INFORMATION AND CONTROL 15, 510–526 (1969)

A Variant of the Kolmogorov Concept of Complexity*

D. W. LOVELAND

Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

Kolmogorov in 1965 proposed two related measures of information content (alternately, measures of complexity) based on the size of a program which when processed by a suitable algorithm (machine) yields the desired object. The main emphasis was placed on a conditional complexity measure. In this paper a simple variation of the (restricted) conditional complexity measure investigated by Martin-Löf is noted because of interesting characteristics not shared by the measures proposed by Kolmogorov. The characteristics suggest situations in which this variant is the most desirable measure to employ. The interpretation of the measure offers some desirable general qualities; also the measure is relatively advantageous when working with entities of low complexity and maintains the important properties of the Kolmogorov conditional complexity measure when concerned with high complexity.

1. We consider the measures of complexity introduced by Kolmogorov [3]. The domain of entities of concern here is the domain X of finite binary sequences. A finite (binary) sequence of length n is denoted by x^n with subscripts employed when it is necessary to distinguish among sequences of length n. Likewise, p denotes a finite binary sequence (we omit the length denotation here) and $l(p)$ = length of p, so $l(x^n) = n$. The capital letters A and B denote effectively computable functions from X to X (or $X \times N$ to X in proper context, where N is the set of positive integers). It is convenient to regard A as either a recursive function having as an argument a suitable encoding of p, or as a computing machine with input sequence p; the choice is determined by context.

The *Kolmogorov complexity* of x^n with respect to algorithm A is given by

$$K_A(x^n) = \min_{A(p)=x^n} l(p),$$

* The author was partially supported by Air Force grant AF-AFOSR-995-67 during the writing of this paper.

if there exists a program p such that $A(p) = x^n$; otherwise

$$K_A(x^n) = \infty.$$

Likewise, the (restricted) *Kolmogorov conditional complexity* (or, simply, the *conditional complexity*) of x^n with respect to algorithm A is given by

$$K_A(x^n \mid n) = \min_{A(p,n)=x^n} l(p),$$

if there exists a program p such that $A(p, n) = x^n$; otherwise

$$K_A(x^n \mid n) = \infty.$$

Kolmogorov suggested that the finite sequences of relatively high complexity should be considered random sequences (if the algorithm A is of a suitable "universal" class, see below) and that this identification conforms to our intuitive concept of the phrase "random sequence". In an important paper Martin-Löf [4] outlined a mechanism using what we shall call Martin-Löf tests whereby many important properties of randomness could be shown to hold for finite sequences of sufficiently high conditional complexity. We shall assume the leader is familiar with the Martin-Löf paper [4] which shall be sufficient background for reading this paper. Other than in the brief discussion below, we will not consider further (nor does Martin-Löf) the (nonconditional) Kolmogorov complexity measure. (A complexity measure similar to the Kolmogorov complexity measure is independently introduced and studied by Chaitin ([1] and [2])).

We recall a key property, here stated for the conditional complexity but true for all variants to be discussed. *There exists a universal algorithm A such that for arbitrary algorithm B*

$$K_A(x^n \mid n) \leqq K_B(x^n \mid n) + c_B$$

where c_B is a constant independent of x^n and n. A proof of this is given in Kolmogorov [3]. As a corollary, we have

$$| K_A(x^n \mid n) - K_B(x^n \mid n) | \leqq c$$

for any two universal algorithms A and B, where c is dependent only on A and B. Because the complexity of x^n with respect to an algorithm A is sensitive to the choice of A only up to a constant, we follow the convention of Martin-Löf by choosing a fixed universal algorithm as standard and dropping notational reference when convenient; thus, we write $K(x^n \mid n)$ for $K_A(x^n \mid n)$.

A secondary corollary of the stated key property is the existence of a constant c such that

$$K(x^n \mid n) \leqq n + c$$

for all x^n. This follows from the existence of an algorithm A such that $A(x^n, n) = x^n$. By a counting argument we derive the other basic relationship: sequences of length n for which

$$K(x^n \mid n) < n - c$$

are less than 2^{n-c} in number.

If i, j are positive integers then by $x^i < x^j$ we mean that $i \leqq j$ and x^i consists of the first i bits of x^j. x^i is then the *i prefix* of x^j. We now introduce the modified conditional complexity which we shall label the uniform complexity.

The *uniform complexity* of x^n with respect to A is given by

$$K_A(x^n; n) = \min_{D(A,x^n)} l(p),$$

where $D(A, x^n) = \{p \mid A(p, i) = x^i, x^i < x^n, \text{all } i \leqq n\}$ if $D(A, x^n)$ is non-empty; otherwise,

$$K_A(x^n; n) = \infty.$$

This type of "uniformity" condition is a frequent condition in mathematical definitions. Before discussing the interpretation of this modification, we note several basic properties.

