Theory Comput. Systems **32**, 517–529 (1999)

Theory of Computing Systems

Randomness, Stochasticity, and Approximations*

Y. Wang

Department of Combinatorics and Optimization,
University of Waterloo, Waterloo,
Ontario, Canada N2L 3G1
wang@cs.uwm.edu

Abstract. Polynomial time unsafe approximations for intractable sets were introduced by Meyer and Paterson [9] and Yesha [19], respectively. The question of which sets have optimal unsafe approximations has been investigated extensively; see, e.g., [1], [5], [15], and [17]. Recently, Wang [15], [17] showed that polynomial time random sets are neither optimally unsafe approximable nor Δ-levelable. In this paper we show that: (1) There exists a polynomial time stochastic set in the exponential time complexity class which has an optimal unsafe approximation. (2) There exists a polynomial time stochastic set in the exponential time complexity class which is Δ-levelable. The above two results answer a question asked by Ambos-Spies and Lutz [2]: What kind of natural complexity property can be characterized by p-randomness but not by p-stochasticity? Our above results also extend Ville's [13] historical result. The proof of our first result shows that, for Ville's stochastic sequence, we can find an optimal prediction function f such that we will never lose our own money betting according to f (except the money we have earned), that is to say, if at the beginning we have only \$1 and we always bet \$1 that the next selected bit is 1, then we always have enough money to bet on the next bit. Our second result shows that there is a stochastic sequence for which there is a betting strategy f such that we will never lose our own money betting according to f (except the money we have earned), but there is no such optimal betting strategy. That is to say, for any such betting strategy, we can find another betting strategy which could be used to make money more quickly.

* A preliminary version of this paper has appeared in the *Proceedings of RANDOM* '97, Springer-Verlag, New York.

1. Introduction

Random sequences were first introduced by von Mises [10] as a foundation for probability theory. Von Mises considered a random sequence to be a type of disordered sequences, called "Kollektivs." The two features characterizing a Kollektiv are: the existence of limiting relative frequencies within the sequence and the invariance of these limits under the operation of an "admissible place selection rule." Here an admissible place selection rule is a procedure for selecting a subsequence of a given sequence ξ in such a way that the decision to select a bit $\xi[n]$ does not depend on the value of $\xi[n]$. However, von Mises' definition of an "admissible place selection rule" is not rigorous according to modern mathematics. After von Mises introduced the concept of "Kollektivs," the first question raised was whether this concept is consistent. Wald [14] answered this question affirmatively by showing that, for each countable set of admissible place selection rules, the corresponding set of "Kollektivs" has Lebesgue measure 1. The second question raised was whether all "Kollektivs" satisfy the standard statistical laws. For a negative answer to this question, Ville [13] constructed a counterexample in 1939. He showed that, for each countable set of admissible place selection rules, there exists a "Kollektiv" which does not satisfy the law of the iterated logarithm. The example of Ville defeated the plan of von Mises to develop probability theory based on "Kollektivs," that is to say, to give an axiomatization of probability theory with "random sequences" (i.e., "Kollektivs") as a primitive term. Later, admissible place selection rules were further developed by Tornier, Wald, Church, Kolmogorov, Loveland, and others. This approach of von Mises to define random sequences is now known as the "stochastic approach."

A completely different approach to the definition of random sequences was proposed by Martin-Löf [8]. He developed a quantitative (measure-theoretic) approach to the notion of random sequences. This approach is free from those difficulties connected with the frequency approach of von Mises. Later, Schnorr [11] used the martingale concept to give a uniform description of various notions of randomness. In particular, he gave a characterization of Martin-Löf's randomness concept in these terms.

Using the martingale concept, Schnorr [11] introduced resource-bounded randomness concepts, and later Lutz [7] introduced a kind of resource-bounded measure theory. Resource-bounded versions of stochasticity concepts were also introduced by several authors, e.g., Wilber [18], Ko [6], and Ambos-Spies et al. [3].

The notion of unsafe approximations was introduced by Yesha in [19]: an unsafe approximation algorithm for a set A is just a standard polynomial time bounded deterministic Turing machine M with outputs 1 and 0. Duris and Rolim [5] further investigated unsafe approximations and introduced a levelability concept, Δ-levelability, which implies the nonexistence of optimal polynomial time unsafe approximations. They showed that complete sets for $\mathbf{E}$ are Δ-levelable and there exists an intractable set in $\mathbf{E}$ which has an optimal safe approximation but no optimal unsafe approximation. However, they did not succeed in producing an intractable set with optimal unsafe approximations. Ambos-Spies [1] defined a concept of weak Δ-levelability and showed that there exists an intractable set in $\mathbf{E}$ which is not weakly Δ-levelable (hence it has an optimal unsafe approximation). In [15], [16], and [17], Wang extended Ambos-Spies's results by showing that both the class of Δ-levelable sets and the class of sets which have optimal polynomial time unsafe approximations have p-measure 0. Wang's results show that

Δ-levelable sets and optimally approximable sets could not be p-random. However, in this paper, we show the following results:

- There is a p-stochastic set in $\mathbf{E}_2$ which has an optimal unsafe approximation.
- There is a p-stochastic set in $\mathbf{E}_2$ which is Δ-levelable.

