THE REVIEW OF SYMBOLIC LOGIC
Volume 1, Number 1, June 2008

RECOGNIZING STRONG RANDOM REALS

DANIEL OSHERSON
Princeton University
and
SCOTT WEINSTEIN
University of Pennsylvania

Abstract. The class of strong random reals can be defined via a natural conception of effective null set. We show that the same class is also characterized by a learning-theoretic criterion of 'recognizability'.

1. Characterizing randomness. Consider a physical process that, if suitably idealized, generates an indefinite sequence of independent random bits. One such process might be radioactive decay of a lump of uranium whose mass is kept at just the level needed to ensure that the probability is one-half that no alpha particle is emitted in the nth microsecond of the experiment. Let us think of the bits as drawn from $\{0, 1\}$ and denote the resulting sequence by x with coordinates $x_0, x_1, \ldots$. *Now wouldn't it be odd* if there were a computer program P with the following property?

1. *For any input i, P enters a nonterminating routine that writes a nonempty, finite sequence $b_1, \ldots, b_m$ with $b_m = x_i$ (m depends on i).*[1]

The program will not, in general, allow prediction of x_i inasmuch as there is no requirement that the ultimate bit b_m written by $P(i)$ be marked as final. Nonetheless, shouldn't randomness exclude any computational process from having the kind of intimate knowledge of x_i described in 1?

The tension engendered by 1 afflicts a celebrated theory of randomness developed over the last half century.[2] The theory offers diverse criteria, each well motivated, for the concept 'infinite sequence of random bits'. Remarkably, the criteria yield the same collection of sequences – a collection, moreover, of measure 1 with respect to the 'coin flip' measure on the collection of infinite binary sequences. Despite this evidence for theoretical adequacy, some of the sequences labeled 'random' can be associated with a program P satisfying 1.

One response to this state of affairs has been to modify (in a simple and satisfying way) the randomness criteria originally proposed by Martin-Löf (1966). The resulting collection

Received: September 22, 2007

[1] The sequences satisfying 1 correspond to the *limit recursive sets* of Gold (1965) and *trial and error predicates* of Putnam (1965); such limits were introduced earlier by Shoenfield (1959) to characterize the sets recursive in $0'$.

[2] For overviews, see Li & Vitányi (1997), Uspenskii *et al.* (1990), and Downey *et al.* (2006a, §1–3). An older approach to randomness, due to von Mises (1919), is discussed in Lieb *et al.* (2006).

© 2008 Association for Symbolic Logic
doi:10.1017/S1755020308080076

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

of 'strong random sequences' still has measure 1 but none of its members have the defect 1. These developments are summarized in Downey *et al.* (2006b).

The goal of the present note is to offer a converging criterion for the strong random sequences. The criterion is based on the intuition that computable devices should not be able to *recognize* a random sequence if exposed to it bit by bit. Our approach is inspired by a similar idea due to Goldreich (2001) in the theory of pseudorandom numbers. We first review Martin-Löf's criterion and its proposed modification and then discuss recognition.

2. Random reals and effective null sets. Following standard terminology, infinite sequences over $\{0, 1\}$ will be called *reals*. The initial finite sequence of length n in x is written $x[n]$. (We count from 0, so $x[3] = (x_0, x_1, x_2)$.) The set of finite binary sequences is called B. The space of all reals is denoted $\mathcal{C}$ (for Cantor) and given the tree topology with basic open (and closed) sets $[\sigma] = \{x \in \mathcal{C} : \sigma \subset x\}$, for $\sigma \in \mathsf{B}$. For $S \subseteq \mathsf{B}$, let $[S]$ denote the open subset of $\mathcal{C}$ determined by S, that is, $[S] = \bigcup_{\sigma \in S} [\sigma]$. We let μ be the the unique natural measure on $\mathcal{C}$ satisfying $\mu([\sigma]) = 1/2^{|\sigma|}$ for all $\sigma \in \mathsf{B}$, where $|\sigma|$ is the length of σ. The following fact will be used (Oxtoby, 1971, Thm. 3.17).