The key property, the existence of a universal algorithm A such that for any algorithm B

$$K_A(x^n; n) \leqq K_B(x^n; n) + c_B,$$

where c_B is independent of x and n, is still valid. The essential point of Kolmogorov's proof of this property for conditional complexity measures is the existence of a universal algorithm A which can simulate any algorithm B given a proper "translation" program. Thus to a sequence (or program) p such that $B(p, n) = x^n$ there corresponds a sequence of form bp, that is, a (binary) sequence b followed by p, such that $A(bp, n) = x^n$. Here b depends only on A and B. Because b does not depend on p or n, there is a fixed sequence q, which is bp, such that $B(p, i) = A(q, i) = x^i$, $i \leqq n$ so $q \in D(A, x^n)$ if $p \in D(B, x^n)$ and $l(q) \leqq l(p) + c_B$.

Again, it follows as a corollary that the choice of universal algorithm alters the uniform complexity measure by at most an added constant. From the manner of construction of a universal algorithm just outlined, it is clear that an algorithm universal with respect to both the conditional complexity and the uniform complexity exists. We choose one such algorithm as standard for both complexity measures. Again, in the complexity notation we will in general omit reference to the underlying algorithm. Because the underlying universal algorithm is the same for both the conditional and uniform complexity measures, the inequality

$$K(x^n \mid n) \leqq K(x^n; n)$$

holds. This follows from the definitions. Moreover, there exists a constant c such that $K(x^n; n) \leqq n + c$ for all x^n; this follows as before as the pertinent program satisfies the uniformity condition. The analog to the last basic property stated for conditional complexity, that less than 2^{n-c} sequences x^n have $K(x^n - n) < n - c$, follows from the same property for conditional complexity using the above inequality.

Each of the three variant measures expresses a slightly different quality of the sequence x^n in assessing its information content. The quantity $K(x^n)$, measured with respect to some universal algorithm, gives the (minimum) length of programs for x^n which must contain in addition to the distribution of characters, here 0's and 1's, in x^n also information about the length n. The integer n can generally be expected to use about length $\log_2 n$ of the binary sequence p which is a "program" for x^n. However, for values of n quite easy to compute the requirement is much less. (Hereafter, we write $\operatorname{Log} n$ for $\log_2 n$.) The quantity $K(x^n \mid n)$ reports the minimum length of a program which need not contain information on the length n but "merely" determine the distribution of 0's and 1's in x^n. $K(x^n)$ has appeal intuitively because it reports the "entire information content" necessary to generate x^n (with respect to the given standard algorithm). It must then suffer this property as a disadvantage when concern centers on the distribution (of 0's and 1's) and when for reasons determined by context the length n can be assumed known. This distinction is dramatic at the low complexity end of the scale where the information needed to determine the distribution is less than $\operatorname{Log} n$. Comparison between x_1^n and x_2^n as to distributions may then be lost by $K(x^n)$ in the need to "report" that each is of length n. This aspect should be far less troublesome in the higher complexity region. Also, for mathematical reasons it is interesting to consider infinite sequences, either for their own sake or as an "approximation" to very large finite sequences. If we denote by x an infinite binary sequence and let x^n denote its n prefix, then we can discuss the information content of x using the measures introduced by associating with the information content of x the function $K(x^n)$ [$K(x^n \mid n)$ or $K(x^n; n)$, respectively] viewed as a function of n. In this instance we are clearly interested solely in the distribution of 0's and 1's (the "pattern") in any n prefix. The measure $K(x^n \mid n)$ is then preferable to $K(x^n)$ in this instance. It should be made clear that no stand is being taken here as to whether the length of a sequence is or is not to be an integral part of the information content associated with the sequence. This seems a matter of context.

If we agree that our interest is in measuring the *pattern* of 0's and 1's *without concern for the length of the pattern*, then we should choose between $K(x^n \mid n)$ and $K(x^n; n)$. Here the latter has some distinct advantages. Recall that the existence of a program p such that $A(p, n) = x^n$ is sufficient to assure $K(x^n \mid n) \leqq l(p)$. But here p may make heavy use of n in generating the pattern as well as determining the length of x^n. For example, there is clearly a program p_0 such that $A(p_0, n) = \bar{n}00 \cdots 0 = x^n$ where $\bar{n}$ is the binary expansion of integer n. The string of following 0's has length determined to meet the length requirement for x^n. Such a sequence we shall call an n string. For example, the n string for $n = 5$ is 10100. All n strings have an upper bound of $l(p_0)$, a constant, on their conditional complexity measure although the patterns are intuitively becoming more complex. (This is more dramatic if instead of trailing zeros, x^n consists of $\bar{n}$ iterated sufficiently often to fill the length requirement. A program p' exists which accomplishes this.) This measure also has the somewhat counter intuitive property that a given *pattern* may be judged considerably more complex than the same pattern followed by a finite sequence of 0's. For choose any sequence x^n with $K(x^n \mid n) \gg l(p_0)$ such that the first bit of x^n is a 1. Then $x^n = \bar{r}$ for some integer r. The r string $x^r = \bar{r}00 \cdots 0$ satisfies $x^n \prec x^r$ and $K(x^r \mid r) \ll K(x^n \mid n)$. (This should not be confused with "density of information" measures such as K/n, where K is the complexity of a sequence of length n. It is to be expected that the information density of x^r should be less than that of x^n if x^r is an extension by 0's of x^n).

