HAL
archives-ouvertes.fr

Kolmogorov Complexity and Solovay Functions

Laurent Bienvenu, Rodney Downey

▶ To cite this version:

Laurent Bienvenu, Rodney Downey. Kolmogorov Complexity and Solovay Functions. 26th International Symposium on Theoretical Aspects of Computer Science STACS 2009, Feb 2009, Freiburg, Germany. pp.147-158. inria-00359056

HAL Id: inria-00359056
https://hal.inria.fr/inria-00359056

Submitted on 5 Feb 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Symposium on Theoretical Aspects of Computer Science 2009 (Freiburg), pp. 147–158
www.stacs-conf.org

KOLMOGOROV COMPLEXITY AND SOLOVAY FUNCTIONS

LAURENT BIENVENU [1] AND ROD DOWNEY [1]

[1] School of Mathematics, Statistics and Computer Science
Victoria University
P.O. Box 600
Wellington, New Zealand
E-mail address: {laurent.bienvenu,rod.downey}@mcs.vuw.ac.nz

Abstract. Solovay [19] proved that there exists a computable upper bound f of the prefix-free Kolmogorov complexity function K such that $f(x) = K(x)$ for infinitely many x. In this paper, we consider the class of computable functions f such that $K(x) \leq f(x)+O(1)$ for all x and $f(x) \leq K(x) + O(1)$ for infinitely many x, which we call Solovay functions. We show that Solovay functions present interesting connections with randomness notions such as Martin-Löf randomness and K-triviality.

1. Introduction

The plain and prefix-free Kolmogorov complexities (which we denote respectively by C and K) are both non-computable functions, but they do admit computable *upper bounds*. How good can these upper bounds be? That is, how close to C (resp. K) can a computable upper bound of C (resp. K) be? It can be easily proven that no computable upper bound of C can be close to C on all values, i.e. given any computable upper bound f of C, the ratio $f(x)/C(x)$ is not bounded. To see this, we use a variation of Berry's paradox: take a computable upper bound f of C, and define, for all $n \in \mathbb{N}$, x_n to be the smallest string x such that $f(x) \geq n$. Since f is computable, x_n can be computed from n, hence $C(x_n) \leq \log(n) + O(1)$. Thus, $f(x_n)/C(x_n) \geq n/(\log(n) + O(1))$ which proves the result. The exact same argument shows that no computable upper bound of K approximates K well on all values.

Therefore, one may ask the natural question: are there computable upper bounds for C or K that are good approximations on infinitely many values? The answer is trivially yes for C. Indeed, for most strings x, we have $C(x) = |x| + O(1)$ (see for example Downey and Hirschfeldt [7]), hence for some constant c, the function f defined by $f(x) = |x| + c$ is a computable upper bound of C, and $f(x) = C(x) + O(1)$ for infinitely many strings x. The case of K is less clear: indeed, the maximal prefix-free complexity of a string x of length n

1998 ACM Subject Classification: F.4.1.
Key words and phrases: Algorithmic randomness, Kolmogorov complexity, K-triviality.
The authors are supported by a grant from the Marsden fund of New Zealand.

SYMPOSIUM ON THEORETICAL ASPECTS OF COMPUTER SCIENCE

© L. Bienvenu and R. Downey
Creative Commons Attribution-NoDerivs License

(attained by most strings of that length) is $n + K(n) + O(1)$. And giving a good upper bound of this last expression already necessitates a good upper bound on K! Solovay [19] nonetheless managed to construct a computable upper bound f of K such that $f(x) = K(x)$ for infinitely many x. In this paper, we consider the class of computable functions f such that $K(n) \leq f(n) + O(1)$ for all n and $f(n) \leq K(n) + O(1)$ for infinitely many n, which we call Solovay functions.

Our first main result (Theorem 2.5) is that Solovay functions have a very simple characterization: they correspond to the computable functions f such that $\sum_x 2^{-f(x)}$ is finite and is a Martin-Löf random real.

Then, we discuss the role of Solovay functions in the characterization of randomness notions. In particular, we show (Theorem 3.4) that Solovay functions are particularly relevant to the Miller-Yu characterization of Martin-Löf random sequences via the plain Kolmogorov complexity of the initial segments. We prove along the way a theorem of independent interest (Theorem 3.5) showing that the Levin-Schnorr characterization of Martin-Löf randomness by prefix-free Kolmogorov complexity is very sharp, and derive several interesting consequences of this result.

Finally, we study two triviality notions that relate to computable upper bounds of prefix-free Kolmogorov complexity and Solovay functions. In the spirit of the Miller-Yu theorem, we obtain (Theorem 4.3) a characterization of K-triviality via computable upper bounds of K.

