Journal of Universal Computer Science, vol. 1, no. 1 (1995), 48-66
submitted: 21/10/94, accepted: 3/1/95, appeared: 28/1/95©Springer Pub. Co.

What Is a Random String? [1] [2]

Cristian Calude
(Computer Science Department
The University of Auckland, New Zealand
email: cristian@cs.auckland.ac.nz)

Abstract: Chaitin's algorithmic definition of **random strings**—based on the complexity induced by self-delimiting computers—is critically discussed. One shows that Chaitin's model satisfy many natural requirements related to randomness, so it can be considered as an **adequate model** for finite random objects. It is a better model than the original (Kolmogorov) proposal. Finally, some open problems will be discussed.

Keywords: Blank-endmarker complexity, Chaitin (self-delimiting) complexity, random strings.
Category: G.3

1 Motivation

Suppose that persons A and B give us a sequence of 32 bits each, saying that they were obtained from independent coin flips. If A gives the string

$$u = 01001110100111101001101001110101$$

and B gives the string

$$v = 00000000000000000000000000000000,$$

then we would tend to believe A and would not believe B: the string u *seems* to be random, but the string v does not. Further on, if we change the value of a bit (say, from 1 to 0) in a (non) "random" string, then the result is still a (non) "random" string. If we keep making such changes in a "random" string, then we will eventually complete distroy randomness.

Laplace [21], pp.16-17 was, in a sense, aware of the above paradox, as it may be clear from the following phrase:

> *In the game of heads and tails, if head comes up a hundred times in a row then this appears to us extraordinary, because after dividing the nearly infinite number of combinations that can arise in a hundred throws into regular sequences, or those in which we observe a rule that is easy to grasp, and into irregular sequences, the latter are incomparably more numerous.*

[1] This paper is an expanded version of an invited lecture given to the International Symposium *The Foundational Debate*, Vienna, 15-17 September, 1994.

[2] This work has been partially supported by Auckland University Research Grant A18/XXXXX/62090/F3414022.

In other words: non random strings are strings possessing some kind of regularity, and since the number of all those strings (of a given length) is **small**, the occurrence of such a string is **extraordinary**.

Furthermore, regularity is a good basis for compression. Accordingly, randomness means the absence of any compression possibility; it corresponds to maximum information content (because after dropping any part of the string, there remains no possibility of recovering it). As we shall prove in Section 5, most strings have this property. In opposition, *most strings we deal with do not.*

The information content of a phrase in a natural language (English, for example) can be recovered even some letters (words) are omitted. The reason comes from the redundancy of most spoken languages. As a consequence, there exist many efficient programs to compress texts written in natural languages. It is important to emphasize that all these methods work very well on texts written in some natural language, but they do not work well on average, i.e. on *all* possible combinations of letters of the same length. Redundancy is also a very powerful handle to readers of mathematical books (and, in general, of scientific literature), and also to cryptanalysts (for example, Caesar's ciphers—just permutations of letters—can be broken by frequency analysis; see more on this topic in Salomaa [27]). A hypothetical language in which there are only strings with maximum information content gives no preference to strings (i.e. they have equal frequency); this makes the cipher impossible to break. However, such languages do not exist (and cannot be constructed, even with the help of the best computers, available now and in the future); redundancy is essential and inescapable in a spoken language (and to a large extent in most artificial languages; see Marcus [25]).

Before passing to some the formal treatment it is natural to ask the following question: Are there any random strings? Of course, we do not have yet the necessary tools to properly answer this question, but we may try to approach it informally. Let us call *canonical program* the smallest program generating a string. We claim that *every canonical program should be random*, independently if it generates or not a random output. Indeed, assume that x is a canonical program generating y. If x is not random, then there exists a program z generating x which is substantially smaller than x. Now, consider the program

from z calculate x, then from x calculate y.

This program is only a few letters longer than z, and thus it should be much shorter than x, which was supposed to be canonical. We have reached a contradiction.

Borel [1, 2] was the first author who systematically studied random sequences. The complexity-theoretic approach was independently initiated by Kolmogorov [22] and Chaitin [9]. For more historical facts see Chaitin [17] (A Life in Math), Uspensky [31], Li and Vitányi [23] and Calude [4].

2 Computers and Complexities

Denote by N the set of natural numbers; $N_+ = N \setminus \{0\}$. If S is a finite set, then $\#S$ denotes the cardinality of S. We shall use the following functions: i) $rem(m, i)$, the remainder of the integral division of m by i ($m, i \in N_+$), ii) $\lfloor \alpha \rfloor$, the integral part of the real α, iii) $\log_Q$, the base Q logarithm, $\log = \lfloor \log_2 \rfloor$.

Fix $A = \{a_1, \ldots, a_Q\}, Q \geq 2$, a finite alphabet. By A^* we denote the free monoid generated by A (under concatenation). The elements of A^* are called *strings*; λ is the empty string. For x in $A^*, |x|$ is the length of $x(|\lambda| = 0)$. For m in N, $A^m = \{x \in A^* \mid |x| = m\}$. For every $x \in A^*$ and natural n put $x^n = xx\ldots x$, (n times); $x^0 = \lambda$.

Every total ordering on A, say $a_1 < a_2 < \cdots < a_Q$, induces a quasi-lexicographical order on $A^* : \lambda < a_1 < \cdots < a_Q < a_1a_1 < \cdots < a_1a_Q < a_Qa_Q < \cdots < a_1a_1a_1 < \cdots$. We denote by $string(n)$ the nth string in A^* according to the quasi-lexicographical order. The induced order on each set A^m coincides with the lexicographical order.

Working with *partial recursive (p.r.) functions* $\varphi : A^* \times A^* \xrightarrow{o} A^*$ (called sometime *blank-endmarker computer*—see Chaitin [15]) we adopt the notations from Calude [3]. If $x \in dom(\varphi)$, that is x is in the domain of φ, then we write $\varphi(x) < \infty$. A *Chaitin computer* is a p.r. function $C : A^* \times A^* \xrightarrow{o} A^*$ with a *prefix-free domain* (i.e. for every string z, there is no pair of distinct strings x, y such that $U(x, z) < \infty$, $U(y, z) < \infty$, and x is a prefix of y). To a Chaitin computer C one associates the *self-delimiting complexity* or *Chaitin complexity*

$$H_C : A^* \xrightarrow{o} N,\ H_C(x/y) = \min\{|z| \mid z \in A^*, C(z, y^*) = x\},$$

with the convention $\min \emptyset = \infty$; here $y^* = \min\{w \in A^* \mid U(w, \lambda) = y\}$, the operator min being taken according to the quasi-lexicographical order.

