

# ON THE LENGTH OF PROGRAMS FOR COMPUTING FINITE BINARY SEQUENCES: STATISTICAL CONSIDERATIONS

Journal of the ACM 16 (1969),  
pp. 145–159

Gregory J. Chaitin<sup>1</sup>  
*Buenos Aires, Argentina*

## Abstract

*An attempt is made to carry out a program (outlined in a previous paper) for defining the concept of a random or patternless, finite binary sequence, and for subsequently defining a random or patternless, infinite binary sequence to be a sequence whose initial segments are all random or patternless finite binary sequences. A definition based on*

*the bounded-transfer Turing machine is given detailed study, but insufficient understanding of this computing machine precludes a complete treatment. A computing machine is introduced which avoids these difficulties.*

### Key Words and Phrases:

computational complexity, sequences, random sequences, Turing machines

### CR Categories:

5.22, 5.5, 5.6

## 1. Introduction

In this section a definition is presented of the concept of a random or patternless binary sequence based on 3-tape-symbol bounded-transfer Turing machines.<sup>2</sup> These computing machines have been introduced and studied in [1], where a proposal to apply them in this manner is made. The results from [1] which are used in studying the definition are listed for reference at the end of this section.

An  $N$ -state, 3-tape-symbol bounded-transfer Turing machine is defined by an  $N$ -row, 3-column table. Each of the  $3N$  places in this table must contain an ordered pair  $(i, j)$  of natural numbers where  $i$  takes on values from  $-b$  to  $b$ , and  $j$  from 1 to 5.<sup>3</sup> These entries constitute, when specified, the program of the  $N$ -state, 3-tape-symbol bounded-transfer Turing machine and are to be interpreted as follows. An entry  $(i, j)$  in the  $k$ th row and the  $p$ th column of the table means that when the machine is in its  $k$ th state, and the square of its one-way infinite tape

---

<sup>1</sup>Address: Mario Bravo 249, Buenos Aires, Argentina.

<sup>2</sup>The choice of 3-tape-symbol machines is made merely for the purpose of fixing ideas.

<sup>3</sup>Here  $b$  is a constant whose value is to be regarded as fixed throughout this paper. Its exact value is not important as long as it is not "too small." For an explanation of the meaning of "too small," and proofs that  $b$  can be chosen so that it is not too small, see [1, Secs. 2.1 and 2.2]. ( $b$  will not be mentioned again.)



Turing machine is said to calculate a particular finite binary sequence (e.g. 01111000) if the machine stops with that sequence written at the end of its tape, with all other squares of the tape blank, and with its scanner on the first blank square of the tape. Figure 2 illustrates a machine which has calculated the particular sequence mentioned above.

Before proceeding we would like to make a comment from the point of view of the programmer. The logical design of the bounded-transfer Turing machine provides automatically for relocation of programs, and the preceding paragraph establishes linkage conventions for subroutines which calculate finite binary sequences.

Two functions are now defined which play fundamental roles in all that follows.  $L$ , the first function,<sup>4</sup> is defined on the set of all finite binary sequences  $S$  as follows: An  $N$ -state, 3-tape-symbol bounded-transfer Turing machine can be programmed to calculate  $S$  if and only if  $N \geq L(S)$ .

The second function  $L(C_n)$  is defined as

$$L(C_n) = \max_{S \text{ of length } n} L(S)$$

where the maximum is taken (as indicated) over all binary sequences  $S$  of length  $n$ . Also denote by<sup>5</sup>  $C_n$  the set of all binary sequences  $S$  of length  $n$  satisfying  $L(S) = L(C_n)$ .

An attempt is made in [1, Sec. 3.1] to make it plausible, on the basis of various philosophical considerations, that the patternless or random finite binary sequences of length  $n$  are those sequences  $S$  for which  $L(S)$  is approximately equal to  $L(C_n)$ . Here an attempt is made to clarify this (somewhat informal) definition and to make it plausible by proving various results concerning what may be termed statistical properties of such finite binary sequences. The set  $C_\infty$  of patternless or random, infinite binary sequences is formally defined to be the set of all infinite binary sequences  $S$  which satisfy the following inequality for all sufficiently large values of  $n$ :

$$L(S_n) > L(C_n) - f(n)$$

---

<sup>4</sup>Use of the letter " $L$ " is suggested by the phrase "the *Length* of program necessary for computing...".

<sup>5</sup>Use of the letter " $C$ " is suggested by the phrase "the most *Complex* binary sequences of length...".

where  $f(n) = 3 \log_2 n$  and  $S_n$  is the sequence of the first  $n$  bits of  $S$ .

This definition, unlike the first, is quite precise but is also somewhat arbitrary. The failure to state the exact cut-off point at which  $L(S)$  becomes too small for  $S$  to be considered random or patternless gives to the first definition its informal character. But in the case of finite binary sequences, no gain in clarity is achieved by arbitrarily settling on a cut-off point, while the opposite is true for infinite sequences.

The results from [1] which we need are as follows:

$$L(S * S') \leq L(S) + L(S') \tag{1}$$

where  $S$  and  $S'$  are finite binary sequence and  $*$  is the concatenation operation.

$$L(C_{n+m}) \leq L(C_n) + L(C_m) \tag{2}$$

There exists a positive real constant  $a^*$  such that (3)

$$(L(C_n)/n) \geq a^* \tag{3a}$$

$$\lim_{n \rightarrow \infty} (L(C_n)/n) = a^*. \tag{3b}$$

There exists an integer  $c$  such that there are less than  $2^{n-m}$  binary sequences  $S$  of length  $n$  satisfying the inequality

$$L(S) \leq L(C_n) - \log_2 n - m - c. \tag{4}$$

Inequalities (1), (2), and (3a) are used only in Section 6, and (4) is used only in Section 7. For the proofs of (1), (2), and (3) see [1, Sec. 2.3]. The validity of inequality (4) is easily demonstrated using the method of [1, Sec. 2.4].