We assume that the reader is familiar with the field of algorithmic randomness. If not, one can consult Downey and Hirschfeldt [7] or Nies [18]. We denote by $2^{<\omega}$ and 2^ω the set of binary sequences (or "strings") and binary infinite sequences respectively. For a binary sequence x (finite or infinite), we denote by $x(i)$ the $(i+1)$-th bit of x, and by $x \upharpoonright i$ the string made of the first i bits of x (that is, $x \upharpoonright i = x(0)x(1)\dots x(i-1)$). The length of a string x is denoted by $|x|$. Throughout this paper, we identify $2^{<\omega}$ with $\mathbb{N}$, via the usual length-lexicographic bijection: $0 = \epsilon$ (ϵ being the empty string), $1 = 0$, $2 = 1$, $3 = 00$, $4 = 01$, $5 = 10\dots$. We also identify any element $r \in [0,1]$ to an element $\alpha \in 2^\omega$ such that $r = \sum_n \alpha(n)2^{-n+1}$. If r is not dyadic then α is unique; if r is dyadic, there are two possible choices for $\alpha \in 2^\omega$ and which one we choose does not matter in this paper. We say that a real number is left-c.e. if it is the limit of a computable nondecreasing sequence of rational numbers. Given a nondrecreasing unbounded function $f : \mathbb{N} \to \mathbb{N}$, we denote by f^{-1} the function defined by $f^{-1}(k) = \min\{n \in \mathbb{N} \mid f(n) \geq k\}$. As we stated earlier, we denote by $C(x)$ and $K(x)$ the plain Kolmogorov complexity and prefix-free Kolmogorov of a string x. Since C and K are enumerable from above (i.e. their upper graph is a c.e. set), for a fixed enumeration, let $C_s(x)$ and $K_s(x)$ be the value of $C(x)$ and $K(x)$ at the s-th stage of the enumeration. In particular, this means that the function $(x, s) \mapsto K_s(x)$ is computable and, for any fixed x, $s \mapsto K_s(x)$ is nonincreasing and converges to $K(x)$ (and the same is true for C).

2. Computable upper bounds of Kolmogorov complexity

The class of computable upper bounds of Kolmogorov complexity has been studied in Bienvenu and Merkle [2] (in the setting of "decidable machines"), where they were used to give characterizations of a wide variety of randomness notions of randomness, such as

Martin-Löf randomness, Schnorr randomness, Kurtz randomness or computable dimension. Here, we are interested in the class of Solovay functions, which is a subclass of computable upper bounds of K (here and from now on, we use a slight abuse of terminology, calling "upper bound" of K a function f such that $K \leq f + O(1)$).

Let us first mention that computable upper bounds of K admit a very simple characterization.

Lemma 2.1. *Let $f : \mathbb{N} \to \mathbb{N}$ be a computable function. The following are equivalent:*
(i) $K \leq f + O(1)$
(ii) The sum $\sum_{n\in\mathbb{N}} 2^{-f(n)}$ is finite.

Proof. $(i) \Rightarrow (ii)$ is trivial as $\sum_n 2^{-K(n)} \leq 1$. For $(ii) \Rightarrow (i)$, let c be such that $\sum_n 2^{-f(n)} \leq 2^c$. Using the Kraft-Chaitin theorem, we effectively construct a prefix-free c.e. set of strings $\{x_n \mid n \in \mathbb{N}\}$ with $|x_n| = f(n) + c$ for all n. Then, we define a (computable) function F by $F(x_n) = n$ for all n. Since F has prefix-free domain and is computable, we have $K(n) \leq |x_n| + O(1)$ for all n, hence $K(n) \leq f(n) + c + O(1)$. ∎

Definition 2.2. We denote by $\mathfrak{K}$ the class of computable functions such that $\sum_n 2^{-f(n)} < +\infty$ (or equivalently, the computable functions f such that $K \leq f + O(1)$).
We call *Solovay function* any function $f \in \mathfrak{K}$ such that $\liminf_{n\to+\infty} f(n) - K(n) < +\infty$ (or equivalently, any function $f \in \mathfrak{K}$ such that for some c, $f(n) \leq K(n) + c$ for infinitely many n).

Theorem 2.3 (Solovay [19])**.** *Solovay functions exist.*

Proof. Let us start by an observation. Given $x \in 2^{<\omega}$, and some p such that $\mathbb{U}(p) = x$, if we call t the computation time of $\mathbb{U}(p)$, we have

$$K\left(\langle x, p, t\rangle\right) \leq |p| + O(1)$$

(where $\langle ., ., .\rangle$ is a computable bijection from $2^{<\omega} \times 2^{<\omega} \times 2^{<\omega}$ to $2^{<\omega}$). Indeed, given p only, one can easily compute x and t. Suppose now that p is a *shortest* $\mathbb{U}$-program for x i.e. $\mathbb{U}(p) = x$ and $K(x) = |p|$. We then have:

$$|p| = K(x) \leq K\left(\langle x, p, t\rangle\right) \leq |p| + O(1)$$

Thus, let f be the function defined by:

$$f(\langle x, p, t\rangle) = \begin{cases} |p| \text{ if } \mathbb{U}(p) \text{ outputs } x \text{ in exactly } t \text{ steps of computation} \\ +\infty \text{ otherwise} \end{cases}$$

(here we use the value $+\infty$ for convenience, but any coarse upper bound of $K(\langle x, p, t\rangle)$, like $2|x| + 2|p| + 2\log t$, would do). By the above discussion, we have $K \leq f + O(1)$ and $f(\langle x, p, t\rangle) \leq K(\langle x, p, t\rangle) + O(1)$ for all triples (x, p, t) such that p is a shortest $\mathbb{U}$-program for x and $\mathbb{U}(p)$ outputs x in exactly t steps of computation. Thus, f is as desired. ∎

Remark 2.4. In fact, what Solovay actually proved is that there exists a computable function f such that $K \leq f$ and $K(n) = f(n)$ for infinitely many n. This can be easily deduced from Theorem 2.3. Indeed, given a computable function f such that $K \leq f + O(1)$ and $c = \liminf_{n\to+\infty} f(n) - K(n) < +\infty$, the (computable) function $f' = f - c$ is such that $f'(n) = K(n)$ for infinitely many n, and $f'(n) \geq K(n)$ for almost all n. Hence, up to modifying only finitely many values of f', we may assume that $f'(n) \geq K(n)$ for all n.