The uniform complexity avoids these characteristics in an obvious way. If $p \in D(A, x^n)$ then $A(p, i) = x^i \prec x^n$, $i < n$, which assures that n explicitly influences the pattern itself for at most the last bit. Thus n as the second argument in $A(p, n)$ serves solely to determine the length of the sequence. It is clear that no single program that generates all n strings is acceptable for the uniform complexity measure. We show below that no constant c exists such that $K(x^n; n) \leqq c$ for all n strings x^n. Also, if $n < r$ and $x^n \prec x^r$ then $K(x^n; n) \leqq K(x^r; r)$. The objections mentioned previously are thus not valid here. To summarize, when concern is centered on the pattern (shape) exclusively, the interpretation of the uniform measure has some desirable properties not found in the interpretation of the conditional measure of complexity.[1]

Let us look further at the behavior of $K(x^n \mid n)$ and $K(x^n; n)$ when x^n is of "low" complexity. It is almost immediate that *if there exists a constant c such that for every n prefix x^n of an infinite sequence x, $K(x^n; n) \leqq c$ then x is effectively computable (recursive)*. It suffices to show there exists a single program p_1 such that for any n $A(p_1, n) = x^n$ where $x^n \prec x$. There are less than 2^{c+1} programs p such that $l(p) \leqq c$. Thus, for each n we know that $A(p, i) = x^i \prec x^n$, all $i < n$ for some program p of a finite number of programs. Then at least one program p' satisfies $A(p', i) = x^i \prec x^n$, all $i < n$ for infinitely many n, hence for all n. Then $p' = p_1$. Thus we also know $l(p_1) \leqq c$.

The statement above also holds for $K(x^n \mid n)$ although in general $l(p_1) \nleqq c$ for the desired program p_1. The last Section gives a proof due to A. R. Meyer that the statement holds for $K(x^n \mid n)$.

The justification above of the statement concerning $K(x^n; n)$ yields a stronger statement. If there exists a constant c such that *for infinitely many n prefixes x^n of an infinite sequence x, $K(x^n; n) \leqq c$ then x is effectively computable (recursive)*. This does not hold for the measure $K(x^n \mid n)$. In fact, we have the following Theorem.

There exists a constant c such that the set of infinite sequences x for which $K(x^n \mid n) \leqq c$ for infinitely many n prefixes x^n has the cardinality of the continuum.

The theorem is proven by providing a 1–1 map from the subsets of the positive integers to the infinite sequences x such that $K(x^n \mid n) \leqq c$ infinitely often. We choose $c = l(p_0)$, i.e., the length of a program which generates all n strings. Let $^*: N \to X$ be the function that maps integer n

[1] For a further comparison of the uniform and conditional complexities see Loveland, D. W. "On minimal-program complexity measures", Conference Record of the *ACM Symposium on theory of computing*, May, 1969. It is observed there that the conditional complexity measure enjoys a symmetry property not valid for the uniform complexity measure, a desirable property in some instances.

into the binary sequence formed by placing a 1 to the right of every symbol in the binary expansion of n. For example, $2 = 1101$. If $b \in X$ and the first bit of b is a 1, let (b) denote the n string formed by adding a correct number of 0's to the right of b. Thus, $(2^*) = 1101000000000$. Let $P \subseteq N$. Let $m_1, m_2, \cdots$ be an enumeration of the elements of P. If P is finite let m_k be the last element enumerated. For P infinite form the infinite sequence $\cdots (((m_1{}^*)m_2{}^*)m_3{}^*) \cdots$ where $(a)b$ is the concatenation of sequences (a) and b. This infinite sequence is uniquely associated with P as the numbers m_i may be determined from the sequence by noting that consecutive 0's act as spacers between subsequences denoting the $m_i{}^*$. If P is finite the finite sequence $(\cdots((m_1{}^*)m_2{}^*) \cdots)m_k{}^*)$ is formed and the desired infinite sequence then defined to be this finite sequence followed by a sequence of 0's. In each case, P finite or infinite, $K(x^n \mid n) \leqq c$ holds for infinitely many n prefixes x^n.

We now establish a statement made earlier and derive an interesting corollary from the method of proof.

For no constant c does $K(x^n; n) \leqq c$ for all n strings x^n. Choose $k \geqq 2$. Consider the set W of n-strings with $2^{k-1} \leqq n < 2^k$. Note that n has a k-bit binary expansion within this range. Let $S \subset W$ denote the subset of n strings which also have a 1 in the k-th bit; S has 2^{k-2} members. However, no program acceptable for the uniform complexity measure (acceptable programs are hereafter called "uniform programs") can compute two members of S. In order for a uniform program to compute two members of S one member of S must be a prefix of another member of S which is impossible. Thus, 2^{k-2} distinct uniform programs are needed to express the members of S. As the total number of (binary) sequences of length less than $k - 2$ is less than 2^{k-2} at least one program of length $k - 2$ is needed to express a member of S. But k was chosen arbitrarily. This establishes the statement to be proven. As a corollary, we have the following statement.

There exists a constant c such that for infinitely many n

$$K(x^n; n) - K(x^n \mid n) > \operatorname{Log} n - c$$

for some sequence x^n.

