J. Cryptology (1993) 6: 119–133

Journal of
CRYPTOLOGY

Universal Tests for Nonuniform Distributions

A. W. Schrift and A. Shamir
Department of Applied Mathematics and Computer Science,
The Weizmann Institute of Science, Rehovot 76100, Israel

Communicated by Claude Crépeau

Received 11 January 1991 and revised 23 January 1992

Abstract. The next bit test as introduced by Blum and Micali was shown by Yao to be a universal test for sources of unbiased independent bits. The aim of this paper is to provide a rigorous methodology for testing sources whose output distributions are not necessarily uniform. We first show that the natural extension of the next bit test, even in the simplest case of biased independent bits, is no longer universal: we construct a source of biased bits, whose bits are obviously dependent and yet none of these bits can be predicted with probability of success greater than the bias. To overcome this difficulty, we develop new universal tests for arbitrary models of (potentially imperfect) sources of randomness. These new tools contribute to the theoretical as well as practical study of sources of randomness.

Key words. Universal test, Next bit test, Nonuniform distribution, Source of randomness, Independent biased source.

1. Introduction

Randomness is an essential resource in many scientific areas, and pseudorandomness is a good substitute in many applications. Blum and Micali [3] were first to show that the ability to predict some bit of a given source (the *next bit test*) can be used to characterize pseudorandom generators. In his seminal paper Yao [14] formally defines the notion of perfect pseudorandom bits, i.e., bits that are indistinguishable from truly random bits by any probabilistic polynomial-time observer, and shows that the next bit test serves as a universal test for randomness: a natural or pseudorandom source is perfect iff no probabilistic polynomial-time algorithm can, given any prefix of bits, predict the next bit of the source with probability of success significantly greater than 1/2.

Several models of natural sources of randomness have been suggested and investigated in many works, such as [13], [2], [8], and [5]. In all the models the output distribution of natural sources is not uniform: In [13] a natural source outputs biased independent bits, in [2] a source is modeled by a Markov chain, and in [8]

and [5] the outcomes of the source are controlled by a powerful adversary. Nonuniform distributions appear also in some applications which require sources of randomness with independent yet biased bits (see, for example, [12] and [6]). Nevertheless, no rigorous methodology of how to verify the assumed properties of a source of randomness with a nonuniform output distribution has been given. The aim of our paper is to provide such a formalization.

Consider, for example, the roulette in your favorite casino, where you are in the habit of placing a variety of bets on 17 with a 1/37 probability of winning each time. However, after an unfortunate series of losses you begin to suspect that the roulette has been tampered with. You can easily check that the overall probability of 17 is close to 1/37, but that does not rule out the possibility that the outcomes of the roulette are artificially determined in a way that maintains the overall bias but inhibits 17 from appearing whenever the bets are high. How can you verify that indeed the outcomes of the roulette are independent, and that it is only your bad luck that brought you to the edge of bankruptcy? Clearly, the well-known next bit test cannot be employed here since you deal with a biased event.

Using the known notion of polynomial indistinguishability we define the notions of *perfect independence* and in general *perfect simulation* of a source by a mathematical model. We then move to the question of specifying the universal tests for these notions, which let only perfect sources pass the universal test. Surprisingly, the natural extension of the next bit test fails, even for the simplest case of independent biased bits. In other words, the extended next bit test for biased bits, which requires that no observer succeeds in predicting the bits of the source with probability greater than the bias, is no longer a universal test for independence. We introduce a new test of independence, which we call *the predict or pass* (*POP*) *test* and prove its universality. We also discuss several alternative tests, and in particular the test we call the *weighted success rate* (*WSR*) *test*.

For general sources of randomness we present a universal test that determines whether a certain mathematical model perfectly simulates a given source. This test is the *comparative version* of the next bit test. The standard next bit test as well as the POP and WSR tests can be derived from the comparative next bit test as special cases. Our proof of the universality of the test is a generalization of Yao's original proof: while the original proof techniques cannot be implemented directly, our refined techniques apply also to the proof of universality of Yao's next bit test.

The rigorous treatment of our universal tests has several theoretical as well as practical applications. In [9] we present some new results that the new tools make possible: An improved definition of the quality of natural bits is given. This in turn is used to measure the tradeoff between the quality of bits extracted from a given source and their quantity. Another application is in the modeling of natural sources of randomness from an external point of view without knowledge of their internal structure. It is also possible to apply the universal test of independence to every biased predicate and use a hard biased predicate to construct a generator of independent biased bits. For constant output length this is a more efficient construction of perfect independent biased bits than the obvious construction of rebiasing the outputs of perfect pseudorandom (unbiased) generators. Finally, in [10] the individual security of every bit of the discrete logarithm modulo a composite is

proven. The known definitions of unpredictability cannot be applied to the most significant bits, since for moduli which are not powers of 2 these bits are biased toward 0 by definition. For these bits it is necessary to use our new definitions in order to define and prove their security.

2. Definitions and Notations

Our definitions follow the original definitions of Blum and Micali [3] and Yao [14]. The notions of a probability distribution, independence, etc., are the standard notions from probability theory. All our results are stated in terms of probabilistic polynomial-time algorithms but can be restated in terms of polynomial-size Boolean circuits.