It turns out that, among the computable functions f such that the sum $\sum_n 2^{-f(n)}$ is finite, the Solovay functions are precisely those for which this sum is not only finite but also a Martin-Löf random real.

Theorem 2.5. *Let f be a computable function. The following are equivalent:*
(i) f is a Solovay function.
(ii) The sum $\sum_n 2^{-f(n)}$ is finite and is a Martin-Löf random real.

Proof. $(i) \Rightarrow (ii)$. If f is a Solovay function, we already know by definition that $\alpha = \sum_n 2^{-f(n)}$ is finite. Let us now prove that α is a Martin-Löf random real. Suppose it is not. Then for arbitrarily large c there exists k such that $K(\alpha \upharpoonright k) \leq k - c$ (this because of the Levin-Schnorr theorem, see next section). Given $\alpha \upharpoonright k$, one can effectively find some s such that

$$\sum_{n>s} 2^{-f(n)} \leq 2^{-k}$$

Thus, by a standard Kraft-Chaitin argument, one has $K(n|\alpha \upharpoonright k) \leq f(n) - k + O(1)$ for all $n > s$. Thus, for all $n > s$:

$$K(n) \leq f(n) + K(\alpha \upharpoonright k) - k + O(1) \leq f(n) + (k - c) - k - O(1) \leq f(n) - c - O(1)$$

And since c can be taken arbitrarily large, this shows that $\lim_{n\to+\infty} f(n) - K(n) = +\infty$ i.e. f is not a Solovay function.

$(ii) \Rightarrow (i)$. Suppose now for the sake of contradiction that f is not a Solovay function and that α is Martin-Löf random. We will prove that under these assumptions, the number $\Omega = \sum_n 2^{-K(n)}$ is not Martin-Löf random, which indeed is a contradiction (see for example Downey and Hirschfeldt [7]).

Since α is Martin-Löf random and is left-c.e., we can apply the Kučera-Slaman theorem [13]. This theorem states that given a Martin-Löf random left-c.e. real η, for any left-c.e. real ξ, there exists a constant d and a partial recursive function φ such that for every rational $q < \eta$, $\varphi(q)$ is defined and $\xi - \varphi(q) < 2^d(\eta - q)$. We will use this fact for $\eta = \alpha$ and $\xi = \Omega$ and also call d and φ the associated constant and partial recursive function.

Now, let c be a large integer (to be specified later). Suppose also that $\alpha \upharpoonright k$ is given for some k. Since $\alpha - (\alpha \upharpoonright k) < 2^{-k}$, by the Kučera-Slaman theorem:

$$\Omega - \varphi(\alpha \upharpoonright k) < 2^{-k+d}$$

Thus, from $\alpha \upharpoonright k$, one can effectively compute some $s(k)$ such that

$$\sum_{n>s(k)} 2^{-K(n)} \leq 2^{-k+d}$$

If k is large enough, then $s(k)$ is large enough and hence $n > s(k) \Rightarrow K(n) \leq f(n) - c - d$ (this because f is not a Solovay function). Thus, for all k large enough:

$$\sum_{n>s(k)} 2^{-f(n)} \leq 2^{-c-d} \sum_{n>s(k)} 2^{-K(n)} \leq 2^{-c-d} \cdot 2^{-k+d} \leq 2^{-k-c}$$

This tells us that for k large enough, knowing $\alpha \upharpoonright k$ suffices to compute $s(k)$ and then (by the above inequality) effectively compute an approximation of α by at most 2^{-k-c}. In other words, $\alpha \upharpoonright (k + c)$ can be computed from $\alpha \upharpoonright k$ and c. Therefore, for all k large enough:

$$K(\alpha \upharpoonright (k + c)) \leq K(\alpha \upharpoonright k, c) + O(1) \leq K(\alpha \upharpoonright k) + 2\log c + O(1)$$

The constant d is fixed, and c can be taken arbitrarily large. Choose c such that the expression $2\log c + O(1)$ in the above inequality is smaller than $c/2$. Then, for all k large enough,

$$K(\alpha \upharpoonright (k+c)) \leq K(\alpha \upharpoonright k) + c/2$$

An easy induction then shows that $K(\alpha \upharpoonright k) \leq O(k/2)$, contradicting the fact that α is random. ∎

An interesting corollary of this theorem is that there are nondecreasing Solovay functions (which is not really obvious from the definition). To see that it is the case, it suffices to take a computable sequence $(r_n)_{n\in\mathbb{N}}$ of rational numbers such that every r_n is a negative power of 2, the r_n are nonincreasing and $\sum_n r_n$ is a Martin-Löf random number (it is very easy to see that such sequences exist). Then, take $f(n) = -\log(r_n)$ for all n. The function f is computable, nondecreasing and by Theorem 2.5 is a Solovay function.

3. Solovay functions and Martin-Löf randomness

One of the most fundamental theorems of algorithmic randomness is the Levin-Schnorr theorem, proven independently by Levin and Schnorr in the 1970's. It characterizes Martin-Löf random sequences by the prefix-free Kolmogorov complexity of their initial segments. More precisely, a sequence $\alpha \in 2^\omega$ is Martin-Löf random if and only if

$$K(\alpha \upharpoonright n) \geq n - O(1)$$

This theorem left open a fundamental question: is there a similar characterization of Martin-Löf randomness in terms of plain Kolmogorov complexity?

3.1. The Miller-Yu theorem

This question remained open for almost three decades. It was finally answered positively in a recent paper of Miller and Yu [16].

