Representation of Left-Computable ε–Random Reals[☆]

Cristian S. Calude[*,a,1], Nicholas J. Hay[a,1], Frank Stephan[b,2]

[a] *Department of Computer Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand*
[b] *Department of Mathematics and School of Computing, National University of Singapore, Singapore 117543*

Abstract

In this paper we introduce the notion of ε–universal prefix-free Turing machine (ε is a computable real in (0,1]) and study its halting probability. The main result is the extension of the representability theorem for left-computable random reals to the case of ε–random reals: *a real is left-computable ε–random iff it is the halting probability of an ε–universal prefix-free Turing machine.* We also show that left-computable ε–random reals are provable ε–random in Peano Arithmetic. The theory developed here parallels to a large extent the classical theory, but not completely. For example, random reals are Borel normal (in any base), but for $\varepsilon \in (0, 1)$, some ε–random reals do not contain even arbitrarily long runs of 0s.

Key words: ε–universal prefix-free Turing machine, halting probability, ε–random real, Peano Arithmetic

1. Introduction

A real α is left-computable (or recursively/computably enumerable) if there is a computable increasing sequence of rationals which converges to α. Left-

[☆] A preliminary version of this paper was presented at the Joint AMS–NZMS Meeting, Wellington, NZ, December 2007.

[*] Corresponding author.

Email addresses: `cristian@cs.auckland.ac.nz` (Cristian S. Calude), `nickjhay@gmail.com` (Nicholas J. Hay), `fstephan@comp.nus.edu.sg` (Frank Stephan)

[1] Supported in part by a Hood Fellowship.

[2] Supported in part by NUS grant numbers R146-000-114-112 and R252-000-308-112.

computable random reals can be characterised using various tools including prefix-complexity, Martin-Löf tests, martingales, Chaitin Omega numbers and universal probability [1, 3, 5, 6, 8, 11, 15].

Some left-computable reals are not random, but "partially random." For example, inserting a 0 in between adjacent bits of a (left-computable) random sequence produces a non-random sequence, having some weak randomness properties: this sequence is, as intuition suggests, left-computable (because it is left-approximated by approximations of the original sequence in which a 0 was inserted in between each adjacent bits) and 1/2–random.

The papers [4, 12, 16, 17, 18, 19] have studied the degree of randomness of reals (or sequences) by measuring their "degree of compression." In what follows ε is a fixed computable real number with $0 < \varepsilon \leq 1$. We study ε–randomness of reals, both intrinsically and in relation to the classical notion of randomness (which corresponds to $\varepsilon = 1$, here referred to as 1–randomness or simply randomness).

Our main tool is the ε–universal prefix-free Turing machine, a machine that can simulate any other prefix-free machine: the length of the simulating program on the ε–universal machine is bounded up to a fixed constant by the length of the simulated program divided by ε. In case $\varepsilon = 1$ we get the classical notion of universal machine.

We show that the halting probability of an ε–universal prefix-free Turing machine is left-computable and ε–random. Generalising the corresponding representability theorem of left-computable random reals [1, 3, 8, 11] we show that the converse is also true: every left-computable ε–random real is the halting probability of an ε–universal prefix-free Turing machine. A specific ε–universal Turing machine V_ε is obtained via Equation (1) below; the main principle is to "dilute" a universal Turing machine V. This machine plays an important role as its halting probability is the least with respect to H–reducibility of all ε–random reals.

The theory developed here parallels to a large extent the classical theory, but not completely. The following two results show interesting differences: (a) the prefix-free complexities induced by universal machines differ by at most an additive constant, but the difference between prefix-free complexities induced by ε–universal machines may be unbounded, (b) random reals are Borel normal (in any base), but some ε–random reals do not contain even arbitrarily long runs of 0s.

The paper is organised as follows. In Section 2 we present the necessary notation and previous results. In Section 3 we introduce and study

ε–universal machines and their halting probabilities. In Section 4 we study left-computable ε–random reals and in Section 5 we present the representability theorem for left-computable ε–random reals. In Section 6 we discuss the provability in Peano Arithmetic of ε–randomness for left-computable reals. In Section 7 we disprove Stay's conjecture regarding the 1–randomness (with respect to U) of the halting probability of an ε–universal machine U. We end with a few conclusions.

2. Notation and background

Let $\Sigma = \{0, 1\}$ and denote by Σ^n and Σ^* the set of all bit-strings of length n and the set of all bit-strings, respectively. The length of $\sigma \in \Sigma^*$ is denoted by $|\sigma|$. By $\log n$ we abbreviate the function $\lfloor \log_2(n+1) \rfloor$. Let $\mathbb{N} = \{1, 2, \ldots\}$ and let $\mathrm{bin} : \mathbb{N} \to \Sigma^*$ be the bijection which associates to every $n \geq 1$ its binary expansion without the leading 1.

To every infinite binary sequence $\alpha_1\alpha_2 \cdots \alpha_n \cdots$ we associate the real number $\alpha = 0.\alpha_1\alpha_2 \cdots \alpha_n \cdots$ in (0,1]. We denote by $\alpha \restriction n = \alpha_1\alpha_2 \cdots \alpha_n$ the prefix of length n of α's expansion. In this way, reals are identified with infinite binary sequences. Similarly, if $\mathbf{x} = x_1x_2 \cdots x_n \cdots$ is an infinite sequence, $\mathbf{x} \restriction n = x_1x_2 \cdots x_n$.