Let s_1^n denote a binary string of length n in $\{0, 1\}^n$. The ith bit of the string is denoted by s_i. The substring starting with the jth bit and ending with the kth bit $(1 \leq j < k \leq n)$ is denoted by s_j^k. We use the notation $O(v(n))$ for any function $f(n)$ that vanishes faster than any polynomial, i.e., for every polynomial $poly(n)$ and n large enough, $f(n) < 1/poly(n)$. Such functions are called *negligible*.

Definition 1 (Source Ensemble). A source ensemble S is a sequence $\{S_n\}$, where S_n is a probability distribution on $\{0, 1\}^n$.

We denote by $\Pr_S(E)$ the probability of an event E taking place when the probability distribution is defined by the source ensemble S. Whenever we refer to events that involve a probabilistic algorithm, we explicitly denote only the source ensemble S, and implicitly assume the probability of the event to be induced by S and by the independent unbiased coin flips of the algorithm.

Definition 2 (Uniform Source Ensemble). A source ensemble S is uniform if, for every n, S_n is the uniform probability distribution, i.e., for every $\alpha \in \{0, 1\}^n$,

$$\Pr_S(s_1^n = \alpha) = 2^{-n}.$$

We denote the uniform source ensemble by U.

Definition 3 (Biased Source Ensemble). A source ensemble S is biased toward 1 with a fixed bias $\frac{1}{2} \leq b < 1$ if, for every i, $\Pr_S(s_i = 1) = b$.

Note that by our restriction on the bias, the output bits of a biasd source have a nonzero probability of being both 0 and 1. This ensures that the definitions of conditional probabilities, dependencies, etc., remain meaningful.

Definition 4 (Independent Biased Ensemble). A source ensemble S is independent biased if it is a biased source ensemble and all the bits are independent, i.e., for every binary string $\alpha \in \{0, 1\}^n$,

$$\Pr_S(s_1^n = \alpha) = b^\rho \cdot (1 - b)^{n-\rho},$$

where $0 \leq \rho \leq n$ denotes the number of 1's in α.

We denote by B the independent biased source ensemble.

Definition 5 (Polynomial Indistinguishability). Two source ensembles S_1 and S_2 are polynomially indistinguishable if, for every probabilistic polynomial-time algorithm (distinguisher) D: $\{0, 1\}^n \to \{0, 1\}$,

$$|\Pr_{S_1}(D = 1) - \Pr_{S_2}(D = 1)| \leq O(v(n)).$$

Definition 6 (Constant Algorithm). A probabilistic polynomial-time algorithm is constant if, for some value v, $\Pr(\textit{algorithm} = v) > 1 - O(v(n))$ for all inputs.

3. Universal Tests of Independence

In this section we deal with universal tests of independence.

Definition 7 (Perfect Independence). A source S is a perfect independent biased source with some fixed bias b if it is polynomially indistinguishable from the independent biased source ensemble B with the same bias b.

We first construct what seems to be the natural extension of the standard next bit test. We then show that there exist imperfect sources of randomness that pass the extended next bit test, thus disproving its universality. Our proof is based on the following intuition: dependencies between the bits of an imperfect source will result in 1 having in some cases probability greater than the bias and in other cases probability smaller than the bias. It is possible, however, for the biased source to be imperfect with 1 remaining always more probable than 0, regardless of the preceding bits. Hence, deterministically predicting 1 is the optimal prediction strategy but its probability of success cannot exceed the bias.

In the following subsections we assume without loss of generality that all our sources are biased toward 1 with some fixed bias b. It is easy to extend our results to the case where each bit has a different bias. It is worthwhile emphasizing that since we are interested in detecting dependencies among bits that have a particular bias, our basic WSR test may fail to detect imperfectness that results simply from a different overall bias. Testing the condition that the bits of a source have a certain bias can be done easily in polynomial time and with high accuracy using the law of large numbers. The POP test has the additional feature that any deviation from the *a priori* known bias is automatically detected.

3.1. *The Extended Next Bit Test*

Trying to extend the definition of the well-known next bit test to biased sources we must take into consideration the fact that the bits of an independent biased source can be trivially predicted with probability of success b, simply by always predicting 1.

Definition 8 (To Pass the Extended Next Bit Test). A biased source S passes the extended next bit test if, for every $1 \leq i \leq n$ and for every probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$,

$$\Pr_S(A(s_1^{i-1}) = s_i) \leq b + O(v(n)).$$

Theorem 1. *The extended next bit test is not a universal test of independence.*

Proof. Fix a bias $b: \frac{1}{2} + 1/n^{t_1} \leq b \leq 1 - 1/n^{t_2}$ for some constants t_1 and t_2. We construct a source which is biased toward 1 with bias b. We then show that it is imperfect and yet it passes the extended next bit test. The source is the following:

$$\Pr_S(s_i = 1) = \begin{cases} b & \text{for } 1 \leq n \leq n-1, \\ b+\delta & \text{for } i = n \text{ and } s_1^2 = 01, \\ b-\delta & \text{for } i = n \text{ and } s_1^2 = 10, \\ b & \text{for } i = n \text{ and } s_1^2 = 00 \text{ or } 11, \end{cases}$$

where $1/n^q \leq \delta < \min(b - \frac{1}{2}, 1 - b)$ for some constant $q > \max\{t_1, t_2\}$.