Note that our above results extend Ville's [13] historical result. Ville's result says that: for every countable set of admissible place selection rules, we can construct a stochastic sequence ξ which has more ones than zeros in its initial segments. As we show in Theorem 4.10, for this stochastic sequence ξ, the prediction function $f(x) = 1$ will be the optimal prediction strategy since, for every other prediction function g, there is a $k \in N$ such that $\|\{i < n\colon g(\xi[0..i-1]) = \xi[i]\}\| \leq \|\{i < n\colon f(\xi[0..i-1]) = \xi[i]\}\| + k$ for almost all $n \in N$. Our second result (Lemma 4.11 and Theorem 4.12) says that: for every countable set of admissible place selection rules, we can construct a stochastic sequence ξ such that there is no optimal prediction strategy for this sequence. That is to say, for every prediction function f, there is another prediction function g and an unbounded nondecreasing function $r(n)$ such that $\|\{i < n\colon g(\xi[0..i-1]) = \xi[i]\}\| \geq \|\{i < n\colon f(\xi[0..i-1]) = \xi[i]\}\| + r(n)$ for almost all $n \in N$. We prove our results for the resource-bounded case only, but all of these results hold for the classical case also.

The outline of the paper is as follows. In Section 3 we review the relations between the concept of resource-bounded randomness and the concept of polynomial time unsafe approximations. In Section 4 we establish the relations between the concept of resource-bounded stochasticity and the concept of polynomial time unsafe approximations.

2. Definitions

Let N and $Q(Q^+)$ denote the set of natural numbers and the set of (nonnegative) rational numbers, respectively. $\Sigma = \{0, 1\}$ is the binary alphabet, Σ^* is the set of (finite) binary strings, Σ^n is the set of binary strings of length n, and Σ^∞ is the set of infinite binary sequences. The length of a string x is denoted by $|x|$. $<$ is the length-lexicographical ordering on Σ^* and z_n ($n \geq 0$) is the nth string under this ordering. λ is the empty string. For strings $x, y \in \Sigma^*$, xy is the concatenation of x and y. For a sequence $x \in \Sigma^* \cup \Sigma^\infty$ and an integer number $n \geq -1$, $x[0..n]$ denotes the initial segment of length $n+1$ of x ($x[0..n] = x$ if $|x| < n+1$) and $x[i]$ denotes the ith bit of x, i.e., $x[0..n] = x[0]\cdots x[n]$. Lowercase letters $\ldots, k, l, m, n, \ldots, x, y, z$ from the middle and the end of the alphabet denote numbers and strings, respectively. The letter b is reserved for elements of Σ, and lowercase Greek letters $\xi, \eta, \ldots$ denote infinite sequences from Σ^∞.

A subset of Σ^* is called a language or simply a set. Capital letters are used to denote subsets of Σ^* and boldface capital letters are used to denote subsets of Σ^∞. The cardinality of a language A is denoted by $\|A\|$. We identify a language A with its characteristic function, i.e., $x \in A$ iff $A(x) = 1$. The characteristic sequence χ_A of a language A is the infinite sequence $\chi_A = A(z_0)A(z_1)A(z_2)\cdots$. We freely identify a language with its characteristic sequence and the class of all languages with the set Σ^∞. For a language $A \subseteq \Sigma^*$ and a string $z_n \in \Sigma^*$, $A\lceil z_n$ denotes the finite initial segment of A below z_n, i.e., $A\lceil z_n = \{x\colon x < z_n \& x \in A\}$. For languages A and B, $\bar{A} = \Sigma^* - A$ is the complement of A, and $A\Delta B = (A - B) \cup (B - A)$ is the symmetric difference of A and B.

We fix a standard polynomial time computable and invertible pairing function $\lambda x, y\langle x, y\rangle$ on Σ^*. We use **P**, **E**, and $\mathbf{E}_2$ to denote the complexity classes *DTIME*(*poly*), $DTIME(2^{linear})$ and $DTIME(2^{poly})$, respectively. Finally, we fix a recursive enumeration $\{P_e\colon e \geq 0\}$ of **P** such that $P_e(x)$ can be computed in $O(2^{|x|+e})$ steps (uniformly in e and x).

We close this section by introducing a fragment of Lutz's effective measure theory which will be sufficient for our investigation.

Definition 2.1. A *martingale* is a function $F\colon \Sigma^* \to Q^+$ such that, for all $x \in \Sigma^*$,

$$F(x) = \frac{F(x1) + F(x0)}{2}.$$

A martingale F *succeeds* on a set $A \subseteq \Sigma^*$ if $\limsup_n F(A\restriction z_n) = \infty$.