We assume that the reader is familiar with algorithmic information theory, cf. [1, 8] and present only a few notions to fix the notation.

If the Turing machine T is defined on σ we write $T(\sigma) < \infty$; the domain of T is the set $\mathrm{dom}(T) = \{\sigma \in \Sigma^* \ : \ T(\sigma) < \infty\}$. A prefix-free (Turing) machine is a Turing machine whose domain is a prefix-free set of strings. The prefix complexity of a string induced by a prefix-free machine W is $H_W(\sigma) = \inf\{|p| \ : \ W(p) = \sigma\}$. From now on *all Turing machines will be prefix-free and will be referred to as machines.*

We use several times the Kraft-Chaitin Theorem: given a computable enumeration of positive integers n_i such that $\sum_i 2^{-n_i} \leq 1$, we can effectively construct a prefix-free set of binary strings $\{x_i\}$ such that $|x_i| = n_i$, for all $i \geq 1$.

Throughout the whole paper ε is assumed to be a computable real in the interval (0,1]. Fix a machine W. A sequence $\mathbf{x}$ is Chaitin (ε, W)–random if there is a constant $c > 0$ such that for every $n \geq 1$, $H_W(\mathbf{x} \restriction n) \geq \varepsilon \cdot n - c$; $\mathbf{x}$ is strictly Chaitin (ε, W)–random if $\mathbf{x}$ is Chaitin (ε, W)–random, but not Chaitin (δ, W)–random for any δ with $\varepsilon < \delta \leq 1$.

If W is universal (from now on called 1–universal), then we get Tadaki's definition of weak Chaitin ε–randomness (see [4, 18]). If W is 1–universal and $\varepsilon = 1$, then we get Chaitin's classical definition of randomness [5, 6]. A real is Chaitin (ε, W)–random (shortly, (ε, W)–random) if its binary expansion is Chaitin (ε, W)–random.

For any prefix-free set $A \subset \Sigma^*$ we define $\Omega_A = \sum_{x \in A} 2^{-|x|}$. The halting probability of a machine W is $\Omega_W = \sum_{x \in \mathrm{dom}(W)} 2^{-|x|}$.

Following Tadaki [18], for any (not necessarily prefix-free) set $W \subseteq \Sigma^*$ and computable $\delta > 0$ we write $\mu^\delta(W) = \sum_{w \in W} 2^{-\delta \cdot |w|}$. If $\delta > 1$ and W is prefix-free, then $\mu^\delta(W) \leq \Omega_W \leq 1$. However, if $0 < \delta < 1$ then we can have $\mu^\delta(W) = \infty$ even for prefix-free W (for example, for $W = \{1^{\log n} 0 \mathrm{bin}(n) : n > 0\}$ and $0 < \delta < 1/2$).

3. ε–universal machines

In this section we introduce and study the notion of ε–universal machine.

In analogy with the classical case we say, following Stay [14], that a machine U is *ε–universal* if for every machine T there exists a constant $c_{U,T}$ such that for each program $\sigma \in \mathrm{dom}(T)$ there exists a program $p \in \mathrm{dom}(U)$ such that

$$U(p) = T(\sigma) \text{ and } \varepsilon \cdot |p| \leq |\sigma| + c_{U,T}.$$

If $\varepsilon = 1$ we get the classical notion of universal machine. Every universal machine is ε–universal, but the converse is not true (see Theorem 2).

A machine U is *strictly* ε–universal if U is ε–universal but not δ–universal for any δ with $\varepsilon < \delta \leq 1$.

Lemma 1. *The machine U is ε–universal iff there exists a 1–universal machine V and a constant $c_{U,V}$ such that for all $\sigma \in \Sigma^*$ we have $\varepsilon \cdot H_U(\sigma) \leq H_V(\sigma) + c_{U,V}$.*

Theorem 2. *Let V be a 1–universal machine and define*

$$V_\varepsilon(p 0^{\lfloor (1/\varepsilon - 1)|p| \rfloor}) = V(p). \tag{1}$$

Then:

(a) *V_ε is a machine and for all $\sigma \in \Sigma^*$ we have $H_{V_\varepsilon}(\sigma) = \lfloor H_V(\sigma)/\varepsilon \rfloor$,*
(b) *V_ε is strictly ε–universal.*

Proof. Clearly V_ε is a machine and the equality in (a) can be directly checked. From (a) and Lemma 1 we deduce the ε–universality of V_ε. If there were a constant c such that for all $\sigma \in \Sigma^*$, $\delta \cdot H_{V_\varepsilon}(\sigma) \leq H_V(\sigma) + c$, for some $\varepsilon < \delta \leq 1$, then in view of (a) we would have $(\delta/\varepsilon - 1) \cdot H_V(\sigma) \leq c + \delta$, for all $\sigma \in \Sigma^*$, a contradiction (H_V is unbounded). So, V_ε is strictly ε–universal. □

Theorem 3. *Let V be a 1–universal machine. Then for every ε–universal machine U, Ω_U is (ε, V)–random.*

Proof. Let f be a computable one-to-one function which enumerates $\mathrm{dom}(U)$. Let $\omega_k = \sum_{j=1}^{k} 2^{-|f(j)|}$. Clearly, (ω_k) is a computable, increasing sequence of rationals converging to Ω_U. Consider the binary expansion of $\Omega_U = 0.\Omega_1\Omega_2\cdots$.