Let a polynomial-time distinguisher D be defined by $D = 1$ iff $s_1^2 = 01$ and $s_n = 1$. Clearly, $\Pr_S(D = 1) = c \cdot (b + \delta)$, while $\Pr_B(D = 1) = c \cdot b$, where $c = b \cdot (1 - b) \geq 1/2n^{t_2}$. Therefore, $\Pr_S(D = 1) - \Pr_B(D = 1) = c \cdot \delta \geq 1/2n^{q+t_2}$, and by definition the source is imperfect. Nevertheless, the source passes the extended next bit test: the nth bit is always biased toward 1, so the best prediction strategy is to predict 1 deterministically regardless of the known values of the first two bits. It is easy to check that the probability of success of this optimal strategy remains b. □

3.2. *The Predict Or Pass Test*

Definition 9 (To Pass the POP Test). A biased source S passes the predict or pass (POP) test if, for every $1 \leq i \leq n$, for every fixed l, and every probabilistic polynomial-time algorithm $A: \{0, 1\}^{i-1} \to \{0, 1, *\}$, if $\Pr_S(A(s_1^{i-1}) \neq *) \geq 1/n^l$, then

$$|\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) \neq *) - b| \leq O(v(n)).$$

The POP test allows a predictor to be successful only on some nonnegligible fraction of its inputs. Despite the fact that this formal definition is novel (as far as we know), known constructions of pseudorandom bit generators often prove their perfectness by showing that they pass what is essentially a POP test (i.e., it is impossible to predict correctly the output bits of the generator even on a nonnegligible fraction of the output strings). In what follows we prove the POP test to be a universal test of independence. Before formally stating and proving this result we first introduce some alternative tests and in particular the WSR test. In fact, we first prove the universality of the WSR test and only then derive the universality of all other tests.

3.3. *Alternative Versions*

Definition 10 (Weighted Success Rate). Fix $1 \leq i \leq n$. The weighted success rate of any nonconstant probabilistic polynomial-time algorithm $A: \{0, 1\}^{i-1} \to \{0, 1\}$

in predicting the ith bit of a biased source S is

$$ws(A, S, i) = \frac{\Pr_S(A(s_1^{i-1}) = s_i | s_i = 1)}{\Pr_S(A(s_1^{i-1}) = 1)} + \frac{\Pr_S(A(s_1^{i-1}) = s_i | s_i = 0)}{\Pr_S(A(s_1^{i-1}) = 0)}$$

$$= \frac{1}{b} \cdot \Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1)$$

$$+ \frac{1}{1-b} \cdot \Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0).$$

Definition 11 (To Pass the WSR Test). A biased source S passes the weighted success rate test, if for every $1 \le i \le n$ and every nonconstant probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$,

$$ws(A, S, i) \le 2 + O(v(n)).$$

Note. 1. We give two alternative expressions for the weighted success rate. They are equivalent since

$$\Pr_S(A = s_i | s_i = 1) \cdot \Pr_S(s_i = 1) = \Pr_S(A = s_i | A = 1) \cdot \Pr_S(A = 1),$$

$$\Pr_S(A = s_i | s_i = 0) \cdot \Pr_S(s_i = 0) = \Pr_S(A = s_i | A = 0) \cdot \Pr_S(A = 0),$$

and

$$\Pr_S(s_i = 1) = b \qquad (\Pr_S(s_i = 0) = 1 - b).$$

2. The above definitions do not allow constant prediction algorithms. Remember that we assume that indeed all the tested sources of randomness have a bias b. Since constant algorithms can only detect that the overall bias is other than b, which is not the case, it is possible without loss of generality to ignore them.

Definition 12 (To Pass the Modified WSR Test). A biased sources S passes the modified WSR test if, for every $1 \le i \le n$ and every nonconstant probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$,

$$\max\left\{\frac{\Pr_S(A(s_1^{i-1}) = s_i | s_i = 1)}{\Pr_S(A(s_1^{i-1}) = 1)}, \frac{\Pr_S(A(s_1^{i-1}) = s_i | s_i = 0)}{\Pr_S(A(s_1^{i-1}) = 0)}\right\}$$

$$= \max\left\{\frac{1}{b}\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1), \frac{1}{1-b}\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0)\right\}$$

$$\le 1 + O(v(n)).$$

Definition 13 (To Pass the Behavior Test). A biased source S passes the behavior test if, for every $1 \le i \le n$ and every probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$,

$$|\Pr_S(A(s_1^{i-1}) = 1 | s_i = 1) - \Pr_S(A(s_1^{i-1}) = 1 | s_i = 0)| \le O(v(n)).$$

Note that the WSR test can detect deviations from the bias since we separately compute the probabilities of success in predicting the 0 and 1 values of a next bit,

and compose the two terms with appropriate weights into a single measure. While the preceding three definitions are all closely related, the POP test presents an entirely different approach, which stems from the fact that if a source is imperfect it is possible to detect a nonnegligible fraction of the events in which 1 is more probable than the given bias, and ignore all other events.

Theorem 2. *The following conditions are equivalent*:

1. *A biased source is a perfect independent biased source.*
2. *A biased source passes the POP test.*
3. *A biased source passes the WSR test.*
4. *A biased source passes the modified WSR test.*
5. *A biased source passes the behavior test.*

The above equivalence holds only for biased sources that were *a priori* tested to have a certain bias. Otherwise, the POP test and the behavior test behave differently from the other tests. Their definitions allow constant as well as nonconstant prediction algorithms. More important is the fact that unlike the WSR test, these tests succeed in detecting imperfectness that results merely from a different overall bias.