Definition 2.2 [7]. A class **C** of sets has *p-measure* 0 ($\mu_p(\mathbf{C}) = 0$) if there is a polynomial time computable martingale $F\colon \Sigma^* \to Q^+$ which succeeds on every set in **C**. The class **C** has *p-measure* 1 ($\mu_p(\mathbf{C}) = 1$) if $\mu_p(\bar{\mathbf{C}}) = 0$ for the complement $\bar{\mathbf{C}} = \{A \subseteq \Sigma^*\colon A \notin \mathbf{C}\}$ of **C**.

Definition 2.3 [11]. A set A is n^k*-random* if, for every n^k-time computable martingale F, F does not succeed on A. A set A is *p-random* if A is n^k-random for all $k \in N$.

The following theorem is useful in the study of p-measure theory.

Theorem 2.4. *A class* **C** *of sets has* p-measure 0 *if and only if there exists a number* $k \in N$ *such that there is no* n^k*-random set in* **C**.

Proof. See, e.g., [15]. □

3. Resource-Bounded Randomness versus Polynomial Time Unsafe Approximations

For completeness, in this section we review the results in [15] and [17] which show the relations between the resource-bounded randomness concept and polynomial time unsafe approximation concepts.

Definition 3.1 [5], [19]. A *polynomial time unsafe approximation* of a set A is a set $B \in \mathbf{P}$. The set $A\Delta B$ is called the *error* set of the approximation. Let f be a function defined on the natural numbers such that $\limsup_{n\to\infty} f(n) = \infty$. A set A is Δ*-levelable with density* f if, for any set $B \in \mathbf{P}$, there is another set $B' \in \mathbf{P}$ such that

$$\|(A\Delta B)\restriction z_n\| - \|(A\Delta B')\restriction z_n\| \geq f(n) \tag{1}$$

for almost all $n \in N$. A set A is Δ-*levelable* if A is Δ-levelable with density f for some f.

Definition 3.2 [1]. A polynomial time unsafe approximation B of a set A is *optimal* if, for any approximation $C \in \mathbf{P}$ of A,

$$\exists k \in N, \quad \forall n \in N, \quad \|(A \Delta B) \lceil z_n\| < \|(A \Delta C) \lceil z_n\| + k. \tag{2}$$

A set A is *weakly* Δ-*levelable* if, for any polynomial time unsafe approximation B of A, there is another polynomial time unsafe approximation B' of A such that

$$\forall k \in N, \quad \exists n \in N, \quad \|(A \Delta B) \lceil z_n\| > \|(A \Delta B') \lceil z_n\| + k. \tag{3}$$

It should be noted that our above definitions are a little different from the original definitions of Ambos-Spies [1], Duris and Rolim [5], and Yesha [19]. In the original definitions, they considered the errors on strings up to a certain length (i.e., $\|(A\Delta B)^{\leq n}\|$) instead of errors on strings up to z_n (i.e., $\|(A\Delta B) \lceil z_n\|$).

Lemma 3.3 [1].

1. *A set A is weakly Δ-levelable if and only if A does not have an optimal polynomial time unsafe approximation*.
2. *If a set A is Δ-levelable, then it is weakly Δ-levelable.*

In [15] and [17], we have established the following relations between the p-randomness concept and unsafe approximation concepts.

Theorem 3.4 [15], [17]. *The class of Δ-levelable sets has p-measure* 0.

Theorem 3.5 [15], [17]. *The class of sets which have optimal polynomial time unsafe approximations has p-measure* 0.

Corollary 3.6 [15], [17]. *The class of sets which are weakly Δ-levelable but not Δ-levelable has p-measure* 1.

Corollary 3.7 [15], [17]. *Every p-random set is weakly Δ-levelable but not Δ-levelable.*

4. Resource-Bounded Stochasticity versus Polynomial Time Unsafe Approximations

As we have mentioned in the Introduction, the first notion of randomness was proposed by von Mises [10]. He called a sequence random if every subsequence obtained by an admissible selection rule satisfies the law of large numbers. A formalization of this notion, based on formal computability, was given by Church [4] in 1940. Following Kolmogorov (see [12]) we call randomness in the sense of von Mises and Church stochasticity.

For a formal definition of Church's stochasticity concept, we first formalize the notion of a selection rule.

Definition 4.1. A *selection function* f is a partial recursive function $f\colon \Sigma^* \to \Sigma$. A selection function f is *dense along* A if $f(\chi_A[0..n-1])$ is defined for all n and $f(\chi_A[0..n-1]) = 1$ for infinitely many n.

By interpreting A as the infinite 0–1-sequence χ_A, a selection function f selects the subsequence $\chi_A[n_0]\chi_A[n_1]\chi_A[n_2]\cdots$ of χ_A where $n_0 < n_1 < n_2 < \cdots$ are the numbers n such that $f(\chi_A[0..n-1]) = 1$. In particular, f selects an infinite subsequence ξ of χ_A iff f is dense along A. So Church's stochasticity concept can be defined as follows.