We define a machine T as follows: on input $\sigma \in \Sigma^*$, T first "tries to compute" the smallest number t with $\omega_t \geq 0.\sigma$. If successful, $T(\sigma)$ is the first (in quasi-lexicographical order) string not belonging to the set $\{U(f(1)), U(f(2)), \ldots, U(f(t))\}$; if no such t exists then $T(\sigma) = \infty$.

Fix a number $m \geq 1$ and note that T is defined on $\Omega_U \restriction m$. Let t be the smallest number (computed in the first step of the computation of T) with $\omega_t \geq 0.\Omega_U \restriction m$. We have

$$0.\Omega_U \restriction m \leq \omega_t < \omega_t + \sum_{s=t+1}^{\infty} 2^{-|f(s)|} = \Omega_U \leq 0.\Omega_U \restriction m + 2^{-m}\,.$$

Hence, $\sum_{s=t+1}^{\infty} 2^{-|f(s)|} \leq 2^{-m}$, which implies $|f(s)| \geq m$, for every $s \geq t+1$. From the construction of T we conclude that

$$H_U(T(\Omega_U \restriction m)) \geq m. \tag{2}$$

Since T is a partially computable function, we get a constant c' such that for all $\sigma \in \Sigma^*$ for which $T(\sigma) < \infty$ we have:

$$H_V(T(\sigma)) \leq H_V(\sigma) + c'. \tag{3}$$

Using (2), the ε–universality of U, and (3) we obtain

$$\begin{array}{rcl} \varepsilon \cdot m & \leq & \varepsilon \cdot H_U(T(\Omega_U \restriction m)) \\ & \leq & H_V(T(\Omega_U \restriction m)) + c \\ & \leq & H_V(\Omega_U \restriction m) + c + c', \end{array}$$

which proves that Ω_U is (ε, V)–random. □

Corollary 4. *If V be a 1–universal machine, then Ω_{V_ε} is (ε, V)–random and $(1, V_\varepsilon)$–random.*

Proof. The halting probability Ω_{V_ε} is (ε, V)–random because of Theorem 2 (b) and Theorem 3. Using this fact and Theorem 2 (a) we deduce that Ω_{V_ε} is $(1, V_\varepsilon)$–random. □

Next we present a mechanism for producing examples of ε–universal machines.

Let A, B be infinite, prefix-free (recursively/computably) enumerable sets. Generalising the strong simulation in [3], we say that the set A *ε–strongly simulates* the set B (write $B \leq_\varepsilon A$) if there is a constant $c > 0$ and a partial computable function $f : \Sigma^* \stackrel{o}{\to} \Sigma^*$ satisfying the following three conditions:

(a) $A = \mathrm{dom}(f)$,
(b) $B = f(A)$ and
(c) $\varepsilon \cdot |\sigma| \leq |f(\sigma)| + c$, for all $\sigma \in A$.

The function f is called an ε–strong simulation of A onto B.

Proposition 5. *If V is a 1–universal machine and f is an ε–strong simulation of* $\mathrm{dom}(V)$ *onto a prefix-free computably enumerable set A, then $V \circ f$ is an ε–universal machine with domain A.*

Proof. Recall that $(V \circ f)(p) = V(f(p))$ for all $p \in \Sigma^*$. Fix a machine T. Since V is 1–universal there exists a constant c_T such that for each $p \in \mathrm{dom}(T)$ there exists a $\sigma \in \mathrm{dom}(V)$ satisfying $|\sigma| \leq |p| + c_T$ and $V(\sigma) = T(p)$. Since f is onto there exists $\tau \in A$ such that $f(\tau) = \sigma$. Since f is an ε–strong simulation we have $\varepsilon \cdot |\tau| \leq |f(\tau)| + c = |\sigma| + c$. Combining the previous two equations we deduce that for every $p \in \mathrm{dom}(T)$ there exists an $\tau \in A$ such that $\varepsilon \cdot |\tau| \leq |p| + c_T + c$ and $V(f(\tau)) = T(p)$, so $V \circ f$ is ε–universal. □

It may seem that the difference between the cases $\varepsilon = 1$ and $0 < \varepsilon < 1$ is just technical. Here is a deeper difference. If V and V' are 1–universal machines, then their complexities H_V and $H_{V'}$ differ by at most an additive constant [1]. *This result is not true for ε–universal machines.* To prove the claim we construct the following sequence of machines $V_{\varepsilon,k}$ by means of a fixed 1–universal machine V. We let

$$f_{\varepsilon,k}(p) = \begin{cases} p0^{\lfloor(1/\varepsilon-1)|p|-k\cdot\log(|p|)\rfloor}, & \text{if } (1/\varepsilon-1)|p|-k\cdot\log|p|\ge 1,\\ p1, & \text{otherwise,}\end{cases} \quad (4)$$

$$V_{\varepsilon,k}\circ f_{\varepsilon,k} = V. \quad (5)$$

Note that only for finitely many strings p the value $f_{\varepsilon,k}(p)$ is defined by the otherwise-case. Furthermore, equation (5) means that $V_{\varepsilon,k}(f_{\varepsilon,k}(p)) = V(p)$ for all $p \in \mathrm{dom}(V)$ and $V_{\varepsilon,k}(q)$ is undefined for all $q \notin \{f_{\varepsilon,k}(p) : p \in \mathrm{dom}(V)\}$.