The basic result obtained by Chaitin [9] (called the *Invariance Theorem*) states the existence of a Chaitin computer U (called *universal Chaitin computer*) such that for every Chaitin computer C there exists a constant c (depending upon U and C) such that

$$H_U(x/y) \leq H_C(x/y) + c,$$

for all $x, y \in A^*$.[3] The complexity induced by a blank-endmarker computer φ, K_φ is defined by $K_\varphi(x/y) = \min\{|z| \mid z \in A^*, \varphi(z, y) = x\}$. A similar Invariance Theorem holds true for blank-endmarker computers. See also Chaitin [9, 10], Kolmogorov [22], Martin-Löf [26], Calude [3].

For this paper we fix a universal Chaitin computer U and denote by H the induced complexity. Also, fix a universal blank-endmarker computer $\psi : A^* \times A^* \xrightarrow{o} A^*$ and denote by K the induced complexity. By $H(x)$, $K(x)$ we denote the complexities $H(x/\lambda)$, $K(x/\lambda)$, respectively.

Let $f, g, h : A^* \to [0, \infty)$ be three functions. We write $f \leq g + \mathrm{O}(h)$ in case there exists $C > 0$ such that $f(x) \leq g(x) + Ch(x)$, for almost all strings x. We write $f = g + \mathrm{O}(h)$ in case $f \leq g + \mathrm{O}(h)$ and $g \leq f + \mathrm{O}(h)$; $f \asymp g$ means that there exists two positive reals α, β such that

$f(x) \leq \alpha g(x)$ and $g(x) \leq \beta f(x)$, for almost all strings x.

3 Chaitin Random Strings

To motivate our approach we use the analogy between "tallness" and "randomness". To appreciate if a person is or is not tall we proceed as follows. We choose a unity measure (say, centimetre) and we evaluate the height. We get an *absolute*

[3] Exact values for all additive constants discussed in this paper have been recently computed by Chaitin [19]—using a Lisp model of computation.

value. Next, we establish "a set of people of reference". For instance, if we have to appreciate how tall is a little girl we fix an age and we relate her height to the average height of girls of that age. But, if we discuss the same question for a teenager, the situation is completely different. It follows that the adjective tall is *relative*. To correctly appreciate it we need both components: the exact one (height) and the relative one (comparison within a fixed set). It is fortunate that in English we have two words to express this: height and tall.

For randomness we proceed in a similar way, trying to capture, as best as possible, the idea that *a string is random if it cannot be algorithmically compressed*. First we use a measure of complexity for strings (H); this represents the "absolute component". Secondly, we define randomness "relative to a set"—the relative component. In our case we appreciate the degree of randomness of a string with respect to the set of all strings, over a fixed alphabet, having the same length.[4]

Of course, the success or failure of the approach depends upon the measure of complexity we are adopting. The use of self-delimiting programs instead of blank-endmarker programs is motivated by Chaitin [16] as follows:

> *A key point that must be stipulated ... is that an input program must be self-delimited: its total length (in bits) must be given within the program itself. (This seemingly minor point, which paralyzed progress in the field for nearly a decade, is what entailed the redefinition of algorithmic randomness.) Real programming languages are self-delimiting, because they provide constructs for beginning and ending a program. Such constructs allow a program to contain well-defined subprograms nested in them. Because a self-delimiting program is built up by concatenation and nesting self-delimiting subprograms, a program is syntactically complete only when the last open subprogram is closed. In essence the beginning and ending constructs for programs and subprograms function respectively like left and right parentheses in mathematical expressions.*

We first recall the asymptotical behaviour of the complexity H (see Chaitin [12, 14]):

Theorem 3.1. *Let $f : N \to A^*$ be an injective, recursive function. a) One has:*

$$\sum_{n \geq 0} Q^{-H(f(n))} \leq 1.$$

b) Consider a recursive function $g : N_+ \to N_+$.[5]

i) If $\sum_{n\geq 1} Q^{-g(n)} = \infty$, then $H(f(n)) > g(n)$, for infinitely many $n \in N_+$.
ii) If $\sum_{n\geq 1} Q^{-g(n)} < \infty$, then $H(f(n)) \leq g(n) + O(1)$.

Proof. a) It is plain that:

$$\sum_{n \geq 0} Q^{-H(f(n))} \leq \sum_{x \in A^*} Q^{-H(x)} \leq \sum_{x \in A^*} P(x) \leq 1.$$

[4] So, the "context" is determined by the length and the size of the alphabet.
[5] Actually, this part is still true in case g is a function semi-computable in the limit from above.

b) i) Assume first that $\sum_{n\geq 1} Q^{-g(n)} = \infty$. If there exists a natural N such that

$$H(f(n)) \leq g(n), \text{ for all } n \geq N,$$

then we get a contradiction:

$$\infty = \sum_{n\geq N} Q^{-g(n)} \leq \sum_{n\geq N} Q^{-H(f(n))} \leq \sum_{n\geq 0} Q^{-H(f(n))} \leq 1.$$

In view of the hypothesis in b) ii), there exists a natural N such that $\sum_{n\geq N} Q^{-g(n)} \leq 1$. We can use Kraft-Chaitin Theorem in order to construct a Chaitin computer $C : A^* \times A^* \xrightarrow{o} A^*$ with the following property: For every $n \geq N$ there exists $x \in A^*$ with $|x| = g(n)$ and $C(x, \lambda) = f(n)$. So, there exists a natural c such that for all $n \geq N$,

$$H(f(n)) \leq H_C(f(n)) + c \leq g(n) + c.$$

□

Example 3.2. $\sum_{n\geq 0} Q^{-H(string(n))} \leq 1$.

Example 3.3. *i) Take $g(n) = \lfloor \log_Q n \rfloor$. It is seen that $\sum_{n\geq 1} Q^{-g(n)} = \infty$, so $H(string(n)) > \lfloor \log_Q n \rfloor$, for infinitely many $n \geq 1$.*
ii) For $g(n) = 2\lfloor \log_Q n \rfloor$, one has:

$$\sum_{n\geq 1} Q^{-g(n)} \leq Q \sum_{n\geq 1} \frac{1}{n^2} < \infty,$$

so $H(string(n)) \leq 2\lfloor \log_Q n \rfloor + O(1)$. For $Q > 2$ and $g(n) = \lfloor \log_{Q-1} n \rfloor$, one has:

$$\sum_{n\geq 1} Q^{-g(n)} \leq Q \sum_{n\geq 1} \frac{1}{n^{\log_{Q-1} Q}} < \infty,$$

so $H(string(n)) \leq \lfloor \log_{Q-1} n \rfloor + O(1)$.