The following notational conventions are used throughout this paper:

- (a)  $*$  denotes the concatenation operation.
- (b) Let  $S$  be a finite binary sequence.  $S^n$  denotes the result of concatenating  $S$  with itself  $n - 1$  times.
- (c) Let  $m$  be a positive integer.  $B(m)$  denotes the binary sequence which is the numeral representing  $m$  in base-two notation; e.g.  $B(37) = 100101$ . Note that the bit at the left end of  $B(m)$  is always 1.

- (d) Let  $x$  be a real number.  $\lceil x \rceil$  denotes the least integer greater than the enclosed real  $x$ . Note that this is not the usual convention, and that the length of  $B(m)$  is equal to  $\lceil \log_2 m \rceil$ . This last fact will be used but not explicitly mentioned.
- (e)  $\epsilon_x$  denotes a (not necessarily positive) function of  $x$ , and possibly other variables, which approaches zero as  $x$  approaches infinity with any other variables held fixed.
- (f) Let  $S$  be an infinite binary sequence.  $S_k$  denotes the binary sequence consisting of the first  $k$  bits of  $S$ .

## 2. The Fundamental Theorem

All of our results concerning the statistical properties of random binary sequences will be established by applying the result which is proved in this section.

**Theorem 1.** Let  $q$  be an effective ordering of the finite binary sequences of any given length among themselves; i.e. let  $q$  be an effectively computable function with domain consisting of the set of all finite binary sequences and with range consisting of the set of all positive integers, and let the restriction of  $q$  to the domain of the set of all binary sequences of length  $n$  have the range  $\{1, 2, 3, \dots, 2^n\}$ . Then there exists a positive integer  $c$  such that for all binary sequences  $S$  of length  $n$ ,

$$L(S) \leq L(C_{\lceil \log_2 q(S) \rceil}) + L(C_{\lceil \log_2 n \rceil}) + c.$$

*Proof.* The program in Figure 3 calculates  $S$  and consists of the following number of rows:

$$1 + L(B(q(S))) + 1 + L(B(n)) + (c - 2) \leq L(C_{\lceil \log_2 q(S) \rceil}) + L(C_{\lceil \log_2 n \rceil}) + c.$$

## 3. An Application: Matching Pennies

The following example of an application of Theorem 1 concerns the game of Matching Pennies.

<p><i>Section I:</i> 1,4 1,4 1,4</p>
<p><i>Section II</i> consists of <math>L(B(q(S)))</math> rows. It is a program for calculating <math>B(q(S))</math> consisting of the smallest possible number of rows.</p>
<p><i>Section III:</i> 1,4 1,4 1,4</p>
<p><i>Section IV</i> consists of <math>L(B(n))</math> rows. It is a program for calculating <math>B(n)</math> consisting of the smallest possible number of rows.</p>
<p><i>Section V</i> consists by definition of <math>c - 2</math> rows. It calculates the effectively computable function <math>q^{-1}(q(S), n) = S</math>; it finds the two arguments on the tape.</p>

Figure 3. Proof of Theorem 1

According to the von Neumann and Morgenstern theory [2] of the mixed strategy for nonstrictly determined, zero-sum two-person games, a rational player will choose heads and tails equiprobably by some “device subject to chance”.

**Theorem 2.** Let  $S_1, S_2, S_3, \dots$  be a sequence of distinct, finite binary sequences of lengths, respectively,  $n_1, n_2, n_3, \dots$  which satisfies  $L(S_k) \sim L(C_{n_k})$ . Let  $st$  be an effectively computable binary function defined on the set of all finite binary sequences and the null sequence. For each positive integer  $k$  consider a sequence of  $n_k$  plays of the game of Matching Pennies. There are two players:  $A$  who attempts to avoid matches and  $B$  who attempts to match  $A$ 's penny. The players employ the following strategies for the  $m$ th ( $1 \leq m \leq n_k$ ) play of the sequence.  $A$ 's strategy is to choose heads (tails) if the  $m$ th bit of  $S_k$  is 1 (0).  $B$ 's strategy is to choose heads (tails) if  $1$  (or  $0$ ) =  $st$ (the sequence consisting of the  $m - 1$  successive choices of  $A$  up to this point on this sequence of plays, heads being represented by 1 and tails by 0). Then as  $k$  approaches infinity, the ratio of the two quantities (the sum of

the payoffs to  $A$  ( $B$ ) during the  $k$ th sequence of plays of the game of Matching Pennies) approaches the limit 1.

In other words, a random or patternless sequence of choices of heads or tails will be matched about half the time by an opponent who attempts to predict the next choice in an effective manner from the previous choices.

The proof of Theorem 2 is similar to the proof of Theorem 3 below, and therefore is omitted.

## 4. An Application: Simple Normality

In analogy to Borel's concept of the simple normality of a real number  $r$  in the base  $b$  (see [3, Ch. 9] for a definition of this concept of Borel), let a sequence  $S_1, S_2, S_3, \dots$  of finite  $b$ -ary sequences be called simply normal if

$$\lim_{k \rightarrow \infty} \frac{\text{the number of occurrences of } a \text{ in } S_k}{\text{the length of } S_k} = \frac{1}{b}$$

for each of the  $b$  possible values of  $a$ . The application of Theorem 1 given in this section concerns the simple normality of a sequence  $S'_1, S'_2, S'_3, \dots$  of finite  $b$ -ary sequences in which each of the  $S'_k$  is associated with a binary sequence  $S_k$  in a manner defined in the next paragraph. It will turn out that  $L(S_k) \sim L(C_{n_k})$ , where  $n_k$  is the length of  $S_k$ , is a sufficient condition for the simple normality of the sequence of associated sequences.