The proof of the theorem is given in the Appendix. In this proof we present and prove the following useful lemma:

Prediction Lemma. *For any biased source and any nonconstant probabilistic polynomial-time algorithm* $A: \{0,1\}^{i-1} \to \{0,1\}$ *there exists a constant* k_1 *such that*

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq b + \frac{1}{n^{k_1}}$$

iff there exists a constant k_2 *such that*

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq 1 - b + \frac{1}{n^{k_2}}.$$

3.4. *Comparison with the Next Bit Test*

For unbiased ($b = \frac{1}{2}$) independent bits the POP test and its variatons all serve as alternative universal tests to the next bit test. We can, however, show an even stronger equivalence between the tests, namely that the same algorithm that succeeds in the prediction of a certain bit with probability significantly greater than 1/2 (thus proving the source of the bits to be imperfect by the well-known next bit test) has a weighted success rate that is significantly greater than 2 (thus proving the source to be imperfect by the WSR test). We can also show that our new universal tests are superior to the next bit test in terms of the conditional probability of correct predictions (provided that the test does not pass).

Proposition 3. *For any unbiased source* ($b = \frac{1}{2}$) *and any nonconstant probabilistic polynomial-time algorithm* $A: \{0,1\}^{i-1} \to \{0,1\}$ *there exists a constant* k_1 *such that* $\Pr_S(A(s_1^{i-1}) = s_i) \geq \frac{1}{2} + 1/n^{k_1}$ *iff there exists a constant* k_2 *such that* $ws(A, S, i) \geq 2 + 1/n^{k_2}$.

Proof. Let A be any nonconstant probabilistic polynomial-time algorithm: $A: \{0,1\}^{i-1} \to \{0,1\}$. Clearly,

$$\Pr_S(A(s_1^{i-1}) = s_i) = \Pr_S(A(s_1^{i-1}) = 1) \cdot \Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) + \Pr_S(A(s_1^{i-1}) = 0) \cdot \Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0). \quad (1)$$

The proposition results from the following two easily proved equivalences. We sketch their proofs in parentheses:

1. $\Pr_S(A(s_1^{i-1}) = s_i) \geq \frac{1}{2} + 1/n^{k_1}$ iff $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq \frac{1}{2} + 1/n^{l_1}$ and $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq \frac{1}{2} + 1/n^{l_2}$ for some constants l_1 and l_2. (If $\Pr_S(A(s_1^{i-1}) = s_i) \geq \frac{1}{2} + 1/n^{k_1}$, then by (1) there exists a value $\sigma \in \{0,1\}$ such that $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = \sigma) \geq \frac{1}{2} + 1/n^{k_1}$. This in turn implies the equivalence according to the Prediction Lemma. The other direction is an immediate consequence of (1).)
2. $ws(A, S, i) \geq 2 + 1/n^{k_2}$ iff $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq \frac{1}{2} + 1/n^{l_1}$ and $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq \frac{1}{2} + 1/n^{l_2}$ for some constants l_1 and l_2. (A direct result from the proof of Theorem 2.) □

Proposition 4. *For any unbiased source and any next bit test T there exists a POP test A, such that, for every $1 \leq i \leq n$,*

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) \neq *) \geq \Pr_S(T(s_1^{i-1}) = s_i).$$

Furthermore, for some unbiased sources there exists a POP test A such that, for every $1 \leq i \leq n$,

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) \neq *) > \Pr_S(T(s_1^{i-1}) = s_i).$$

Proof. It is obvious that a POP test can always simulate a next bit test (without ever outputting $*$). We now show that inequality is also possible. To do so we construct an imperfect source S and demonstrate a POP test that does better than any next bit test. The source is the following:

1. The first $n-1$ bits are independent unbiased coin flips.
2. Fix any $0 \leq \delta \leq \frac{1}{2}$.

$$\Pr_S(s_n = 1) = \begin{cases} \frac{1}{2} + \delta & \text{if } s_1^2 = 00, \\ \frac{1}{2} - \delta & \text{if } s_1^2 = 01, \\ \frac{1}{2} & \text{if } s_1 = 1. \end{cases}$$

Since the next bit test is a global test, for any next bit test T:

$$\Pr_S(T(s_1^{n-1}) = s_n) \leq \tfrac{1}{4} \cdot (\tfrac{1}{2} + \delta) + \tfrac{1}{4} \cdot (\tfrac{1}{2} + \delta) + (\tfrac{1}{2})^2 = \frac{1}{2} + \frac{\delta}{2}.$$

The POP test A we use is $A = 1$ iff $s_1^2 = 00$; else $A = *$. Clearly,

$$\Pr_S(A(s_1^{n-1}) = s_n | A(s_1^{n-1}) \neq *) = \tfrac{1}{2} + \delta.$$

□

4. Perfectness with Respect to Arbitrary Models

In this section we consider an arbitrary source S, which we believe to have a certain distribution described by a mathematical model ensemble M. As in the previous tests of randomness and independence we search for a convenient universal test, based on the probability of correct predictions:

Definition 14 (Perfect Simulation). A model M is a perfect simulation of a source S if S and M are polynomially indistinguishable.