Remark. Chaitin complexity H can be characterized as a minimal function, semi-computable in the limit from above, that lies on the borderline between the convergence and the divergence of the series

$$\sum_{n\geq 0} Q^{-H(string(n))}.$$

Put $H_C(x, y) = H_C(< x, y >)$, where $<,>: A^* \times A^* \to A^*$ is a recursive bijection.

The following three formulae come from Chaitin [12, 13, 14]:

Lemma 3.4. *There exists a natural c such that for all string s $x, y \in A^*$ one has*

$$H(x) \leq H(x, y) + c,$$
$$H(x, y) \leq H(x) + H(y/x) + c,$$
$$H(x, y) \leq H(x) + H(y) + c.$$

We are in the position to evaluate the complexity of the most complex strings of a given length (first obtained in Chaitin [12]).

Theorem 3.5. *For every $n \in N$, one has:*

$$\max_{x \in A^n} H(x) = n + H(string(n)) + O(1).$$

Proof. In view of Lemma 3.4, for every string x of length n,

$$H(x) \leq H(string(n), x) + \mathrm{O}(1) \leq H(string(n)) \\ + H(x/string(n)) + \mathrm{O}(1).$$

To get the relation

$$\max_{x \in A^n} H(x) \leq n + H(string(n)) + \mathrm{O}(1)$$

we shall prove that for every string x of length n,

$$H(x/string(n)) \leq n + \mathrm{O}(1).$$

Fix $n \geq 0$ and define the Chaitin computer $C_n : A^n \times A^* \xrightarrow{o} A^*$ by

$$C_n(x, y) = x \text{ if } U(y, \lambda) \neq \infty.$$

Accordingly, $U((string(n))^*, \lambda) = string(n)$ and

$$\begin{aligned} H(x/string(n)) &\leq H_{C_n}(x/string(n)) + \mathrm{O}(1) \\ &= \min\{|z| \mid z \in A^*, C_n(z, (string(n))^*) = x\} + \mathrm{O}(1) \\ &\leq n + \mathrm{O}(1). \end{aligned}$$

To prove the converse relation we need the following
Intermediate Step. For every $n \geq 0$,

$$\#\{x \in A^n \mid H(x) < n + H(string(n)) - t + O(1)\} < Q^{n-t+O(1)}.$$

By Lemma ?? one has:

$$H(x) < n + H(string(n)) - t + \mathrm{O}(1) \Longleftrightarrow H(x/string(n)) < n - t + \mathrm{O}(1),$$

so

$$\#\{x \in A^n \mid H(x) < n + H(string(n)) - t + \mathrm{O}(1)\} = \\ \#\{x \in A^n \mid H(x/string(n)) < n - t + \mathrm{O}(1)\} < Q^{n-t+\mathrm{O}(1)}.$$

Accordingly, not all strings of length n have the complexity less than $n + H(string(n)) + \mathrm{O}(1)$, i.e. $\max_{x \in A^n} H(x) \geq n + H(string(n)) + \mathrm{O}(1)$. □

The above discussion may be concluded with the following definition. Let $\Sigma : N \to N$ be the function defined by

$$\Sigma(n) = \max_{x \in A^n} H(x).$$

In view of Theorem 3.5, $\Sigma(n) = n + H(string(n)) + \mathrm{O}(1)$. We define the random strings of length n to be the strings with maximal self-delimiting complexity among the strings of length n, i.e. the strings $x \in A^n$ having $H(x) \approx \Sigma(n)$.

Definition 3.6. *A string $x \in A^*$ is **Chaitin m-random** (m is a natural number) if $H(x) \geq \Sigma(|x|) - m$; x is **Chaitin random** if it is 0-random.*

The above definition depends upon the fixed universal computer U; the generality of the approach comes from the Invariance Theorem.

Obviously, for every length n and for every $m \geq 0$ there exists a Chaitin m-random string x of length n. Denote by $RAND_m^C, RAND^C$, respectively, the sets of Chaitin m-random strings and random strings.

It is worth to note that the property of Chaitin m-randomness is asymptotic. Indeed, for $x \in RAND_m^C$, the larger is the difference between $|x|$ and m, the more random is x. There is no sharp dividing line between randomness and pattern, but it looks as though all $x \in RAND_m^C$ with $m \leq H(string(|x|))$ have a true random behaviour.

How many strings $x \in A^n$ have maximal complexity, i.e. $H(x) = \Sigma(|x|)$? The answer was given by Chaitin [18]:

Theorem 3.7. *There exists a natural constant $c > 0$ (which depends upon the size of the underlying alphabet, Q) such that*

$$\gamma(n) = \#\{x \in A^n \mid H(x) = \Sigma(|x|)\} > Q^{n-c},$$

for all natural n.

How large is c? Out of Q^n strings of length n, at most $Q + Q^2 + \cdots + Q^{n-m-1} = (Q^{n-m} - 1)/(Q - 1)$ can be described by programs of length less than $n - m$. The ratio between $(Q^{n-m} - 1)/(Q - 1)$ and Q^n is less than 10^{-i} as $Q^m \geq 10^i$, irrespective of the value of n. For instance, this happens in case $Q = 2, m = 20, i = 6$; it says that *less than one in a million among the binary strings of any given length is not Chaitin 20-random.*

So, in a strictly quantitative sense, almost all strings are Chaitin random.

Problem. Denote by $(c_Q)_{Q\geq 2}$ the sequence of constants appearing in Theorem 3.7. Is this sequence bounded?

The rest of this paper will be devoted to the analysis of the adequacy of Chaitin's definition of randomness.

4 A Statistical Analysis Random Strings

In this section we confront Chaitin's definition of randomness with the probability point of view. As we have already said, the present proposal identifies *randomness* with *incompressibility*. In order to justify this option we have to show that the strings that are incompressible justify the various properties of stochasticity identified by the classical Probability Theory. It is not so difficult, although tedious, to check *separately* such a single property. However, we may proceed in a better way, due to the celebrated theory developed by Martin-Löf: We demonstrate that the incompressible strings do possess all conceivable effectively testable properties of stochasticity. Here we include the known properties, but also the possible unknown ones. A general transfer principle will emerge, by virtue of which various results from classical probability theory carry automatically for random strings.