For each k chosen for the above argument, a (different) n string of length between 2^{k-1} and 2^k is obtained which requires a uniform program of length at least $k - 2$. This infinite collection of n-strings are the x^n's which satisfy the corollary. Recall that $K(x^n \mid n) \leqq c$ for an appropriate c for any n-string x^n.

This is as strong a divergence between these measures as one could expect for if there is a program p such that $A(p, n) = x^n$ where p uses n, a uniform program q can be built from p which includes the integer n as information. This adds at most approximately Log n to the length of p.

2. We now shift our attention to the characteristics of the uniform complexity measure in the region of high complexity. Much is known about the qualities of sequences x^n whose conditional complexity is high from the work of Martin-Löf [4]. Such complex sequences exhibit properties of "randomness". We show that sequences of sufficiently high uniform complexity also share many properties associated with randomness by showing that the technique developed by Martin-Löf for establishing this quality for sequences of high conditional complexity carries over to the uniform case. In doing so, we establish that *for every* $c > 0$ *there is a* $c_1 > 0$ *such that*

$$\{x^n \mid K(x^n; n) \geqq n - c\} \subseteq \{x^n \mid K(x^n \mid n) \geqq n - c_1\}$$

holds for all $n \in N$.

This is from one viewpoint an unexpected result. Recall the measures differ by approximately Log n for some sequences x^n at the low end of the complexity scale; at the high end of the complexity scale where a difference of Log n is small relative to the complexity of the sequence itself, the theorem asserts that the difference between $K(x^n; n)$ and $K(x^n \mid n)$ is no longer a function of n.

Before we give the proof of this theorem (which includes some remarks concerning the randomness properties associated with sequences x^n such that $K(x^n \mid n) \geqq n - c$), we mention another property of the measure $K(x^n; n)$. The property follows from a result of Martin-Löf [5]. Consider a recursive function f such that $\sum 2^{-f(n)} = +\infty$. Then, for all x, $K(x^n; n) \leqq n - f(n)$ for infinitely many n prefixes x^n of x. For example, $K(x^n; n) \leqq n -$ Log n for infinitely many $n \in N$. This property was shown by Martin-Löf for the complexity measure $K(x^n)$. However, we have $K(x^n; n) \leqq K(x^n) + c$ as for any algorithm A we have an algorithm B such that if $A(p) = x^n$ then $B(p, i) = x^i$, all $i \leqq n$. B merely acts as A and allows only the first i bits to "print out". So $K_B(x^n; n) \leqq K(x^n)$. But $K(x^n; n) \leqq K_B(x^n; n) + c$, some constant c. This constant is easily absorbed in the function f to give the result for the uniform measure. (Chaitin [2] establishes a similar but less strong result than that of Martin-Löf's used here.)

We will need the notion of a uniform test which is a special type of Martin-Löf test. We recall the definition of a Martin-Löf test.

A *Martin-Löf test* V is a subset of $N \times X$ with the following properties:

(1) V is recursively enumerable.

(2) $V_{m+1} \subseteq V_m$, where $V_m = \{x^n \mid (m, x^n) \in V, \text{all } n \in N\}$.

(3) The number of sequences of length n in V_m is $\leqq 2^{n-m}$.

If V is a Martin-Löf test and $x^n \in V_m$ then x^n is *terminal of class c at m*, $c = n - m$, if there does not exist a $y^{n+1} \in V_{m+1}$ such that $x^n < y^{n+1}$. Let $T_c(V)$ denote the set of terminal sequences in class c in V. Let $T_c^m(V)$ denote the set of terminal sequences in class c at r for $1 \leqq r < m$, in V. The number of sequences in a set S is notated ${}^{\#}S$ while ${}^{\#n}S$ denotes the number of sequences of length n in S. We now can write the added condition for a uniform test. A Martin-Löf test is a *uniform test* if:

(4) ${}^{\#n}V_m + {}^{\#}T_c^m(V) \leqq 2^{n-m}$ for $m \geqq 2$ where $c = n - m$.

Condition (3) for the Martin-Löf test, ${}^{\#n}V_m \leqq 2^{n-m}$, is subsumed by condition (4). This is immediate for all values of m except $m = 1$ for condition (3). The $m = 1$ case should become clear after consideration below of the intuitive meaning of the four conditions. It thus suffices to show conditions (1), (2) and (4) to establish a set of finite binary sequences is a uniform test.

Martin-Löf defined for any test V the set V_0 to be the set X. This is convenient when working with the critical levels. We do the same for the uniform test. This definition is compatible with the requirements of uniform test as condition (4) is void at $m = 0$. The definition of $T_c^m(V)$ (and $T_c(V)$) are not extended to include sequences terminal at $m = 0$.