Theorem 6. *The following properties are true:*

(a) $V_{\varepsilon,k}$ *is a machine and* $H_{V_{\varepsilon,k}}(\sigma) = \lfloor H_V(\sigma)/\varepsilon - k\cdot\log H_V(\sigma)\rfloor$, *for almost all strings* σ,
(b) $V_{\varepsilon,k}$ *is strictly* ε*–universal,*
(c) *We have* $H_{V_{\varepsilon,k}}(\sigma) - H_{V_{\varepsilon,k+1}}(\sigma) \ge \log H_V(\sigma) - 1 \to \infty$ *whenever* $|\sigma| \to \infty$,
(d) $\Omega_{V_{\varepsilon,k}}$ *is* (ε, V)*–random.*

Proof. Properties (a)–(c) follow from (4) and (5) using the technique presented in the proof of Theorem 2. In detail, the equality in (a) can be directly checked; ε–universality follows from (a) and Lemma 1. To show that $V_{\varepsilon,k}$ is strictly ε–universal we suppose, by absurdity, that there exist two constants c, δ such that $c > 0$, $1 > \delta > \varepsilon$ and $\delta \cdot H_{V_{\varepsilon,k}}(\sigma) \le H_V(\sigma) + c$ for all $\sigma \in \Sigma^*$. Then given the equality (a) we would have $(\delta/\varepsilon - 1) \le H_V(\sigma) \le \delta \cdot k \cdot \log H_V(\sigma) + c + \delta$, for almost all strings σ, a contradiction since H_V is unbounded. Property (c) follows from (a) and property (d) follows from (b) and Theorem 3. □

4. Left-computable (ε, V)–random reals

We now study (ε, V)–random reals with the following reducibility relation: a real α is *H–reducible* to a real β, written $\alpha \le_H \beta$, if there exists a 1–universal machine V and a constant $c > 0$ such that for all $n \ge 1$, we have $H_V(\alpha \restriction n) \le H_V(\beta \restriction n) + c$. Of course, the choice of the 1–universal machine V is irrelevant. Two reals α, β are *H–equivalent* if $\alpha \le_H \beta$ and $\beta \le_H \alpha$.

Recall that a real γ is *ε–convergent* [18] if there exists an increasing computable sequence of rationals $\{a_n\}$ converging to γ such that $\sum_{n=1}^{\infty}(a_{n+1} - a_n)^{\varepsilon} < \infty$.

Theorem 7. *Let V be a 1–universal machine. For every left-computable (ε, V)–random real α, $\Omega_{V_\varepsilon} \leq_H \alpha$.*

Proof. Tadaki [19, Theorem 4.6 (i) and (iv)] shows the following equivalence: a left-computable real α is (ε, V)-random iff for every left-computable ε–convergent real β there exists a constant c such that for all n, $H_V(\beta \upharpoonright n) \leq H_V(\alpha \upharpoonright n) + c$.

Now start with left-computable (ε, V)–random real α. Because Ω_{V_ε} is left-computable and ε–convergent we can apply the above mentioned equivalence to deduce the existence of a constant c such that $H_V(\Omega_{V_\varepsilon} \upharpoonright n) \leq H_V(\alpha \upharpoonright n) + c$, i.e. $\Omega_{V_\varepsilon} \leq_H \alpha$. □

Comment 8. Theorem 7 shows that Ω_{V_ε} is up to H–equivalence the least of all (ε, V)–random reals. In fact, there is one left-computable real below all other left-computable (ε, V)–random reals.

Proposition 9. *Let V be a 1–universal machine. Assume that $\varepsilon \in (0, 1)$ is computable. Then, for almost all constants c, and for every string x there exist two strings y, z of length c such that*

1. $H_V(xy) \geq H_V(x) + \varepsilon \cdot c + 1$,
2. $H_V(x) - \varepsilon \cdot c + 1 \leq H_V(xz) \leq H_V(x) + \varepsilon \cdot c - 1$.

Furthermore, z can be chosen as 0^c.

Proof. The proof follows mainly along the lines of Lemma 1 in [12] (with $\rho(x) = 2^{-\varepsilon|x|}$).

For item 1, given x and c we find y of length c such that $H_V(y|(x, H_V(x))) \geq c$; such an y exists by the pigeon hole principle. Then $H_V((x, y)) \geq H_V(x) + c - d$ and $H_V(xy) \geq H_V(x) + c - H_V(c) - d$, for some constant d independent of c. The first inequality follows from Theorem 2.3.6 in [8] and the second inequality follows from the first one by noting that (x, y) can be computed from xy and c; the constant d is taken such that it satisfies both inequalities. Now all sufficiently large c satisfy $c - H_V(c) - d \geq \varepsilon \cdot c + 1$.

For item 2, note that $H_V(x)$ and $H_V(x0^c)$ differ at most by $H_V(c) + d'$ from each other, where d' is again a constant independent of c. The function $c \mapsto H_V(c) + d'$ is dominated by the function $c \mapsto \varepsilon \cdot c - 1$, hence the given inequalities hold. □

Theorem 10. *Let V be a 1–universal machine. Assume that ε is a computable real in* (0,1)*. There exists a left-computable α and a constant C such that for all $n \geq 1$, $|H_V(\alpha \restriction n) - n \cdot \varepsilon| \leq C$.*

Proof. In view of Proposition 9 there is a constant c such that for all $\sigma \in \Sigma^*$:

1. σ has an extension τ of length $|\sigma|+c$ such that $H_V(\tau) > H_V(\sigma)+\varepsilon \cdot c+1$,
2. $H_V(\sigma) - c < H_V(\sigma 0^c) < H_V(\sigma) + \varepsilon \cdot c - 1$.

