THE KLEENE HIERARCHY CLASSIFICATION OF RECURSIVELY RANDOM SEQUENCES(1)

BY
D. W. LOVELAND

Introduction. In [1], Church suggests a precise formulation for the notion of random sequence as conceived by von Mises. This paper introduces a modified formulation to avoid some objections inherent in the "classical" case (see Loveland [8] or [9]), and defines corresponding sets which are then located in the Kleene hierarchy of recursive unsolvability. More precisely, we locate a collection of classes in the hierarchy containing some of the defined sets, including the "lowest" class in which these sets may be found. We confine our attention to sequences of 0's and 1's and then make the natural correspondence between sequences and sets by associating with the infinite sequence $a_0, a_1, a_2, \ldots$ the set $\{i \mid a_i = 1\}$. (We only consider infinite sequences.)

Unlike the modern concept of "random sequence," the von Mises theory allows the label "random" or "nonrandom" to be applied to a specific sequence of outcomes of events, according as to whether or not the sequence has a given structure. In essence, if a sequence $A = (a_0, a_1, a_2, \ldots)$ of 0's and 1's has a certain limiting relative frequency p of 1's to number of places, then every (infinite) sequence, composed of members of the sequence A, which could represent a betting scheme played by someone attempting to "beat" the system, (i.e., achieve a different limiting relative frequency) must indeed result in the limiting relative frequency p. The intuitive concept of "betting scheme" is this: if we regard the indices of the given sequence as indicating the order of performance of the "events" (such as coin-tossing) and the corresponding sequence entries as the outcomes of these events, the "betting scheme" is some effective rule wherein the better observes outcomes of certain events and then uses this information to select an event on which to bet, the selection being made without knowledge of the outcome of that event. If the *selection rule* allows an infinite number of selections (bets) to be made over the infinite sequence, it is called a *proper selection rule*, otherwise it is *improper* with respect to that sequence. Only betting schemes that are proper selection rules shall be of concern in identifying random sequences. ("Effective rule" here means

Received by the editors April 1, 1966.

(1) This paper is, in essence, the final chapter of the author's doctoral dissertation presented at New York University. The author is indebted to Martin Davis for the suggestion of the problem and his subsequent guidance.

The author was supported by an NSF cooperative fellowship during the writing of the dissertation.

constructively executable and reproducible; a person who bets every time his ear itches would be following a rule executable but not reproducible.)

One such "betting scheme" for coin tossing would be to bet "heads" after each consecutive run of seven "tails." A more unorthodox scheme would permit the gambler to observe only every other trial (say, the even-numbered trials) allowing him to "return" at intervals to bet on the unobserved trials. The gambler might, for example, choose to bet "heads" on odd-numbered trials occurring in the middle of each run of six "heads" in the even-numbered trials observed.

Various properties were established concerning von Mises' random sequences under certain informal interpretations of the term "betting scheme," or selection rule (see e.g., Wald [14]). The problem of formalization centered on the concept of "all possible effective selection rules." In [1], Church proposes a solution to this problem and gives an explicit definition for randomness for sequences of 0's and 1's. The role of "effectively calculable" functions is assumed by recursive functions, a precisely defined class of "effectively calculable" functions which has proved to date to include all functions we are willing to call "effectively calculable." (The equivalence of the class of "effectively calculable" functions with the class of recursive functions is called *Church's thesis*. A discussion of this relationship appears in Davis [2, p. 10]; on p. 41 Davis gives a formal definition of the class of recursive functions. This paper will assume a familiarity with recursive function theory to the extent given in [2], when convenient, however, using the more conventional notation as appears in Rogers [13]. Certain basic properties of the Kleene hierarchy using this notation are stated in the next section.)

One reason for interest in sequences satisfying Church's definition is that if the class of effective selection rules is properly formalized using recursive functions, then we have at hand "models" of random sequences in the sense that we have sequences sharing properties possessed by "almost all" sequences (with respect to a suitable measure) as regards the impossibility of a successful gambling system (see Doob [4]; for the generalization treated here see Loveland [8], [9]). The question considered here can be then phrased: how "noneffective" (measured via the Kleene hierarchy) do these "models" have to be? Clearly, these "models" cannot be effectively enumerable for then one would use the enumerating rule to "place one's bets" and score perfectly. In a certain sense, however, these sequences are "almost effectively enumerable." This remark will be clarified below.

We present the formal definitions and results in the next section. Proofs are given in the final section.

Basic definitions and results. In defining the notion of random sequence to be used here (with the accompanying notion of random set), it is desirable to modify Church's definition to make the sequences more accurately reflect our notion of "randomness." (This weakness is not due to Church's formalization of the von

Mises system but was inherited from previous formulations. Church's formulation, for example, does not force sequences defined by selection rules of the type illustrated by the second example to give the same limiting relative frequency as the original sequence, although the type illustrated by the first example is controlled. See Loveland [8] or [9](2).) Unfortunately, the modification necessary adds to the complexity of the definition, due to the more complex basis for describing the general selection rule. The selection rule format is composed of two parts: (i) an *observed sequence* is enumerated which orders the trials as the better (or "game" format) dictates (as the observed trials may differ in order and/or selection from the original order of execution of the events); (ii) a final, or *betting*, subsequence of the observed sequence which is the sequence constrained in its limiting relative frequency. In selecting both the observed and betting sequences, the better has knowledge of the preceding members of the *observed* sequence (hence also knowledge of the preceding members of the betting sequence). The c_i, $i=0, 1, 2, \ldots$ used in the definition below encode the information available preceding the $(i+1)$th selection in the observed sequence; this includes the "index" numbers of previous trials observed (i.e., order of execution of the events) and the outcome of each observed trial. For clarity we use $a(i)$ for a_i in the definition. By "partial recursive function with adequate domain" we mean the domain of the partial function is some subset of the nonnegative integers which includes the integers c_i, $i=0, 1, 2, \ldots$.