Given a finite binary sequence, we may place a binary point to its left and consider it to be the base-two notation for a nonnegative real number  $r$  less than 1. Having done so it is natural to consider, say, the ternary sequence used to represent  $r$  to the same degree of precision in base-three notation. Let us define this formally for an arbitrary base  $b$ . Suppose that the binary sequence  $S$  of length  $n$  represents a real number  $r$  when a binary point is affixed to its left. Let  $n'$  be the smallest positive integer for which  $2^n \leq b^{n'}$ . Now consider the set of all reals written in base- $b$  notation as a "decimal" point followed by any of the  $b^{n'}$   $b$ -ary sequences of length  $n'$ , including those with 0's at the right end. Let  $r'$  be the greatest of these reals which is less than or

equal to  $r$ , and let the  $b$ -ary sequence  $S'$  be the one used to represent  $r'$  in base- $b$  notation.  $S'$  is the  $b$ -ary sequence which we will associate with the binary sequence  $S$ . Note that no two binary sequences of the same length are associated with the same  $b$ -ary sequence.

It is now possible to state the principal result of this section.

**Theorem 3.** Let  $S_1, S_2, S_3, \dots$  be a sequence of distinct, finite binary sequences of lengths, respectively,  $n_1, n_2, n_3, \dots$  which satisfies  $L(S_k) \sim L(C_{n_k})$ . Then the sequence  $S'_1, S'_2, S'_3, \dots$  of associated  $b$ -ary sequences is simply normal.

We first prove a subsidiary result.

**Lemma 1.** For any real number  $e > 0$ , any real number  $d > 1$ ,  $b$ , and  $0 \leq j < b$ , for all sufficiently large values of  $n$ , if  $S$  is a binary sequence of length  $n$  whose associated  $b$ -ary sequence  $S'$  of length  $n'$  satisfies the following condition

$$\left| \frac{\text{the number of occurrences of } j \text{ in } S'}{n'} - \frac{1}{b} \right| > e, \quad (5)$$

then

$$L(S) < L(C_{\lfloor nd \frac{H(\frac{1}{b} - \frac{e}{b-1}, \dots, \frac{1}{b} - \frac{e}{b-1}, \frac{1}{b} + e)}{\log_2 b} \rfloor}).$$

Here

$$H(p_1, p_2, \dots, p_b) (p_1 \geq 0, p_2 \geq 0, \dots, p_b \geq 0, \sum_{i=1}^b p_i = 1)$$

is defined to be equal to

$$-\sum_{i=1}^b p_i \log_2 p_i$$

where in this sum any terms  $0 \log_2 0$  are to be replaced by 0.

The  $H$  function occurs because the logarithm to the base two of

$$\sum_{\left| \frac{k}{n'} - \frac{b-1}{b} \right| > e} (b-1)^k \binom{n'}{k},$$

the number of  $b$ -ary sequences  $S'$  of length  $n'$  which satisfy (5) is asymptotic, as  $n$  approaches infinity, to

$$n'H \left( \frac{1}{b} - \frac{e}{b-1}, \dots, \frac{1}{b} - \frac{e}{b-1}, \frac{1}{b} + e \right),$$

which is in turn asymptotic to  $nH/\log_2 b$ , for  $n' \sim n/\log_2 b$ . This may be shown by considering the ratio of successive terms of the sum and using Stirling's approximation,  $\log(n!) \sim n \log n$  [4, Ch. 6, Sec. 3].

To prove Lemma 1 we first define an ordering  $q$  by the following two conditions:

- (a) Consider two binary sequences (of length  $n$ )  $S$  and  $T$  whose associated  $b$ -ary sequences (of length  $n'$ )  $S'$  and  $T'$  contain, respectively,  $s$  and  $t$  occurrences of  $j$ .  $S$  comes before (after)  $T$  if

$$\left| \frac{s}{n'} - \frac{1}{b} \right| \text{ is greater (less) than } \left| \frac{t}{n'} - \frac{1}{b} \right|.$$

- (b) If condition (a) doesn't settle which of the two sequences of length  $n$  comes first, take  $S$  to come before (after)  $T$  if  $S'$  represents (ignoring 0's to the left) a larger (smaller) number in base- $b$  notation than  $T'$  represents.<sup>6</sup>

*Proof.* We now apply Theorem 1 to any binary sequence  $S$  of length  $n$  such that its associated  $b$ -ary sequence  $S'$  of length  $n'$  satisfies (5). Theorem 1 gives us

$$L(S) \leq L(C_{[\log_2 q(S)]}) + L(C_{[\log_2 n]}) + c \quad (6)$$

where, as we know from the paragraph before the last, for all sufficiently large values of  $n$ ,

$$\log_2 q(S) < \left( 1 + \frac{1}{4}(d-1) \right) \frac{nH}{\log_2 b}. \quad (7)$$

From (3b) and (7) we obtain for large values of  $n$ ,

$$L(C_{[\log_2 q(S)]}) < a^* \left( 1 + \frac{1}{2}(d-1) \right) \frac{nH}{\log_2 b}. \quad (8)$$

---

<sup>6</sup>This condition was chosen arbitrarily for the sole purpose of "breaking ties."

And eq. (3b) implies that for large values of  $n$ ,

$$L(C_{\lfloor \log_2 n \rfloor}) + c < a^* \frac{1}{4} (d-1) \frac{nH}{\log_2 b}. \quad (9)$$

Adding ineqs. (8) and (9), we see that ineq. (6) yields, for large values of  $n$ ,

$$L(S) < a^* \left(1 + \frac{3}{4}(d-1)\right) \frac{nH}{\log_2 b}.$$

Applying eq. (3b) to this last inequality, we see that for all sufficiently large values of  $n$ ,

$$L(S) < L(C_{\lfloor nd \frac{H}{\log_2 b} \rfloor}),$$

which was to be proved.