Definition 15 (To Pass the Comparative Next Bit Test). A source S passes the comparative next bit test with respect to a model M if, for every $1 \leq i \leq n$ and every probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$,

$$|\Pr_S(A(s_1^{i-1}) = s_i) - \Pr_M(A(s_1^{i-1}) = s_i)| \leq O(v(n)).$$

Note that the comparative next bit test enables us to avoid performing any *a priori* tests on either source. The test is easiest to implement when the model is described in such a way that the exact probability of bit predictions in the model can be efficiently computed. Yet we can perform the test even when the model is completely unknown and given to us as a black box. In that case the test simply involves a comparison between two boxes: one containing the tested source and the other containing the model.

It is instructive to examine simple examples of the comparative next bit test, when the model source is explicitly known:

1. $M = U$, i.e., the model is the uniform source of unbiased independent bits. In that case we know that no matter which algorithm is used $\Pr_U(A = s_i) = \frac{1}{2}$ and we can immediately derive the well-known next bit test.
2. $M = B$, i.e., the model is a source of biased independent bits. Here we know that for any nonconstant algorithm $\Pr_B(A = s_i | A = 1) = b$ and that $\Pr_B(A = s_i | A = 0) = 1 - b$ so that the predictions must be evaluated separately according to the value that is being predicted. This gives rise to the POP or WSR test.
3. M is a source with a one-bit memory, in which the probability of the ith bit is determined according to the outcome of the $(i-1)$th bit. Let $b_i(0) = \Pr_M(s_i = 1 | s_{i-1} = 0)$ and $b_i(1) = \Pr_M(s_i = 1 | s_{i-1} = 1)$. Then it is easy to see that the performance of any algorithm must be evaluated not only according to the value of s_i but also according to the value of s_{i-1}. We therefore get that M is a perfect simulation of a source S if, for every $1 \leq i \leq n$ and every probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1, *\}$ such that $\Pr_S(A(s_1^{i-1}) \neq * | s_{i-1} = 0) \geq 1/n^{l_1}$ and $\Pr_S(A(s_1^{i-1}) \neq * | s_{i-1} = 1) \geq 1/n^{l_2}$ for some constants l_1, l_2,

$$\max\{|\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) \neq *, s_{i-1} = 0) - b_i(0)|,$$
$$|\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) \neq *, s_{i-1} = 1) - b_i(1)|\} \leq O(v(n)).$$

It is easy to see that similar analysis holds for any M which is a Markov chain [2], where predictions must be evaluated according to the output value and to the state (which determines the bias).

Theorem 5. *A model M is a perfect simulation of a source S iff S passes the comparative next bit test with respect to M.*

Proof. It is easy to see that if a source S fails the comparative next bit test it is distinguishable from the model source M. Assume now that we are given that S and M are distinguishable and need to prove that S fails the comparative next bit test with respect to M. We cannot implement the proof techniques used for proving Theorem 2 since they inherently assume independence in concatenating a random prefix of bits taken out of the tested source with a random suffix of bits generated according to the model distribution. We overcome the problem by using an additional truly random source for the concatenation. The following can therefore be considered a generalization of the proof of universality of the next bit test as given for example in [4].

Let D: $\{0, 1\}^n \to \{0, 1\}$ be a distinguisher between S and M for which

$$|\Pr_S(D(s_1^n) = 1) - \Pr_M(D(s_1^n) = 1)| \geq 1/n^k \qquad \text{for some constant } k.$$

Let p_i^S (p_i^M) denote the probability that D outputs 1 when the first i bits of its input are taken out of S (M) and the rest are independent unbiased coin flips. Let $d_i = p_i^S - p_i^M$. Note that $p_n^S = \Pr_S(D(s_1^n) = 1)$, $p_n^M = \Pr_M(D(s_1^n) = 1)$, and $p_0^S = p_0^M = \Pr_U(D(s_1^n) = 1)$. Since $d_0 = 0$ and $|d_n| = |p_n^S - p_n^M| \geq 1/n^k$, by the pigeonhole principle there exists an i for which p_i^S and p_i^M significantly differ, i.e., $|d_i - d_{i-1}| \geq 1/n^{k+1}$. We can assume without loss of generality that $d_i > 0$. The comparative next bit test A submits to D the string $s_1^n = s_1^{i-1}s_i^n$, where $s_1^{i-1} \in S$ or M and $s_i^n \in U$. If $D(s_1^n) = 1$, then $A(s_1^{i-1})$ outputs s_i, else $A(s_1^{i-1})$ outputs $1 - s_i$.

Let s_1^i denote a sequence of bits taken out of S or M. Let q^S (q^M) denote the probability that D outputs 1 when the first $i - 1$ bits of its input equal those taken out of S (M), the ith bit of its input is $1 - s_i$ and the rest are independent unbiased coin flips. It is easy to see that

$$p_{i-1}^S = \frac{p_i^S + q^S}{2},$$

$$p_{i-1}^M = \frac{p_i^M + q^M}{2}.$$

Therefore,

$$\Pr_S(A(s_1^{i-1}) = s_i) = \tfrac{1}{2}p_i^S + \tfrac{1}{2}(1 - q^S) = \tfrac{1}{2} + p_i^S - p_{i-1}^S.$$

While

$$\Pr_M(A(s_1^{i-1}) = s_i) = \tfrac{1}{2}p_i^M + \tfrac{1}{2}(1 - q^M) = \tfrac{1}{2} + p_i^M - p_{i-1}^M.$$

Hence,

$$\Pr_S(A(s_1^{i-1}) = s_i) - \Pr_M(A(s_1^{i-1}) = s_i) \geq \frac{1}{n^{k+1}}.$$

□

Appendix. Proof of Theorem 2

We prove the theorem by first proving the WSR test to be a universal test of independence (Proposition 2.1). We then use this proof to show the universality of the other tests (Propositions 2.2–2.4).