Definition. An infinite sequence $a(0), a(1), a(2), \ldots$ of 0's and 1's is a *recursively random sequence* if the following conditions hold:

(1) If $f(r)$ is the number of 1's among the first r terms of $a(0), a(1), a(2), \ldots$ then $f(r)/r$ approaches a limit p, $0<p<1$, as r approaches infinity.

(2) Let $\phi(x)$ be any partial recursive function with adequate domain and $c_0=1$, $c_{n+1}=\Pr(n+1)^{b_n}\cdot c_n$ where $\Pr(x)$ is the xth prime number, and $b_n=2k_n+a(k_n)$. The k_n's are determined by $\phi(c_n)=k_n$. If the integers $k_0, k_1, k_2, \ldots$ form a nonrepetitive infinite sequence then the new sequence $a(k_0), a(k_1), a(k_2), \ldots$ (*the observed sequence*) must satisfy the following condition:

If $\psi(x)$ is any partial recursive function with adequate domain and the $a(k_n)$ such that $\psi(c_n)=0$ form an infinite subsequence $a(k_{n_0}), a(k_{n_1}), a(k_{n_2}), \ldots$ (*the betting sequence*) of the observed sequence, then if $g(r)$ is the number of 1's among the first r terms of $a(k_{n_0}), a(k_{n_1}), a(k_{n_2}), \ldots$ then $g(r)/r$ approaches the same limit p as r approaches infinity.

Definition. Given the recursively random sequence $a_0, a_1, a_2, \ldots$ the associated *recursively random set* S (or, simply, *random* set) is given by $S=\{i \mid a_i=1\}$.

Clearly, if S is a random set so is $\bar{S}$, the complement of S. (The "universe" is taken to be the nonnegative integers; thus, $\bar{S}=\{i \mid a_i=0\}$ for the sequence defining S.)

For the remainder of the paper we shall assume the limiting relative frequency

(2) *Added in proof.* A similar modification is independently noted and employed in A. N. Kolmogorov, *On tables of random numbers*, Sankhyā Ser. A **25** (1963), 369–376.

to be 1/2. Conversion from a recursively random sequence associated with 1/2 to a recursively random sequence associated with p, for p a recursive real number, may be effectively carried out given the original sequence. (See e.g., [3, p. 209].)

We shall report the results in two ways, via the Kleene hierarchy and by use of Putnam's *trial-and-error predicates* (see Putnam [12]). The latter are linked to the Kleene hierarchy and is the tool used in obtaining the hierarchy results. However, they have an independent appeal as a certain very restrictive type of noneffective process. We consider these notions and their relationship before stating our results.

The Kleene hierarchy of predicate forms classifies the "arithmetrical" sets in classes Σ_n, Π_n, $n=0, 1, 2, \ldots$ defined as follows: Σ_n is the class of all sets A of the form $A=\{(a_1, \ldots, a_m) \mid (Q_1x_1)(Q_2x_2)\cdots(Q_nx_n)P(a_1, \ldots, a_m, x_1, \ldots, x_n)\}$ where $P(a_1, \ldots, a_m, x_1, \ldots, x_n)$ is a recursive predicate, the Q_{2k+1} are existential quantifiers and the Q_{2k} are universal quantifiers. Π_n is the class of all sets A as above except that the Q_{2k+1} are universal quantifiers and the Q_{2k} are existential quantifiers. We summarize some basic properties of this hierarchy of sets (see Rogers [13]):

(1) $\Sigma_0=\Pi_0=\Sigma_1 \cap \Pi_1=$ the collection of all recursive sets;
(2) $A \in \Sigma_n \Leftrightarrow \bar{A} \in \Pi_n$, all n;
(3) $\Sigma_n \subset \Sigma_{n+1}$, $\Pi_n \subset \Sigma_{n+1}$, $\Sigma_n \subset \Pi_{n+1}$ and $\Pi_n \subset \Pi_{n+1}$, all n;
(4) $\Sigma_n \not\subset \Pi_n$ and $\Pi_n \not\subset \Sigma_n$ for $n>0$;
(5) $\Sigma_n \cup \Pi_n \subset \Sigma_{n+1} \cap \Pi_{n+1}$ for $n>0$ and containment is proper.

We now define the notion of trial-and-error predicate. We include here an immediate generalization of the concept of Putnam's which is useful in the following section.

DEFINITION. The predicate "$\lim_{j\to\infty} Q(i,j)$" is *well defined* if for each i there exists a $j_0(i)$ such that either $Q(i,j)$ is true for all $j \geqq j_0(i)$ or $Q(i,j)$ is false for all $j \geqq j_0(i)$. The value of the predicate is $Q(i, j_0(i))$.

DEFINITION. $P(i)$ is a *trial-and-error predicate over* $Q(i,j)$ if $P(i) \equiv \lim_{j\to\infty} Q(i,j)$ is well defined. $P(i)$ is a *trial-and-error predicate* if $Q(i,j)$ is recursive.