Let T be the tree of all strings $\sigma \in \Sigma^*$ whose all prefixes η with $|\eta|$ are a multiple of c have the property $H_V(\eta) \geq \varepsilon \cdot |\eta|$. Note that whenever σ is a node of length $n \cdot c$, by the first condition, there is an extension of σ in T of length $n \cdot c + c$.

Let α be the left-most infinite branch of T, hence left-computable. If $H_V(\alpha \restriction (n \cdot c)) > n \cdot c \cdot \varepsilon + 2c + 1$, then $\alpha \restriction (n \cdot c) 0^c$ is in T as

$$H_V(\alpha \restriction (n \cdot c) 0^c) > n \cdot c \cdot \varepsilon + c + 1 > (n \cdot c + c) \cdot \varepsilon.$$

As α is the leftmost infinite branch, $\alpha \restriction (n \cdot c + c) = \alpha \restriction (n \cdot c) 0^c$. Consequently, by the second condition,

$$H_V(\alpha \restriction (n \cdot c + c)) < H_V(\alpha \restriction (n \cdot c)) + \varepsilon \cdot c - 1,$$

hence $H_V(\alpha \restriction (n \cdot c + c))$ is at least by 1 less than the target $H_V(\alpha \restriction (n \cdot c))$. From this it follows that $|H_V(\alpha \restriction (n \cdot c)) - n \cdot c \cdot \varepsilon|$ is bounded by a constant.

The tree T is the intersection of trees $T_0, T_1, T_2, \ldots$ where each T_s contains an infinite branch β iff the initial segments σ of β of length $0 \cdot c$, $1 \cdot c$, …, $s \cdot c$ satisfy the $H_{V,s}(\sigma) \geq \varepsilon \cdot |\sigma|$. The left-most branches α_s of T_s are uniformly computable and approximate α from the left, hence α is left-computable. $\Box$

Comment 11. Note that in [12] an essentially similar construction for the real α in Theorem 10 was given; the new fact in Theorem 10 is the property of α to be left-computable. If one does not need the left-computable part of the result, one can construct α by a simple induction: append the corresponding strings previously obtained and keep the complexity of $\alpha(0)\alpha(1)\ldots\alpha(n)$ to be $\varepsilon \cdot n$ up to an additive constant. This method does not work with $\varepsilon = 1$ as it is known that whenever $H_V(\alpha(0)\alpha(1)\ldots\alpha(n)) \geq n - c$ for all n then

$$\forall d \forall^\infty n [H_V(\alpha(0)\alpha(1)\ldots\alpha(n)) \geq n + d].$$

Hence the existence of α in Theorem 10 holds only $0 < \varepsilon < 1$, another difference between 1-randomness and ε–randomness.

Corollary 12. *Assume that ε is computable in* (0,1) *and V is a 1–universal machine. a) There is a constant C such that for all n, $|H_V(\Omega_{V_\varepsilon} \upharpoonright n) - \varepsilon \cdot n| \leq C$. b) The real Ω_{V_ε} is strictly (ε, V)-random.*

Proof. From Corollary 4, Ω_{V_ε} is (ε, V)–random. In view of Theorem 10 there exists a left-computable real α and a constant C such that for all n, $|H_V(\alpha \upharpoonright n) - \varepsilon \cdot n| \leq C$. In particular, α is left-computable and (ε, V)-random, so by Theorem 7 there exists a constant c such that for all n:

$$H_V(\Omega_{V_\varepsilon} \upharpoonright n) \leq H_V(\alpha \upharpoonright n) + c \leq \varepsilon \cdot n + c + C. \quad (6)$$

The converse inequality comes from Corollary 4.

Finally, b) is a consequence of (6). □

It is well known that Ω_V is Borel absolutely normal[3] [1] If $\alpha = 0.\alpha_1\alpha_2\cdots$ is $(1, V)$–random then the real $\beta = 0.\alpha_1 0 \alpha_2 0 \cdots$ is $(1/2, V)$–random and not Borel normal (because in its binary expansion, in the limit, the frequency of 0's is three times larger than the frequency of 1s).

We show not that Ω_{V_ε} is more than not Borel normal:

Proposition 13. *Let V be a 1–universal machine. Assume that ε is computable in* (0,1) *and α is a left-computable real such that there is a constant C such that for all n, $|H_V(\alpha \upharpoonright n) - \varepsilon \cdot n| \leq C$. Then, for every binary string τ there is a constant c such that τ^c is not an infix of α.*

Proof. The proof follows along the lines of the proof of Proposition 9. To see this, note that for every strings σ, τ and positive integer c we have

$$H_V(\sigma\tau^c) \leq H_V(\sigma) + H_V(\tau) + H_V(c) + d,$$

for some constant d independent of σ, τ, c. Hence, if all prefixes σ of α satisfy the inequality

$$|\sigma| \cdot \varepsilon - c' \leq H_V(\sigma) \leq |\sigma| \cdot \varepsilon + c',$$

[3] A real is absolutely Borel normal if its digits, in every base, follow the uniform distribution: all digits are equally likely, all pairs of digits are equally likely, all triplets of digits are equally likely, etc.

then for every τ there is a value for c such that

$$H_V(\tau) + H_V(c) + d < \varepsilon \cdot c - 2c',$$

and thus whenever $H_V(\sigma) \leq |\sigma| \cdot \varepsilon + c'$ we have $H_V(\sigma\tau^c) < |\sigma\tau^c| \cdot \varepsilon - c'$ and $\sigma\tau^c$ is not a prefix of α. □

Corollary 14. *For every binary string τ there is a constant c such that τ^c does not occur in Ω_{V_ε} as a substring.*

Proof. Use Corollary 12 (a) and Proposition13. □

Comment 15. If $\alpha = 0.\alpha_1\alpha_2\cdots$ is $(1, V)$–random then the real $\beta = 0.\alpha_1\alpha_1\alpha_2\alpha_2\cdots$ is $(1/2, V)$–random but doesn't satisfy the hypothesis of Proposition 13.