The ideas of Martin-Löf's theory are rooted in the statistical practice. We are given an element x of some sample space (associated to some distribution) and we want to test the hypothesis *x is a typical outcome.* Being typical means "belonging to every reasonable majority". An element x will be "random" just in case x lies in the intersection of all such majorities.

A level of a statistical test is a set of strings which are found relatively non-random (by the test). Each level is a subset of the previous level, containing less and less strings, considered more and more non-random. The number of strings decreases exponentially fast at each level. In the binary case, a test contains at level 0 all possible strings, at level two only at most 1/2 of the strings, at level three only 1/4 of all strings, and so on; accordingly, at level m the test contains at most 2^{n-m} strings of length n.

We give now the formal definition.

Definition 4.1. *An r.e. set $V \subset A^* \times N_+$ is called a* ***Martin-Löf test*** *if the following two properties hold true:*
1) $V_{m+1} \subset V_m$, for all $m \geq 1$ (here $V_m = \{x \in A^* \mid (x, m) \in V\}$ is the m-section of V),
2) $\#(A^n \cap V_m) < Q^{n-m}/(Q-1)$, for all $n \geq m \geq 1$.

By definition, the empty set is a Martin-Löf test.

The set V_m is called the *critical region at level* $Q^{-m}/(Q-1)$. (Getting an outcome string x in V_m means the rejection of the randomness hypothesis

for x.) A string x is declared "random" at level m by V in case $x \notin V_m$ and $|x| > m$.

The following example models the following simple idea: If a binary string x has too many ones (zeros), then it cannot be random.

Example 4.2. *The set*

$$V = \{(x, m) \in A^* \times N_+ \mid | \frac{N_i(x)}{|x|} - \frac{1}{Q} | > Q^m \frac{1}{\sqrt{|x|}}\},$$

where $N_i(x)$ is the number of occurrences of the letter a_i in x, is a Martin -Löf test.

Proof. Clearly, V is r.e. and satisfies condition 1). In view of the formula:

$$\#\{x \in A^n \mid | \frac{N_i(x)}{|x|} - \frac{1}{Q} | > \varepsilon\} \leq \frac{Q^{n-2}(Q-1)}{n\varepsilon^2},$$

one gets

$$\#(A^n \cap V_m) = \#\{x \in A^n \mid | \frac{N_i(x)}{|x|} - \frac{1}{Q} | > Q^m \frac{1}{\sqrt{|x|}}\}$$

$$= \frac{Q^{n-2}(Q-1)}{Q^{2m}} = Q^{n-2-2m}(Q-1) \leq \frac{Q^{n-m}}{Q-1}.$$

□

Definition 4.3. *To every Martin-Löf test V we associate the* ***critical level*** $m_V : A^* \to N$,

$$m_V(x) = \begin{cases} \max\{m \geq 1 \mid (x,m) \in V\}, & \textit{if } (x,1) \in V, \\ 0, & \textit{otherwise.} \end{cases}$$

A string x is declared ***random*** *by a Martin-Löf test V if $x \notin V_q$, for some $q < |x|$, or, equivalently, if $m_V(x) < |x| - 1$.*

Definition 4.4. *A Martin-Löf test $\mathcal{U}$ is called* ***universal*** *in case for every Martin-Löf test V, there exists a constant c (depending upon $\mathcal{U}$ and V) such that*

$$V_{m+c} \subset \mathcal{U}_m, m = 1, 2, \ldots$$

It is easy to see that a *Martin -Löf test $\mathcal{U}$ is universal iff for every Martin-Löf test V there exists a constant c (depending upon $\mathcal{U}$ and V) such that*

$$m_V(x) \leq m_{\mathcal{U}}(x) + c, \textit{ for all } x \in A^*.$$

Our aim is to prove that almost all Chaitin random strings pass all conceivable effective tests of stochasticity, i.e. they are declared random by every Martin-Löf test.

Theorem 4.5. *Fix $t \in N$. Almost all strings in $RAND_t^C$ will be declared eventually* random *by every Martin-Löf test.*

Proof. If K is the complexity induced by a fixed universal blank-endmarker computer, then by a result in Calude, Chitescu [7], the set $\mathcal{U} = \{(x,m) \in A^* \times N_+ \mid K(x) < |x| - m\}$ is a universal Martin-Löf test. Every section $\mathcal{U}_t$ is r.e., so when a string x belongs to $\mathcal{U}_t$, then
we can eventually discover this. Moreover, there are less than

$$Q^{n-t}/(Q-1)$$

strings of length n having this property. Thus if we are given $|x| = n$ and t we need to only know $n - t$ digits to pick out any particular string $x \in A^n$ with this property. I.e., as the first x that we discover has this property, the second x that we discover has this property,...,the ith x that we discover has this property, and $i < Q^{n-t}/(Q-1)$. Accordingly, every string $x \in A^n \cap \mathcal{U}_t$ has the property that

$$H(x/ < string(n), string(t) >) < n - t + \mathrm{O}(1).$$

So, by Lemma 3.4:

$$\begin{aligned} H(x) &< H(x/ < string(n), string(t) >) + H(< string(n), \\ &\quad string(t) >) + \mathrm{O}(1) \\ &< n - t + H(string(n), string(t)) + \mathrm{O}(1) \\ &< n - t + H(string(n)) + H(string(t) + \mathrm{O}(1) \\ &< n - t + H(string(n)) + \mathrm{O}(\log_Q t), \end{aligned}$$

since in general $H(string(m)) < \mathrm{O}(\log_Q m)$.

Finally, take $x \in RAND_t^C$, i.e. $H(x) \geq \Sigma(|x|) - t$, and fix an arbitrary Martin-Löf test V. Take a natural T such that $T - \mathrm{O}(\log_Q T) \geq t$. It follows that $x \notin \mathcal{U}_T$—otherwise we would have

$$H(x) < \Sigma(|x|) - T - \mathrm{O}(\log_Q T),$$

which means $T - \mathrm{O}(\log_Q T) < t$. By virtue of the universality of $\mathcal{U}$ we get a constant c such that $V_{m+c} \subset \mathcal{U}_m$, $m = 1, 2, \ldots$, i.e. $x \notin V_{c+T}$. The constant $c+T$ is *independent* of x. So, for large enough random strings x, one has $x \notin V_m$ and $|x| > m$. □

5 A Computational Analysis Random Strings

We pursue the analysis of the relevance of Chaitin's definition by confronting it with a natural, computational requirement: *there should be no algorithmic way to recognize what strings are random.*

The following result, due to Chaitin [11], Theorem 4.1, was used for deriving the first information-theoretic incompleteness results (see also Chaitin [17]). We use it now to discuss the uncomputability of $RAND_t^C$.