Lemma 2. *Suppose that A_i is a descending chain of measurable subsets of $\mathcal{C}$.*[3] *Then,*

$$\lim_{i \to \infty} \mu(A_i) = \mu\left(\bigcap_i A_i\right).$$

A *Martin-Löf test* is a computable function f such that for all $i \geq 0$, $f(i)$ is a program for enumerating a subset $W_{f(i)}$ of B with $\mu([W_{f(i)}]) \leq 1/2^i$. A real x is *Martin-Löf random* just in case for all Martin-Löf tests f, $x \notin \bigcap_i [W_{f(i)}]$. The class of Martin-Löf random reals has μ-measure one since its definition exhibits it as the complement of a countable union of μ-measure null sets.[4]

A real x is *computably approximable* just in case there is a program P satisfying 1 above.[5]

Proposition 3. *Some Martin-Löf random reals are computably approximable.*

Proposition 3 is an immediate corollary of the well-known fact that some Martin-Löf random reals are *left computably enumerable*, that is, limits of computable increasing sequences of rational numbers.[6] Indeed, this fact implies that the Martin-Löf random reals contain computably approximable reals of special kinds, even less congruent with the intuition that random sequences lack structure. The following theorem of Kučera (1985) and Gács (1986) shows that Martin-Löf random reals may carry information accessible to a Turing machine. This provides further motivation for a more constrained concept of random real.

[3] That is, for all i, $A_i \supseteq A_{i+1}$. The subsets of reals that arise in this paper are all measurable since they are unions of basic open sets, intersections of such unions, or complements thereof.

[4] All of this is developed in Martin-Löf (1966). Alternative definitions of the same class are advanced in Schnorr (1973) and Levin (1973). See the references in footnote 2 for much fuller treatment.

[5] See Ambos-Spies *et al.* (2000) for alternate formulations of computable approximation.

[6] See, for example, Downey *et al.* (2006a, §4).

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

Proposition 4. *Every real is computable from a Martin-Löf random real.*[7]

A *generalized Martin-Löf test* is a computable function f such that for all $i \geq 0$:

(a) $f(i)$ is a program for enumerating a subset $W_{f(i)}$ of B,
(b) $\lim_{n\to\infty} \mu([W_{f(n)}]) = 0$.

If f is a generalized Martin-Löf test, $\bigcap_i [W_{f(i)}]$ is called an *effective null set*. Note that the effective null sets are exactly the Π^0_2 subsets of $\mathcal{C}$ with measure 0. It is easy to see the following.

Lemma 5. *For every generalized Martin-Löf test f, there is a generalized Martin-Löf test g such that $[W_{g(i)}] \supseteq [W_{g(i+1)}]$ and $\bigcap_i [W_{g(i)}] = \bigcap_i [W_{f(i)}]$.*

A real x is *strongly random* just in case for all generalized Martin-Löf tests f, $x \notin \bigcap_i [W_{f(i)}]$. The class of strongly random reals has μ-measure 1 (for the same reason as before).[8]

Proposition 6. *No strongly random real is computably approximable.*

Proof. Let program P computably accompany real x. We will exhibit a generalized Martin-Löf test f with $x \in \bigcap_i [W_{f(i)}]$. Given $j > 0$, let $\langle j \rangle$ be the finite sequence of length j whose ith bit β is defined as follows (for all $i \leq j$).

$$\beta = \begin{cases} \text{the last output of } P \text{ run on } i \text{ for } j \text{ steps, if there is any output;} \\ 0 \text{ otherwise.} \end{cases}$$

Let f be such that $W_{f(0)} = \mathsf{B}$ and for all $n > 0$,

$$W_{f(n)} = \{\sigma \in \mathsf{B} : |\sigma| = n \wedge \exists j > n(\sigma \subset \langle j \rangle)\}.$$

It is clear that f is computable and that:

7.

(a) For all n, $[W_{f(n)}] \supseteq [W_{f(n+1)}]$.
(b) Let $n > 0$ be given, and let $b = x[n]$. Then, for all $i \geq n$, some extension of b belongs to $W_{f(i)}$.
(c) Let $n > 0$ be given, and let $b \in \mathsf{B}$ be such that $b \neq x[n]$. Then, for only finitely many $i \geq n$, some extension of b belongs to $W_{f(i)}$.