Definition 4.2. An infinite sequence $\xi \in \Sigma^\infty$ satisfies the law of large numbers if the following condition holds:

$$\lim_n \frac{\sum_{i=0}^{n} \xi[i]}{n+1} = \frac{1}{2}.$$

Definition 4.3 [4]. A set A is *stochastic* if, for every selection function f which is dense along A, f selects an infinite subsequence ξ (of χ_A) which satisfies the law of large numbers.

For the resource-bounded version of Church stochasticity, Ambos-Spies et al. [3] introduced the following n^k-stochasticity notion.

Definition 4.4 [3]. An n^k*-selection function* is a total selection function f such that $f \in \mathit{DTIME}(n^k)$. A set A is n^k*-stochastic* if, for every n^k-selection function f which is dense along A, f selects an infinite subsequence ξ (of χ_A) which satisfies the law of large numbers.

These concepts can also be characterized in terms of prediction functions. A prediction function f is a procedure which, given a finite initial segment of a 0–1-sequence, predicts the value of the next member of the sequence. We will show that a set A is stochastic iff, for every partial prediction function which makes infinitely many predictions along A, the numbers of the correct and incorrect predictions are asymptotically the same.

Definition 4.5 [3]. A *prediction function* f is a partial function $f\colon \Sigma^* \to \Sigma$. An n^k*-prediction function* f is a prediction function f such that $f \in \mathit{DTIME}(n^k)$ and $\mathit{domain}(f) \in \mathit{DTIME}(n^k)$. A prediction function f is *dense along* A if $f(\chi_A[0..n-1])$ is defined for infinitely many n. *A meets* (*avoids*) f *at* z_n if $f(\chi_A[0..n-1])$ is defined and $f(\chi_A[0..n-1]) = \chi_A[n]$ ($f(\chi_A[0..n-1]) = 1 - \chi_A[n]$). *A meets* f *balancedly* if

$$\lim_n \frac{\|\{i < n\colon\ f(\chi_A[0..i-1]) = \chi_A[i]\}\|}{\|\{i < n\colon\ f(\chi_A[0..i-1])\downarrow\}\|} = \frac{1}{2}. \tag{4}$$

Theorem 4.6 [3]. *For any set A, the following are equivalent*:

1. *A is n^k-stochastic* (p*-stochastic*).
2. *A meets balancedly every n^k-prediction* (p*-prediction*) *function which is dense along A*.

The following theorem is straightforward.

Theorem 4.7 [3]. *If a set A is n^k-random, then it is n^k-stochastic*.

We first show that neither Δ-levelability nor optimal approximability implies p-stochasticity.

Theorem 4.8.

1. *There is a non-p-stochastic set B in $\mathbf{E}_2$ which has an optimal unsafe approximation*.
2. *There is a non-p-stochastic set B in $\mathbf{E}_2$ which is Δ-levelable*.

Proof. 1. Let $A \in \mathbf{E}_2$ be a set which has an optimal unsafe approximation (the existence of such A has been shown by Ambos-Spies [1]), and let $B = \{z_{2n}, z_{2n+1} : z_n \in A\}$. Then B has an optimal unsafe approximation and the prediction function f defined by

$$f(x) = \begin{cases} x[|x|-1] & \text{if } |x| \text{ is odd,} \\ \uparrow & \text{otherwise} \end{cases}$$

witnesses that B is not p-stochastic.

2. The proof is the same as that of 1. □

Before we prove our main theorems, we prove the following lemma which presents the basic idea underlying Ville's construction.

Lemma 4.9. *Let f_0, f_1 be two n^k-selection functions. Then there is a set A in $\mathbf{E}_2$ such that*

$$\begin{aligned} &\|\{i < n\colon\ f_b(\chi_A[0..i-1]) = 1 = \chi_A[i]\}\| \\ &\qquad > \|\{i < n\colon\ f_b(\chi_A[0..i-1]) = 1 = 1 - \chi_A[i]\}\| \end{aligned}$$

for all $n \in N$ and $b \in \Sigma$.

Proof. The construction of A is as follows.

Let $\xi_{0,0} = \xi_{0,1} = \xi_{1,0} = \xi_{1,1} = 110101010\cdots\cdots \in \Sigma^\infty$. For $i \in N$, assume that $\chi_A[0..i-1]$ has already been defined. If $(b_0, b_1) = (f_0(\chi_A[0..i-1]),\ f_1(\chi_A[0..i-1]))$, then let $\chi_A[i]$ be the first bit in the sequence ξ_{b_0,b_1} that has not been used.