Let $<>$: $A^* \times N \to A^*$ be a recursive bijective function.

Theorem 5.1. *We can effectively compute a constant d such that if*

$$W_e \subset \{< w, m > \in A^* \times N \mid H(w) > m\},$$

then $n \leq H(e) + d$, for all $< u, n > \in W_e$.

Recall that a subset $X \subset A^*$ is immune iff it is infinite and has no infinite r.e. subsets.

Corollary 5.2. *Let $g : N \to N$ be a recursive function such that the set $S = \{w \in A^* | H(w) > g(|w|)\}$ is infinite. Then, the set S is immune.*

Proof. Let $W_e \subset \{w \in A^* | H(w) > g(|w|)\}$. Put $V_e = \{< w, g(|w|) > | w \in W_e\}$; clearly, V_e is r.e. and $V_e = W_{f(e)}$, for some recursive function $f : A^* \to A^*$. So,

$$V_e = W_{f(e)} \subset \{< w, m > | H(w) > m\}$$

and in view of Theorem 5.1 from $< w, g(|w|) > \in V_e = W_{f(e)}$ we deduce $g(|w|) \leq H(f(e)) + d$, i.e., V_e is finite. This shows that W_e itself is finite. □

Corollary 5.3. *The set $RAND_t^C$ is immune for every $t \geq 0$.*

Proof. Fix $t \in N$. It is plain that

$$RAND_t^C \subset \{x \in A^* | H(x) \geq |x| - t\}.$$

The set $\{x \in A^* | H(x) \geq |x| - t\}$ is immune by Corollary 5.2 and every infinite subset of an immune set is itself immune. □

The above theorem can be expressed as:

$$(\forall B \subset A^*)(B \text{ infinite and r.e. } \Rightarrow B \setminus RAND_t^C \neq \emptyset).$$

There are two (classically equivalent) ways to represent the above statement:

1. $(\forall x \in A^*)$ $(W_x$ infinite $\Rightarrow \exists y \in A^* : y \in W_x \to etminusRAND_t^C)$,
2. $\forall x \in A^* : (W_x \subset RAND_t^C \Rightarrow (\exists n \in N)\,\#(W_x) \leq n)$.

Based on theses statements we can formulate two constructive versions of immunity:

> The set $R \subset A^*$ is called *constructively immune* (Li [24]) if there exists a p.r. function $\varphi : A^* \xrightarrow{o} A^*$ such that for all $x \in A^*$, if W_x is infinite, then $\varphi(x) \neq \infty$ and $\varphi(x) \in W_x \setminus R$.
>
> The set $R \subset A^*$ is called *effectively immune* (Smullyan [30]) if there exists a p.r. function $\sigma : A^* \xrightarrow{o} N$ such that for all $x \in A^*$, if $W_x \subset R$, then $\sigma(x) \neq \infty$ and $\#(W_x) \leq \sigma(x)$.

It is worth noticing that there exist constructively immune sets which are not effectively immune and vice-ve rsa. Moreover, if the complement of an immune set is r.e., then that set is constructively immune. Hence, we get:

Theorem 5.4. *For every $t \geq 0$, $RAND_t^C$ is constructively immune.*

Proof. Fix $t \in N$. The set $\{x \in A^* | H(x) \geq |x| - t\}$ is constructively immune since its complement is r.e. As $RAND_t^C$ is an infinite subset of a constructive immune set it follows that itself is constructive immune. □

As a scholium of Corollary 5.2 one obtains:

Scholium 5.5. *If $g : N \to N$ is a recursive function with $\lim_{n\to\infty} g(n) = \infty$, recursively (i.e. there exists an increasing recursive function $r : N \to N$, such that if $n \geq r(k)$, then $g(n) \geq k$) and the set $S = \{w \in A^* \mid H(w) > g(|w|)\}$ is infinite, then S is effectively immune.*

Proof. In the context of the proof of
Corollary 5.2, $\#W_e = \#W_{f(e)}$ and if $w \in W_e \subset \{u \in A^* \mid H(u) > g(|u|)\}$, then

$$g(|w|) < H(w) \leq H(f(e)) + d \leq |f(e)| + 2\log|f(e)| + d + c + c'.$$

If $|w| \geq r(|f(e)|+2\log|f(e)|+d+c+c')$, then $g(|w|) > |f(e)|+2\log|f(e)|+d+c+c'$, so $w \notin W_e$. Accordingly, if $w \in W_e$, then $|w| < r(|f(e)|+2\log|f(e)|+d+c+c')$, i.e.

$$\#W_e \leq (Q^{r(|f(e)|+2\log|f(e)|+d+c+c')} - 1)/(Q-1),$$

and the upper bound is a recursive function of e. □

Corollary 5.6. *For all $t \geq 0$, $RAND_t^C$ is effectively immune.*

Proof. An infinite subset of an effectively immune set is effectively immune. □

6 Random Strings Are Borel Normal

Another important restriction pertaining a good definition of randomness concerns the frequency of letters and blocks of letters. In a "true random" string each letter has to appear with approximately the same frequency, namely Q^{-1}. Moreover, the same property should extend to "reasonably long" substrings.

These ideas have been stated by Borel [1, 2] for sequences. In Chaitin [10] one shows that Chaitin Omega Number representing the halting probability of a universal self-delimiting computer is Borel normal.

Motivated by these facts we formalize the Borel normality property for strings. First, let $N_i(x)$ be the number of occurrences of the letter a_i in the string x, $1 \leq i \leq Q$. Accordingly, the ratio $N_i(x)/|x|$ is the relative frequency of the letter a_i in the string x.

For strings of length $m \geq 1$ we proceed as follows. We consider the alphabet $B = A^m$ and construct the free monoid $B^* = (A^m)^*$. Every $x \in B^*$ belongs to A^*, but the converse is false. For $x \in B^*$ we denote by $|x|_m$ the length of x (according to B) which is exactly $|x|m^{-1}$.

For every $1 \leq i \leq Q^m$ denote by N_i^m the number of occurrences of y_i in the string $x \in B^*, B = \{y_1, \ldots, y_{Q^m}\}$. For example, take $A = \{0,1\}, m = 2, B = A^2 = \{00, 01, 10, 11\} = \{y_1, y_2, y_3, y_4\}, x = y_1y_3y_3y_4y_3 \in B^*(x = 0010101110 \in A^*)$. It is easy to see that $|x|_2 = 5, |x| = 10, N_1^2(x) = 1, N_2^2(x) = 0, N_3^2(x) = 3, N_4^2(x) = 1$. Note that the string $y_2 = 01$ appears three times into x, but not on the right positions.