Immediately from 7, we obtain:

8. $\bigcap_i [W_{f(i)}] = \{x\}$.

Hence, $\mu(\bigcap_i [W_{f(i)}]) = 0$, which implies via 7a and 2 that $\lim_{i\to\infty} \mu([W_{f(i)}]) = 0$. With 8, the latter fact exhibits f as the desired generalized Martin-Löf test.[9] □

Proposition 6 gives reason to identify the intuitive concept of randomness with the technical definition of 'strong' randomness. Indeed, Downey *et al.* (2006b) establish the

[7] In more detail: For every real x, there is a Martin-Löf random real y and a Turing machine T such that T using y as oracle enumerates x. (The nontrivial case is that x is neither recursive nor Martin-Löf random.)

[8] Strong randomness, also referred to as *weak 2-randomness*, was introduced and studied by Kurtz (1981).

[9] An alternative proof can be obtained from results in Ambos-Spies *et al.* (2000, §4) and an argument attributed to D. A. Martin in Downey & Hirschfeldt (to appear, §8.3).

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

following result, which both implies Proposition 6 and shows that no analog to Proposition 4 can be proved for strong randomness.[10]

Proposition 9. *Every strongly random real forms a minimal pair with* $\mathbf{0}'$.

In the remainder of the present paper, we attempt to provide additional motivation for identifying the intuitive concept of randomness with strong randomness.

3. A learning theoretic criterion for randomness. Let x be a computable real, and suppose that you are progressively examining a possibly different real r bit by bit, responding either Yes or No at each stage.[11] Relying on an algorithm for computing x, it is easy to 'recognize' whether $r = x$ in the sense of responding Yes cofinitely often just in case $x = r$.[12] In contrast, if x were random, intuition suggests that such behavior would exceed your abilities; for a random real cannot be memorized, nor does it include any comprehensible pattern that distinguishes it from all other reals. The same limitation would seem to apply to any computable agent who takes your place. Can these ideas be used to define a set of reals that coincides with strong randomness?

As shown in the next section, the concept of 'recognition' just introduced is too demanding to serve our purposes. We liberalize it by requiring that $r = x$ iff the agent responds Yes infinitely often (instead of cofinitely often). The agent is also allowed to produce this behavior mistakenly on a small set of reals (measure 0). The naturalness of this criterion for characterizing randomness is taken up in the Discussion section. Here we proceed as follows. A function $\ell : \mathsf{B} \to \{\mathsf{Yes}, \mathsf{No}\}$ is called a *learner*. We say that a learner ℓ *recognizes* a real x just in case x belongs to some $S \subseteq \mathcal{C}$ of μ-measure 0 such that for all reals r, $r \in S$ iff $\{n : \ell(r[n]) = \mathsf{Yes}\}$ is infinite.[13]

Proposition 10. *A real x is strongly random iff no computable learner recognizes x.*

Proof. For the left-to-right direction, suppose that computable learner ℓ recognizes real x. It suffices to show that x is not strongly random. By the definition of recognition, let V be such that:

11.

(a) *V is the set of reals r such that $\{n : \ell(r[n]) = \mathsf{Yes}\}$ is infinite.*

(b) $\mu(V) = 0$.

(c) $x \in V$.

For $n \geq 0$, let U_n be the set of $b \in \mathsf{B}$ such that for some $c \subset b$ of length at least n, $\ell(c) = \mathsf{Yes}$. It is clear that there is computable f with $W_{f(n)} = U_n$ for all n. Because the U_n are a descending chain, so are the $[W_{f(n)}]$. By 11a, $\bigcap_n [W_{f(n)}] = V$, so by 11b, $\mu(\bigcap_n [W_{f(n)}]) = 0$. It follows from 2 that $\lim_{n\to\infty} \mu([W_{f(n)}]) = 0$ and hence that f is a generalized Martin-Löf test. By 11c, this exhibits x as not strongly random.