Theorem 3.1 (Miller and Yu[1] [16]). *Let $\alpha \in 2^\omega$. The following are equivalent:*
(i) α is Martin-Löf random.
(ii) $C(\alpha \upharpoonright n) \geq n - K(n) - O(1)$.
(iii) For all functions $f \in \mathfrak{K}$, $C(\alpha \upharpoonright n) \geq n - f(n) - O(1)$.

Remarkably, Miller and Yu showed that in the item (iii) above, the "for all f" part can be replaced by a *single* function:

Theorem 3.2 (Miller and Yu [16]). *There exists a function $g \in \mathfrak{K}$ such that for all $\alpha \in 2^\omega$:*

$$\alpha \text{ is Martin-Löf random } \Leftrightarrow \ C(\alpha \upharpoonright n) \geq n - g(n) - O(1) \tag{3.1}$$

Informally, the function $g \in \mathfrak{K}$ in this last proposition is a "good" upper bound of K, in the sense that it is close enough to K to make possible the replacement of K by g in the equivalence $(i) \Leftrightarrow (ii)$ of Theorem 3.1. This reminds us of the Solovay functions which are also "good" upper bounds of K in their own way. And indeed, the function g constructed by Miller and Yu to make the equivalence (3.1) true is a Solovay function. We will show that this is not a coincidence, as *all* functions g satisfying (3.1) are Solovay functions. But before that, we state a related theorem:

[1] Gács [10] proved the equivalence $(i) \Leftrightarrow (ii)$ of Theorem 3.1

Theorem 3.3 (Bienvenu and Merkle [2])**.** *A sequence α is Martin-Löf random if and only if for all $f \in \mathfrak{K}$, $f(\alpha \upharpoonright n) \geq n - O(1)$. Moreover, there exists a unique function $g \in \mathfrak{K}$ such that*

$$\alpha \text{ is Martin-Löf random } \Leftrightarrow \ g(\alpha \upharpoonright n) \geq n - O(1) \tag{3.2}$$

We will prove:

Theorem 3.4. *Any function g satisfying the equivalence (3.1) of Theorem 3.2 is a Solovay function. The same is true for any function g satisfying the equivalence (3.2) of Theorem 3.3.*

In order to prove this theorem, we show that in both characterizations of Martin-Löf randomness ($K(\alpha \upharpoonright n) \geq n - O(1)$ and $C(\alpha \upharpoonright n) \geq n - K(n) - O(1)$) the lower bound on complexity is very sharp, that is there is no "gap phenomenon".

3.2. A "no-gap" theorem for randomness

Chaitin [4] proved an alternative characterization of Martin-Löf randomness: $\alpha \in 2^\omega$ is Martin-Löf random if and only if $K(\alpha \upharpoonright n) - n$ tends to infinity. Together with the Levin-Schnorr characterization, this shows a dichotomy: given a sequence $\alpha \in 2^\omega$, either α is not Martin-Löf random, in which case $K(\alpha \upharpoonright n) - n$ takes arbitrarily large negative values, or α is Martin-Löf random, in which case $K(\alpha \upharpoonright n) - n$ tends to $+\infty$. This means for example that there is no sequence $\alpha \in 2^\omega$ such that $K(\alpha \upharpoonright n) = n + O(1)$. One may ask whether this dichotomy is due to a gap phenomenon, that is: is there a function h that tends to infinity, such that for every Martin-Löf random sequence α, $K(\alpha \upharpoonright n) \geq n + h(n) - O(1)$? Similarly, is there a function h' that tends to infinity such that for every sequence α, $K(\alpha \upharpoonright n) \geq n - h'(n) - O(1)$ implies that α is Martin-Löf random? We answer both these questions (and their plain complexity counterpart) negatively.

Theorem 3.5. *There exists no function $h : \mathbb{N} \to \mathbb{N}$ (computable or not) which tends to infinity and such that*

$$K(\alpha \upharpoonright n) \geq n - h(n) - O(1)$$

is a sufficient condition for α to be Martin-Löf random (in fact, not even for α to be Church stochastic).

Similarly, there is no function $h : \mathbb{N} \to \mathbb{N}$ which tends to infinity and such that

$$C(\alpha \upharpoonright n) \geq n - K(n) - h(n) - O(1)$$

is a sufficient condition for α to be Martin-Löf random (in fact, not even for α to be Church stochastic).

Proof. First, notice that since we want to prove this for *any* function that tends to infinity, we can restrict our attention to the nondecreasing ones. Indeed, if h is a function that tends to infinity, the function

$$\tilde{h}(n) = \min\{f(k) \mid k \geq n\}$$

also tends to infinity and $\tilde{h} \leq h$.

Now, assume we are in the simple case where the function h is nondecreasing and computable. A standard technique to get a non-random binary sequence β such that $K(\beta \upharpoonright n) \geq n - h(n) - O(1)$ is the following: take a Martin-Löf random sequence α,

and insert zeroes into α in positions $h^{-1}(0), h^{-1}(1), h^{-1}(2), \ldots$. It is easy to see that the resulting sequence β is not Martin-Löf random (indeed, not even Church stochastic), and that the Kolmogorov complexity of its initial segments is as desired. This approach was refined by Merkle et al. [15] where the authors used an insertion of zeroes on a co-c.e. set of positions in order to construct a left-c.e. sequence β that is not Mises-Wald-Church stochastic, but has initial segments of very high complexity.

