Complex Systems **9** (1995) 1–10

Sequences of Pseudorandom Numbers with Arbitrarily Large Periods

P. S. Joag
Department of Physics, University of Poona,
Ganeshkhind, Pune-411007, India

Abstract. We show that the restriction of the Bernoulli shift map $x \leftarrow rx \pmod 1$ to the set of rationals between 0 and 1 can generate sequences of uniformly distributed pseudorandom numbers whose periods exceed any given $k > 0$, however large. Here $r > 1$ is an integer.

1. Introduction

In this paper, we present an algorithm (Theorem 1) that generates sequences of pseudorandom numbers, comprising rational numbers between 0 and 1, with arbitrarily large periods. For any given positive integer k, however large, this algorithm can be used to generate pseudorandom number sequences with period greater than k. Moreover, at least for large periods, these pseudorandom numbers are approximately uniformly distributed, the approximation getting better as the period increases. These pseudorandom sequences are generated using finite strings of r-ary digits which are Kolmogorov random, that is, whose Kolmogorov complexity is close, in some sense, to the length of the string. We explain this concept precisely in section 2.

In order for the main result of Theorem 1 to be of any utility, we need Lemma 1 stated below, which is probably well known, but I can find no exact reference. Before stating the lemma, we give meaning to the phrase "almost all rationals."

Let S denote an infinitely denumerable set with an operation of multiplication defined on it. Define a real-valued mapping on S, denoted $\|a\|$, $a \in S$ such that:

(i) $\|ab\| \leq \|a\| \, \|b\| \quad \forall a, b \in S$;

(ii) the total number $N_S(x)$ of elements $a \in S$ with $\|a\| \leq x$ is finite for each real $x > 0$.

Let

$$S[x] = \{ a \in S \mid \|a\| \leq x \}$$

Denote the cardinality of $S[x]$ by card $S[x]$. Obviously, card $S[x] = N_S(x)$.

Consider $E \subset S$ and let

$$E[x] = \{ a \in E \mid \|a\| \leq x \}.$$

Let card $E[x] = N_E(x)$. We define the relative density of $E[x]$ in $S[x]$ as $N_E(x)/N_S(x)$. The asymptotic relative density of E in S is then defined as

$$d_S(E) = \lim_{x \to \infty} \frac{N_E(x)}{N_S(x)},$$

provided that this limit exists.

A property is said to hold for almost all elements of S if it is valid for all elements in some subset of S having asymptotic relative density 1 in S. We specialize this to the set of rational numbers between 0 and 1. Let

$$Q = \{x \mid x = p/q \text{ where } 0 < p < q \text{ are integers}\}.$$

Define a real-valued map on Q to be

$$\left\| \frac{p}{q} \right\| = q. \tag{1}$$

Obviously, (1) satisfies the properties (i) and (ii) above.

We know that every rational number has either terminating or recurring representation with respect to any base $r > 1$. We now state and prove Lemma 1.

Lemma 1. *Given any integer $r > 1$, almost all rationals have recurring representation with regard to the base r (r-ary representation).*

Proof. We prove this by showing that the set of rationals having terminating r-ary representation have zero asymptotic density in Q. We call such a set T. A given rational number p/q has a terminating r-ary representation provided

$$q = a_1^{m_1} a_2^{m_2} \ldots a_s^{m_s} \tag{2}$$

where $a_1, a_2, \ldots, a_s$ are the prime factors of r and $m_1, m_2, \ldots, m_s$ are integers ≥ 0. Let a_L and a_ℓ denote the largest and smallest prime factors of r, respectively. For a given integer $m > 0$, the set M of integers $q \leq a_L^m$ which can be factorized as in equation (2), is equinumerous to the set of all s tuples satisfying

$$m_1 + \cdots + m_s \leq m \left\lceil \frac{\log a_L}{\log a_\ell} \right\rceil.$$

The cardinality of this set is given by

$$\sum_{k=1}^{m\left\lceil \frac{\log a_L}{\log a_\ell} \right\rceil} \binom{k+s-1}{s-1} \tag{3}$$

where the summand stands for the number of s tuples [6] satisfying the equation

$$m_1 + m_2 + \cdots + m_s \;=\; k.$$

Expression (3) is a polynomial in m, say $P(m)$, and is an upper bound on the number of integers $q \;\leq\; a_L^m$ satisfying equation (2).