For the right-to-left direction, suppose that x is not strongly random. It suffices to exhibit a computable learner that recognizes x. Since x is not strongly random, let f be a generalized Martin-Löf test with

---

[10] We thank Denis Hirschfeldt for these observations.

[11] A real x is computable iff the set of coordinates set to 1 in x is a decidable set.

[12] By responding Yes 'cofinitely often' is meant responding Yes except for finitely many exceptions.

[13] It is crucial to our definition that S have μ-measure 0. For any $\alpha > 0$, allowing S to have measure α renders all reals recognizable.

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

12.

(a) $\lim_{n\to\infty} \mu([W_{f(n)}]) = 0$ *and*
(b) $x \in \bigcap_n [W_{f(n)}]$.

By 5, we may assume

13. $[W_{f(i)}] \supseteq [W_{f(i+1)}]$.

Hence, by 12a, 13, and 2:

14. $\mu(\bigcap_n [W_{f(n)}]) = 0$.

So by 12b and 14, it suffices to exhibit a computable learner ℓ such that

15. $\bigcap_n [W_{f(n)}] = \{x : \ell(x[n]) = \mathsf{Yes}$ *for infinitely many* $n\}$.

For this purpose, given $b \in \mathsf{B}$, let $\langle b\rangle$ be the greatest $j \leq |b|$ such that some initial segment of b appears within $|b|$ steps in the standard enumeration of $W_{f(j)}$. Let $\ell(\emptyset) = \mathsf{No}$, and for all $b \in \mathsf{B}$ and $\beta \in \{0, 1\}$, let

$$\ell(b\beta) = \begin{cases} \mathsf{Yes} & \text{if } \langle b\beta\rangle > \langle b\rangle \\ \mathsf{No} & \text{otherwise.} \end{cases}$$

It is clear that ℓ is computable. Let real x be given, and suppose that $\ell(x[n]) = \mathsf{Yes}$ for only finitely many n. Then, there are (cofinitely many) k with $x \notin [W_{f(k)}]$. Suppose alternatively that $\ell(x[n]) = \mathsf{Yes}$ for infinitely many n. Then, there are infinitely many k with $x \in [W_{f(k)}]$, hence by 13, $x \in \bigcap_n [W_{f(n)}]$. Thus, for every real x, $\ell(x[n]) = \mathsf{Yes}$ for infinitely many n iff $x \in \bigcap_n [W_{f(n)}]$, which verifies 15. □

It follows immediately from Propositions 6 and 10 that every computably approximable real is computably recognized.

4. Strong recognition. We develop the 'cofinite' version of recognition mainly to document its limitations in the present context. We say that a learner ℓ *strongly recognizes* a real x just in case x belongs to some $S \subseteq \mathcal{C}$ of μ-measure 0 such that for all reals r, $r \in S$ iff $\{n : \ell(r[n]) = \mathsf{Yes}\}$ is cofinite. Strong recognizability is related to another notion of 'randomness' due to Kurtz (1981). A *Kurtz test* is a computably enumerable set $S \subseteq \mathsf{B}$ such that $\mu([S]) = 1$. A real x is *Kurtz random* iff $x \in [S]$ for every Kurtz test S. It is well-known that the Kurtz random reals are a proper superset of the Martin-Löf random reals and that there are computably approximable reals that fall in the former set but not the latter (Downey *et al.*, 2006a, §10.3). Therefore, the next proposition marks a sharp contrast to Proposition 10.

Proposition 16. *A real x is Kurtz random iff no computable learner strongly recognizes x.*

Proof. For the right-to-left direction of the proposition, suppose that the real x is not Kurtz random. Then, there is a Kurtz test S such that $x \notin [S]$. For every $b \in \mathsf{B}$, let $\ell(b) = \mathsf{No}$ if some initial segment of b appears within $|b|$ steps in the standard enumeration of S; otherwise, let $\ell(b) = \mathsf{Yes}$. It is clear that ℓ is computable and that for all reals r, $\{n : \ell(r[n]) = \mathsf{Yes}\}$ is cofinite iff $r \in \mathcal{C} - [S]$; moreover, $x \in \mathcal{C} - [S]$, and $\mu(\mathcal{C} - [S]) = 0$ since S is a Kurtz test.