Condition (4) is chosen so that uniform tests relate to the uniform complexity measure as Martin-Löf tests relate to conditional complexity. The condition is forced if a key theorem relating Martin-Löf tests and the conditional complexity measure is to be preserved. This condition along with conditions (2) and (3) can be illustrated by presenting a partial "picture" of a test V given by Fig. 1. Each box represents for a fixed n and m the set of x^n in V_m. V_m is the union of all boxes in the m-th row from the bottom; the set of x^n in V is given by the n-th column from the left. The number in the lower lefthand corner of each box gives the maximum number of sequences permitted in the box. This expresses condition (3). Condition (2) demands that any sequence in a box B appear also in the box below B. Condition (4) is a condition on the diagonal line

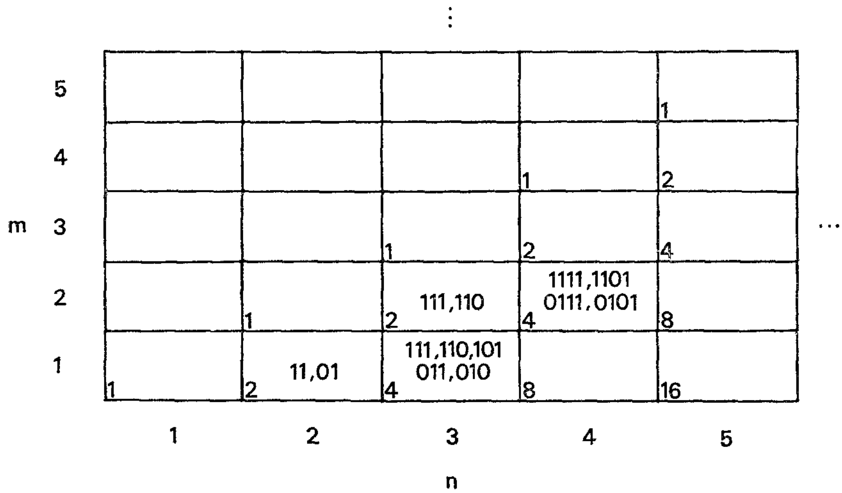

FIG. 1. A structure violating some conditions required of a uniform test.

of boxes containing the same left-hand corner number. A sequence x^n is a terminal sequence if the next higher diagonal box does not contain an extension of x^n. To satisfy condition (4), the number of sequences in a box plus the number of terminal sequences "below" it on the diagonal must not exceed the lefthand corner number. In Fig. 1, boxes $m = 1, n = 2$ and $m = 2, n = 3$ together present a violation of condition (4). Boxes $m=1, n=3$, and $m=2, n=4$ together also violate condition (4) and illustrate the reason condition (4) implies condition (3). The proof is left to the reader.

The properties shown to hold for Martin-Löf tests in Martin-Löf [4] also hold for uniform tests. Proofs of statements below are similar to the proofs for the corresponding results for Martin-Löf tests.

There exists a universal uniform test U such that for every uniform test V

$$V_{m+c} \subseteq U_m \qquad m = 1, 2, 3, \cdots,$$

where c is a constant (dependent on U and V).

Recall that the *critical level* function $m_V(x^n)$ is defined for each Martin-Löf test by

$$m_V(x^n) = \max_{x^n \varepsilon V_m} m,$$

for all x^n, $0 \leqq m_V(x^n) \leqq n$. The following fact follows directly from the theorem on universal tests.

If U is a universal uniform test then given uniform test V, there exists a c such that

$$m_V(x^n) \leqq m_U(x^n) + c.$$

If U and V are universal uniform tests then $|m_V(x^n) - m_U(x^n)| \leqq c$ for a suitable constant c, so the critical level function depends on the choice of universal uniform test only as to addition of a constant. We choose some universal uniform test as a standard and again suppress the indication of the test when the standard is used. We indicate that a uniform universal test is the standard by writing the critical level function as $\bar{m}$.

The important theorem that ties the conditional complexity measure to Martin-Löf tests carries over to the uniform case. The uniform tests are defined as such to enable this theorem to go through.

There exists a constant c such that $|n - K(x^n; n) - \bar{m}(x^n)| \leqq c$, *all binary strings* x^n.

In one direction, we define

$$V = \{(m, x^n) \mid K(x^n; n) < n - m\}.$$

It is easily checked that this defines a uniform test V. We then obtain

$$n - K(x^n; n) - \bar{m}(x^n) \leqq c_1,$$

for a suitable c_1, in the same manner as for Martin-Löf tests. For the converse inequality, we assume given the standard universal uniform test enumerated by a recursive function $f: N \to N \times X$. (The function f is total but not necessarily 1–1). We give the s-th stage in defining an algorithm A. The algorithm will be seen to be an effectively calculable function with domain a subset of $N \times X$.

s-th stage:

(1) Evaluate $f(k) = (m_k, x^{n_k})$, $k \leqq s$.

(2) If there exists a $k < s$ such that $x^{n_s} \prec x^{n_k}$ and $n_s - m_s = n_k - m_k$ then go to stage $(s + 1)$.

(3) If (2) does not hold and there exists a $k < s$ with $n_s - m_s = n_k - m_k$ and a program p such that $A(p, n_k) = x^{n_k} \prec x^{n_s}$ and $A(p, n_k + 1)$ has not yet been defined, then define $A(p, i) = x^i \prec x^{n_s}$ for $n_k < i \leqq n_s$.