Having demonstrated Lemma 1 we need only point out that Theorem 3 follows immediately from Lemma 1, eq. (3b), and the fact that

$$H(p_1, p_2, \dots, p_b) \leq \log_2 b,$$

with equality if and only if

$$p_1 = p_2 = \dots = p_b = \frac{1}{b}$$

(for a proof of this inequality, see [5, Sec. 2.2]).

## 5. Applications of a von Mises Place Selection V

In this section we consider the finite binary sequence  $S'$  resulting from the application of a von Mises place selection  $V$  to a finite binary sequence  $S$  which is random in the sense of Section 1. For  $S$  not to be rejected as random in the sense of von Mises [6] (i.e. in von Mises' terminology, for  $S$  not to be rejected as a collective<sup>7</sup>),  $S'$  must contain about as many 0's as 1's.

---

<sup>7</sup>Strictly speaking we cannot employ von Mises' terminology here for von Mises was interested only in infinite sequences. Kolmogorov [7] considers finite sequences.

A place selection  $V$  is defined to be a binary function (following Church [8], it must be effectively computable) defined on the set of all finite binary sequences and the null sequence. If  $S = S' * S''$  is a finite binary sequence, then  $V(S') = 0$  (1) is interpreted to mean that the first (i.e. the leftmost) bit of  $S''$  is not (is) selected from  $S$  by the place selection  $V$ .

By applying Theorem 1 and eq. (3b) we obtain the principal result of this section.

**Theorem 4.** Let  $S_1, S_2, S_3, \dots$  be a sequence of distinct finite binary sequences of lengths, respectively,  $n_1, n_2, n_3, \dots$  which satisfies  $L(S_k) \sim L(C_{n_k})$ . Let  $V$  be any place selection such that

$$\inf \left( \frac{\text{length of subsequence of } S \text{ selected by } V}{\text{length of } S} \right) > 0 \quad (10)$$

where the infimum is taken over all finite binary sequences  $S$ . Then as  $k$  approaches infinity, the ratio of the number of 0's in the subsequence of  $S_k$  which is selected by  $V$  to the number of 1's in this subsequence tends to the limit 1.

Before proceeding to the proof it should be mentioned that a similar result can be obtained for the generalized place selections due to Loveland [9–11].

The proof of Theorem 4 runs parallel to the proof of Theorem 3. The subsidiary result which is proved by taking in Theorem 1 the ordering  $q$  defined below is

**Corollary 1.** Let  $e$  be a real number greater than 0,  $d$  be a real number greater than 1,  $S$  be a binary sequence of length  $n$ , and let  $V$  be a place selection which selects from  $S$  a subsequence  $S'$  of length  $n'$ . Suppose that

$$\left| \frac{\text{the number of 0's in } S'}{n'} - \frac{1}{2} \right| > e. \quad (a)$$

Then for  $n'$  greater than  $N$  we have

$$L(S) \leq L(C_{[\log_2 q(S)]}) + L(C_{[\log_2 n]}) + c,$$

where

$$\log_2 q(S) < n'dH\left(\frac{1}{2} + e, \frac{1}{2} - e\right) + (n - n').$$

Here  $N$  depends only on  $e$  and  $d$ , and  $c$  depends only on  $V$ .

**Definition.**<sup>8</sup> Let  $S$  be a binary sequence of length  $n$ , let  $S'$  of length  $n'$  be the subsequence of  $S$  selected by the place selection  $V$ , and let  $S''$  be the subsequence of  $S$  which is not selected by  $V$ . Let<sup>9</sup>

$$Q = F(S') * S'' * 01 * B_1^2 * B_2^2 * B_3^2 * \dots$$

where each  $B_i$  is a single bit and

$$1 * B_1 * B_2 * B_3 * \dots = B(\text{the length of } F(S')).$$

We then define  $q(S)$  to be the unique solution of  $B(q(S)) = Q$ .

**Definition.** (Let us emphasize that  $F(S')$  is never more than about

$$n'H\left(\frac{1}{2} + e, \frac{1}{2} - e\right)$$

bits long for  $S'$  which satisfy supposition (a) of Cor. 1: this is the crux of the proof.) Consider the “padded” numerals for the integers from 0 to  $2^{n'} - 1$ ; padded to a length of  $n'$  bits by adding 0’s on the left. Arrange these in order of decreasing

$$\left| \frac{m}{n'} - \frac{1}{2} \right|$$

where  $m$  is the number of 0’s in the padded numeral, and when this does not decide the order, in numerical order (e.g. the list starts  $0^{n'}, 1^{n'}, 0^{n'-1} * 1, 0^{n'-2} * 10, \dots$ ). Suppose that  $S'$  is the  $k$ th padded numeral in the list. We define  $F(S')$  to be equal to  $B(k)$ . Further details are omitted.

---

<sup>8</sup>Strictly speaking, this definition is incorrect.  $S'$ , reconstructed from  $F(S')$  and  $n$ , and  $S''$  can be “pieced together” to form  $S$  using  $V$  to dictate the intermixing, and thus  $q(S) = q(T)$  for  $S$  and  $T$  of the same length only if  $S = T$ . But  $q(S)$  is greater than  $2^n$  for some binary sequences  $S$  of length  $n$ . To correct this it is necessary to obtain the “real” ordering  $q'$  from the ordering  $q$  that we define here by “pressing the function  $q$  down so as to eliminate gaps in its range.” Formally, consider the restriction of  $q$  to the domain of all binary sequences of length  $n$ . Let the  $k$ th element in the range of this restriction of  $q$ , ordered according to magnitude, be denoted by  $r_k$ . Let  $S$  satisfy  $q(S) = r_k$ . We define  $q'(S)$  to be equal to  $k$ . As, however, the result of this redefinition is to decrease the value of  $q(S)$  for some  $S$ , this is a quibble.