Let $Q(i,j)$ be a recursive predicate. Then its truth-value is effectively calculable. If we regard $Q(i,j)$ as the jth stage in the process of determining the value of $P(i)$, then $P(i)$ is "eventually determined," *but at a time* (*stage*) *not necessarily effectively determinable*. That is, for each i, from some (unpredictable) stage onward, $P(i)$ has been determined by a given effective process. One might regard this as an "asymptotically effective" process. (For a fuller discussion of this type of predicate see Putnam [12].)

The following lemma is a direct extension of an observation of Putnam's and links the two notions defined above. Its proof, as well as the theorems following, appears in the next section.

LEMMA (PUTNAM). *If* $Q=\{(i,j) \mid Q(i,j)\}$ *and* $Q \in \Sigma_n \cap \Pi_n$, $n=1, 2, 3, \ldots,$ *if* $P(i)$ *is a trial-and-error predicate over* $Q(i,j)$ *and* $P=\{i \mid P(i)\}$ *then* $P \in \Sigma_{n+1} \cap \Pi_{n+1}$. *In particular, if* $P(i)$ *is a trial-and-error predicate then* $P \in \Sigma_2 \cap \Pi_2$.

Our results may now be stated:

THEOREM 1. *There is a trial-and-error predicate which defines a recursively random set (i.e., there exists a trial-and-error predicate* $P(i)$ *such that* $P=\{i|P(i)\}$ *is a recursively random set).*

THEOREM 2. *There exist recursively random sets properly in* $\Sigma_i \cap \Pi_i$, *for* $i=2, 3, 4, \ldots$ *(i.e., in* $\Sigma_i \cap \Pi_i$ *but not in* $\Sigma_{i-1} \cup \Pi_{i-1}$).

For completeness, we present a result of Markwald [10] concerning the Mostowski field of sets, i.e., the class of sets containing Σ_1 and closed under a finite number of applications of complement, union, and intersection. Markwald proves that for every set M in the Mostowski field, M or $\overline{M}$ contains an infinite subset in Σ_1, which yields immediately the following theorem.

THEOREM 3 (MARKWALD). *There exists no recursively random set in the Mostowski field of sets.*

It remains an open question as to whether or not every class of the Kleene hierarchy containing $\Sigma_2 \cap \Pi_2$ "properly" contains a recursive random set (i.e., whether in addition to the above result, random sets exist in the class Σ_n but not Π_n, $n \geqq 2$, and conversely).

Proofs of stated results. We first prove the lemma.

Proof (of Lemma). It is convenient to establish the result for $Q(i,j)$ recursive first. Let $P=\{i|P(i)\}$.

(1) $P(i) \equiv (\exists y)(z)[z<y \vee Q(i, z)]$ so $P \in \Sigma_2$.

Given that $P(i)$ is a trial-and-error predicate, we also have

(2) $P(i) \equiv (y)(\exists z)[z>y \text{ and } Q(i, z)]$ so $P \in \Pi_2$.

If $Q=\{(i,j)|Q(i,j)\} \in \Sigma_n \cap \Pi_n$, then

(3) $Q(i,j) \equiv [k]R(i, j, u_1, \ldots, u_n)$ and also

(4) $Q(i,j) \equiv [m]S(i, j, u_1, \ldots, u_n)$

where $[k]$, $[m]$ represent strings of n alternating quantifiers, the former beginning with an existential quantifier, the latter beginning with a universal quantifier, and where $R(i, j, u_1, \ldots, u_n)$ and $S(i, j, u_1, \ldots, u_n)$ are recursive predicates. Using (4) for $Q(i, z)$ in (1) bringing forward the quantifiers and combining the adjacent universal quantifiers (see Rogers [13] for technique), it is seen that $P \in \Sigma_{n+1}$. Likewise, replacing $Q(i, z)$ by its equivalent form of (3) in (2), bringing forward the quantifiers and combining adjacent existential quantifiers yields $P \in \Pi_{n+1}$. Q.E.D.

The main task is the demonstration of the existence of a recursive predicate $Q(i,j)$ yielding the necessary trial-and-error predicate $P(i)$ for Theorem 1. A slight modification then yields Theorem 2. We choose to define $Q(i,j)$ by describing an effectively calculable process with the desired properties and then invoking Church's thesis to establish that the process defining $Q(i,j)$ is represented by some recursive function. We present an infinite string of sequences $A_0, A_1, A_2, \ldots$, where $A_j=a_j(0)$,

$a_j(1), a_j(2), \ldots$ is a sequence of 0's and 1's, and define $Q(i,j) \equiv a_j(i) = 1$, for $i, j = 0, 1, 2, 3, \ldots$. Thus the condition we seek is that for each i, $\lim_{j\to\infty} a_j(i) = 0$ or $\lim_{j\to\infty} a_j(i) = 1$; intuitively, the sequences can be viewed as successive approximation to the recursively random sequence (and likewise for the corresponding sets).

Notation. The observed sequence associated with sequence A (of 0's and 1's) will be denoted by $A(g)$, where g is (the Gödel number of) the recursive function which determines $A(g)$ from A. Likewise, $A(g, h)$ denotes the betting sequence derived from A via $A(g)$, with h the (Gödel number of) the recursive function which determines $A(g, h)$ from $A(g)$. Thus $A(g) = a(k_0), a(k_1), a(k_2), \ldots$ and $A(g, h) = a(k_{n_1}), a(k_{n_2}), a(k_{n_3}), \ldots$ where the k_i's and k_{n_i}'s are determined as given in the definition of recursively random sequence.