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

For the left-to-right direction, suppose that some computable learner ℓ strongly recognizes the real x. We will exhibit a Kurtz test that excludes x, implying that x is not random. Let $Y \subseteq \mathcal{C}$ be such that for every real r, $r \in Y$ iff $\{n : \ell(r[n]) = \mathsf{Yes}\}$ is cofinite. Then, $\mu(Y) = 0$ and $x \in Y$. For each m, let $Y_m = \{r \in \mathcal{C} : \forall n > m\ (\ell(r[n]) = \mathsf{Yes})\}$. It is clear that $Y = \bigcup_m Y_m$, from which it follows that

17.

(a) for each m, $\mu(Y_m) = 0$ and
(b) $x \in Y_k$, for some k.

Let $S_m = \{b \in \mathsf{B} : |b| > m \text{ and } \ell(b) = \mathsf{No}\}$. Then:

18. *For every m,*

(a) S_m is computably enumerable and
(b) $[S_m] = \mathcal{C} - Y_m$.

It follows immediately from 17 and 18 that S_k is a Kurtz test with $x \notin [S_k]$. □

5. Discussion. We began by conceiving the output of a uranium mass m as a random real, but now let us acknowledge that there are scant grounds for identifying any proper subset of $\mathcal{C}$ with m's potential behavior. For even the constant 1-sequence seems as possible for m to produce as any other real. (It would be surprising to discover that m keeps track of its past behavior or looks ahead.) In a sense familiar from earlier literature, m embodies a random *process* rather than delimiting a special random *product* characterized in advance.[14] What empirical significance, then, can be attached to the technical definition of 'strong random real', given that it *does* carve out a proper subset of $\mathcal{C}$ as the product of randomness? Alternative answers to this question have exploited different intuitions about randomness, e.g., disorder versus typicality (Uspenskii *et al.*, 1990). Here we attempt to interpret product randomness in epistemic terms.[15]

We suggest that random reals are *anonymous* in the sense that human minds cannot identify them. To fill out this idea, fix a real r and suppose that another real x (possibly r itself) is presented to you bit by bit. Your task is to indicate whether x is r. You might wait for a large number of bits to be shown, then say "Yes, I think this real is r." Prudent, you examine many more bits before issuing another statement of the same character. If your conviction wanes, you stop responding and sink into eternal silence (rather than giving voice to your doubts). An infinite number of affirmations therefore signals having spotted r, whereas finite responding serves to deny r's presence. Of course, this is not enough for genuine ability to identify r since you must also be selective. A natural interpretation of the latter requirement is that the probability be 0 of producing infinitely many affirmations in response to an arbitrary real that is not r. Our suggestion that random reals are anonymous comes to this: a human mind can identify a given real in the foregoing sense just in case the real is not random.

Unfortunately, putting this picture into sharp focus requires more information about human cognition than currently available. Turing machines suggest themselves as substitutes

[14] For discussion of the process–product distinction, see Earman (1986, Ch. 8) and Eagle (2005, §3.1). Earman uses the term 'performance' instead of 'product'.

[15] An epistemic perspective different from ours is offered in Eagle (2005, §6).

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

for the role of epistemic agent given the popular hypothesis that brains embody computable functions relative to some intelligible convention about inputs and outputs.[16] We follow suit for want of evidence (at present) that neural interactions implement uncomputable operations.[17]

Now it should be clear that our criterion of 'identifiability' corresponds to the definition of *recognition* presented earlier, where absent responses amount to No's. Proposition 10 thus shows that the strongly random reals (defined in terms of effective null sets) are exactly the anonymous reals in the sense developed above. Whatever its physical significance, the strongly random subset of $\mathcal{C}$ thus appears to be a natural collection, lying at the confluence of measure-theoretic and epistemic perspectives on reals.

6. Acknowledgments. We are grateful for valuable comments from Adam Elga, Denis Hirschfeldt, and Stuart Kurtz.

BIBLIOGRAPHY

Ambos-Spies, K., Weihrauch, K., & Zheng, X. (2000). Weakly computable real numbers. *Journal of Complexity*, **16**, 676–690.