For the above constructed set A, every initial segment of the sequence selected by f_0 (f_1) from χ_A is a "mixture" of the initial segments of $\xi_{1,0}$ and $\xi_{1,1}$ ($\xi_{0,1}$ and $\xi_{1,1}$). Hence it satisfies the requirements of the lemma. □

Theorem 4.10. *There is a p-stochastic set $A \in \mathbf{E}_2$ satisfying the following properties*:

1. *For every p-selection function f which is dense along A, there is an unbounded nondecreasing function $r(n)$ such that*

$$\begin{aligned} &\|\{i < n\colon\ f(\chi_A[0..i-1]) = 1 = \chi_A[i]\}\| \\ &\qquad \geq \|\{i < n\colon\ f(\chi_A[0..i-1]) = 1 = 1 - \chi_A[i]\}\| + r(n) \end{aligned} \tag{5}$$

   for almost all $n \in N$.
2. *A has an optimal unsafe approximation.*

Proof. Let $f_0, f_1, \ldots$ be an enumeration of all p-selection functions. The construction of A is a modification of the construction in Lemma 4.9. The detailed construction is as follows.

Let $n_j = 2^{2j}$ for all $j \in N$, and let $\xi_w = 1110101010\cdots\cdots \in \Sigma^\infty$ for all $w \in \Sigma^*$. For $i \in N$, assume that $\chi_A[0..i-1]$ has already been defined. Let $x = f_0(\chi_A[0..i-1])\, f_1(\chi_A[0..i-1]) \cdots f_{i-1}(\chi_A[0..i-1])$ and let j be the least integer such that we have used less than n_j bits from $\xi_{x[0..j]}$. Then let $\chi_A[i]$ be the first bit in $\xi_{x[0..j]}$ that we have not used.

The basic idea underlying the above construction is the same as that underlying the construction in Lemma 4.9. However, here there are countably many selection rules whence each bit of the constructed sequence is characterized by an infinite binary sequence $b_0 b_1 \cdots\cdots$ ($b_i = 1$ if f_i selects this bit). In other words, each bit is characterized by an infinite path in a binary tree. Nevertheless, we only use an initial segment of this path. More precisely, at each stage of our construction one of the vertices of the binary tree is called *active*. To find out the active vertex we start from the root and follow the path until we find a vertex $x_{[0..j]}$ which was active less than n_j times.

We show that the above constructed set A satisfies our requirements by establishing two claims.

Claim 1. *Let f be a p-selection function. Then the subsequence selected by the selection function f satisfies the law of large numbers and there is an unbounded nondecreasing function $r(n)$ satisfying* (5).

Proof. Let $b_0 b_1 \cdots\cdots$ be the infinite subsequence obtained by the application of the selection function $f = f_n$. We consider an arbitrary initial segment $b_0 b_1 \cdots b_t$ of the sequence $b_0 b_1 \cdots\cdots$ and the vertices (strings) of the binary tree corresponding to these bits. Let x be one of the longest strings among these strings (vertices) corresponding to the bits in $b_0 b_1 \cdots b_t$. Without loss of generality, we may assume that $|x| \gg 2^0 2^0 + \cdots + 2^n 2^{2n} > n + 1$. First we give a lower bound of t as a function of $|x|$. If the string

x on the tree T is active, then the string $x' = (x$ without the last bit) on T has been active $2^{2(|x|-2)}$ times. The nth bit of x' is equal to 1 (we assume that $|x| > n+1$), hence all the bits corresponding to x' are selected by f. So the length $t+1$ of $b_0 b_1 \cdots b_t$ is at least $2^{2(|x|-1)}$. Now $b_0 b_1 \cdots b_t$ can be divided into two groups. For the bits in one group the corresponding strings (vertices) have length at most n, the total number of such bits is bounded by $2^0 2^0 + \cdots + 2^n 2^{2n}$, so we may ignore them. For other bits the corresponding strings (vertices) have length greater than n and the nth bit is equal to 1. So the total number of such kind of strings (vertices) used does not exceed $(1+2+\cdots+2^{|x|}) < 2^{|x|+2}$. The difference between the number of zeros and the number of ones in each sequence corresponding to each string (vertex) is at most 3. Thus the difference between the number of ones and the number of zeros in $b_0 b_1 \cdots b_t$ does not exceed $3 \cdot 2^{|x|+2}$. Hence the frequency of ones in $b_0 b_1 \cdots b_t$ is close to $1/2$ (the difference is less than $(3 \cdot 2^{|x|+2})/(2^{2(|x|-2)})$ and tends to zero).

It remains to show the existence of an unbounded nondecreasing function $r(n)$ satisfying (5). This is straightforward because each base sequence in our construction is $111010 \cdots\cdots$. □

Claim 2. *$B = \Sigma^*$ is an optimal unsafe approximation of A. That is to say, for every set $C \in$ **P** such that $\|C \Delta B\| = \infty$,* (2) *holds.*

Proof. Define a p-selection function f by letting

$$f(x) = \begin{cases} 1 & \text{if} \quad C(z_{|x|}) = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then, by (5),

$$\begin{aligned} &\|(A\Delta C)\restriction z_n\| - \|(A\Delta B)\restriction z_n\| \\ &\quad = \|\{i < n\colon\ f(\chi_A[0..i-1]) = 1 = \chi_A[i]\}\| \\ &\qquad - \|\{i < n\colon\ f(\chi_A[0..i-1]) = 1 = 1 - \chi_A[i]\}\| \\ &\quad > 0 \end{aligned}$$

for almost all $n \in N$. Hence (2) holds. □

Before proving the second main theorem of our paper, we first prove a preliminary lemma.