$A(\ ; m)$, where m is a positive integer, shall denote the sequence A truncated after m terms, i.e., $a(0), a(1), \ldots, a(m-1)$. $A(g; m)$ denotes the maximum observed sequence derivable from $A(\ ; m)$, that is, $a(k_0), a(k_1), \ldots, a(k_j)$ where $k_{j+1} \geqq m$, and $k_0, k_1, \ldots, k_j < m$. $A(g, h; m)$, likewise denotes the maximum initial sequence of $A(g, h)$ derivable from $A(g; m)$. $M(S, p)$ shall denote the initial p terms of sequence S (S not necessarily a binary sequence), then $M(A, p) = A(\ ; p)$ for binary sequence A.

Finally, let $I(A(\ ; m))$ denote the finite string of indices of the truncated sequence, e.g., $I(A(g; m)) = k_0, k_1, \ldots, k_j$.

Let g, h be arbitrary, given positive integers. Let us denote by $g(x)$ and $h(x)$ the partial recursive functions defined when g and h are regarded as Gödel numbers of functions. (When g or h is not a Gödel number, the corresponding functions are considered as functions over the empty domain.) Given a sequence A (of 0's and 1's), the domain of $g(x)$ (resp. $h(x)$) may not include the set $\{c_0, c_1, c_2, \ldots\}$ (see definition of recursively random sequence) needed to define $A(g)$ or $A(g; m)$ (resp. $A(g, h)$ or $A(g, h, m)$), or $g(x)$ may not be 1-1 over the domain required. $A(g)$ and $A(g; m)$ are then calculated to the first irregularity in the procedure, and the resulting finite sequence is defined to be $A(g)$ and $A(g; m)$ respectively. (The same holds for $A(g, h)$ and $A(g, h; m)$.)

By a binary string we mean a finite sequence of 0's and 1's. We also use the term "sequence" for finite as well as infinite sequences.

Recall that n binary strings of length p comprise the fraction $n \cdot 2^{-p}$ of the total number of such binary strings. The following lemma asserts that, given n binary strings of length p, the fraction of sequences $A(\ ; m)$, for any $m \geqq p$, which map by a given selection rule into sequences $A(g, h; m)$ having at least p elements and such that the initial p elements form one of the given binary strings, is $\leqq n \cdot 2^{-p}$. The corollary considers the situation when binary strings are juxtapositioned.

LEMMA. *Let g, h be any two positive integers. Given positive integer p, for any $m \geqq p$ there are $\leqq 2^{m-p}$ sequences $A(\ ; m)$ such that $M(A(g, h, m), p) = b$, a given binary string of length p.*

Proof. Let $g(x)$, $h(x)$ be the corresponding partial recursive functions. The proof is by induction.

(1) $p=1$.

Assume $b=1$, to show that $\leqq 2^{m-1}$, or not more than 1/2 of the sequences of form $A(\ ;m)$, give $M(A(g,h;m),1)=1$. We show for each $A(\ ;m)$ such that $M(A(g,h;\ m),1)=1$ there is a sequence $B(\ ;m)$ such that $M(B(g,h;m),1)=0$; namely, if $M(I(A(g,h;m)),1)=k$ then define $b(i)=a(i)$, $i=0,1,\ldots,m-1$, $i\neq k$, and $b(k)=0$ (recall $b=1$ states that $a(k)=1$).

$B(\ ;m)$ is clearly the desired sequence as the selection of k by $A(g,h;m)$ depends only on $A(g;m)$ and $I(A(g;m))$ up to the point of selection of k as the first member of $I(A(g,h;m))$. But $B(g;m)$ and $I(B(g;m))$ are identical with their counterparts over the critical range. The case $b=0$ is analogous.

(2) Assume true for $p=n$, to show true for $p=n+1$. By induction hypothesis, it is known that there are $\leqq 2^{m-n}$ sequences $A(\ ;m)$ such that $M(A(g,h;m),n)=\bar{b}$, where $\bar{b}$ denotes the binary string of length n obtained by deleting the last "digit" of b. Denote this class of sequences $A(\ ;m)$ by $\mathscr{A}$. It is sufficient to show that not more than one-half of the sequences $A(\ ;m)$ of $\mathscr{A}$ produce string b. This is established as before by a 1-1 correspondence within $\mathscr{A}$. If $b=\bar{b}1$, then $B(\ ;m)\in\mathscr{A}$ is formed by $a(i)=b(i)$, $i=0,1,\ldots,m-1$, $i\neq k$ and $b(k)=0$, where k is the $n+1$th (or last) term of $M(I(A(g,h;m)),n+1)$. As before, as no other term than $a(k)$ is altered to form $B(\ ;m)$, $b(k)$ must be selected by $B(g,h;m)$ at the corresponding point that $A(g,h;m)$ selects $a(k)$ so $M(B(g,h;m),n+1)=\bar{b}0$. The case $b=\bar{b}0$ is analogous. Q.E.D.