Of course, the problem here is that the function h in the hypothesis may be non-computable, and in particular may grow slower than any computable nondecreasing function. In that case, the direct construction we just described does not necessarily work: indeed, inserting zeroes at a noncomputable set of positions may not affect the complexity of α. To overcome this problem, we invoke the Kučera-Gács theorem (see Kučera [12], Gács [11], or Merkle and Mihailovic [14]). This theorem states that any subset of $\mathbb{N}$ (or function from $\mathbb{N}$ to $\mathbb{N}$) is Turing-reducible to a Martin-Löf random sequence. Hence, instead of choosing *any* Martin-Löf sequence α, we pick one that computes the function h^{-1} and then insert zeroes into α at positions $h^{-1}(0), h^{-1}(1), \ldots$. Intuitively, the resulting sequence β should not be random, as the bits of α can be used to compute the places where the zeroes have been inserted. This intuition however is not quite correct, as inserting the zeroes may destroy the Turing reduction Φ from α to h^{-1}. In other words, looking at β, we may not be able to distinguish the bits of α from the inserted zeroes.

The trick to solve this last problem is to delay the insertion of the zeroes to "give enough time" to the reduction Φ to compute the positions of the inserted zeroes. More precisely, we insert the k-th zero in position $n_k = h^{-1}(k) + t(k)$ where $t(k)$ is the time needed by Φ to compute $h^{-1}(k)$ from α. This way, n_k is computable from $\alpha \upharpoonright n_k$ in time at most n_k. From this, it is not too hard to construct a computable selection rule that selects precisely the inserted zeroes, witnessing that β is not Church stochastic (hence not Martin-Löf random). Moreover, since the "insertion delay" only makes the inserted zeroes more sparse, we have $K(\beta \upharpoonright n) \geq n - h(n) - O(1)$. And similarly, since α is Martin-Löf random, we have by the Miller-Yu theorem: $C(\alpha \upharpoonright n) \geq n - K(n) - h(n) - O(1)$.

The formal details are as follows. Let h be a nondecreasing function. By the Kučera-Gács theorem, let α be a Martin-Löf random sequence and Φ be a Turing functional such that $\Phi^\alpha(n) = h^{-1}(n)$ for all n. Let $t(n)$ be the computation time of $\Phi^\alpha(n)$ (we can assume that t is a nondecreasing function). Let $\beta \in 2^\omega$ be the sequence obtained by inserting zeroes into α in positions $h^{-1}(n) + t(n)$. To show that β is not Church stochastic, we construct a (total) computable selection rule that filters the inserted zeroes from β. Let S be the selection rule that works as follows on a given sequence $\xi \in 2^\omega$. We proceed by induction; we call k_n the number of bits selected by S from $\xi \upharpoonright n$ and x_n the prefix $\xi \upharpoonright n$ of ξ from which these k_n bits are deleted (x_0 is thus the empty string, and $k_0 = 0$).

At stage $n+1$, having already read $\xi \upharpoonright n$, S computes $\Phi_n^{x_n}(k_n)$. If the computation halts after s steps, S checks whether $\Phi_n^{x_n}(k_n) + s$ returns n. If so, S selects the n-th bit of $\xi(n)$ of ξ and then sets $x_{n+1} = x_n$ and $k_{n+1} = k_n + 1$. Otherwise, S just reads the bit $\xi(n)$, and sets $x_{n+1} = x_n\xi(n)$ and $k_{n+1} = k_n$.

It is clear that S is a total computable selection rule. Now suppose that we run it on β. We argue that S selects exactly the zeroes that have been inserted into α to get β. We prove

this by induction. If S has already selected from β the first i inserted zeroes, then the next selected bit is the bit in position $n = \Phi^{x_n}(k_n) + s$ where $\Phi^{x_n}(k_n)$ is computed in s steps. But since the selected bits are exactly the zeroes that were inserted in α, we have $k_n = i$ and $x_n = \alpha \upharpoonright n - i$, and thus s is the computation time of $\Phi^{x_n}(k_n) = \Phi^{\alpha\upharpoonright n-i}(i)$, which we called $t(i)$. And by definition of Φ, $\Phi^{\alpha\upharpoonright n-i}(i) = h^{-1}(i)$. Therefore, $n = h^{-1}(i) + t(i)$, i.e. the selected bit was an inserted zero. This proves that S only selects bits that belong to the zeroes that were inserted into α. Conversely, we need to prove that all such bits are indeed selected by S. Let $i \in \mathbb{N}$. The $i+1$-th inserted zero is in position $n = h^{-1}(i) + t(i)$. At stage n, we have by the induction hypothesis $x_n = \alpha \upharpoonright n - i$ and $k_n = i$. Thus, $\Phi_n^{x_n}(k_n) = \Phi_{h^{-1}(i)+t(i)}^{\alpha\upharpoonright t(i)+h^{-1}(i)-i}(i)$, which has to halt because both quantities $t(i) + h^{-1}(i) - i$ and $h^{-1}(i)+t(i)$ are greater than $t(i)$, which is the computation time of $\Phi^{\alpha}(i)$. Thus the bit in position n is indeed selected. Therefore, S satisfies the desired properties, and witnesses the fact that β is not Church stochastic.

Finally, for all n, calling i the number of inserted zeroes in $\beta \upharpoonright n$, we easily see that $\beta \upharpoonright n$ and $\alpha \upharpoonright n - i$ can each be computed from the other one (by insertion or deletion of zeroes). Thus: $K(\beta \upharpoonright n) = K(\alpha \upharpoonright n - i) \geq n - i$ (since α is Martin-Löf random). And by definition of the positions where the zeroes are inserted, we have $n \geq h^{-1}(i-1) + t(i-1)$, hence $i \leq h(n) + O(1)$. Therefore:

$$K(\beta \upharpoonright n) \geq n - i \geq n - h(n) + O(1)$$

for all n. This completes the proof. ∎

As a consequence of the construction performed in this proof, we get the dual version of Theorem 3.5:

Proposition 3.6. *There exists no function $h : \mathbb{N} \to \mathbb{N}$ (computable or not) which tends to infinity and such that*

$$K(\alpha \upharpoonright n) \geq n + h(n) - O(1)$$

is a necessary condition for α to be Martin-Löf random.