(4) If (2) and (3) do not hold choose the first binary sequence p of length n_s–m_s not assigned at an earlier step and define

$$A(p, i) = x^i, 1 \leqq i \leqq n_s, \quad \text{where } x^i \prec x^{n_s}.$$

This completes the definition of algorithm A. Condition (4) of the uniform test is used to insure an available sequence p exists of length $n_s - m_s$ if step (4) of stage s requires such a p. It is here that the form of condition (4) is determined. It follows that

$$K_A(x^n; n) = n - \bar{m}(x^n),$$

or $K(x^n; n) \leqq n - \bar{m}(x^n) + c_2$, so $-c_2 \leqq n - K(x^n; n) - \bar{m}(x^n)$ for a suitable c_2. The theorem follows.

Let s_n denote the number of 1's in x^n. Let $f(m, n)$ be determined so that $|2s_n - n| > f(m, n)$ holds for less than 2^{n-m} sequences x^n but that $f(m, n)$ cannot be decreased without violating this condition. Martin-Löf uses the Martin-Löf test $V = \{(m, x^n) \mid |2s_n - n| > f(m, n)\}$ to establish that for sequences x^n with conditional complexity near n, $|2s_n - n|$ has a bound of the order of $\sqrt{n}$. More precisely, for an arbitrary constant c if $G_c = \{x^n \mid K(x^n \mid n) \geqq n - c\}$ then any sequence

$$x_1^{n_1}, x_2^{n_2}, x_3^{n_3}, \cdots$$

of members of G_c such that $n_i < n_{i+1}$ yields

$$\limsup_{i\to\infty} \frac{|2s_{n_i} - n_i|}{\sqrt{n_i}} \leqq k,$$

where k depends on c. We shall label this property of G_c the "weak central limit property."

Many other limit properties of probability theory such as the law of the iterated logarithm or von Mises' "impossibility of a successful gambling system" axiom can be shown valid for sequences of high conditional complexity by using the same technique as referenced above. It is certainly desirable, therefore, to establish the same property for sequences of high uniform complexity. We undertake this now. Let $H_c = \{x^n \mid K(x^n; n) \geqq n - c\}$. Then the following theorem holds.

For any constant c there exists a constant c_1 such that $H_c \subseteq G_{c_1}$.

We already have observed in Section 1 that $G_c \subseteq H_c$ as $K(x^n \mid n) \leqq K(x^n; n)$ under the assumption that the underlying universal algorithm is the same.

As an example of the use of this theorem we note that the weak central limit property holds for every H_c. In a similar manner all such statements that hold for the class of G_c sets also hold for the class of H_c sets.

The theorem is established by relating Martin-Löf tests to uniform tests. We might remark here that the Martin-Löf tests formed to establish individual properties are usually not uniform. For example, the test $V = \{(m, x^n) \mid |2s_n - n| > f(m, n)\}$ is not a uniform test. It does not seem that the formulation of (a) uniform test(s) from V is any easier in this singular case than the general method we consider in the proof we give now.

The basis of the proof is to represent any given Martin-Löf test V by an infinite collection of uniform tests $V(b)$ each of which faithfully represents a particular "subset" of V, namely V_{b+1}. The tests are then put together in a manner similar to constructing a universal test.

Given Martin-Löf test V we define *uniform test* $V(b)$ *defined from* V *with base* b, $b \geqq 1$, to be the uniform test satisfying

(1) $V_m(b)$ is empty if $m > b$.

(2) If $m \leqq b$ then $x^n \in V_m(b)$ if and only if $x^n \in V_{b+1}$.

We must justify that $V(b)$ is a uniform test. Condition (1): $V(b)$ is recursively enumerable as V is given recursively enumerable. Condition (2): to show $V_{m+1}(b) \subseteq V_m(b)$. This is immediate. Condition (4) [which implies condition (3)]: to show ${}^{\#n}V_m(b) + {}^{\#}T_c^m[V(b)] \leqq 2^{n-m}$, where $c = n - m$. We first consider $m \leqq b$. We have ${}^{\#n}V_m(b) \leqq 2^{-(b+1)}$. The number of sequences x^k terminal in class c at m' is $\leqq 2^{k-(b+1)}$. Thus

$$ {}^{\#}T_c^m[V(b)] \leqq \sum_{j=1}^{m-1} 2^{(n-j)-(b+1)} \qquad \text{where } m \geqq 2. $$

Putting both estimates together we have

$$ \begin{aligned} {}^{\#n}V_m(b) + {}^{\#}T_c^m[V(b)] &\leqq 2^{-(b+1)} \cdot \sum_{j=0}^{m-1} 2^{n-j} \\ &\leqq 2^{-(b+1)} \cdot 2^{n+1} \\ &\leqq 2^{n-b} \\ &\leqq 2^{n-m} \quad \text{for} \quad m \leqq b. \end{aligned} $$

For $m > b$, let us define k by $m - n = k - b = c$. Then ${}^{\#n}V_m(b) + {}^{\#}T_c^m[V(b)] = {}^{\#}T_c^m[V(b)] = {}^{\#k}V_b(b) + {}^{\#}T_c^b[V(b)] \leqq 2^{k-b} = 2^{n-m}$. Thus $V(b)$ is a uniform test.

We make use of the following Lemma.