Not every string $x \in A^*$ belongs to B^*. However, there is a possibility "to approximate" such a string by a string in B^*. We proceed as follows. For $x \in A^*$ and $1 \leq j \leq |x|$ we denote by $[x; j]$ the prefix of x of length $|x| - rem(|x|, j)$ (i.e. $[x; j]$ is the longest prefix of x whose length is divisible by j). Clearly, $[x; 1] = x$ and $[x; j] \in (A^j)^*$. We are now in a position to extend the functions N_i^m from B^* to A^*: put $N_i^m(x) = N_i^m([x; m])$, in case $|x|$ is not divisible by m. Similarly, $|x|_m = |[x; m]|_m$.

Definition 6.1. *A non-empty string $x \in A^*$ is called **ε-limiting** (ε is a fixed positive real) if for all $1 \leq i \leq Q, x$ satisfies the inequality:*

$$\left| \frac{N_i(x)}{|x|} - Q^{-1} \right| \leq \varepsilon.$$

Definition 6.2. *A string $x \in A^*$ is called **Borel normal** iff for every natural $m, 1 \leq m \leq \log_Q \log_Q |x|$,*

$$\left| \frac{N_j^m(x)}{|x|_m} - Q^{-m} \right| \leq \sqrt{\frac{\log_Q |x|}{|x|}},$$

for every $1 \leq j \leq Q^m$.

In Calude [5] one proves the following result:

Theorem 6.3. *For every natural $t \geq 0$ we can effectively compute a natural number M_t (depending upon t) such that every string of length greater than M_t in $RAND_t^C$ is Borel normal.*

Theorem 6.3 can be used to prove the following result (a weaker version was obtained in Calude, Câmpeanu [6]):

Theorem 6.4. *For every natural t and for every string x we can find two strings u, v such that $uxv \in RAND_t^C$.*

Proof. Fix $t \in N, x \in A^i, i \geq 1$. Almost all strings $z \in RAND_t^C$ are Borel normal by Theorem 6.3, i.e. they satisfy the inequality

$$\left| \frac{N_j^m(z)}{\lfloor n/m \rfloor} - Q^{-m} \right| \leq \sqrt{\frac{\log_Q n}{n}},$$

for every $1 \leq j \leq Q^m, 1 \leq m \leq \log_Q \log_Q n; n = |z|$.
Take $m = i$, x to be the jth string of length i and pick a string $z \in RAND_t^C$ such that $n = |z| = Q^{Q^{2i+1}}$. It follows that

$$\left| \frac{N_j^i(z)}{\lfloor n/i \rfloor} - Q^{-m} \right| \leq \sqrt{\frac{\log_Q n}{n}},$$

in particular,

$$Q^{-i} - \sqrt{\frac{\log_Q n}{n}} \leq \frac{N_j^i(z)}{\lfloor n/i \rfloor}.$$

To prove that $N_j^i(z) > 0$ it is enough to show that

$$Q^{-i} > \sqrt{\frac{\log_Q n}{n}},$$

which is true because

$$\sqrt{\frac{\log_Q n}{n}} = Q^{\frac{2i+1}{2} - \frac{1}{2}Q^{2i+1}},$$

and

$$\frac{1}{2}Q^{2i+1} \geq \frac{1}{2}2^{2i+1} = 4^i > 2i + \frac{1}{2}.$$

□

7 Extensions of Random Strings

In this section we deal with the following problem: To what extent is it possible to extend an arbitrary string to a Chaitin random or no n-random string ?

Theorem 6.4 says that every string x can be embedded into a Chaitin random string. The next results will put some more light on this phenomenon.

Theorem 7.1. *For every natural t and every string $x \in A^*$ there exists a string $u \in A^*$ such that for every string $z \in A^*$, $xuz \notin RAND_t^C$.*

Proof. Fix $t \in N$ and $x \in A^*$. Append to x a string u such that $xu = string(m)$ and $1+T < |string(m)|$, where T is a natural number such that $T - \mathrm{O}(\log T) \geq t$. Put

$$y = xua_1^m.$$

Next define the Chaitin computer $C(a_1^m a_2(z)) = yz = xua_1^m z$, where $d(z)$ is a self-delimiting program for z. It is seen that

$$\begin{aligned} H(yz) &\leq m + 1 + H(z) + \mathrm{O}(1) \\ &= |yz| + (H(z) - |z|) + m - |y| + \mathrm{O}(1) \\ &= |yz| + (H(z) - |z|) - |string(m)| + \mathrm{O}(1) \\ &\leq \Sigma(|yz|) - t. \end{aligned}$$

This shows that every extension xuz of x lies in $A^* \setminus RAND_t^C$. □

Corollary 7.2. *For every natural t we can find a string x no extension of which is in $RAND_t^C$.*

The above result shows that in Theorem 6.4 we need both the prefix u and the suffix v, i.e. it is not possible to fix $u = \lambda$ and then find an appropriate w. However, such a possibility is regained—conforming with the probabilistic intuition—as far as we switch from $RAND_t^C$ with a *fixed* t to $RAND_t^C$ with an *appropriate, small* t.

We start first with a preliminary result:

Lemma 7.3. *Let $<$ be a partial order on A^* which is recursive and unbounded. Assume that $<$ satisfies the relation*

$$\sum_{x<w} Q^{-|w| - \lfloor \log_Q |w| \rfloor} = \infty, \text{ for all } x \in A^*.$$

Then, for every string x we can find a string y such that $x < y$ and $H(y) \geq |y| + \lfloor \log_Q |y| \rfloor$.