Similarly, there is no function $h : \mathbb{N} \to \mathbb{N}$ which tends to infinity and such that

$$C(\alpha \upharpoonright n) \geq n - K(h) + h(n) - O(1)$$

is a necessary condition for α to be Martin-Löf random.

Proof. Suppose for the sake of contradiction that there exists a function h' which tends to infinity and such that $K(\alpha \upharpoonright n) \geq n + h'(n) - O(1)$ is a necessary condition for α to be Martin-Löf random. Once again, we can assume that h' is non-decreasing. Then, we perform the exact same construction as in the proof of Theorem 3.5 for a given function h. Then, at the end of proof, when evaluating the complexity of β, we have $K(\beta \upharpoonright n) = K(\alpha \upharpoonright n - i) + O(1)$, with $i \leq h(n) + O(1)$, and since α is Martin-Löf random, $K(\alpha \upharpoonright n - i) \geq (n-i) + h'(n-i) - O(1)$. It follows that

$$K(\beta \upharpoonright n) \geq n - h(n) + h'(n - h(n)) - O(1)$$

Thus, if we take h to be sufficiently slow growing (for example $h(n) = \log(h'(n))$), we have $K(\beta \upharpoonright n) \geq n - O(1)$ for all n. This is a contradiction since by the Levin-Schnorr theorem, this would imply that β is Martin-Löf random, which it is not by construction. The proof of the second part of the proposition is almost identical. ∎

Theorem 3.4 now easily follows:

Proof (of Theorem 3.4). Let g be a function satisfying the equivalence (3.1) of Theorem 3.2. Suppose that g is not a Solovay function. This means, by definition, that $h(n) = g(n) - K(n)$ tends to infinity. Then, we can rewrite the equivalence (3.1) as:

$$\alpha \text{ is Martin-Löf random } \Leftrightarrow\ C(\alpha \upharpoonright n) \geq n - K(n) - h(n) - O(1)$$

which contradicts Theorem 3.5. Similarly, if a function g satisfies the condition (3.2) of Theorem 3.3, and is such that $h(n) = g(n) - K(n)$ tends to infinity, then for all $\alpha \in 2^\omega$, α is Martin-Löf random if and only if $K(\alpha \upharpoonright n) \geq n - h(n)$, contradicting Theorem 3.5. ∎

The consequences of Theorem 3.5 go beyond its applications to Solovay functions. For example, it gives an alternative proof of the fact that Schnorr randomness does not imply Church stochasticity (a result originaly proven by Wang [20]). Indeed, it is rather well-known that if h tends to infinity slower than any computable nondecreasing function, then the condition $K(\alpha \upharpoonright n) \geq n - h(n) - O(1)$ is sufficient for α to be Schnorr random (see for example Bienvenu and Merkle [2]), whereas we just saw that it was not sufficient for α to be Church stochastic.

One can also adapt the proof of Theorem 3.5 to separate Church stochasticity from Schnorr randomness within the left-c.e. reals. Informally, this is done as follows. Take a left-c.e. Martin-Löf random sequence $\alpha \in 2^\omega$. Call $t(n)$ the settling time of $\alpha \upharpoonright n$, i.e. given a computable nondecreasing sequence $(q_s)_{s\in\mathbb{N}}$ that converges to α, $t(n)$ is the smallest s such that $|\alpha - q_s| < 2^{-n}$. It is easy to see that t is enumerable from below. Thus, the sequence $\beta \in 2^\omega$ which we obtain from α by inserting zeroes in positions $t(0) < t(1) < \ldots$ is left-c.e. and for the same reason as above, is not Church stochastic. And the same kind of computation as above shows that $K(\beta \upharpoonright n) \geq n - t^{-1}(n) - O(1)$. Since it can easily be shown that t grows faster than any computable function, it follows that t^{-1} tends to infinity more slower than any nondecreasing unbounded computable function. Thus, β is not Church stochastic. This improves a result of Merkle et al. [15] (Theorem 26), who proved an equivalent fact for a weaker notion of stochasticity. For details on that result, see Bienvenu [1].

4. Solovay functions and triviality notions

A very successful line of research in algorithmic randomness over the last years concerns triviality and lowness notions. Informally, a sequence $\alpha \in 2^\omega$ is trivial if its Kolmogorov complexity is minimal or quasi-minimal, while a sequence α is low for randomness if it has little computation power, i.e. if relativizing the definition of random sequences to the oracle α does not change the class of random sequences. Perhaps the most important result in this direction was given by Nies [17]: a sequence $\alpha \in 2^\omega$ is low for Martin-Löf randomness (i.e. Martin-Löf randomness relativized to α coincides with standard Martin-Löf randomness) if and only if α is K-trivial (i.e. $K(\alpha \upharpoonright n) \leq K(n) + O(1)$). Other interesting notions of triviality have been studied, like Schnorr triviality, introduced by Downey et al. [6]: a sequence α is Schnorr trivial if for every prefix-free machine M whose domain has measure 1, there exists a machine M' whose domain also has measure 1, and such that $K_{M'}(\alpha \upharpoonright n) \leq K_M(n) + O(1)$. This notion was extensively studied by Franklin [8, 9].