For each universal uniform test U and each Martin-Löf test V there exists a monotone increasing function $c(m)$ such that

$$ V_{m+c(m)} \subseteq U_m, \quad \textit{for all } m. $$

To prove this Lemma we construct a uniform test V^* such that for any test V

$$V_{m+c_1(m)} \subseteq V_m^*,$$

where $c_1(m)$ is monotone increasing. However, for each universal uniform test U, there is a constant c such that $V_{m+c} \subseteq U_m$, all m. Putting these two inequalities together gives the Lemma. It, therefore, suffices to construct V^*.

First, a set T^*, a subset of $N \times N \times X$, is defined. Recall that Martin-Löf [4] proves a lemma stating there exists a recursively enumerable set $T \subseteq N \times N \times X$ such that V is a Martin-Löf test if and only if $V = \{(m, x^n) \mid (i, m, x^n) \in T\}$ for some $i = 1, 2, \cdots$. By V^i we shall mean the Martin-Löf test with index i as determined by T. Thus $V^1, V^2, \cdots$ is a (repetitive) effective enumeration of the Martin-Löf tests.

We consider a different effective enumeration of the V^i's for the definition of T^*. The rule of formation is: V^k occurs at the i-th place in the enumeration if and only if $i = (k-1) + 2^n$ where n is chosen such that $k \leqq 2^n$. The enumeration begins (indicating indices of V) 1, 1, 2, 1, 2, 3, 4, 1, $\cdots$, 7, 8, 1, 2, $\cdots$. However, in place of V^k at position i in the enumeration just given we enter uniform test $V^k(i+j)$, the uniform test defined from V^k with base $i + j$. Here, j is the number of occurrences of $V^k(b)$, any b, to (and including) position i. In particular, the first occurrence of a $V^k(b)$ is $V^k(i+1)$. We define

$$(i, m, x^n) \in T^* \Leftrightarrow (m, x^n) \in V^k(i+j),$$

where j and k are determined from i as stated above.

We now define V^* as follows:

$$(m, x^n) \in V^* \Leftrightarrow (i, i+m, x^n) \in T^* \text{ for some } i \geqq 1.$$

It is easily checked that conditions (1), (2) and (4) hold for V^* so V^* is indeed a uniform test. Because $V_{b+1}^k = V_b^k(b)$, using the definition of V^* (and T^*) it follows that for each $j > 0$, $k > 0$,

$$V_{i+j+1}^k = V_{i+j}^k(i+j) \subseteq V_j^*$$

for some i depending on j. For fixed k, i increases as j increases. We write $i(j)$ for i to emphasize dependency on j. Then the function $c(j)$ such that

$$V_{j+c(j)}^k \subseteq V_j^*,$$

is given by $c(j) = i(j) + 1$ which is monotone increasing in j. This establishes the Lemma.

Let V be the "standard" universal Martin-Löf test and let U be the "standard" universal uniform test. To complete the proof of the theorem we seek a relation between the critical levels $m(x^n)$ of V and $\bar{m}(x^n)$ of U for arbitrary x^n. Let $c(m)$ be a function such that

$$V_{m+c(m)} \subseteq U_m, \qquad \text{all } m,$$

assured by the above Lemma. Fix x^n. Let

$$m_0 = \max \{m \mid m + c(m) \leqq m(x^n)\}.$$

Then $x^n \in V_{m_0+c(m_0)}$ so $x^n \in U_{m_0}$ by the Lemma. Thus $\bar{m}(x^n) \geqq m_0$. Because

$$m_0 + c(m_0) \leqq m(x^n) < m_0 + 1 + c(m_0 + 1),$$

we have

$$m(x^n) \leqq \bar{m}(x^n) + c(m_0 + 1).$$

Now,

$$c[\bar{m}(x^n) + 1] \geqq c(m_0 + 1)$$

because $c(m)$ is an increasing function. We have

$$m(x^n) \leqq \bar{m}(x^n) + c[\bar{m}(x^n) + 1] = c_2(\bar{m})$$

for an appropriate increasing function $c_2(\bar{m})$ which depends only on $\bar{m}(x^n)$ (and on the choice of "standard" tests U and V).

Consider a given $H_c = \{x^n \mid K(x^n; n) \geqq n - c\}$. Using the relation $|n - K(x^n; n) - \bar{m}(x^n)| \leqq c_3$ for a suitable c_3, we observe that if $x^n \in H_c$ then $\bar{m}(x^n) \leqq c + c_3 = c_4$. Then $m(x^n) \leqq c_2(c_4)$ as $c_2(\bar{m})$ is increasing in $\bar{m}$. Now using $|n - K(x^n \mid n) - m(x^n)| \leqq c_5$, we have $n - c_2(c_4) - c_5 \leqq K(x^n \mid n)$, or, with $c_1 = c_2(c_4) + c_5$, $x^n \in G_{c_1} = \{x^n \mid K(x^n \mid n) \geqq n - c_1\}$. The Theorem is proved.