Now let,

$$Q[y] \;=\; \{x \;\in\; Q \mid \|x\| \;\leq\; y\}$$

where $\|x\|$ is defined by (1) and $y \;>\; 1$ is real. We easily get

$$\text{card } Q[a_L^m] \;=\; N_Q(a_L^m) \;=\; \sum_{q=1}^{a_L^m} \phi(q) \;\geq\; c\, a_L^{2m}$$

where c is a constant and $\phi(q)$ is the Euler function [8]. The last inequality is proved in [2].

Next we define

$$T[y] \;=\; \{x \;\in\; Q \cap T \mid \|x\| \;\leq\; y\}\,.$$

Obviously,

$$\text{card } T[a_L^m] \;=\; N_T(a_L^m) \;=\; \sum_{q \in M} \phi(q) \;<\; a_L^m\, P(m).$$

Now, given $y \;>\; 1$, choose m such that $a_L^{m-1} \;\leq\; y \;<\; a_L^m$. This gives,

$$N_T(y) = \text{card } T[y] \;<\; N_T(a_L^m) \;<\; a_L^m\, P(m)$$
$$N_Q(y) = \text{card } Q[y] \;\geq\; N_Q(a_L^{m-1}) \;\geq\; c\, a_L^{2(m-1)}.$$

Thus the relative density of $T[y]$ in $Q[y]$ is bounded above by

$$\frac{N_T(y)}{N_Q(y)} \;\leq\; \frac{P(m)}{c\, a_L^{m-2}}.$$

Taking the limit as $y \to \infty$, the left-hand side of the previous inequality gives the asymptotic relative density $d_Q(T)$. Since m increases monotonically with y and $P(m)$ diverges as a polynomial in m while the denominator diverges exponentially, the right-hand side of the above inequality tends to zero as $y \to \infty$. Thus $d_Q(T) \not> 0$. Since by definition $d_Q(T) \geq 0$, we must have $d_Q(T) \;=\; 0$. ■

Thus almost all rationals have recurring r-ary representations. There are two cases of recurring representation : purely and mixed. A purely recurring representation consists of an infinite periodic sequence of digits with period n. A mixed recurring representation consists of finitely many nonperiodic digits followed by an infinite periodic sequence of digits.

For a given r, we call any string constructed out of $\{0, 1, \ldots, r-1\}$ a r-ary string. The digits are called r-ary digits. A string x having n digits is said to be of length n, denoted $|x|$.

2. Random strings of digits: Kolmogorov complexity

We briefly review the concept of Kolmogorov complexity of a finite string x of length n and use it to define a finite random string of r-ary digits [3,5,10–15]. Without losing any generality, we take $r = 2$ while discussing this concept.

Kolmogorov complexity concerns the problem of describing a finite object x. Since a finite object can be coded in terms of a finite binary string, we can take this string to be our object. It is useful to think that the complexity of specifying an object can be facilitated when another object is already specified. Thus we define the complexity of an object x given an object y. Let $p \in \{0,1\}^*$. We call p a *program.* Any computable function f together with strings p and y such that $f(p,y) = x$ is a description of x. We call f the *interpreter* or *decoding* function. The complexity K_f of x with respect to f, conditional to y, is defined by

$$K_f(x|y) = \min\{|p| : p \in \{0,1\}^* \text{ and } f(p,y) = x\}.$$

If there is no such p, then $K_f(x|y) = \infty$.

The invariance theorem [3,10,11,15] asserts that each finite object has an intrinsic complexity that is independent of the means of description. Namely, there exist asymptotically optimal functions (universal Turing machines) such that the description length with respect to them minorizes the description length with respect to any other function, apart from an additive constant, for all finite objects.

Invariance Theorem. *There exists a partial recursive function f_0, such that, for any other partial recursive function f, there is a constant c_f such that for all strings $x, y, K_{f_0}(x|y) \leq K_f(x|y) + c_f$.*

Clearly, any function f_0 that satisfies the invariance theorem is *optimal* in the sense discussed previously. Therefore, we are justified to fix a particular partial recursive function f_0 and drop the subscripts on K. We define the *conditional Kolmogorov complexity* $K(x|y)$ of x under condition of y to be equal to $K_{f_0}(x|y)$ for this fixed optimal f_0. We can now define the *unconditional Kolmogorov complexity* (or Kolmogorov complexity) of x as $K(x) = K(x|\epsilon)$ where ϵ denotes the empty string ($|\epsilon| = 0$).