In the same spirit, we can consider the class of sequences α such that for all computable upper bounds f of K, there exists a computable upper bound f' of K such that $f'(\alpha \upharpoonright n) \leq f(n) + O(1)$. However, because of the existence of Solovay functions, only computable sequences have this property.

Proposition 4.1. *Let $\alpha \in 2^\omega$. Suppose that for all $f \in \mathfrak{K}$, there exists $f' \in \mathfrak{K}$ such that*

$$f'(\alpha \upharpoonright n) \leq f(n) + O(1)$$

Then α is computable.

To prove this proposition, we need the following lemma:

Lemma 4.2 (Chaitin [3])**.** *For every $n, c \in \mathbb{N}$:*

$$\#\left\{w \in 2^{<\omega} \mid |w| = n \ \wedge \ K(w) \leq K(n) + c\right\} \leq 2^{c+O(1)}$$

where the $O(1)$ term does not depend on n or c.

Proof (of Proposition 4.1). Let $\alpha \in 2^\omega$ satisfy the hypothesis of the proposition. Let f be a Solovay function. By the assumption on α, there is a function $f' \in \mathfrak{K}$ and a constant c such that $f'(\alpha \upharpoonright n) \leq f(n) + c$ for all n. Let d be a constant such that $K \leq f' + d$. Since f is a Solovay function, there exists a constant e such that $f(n) \leq K(n) + e$ for infinitely many n. For any such n, we have:

$$\begin{aligned}
&\#\left\{w \in 2^{<\omega} \mid |w| = n \ \wedge \ f'(w) \leq f(n) + c\right\} \\
&\leq \#\left\{w \in 2^{<\omega} \mid |w| = n \ \wedge \ K(w) \leq K(n) + c + d + e\right\} \\
&\leq 2^{c+d+e+O(1)}
\end{aligned}$$

(the last inequality comes from Lemma 4.2). From this, we see that the Π^0_1 class

$$\{\xi \in 2^\omega \mid \forall n \ f'(\xi \upharpoonright n) \leq f(\xi \upharpoonright n) + c\}$$

to which α belongs, has only finitely many elements (at most $2^{c+d+e+O(1)}$), hence all these elements are computable. ∎

Another thing we can do is to study a weakened version of K-triviality: we consider the class of sequences α such that for any $f \in \mathfrak{K}$, $K(\alpha \upharpoonright n) \leq f(n) + O(1)$. As we shall now see, this is equivalent to K-triviality, hence we obtain an analogue of the Miller-Yu theorem for K-triviality.

Theorem 4.3. *Let $\alpha \in 2^\omega$. Then, α is K-trivial if and only if for all functions $f \in \mathfrak{K}$, $K(\alpha \upharpoonright n) \leq f(n) + O(1)$. Moreover, there exists a unique function $g \in \mathfrak{K}$ such that for all $\alpha \in 2^\omega$:*

$$\alpha \text{ is K-trivial} \Leftrightarrow K(\alpha \upharpoonright n) \leq g(n) + O(1) \tag{4.1}$$

Proof. By Lemma 2.1, it is obvious that any K-trivial α satisfies $K(\alpha \upharpoonright n) \leq f(n) + O(1)$ for all $f \in \mathfrak{K}$. Thus, all we have to do to prove this theorem is to construct a function g such that the implication "$\Leftarrow$" of equation (4.1) holds. In fact, we just take for g the function f constructed in the proof of Theorem 2.3.

Let then α be a sequence such that $K(\alpha \upharpoonright n) \leq g(n) + c$ for some constant c and all n. As usual, we prove that α is K-trivial by building a c.e. set L of pairs $(w_i, k_i)_{i \in \mathbb{N}}$ (with $w_i \in 2^{<\omega}$ and $k_i \in \mathbb{N}$) such that $\sum_i 2^{-k_i} < +\infty$ and for all n, some pair $(\alpha \upharpoonright n, K(n) + O(1))$ belongs

to L.

Let n be a fixed integer. We describe the strategy to enumerate strings of length n into L. We proceed by stages. At stage s, we look at the value of $K_s(n)$, and work under the assumption that $K_s(n) = K(n)$ (this assumption might turn out to be incorrect, we shall see below what to do when this happens). We then effectively find a $\mathbb{U}$-program p of length at most $K_s(n)$ such that $\mathbb{U}(p) = n$. By definition of g, if we call t the computation time of $\mathbb{U}(p)$, we have $g(\langle n, p, t\rangle) = |p| \leq K_s(n)$ (by definition of the function g), which, under the assumption $K_s(n) = K(n)$ implies $K(\langle n, p, t\rangle) = K(n)+O(1) = g(\langle n, p, t\rangle)+O(1)$. In other words, at every stage s, we can find a witness $m_s = \langle n, p, t\rangle$ such that $K(m_s) = g(m_s)+O(1)$, provided $K_s(n) = K(n)$.

Then, we enumerate all strings w of length m_s such that $K(w) \leq g(m_s) + c$, and for each such string we find, we put $(w \upharpoonright n, K_s(n))$ into L (without repetitions). Under the assumption $K_s(n) = K(n)$, we have $g(m_s) = K(m_s) + O(1)$ hence by Lemma 4.2, there are at most $d = 2^{c+O(1)}$ different strings w of length m_s such that $K(w) \leq g(m_s) + c$, hence at most d pairs of type $(w \upharpoonright n, K_s(n))$ enter L.