Martin-Löf [4] also introduces the notion of sequential test as an extension of the notion of (Martin-Löf) test for the purpose of defining "random sequence" as applied to infinite sequences. The sequential test differs from the Martin-Löf test by the addition of the condition that if $x^n \in V_m$ and $x^n < y^{n+i}$, $i = 1, 2, 3, \cdots$ then $y^n \in V_m$. The development of the calculus for sequential tests parallels that for Martin-Löf tests. A "standard" universal sequential test is chosen and the level $m(x^n)$ defined. We define the critical level $m(x)$ for an infinite sequence x as $m(x) = \lim_{n\to\infty} m(x^n)$ where x^n is the n prefix of x. Here the extended value ∞ is permitted in the range of $m(x)$. The limit is well-defined (in the extended sense) due to the added condition imposed on sequential tests. Martin-Löf defines an infinite binary sequence as random if $m(x) < \infty$.

The concept of uniform test can be extended to uniform sequential test by addition of the same condition. The critical level $\bar{m}(x^n)$ and $\bar{m}(x)$ are defined in the analogous manner. It may be shown by an argument directly parallel to the preceding work that $\bar{m}(x) = \infty \Leftrightarrow m(x) = \infty$. Thus, in the formal definition of infinite random sequence as given by Martin-Löf, it makes no difference if the sequential test or the uniform sequential test forms the basis of the definition.

3. We now give the proof of the Theorem on conditional complexity stated in Section 2. The proof given here is a modification of the proof originally given by A. R. Meyer.

If x is an infinite binary sequence for which there exists a constant $c > 0$ such that $K(x^n \mid n) \leqq c$, all n, then x is recursive.

Let A be the underlying universal algorithm for $K(x^n \mid n)$. Then A is a recursive function. By hypothesis, there exists a set of programs $p_1, \cdots, p_m$ such that for each $n > 0$, $A(p_i, n) = x^n$ for some $i \leqq m$, where $m < 2^{c+1}$. We let x^n denote the n prefix of the given sequence x. Also, we denote $A(p_i, n)$ by x_i^n if $A(p_i, n)$ is defined. We must prove there exists a single program p such that $A(p, n) = x^n$, all n. We construct (nonuniformly from $p_1, \cdots, p_m$) such a program p. First we need some definitions.

Let $k(b)$ denote the number of distinct x_i^b defined at b for $1 \leqq i \leqq m$. Let $k = \limsup_{b\to\infty} k(b)$. Let r be an integer chosen so that $r \leqq b$ implies $k(b) \leqq k$. Define $S = \{b \mid r \leqq b \text{ and } k(b) = k\}$. Then S is an infinite recursively enumerable set. We say programs $p_1', \cdots, p_k'$ *S-define y with prefix z^h* if for all $b \in S$ satisfying $b \geqq h$ there is an $i \leqq k$ such that $A(p_i', b) = y^b$ and $z^h \prec y^b$. The programs $p_1, \cdots, p_m$ clearly S-define x with prefix x^r.

We show that there exists a $t > 0$ such that programs $p_1, \cdots, p_m$ S-define only x with prefix x^t. For suppose $p_1, \cdots, p_m$ S-defines $d + 1$ distinct infinite sequences $x, y_1, \cdots, y_d$ with the empty prefix requirement. Then there exists a w such that x^s, y_1^s, $\cdots$ y_d^s are all distinct for all $s \in S$ such that $s > w$. Because only x satisfies $x^w \prec x$ we may choose $t = w$.

We now present an outline of the computation of $A(p, n)$, which determines the desired p. First, for $n < t$ set $A(p, n) = x^n$ where t is chosen so that $p_1, \cdots, p_m$ S-defines only x with prefix x^t. We enumerate S and show how to define $A(p, b)$ for each $b \in S$ not already determined. We note that if $A(p, b)$ is determined then so is $A(p, n)$ for all $n < b$. There are two possibilities for $A(p, b)$ with $b \in S$ and $b \geqq t$. If only one x_i^b satisfies (i) $x^t \prec x_i^b$ and (ii) for every $d \in S$ so far enumerated, either $x_i^b \prec x_j^d$ or $x_j^d \prec x_i^b$ for some j, $1 \leqq j \leqq m$, then $x_i^b = x^b$ and $A(p, b)$ is determined. The second possibility is that there exist more than one x_i^b satisfying the conditions just mentioned. But it is then simply necessary to keep enumerating S until condition (ii) holds for only one x_i^b such that $x^t \prec x_i^b$. This must occur for otherwise more than one infinite sequence will be S-defined with prefix x^t. This contradicts our assumption. Thus, x^b and $A(p, b)$ are determined. This concludes the proof.

ACKNOWLEDGMENT

The author would like to thank Albert Meyer for discussions which aided his understanding of the behavior of the complexity measures.

RECEIVED: February 24, 1969

REFERENCES

1. G. J. CHAITIN, (1966). On the length of programs for computing finite binary sequences, *J. ACM* **13**, 547–569.
2. G. J. CHAITIN, (1969). On the length of programs for computing finite binary sequences: statistical considerations, *J. ACM* **16**, 145–159.
3. A. N. KOLMOGOROV, (1965). Tri podhoda k opredeleniju ponjatija 'količestvo informacii', *Problemy Peredači Informacii* **1**, 3–11.
4. P. MARTIN-LÖF, (1966). The definition of random sequences, *Information and Control* **9**, 602–619.
5. P. MARTIN-LÖF, O kolebanii složnosti beskonečnykh dvoičnykh posledovatel'nostej, unpublished.