Proof. Assume, by absurdity, that there exists $x \in A^*$ such that for all $x < u$ one has $H(u) < |u| + \lfloor \log_Q |u| \rfloor$. Put

$$X = \{v \in A^* \mid x < v, H(v) < |v| + \lfloor \log_Q |v| \rfloor\}$$

and notice that

$$\sum_{w \in X} Q^{-H(w)} \leq \sum_{v \in domU_\lambda} Q^{-|v|} < 1.$$

So,

$$1 > \sum_{w \in X} Q^{-H(w)} \geq \sum_{w \in X} Q^{-|w| - \lfloor \log_Q |w| \rfloor} = \sum_{x<w} Q^{-|w| - \lfloor \log_Q |w| \rfloor} = \infty,$$

since

$$\{w \in A^* \mid x < w, H(w) \geq |w| + \lfloor \log_Q |w| \rfloor\} = \emptyset.$$

We have got a contradiction. □

Theorem 7.4. *For every string x and natural n we can find a string u such that: i) $|xu| \geq n$, ii) for some natural t (which is about $\lfloor \log_Q |xu| \rfloor$), $xu \in RAND_t^C$.*

Proof. Let $< = <_p$ be the prefix order ($x <_p y$ in case $y = xz$, for some string z). First it is seen that all conditions in Lemma 7.3 are verified: $<_p$ is recursive, unbounded and

$$\begin{aligned}\sum_{x<_p w} Q^{-|w|-\lfloor \log_Q |w| \rfloor} &= \sum_{y \in A^*} Q^{-|xy|-\lfloor \log_Q |xy| \rfloor} \\ &\geq \sum_{y \in A^*} Q^{-|xy|-\log_Q |xy|} \\ &\geq Q^{-|x|-\log_Q |x|} \sum_{n=1}^{\infty} \sum_{y \in A^n} Q^{-|y|-\log_Q |y|} \\ &\geq Q^{-|x|-\log_Q |x|} \sum_{n=1}^{\infty} \frac{1}{n} \\ &= \infty.\end{aligned}$$

Now take $x \in A^*$ and $n \geq 1$. Let y be an arbitrary string such that $|xy| \geq n$. By virtue of the preceding argument we can find a string w such that $xy <_p w$ and

$$H(w) \geq |w| + \lfloor \log_Q |w| \rfloor = \Sigma(|w|) - t,$$

where $t = H(string(|w|)) - \lfloor \log_Q |w| \rfloor$ is about $\lfloor \log_Q |w| \rfloor$. $\square$

8 Chaitin's Model vs Kolmogorov's Model

The original definition of random strings (see Kolmogorov [22], Chaitin [9, 10, 15]) is motivated by the fact that

$$\max_{|x|=n} K(x) = |x| + \mathrm{O}\ (1);$$

accordingly, x is called *Kolmogorov t-random* if $K(x) \geq |x| - t$; $RAND_t^K$ stands for the set of Kolmogorov t-random strings.[6]

All results proven in this paper concerning the adequacy of Chaitin's definition of random strings actually hold true for Kolmogorov's model of random strings.[7] To the best of our knowledge there are no "natural" properties associated with randomness valid for one model and not valid for the other one. The underlying complexities H and K are "asymptotical equivalent". Indeed, a crude relation between H and K is the following:

$$H(x) \asymp K(x).$$

[6] Martin-Löf [26] used the blank-endmarker complexity of a string relative to its length to measure the degree of randomness of a string "within" the context of all strings having the same length.

[7] See Chaitin [11, 12, 13, 14, 15], Martin-Löf [26], Solovay [28], Calude [3, 4], Li and Vitányi [23] for a more detailed discussion.

A more exact relation was obtained by Solovay [28]. Put:

$$K^1(x) = K(x),\ K^{n+1}(x) = K(string(K^n(x))),$$

$$H^1(x) = H(x),\ H^{n+1}(x) = H(string(H^n(x))).$$

Theorem 8.1. *The following relations hold true:*

$$H(x) = K(x) + K^2(x) + O(K^3(x)),$$

$$K(x) = H(x) - H^2(x) + O(H^3(x)).$$

In view of Theorem 8.1 it might be the case that the set of Kolmogorov random strings actually coincides with the set of Chaitin random strings. **This is not the case!**

Using the proof of Theorem 3.5 one can show that every Chaitin random string is Kolmogorov random. However, the converse is not true as Solovay [28] has shown. Actually, Solovay [29] conjectures that there exists a constant L such that for all sufficiently large n, there are at least $Q^{n/2}$ strings of length n, s, such that:

$$K(s) \geq |x| - L,$$

$$H(s) \leq |s| + H(string(n)) - \frac{1}{2}K^2(string(n)).$$

So, many Kolmogorov random strings only "look" random, but in fact, they are not. It is an **open question** to find out "natural" properties related to the informal notion of randomness which hold true for Chaitin random strings, but fail to be true for Kolmogorov random strings. Martin-Löf analysis, developed in Section 4, is not fine enough for this problem.

9 The Role of the Underlying Alphabet

It seems that there is a wide spread feeling that the binary case encompasses the whole strength and generality of coding phenomena, at least from an algorithmic point of view. The problem is the following: Does there exist a *binary* asymptotical optimal coding of all strings over an alphabet with $q > 2$ elements? Surprisingly, the answer is *negative*. The answer is negative for both complexities K and H. As our main interest is directed to Chaitin complexity we shall outline the results for this complexity measure.

Let $q > p \geq 2$ be naturals, and fix two alphabets, A, X, having q and p elements, respectively. The lengths of $x \in A^*$ and $y \in X^*$ will be denoted by $|x|_A$ and $|y|_X$, respectively. Fix a universal Chaitin computer $U : A^* \times A^* \xrightarrow{o} A^*$ and denote by H its induced complexity.

Does there exist a Chaitin computer $C : X^* \times A^* \xrightarrow{o} A^*$ which is universal for the class of all Chaitin computers acting on A^*?

The upshot is the following result (see Calude [4], Calude, Jürgensen, and Salomaa [8]):

Theorem 9.1. *There is no Chaitin computer $C : X^* \times A^* \xrightarrow{o} A^*$ which is universal for the class of all Chaitin computers acting on A^*.*

The proof is based on the fact that Chaitin complexity cannot be optimized better than linearly, i.e. the Invariance Theorem is the best possible result in this direction: For every fix a real number $0 < \alpha < 1$, there is no Chaitin computer $C : A^* \times A^* \xrightarrow{o} A^*$ such that for all Chaitin computers D one has:

$$H_C(x) \leq \alpha H_D(x) + \mathrm{O}(1).$$

Let us study Chaitin complexity acting on alphabets of different size. We need some more notation. For every natural $i \geq 2$ put $A_i = \{0, 1, \ldots, i-1\}$, and let us denote by $string_i(n)$ the nth string in A_i^* (according to the quasi-lexicographical order induced by $0 < 1 < \cdots < i-1$); let $H_i : A_i^* \to N$ be Chaitin complexity.