5. Representability of left-computable (ε, V)–random reals

In this section we generalise the representability of left-computable random reals [3, 11] for the case of left-computable (ε, V)–random reals.

Theorem 16. *Let V be a 1–universal machine. Every left-computable (ε, V)–random number in (0,1] is the halting probability of an ε–universal machine.*

Proof. Given V and ε we consider the machine V_ε defined by (1). Recall that $\mathrm{dom}(V_\varepsilon)$ is the set of all strings $p0^{\lfloor (1/\varepsilon - 1)|p| \rfloor}$ with $p \in \mathrm{dom}(V)$. Now Ω_{V_ε} can be represented by the sum $\sum_{q \in \mathrm{dom}(V_\varepsilon)} 2^{-|q|}$. This sum is ε–convergent, as

$$\sum_{q \in \mathrm{dom}(V_\varepsilon)} (2^{-|q|})^\varepsilon \leq \sum_{p \in \mathrm{dom}(V)} (2^{1-|p|/\varepsilon})^\varepsilon \leq \sum_{p \in \mathrm{dom}(V)} 2^{\varepsilon - |p|} \leq 2^\varepsilon \cdot \Omega_V < \infty.$$

Hence Ω_{V_ε} is ε–convergent.

By Theorem 4.6 (i,v) in [19], given a left-computable and (ε, V)–random real α we can construct a left-computable real $\beta \geq 0$ and a rational $q > 0$ (in fact, we can take q to be 2^{-m}, for some $m > 0$) such that $\alpha = \beta + 2^{-m} \cdot \Omega_{V_\varepsilon}$, hence

$$\begin{aligned}\alpha &= \beta + 2^{-m} \cdot \sum_{p\in \mathrm{dom}(V_\varepsilon)} 2^{-|p|} \\ &= 2 \cdot \sum_{r\in \mathrm{dom}(T)} 2^{-|r|-1} + \sum_{p\in \mathrm{dom}(V_\varepsilon)} 2^{-|p|-m}\end{aligned}$$

where the machine T is constructed from the left-computable real β using the Kraft-Chaitin Theorem.

Define now the ε-universal machine W by the formula:

$$W(s) = \begin{cases} 0, & \text{if } s = 1r \text{ and } r \in \mathrm{dom}(T), \\ V_\varepsilon(s), & \text{if } s = 0^m p \text{ and } p \in \mathrm{dom}(V_\varepsilon), \\ \infty, & \text{otherwise,} \end{cases}$$

and notice that its domain is the disjoint union of the sets $\{1r : r \in \mathrm{dom}\, T\} \cup \{0^m p : p \in \mathrm{dom}(V_\varepsilon)\}$, hence

$$\alpha = \sum_{s\in \mathrm{dom}(W)} 2^{-|s|} = \Omega_W.$$

This completes the proof. □

6. Provability of left-computable (ε, V)–random reals

Peano Arithmetic (see [10], shortly, PA) is the first-order theory given by a set of 15 axioms defining discretely ordered rings, together with induction axioms for each formula $\varphi(x, y_1, \ldots, y_n)$: $\forall \overline{y}(\varphi(0, \overline{y}) \wedge \forall x(\varphi(x, \overline{y}) \to \varphi(x + 1, \overline{y})) \to \forall x(\varphi(x, \overline{y})))$.

The proof in [2] can be adapted to show that *every left-computable* (ε, V)*–random real is provable* (ε, V)*–random in* PA. This means the following: if PA is given an algorithm for computing the computable real ε, an algorithm for a machine U, a proof that U is prefix-free and ε–universal, then it can prove that Ω_U is left-computable and (ε, V)–random. This proof requires ε to be defined in terms of primitive recursive functions, which is always possible by a result in [13].[4]

[4] The proof in [2] has been precisely formalised and mechanically proved in the interactive theorem prover Isabelle [7]. It should be straight forward to adapt this proof to the more general ε–random case.

Another representation which can be used to prove (ε, V)–randomness is the following: if PA is given an algorithm for computing the computable real ε, an algorithm for a machine V, a proof that V is prefix-free and ε–universal, a positive integer c, and a computable increasing sequence of rationals converging to a real $\gamma > 0$, then PA can prove that $\alpha = 2^{-c} \cdot \Omega_V + \gamma$ is (ε, V)–random.