Downey, R., & Hirschfeldt, D. R. (2008). *Algorithmic Randomness and Complexity*. New York: Springer-Verlag.

Downey, R., Hirschfeldt, D. R., Nies, A., & Terwijn, S. A. (2006a). Calibrating randomness. *Bulletin of Symbolic Logic*, **12**, 411–491.

Downey, R., Nies, A., Weber, R., & Yu, L. (2006b). Lowness and Π^0_2 nullsets. *The Journal of Symbolic Logic*, **71**, 1044–1052.

Eagle, A. (2005). Randomness is unpredictability. *British Journal for the Philosophy of Science*, **56**, 749–790.

Earman, J. (1986). *A Primer on Determinism*. Dordrecht, The Netherlands: D. Reidel.

Gács, P. (1986). Every sequence is reducible to a random one. *Information and Control*, **70**, 186–192.

Gold, E. M. (1965). Limiting recursion. *The Journal of Symbolic Logic*, **30**, 20–48.

Goldreich, O. (2001). *Foundations of Cryptography: vol. 1, Basic Tools*. New York: Cambridge University Press.

Kučera, A. (1985). Measure, Π^0_1 classes, and complete extensions of PA. In *Proceedings of a Conference held in Oberwolfach*, *West Germany*, April 15–21, 1984 Series: *Lecture Notes in Mathematics*, vol. 1141. Ebbinghaus, H.-D., Müller, G. H., and Sacks, G. E., editors. Springer, pp. 245–259.

Kurtz, S. A. (1981). Randomness and genericity in the degrees of unsolvability. PhD dissertation, University of Illinois.

Levin, L. A. (1973). The concept of random sequence. *Doklady Akademii Nauk SSSR*, **212**, 548–550. Cited in Uspenskii *et al.* [1990].

Li, M., & Vitányi, P. B. M. (1997). *An Introduction to Kolmogorov Complexity and Its Applications* (second edition). New York: Springer.

Lieb, E. H., Osherson, D., & Weinstein, S. (2006). Elementary proof of a theorem of Jean Ville. Available online from: http:arxiv.org/PS_cache/cs/pdf/0607/0607054.pdf.

[16] See, for example, Medin *et al.* (2005, pp. 22–23).

[17] See Moser (1973) for demonstration that even Newtonian systems can exhibit discrete behavior that is not Turing simulable. (Three-body problems can be set to code arbitrary binary sequences.)

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press

Martin-Löf, P. (1966). The definition of random sequences. *Information and Control*, **9**, 602–619.
Medin, D. L., Ross, B. H., & Markman, A. B. (2005). *Cognitive Psychology* (fourth edition). New York: John Wiley & Sons.
Moser, J. (1973). *Stable and Random Motions in Dynamical Systems (With Special Emphasis on Celestial Mechanics)*. Princeton, NJ: Princeton University Press.
Oxtoby, J. C. (1971). *Measure and Category: A Survey of the Analogies between Topological and Measure Spaces*. New York: Springer-Verlag.
Putnam, H. (1965). Trial and error predicates and the solution to a problem of Mostowski. *The Journal of Symbolic Logic*, **30**, 49–57.
Schnorr, C. P. (1973). Process complexity and effective random tests. *Journal of Computer and System Sciences*, **7**, 376–378.
Shoenfield, J. R. (1959). On degrees of unsolvability. *Annals of Mathematics*, **69**, 644–653.
Uspenskii, V. A., Semenov, A. L., & Shen, A. K. (1990). Can an individual sequence of zeros and ones be random? *Russian Mathematical Surveys*, **45**, 121–189.
von Mises, R. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, **5**, 52–99.

DEPARTMENT OF PSYCHOLOGY
PRINCETON UNIVERSITY
PRINCETON, NJ 08544, USA
E-mail: osherson@princeton.edu

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PA 19104, USA
E-mail: Weinstein@cis.upenn.edu

https://doi.org/10.1017/S1755020308080076 Published online by Cambridge University Press