Theorem 9.2.
Let $2 \leq q < Q$. Then, there exists a constant α (which depends upon q, Q) such that for all $x \in A_q^*$ we have:

$$|H_q(x) - (\log_q Q) H_q(x)| \leq \alpha.$$

Proposition 9.3. *For every $2 \leq q < Q$ and all $x \in A_q^*$,*

$$H_Q(x) < |x| + O(1).$$

So, no string $x \in A_q^*$ is random over A_Q^*.[8] In the binary case we have only two such strings, namely

$$00\ldots0 \text{ and } 11\ldots1,$$

which are obviously non-random. In the non-binary case we have

$$\sum_{i=2}^{Q-1} i^n \binom{Q}{i}$$

strings over the alphabet A_Q which are non-binary because they do not contain all Q letters. For instance, for $Q = 3$ one has 3×2^n such strings, some of them (in fact, according to Theorem 3.7, more then $3 \times 2^{n-c_2}$, where c_2 is a constant which depends on the size of the alphabet but not on the length n) are random as *binary* strings. So, it is shown once again, that randomness is a contextual property.

10 Conclusion

In view of the above discussion we conclude that Chaitin's model of random strings satisfy many natural requirements related to randomness, so it can be considered as an **adequate model** for finite random objects. It is a better model than the original (Kolmogorov) proposal. The distinction between Chaitin and Kolmogorov models is far from being elucidated, e.g. no property—naturally associated with randomness—holding true for Chaitin random strings and failing to be satisfied by Kolmogorov random strings is actually known. All descriptional complexities in the binary and non-binary cases have crucial differences, so it appears that it is only natural to discuss the complexity and randomness of finite objects in a non-necessarily binary framework.

[8] This result follows also from Theorem 6.3.

Acknowledgment

I wish to warmly thank Greg Chaitin for many stimulating discussions on random strings (by email, in Auckland, New Zealand and Bar Harbor, Maine, US) and for kindly permitting to incorporate his proofs for Theorems 4.5 and 6.4. I express my gratitude to Helmut Jürgensen, Per Martin-Löf, Charles Rackhoff, Arto Salomaa, and Bob Solovay for their illuminating comments.

References

1. É. Borel. Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Mat. Palermo* 27(1909), 247-271.
2. É. Borel. *Lecons sur la théorie des fonctions*, Gauthier-Villars, Paris, 2nd ed., 1914.
3. C. Calude. *Theories of Computational Complexity*, North-Holland, Amsterdam, New York, Oxford, Tokyo, 1988.
4. C. Calude. *Information and R andomness. An Algorithmic Perspective*, Springer-Verlag, Berlin, 1994. (Forewords by G. J. Chaitin and A. Salomaa)
5. C. Calude. Borel normality and algorithmic randomness, in G. Rozenberg, A. Salomaa (eds.). *Developments in Language Theory*, World Scientific, Singapore, 1994, 113-129. (With a note by G. J. Chaitin)
6. C. Calude, C. Câmpeanu. Note on the topological structure of random strings, *Theoret. Comput. Sci.* 112(1993), 383-390.
7. C. Calude, I. Chitescu. A class of universal P. Martin-Löf tests, *EATCS Bull.* 25 (1984), 14-19.
8. C. Calude, H. Jürgensen, A. Salomaa. *Coding without Tears*, manuscript, February 1994, 15 pp.
9. G. J. Chaitin. On the length of programs for computing finite binary sequences, *J. Assoc. Comput. Mach.*. 13(1966),547-569. (Reprinted in: Chaitin [15], 369-410.)
10. G. J. Chaitin. On the length of programs for computing finite binary sequences: statistical considerations, *J. Assoc. Comput. Mach.* 16(1969), 145-159. (Reprinted in: Chaitin [15], 411-434.)
11. G. J. Chaitin. Information-theoretic limitations of formal systems, *J.Assoc. Comput. Mach.* 21(1974), 403-424. (Reprinted in: Chaitin [15], 291-333.)
12. G. J. Chaitin. A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.* 22(1975), 329-340. (Reprinted in: Chaitin [15], 197-223.)
13. G. J. Chaitin. Algorithmic information theory, *IBM J. Res. Develop.* 21(1977), 350-359,496. (Reprinted in: Chaitin [15], 83-108.)
14. G. J. Chaitin. *Algorithmic Information Theory*, Cambridge University Press, Cambridge,1987. (third printing 1990)
15. G. J. Chaitin. *Information, Randomness and Incompleteness, Papers on Algorithmic Information Theory*, World Scientific, Singapore, New Jersey, Hong Kong, 1987.(2nd ed., 1990)
16. G. J. Chaitin. Randomness in arithmetic, *Scientific American* 259(1988), 80-85. (Reprinted in: Chaitin [15], 14-19.)
17. G. J. Chaitin. *Information-Theoretic Incompleteness*, World Scientific, Singapore, New Jersey, Hong Kong, 1992.
18. G. J. Chaitin. On the number of N-bit strings with maximum complexity, *Applied Mathematics and Computation* 59(1993), 97-100.
19. G. J. Chaitin. *The Limit s of Mathematics*, IBM Watson Center, Yorktown Heights, Draft July 23, 1994, 219 pp.

20. P. Gács. *Lecture Notes on Descriptional Complexity and Randomness*, Boston University, 1988, manuscript, 62 pp.
21. P. S. Laplace. *A Philosophical Essay on Probability Theories*, Dover, New York, 1951.
22. A. N. Kolmogorov. Three approaches for defining the concept of "information quantity", *Problems Inform. Transmission* 1(1965), 3-11.
23. M. Li, P. M. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, Berlin, 1993.
24. X. Li. Effective immune sets, program index sets and effectively simple sets - generalizations and applications of the recursion theorem, in C. -T. Chong, M. J. Wicks (eds.). *South-East Asian Conference on Logic*, Elsevier, Amsterdam, 1983, 97-106.
25. S. Marcus (ed.). *Contextual Ambiguities in Natural & Artificial Languages*, Vol. 2, Ghent, Belgium, 1983.
26. P. Martin-Löf. The definition of random sequences, *Inform. and Control* 9(1966), 602-619.
27. A. Salomaa. *Public-Key Cryptography*, Springer Verlag, Berlin, 1990.
28. R. 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.
29. R. M. Solovay. Email to C. Calude, August 13, 1994.
30. R. M. Smullyan. Effectively simple sets, *Proc. Amer. Math. Soc.* 15(1964), 893-895.
31. V. A. Uspensky. Kolmogorov and mathematical logic,*J . Symbolic Logic* 57(1992), 385-412.