the minimal-program complexity hierarchies, J. Comput. System Sci. 9 (1974) 151–163.

[5] R. Downey, D. Hirschfeldt, Algorithmic Randomness and Complexity, Springer-Verlag (in preparation).

[6] K. Falconer, Fractal Geometry. Mathematical Foundations & Applications, Wiley & Sons, New York, 1990.

[7] L. Fortnow, Personal communication to L. Staiger, May 2003.

[8] F. Hausdorff, Dimension und äußeres Maß, Math. Ann. 79 (1919) 157–179.

[9] J.H. Lutz, The dimensions of individual strings and sequences, Inform. and Comput. 187 (2003) 49–79.

[10] P. Martin-Löf, The definition of random sequences, Inform. and Control 9 (1966) 602–619.

[11] A. Nies, Personal communication to L. Staiger, May 2003.

[12] A. Nies, F. Stephan, S.A. Terwijn, Randomness, relativization and Turing degrees, J. Symbolic Logic (in press).

[13] C.-P. Schnorr, Zufälligkeit und Wahrscheinlichkeit, in: Lect. Notes in Math., vol. 218, Springer-Verlag, Heidelberg, 1971.

[14] L. Staiger, Kolmogorov complexity and Hausdorff dimension, Inform. and Comput. 103 (1993) 159–194.

[15] L. Staiger, A tight upper bound on Kolmogorov complexity and uniformly optimal prediction, Theory Comput. Syst. 31 (1998) 215–229.

[16] K. Tadaki, A generalization of Chaitin's halting probability Ω and halting self-similar sets, Hokkaido Math. J. 31 (2002) 219–253.

[17] S.A. Terwijn, Complexity and randomness, Rendiconti del Seminario Matematico di Torino 62 (1) (2004) 1–38.

[18] V.A. Uspensky, A.Kh. Shen, Relations between varieties of Kolmogorov complexities, Math. Systems Theory 29 (1996) 271–292.

[19] J. Ville, Étude critique de la notion de collectif, Gauthier-Villars, Paris, 1939.