Lemma 4.11. *Let $B_{0,0}, B_{0,1}, B_{1,0}, B_{1,1}, B_{2,0}, B_{2,1}, \ldots$ be a sequence of mutually disjoint sets which has a universal characteristic function in* **E** *such that $\bigcup_{i \in N} \bigcup_{b=0,1} B_{i,b} = \Sigma^*$. Then there is a p-stochastic set $A \in$ $\mathbf{E}_2$ satisfying the following properties.*

1. *For each $i \in N$, let $\alpha_{i,0} = b_0 b_1 b_2 \cdots\cdots$, where*

$$b_j = \begin{cases} A(z_j) & \textit{if} \quad z_j \in B_{i,0}, \\ \lambda & \textit{if} \quad z_j \notin B_{i,0}. \end{cases}$$

If $|\alpha_{i,0}|$ *is infinite, then there is an unbounded nondecreasing function* $r_{i,0}(n)$ *such that* $\|\{j < n\colon \alpha_{i,0}[j] = 0\}\| \ge \|\{j < n\colon \alpha_{i,0}[j] = 1\}\| + r_{i,0}(n)$ *for almost all* $n \in N$.

2. *For each* $i \in N$, *let* $\alpha_{i,1} = b_0 b_1 b_2 \cdots\cdots$, *where*

$$b_j = \begin{cases} A(z_j) & \text{if } \; z_j \in B_{i,1}, \\ \lambda & \text{if } \; z_j \notin B_{i,1}. \end{cases}$$

If $|\alpha_{i,1}|$ *is infinite, then there is an unbounded nondecreasing function* $r_{i,1}(n)$ *such that* $\|\{j < n\colon \alpha_{i,1}[j] = 1\}\| \ge \|\{j < n\colon \alpha_{i,1}[j] = 0\}\| + r_{i,1}(n)$ *for almost all* $n \in N$.

Proof. Let $f_0, f_1, \ldots$ be an enumeration of all p-selection functions. The proof is a nested combination of infinitely many copies of the construction in the proof of Theorem 4.10. That is to say, for each $B_{i,b}$, we construct $\alpha_{i,b}$ in the same way as in the construction of A in the proof of Theorem 4.10. The formal construction is given below.

Let $n_j = 2^{3^j}$ for all $j \in N$, and let

$$\xi_w = 10101010\cdots\cdots \in \Sigma^\infty,$$
$$\xi_{w,j,1} = 1110101010\cdots\cdots \in \Sigma^\infty,$$
$$\xi_{w,j,0} = 00010101010\cdots\cdots \in \Sigma^\infty,$$

for all $w \in \Sigma^*$ and $j \in N$. For $i \in N$, assume that $\chi_A[0..i-1]$ has already been defined. Now we show how to define $\chi_A[i]$. Let j, b be the unique numbers such that $z_i \in B_{j,b}$. If the condition

- for all $k \le j$ such that $f_k(\chi_A[0..i-1]) = 1$, there is a stage $u < i$ such that $f_k(\chi_A[0..u-1]) = 1$ and $\chi_A[u]$ was constructed from $\xi_{w,m,b'}$ or ξ_w for some $|w| \ge 3j$

holds, then we construct $\chi_A[i]$ according to process 2 following, otherwise construct $\chi_A[i]$ according to process 1.

1. Let $x = f_0(\chi_A[0..i-1])\, f_1(\chi_A[0..i-1]) \cdots f_{i-1}(\chi_A[0..i-1])$ and let s be the least integer such that we have used less than n_s bits from $\xi_{x[0..s]}$. Then let $\chi_A[i]$ be the first bit in $\xi_{x[0..s]}$ that we have not used.
2. Let $x = f_0(\chi_A[0..i-1])\, f_1(\chi_A[0..i-1]) \cdots f_{i-1}(\chi_A[0..i-1])$ and let s be the least integer such that we have used less than n_s bits from $\xi_{x[0..s],j,b}$. Then let $\chi_A[i]$ be the first bit in $\xi_{x[0..s],j,b}$ that we have not used.

In the above construction, we have a base tree of binary strings where each vertex corresponds to the infinite binary sequence $1010\cdots\cdots$ and for each $B_{j,b}$ ($j \in N$, $b \in \Sigma$) we have a tree of binary strings where each vertex corresponds to the infinite binary sequence $111010\cdots\cdots$ if $b = 1$ and $0001010\cdots\cdots$ otherwise. At each stage of our constuction, one tree will be called *active*, and one vertex on the active tree will be called *active*. To find out the active tree, first we compute the unique numbers j, b such that $z_i \in B_{j,b}$. If the condition

- for all $k < j$ such that $f_k(\chi_A[0..i-1]) = 1$, there is a stage $u < i$ such that $f_k(\chi_A[0..u-1]) = 1$ and $\chi_A[u]$ was constructed from $\xi_{w,m,b'}$ or ξ_w for some $|w| \geq 3j$

holds, then the tree corresponding to $B_{j,b}$ will be active at stage i, otherwise the base tree will be active. To find out the active vertex on the active tree, it is the same as in the proof of Theorem 4.10.