COROLLARY. *Let U, V be sets of binary strings of length s, t respectively having x, y members respectively. Let $p=x/2^s$ and $g=y/2^t$. Define the prefix of w, where w is a binary string of length $s+t$, to be the leftmost string of s symbols of w, and the suffix of w to be the rightmost string of t symbols of w. Let r be the fraction of sequences $A(\ ;m)$ such that $M(A(g,h;m),s+t)\neq w$ where w has either prefix $u\in U$ or suffix $v\in V$. Then $r\geqq(1-p)(1-q)$, all m.*

Proof. If $m<s+t$ then $r=1$ as $M(A(g,h;m),s+t)$ must be a string of less than $s+t$ members.

For $m\geqq s+t$, the equivalent statement $1-r\leqq 1-(1-p)(1-q)=p+q-pq$ will be established. If z is the number of sequences $A(\ ;m)$ which have map w where w has prefix $u\in U$ or suffix $v\in V$ then $1-r=z/2^m$. The number of binary strings of length $s+t$ of form w is $x\cdot 2^t+2^s\cdot y-x\cdot y$. By the lemma

$$z \leqq 2^{m-(s+t)}(x\cdot 2^t+2^s\cdot y-x\cdot y),$$

$$1-r \leqq 2^{-(s+t)}(x\cdot 2^t+2^s\cdot y-x\cdot y),$$

$$1-r \leqq x\cdot 2^{-s}+y\cdot 2^{-t}-x\cdot y\cdot 2^{-s}\cdot 2^{-t} = p+q-pq.$$

Q.E.D.

The basic idea of the algorithm is to discard from consideration collections of binary sequences, each collection of sequences specified by a common initial segment, if the initial segment is transformed by some selection rule into a binary string having too many 0's or 1's. Discarded sequences are never reconsidered as candidates for the random sequence. Larger and larger initial segments are considered with respect to more and more selection rules with care taken never to let the total of discarded sequences become too large a fraction of the total collection of binary sequences (in the obvious measure sense). The approximations to the random sequence are then taken from the collection of still eligible sequences at appropriate times in the process.

To construct the algorithm outlined, we use a procedure which, for a given selection rule, discards less than a specified fraction (dependent on the selection rule) of the total collection of binary sequences and yet leaves only those sequences which transform under the selection rule to betting sequences having the same limiting relative frequency as the original sequence, i.e., limiting frequency 1/2. This procedure is now considered.

Let us consider binary strings of length m. If we decide to discard those sequences whose initial segment of length m map into given betting sequences of length $p \leqq m$, then by the lemma we discard no larger fraction of the total number of sequences considered than the ratio of the number of specified sequences of length p to the total number of sequences of length p. (Hereafter we write "discard the initial segment b" for "discard the sequences with initial segment b.") We now specify n_1 binary strings of length x_1, n_2 strings of length $x_2, \ldots, n_k$ strings of length x_k, and discard any initial segment whose betting sequence when split into successive sections of length $x_1, x_2, \ldots, x_k$ respectively, has any one section match a binary string specified for that section. The corollary of the lemma (iterated a sufficient number of times) states that for fixed m, the fraction of initial segments of length m retained is at least as large as the fraction of binary strings of length $x_1+x_2+\cdots+x_k$ retained as acceptable initial segments for the betting sequence, namely

$$\prod_{i=1}^{k} (1-n_i \cdot 2^{-x_i}).$$

We shall discard initial segments whose betting sequences have a section containing too many 0's or 1's; we have just noted that we need only consider the fraction of binary strings of length $x_1+x_2+\cdots+x_k$ not violating the given conditions, as this fraction will never exceed the fraction of initial segments retained. Thus we can consider the discard problem for binary strings with no transformation problems involved.

Let m_i denote the length of the first i sections of a binary string, i.e.,

$$m_1 = x_1+x_2+\cdots+x_i.$$

It will develop that the fraction of strings to be discarded determines a positive number t_0 such that $x_i=(t_0+t)^6$. The discarded strings are those whose ith section has s_i or fewer 0's or 1's, where $s_i=[x_i/(2+1/t)]$ with $t=t_0+i$. $[a]$ denotes the integral part of a. Note that $s_i<x_i/2$ but that $s_i/(x_i/2)\to 1$ as $i\to\infty$. A conventional argument shows that once t_0 is chosen the binary strings which are not discarded at some point approach the limiting frequency 1/2 uniformly, i.e., if $f(r)$ counts the number of 1's in the first r places of the binary string, then, given t_0, for any $\varepsilon>0$ there exists an N such that

$$\left|\frac{t(r)}{r}-\frac{1}{2}\right| < \varepsilon$$

holds for all $r>m_N$, where N is independent of the binary string chosen. (Of course, r is not confined to the values m_i hence the counting function $f(r)$ may stop counting in the middle of an interval, but this causes no fundamental problem.) The choice of x_i and s_i will now be justified and the method given for computing t_0 from a given lower bound for the fraction of binary strings to be retained.

Let $b(m,n)=n!/(m!(n-m)!)$, the number of binary strings of length n that have exactly m 0's. Then it can be shown (see Feller [5, p. 140, equation 3.6], where we use $b(p,n)$ for Feller's $b(p;n,1/2)$) that

$$\sum_{p=0}^{s} b(p,n) < b(s,n)\frac{n-s+1}{n+1-2s} \qquad \text{when } s \leqq n/2.$$

Let $k=n/s$ so $s=n/k$ and $k>2$, then

$$\begin{aligned}\sum_{p=0}^{s} b(p,n) &< b(s,n)\frac{n-n/k+1}{n+1-2(n/k)}\\ &< b(s,n)\left(\frac{n((k-1)/k)+1}{n\cdot((k-2)/k)+1}\right)\\ &< b(s,n)\left(\frac{k-1}{k-2}\right).\end{aligned}\tag{1}$$

Thus the number of strings of length n with less than or equal to s 0's is bounded in terms of the number of such strings with exactly s 0's, as long as $s<n/2$, as is true for 1's in place of 0's.