Proposition 2.1. *A biased source is a perfect independent biased source iff it passes the WSR test.*

Proof. Given that a source fails the weighted success rate test, it is easy to construct a distinguisher between the source S and the independent biased source B by examining the predictions of the test. Formally assume that we are given a non-constant probabilistic polynomial-time algorithm A: $\{0,1\}^{i-1} \to \{0,1\}$ for the ith bit of a source S such that $ws(A, S, i) \geq 2 + 1/n^k$ for some constant k. We use A to construct two possible distinguishers and show that for one of them $|\Pr_S(D=1) - \Pr_B(D=1)| \geq 1/n^{k'}$ for some constant k'. Given s_1^n, both D's submit s_1^{i-1} to A and examine A's output. $D_1(s_1^n) = 1$ iff $A(s_1^{i-1}) = s_i = 1$. $D_2(s_1^n) = 1$ iff $A(s_1^{i-1}) = 1$. If the overall behavior of A is the same for S and for B, i.e., $|\Pr_S(A(s_1^{i-1})=1) - \Pr_B(A(s_1^{i-1})=1)| \leq O(v(n))$, then D_1 distinguishes between S and B. Otherwise D_2 distinguishes. Hence S is imperfect.

To prove the other direction, we show how to construct a weighted success rate test using any distinguisher D for an imperfect source. Let p_i denote the probability that $D(s_1^n) = 1$ when the first i input bits are taken out of S and the rest are independent biased coin flips (taken from B). Note that $p_n = \Pr_S(D(s_1^n) = 1)$, while $p_0 = \Pr_B(D(s_1^n) = 1)$. Since D distinguishes between S and B, $|p_0 - p_n| \geq 1/n^k$ for some k. By the pigeonhole principle there exists a bit i for which $|p_i - p_{i-1}| \geq 1/n^{k+1}$. We assume without loss of generality that $p_i - p_{i-1} > 0$. Explicitly,

$$\begin{aligned}
p_i &= \sum_{s_1^n} \Pr(D(s_1^n)=1)\cdot \Pr_S(s_1^i)\cdot \Pr_B(s_{i+1}^n) \\
&= \sum_{s_1^{i-1}, s_{i+1}^n} [\Pr(D(s_1^{i-1}1 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_S(s_i=1|s_1^{i-1})\cdot \Pr_B(s_{i+1}^n) \\
&\qquad + \Pr(D(s_1^{i-1}0 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_S(s_i=0|s_1^{i-1})\cdot \Pr_B(s_{i+1}^n)], \\
p_{i-1} &= \sum_{s_1^n} \Pr(D(s_1^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_B(s_i^n) \\
&= \sum_{s_1^{i-1}, s_{i+1}^n} [\Pr(D(s_1^{i-1}1 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_B(s_i=1)\cdot \Pr_B(s_{i+1}^n) \\
&\qquad + \Pr(D(s_1^{i-1}0 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_B(s_i=0)\cdot \Pr_B(s_{i+1}^n)].
\end{aligned}$$

Since $p_i - p_{i-1} \geq 1/n^{k+1}$, one of the following two equations hold:

$$\sum_{s_1^{i-1}, s_{i+1}^n} [\Pr(D(s_1^{i-1}1 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_S(s_i=1|s_1^{i-1})\cdot \Pr_B(s_{i+1}^n)$$
$$- \Pr(D(s_1^{i-1}1 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_B(s_i=1)\cdot \Pr_B(s_{i+1}^n)] \geq \frac{1}{2n^{k+1}}, \qquad \text{(A.1)}$$

$$\sum_{s_1^{i-1}, s_{i+1}^n} [\Pr(D(s_1^{i-1}0 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_S(s_i=0|s_1^{i-1})\cdot \Pr_B(s_{i+1}^n)$$
$$- \Pr(D(s_1^{i-1}0 s_{i+1}^n)=1)\cdot \Pr_S(s_1^{i-1})\cdot \Pr_B(s_i=0)\cdot \Pr_B(s_{i+1}^n)] \geq \frac{1}{2n^{k+1}}. \qquad \text{(A.2)}$$

By testing D on suitably generated inputs it is possible to decide which of the two holds and construct a WSR test A_j accordingly. Also the value of i for which the pigeonhole principle holds can be determined by sampling.