We are basically concerned with Lemma 2, which is the most important consequence of the invariance theorem [14].

Lemma 2. *There is a fixed constant c' such that for all x of length n,*

$$K(x) \leq n + c'. \tag{4}$$

Thus $K(x)$ is bounded above by the length of x modulo an additive constant. The constant c' turns out to be the number corresponding to the machine T that just copies its input to its output, in some standard enumeration of Turing machines and can be conveniently chosen. We are interested in the binary strings x of length n for which $K(x) \geq n - c$ where $c \geq 0$ is a constant. If c/n is understood to have a fixed fractional value,

we call such strings *c-incompressible*, or simply *incompressible*. Lemma 3 [4] answers the question: How many strings are incompressible?

Lemma 3. *For a fixed $c < n$, out of all possible binary strings of length n at most one in 2^c has $K(x) < n - c$.*

Thus, if we fix c/n to be a small fraction, then the fraction of c-incompressible binary strings increases exponentially with n as $(1 - 2^{-(c/n)n})$. Generally, let $g(n)$ be an integer function. Call a string x of length n *g-incompressible* if $K(x) \geq n - g(n)$. There are 2^n binary strings of length n, and only $2^{n-g(n)} - 1$ possible descriptions shorter than $n - g(n)$. Thus the ratio between the number of strings x of length n with $K(x) < n - g(n)$ and the total number of strings of length n is at most $2^{-g(n)}$, a *vanishing function* when $g(n)$ increases unboundedly with n.

Intuitively, incompressibility implies the absence of regularities, since regularities can be used to compress descriptions. Accordingly, we identify incompressibility with the absence of regularities or *randomness*. In particular, we call c-incompressible strings *c-random*. We do not deal here with the infinite random sequences of digits.

3. Main results

We now state and prove our main results.

Lemma 4. *Suppose an r-ary string x of length n has complexity $K(x)$. Cut this string after the sth digit so that the concatenation of the two resulting substrings (partitions), say x_1 and x_2, can give the original string. Let $K(x_1)$ and $K(x_2)$ denote the complexities of these partitions. Then $K(x_1)+K(x_2) \geq K(x)$.*

Proof. Suppose $K(x_1) + K(x_2) < K(x)$. However, the string x can be produced by using the minimal programs corresponding to x_1 and x_2 in succession. This gives the complexity of x to be $\leq K(x_1) + K(x_2) < K(x)$, which contradicts our premise that the complexity of x is $K(x)$. ∎

Corollary 1. *If a string x of length n is c-random, then any two partitions of x, say x_1 and x_2, are at least $2c$-random. A string generated by concatenating x_1 and x_2 is c-random.*

Proof. Let $|x_1| = n_1$, $|x_2| = n_2$, with $n_1+n_2 = n$. Without losing generality we can choose the constant c' appearing in the inequalities $K(x_1) \leq n_1 + c'$, $K(x_2) \leq n_2 + c'$, and $K(x) \leq n + c'$, which are true by virtue of Lemma 2, to satisfy $c > c'$. Now assume that x_1 is not $2c$-random so that $K(x_1) < n_1 - 2c$. Subtracting $K(x_1)$ from the left-hand side and $n_1 - 2c$ from the right-hand side of the inequality $K(x_1) + K(x_2) \geq K(x) \geq n_1 + n_2 - c$ we get $K(x_2) \geq n_2 + c > n_2 + c'$ which contradicts the requirement $K(x_2) \leq n_2 + c'$ imposed by Lemma 2.

As for the second part of the corollary, if the two partitions are concatenated to produce the original string, there is nothing to prove. Now suppose that the partitions x_1 and x_2 are concatenated in the reverse order, giving a string $x_{\text{inv}} = x_2 x_1$ with complexity

$$K(x_{\text{inv}}) \ll n - c, \tag{5}$$

which means that x_{inv} is not c-random. Then the string x can be printed in the following way. Produce x_{inv} and print the last n_1 digits first and the first n_2 digits next. This will enhance the length of the minimal program producing x by $\min(\log n_1, \log n_2)$ which we take to be $\log n_1$. Thus

$$K(x) \leq K(x_{\text{inv}}) + \log n_1 < n - c. \tag{6}$$

The last inequality follows from (5) and $\log n_1 \ll n - c$. Inequality (6) means that x is not c-random, which contradicts our premise and completes the proof. ■

Definition. *A string x of length n is said to be random if $K(x) \geq n - O(\log n)$.*

Corollary 1 can be easily extended to the case of a random string of length n, as is done in Corollary 2.