For n even, clearly $b(s,n)\leqq b(n/2,n)$, $s\leqq n/2$. By Stirling's formula,

$$\frac{b(n/2,n)}{2^n} \sim (2\pi)^{-1/2}\left(2^{-n}\frac{n^{n+1/2}}{(n/2)^{n+1}}\right) = \left(\frac{2}{\pi n}\right)^{1/2}. \tag{2}$$

The relations $(2\pi)^{1/2}n^{n+1/2}e^{-n}<n!$ of Stirling's holds for all n and the percentage error decreases, with increasing n, monotonically to 0 (see Feller [5, pp. 50–52]). These two facts with (2) yield

$$2^{-n}b(n/2,n) < \left(\frac{2}{\pi n}\right)^{1/2}, \qquad n \text{ even}. \tag{3}$$

Combining (1) and (2), for n even, and $s<n/2$,

$$2^{-n}\sum_{p=0}^{s} b(p,n) < 2^{-n}b(n/2,n)\left(\frac{k-1}{k-2}\right) < \left(\frac{2}{\pi n}\right)^{1/2}\left(\frac{k-1}{k-2}\right). \tag{4}$$

If n is not even, let $s\leqq[n/2]$. Then (4) holds with $b(n/2,n)$ replaced by $b([n/2],n)$.

If $n=x_i$ and $s=s_i$, then twice the left-hand side of (4) represents the fraction of binary strings discarded due to less than or equal to s_i 0's or 1's in the ith section, and thus twice the right-hand side is an upper bound of this quantity. Taking this estimate, let

$$c_t = 2\cdot\left(\frac{2}{\pi t^6}\right)^{1/2}\left(\frac{2+1/t-1}{2+1/t-2}\right) = 2\cdot\left(\frac{2}{\pi t^6}\right)^{1/2}\cdot(t+1) > 0 \tag{5}$$

where $n=t^6$ and $k=n/s=2+1/t$, thus $n=t^6$ and $s=t^6/(2+1/t)$ represent the x_i and s_i respectively. Then c_t bounds the fraction of discarded strings of section i.

It is well known (see Knopp [7, p. 27]) that

$$\frac{\sin \pi z}{\pi z} = \prod_{y=1}^{\infty}(1-z^2/y^2), \qquad \text{all (complex) } z. \tag{6}$$

If $z=2\cdot(2/\pi)^{1/4}$ then with $c_t'=z^2t^{-2}$,

$$\begin{aligned}\prod_{t=1}^{\infty}(1-c_t') &= \prod_{t=1}^{\infty}(1-z^2/t^2) = \frac{\sin \pi z}{\pi z}\\ &= \frac{\sin 2(2\pi^3)^{1/4}}{2(2\pi^3)^{1/4}} = C.\end{aligned} \tag{7}$$

As $1>c_t'=4(2/\pi)^{1/2}t^{-2}>2\cdot(2/\pi)^{1/2}((t+1)/t^3)=c_t$, for $t\geqq 2$, it follows that

$$\prod_{t=u}^{v}(1-c_t') < \prod_{t=u}^{v}(1-c_t) < 1 \tag{8}$$

for any positive integers u, v, $v>u$, so $\prod_{t=2}^{\infty}(1-c_t)$ converges.

For any $l>t_0\geqq 1$, $\prod_{t=t_0+1}^{\infty}(1-c_t')<\prod_{t=t_0+1}^{l}(1-c_t)$, but the right-hand side is the fraction of strings retained after processing the first l sections (with $x_i=t^6$ and $s_i=[t^6/(2+1/t)]$ where $t=t_0+i$). Thus it is sufficient to find for each η, $0<\eta<1$, a t_0 such that

$$\prod_{t=t_0+1}^{\infty}(1-c_t) > \eta. \tag{9}$$

But by (8) this may be solved instead for the infinite product of (7), which can be done constructively simply by calculating a sufficiently by close rational approximation to the constant C, then dividing out a sufficient number of initial terms of the product until a t_0 is found such that

$$\prod_{t=t_0+1}^{\infty}(1-c_t') = \frac{C}{(1-c_1')(1-c_2')\cdots(1-c_{t_0}')} > \eta. \tag{10}$$

This produces the t_0 necessary to specify x_i and s_i in the required manner.

With the main tool now available, that of controlling the fraction of binary sequences discarded per selection rule, we can outline the algorithm which yields the sequence of approximating binary sequences whose "limit" is the desired binary sequence.

We have represented the observed sequence and betting sequence by $A(g)$ and $A(g, h)$ respectively, where g is the Gödel number of the recursive function $g(x)$ which determines $A(g)$ from A, and h is the Gödel number of the recursive function $h(x)$ which determines $A(g, h)$ from $A(g)$. That is, g and h are positive integers such that

$$g(x) = U\left(\min_y T_1(g, x, y)\right); \tag{11}$$

$$h(x) = U\left(\min_y T_1(h, x, y)\right), \tag{12}$$

this representation established by Kleene's normal form theorem (see Davis [2] or Kleene [6]). The predicate $T_1(z, x, y)$ is (primitive) recursive, hence its truth value may be effectively determined; its determination for a particular x, y, z we shall call a "step" in the algorithm. It is necessary to view this as the atomic operation rather than the calculation of $g(x)$ for a given x as the function may not be defined at x.