Is any "representation" of a (ε, V)–random real enough to guarantee PA provability of (ε, V)–randomness? To answer this question we fix an effective enumeration of all left-computable reals in (0,1), $\{\gamma_i\}$. Such an enumeration can be based on an enumeration of all increasing primitive recursive sequences of rationals in (0,1). Our question becomes: based solely on the index i can we always prove in PA that "γ_i is (ε, V)–random real" in case γ_i is (ε, V)–random real? We answer this question in the negative. To this aim we define the following sets:

$$\mathfrak{R}_{\mathrm{lc}} = \{\gamma \in (0,1) \,:\, \gamma \text{ is left-computable}\} = \{\gamma_i\},$$

$$\mathfrak{R}_{\mathrm{lc}}(\varepsilon, V) = \{\gamma \in \mathfrak{R}_{\mathrm{lc}} \,:\, \gamma \text{ is } (\varepsilon, V)\text{–random}\},$$

$$\mathfrak{R}_{\mathrm{lc}}^{\mathrm{PA}}(\varepsilon, V) = \{\gamma \in \mathfrak{R}_{\mathrm{lc}} \,:\, \gamma \text{ is provable } (\varepsilon, V)\text{–random in PA}\}.$$

By enumerating proofs in PA we deduce that the set $\mathfrak{R}_{\mathrm{lc}}^{\mathrm{PA}}(\varepsilon, V)$ is computably enumerable.[5] Is $\mathfrak{R}_{\mathrm{lc}}(\varepsilon, V)$ computably enumerable?

We use Lemma 26 from [2]:

Lemma 17. *If $A \subseteq \mathfrak{R}_{\mathrm{lc}}$ is computably enumerable, then for every left-computable reals $\alpha, \beta \in A$ such that $\beta > \alpha$, then we have $\beta \in A$.*

Theorem 18. *The set $\mathfrak{R}_{\mathrm{lc}}(\varepsilon, V)$ is not computably enumerable, so there exists $\alpha \in \mathfrak{R}_{\mathrm{lc}}(\varepsilon, V) \setminus \mathfrak{R}_{\mathrm{lc}}^{\mathrm{PA}}(\varepsilon, V)$.*

Proof. Consider $\alpha \in \mathfrak{R}_{\mathrm{lc}}(\varepsilon, V)$ and define the left-computable real β in the following way. If $\alpha \geq 1/2$, then the real $\beta = (\alpha \restriction n)11 \cdots 1 \cdots$ (where $\alpha \restriction (n+1) = 1^n 0$); if $\alpha < 1/2$ consider the left-computable real $\beta = (\alpha \restriction n)11 \cdots 1 \cdots$ (where $\alpha \restriction (n+1) = 0^m 1^{n-m} 0$). In both cases $\beta > \alpha$ and

[5] Recall that a set $A \subseteq \mathfrak{R}_{\mathrm{lc}}$ is computably enumerable if the set $\{i \in \mathbb{N} \,:\, \gamma_i \in A\}$ is computable enumerable (as a set of non-negative integers). In such a set we enumerate all indices for all elements in A.

$\beta \notin \mathfrak{R}_{\rm lc}(\varepsilon, V)$, which shows, by Lemma 17, that $\mathfrak{R}_{\rm lc}(\varepsilon, V)$ is not computably enumerable, thus concluding the proof. □

In fact, a more precise result is true:

Theorem 19. *For every $\alpha \in \mathfrak{R}_{\rm lc}(\varepsilon, V)$ there exists an index i such that $\alpha = \gamma_i$ and* PA *cannot prove the statement "γ_i is (ε, V)–random."*

Proof. The set $A_\alpha = \{\gamma_i : \alpha = \gamma_i\} \subset \mathfrak{R}_{\rm lc}(\varepsilon, V) \subset \mathfrak{R}_{\rm lc}$ is not computably enumerable. □

7. Stay's Conjecture

Stay [14] studied generalisations of the statement that Ω_U is random for every 1–universal machine U. In particular he conjectured that Ω_U is $(1, U)$–random for every ε–universal machine U. Although our results show that Ω_U is (ε, V)–random (for a 1-universal machine V; Theorem 3) and the conjecture is true for V_ε (Corollary 4), it turns out that the conjecture itself is too general and does not hold. We provide now a strong counterexample.

Theorem 20. *There exists a $\frac{1}{16}$–universal machine U such that Ω_U is not $(\frac{1}{2}, U)$–random, hence not $(1, U)$–random.*

Proof. Let V be an 1-universal machine. From V and input σ we define $U(\sigma)$ using a parameter τ which satisfies the right-hand side conditions in the following definition:

$$U(\sigma) = \begin{cases} \tau 0^{16n}, & \text{if } \exists n > 0\ [\sigma = 1^n 0 \tau \text{ and } |\tau| = 8n], \\ V(\tau), & \text{if } \exists n, m > 0,\ \tau \in \mathrm{dom}(V) \\ & \qquad [\sigma = 0\tau 0^n 1, |\sigma| = 4^{m+1} \text{ and } |\tau| \leq 4^m], \\ \infty, & \text{otherwise.} \end{cases}$$

Clearly, U is a machine. Given $\tau \in \mathrm{dom}(V)$, let $m_\tau = \min\{k > 0 : |\tau| \leq 4^k\}$ and $n_\tau = 4^{m_\tau+1} - |\tau| - 2$. Then $U(0\tau 0^{n_\tau} 1) = V(\tau)$ and $|0\tau 0^{n_\tau} 1| \leq 16 \cdot |\tau|$, hence U is $\frac{1}{16}$–universal.