Corollary 2. *If a string x of length n is random, then any two partitions of x, say x_1 and x_2, are random. A string generated by concatenating x_1 and x_2 is random.*

Proof. In the previous statement and in the proof of the Corollary 1, if we replace c by a function $f(n) = O(\log n)$, then we can also replace $2c$ by $f(n)$. ■

In Theorem 1 we make use of Corollary 2.

Definition. *We call the map $x \leftarrow rx \pmod 1$ on Q, that is, the Bernoulli shift* [7], *restricted to rationals a shift over rationals (SOR). Here $r > 1$ is an integer.*

Theorem 1. *Let $n = n(r, x_0)$ denote the period of the recurring r-ary representation of $x_0 \in Q$. Then, for almost all $x_0 \in Q$, at least the first n iterations of the SOR generate a sequence of pseudorandom numbers with period $\{x_0, x_1, \ldots, x_{n-1}\}$ such that the first (most significant) n digits of the r-ary representation of each x_i $(i = 0, 1, \ldots, n-1)$ is a random string, provided the first n digits of the r-ary representation of x_0 is a random string.*

Proof. This proof applies to $x_0 \in Q$ having a recurring r-ary representation. By virtue of Lemma 1, this means that the proof applies to almost all $x_0 \in Q$. There are two possibilities: x_0 may have either purely recurring or mixed recurring representation. We deal with these two cases separately.

Case I: Purely recurring representation.

Let $\{d_1, d_2, \ldots, d_n\}$ be the first (most significant) n digits (the first period) of the r-ary representation of x_0. Operate by SOR on x_0 to get x_1 whose first n digits are $\{d_2, d_3, \ldots, d_n, d_1\}$, because SOR removes d_1 from the r-ary representation of x_0 to generate that of x_1. Thus the string of the first n digits of the r-ary representation of x_1 is obtained by partitioning that of x_0 after d_1 and concatenating the two partitions in reverse order. Therefore, by corollary to Lemma 4, the string formed by the first n digits in the r-ary representation of x_1 is random, provided the corresponding string for x_0 was random. By the same argument, further iterations of the SOR, $x_{t+1} \leftarrow x_t \pmod 1$ $(t = 1, 2, \ldots, n-2)$ generate rational numbers $x_2, \ldots, x_{n-1}$ having r-ary representations whose first (most significant) n digits form a random string, provided the corresponding string for x_0 was random. After the nth iteration, the same cycle $\{x_0, x_1, \ldots, x_{n-1}\}$ repeats due to periodicity of the representation.

Case II: Mixed recurring representation.

Let the r-ary representation of x_0 be periodic after the first m nonperiodic digits and let the period be n. Consider the most significant $m+n-1$ digits of $x_1 = rx_0 \pmod 1$. This is the partition of size $m+n-1$ of the string of most significant $m+n$ digits in the r-ary representation of x_0. By corollary to Lemma 4, this partition is a random string, provided the corresponding string for x_0 was random and the next $m-1$ iterations generate $m-1$ more rationals having a random string of digits of length $> n$ as the most significant digits in their r-ary representations, provided the corresponding string for x_0 was random. After m iterations, r-ary representation of x_m is purely recurring and *Case I* applies. ■

We now show that the sequence of pseudorandom numbers produced by the SOR as in Theorem 1, is uniformly distributed [9]. Note that the frequency of occurrence of all the r digits $\{0, 1, \ldots, r-1\}$ in a string of r-ary digits of length n is close to n/r, provided n is large. In fact, the probability that this frequency deviates from n/r by an amount greater than a fraction δ of n is bounded above by

$$L\, n\, \exp\left(\frac{-1}{2} K\, \delta^2\, n\right) \tag{7}$$

where L and K are constants that depend on r [8]. Since rationals are dense in $(0, 1)$, those having r-ary representations with very large periods are abundantly available (see the beginning of section 4).

