ON RELATIONS BETWEEN DIFFERENT ALGORITHMIC DEFINITIONS OF RANDOMNESS

UDC 517.11

A. KH. SHEN'

There exist different versions of algorithmic definition of an (individual) random object. We establish some connections between them. Namely, we show that (1) Kolmogorov-Loveland stochastic sequences can be nonrandom in the sense of Martin-Löf [1], [2] and that (2) the variant of Martin-Löf's definition called "Solovay randomness" in [3] is equivalent to the original Martin-Löf definition. The first statement is proved by the method of van Lambalgen [4].

1. Kolmogorov [5] stated that there exist Kolmogorov-Loveland stochastic sequences (also called Mises-Kolmogorov-Loveland random sequences) of zeros and one with logarithmic increase of entropy (= Kolmogorov complexity): a prefix of length $n$ had entropy $O(\log n)$. (We consider only uniform Bernoulli distribution on the set of all infinite sequences of zeros and ones, unless explicitly stated otherwise.) Such sequences cannot be Martin-Löf random, so Kolmogorov's statement implies that there are Kolmogorov-Loveland stochastic sequences, which are not Martin-Löf random. But the proof of Kolmogorov's assertion was never published and, as Andrei A. Muchnik has shown, it is false. Muchnik proved that each sequence $\omega$ whose prefixes (of length $n$) have entropy $\leq Cn + O(1)$, where $C$ is a constant, $C < 1$, cannot be Kolmogorov-Loveland stochastic. (There exist different versions of Kolmogorov complexity (= entropy), but for our purposes the differences between them are unessential—they differ only by $O(\log n)$; see [2] and [6].)

Nevertheless Kolmogorov-Loveland stochastic sequences which are not Martin-Löf random do exist: we give a proof of this fact by the method of van Lambalgen [4].

**Theorem 1.** There exists a Kolmogorov-Loveland stochastic sequence (with respect to the uniform Bernoulli distribution) which is not Martin-Löf random (with respect to the same distribution).

**Proof.** Following van Lambalgen, let us consider a sequence $p = p_0, p_1, \ldots$ of real numbers, which converges—but very slowly—to 1/2. Let us consider a distribution $\mu_p$ on the set of all infinite sequences of zeros and ones: the $i$th trial has $p_i$ as the probability of 1 and trials are independent. Now we consider a Martin-Löf random sequence with respect to the distribution $\mu_p$. It will be stochastic but not random with respect to the uniform distribution.

Now we give the proof in detail.

**Lemma 1.** Let $p = p_0, p_1, \ldots$ be a computable sequence of computable real numbers, and let $\sum (p_i - 1/2)^2 = +\infty$. Then no sequence can be random with respect both to $\mu_p$ and the uniform distribution.

**Proof.** This is a corollary of a more general result of Vovk [7]. We present his proof in our special case. We have two distributions: $\mu_p$ and the uniform Bernoulli
distribution \( \mu_{1/2} \). Let us consider one more distribution \( \mu' \): it is a Bernoulli distribution and the probabilities of ones will be arithmetic means between \( p_i \) and 1/2. Assume that a sequence \( \omega \) is random with respect to \( \mu_p \) and \( \mu_{1/2} \). Let \( u_i \) denote the probability of the \( i \)th member of \( \omega \) with respect to \( \mu_p \) (so \( u_i = p_i \) if the \( i \)th member of \( \omega \) is 1, and \( u_i = 1 - p_i \) otherwise). Let \( v_i \) and \( w_i \) denote the same probabilities with respect to \( \mu_{1/2} \) and \( \mu' \) (so \( v_1 = 1/2 \) and \( w_i = (u_i + v_i)/2 \)). The criteria of randomness in terms of an a priori probability (see [8]) gives that

\[
\prod_{i=1}^{n} w_i \leq C \prod_{i=1}^{n} u_i, \quad \prod_{i=1}^{n} w_i \leq C \prod_{i=1}^{n} v_i
\]

for some constant \( C \) and for each \( n \) (because \( \mu' \) does not exceed the a priori probability). Multiplying these two inequalities and taking logarithms, we get

\[
\sum \log((u_i + v_i)/2) \leq \sum \log u_i + \log v_i)/2 + O(1).
\]

The difference between the left and right sides is of order \((u_i - v_i)^2\) (convexity of the logarithm function), so \(\sum(u_i - v_i)^2 < \infty\), a contradiction. Lemma 1 is proved.

**Lemma 2.** Let \( p_0, p_1, \ldots \) be a computable sequence of computable real numbers that computably converges to 1/2. Then each sequence Martin-Löf random with respect to \( \mu_p \) is Kolmogorov-Loveland stochastic with respect to the uniform Bernoulli distribution.