Now consider the binary expansion of the halting probability Ω_U. The first bit after the dot is 1 as the strings starting with 1 contribute $\frac{1}{2}$ to the halting probability of U. Furthermore, the strings of length 4^{m+1} starting with a 0 in the domain of U contribute $4^{-m-1} \cdot a_m$ to the halting probability of U; here a_m is the number of strings up to the length 4^m in the domain of V. Because $a_m < 2^{4^m}$ it follows that a_m can be written with 4^m bits. So, in the binary expansion of Ω_U, the bits from the positions $4^m + 1$ until $3 \cdot 4^m$ are all 0; the bits from the positions $3 \cdot 4^{m+1} + 1$ to 4^{m+1} describe the binary value of a_m.

Let $m \geq 4$, $8n = 4^m$ and let τ be the string of the first $8n$ bits of Ω_U after the dot. Then $U(1^n 0\tau) = \tau 0^{16n}$ is a prefix of Ω_U of length $24n$ which is generated by the program $1^n 0 \tau$ of length $9n + 1$ as Ω_U has 0s on the positions $4^m + 1, \ldots, 3 \cdot 4^m$ and $24n \leq 3 \cdot 4^m$. Consequently, the real Ω_U is not $(\frac{1}{2}, U)$–random. □

8. Conclusion

In this paper we have introduced the notion of ε–universal machine and studied its halting probability. An ε–universal machine is capable of simulating every other machine, but less efficiently than a universal machine V. More precisely, the length of the simulating program on the universal machine is bounded up to a fixed constant by the length of the simulated program divided by ε. The halting probability of an ε–universal machine is left-computable and (ε, V)–random. The main result of this paper is the extension of the representability theorem for left-computable random reals to the case of ε–random reals: *a real is left-computable and (ε, V)–random iff it is the halting probability of an ε–universal machine*. Furthermore, we showed that left-computable ε–random reals are provable (ε, V)–random in Peano Arithmetic, for some, but not all of their representations. Finally we refuted Stay's conjecture stating that Ω_U is $(1, U)$–random provided U is ε–universal.

Acknowledgements

We thank Mike Stay for suggesting the definition of ε–universal machine, Kohtaro Tadaki for suggesting a simplification of the definition of V_ε, and both for valuable discussions. We also thank the anonymous referees for many comments and suggestions that improved the presentation.

References

[1] Cristian S. Calude. *Information and Randomness. An Algorithmic Perspective*, 2nd Edition, Revised and Extended, Springer Verlag, Berlin, 2002.

[2] Cristian S. Calude and Nicholas J. Hay. Every computably enumerable random real is provably computably enumerable random, *Logic Jnl. IGPL* 17 (2009,) 325–350.

[3] Cristian S. Calude, Peter Hertling, Bakhadyr Khoussainov and Yongge Wang. Recursively enumerable reals and Chaitin Ω numbers, in: M. Morvan, C. Meinel, D. Krob (eds.), *Proceedings of the 15th Symposium on Theoretical Aspects of Computer Science* (Paris), Springer–Verlag, Berlin, 1998, 596–606. Full paper in *Theoret. Comput. Sci.* 255 (2001), 125–149.

[4] Cristian S. Calude, Ludwig Staiger and Sebastiaan A. Terwijn. On partial randomness, *Annals of Applied and Pure Logic*, 138 (2006), 20–30.

[5] Gregory J. Chaitin. A theory of program size formally identical to information theory, *Journal of the Association of Computing Machinery* 22 (1975), 329–340.

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

[7] Nicholas J. Hay. Isabelle proof of the theorem "left-computable random reals are provable random in Peano Arithmetic", `http://www.cs.auckland.ac.nz/~nickjhay/SolovayRepresentation.thy`

[8] André Nies. *Computability and Randomness*. Oxford Press, Oxford, 2009.

[9] Piergiorgio G. Odifreddi. *Classical Recursion Theory*, vol. 1, Elsevier, 1997.

[10] Richard Kaye. *Models of Peano Arithmetic*, Oxford Press, Oxford, 1991.

[11] Antonin Kučera and Theodore A. Slaman. Randomness and recursive enumerability, *SIAM Journal on Computing*, 31, 1 (2001), 199–211.

[12] Jan Reimann and Frank Stephan. On hierarchies of randomness tests, in S. S. Goncharov, H. Ono, R. Downey (eds.). *Proceedings 9th Asian Logic Conference, "Mathematical Logic in Asia"*, World Scientific, Singapore, 2006, 215–232.

[13] Dimiter Skordev. Characterization of the computable real numbers by means of primitive recursive functions, in J. Blanck, V. Brattka, P. Hertling (eds.). *Proceedings of Computability and Complexity in Analysis 2000*, LNCS 2064, Springer-Verlag, Berlin, 2001, 296–309.

[14] Mike Stay. Personal communication to C. Calude, 7 May 2007.

[15] Rorbert M. Solovay. *Draft of a paper (or series of papers) on Chaitin's work ... done for the most part during the period of Sept.–Dec. 1974*, unpublished manuscript, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, May 1975, 215 pp.

[16] Ludwig Staiger. Kolmogorov complexity and Hausdorff dimension, *Information and Computation* 103 (1993), 159–194.

[17] Ludwig Staiger. A tight upper bound on Kolmogorov complexity and uniformly optimal prediction, *Theory of Computing Systems* 31 (1998), 215–229.

[18] Kohtaro Tadaki. A generalization of Chaitin's halting probability Ω and halting self-similar sets, *Hokkaido Mathematical Journal* 31 (2002), 219–253.

[19] Kohtaro Tadaki. Equivalent characterizations of partial randomness for recursively enumerable real, `arXiv:0805.2691`, 2008. (also at `http://ims.nju.edu.cn/conference/randomness/tadaki.pdf`)