Now divide $[0, 1)$ into r intervals

$$\frac{s}{r} \leq y < \frac{s+1}{r} \quad (s = 0, 1, \ldots, r-1). \tag{8}$$

The $(s+1)$th interval contains just those numbers whose r-ary representation begins with s. Suppose we construct the sequence of pseudorandom numbers $x_0, x_1, \ldots, x_{n-1}$ using Theorem 1. It is easily seen that each one

of the first $n = n(r, x_0)$ digits $\{d_1, d_2, \ldots, d_n\}$ of the r-ary representation of x_0 successively becomes the most significant digit of the r-ary representation of $\{x_1, x_2, \ldots, x_{n-1}\}$. Since all digits occur with equal frequency n/r, (neglecting very small fluctuations decaying exponentially with n), each of the r intervals defined above will contain n/r numbers out of the pseudorandom sequence $\{x_0, x_1, \ldots, x_{n-1}\}$. We now divide each of the intervals given by (8) into r subintervals so that the $(s+1)$th subinterval contains numbers whose r-ary representations have s as their second digit. The fraction of pseudorandom numbers that fall in the interval corresponding to the ordered pair of digits (s, t) equals the number of pairs (s, t) occurring in the string of the first n digits in the r-ary representation of x_0. The number of such pairs is $(n-1)/r^2$ provided that this string has $n = n(x_0, r)$ large enough to make the fluctuations given by (7) negligible. Thus the expected number of rationals from the pseudorandom sequence in each of the r^2 intervals is $(n-1)/r^2 \approx n/r^2$. We can continue dividing $[0, 1)$ in the same way, each time getting $r^3, r^4, \ldots$ intervals with the expected equal occupancy given by $(n-2)/r^3 \approx n/r^3$, $(n-3)/r^4 \approx n/r^4$ ... rationals from the pseudorandom sequence. This shows that the distribution of pseudorandom numbers is uniform over the pattern of intervals described, provided $n = n(x_0, r)$ is large enough to make the fluctuations given by (7) negligible. Since the rationals are dense in $(0, 1)$, for any given k, however large, there exist infinite $x_0 \in Q$ with $n(x_0, r) > k$. Therefore, we can always choose x_0 with $n(x_0, r) > k$ for any given k.

It is not the case that we are getting the uniform distribution of $\{x_0, \ldots, x_{n-1}\}$ due to some special characteristics of the way we are dividing $[0, 1)$ into subintervals. In fact, any division of $[0, 1)$ into subintervals of equal length satisfies inequality (8) for some h, $(h = 2, 3, 4, \ldots)$. Suppose $h \neq r$. Divide the interval $[0, 1)$ into h^ℓ subintervals of equal length for some $\ell > 0$. For every $\epsilon > 0$, it is possible to find integers m and k such that $0 < (h^{-\ell} - mr^{-k}) < \epsilon$. On the given mesh of h^ℓ intervals we now superimpose a mesh of r^j intervals such that $r^{j-1} < mh^\ell < r^j$. We divide this superimposed mesh into partitions of approximately r^j/h^ℓ intervals such that each partition closely overlaps each of the h^ℓ subintervals of the original mesh. Since the distribution of pseudorandom numbers is uniform over the above partitions it is uniform over the h^ℓ subintervals, each with length $h^{-\ell}$.

4. Discussion

Theorem 1 gives an algorithm to generate sequences of pseudorandom numbers $\{x_0, \ldots, x_{n-1}\}$ whose period is not less than that of the r-ary representation of x_0 namely $n = n(r, x_0)$. A very important advantage of this method lies in the fact that we can choose $n = n(r, x_0)$ as large as we please. If we choose $x_0 = p/q$ such that q is relatively prime to both p and r, then the period $n = n(r, x_0)$ is given by the smallest ν satisfying $r^\nu \equiv 1 \pmod{q}$, that is, the smallest ν such that $(r^\nu - 1)$ is divisible by q [8]. For instance, with $r = 3$ and $x_0 = 0.1875000000000001$ $(q = 10^{16})$, this period is of the order

of 10^{14}. In fact, given any values k and r, it is possible to choose $x_0 = p/q$ that has a recurring r-ary representation with period n greater than k. Since rationals are dense in $(0, 1)$, the rationals having r-ary representation with period greater than a given fixed integer are infinitely abundant. Thus the algorithm given by Theorem 1 can generate sequences of uniformly distributed pseudorandom numbers with arbitrarily large periods.