A_1 submits as input to D the string $s_1^n = s_1^{i-1}1s_{i+1}^n$, where $s_1^{i-1} \in S$ is the input to A_1 and $s_{i+1}^n \in B$ are drawn by A_1. If $D(s_1^n) = 1$, then $A_1(s_1^{i-1}) = 1$; else $A_1(s_1^{i-1}) = 0$. Similarly, A_2 submits as input to D the string $s_1^{i-1}0s_{i+1}^n$, where $s_1^{i-1} \in S$ and $s_{i+1}^n \in B$. If $D(s_1^n) = 1$, then $A_2(s_1^{i-1}) = 0$; else $A_2(s_1^{i-1}) = 1$. We now analyze separately the two terms of $ws(A_1, S, i)$ and $ws(A_2, S, i)$. To make the analysis simple we use the second alternative in the definition of the weighted success rate, which compares the probabilities of successful predictions to b and $1 - b$:

$$\Pr_S(A_1(s_1^{i-1}) = s_i | A_1(s_1^{i-1}) = 1)$$

$$= \frac{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr_S(s_i = 1 | s_1^{i-1}) \cdot \Pr(D(s_1^{i-1}1s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr(D(s_1^{i-1}1s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}$$

(assuming (A.1))

$$\geq \frac{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr_B(s_i = 1) \cdot \Pr(D(s_1^{i-1}1s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n) + 1/2n^{k+1}}{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr(D(s_1^{i-1}1s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}$$

($\Pr_B(s_i = 1) = b$)

$$= b + \frac{1/2n^{k+1}}{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr(D(s_1^{i-1}1s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}$$

(the denominator < 1)

$$\geq b + \frac{1}{2n^{k+1}}.$$

Similarly for A_2:

$$\Pr_S(A_2(s_1^{i-1}) = s_i | A_2(s_1^{i-1}) = 0)$$

$$= \frac{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr_S(s_i = 0 | s_1^{i-1}) \cdot \Pr(D(s_1^{i-1}0s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}{\sum_{s_1^{i-1}, s_{i+1}^n} \Pr(D(s_1^{i-1}0s_{i+1}^n) = 1) \cdot \Pr_S(s_1^{i-1}) \cdot \Pr_B(s_{i+1}^n)}$$

$$\geq 1 - b + \frac{1}{2n^{k+1}}.$$

To complete the proof we show that for one of $ws(A_1, S, i)$ and $ws(A_2, S, i)$ (according to whether (A.1) or (A.2) holds) the remaining term (that dos not appear above) is also significantly greater than 1. □

Prediction Lemma. *For any biased source and any nonconstant probabilistic polynomial-time algorithm* $A: \{0, 1\}^{i-1} \to \{0, 1\}$ *there exists a constant* k_1 *such that*

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq b + \frac{1}{n^{k_1}}$$

iff there exists a constant k_2 such that

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq 1 - b + \frac{1}{n^{k_2}}.$$

Proof. Assume that $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq b + \varepsilon_1$, where $\varepsilon_1 = 1/n^{k_1}$. Note that $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) = \Pr_S(s_i = 1 | A(s_1^{i-1}) = 1)$. Since the overall bias of the source is known to be b,

$$\begin{aligned}\Pr_S(s_i = 1 | A(s_1^{i-1}) = 1) \cdot \Pr_S(A(s_1^{i-1}) = 1) \\ + \Pr_S(s_i = 1 | A(s_1^{i-1}) = 0) \cdot \Pr_S(A(s_1^{i-1}) = 0) = b.\end{aligned}$$

Therefore,

$$\Pr_S(s_i = 1 | A(s_1^{i-1}) = 0) \leq \frac{b - (b + \varepsilon_1)\Pr_S(A = 1)}{\Pr_S(A = 0)}.$$

Simple manipulations give

$$\begin{aligned}\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) &= \Pr_S(s_i = 0 | A(s_1^{i-1}) = 0) \\ &= 1 - \Pr_S(s_i = 1 | A(s_1^{i-1}) = 0) \\ &\geq 1 - b + \varepsilon_1 \cdot \frac{\Pr_S(A(s_1^{i-1}) = 1)}{\Pr_S(A(s_1^{i-1}) = 0)},\end{aligned}$$

which is significantly greater than $1 - b$ for nonconstant algorithms (where $\Pr_S(A(s_1^{i-1}) = 1) \geq 1/n^{k_3}$ for some constant k_3).

Similarly, when $\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq 1 - b + \varepsilon_2$, where $\varepsilon_2 = 1/n^{k_2}$, we get, using the same manipulations, that

$$\Pr_S(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq b + \varepsilon_2 \cdot \frac{\Pr_S(A(s_1^{i-1}) = 0)}{\Pr_S(A(s_1^{i-1}) = 1)}. \qquad \square$$

Proposition 2.2. *A biased source is a perfect independent biased source iff it passes the modified WSR test.*

Proof. If a biased source fails the modified WSR test it is easy to construct a distinguisher between the source and the independent biased source ensemble in a similar way to the construction in Proposition 2.1.