It is convenient to introduce the notation $A^-(g)$ and $A^-(g, h)$ as (nonunique) notations for a selection rule; in particular, for given integers g and h, $A^-(g)$ is called an observed sequence generator and $A^-(g, h)$, a betting sequence generator. Each is a function from the set of binary sequences to the set of binary sequences, having as values the observed and betting sequences respectively corresponding to the binary sequence appearing as the argument of the function. If a, b are not Gödel numbers the value of the functions is the empty string; otherwise, it is the string produced up to the first irregularity, e.g., $g(x)$ or $h(x)$ is undefined for an appropriate argument x. We now give the order of enumeration of selection rules as follows: using the function $z = J(x, y) = (1+2+3+\cdots+(x+y))+x$ which (well) orders pairs of positive integers, we say $A^-(a, b) < A^-(c, d)$ if $J(a, b) < J(c, d)$. The procedure will meet the condition that no betting sequence generator $A^-(g, h)$ rejects more than the fraction 2^{-z} of the possible binary sequences, where $z = J(g, h)$. That is, for η of equation (9) as a function $\eta(z)$ of g and h, $\eta(z) = 1 - 2^{-z}$ is used to find the appropriate t_0 for $A^-(g, h)$.

The nth approximation to the desired random sequence is found as follows: calculate the appropriate t_0 for each of the first n betting sequence generators $A^-(g, h)$, under the ordering defined above. $A^-(1, 1)$ is taken to be the first generator. This determines the discarding rules for the "candidate" sequences with respect to the first n selection rules. Then take each of the 2^n binary sequences of length n and apply the first n generators $A^-(g, h)$ in turn to each binary sequence until either 2^n steps have been calculated (per application of a given generator to a given sequence) or $a(x)$ assumes a value greater than n, that is, the observed sequence extends beyond the first n terms determined by the given binary sequence. A binary sequence (of length n) is discarded if the initial segment of a betting sequence so generated violates the conditions previously mentioned for any one of the n generators $A^-(g, h)$. The nth approximation is the sequence $a_n(0)$, $a_n(1)$, $a_n(2) \cdots$ with $a_n(i) = 1$ for $i \geqq n$ and the first n terms determined as discussed below.

As regards the calculation of the betting sequence given A^- and the binary sequence $A(\quad; n)$, this is to be performed according to the definition, using the given (effective) coding of "past" information and including the automatic discard of $A(\quad; m)$ with respect to $A^-(g, h)$ if $g(x)$ is not 1-1 over the pertinent domain. The functions $g(x)$, $h(x)$ are calculated by (11) and (12) respectively in repetition until 2^n steps have been performed in total in the one application. Of course the 2^n steps will be absorbed for any n in any calculation of $g(x)$ or $h(x)$ for which the function is not defined with respect to that argument. Clearly, any representation of a proper selection rule will not terminate in this manner, or in any other irregularity previously mentioned. It is important to notice that the above is an effective procedure capable, for example, of being programmed on a modern digital computer.

In order to determine the nth approximation to the desired sequence (now reduced to finding values for n terms), it is necessary to keep a memory of the binary sequences of length n which have not been discarded by any part of the process determining the nth approximation. (There are initially 2^n sequences in memory, of course.) The process for the nth approximation will clearly remove from this memory any sequence which has an initial segment of length $m < n$ which was removed from memory during the calculation of the mth approximation. Thus, after memory has been culled, a sequence remains only if it is an extension of a sequence in the memory of each previous approximation. We determine the first n terms of the nth approximation by considering the sequences remaining in memory as binary numbers (with the first member of the sequence as highest order bit) and selecting that sequence having the greatest magnitude as a binary number. The selected sequence becomes the initial segment (of length n) of the nth approximating sequences to A. That the sequence comprising the nth approximation is well defined is a consequence of memory being nonempty; the latter fact is established by recalling that no larger fraction than $\Sigma_z 2^{-z}$ of memory is discarded where $z = J(g, h) \geqq J(1, 1) = 4$ for $g, h \geqq 1$. That $a_n(i)$ changes only a finite number of times as n proceeds through the positive integers follows as a consequence of the discarding of the extensions of any previously discarded sequence, as for any n such that $a_{n+1}(i) \neq a_n(i)$ we must have $a_{n+1}(0)a_{n+1}(1) \cdots a_{n+1}(i) < a_n(0)a_n(1) \cdots a_n(i)$ when the segments are viewed as binary numbers. As only 2^{i+1} such sequences of length $i+1$ exist, there must be an n_0 such that $a_{n+1}(i) = a_n(i)$ for all $n > n_0$.

We define $Q(i, n) \equiv [a_n(i) = 1]$. As mentioned earlier, we establish the recursiveness of $Q(i, n)$ by invoking Church's thesis. The justification of the application of

Church's thesis to this process has been the main undertaking of this section, that is, to show that the process is capable of execution by an appropriate Turing machine. To actually detail the instructions for such a machine would only produce a mountain of incomprehensible data; we rest our case on the above outline.

The notion $\lim_{n\to\infty} Q(i, n)$ is clearly well defined as shown above. Thus we can now define $P(i)=\lim_{n\to\infty} Q(i, n)$ where $P(i)$ is a trial-and-error predicate.