**Proof.** Suppose that \( \omega \) is a Martin-Löf random sequence which is not stochastic and \( R \) is the corresponding rule of choice (\( R \) gives a sequence with no limit frequency of ones or with limit frequency of ones not equal to 1/2). Let \( \varepsilon > 0 \) be such that the frequency of ones in the \( R \)-choice subsequence is greater than 1/2 + \( \varepsilon \) infinitely many times. (If it is less than 1/2 - \( \varepsilon \) infinitely many times, the proof goes the same way). We shall reach a contradiction, showing that the set of all sequences for which the \( R \)-choice subsequence has infinitely many prefixes with frequency greater than 1/2 + \( \varepsilon \) has effective measure zero. By \( D_n \), we denote the set of all sequences which give after application of \( R \) a subsequence with at least \( n \) members and the frequency of ones in the first \( n \) members of it is greater than 1/2 + \( \varepsilon \). It is sufficient to show that \( \Sigma \mu(D_n) \) converges computably. (We mention that Theorem 2 below shows that it is sufficient to establish convergence.)

By \( \alpha_{n,k}(q_1, \ldots, q_n) \) we denote the probability of getting more than \( k \) ones in \( n \) independent trials, in the \( i \)th of which the probability of 1 is \( q_i \). The function \( \alpha_{n,k} \) is monotone, and \( \alpha_{n,k} \leq \alpha_{n,l} \) when \( k \leq l \). We claim that \( \mu_p(D_n) \leq \alpha_{n,k}(q_1, \ldots, q_n) \), where \( k = n(1/2 + \varepsilon) \) and

\[
q_1 = \sup_{i} p_i, \quad q_2 = \sup_{i \neq j} \min\{p_i, p_j\}, \ldots, q_i = \sup_{k_1 \neq \cdots \neq k_i} \min(p_{k_1}, \ldots, p_{k_i}).
\]

This implies the convergence of \( \Sigma \mu(D_n) \). Indeed, we can replace the \( q_i \) that are greater than 1/2 + \( \varepsilon/2 \) (let \( s \) be the number of \( q_i \), and the others by 1/2 + \( \varepsilon/2 \). Then

\[
\mu(D_n) \leq \alpha_{n,k}(1, \ldots, 1, 1/2 + \varepsilon/2, \ldots, 1/2 + \varepsilon/2) = \alpha_{n-s,k-s}(1/2 + \varepsilon/2, \ldots, 1/2 + \varepsilon/2).
\]

Now we can use the standard upper bound (obtained, for example, by Stirling’s formula) and see that \( \mu(D_n) \) decreases exponentially as \( n \) tends to infinity.

Now we prove that \( \mu(D_n) \leq \alpha_{n,k}(q_1, \ldots, q_n) \). Let us imagine (as in [9]) that the members of a sequence \( \omega \) are written on cards which lie on an (infinitely long) table (we do not see what is written on a card unless we turn it). The rule of choice is an algorithm that says which card must be turned next and whether it must be
turned only for information or it is selected into the subsequence. We consider also a protocol containing all information about this process (which cards were turned, what for, and what was written on them).

Let \( \pi \) be the initial segment of such a protocol (containing full information about some cards). We denote by \( n(\pi) \) the number of cards included in the subsequence turned during the initial segment \( \pi \), and by \( k(\pi) \) the number of ones on these cards. By \( q_s(\pi) \) we denote

\[
\sup_{k_1 \neq \ldots \neq k_t} \min(r_{k_1}, \ldots, r_{k_t})
\]

where the sequence \( r_1, r_2, \ldots \) is \( p_0, p_1, \ldots \) without members corresponding to the turned (for any purpose) cards during \( \pi \). By \( \mu_p(A|\pi) \) we denote the conditional probability of \( \omega \in A \) with respect to \( \mu_p \) if the protocol of application of \( R \) to \( \omega \) has a prefix \( \pi \). We prove that if \( n(\pi) \leq n \) and \( k = (1/2 + \varepsilon)n \), then

(1)

\[
\mu_p(D_n|\pi) \leq \alpha_{n-n(\pi),k-k(\pi)}(q_1(\pi), \ldots).
\]

(When \( \pi \) is empty, we get the original inequality.) If \( n(\pi) = n \), inequality (1) in fact becomes an equality (both sides are equal to either 0 or 1). Let \( n(\pi) < n \), and let \( t \) be the number of a card which must be turned immediately after \( \pi \). (If no such card exists, \( \mu(D_n|\pi) = 0 \).) Then

(2)

\[
\mu_p(D_n|\pi) = p_t \mu_p(D_n|\pi_1) + (1 - p_t)\mu_p(D_n|\pi_0)
\]

where \( \pi_1 \) and \( \pi_0 \) are protocols which contain (in addition to \( \pi \)) also information about 1 and 0 (respectively) on the \( t \)th card. Let us show that if inequality (1) holds for \( \pi_1 \) and \( \pi_0 \), it holds also for \( \pi \). The right-hand side of (2) does not exceed

(3)

\[
p_t \alpha_{n-n(\pi_1),k-k(\pi_1)}(q_1(\pi_1), \ldots) + (1 - p_t)\alpha_{n-n(\pi_0),k-k(\pi_0)}(q_1(\pi_0), \ldots)
\]