<sup>9</sup>Our superscript notation for concatenation is invoked here for the first time.

## 6. Fundamental Properties of the L-Function

In Sections 3–5 the random or patternless finite binary sequences have been studied. Before turning our attention to the random or patternless infinite binary sequences, we would like to show that many fundamental properties of the  $L$ -function are simple consequences of the inequality  $L(S * S') \leq L(S) + L(S')$  taken in conjunction with the simple normality of sequences of random finite binary sequences.

In Theorem 3 take  $b = 2^k$  and let the infinite sequence  $S_1, S_2, S_3, \dots$  consist of all the elements of the various  $C_n$ 's. We obtain

**Corollary 2.** For any  $e > 0$ ,  $k$ , and for all sufficiently large values of  $n$ , consider any element  $S$  of  $C_n$  to be divided into between  $(n/k) - 1$  and  $(n/k)$  nonoverlapping binary subsequences of length  $k$  with not more than  $k - 1$  bits left over at the right end of  $S$ . Then the ratio of the number of occurrences of any particular one of the  $2^k$  possible binary subsequences of length  $k$  to  $(n/k)$  differs from  $2^{-k}$  by less than  $e$ .

Keeping in mind the hypothesis of Corollary 2, let  $S$  be some element of  $C_n$ . Then we have  $L(C_n) = L(S)$ , and from Corollary 2 with

$$L(S) = L(S' * S'' * S''' * \dots) \leq L(S') + L(S'') + L(S''') + \dots$$

(this inequality is an immediate consequence of (1)) this gives us

$$L(C_n) \leq \frac{n}{k}(1 + \epsilon_n)(2^{-k} \sum L(S)),$$

where the sum is taken over the set of all binary sequences of length  $k$ . That is,

$$(L(C_n)/n)k \leq (1 + \epsilon_n)(2^{-k} \sum L(S)),$$

with which (3a) or  $(L(C_n)/n) \geq a^*$  gives

$$a^*k \leq (1 + \epsilon_n)(2^{-k} \sum L(S)).$$

We conclude from this last inequality the following theorem.

**Theorem 5.** For all positive integers  $k$ ,<sup>10</sup>

$$a^*k \leq 2^{-k} \sum_{S \text{ of length } k} L(S).$$

Note that the right-hand side of the inequality of Theorem 5 is merely the expected value of the random variable  $L = L(S)$  where the sample space is the set of all binary sequences of length  $k$  to which equal probabilities have been assigned. With this probabilistic framework understood, we can denote the right-hand side of the inequality of Theorem 5 by  $E\{L\}$  and use the notation  $\Pr\{\dots\}$  for the probability of the enclosed event. Recalling eq. (3b) and the definition of  $L(C_k)$  as  $\max L$ , we thus have for any  $e > 0$ ,

$$\begin{aligned} a^*k \leq E\{L\} &= \sum \Pr\{S\}L(S) \\ &\leq \Pr\{L \leq (1-e)a^*k\}((1-e)a^*k) + \\ &\quad (1 - \Pr\{L \leq (1-e)a^*k\})L(C_k) \\ &= \Pr\{L \leq (1-e)a^*k\}((1-e)a^*k) + \\ &\quad (1 - \Pr\{L \leq (1-e)a^*k\})((1+\epsilon_k)a^*k), \end{aligned}$$

or

$$\epsilon_k - (e + \epsilon_k) \Pr\{L \leq (1-e)a^*k\} \geq 0.$$

Thus for any real  $e > 0$ ,

$$\lim_{k \rightarrow \infty} \Pr\{L \leq (1-e)a^*k\} = 0. \quad (11)$$

Although eq. (11) is weaker than (4), it is reached by a completely different route. It must be admitted, however, that it is easy to prove Theorem 5 from (4) by taking into account the subadditivity of the right-hand side of the inequality of Theorem 5.

From Theorem 5 we now demonstrate

**Corollary 3.** For all positive integers  $n$ ,  $(L(C_n)/n) > a^*$ .

*Proof.* Since  $L(0^n) \leq L(B(n)) + c \leq [\log_2 n] + c$ , for large  $n$ ,  $L(S) < a^*n$  for at least one binary sequence of length  $n$ , and we therefore may conclude from Theorem 5 that for large  $n$  there must be at least one

---

<sup>10</sup>This statement remains true, as can be proved in several ways, if “<” replaces “≤”.

binary sequence  $S'$  of length  $n$  for which  $L(S') > a^*n$ ; that is, for large  $n$ ,

$$(L(C_n)/n) > a^*. \quad (12)$$

We now finish the proof of Corollary 3 by contradiction. Suppose that Corollary 3 is false, and there exists an  $n_0$  such that

$$(L(C_{n_0})/n_0) = a^*.$$

(( $L(C_{n_0})/n_0 < a^*$  is impossible by (3a).) Then from (2) and (3a) it would follow that for all positive integers  $k$ ,

$$(L(C_{kn_0})/kn_0) = a^*,$$

which contradicts (12).