For each $j \in N$ and $b \in \Sigma$, there is a number $i_{j,b}$ such that the tree corresponding to $B_{j,b}$ will be active at any stage $i > i_{j,b}$ when $z_i \in B_{j,b}$. Hence, in the same way as in the proof of Theorem 4.10, it is easily checked that properties 1 and 2 of the lemma are satisfied.

Now it remains to show that the above constructed set A is p-stochastic. That is to say, we need to show that each selection function f_n selects a balanced subsequence.

Let $b_0 b_1 \cdots$ be the infinite subsequence obtained by the application of the selection function f_n. We consider an arbitrary initial segment $b_0 b_1 \cdots b_t$ of the sequence $b_0 b_1 \cdots$ and the vertices (strings) of the binary trees corresponding to these bits. Let x be one of the longest strings among these strings (vertices) corresponding to the bits in $b_0 b_1 \cdots b_t$. Then, by the construction, the number of trees which correspond to these bits is not greater than $|x|/3$. Without loss of generality, we may assume that $|x| \gg 2^0 2^0 + \cdots + 2^n 2^{2n} > n+1$. First we give a lower bound of t as a function of $|x|$. If the string x on the tree T is used as active, then the string $x' = (x$ without the last bit) on T is used as active for $2^{3(|x|-2)}$ times. The nth bit of x' is equal to 1 (we assume that $|x| > n+1$), hence all the bits corresponding to x' are selected by f_n. So the length $t+1$ of $b_0 b_1 \cdots b_t$ is at least $2^{3(|x|-2)}$. Now $b_0 b_1 \cdots b_t$ can be divided into two groups. For some of them the corresponding strings (vertices) have length at most n, the total number of such bits is bounded by $(2^0 2^0 + \cdots + 2^n 2^{3n}) \cdot |x|/3$, so we may ignore them. For other bits the corresponding strings (vertices) have length greater than n and the nth bit is equal to 1. So the total number of such kind of strings (vertices) used does not exceed $(1 + 2 + \cdots + 2^{|x|}) \cdot |x|/3 < 2^{2|x|}$. The difference between the number of zeros and the number of ones in each sequence corresponding to each string (vertex) is at most 3. Thus the difference between the number of ones and the number of zeros in $b_0 b_1 \cdots b_t$ does not exceed $3 \cdot 2^{2|x|}$. Hence the frequency of ones in $b_0 b_1 \cdots b_t$ is close to 1/2 (the difference is less than $(3 \cdot 2^{2|x|})/(2^{3(|x|-2)})$ and tends to zero). $\square$

Now we are ready to prove our other main theorem.

Theorem 4.12. *There is a p-stochastic set A in $\mathbf{E}_2$ which is Δ-levelable.*

Proof. Let $P_0, P_1, P_2, \ldots$ be an enumeration of all sets in $\mathbf{P}$. For $i \in N$ and $b \in \Sigma$, let $B_{i,b} = \{z_{\langle i,j \rangle} \colon j \in N$ and $P_i(z_{\langle i,j \rangle}) = 1 - b\}$. Let $A \in \mathbf{E}_2$ be the p-stochastic set in Lemma 4.11 corresponding to the sequence $B_{0,0}, B_{0,1}, B_{1,0}, B_{1,1}, B_{2,0}, B_{2,1}, \ldots$ of sets. We have to show that A is Δ-levelable. For each infinite set P_i, define a polynomial time

Table 1. The relations among randomness, stochasticity, and approximations.

	p-Random	p-Stochastic	Δ-Levelable	Weakly Δ-levelable	Optimally approximable
A	Yes	Yes	No	Yes	No
B	No	Yes	Yes	Yes	No
C	No	Yes	No	No	Yes
D	No	No	Yes	Yes	No
E	No	No	No	No	Yes

computable set P_i' by letting

$$P_i'(z_n) = \begin{cases} 1 - P_i(z_n) & \text{if} \quad n = \langle i, j\rangle \quad \text{for some} \quad j \in N, \\ P_i(z_n) & \text{otherwise.} \end{cases}$$