We define the binary sequence A by $[a(i)=1]\equiv P(i)$. Then A is a recursively random sequence for if it were not there would be some betting sequence determined by partial recursive functions $g(x)$, $h(x)$ whose limiting relative frequency is either not defined or not equal to the limiting relative frequency of the sequence A. But the corresponding generator $A^{-}(g, h)$, where g, h are Gödel numbers for $g(x)$, $h(x)$ respectively, appears in the enumeration from stage z onward, where $z=J(g, h)$. For a given generator, the process discards all sequences not yielding acceptable betting sequences derived from the generator with respect to the associated sequence A.

Thus $P(i)$ is a trial-and-error predicate defining a recursively random sequence (or set), which proves Theorem 1.

To establish Theorem 2 for $i=2$, we note by the lemma of Putnam that $P(i)\in\Sigma_2\cap\Pi_2$. However, P cannot be in Σ_1, for any set B in Σ_1 may be identified with a betting sequence as B is the range set of some recursive function. We observe that if P is a set defining a recursively random sequence then $\bar{P}$, the set of non-negative integers not members of P, must also define a recursively random sequence R by $[r(i)=1]\equiv\bar{P}(i)$, in fact $[r(i)=0]\equiv[a(i)=1]$. Thus $P\in\Pi_1$ implies $\bar{P}\in\Sigma_1$ implies that $\bar{P}$, hence P, does not define a recursively random sequence. This establishes the case $i=2$.

For the case $i=2+j$, $j=1, 2, 3, \ldots$, the method for $i=2$ is employed with use of an "oracle" by which we obtain betting sequences defined by sets in Σ_{j+1}.

Using the notation of the previous theorem, let $g(x)$ represent the function which determines the observed sequence $A(g)$ of A. In the preceding theorem $g(x)$ is recursive; replace this condition as expressed by equation (11) by

$$(11') \qquad g(x) = U\Big(\min_y T_1^{S^{(j)}}(g, x, y)\Big) \qquad \text{where } S^{(j)}$$

is defined by $S^{(0)}=\varnothing$ and $S^{(n+1)}=\{x\,|\,(\exists y)T_1^{S^{(n)}}(x, x, y)\}$, and g is a suitable Gödel number. It is well known (see Rogers [13]) that the class of range sets of $g(x)$ then comprises Σ_{j+1}. Using this formulation for developing the betting sequences in the manner of the previous theorem, each set defines a betting sequence over which the resulting sequence A is random. Thus the set $P=\{i \mid a(k)=1\}$ cannot be in Σ_{j+1}; by symmetry (as discussed above for $i=2$) P cannot be in Π_{j+1}, hence P is not in $\Sigma_{j+1}\cup\Pi_{j+1}$.

To see that $P\in\Sigma_{j+2}\cap\Pi_{j+2}$, define $Q(i, n)\equiv[a(i)=1]$, where $a_n(i)$ is the ith term of the nth approximation to A as defined previously. $Q(i, n)$ is then recursive in $S^{(j)}$ by Church's thesis. Thus $Q \in \Sigma_{j+1} \cap \Pi_{j+1}$ where $Q = \{(i, n) | Q(i, n)\}$. Let $P(i) \equiv \lim_{n\to\infty} Q(i, n)$, where this has been seen to be well defined. Then by the lemma $P = \{i | P(i)\} \in \Sigma_{j+2} \cap \Pi_{j+2}$. This proves Theorem 2.

References

1. A. Church, *On the concept of a random sequence*, Bull. Amer. Math. Soc. **46** (1940), 130–135.

2. M. Davis, *Computability and unsolvability*, McGraw-Hill, New York, 1958.

3. K. de Leeuw, E. F. Moore, C. E. Shannon and N. Shapiro, *Computability by probabilistic machines*, Automata Studies, Princeton Univ. Press, Princeton, N. J., 1956, pp. 183–212.

4. J. L. Doob, *Note on probability*, Ann. of Math. **37** (1936), 363–367.

5. W. Feller, *An introduction to probability theory and its applications*, 2nd ed., Wiley, New York, 1957.

6. S. C. Kleene, *Introduction to mathematics*, Van Nostrand, Princeton, N. J., 1952.

7. K. Knopp, *Theory of functions*, Part 2 (1st American ed., transl. by F. Bagemihl from the 4th German ed.) Dover, New York, 1947.

8. D. W. Loveland, *Recursively random sequences*, Ph.D. dissertation, New York Univ., June, 1964.

9. ———, *A new interpretation of the von Mises concept of random sequence*, Z. Math. Logik Grundlagen Math. (to appear).

10. W. Markwald, *Ein Satz über die elementar-arithmetischen Definierbarkeitsklassen*, Arch. Math. Logik Grundlagenforsch. **2** (1956), 78–86.

11. R. von Mises, *Probability, statistics and truth* (2nd English ed., translated from German), Macmillan, New York, 1957.

12. H. Putnam, *Trial and error predicates and the solution to a problem of Mostowski's*, J. Symbolic Logic **30** (1965), 49–57.

13. H. Rogers, Jr., *Computing degrees of unsolvability*, Math. Ann. **138** (1959), 125–140.

14. A. Wald, *Die Widerspruchsfreiheit des Kollectiv-begriffes der Wahrscheinlichkeitsrechnung*, Ergebnisse eines mathematischen Kolloquiums, Vol. 8, 1937, pp. 38–72.

New York University