The final topic of this section is a derivation of

**Theorem 6.**  $L(C_n) - a^*n$  is unbounded.

*Proof.* Consider some particular binary sequence  $S$  which is a member of  $C_n$ . Then from Corollary 2, for large values of  $n$  there must certainly be a sequence of  $k$  consecutive 0's in  $S$ . Suppose that  $S = R * 0^k * T$ . Then we have

$$\begin{aligned} L(C_n) = L(S) = L(R * 0^k * T) &\leq L(R) + L(0^k) + L(T) \\ &\leq L(R) + L(B(k)) + c + L(T) \\ &\leq L(C_i) + L(C_j) + [\log_2 k] + c \end{aligned}$$

where  $i$  is the length of  $R$ ,  $j$  is the length of  $T$ , and  $n - k = i + j$ . That is,

**Lemma 2.** For any positive integer  $k$ , for all sufficiently large values of  $n$  there exist  $i$  and  $j$  such that

$$L(C_n) \leq L(C_i) + L(C_j) + [\log_2 k] + c,$$

and  $n - k = i + j$ .

Theorem 6 follows immediately from Lemma 2 through proof by contradiction.

## 7. Random or Patternless Infinite Binary Sequences

This section and Section 8 are devoted to a study of the set  $C_\infty$  of random or patternless infinite binary sequences defined in Section 1. Two proofs that  $C_\infty$  is nonempty, both based on (4), are presented here. The first proof is measure theoretic; the measure space employed may be defined in probabilistic terms as follows: the successive bits of an infinite binary sequence are independent random variables which assume the values 0 and 1 equiprobably. The second proof exhibits an element from  $C_\infty$ .

**Theorem 7.**  $C_\infty$  is nonempty.

*First Proof.* From (4) and the Borel-Cantelli lemma, it follows immediately that

$$C_\infty \text{ is a set of measure 1.} \quad (13)$$

*Second Proof.* It is easy to see from (4) that we can find an  $N$  so large that

$$\sum_{k>N} N_k 2^{-k} < 1 \quad (14)$$

where  $N_k$  is the number of binary sequences  $S$  of length  $k$  for which

$$L(S) \leq L(C_k) - 3 \log_2 k.$$

Consider the following process which never terminates (for that would contradict (14)).

Start: Set  $k = 0$ , set  $S =$  null sequence, go to Loop1.

Loop1: Is  $k \leq N$  or  $L(S) > L(C_k) - 3 \log_2 k$ ?

If so, set  $k = k + 1$ , set  $S = S * 0$ , go to Loop1.

If not, go to Loop2.

Loop2: If  $S = S' * 0$ , set  $S = S' * 1$ , go to Loop1.

If  $S = S' * 1$ , set  $k = k - 1$ , set  $S = S'$ .

If  $k \neq 0$ , go to Loop2.

If  $k = 0$ , stop.

Then from Dirichlet's box principle (if an infinity of letters is placed in a finite number of pigeonholes, then there is always a pigeonhole which receives an infinite number of letters) it is clear that from some point on, the first bit of  $S$  will remain fixed; from some point on, the first two bits of  $S$  will remain fixed;  $\dots$ ; from some point on (depending on  $n$ ), the first  $n$  bits of  $S$  will remain fixed;  $\dots$ . Let us denote by  $S_{\text{lim}}$  the infinite binary sequence whose  $n$ th bit is 0 (1) if from some point on, the  $n$ th bit of  $S$  remains 0 (1). It is clear that  $S_{\text{lim}}$  is in  $C_\infty$ .

**Remark.** When  $C_\infty$  was defined in Section 1, we pointed out that this definition contains an arbitrary element, i.e. the choice of  $3 \log_2 n$  as the function  $f(n)$ . In defining  $C_\infty$  it is desirable to choose an  $f$  which goes to infinity as slowly as possible and which results in a  $C_\infty$  of measure 1. We will call such  $f$ 's "suitable." From results in [1, Secs. 2.4 and 2.5], which are more powerful than (5), it follows that there is an  $f$  which is suitable and satisfies the equations

$$\begin{cases} \limsup(f(n)/\log_2 n) = 2a^*, \\ \liminf(f(n)/\log_2 n) = a^*. \end{cases}$$

The question of obtaining lower bounds on the growth of an  $f$  which is suitable will be considered in Section 10, but there a different computing machine is used as the basis for the definition of random or patternless infinite binary sequence.

## 8. Statistical Properties of Infinite, Random or Patternless Binary Sequences

Results concerning the statistical properties of infinite, random or patternless binary sequences follow from the corresponding results for finite sequences. Thus Theorem 8 is an immediate consequence of Theorem 3, and Corollary 1 and eq. (3b) yield Theorem 9.

**Theorem 8.** Real numbers whose binary expansions are sequences in  $C_\infty$  are simply normal in every base.<sup>11</sup>

---

<sup>11</sup>It is known from probability theory that a real  $r$  which is simply normal in every base has the following property. Let  $b$  be a base, and denote by  $a_n$  the  $n$ th "digit" in the base- $b$  expansion of  $r$ . Consider a  $b$ -ary sequence  $c_1, c_2, \dots, c_m$ . As  $n$

**Theorem 9.** Any infinite binary sequence in  $C_\infty$  is a collective with respect to the set of place selections<sup>12</sup> which are effectively computable and satisfy the following condition: For any infinite binary sequence  $S$ ,

$$\liminf \frac{\text{the number of bits in } S_k \text{ which are selected by } V}{k} > 0.$$

## 9. A General Formulation: Binary Computing Machines

Throughout the study of random or patternless binary sequences which has been attempted in the preceding sections, there has been a recurring difficulty. Theorem 1 and the relationship  $L(C_n) \sim a^*n$  have been used as the cornerstones of our treatment, but the assumption that  $L(C_n) \sim a^*n$  does not ensure that  $L(C_n)$  behaves sufficiently smoothly to make really effective use of Theorem 1. Indeed it is conceivable that greater understanding of the bounded-transfer Turing machine would reveal that  $L(C_n)$  behaves rather roughly and irregularly. Therefore a new computing machine is now introduced.<sup>13</sup>