It suffices to show that (1) holds with P_i and P_i' in place of B and B', respectively. Let $\alpha_{i,0}$ and $\alpha_{i,1}$ be defined as in Lemma 4.11. Then at least one of them is an infinite sequence. Without loss of generality, we may assume that $\alpha_{i,0}$ is infinite and $\alpha_{i,1}$ is finite. By Lemma 4.11, there is an unbounded nondecreasing function $r_{i,0}(n)$ such that $\|\{j < n\colon \alpha_{i,0}[j] = 0\}\| \geq \|\{j < n\colon \alpha_{i,0}[j] = 1\}\| + r_{i,0}(n)$ for almost all $n \in N$. Hence

$$\begin{aligned} &\|(A \Delta P_i)\lceil z_n\| - \|(A \Delta P_i')\lceil z_n\| \\ &\qquad \geq \|\{j < n_1\colon \alpha_{i,0}[j] = 0\}\| - \|\{j < n_1\colon \alpha_{i,0}[j] = 1\}\| - |\alpha_{i,1}| \\ &\qquad \geq r_{i,0}(n_1) - |\alpha_{i,1}| \end{aligned}$$

for almost all $n \in N$, where $n_1 = \|\{j < n\colon\ j = \langle i, k\rangle \text{ for some } k \in N \text{ and } P_i(z_j) = 1\}\|$. That is to say, (1) holds with P_i, P_i', and $r_{i,0}(n_1) - |\alpha_{i,1}|$ in place of B, B', and $f(n)$, respectively. □

Our results in this paper show that p-randomness implies weak Δ-levelability, but it implies neither Δ-levelability nor optimal approximability. However, p-stochasticity is independent of weak Δ-levelability, Δ-levelability, and optimal approximability.

As a summary, we list all these relations among randomness, stochasticity, and approximations. There are sets $A, B, C, D, E \subset \Sigma^*$ which satisfy the properties in Table 1.

Acknowledgments

I am grateful to Professor Richard Beigel for some comments on an early version of this paper. I would also like to thank Professor Harald Ganzinger who supported my stay at Max-Planck-Institut für Informatik (Saarbrücken, Germany) where part of this work was done.

References

[1] K. Ambos-Spies. On optimal polynomial time approximations: **P**-levelability vs. Δ-levelability. In *Proc*. 22*nd ICALP*, Lecture Notes in Computer Science, 944, pages 384–392. Springer-Verlag, Berlin, 1995.

[2] K. Ambos-Spies and J. Lutz. The Workshop on Randomness and Information, Dagstuhl, July 15–19, 1996.

[3] K. Ambos-Spies, E. Mayordomo, Y. Wang, and X. Zheng. Resource-bounded balanced genericity, stochasticity and weak randomness. In *Proc*. 13*th STACS*, Lecture Notes in Computer Science, 1046, pages 63-74. Springer-Verlag, Berlin, 1996.

[4] A. Church. On the concept of a random sequence. *Bull*. *Amer. Math*. *Soc*., 45:130–135, 1940.

[5] P. Duris and J. D. P. Rolim. **E**-complete sets do not have optimal polynomial time approximations. In *Proc*. 19*th MFCS*, Lecture Notes in Computer Science, 841, pages 38–51. Springer-Verlag, Berlin, 1994.

[6] K. Ko. On the notion of infinite pseudorandom sequences. *Theoret*. *Comput*. *Sci*., 48:9–33, 1986.

[7] J. H. Lutz. Almost everywhere high nonuniform complexity. *J. Comput*. *System Sci*., 44:220–258, 1992.

[8] P. Martin-Löf. The definition of random sequences. *Inform*. *and Control*, 9:602–619, 1966.

[9] A. R. Meyer and M. S. Paterson. With what frequency are apparently intractable problems difficult? Technical Report TM-126, Laboratory for Computer Science, MIT, 1979.

[10] R. von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Math*. *Z*., 5:52–99, 1919.

[11] C. P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit*. *Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*. Lecture Notes in Mathematics, 218. Springer-Verlag, Berlin, 1971.

[12] V. A. Uspenskii, A. L. Semenov, and A. Kh. Shen. Can an individual sequence of zeros and ones be random? *Russian Math*. *Surveys*, 45:121–189, 1990.

[13] J. Ville. *Ètude Critique de la Notion de Collectif*. Gauthiers-Villars, Paris, 1939.

[14] A. Wald. Sur la notion de collectif dans le calcul des probabilités. *C*. *R*. *Acad*. *Sci*. *Paris*, 202:180–183, 1936.

[15] Y. Wang. Randomness and Complexity. Ph.D. thesis, Heidelberg, 1996.

[16] Y. Wang. The law of the iterated logarithm for p-random sequences. In *Proc*. 11*th Conf*. *on Computational Complexity* (formerly Conf. on Structure in Complexity Theory), pages 180-189. IEEE Computer Society Press, Los Alamitos, CA, 1996.

[17] Y. Wang. Genericity, randomness, and polynomial time approximations. *SIAM J. Comput*. 28(2):394–408, 1999.

[18] R. Wilber. Randomness and the density of hard problems. In *Proc*. 24*th Symp*. *FOCS*, pages 335–342. IEEE Computer Society Press, Los Alamitos, CA, 1983.

[19] Y. Yesha. On certain polynomial-time truth-table reducibilities of complete sets to sparse sets. *SIAM J*. *Comput*., 12:411–425, 1983.

Received May 1997, *and in final form September* 1998.