If a biased source is imperfect, then by the proof of Proposition 2.1 there exists a nonconstant probabilistic polynomial-time algorithm A: $\{0, 1\}^{i-1} \to \{0, 1\}$ such that $\Pr(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 1) \geq b + 1/n^{k_1}$ and $\Pr(A(s_1^{i-1}) = s_i | A(s_1^{i-1}) = 0) \geq 1 - b + 1/n^{k_2}$ for some constants k_1 and k_2. By definition this source fails the modified WSR test. $\square$

Proposition 2.3. *A biased source is a perfect independent biased source iff it passes the behavior test.*

Proof. For notational simplicity let P_1 denote $\Pr(A(s_1^{i-1}) = 1 | s_i = 1)$ and let P_0 denote $\Pr(A(s_1^{i-1}) = 1 | s_i = 0)$. We prove that a biased source passes the behavior

test iff it passes the WSR test. This follows from the close relation between the two measures: For any biased source and any nonconstant probabilistic polynomial-time algorithm $A: \{0,1\}^{i-1} \to \{0,1\}$,

$$ws(A, S, i) = \frac{P_1}{b \cdot P_1 + (1-b) \cdot P_0} + \frac{1 - P_0}{b \cdot (1 - P_1) + (1-b) \cdot (1 - P_0)}.$$

Clearly, if S passes the behavior test, then it also passes the WSR test. If S fails the behavior test, then, for some nonnegligible ε, $|P_1 - P_0| \geq \varepsilon$. Assume without loss of generality that $P_1 > P_0$. Using the relation between the tests we then get

$$ws(A, S, i) \geq 2 + \varepsilon \cdot \left\{ \frac{1-b}{P_1 - \varepsilon \cdot (1-b)} + \frac{b}{1 - P_1 + \varepsilon \cdot (1-b)} \right\}.$$

Finally note that $\varepsilon \cdot b \leq P_1 - \varepsilon \cdot (1-b) \leq 1 - \varepsilon \cdot (1-b)$, so that the term that is added to 2 is indeed nonnegligible. □

Proposition 2.4. *A source is a perfect independent biased source iff it passes the POP test.*

Proof. Given that a source fails the POP test, it is easy to construct a distinguisher between the source and the independent biased source ensemble by examining the predictions of the test, as is done in the proof of Proposition 2.1.

To prove the other direction, assume that S is imperfect and there exists a distinguisher D between S and the independent biased ensemble B. Then by the proof of Proposition 2.1 there exists a nonconstant probabilistic polynomial-time prediction algorithm $T: \{0,1\}^{i-1} \to \{0,1\}$ for the ith bit of S such that $\Pr_S(T(s_1^{i-1}) = s_i | T(s_1^{i-1}) = 1) \geq b + 1/n^k$ for some constant k. From T we construct the following POP test $A: \{0,1\}^{i-1} \to \{0,1,*\}$: $A(s_1^{i-1}) = 1$ iff $T(s_1^{i-1}) = 1$ and $A(s_1^{i-1}) = *$ iff $T(s_1^{i-1}) = 0$. Since T is nonconstant, then $\Pr_S(T(s_1^{i-1}) = 1) = \Pr_S(A(s_1^{i-1}) \neq *) \geq 1/n^l$ for some constant l. We then get that by definition S fails the POP test A. □

References

[1] Alon, N., and Rabin, M. O., Biased Coins and Randomized Algorithms, *Advances in Computing Research*, Vol. 5, ed. S. Micali, JAI Press, Greenwich, CT, 1989, pp. 499–507.

[2] Blum, M., Independent Coin Flips From a Correlated Biased Source: A Finite State Markov Chain, *Proc. 25th FOCS*, 1984, pp. 425–433.

[3] Blum, M., and Micali, S., How To Generate Cryptographically Strong Sequences of Pseudo-Random Bits, *SIAM J. Comput.*, Vol. 13, No. 4, 1984, pp. 850–864. Previous version in *Proc. 26th FOCS*, 1985.

[4] Boppana, R. B., and Hirschfeld, R., Pseudorandom Generators and Complexity Classes, *Advances in Computing Research*, Vol. 5, ed. S. Micali, JAI Press, Greenwick, CT, 1989.

[5] Chor, B., and Goldreich, O., Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity, *Proc. 26th FOCS*, 1985, pp. 429–442.

[6] Feldman, D., Impagliazzo, R., Naor, M., Nisan, N., Rudich, S., and Shamir, A., On Dice and Coins, *Proc. ICALP*, 1989.

[7] Rabin, M. O., Probabilistic Algorithms for Testing Primality, *J. Number Theory*, Vol. 12, 1980, pp. 128–138.

[8] Santha, M., and Vazirani, U. V., Generating Quasi-Random Sequences from Semi-Random Sources, *J. Comput. System Sci.*, Vol. 33, 1986, pp. 75–87. Previous version in *Proc. 25th FOCS*, 1984.

[9] Schrift, A. W., Randomness and Hardness of Bit-Sources, Ph.D. Thesis, The Weizmann Institute of Science, Rehovot, Israel, 1990.

[10] Schrift, A. W., and Shamir, A., The Discrete Log is Very Discrete, *Proc. 22nd STOC*, 1990, pp. 405–415.

[11] Schrift, A. W., and Shamir, A, On the Universality of the Next Bit Test, *Proc. CRYPTO 90.*

[12] Vazirani, U. V., and Vazirani, V. V., Trapdoor Pseudo-Random Number Generator with Applications to Protocol Design, *Proc. 24th FOCS*, 1983, pp. 23–30.

[13] von Neumann, J., Various Techniques Used in Connection with Random Digits, *Appl. Math. Ser.*, Vol. 12, 1951, pp. 36–38. Reprinted in *von Neumann's Collected Works*, Vol. 5, Pergamon Press, New York, 1963, pp. 768–770.

[14] Yao, A. C., Theory and Applications of Trapdoor Functions, *Proc. 23rd FOCS*, 1982, pp. 80–91.