To understand the logical design of this computing machine, it is helpful to provide a general formulation of computing machines for calculating finite binary sequences whose programs are also finite binary sequences. We call these *binary computing machines*. Formally, a binary computing machine is a partial recursive function  $M$  of the finite binary sequences which is finite binary sequence valued. The argument of  $M$  is the program, and the partial recursive function gives the output (if any) resulting from that program.  $L_M(S)$  and  $L_M(C_n)$  (if the

---

approaches infinity the ratio of (the number of those positive integers  $k$  less than  $n$  which satisfy  $a_k = c_1, a_{k+1} = c_2, \dots, a_{k+m-1} = c_m$ ) to  $n$  tends to the limit  $b^{-m}$ .

<sup>12</sup>Wald [12] introduced the notion of a collective with respect to a set of place selections; von Mises had originally permitted "all place selections which depend only on elements of the sequence previous to the one being considered for selection."

<sup>13</sup>The author has subsequently learned of Kolmogorov [13], in which a similar kind of computing machine is used in essentially the same manner for the purpose of defining a finite random sequence. Martin-Löf [14–15] studies the statistical properties of these random sequences and puts forth a definition of an infinite random sequence.

computing machine is understood, the subscript will be omitted) are defined as follows:

$$L_M(S) = \begin{cases} \min_{M(P)=S}(\text{length of } P), \\ \infty \text{ if there are no such } P, \end{cases}$$

$$L_M(C_n) = \max_{S \text{ of length } n} L_M(S).$$

In this general setting the program for the definition of a random or patternless binary sequence assumes the following form: The patternless or random finite binary sequences of length  $n$  are those sequences  $S$  for which  $L(S)$  is approximately equal to  $L(C_n)$ . The patternless or random infinite binary sequences  $S$  are those whose truncations  $S_n$  are all patternless or random finite sequences. That is, it is necessary that for large values of  $n$ ,  $L(S_n) > L(C_n) - f(n)$  where  $f$  approaches infinity slowly.

We define below a binary computing machine  $M^*$  which has, as is easily seen, the following very convenient properties.

- (a)  $L(C_n) = n + 1$ .
- (b) Those binary sequences  $S$  of length  $n$  for which  $L(S) < L(C_n) - m$  are less than  $2^{n-m}$  in number.
- (c) For any binary computer  $M$  there exists a constant  $c$  such that for all finite binary sequences  $S$ ,  $L_{M^*}(S) \leq L_M(S) + c$ .

The computing machine  $M^*$  is constructed from the two-argument partial recursive function  $U(P, M')$ , a universal binary computing machine. That is,  $U$  is characterized (following Turing) by the property that for any binary computer  $M$  there exists a finite binary sequence  $M'$  such that for all programs  $P$ ,  $U(P, M') = M(P)$  where both sides of the equation are undefined whenever one of them is.

**Definition.** If possible<sup>14</sup> let  $P = P' * B$  where  $B$  is a single bit. If  $B = 1$  then we define  $M^*(P)$  to be equal to  $P'$ . If  $B = 0$  then let the following equation be examined for a solution:  $P' = S * T * 01 * B_1^2 * B_2^2 * B_3^2 * \dots$  where each  $B_i$  is a single bit,  $1 * B_1 * B_2 * B_3 * \dots = B(n)$ , and  $T$  is of length  $n$ . If this equation has a solution then the solution must be unique, and we define  $M^*(P)$  to be equal to  $U(S, T)$ .

<sup>14</sup>That is, if  $P$  is a single bit this is not possible.  $M^*(P)$  is therefore undefined.

## 10. Bounds on Suitable Functions

In Section 7 we promised to provide bounds on any  $f$  which is suitable (i.e. suitable for defining a  $C_\infty$  of measure 1). We prove here that

$$\limsup f(k)/\log_2 k \geq 1,$$

the constant being best possible.

We use the result [4 (1950 ed.), p. 163, prob. 4] that the set  $\#$  of those infinite binary sequences  $S$  for which  $r(S_k) > [\log_2 k]$  infinitely often is of measure 1; here  $r$  denotes the length of the run of 0's at the right end of the sequence.<sup>15</sup> As  $\#$  and  $C_\infty$  are both of measure 1, they have an element  $S$  in common. Then for infinitely many values of  $k$ ,

$$\begin{cases} L(S_k) > L(C_k) - f(k), \\ r(S_k) > [\log_2 k]. \end{cases}$$

But taking into account property (c) of  $M^*$ , we see that  $S_k = \dots * 0^{[\log_2 k]}$  implies that  $L(S_k) \leq L(C_{k-[\log_2 k]}) + c$ . Thus for infinitely many values of  $k$ ,

$$L(C_{k-[\log_2 k]}) + c \geq L(S_k) > L(C_k) - f(k)$$

or

$$k - [\log_2 k] + 1 + c \geq L(S_k) > k + 1 - f(k),$$

which implies that  $f(k) > [\log_2 k] - c$ . Hence  $\limsup f(k)/\log_2 k$  must be greater than or equal to 1.

Now it is necessary to show that the constant is the best possible. From the Borel-Cantelli lemma and property (b) of  $M^*$ , we see at once that for  $f$  to be suitable, it is sufficient that

$$\sum_{k=1}^{\infty} 2^{-f(k)}$$

converges. Thus  $f(k) = \log_2(k(\log k)^2)$  is suitable, and this  $f(k)$  is asymptotic to  $\log_2 k$ .

---

<sup>15</sup>We are indebted to Professor Leonard Cohen of the City University of New York for pointing out to us the existence of such results.