Suppose we want to choose r and $x_0 = p/q$ such that $n = n(r, x_0) > k$ where k is some fixed integer. This can be done, for example, as follows. Choose r to be a prime and $x_0 = p/q$ such that p and q are relatively prime and $q = r^k + 1$. Thus q and r are also relatively prime and the period $n(r, x_0)$ is given by the smallest ν such that $r^\nu - 1$ is divisible by $q = r^k + 1$. Therefore, $r^\nu - 1 > r^k + 1$ giving $n = \nu > k$. The actual digits of p can be chosen from a table of random numbers [1].

In order that Theorem 1 applies in practice, we have to compute the successive values of $rx \pmod 1$ using fully rational arithmetic and not by the floating point arithmetic that is available on most computers. This facility is offered by all the computer algebra systems. This slows down the computation, but is not really a bottleneck as one can generate a huge bank of pseudorandom numbers, and issue them to various computing processes (either parallel or sequential) as and when required.

We have carried out statistical tests for randomness as given in [9] for a large sample of pseudorandom sequences generated using Theorem 1, each of the length of a few thousand. Each of these sequences has passed the frequency test and the serial tests satisfactorily. The serial test was repeated by choosing the pairs of numbers with the gap between them increasing from 0 to 23 in steps of 1. Thus, in these sequences, the pairs of numbers with gaps up to 23 are uncorrelated.

As we know, SOR cuts the r-ary representation of a rational number after the first digit and chops it off. It is possible to think of a large class of maps that chop the string of r-ary digits at different places in successive iterations. Although these maps will generate pseudorandom sequences in the sense of Theorem 1, the periods of these sequences will not, in general, be equal to the period of the r-ary representation of the seed x_0, in which case the pseudorandom numbers may not be uniformly distributed.

Acknowledgements

I thank Professor S. R. Adke for the perusal of the manuscript and many useful suggestions. I am grateful to Dr. Mohan Nair and Dr. Mrs. Mangala Naralikar for pointing out mistakes in the proof of Lemma 1, and to Dr. Mohan Nair for suggesting a generalization of this proof. Finally, it is a pleasure to thank Dr. Dominic Welsh for some illuminating discussions.

References

[1] M. Abramowitz and I. A. Stegan (editors), *Handbook of Mathematical Functions* (Dover Publications, New York, 1965).

[2] T. M. Apostol, *Analytic Number Theory* (Narosa, New Delhi, 1972).

[3] G. J. Chaitin, "On the Length of Programs for Computing Finite Binary Sequences: Statistical Considerations," *Journal of the Association for Computing Machinary*, **16** (1969) 145–159.

[4] G. J. Chaitin, "Randomness and Mathematical Proof," *Scientific American*, **232** (1975) 47–52.

[5] G. J. Chaitin, *Algorithmic Information Theory* (Cambridge University Press, Cambridge, 1987).

[6] W. Feller, *An Introduction to Probability Theory and its Applications (Vol. I) (3rd edition)* (Wiley Eastern Ltd., New Delhi, 1968)

[7] J. Ford, "Chaos: Solving the Unsolvable, Predicting the Unpredictable!" in *Chaotic Dynamics and Fractals,* edited by M. F. Barnsley and S. G. Demco (Academic Press, London, 1986).

[8] G. H. Hardy and E. M. Wright, *Theory of Numbers (5th edition)* (Oxford University Press, Oxford, 1978).

[9] D. E. Knuth, *The Art of Computer Programming Volume 2: Seminumerical Algorithms (2nd edition)* (Addison-Wesley, Reading, 1980)

[10] A. N. Kolmogorov, "On Tables of Random Numbers," *Sankhya Series A* **25** (1963) 369–376.

[11] A. N. Kolmogorov, "Three Approaches to the Quantitative Transmission of Information," *Problems Information Transmission* **1** (1965) 1–7.

[12] A. N. Kolmogorov, "Combinatorial Foundations of Information Theory and the Calculus of Probabilities," *Russian Mathematical Surveys* **38** (1983) 29–40.

[13] L. A. Levin, "Randomness Conservation Inequalities: Information and Independence in Mathematical Theories," *Information and Control* **61** (1984) 15–37.

[14] M. Li and P. M. B. Vitanyi, "Kolmogorov Complexity and its Applications," in *Handbook of Theoretical Computer Science,* J. van Leeuwen, editor (Elsevier Science, 1990) 189–253.

[15] R. J. Solomonoff, "A Formal Theory of Inductive Inference, Part 1 and Part 2," *Information and Control* **7** (1964) 1–22 and 224–254.