If the \( t \)th card was turned only for information, then \( n(\pi_0) = n(\pi_1) = n(\pi) \) and \( k(\pi_1) = k(\pi_0) = k(\pi) \), and we use the monotonicity of \( \alpha_{n,k} \) and the inequality \( q_i(\pi_1) = q_i(\pi_0) \leq q_i(\pi) \). If the \( t \)th card was included in the subsequence, \( n(\pi_1) = n(\pi_0) = n(\pi) + 1, k(\pi_1) = k(\pi) + 1, k(\pi_0) = k(\pi) \), and (3) is equal to

\[
\alpha_{n-n(\pi),k-k(\pi)}(p_t, q_1(\pi_0), q_2(\pi_0), \ldots)
\]

and does not exceed (by monotonicity)

\[
\alpha_{n-n(\pi),k-k(\pi)}(q_1(\pi), q_2(\pi), \ldots).
\]

So (1) is proved provided that all prefixes of protocols \( \pi \) with \( n(\pi) \leq n \) have limited length. If this is not so, the above considerations give us a bound for \( \mu_p(D_n^N|\pi) \), where \( D_n^N \) is the set of all sequences for which the rule \( R \) gives, after \( N \) or less turned cards, a subsequence with length \( \geq n \) and frequency of ones among the first \( N \) members greater than \( 1/2 + \varepsilon \). Then we let \( N \) tend to infinity and get (1). Lemma 2 is proved.

Now it remains to fix a computable sequence \( p_0, p_1, \ldots \) of computable real numbers, computably converging to \( 1/2 \) with \( \sum(p_i - 1/2)^2 = +\infty \). We take, for example, \( p_i = 1/2 + (i + 10)^{-1/2} \).

2. Chaitin [3] gives the following definition of a Solovay random sequence of zeros and ones with respect to a given computable distribution on the space of all sequences of zeros and ones. (Strictly speaking, Chaitin gives the definition of a random real number, but this makes no difference for our purposes.)

DEFINITION [3]. A sequence \( \omega \) is called a **Solovay random** with respect to a computable distribution \( \mu \) if for each computable sequence of effectively open sets \( U_i \) with \( \sum \mu(U_i) < \infty \) the sequence \( \omega \) belongs to \( U_i \) only for finitely many \( i \). (An effectively open set is a union of a computable sequence of intervals; an interval is a set of all infinite sequences with a given finite prefix.)
As pointed out in [3], any Solovay random sequence is Martin-Löf random. It is also pointed out there that if we require computable convergence of $\sum \mu(U_i)$ we get a definition equivalent to Martin-Löf’s.

The next theorem says that even without this restriction the Solovay and Martin-Löf definitions are equivalent. (This question is posed in [3].)

**Theorem 2.** Any Martin-Löf random sequence with respect to a computable distribution $\mu$ is Solovay random with respect to $\mu$.

**Proof.** Let $\omega$ be a Solovay nonrandom sequence with respect to $\mu$; $U_0, U_1, \ldots$ is a computable sequence of effectively open sets such that $S = \sum \mu(U_i)$ is finite and $\omega \in U_i$ for infinitely many $i$. We show that $\omega$ is Martin-Löf nonrandom, giving a method of constructing (for each $\varepsilon > 0$) an effectively open set with measure $< \varepsilon$.

Let $n$ be a positive integer, and consider the set $V_n$ of all infinite sequences which belong to at least $n$ different $U_i$. $V_n$ is an effectively open set: if a sequence belongs to $V_n$ this fact can be established by exhibiting $n$ different values of $i$ for which $\omega \in U_i$. It is also easy to see that $\mu(V_n) \leq S/n$. (If $u_i$ is equal to 1 on $U_i$ and to 0 outside, then $\sum u_i = S$ and $\sum u_i \geq n$ on $V_n$.) We suppose that $\omega \in V_n$ for all $n$. So to find an effectively open set containing $\omega$ with measure $< \varepsilon$ it is sufficient to choose $n$ such that $S/n < \varepsilon$ and then use $V_n$. ($S$ can be a noncomputable real, but it does not matter because we can use any upper bound for $S$ instead of $S$ itself.) Theorem 2 is proved.

The author is grateful to E. A. Asarin, N. K. Vereshchagin, V. G. Vovk, V. V. V'yugin, An. A. Muchnick, B. I. Penkov, A. L. Semenov, V. A. Uspensky, and all the participants of A. N. Kolmogorov’s seminar at the Moscow State University.

Institute of Problems of Information Transmission
Academy of Sciences of the USSR
Moscow

**BIBLIOGRAPHY**

8. V. V. V'yugin, Semiotika i Informatika, vyp. 16, VINITI, Moscow, 1981, pp. 14–43. (Russian)
9. A. Kh. Shen', Semiotika i Informatika, vyp. 18, VINITI, Moscow, 1982, pp. 14–42. (Russian)

Translated by the author