## 11. Two Analogues to the Fundamental Theorem

To study the statistical properties of binary sequences which are defined to be random or patternless on the basis of the computing machine  $M^*$ , it is necessary to have, in addition to properties (a) and (b) of  $M^*$ , an analogue to Theorem 1. We state two, the second of which is just a refinement of the first. Both are proved using property (c) of  $M^*$ .

**Theorem 10.** On the hypothesis of Theorem 1, for all binary sequences  $S$  of length  $n$ ,

$$L(S) \leq L(B(q(S)) * B(n) * 01 * B_1^2 * B_2^2 * B_3^2 * \cdots) + c$$

where each  $B_i$  is a single bit and  $1 * B_1 * B_2 * B_3 * \cdots = B([\log_2 n])$ . Thus

$$L(S) \leq L(C_{g(q(S),n)}) + c \leq g(q(S), n) + c'$$

where

$$g(q(S), n) = [\log_2 q(S)] + [\log_2 n] + 2[\log_2 [\log_2 n]].$$

**Theorem 11.** On the hypothesis of Theorem 1, for all binary sequences  $S$  of length  $n$ ,

$$L(S) \leq L(B(q(S)) * B(n + 2 - [\log_2 q(S)]) * 01 * B_1^2 * B_2^2 * B_3^2 * \cdots) + c$$

where each  $B_i$  is a single bit,  $1 * B_1 * B_2 * B_3 * \cdots = B([\log_2 g(q(S), n)])$ , and

$$g(q(S), n) = n + 2 - [\log_2 q(S)].$$

Thus

$$L(S) \leq L(C_{h(q(S),n)}) + c \leq h(q(S), n) + c'$$

where

$$h(q(S), n) = [\log_2 q(S)] + [\log_2 g(q(S), n)] + 2[\log_2 [\log_2 g(q(S), n)]].$$

On comparing property (a) of  $M^*$ , property (b) of  $M^*$ , and Theorem 11 with, respectively, (3b), (4), and Theorem 1, we see that they are analogous but far more powerful. It therefore follows that Sections

3–5, 7, and 8 can be applied almost *verbatim* to the present computing machine. In particular, Theorem 2, Lemma 1, Theorem 4, (13), Theorem 8, and Theorem 9 hold, without any change whatsoever, for the random sequences defined on the basis of  $M^*$ . In all cases, however, much stronger assertions can be made. For example, in place of Theorem 9 we can state that

**Theorem 12.**<sup>16</sup> The set  $C_\infty$  of all infinite binary sequences  $S$  which have the property that for all sufficiently large values of  $k$ ,  $L(S_k) > L(C_k) - \log_2(k(\log k)^2)$ , is of measure 1, and each element of  $C_\infty$  is a collective with respect to the set of place selections  $V$  which are effectively computable and satisfy the following condition:<sup>17</sup> For any infinite binary sequence  $S$ ,

$$\lim_{k \rightarrow \infty} \frac{\text{the number of bits in } S_k \text{ which are selected by } V}{\log_2 k} = \infty.$$

## References

- [1] CHAITIN, G. J. On the length of programs for computing finite binary sequences. *J. ACM* 13, 4 (Oct. 1966), 547–569.
- [2] VON NEUMANN, J., AND MORGENSTERN, O. *Theory of Games and Economic Behavior*. Princeton U. Press, Princeton, N. J., 1953.
- [3] HARDY, G. H., AND WRIGHT, E. M. *An Introduction to the Theory of Numbers*. Oxford U. Press, Oxford, 1962.
- [4] FELLER, W. *An Introduction to Probability Theory and Its Applications, Vol. I*. Wiley, New York, 1964.
- [5] FEINSTEIN, A. *Foundations of Information Theory*. McGraw-Hill, New York, 1958.
- [6] VON MISES, R. *Probability, Statistics, and Truth*. Macmillan, New York, 1939.

---

<sup>16</sup>Compare the last paragraph of Section 10.

<sup>17</sup>In view of Section 10, it apparently is not possible by the methods of this paper to replace the “ $\log_2 k$ ” here by a significantly smaller function.

- [7] KOLMOGOROV, A. N. On tables of random numbers. *Sankhyā* [A], 25 (1963), 369–376.
- [8] CHURCH, A. On the concept of a random sequence. *Bull. Amer. Math. Soc.* 46 (1940), 130–135.
- [9] LOVELAND, D. W. *Recursively Random Sequences*. Ph.D. Diss., N.Y.U., June 1964.
- [10] —. The Kleene hierarchy classification of recursively random sequences. *Trans. Amer. Math. Soc.* 125 (1966), 487–510.
- [11] —. A new interpretation of the von Mises concept of random sequence. *Z. Math. Logik Grundlagen Math.* 12 (1966), 279–294.
- [12] WALD, A. Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung. *Ergebnisse eines mathematischen Kolloquiums* 8 (1937), 38–72.
- [13] KOLMOGOROV, A. N. Three approaches to the definition of the concept “quantity of information.” *Problemy Peredachi Informatsii* 1 (1965), 3–11. (in Russian)
- [14] MARTIN-LÖF, P. The definition of random sequences. Res. Rep., Inst. Math. Statist., U. of Stockholm, Stockholm, 1966, 21 pp.
- [15] —. The definition of random sequences. *Inform. Contr.* 9 (1966), 602–619.
- [16] LÖFGREN, L. Recognition of order and evolutionary systems. In *Computer and Information Sciences—II*, Academic Press, New York, 1967, pp. 165–175.
- [17] LEVIN, M., MINSKY, M., AND SILVER, R. On the problem of the effective definition of “random sequence”. Memo 36 (revised), RLE and MIT Comput. Center, 1962, 10 pp.