As we noted above, we might realize at some point that the assumption $K_s(n) = K(n)$ is incorrect, i.e. there might exist a stage $s' > s$ such that $K_{s'}(n) < K_s(n)$. In this case, we simply compute a new witness $m_{s'}$ and restart the strategy. However, the false assumption $g(m_s) = K(m_s) + O(1)$ may have caused us to enumerate many strings w of length m_s such that $K(w) \leq g(m_s) + c$ hence many pairs $(w \upharpoonright n, K_s(n))$ may enter L. We avoid this situation by only allowing d such pairs to enter L. Indeed, if more than d such pairs ask to enter L, we immediately know that the assumption $K_s(n) = K(n)$ is incorrect, hence we can stop acting and simply wait for a stage s' such that $K_{s'}(n) < K_s(n)$ and only then restart the strategy.

It remains to be verified that this strategy works, i.e. that the set L has the desired properties. For a fixed n, and any $k \geq K(n)$, by construction of L, there are at most d pairs of type $(w \upharpoonright n, k)$ in L. Thus, the total measure of the domain of L is at most

$$\sum_n \sum_{k \geq K(n)} d \cdot 2^{-k} = \sum_n d \cdot 2^{-K(n)+1} \leq 2d$$

hence is finite. Finally, for a given n, at some stage t we do have $K_t(n) = K(n)$. We then have $g(m_t) = K(m_t) + O(1)$ hence for *all* strings w of length m_t satisfying $K(w \upharpoonright m_t) \leq g(m_t) + c$, the pair $(wuhn, K_t(n))$ is enumerated into L (the restriction that at most d such pairs can enter L is not an actual restriction when $g(m_t) = K(m_t) + O(1)$). By definition of α, $K(\alpha \upharpoonright m_t) \leq g(m_t) + c$, hence by assumption $(\alpha \upharpoonright n, K_t(n)) = (\alpha \upharpoonright n, K(n))$ is enumerated into L. This completes the proof. ∎

We would like to end this paper with two questions.

Question 1. *Does* any *Solovay function g make the equivalence (3.1) of Miller-Yu's theorem true?*

Question 2. *Is any computable function g satisfying the equivalence (4.1) of Theorem 4.3 necessarily a Solovay function?*

Note that one cannot invoke a "no-gap" theorem to answer the second question, as it was noted by Csima and Montalbán [5] that there *is* a nondecreasing unbounded function h such that $K(\alpha \upharpoonright n) \leq K(n) + h(n) + O(1)$ implies that α is K-trivial.

Acknowledgement

The first author is grateful to Serge Grigorieff for very useful discussions on this work. The authors also thank the anonymous referees of this paper for helpful suggestions.

References

[1] Laurent Bienvenu. *Game-theoretic characterizations of randomness: unpredictability and stochasticity*. PhD thesis, Université de Provence, 2008.
[2] Laurent Bienvenu and Wolfgang Merkle. Reconciling data compression and Kolmogorov complexity. In *International Colloquium on Automata, Languages and Programming (ICALP 2007)*, volume 4596 of *Lecture Notes in Computer Science*, pages 643–654. Springer, 2007.
[3] Gregory Chaitin. Information-theoretical characterizations of recursive infinte strings. *Theoretical Computer Science*, 2:45–48, 1976.
[4] Gregory Chaitin. Incompleteness theorems for random reals. *Advances in Applied Mathematics*, 8:119–146, 1987.
[5] Barbara Csima and Antonio Montalbán. A minimal pair of K-degrees. *Proceedings of the American Mathematical Society*, 134:1499–1502, 2005.
[6] Rodney Downey, Evan Griffiths, and Geoffrey LaForte. On schnorr and computable randomness, martingales, and machines. *Mathematical Logic Quarterly*, 50(6):613–627, 2004.
[7] Rodney Downey and Denis Hirschfeldt. *Algorithmic randomness and complexity*. Springer, in preparation.
[8] Johanna Franklin. Hyperimmune-free degrees and Schnorr triviality. *Journal of Symbolic Logic*, 73:999–1008, 2008.
[9] Johanna Franklin. Schnorr trivial reals: a construction. *Archive for Mathematical Logic*, 46:665–678, 2008.
[10] Peter Gács. Exact expressions for some randomness tests. *Z. Math. Log. Grdl. M.*, 26:385–394, 1980.
[11] Peter Gács. Every set is reducible to a random one. *Information and Control*, 70:186–192, 1986.
[12] Antonin Kučera. Measure, Π^0_1 classes, and complete extensions of PA. *Lecture Notes in Mathematics*, 1141:245–259, 1985.
[13] Antonin Kučera and Ted Slaman. Randomness and recursive enumerability. *SIAM Journal on Computing*, 31:199–211, 2001.
[14] Wolfgang Merkle and Nenad Mihailovic. On the construction of effective random sets. *Journal of Symbolic Logic*, 69:862–878, 2004.
[15] Wolfgang Merkle, Joseph S. Miller, André Nies, Jan Reimann, and Frank Stephan. Kolmogorov-Loveland randomness and stochasticity. *Annals of Pure and Applied Logic*, 138(1-3):183–210, 2006.
[16] Joseph Miller and Liang Yu. On initial segment complexity and degrees of randomness. *Transaction of the American Mathematical Society*, 360(6):3193–3210, 2008.
[17] André Nies. Lowness properties and randomness. *Advances in Mathematics*, 197(1):274–305, 2005.
[18] André Nies. *Computability and randomness*. Oxford University Press, To appear.
[19] Robert Solovay. Draft of a paper (or series of papers) on Chaitin's work. Unpublished notes, 215 pages, 1975.
[20] Yongge Wang. A separation of two randomness concepts. *Information Processing Letters*, 69(3):115–118, 1999.

This work is licensed under the Creative Commons Attribution-NoDerivs License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/3.0/.