Algorithmic Randomness and Determinacy

Jan Reimann, Penn State University

July 10, 2015

Question

Given an infinite binary sequence

$$X = X_0 X_1 X_2 X_3 \ldots, \qquad X_i \in \{0, 1\},$$

is X random with respect to a (continuous) probability measure?

Goals

We will define what it means for a single infinite binary sequence to be random.

We then ask:

- If a sequence is difficult to describe/define, is it random?
- If a sequence is random, how computationally powerful is it?

Fundamental work by Gödel, Turing, Church, and others has made it clear that there is a strong duality between 'difficult to define' and 'computationally powerful'.

Infinite Binary Sequences

Cantor space

- $2^{\mathbb{N}}$ with standard product topology.
- Clopen basis: cylinder sets

$$[\![\sigma]\!] := \{X \in 2^{\mathbb{N}} : \ \sigma \subset X\}.$$

  where σ is a finite binary string.
- Compatible metric: $d(X, Y) = 2^{-|X \wedge Y|}$, where $X \wedge Y = \min\{k : \ X(k) \neq Y(k)\}$.
- Given a set of strings W, we write $[\![W]\!]$ for the open set induced by W, i.e. $[\![W]\!] = \bigcup_{\sigma \in W} [\![\sigma]\!]$.

Infinite Binary Sequences

Probability Measures on $2^{\mathbb{N}}$

- Determined by values on cylinders.
- $\mu[\![\sigma]\!] = \mu[\![\sigma^\frown 0]\!] + \mu[\![\sigma^\frown 1]\!]$.
- Example: Lebesgue measure $\lambda[\![\sigma]\!] = 2^{-|\sigma|}$.

More general: premeasures

- Premeasure: function $\rho : 2^{<\mathbb{N}} \to [0, \infty)$.
- Can be extended to outer measure

$$\mu_\rho(A) = \inf\{\sum_{\sigma \in W} \rho(\sigma) : [\![W]\!] \text{ open cover of } A\}.$$

  (Set $\mu_\rho(\emptyset) = 0$.)

- Example: Hausdorff premeasures: $\rho(\sigma) = 2^{-|\sigma|s}$, $s \in [0, \infty)$.

Definability and Computability

We identify binary sequences with subsets of $\mathbb{N}$.

- The most simple objects of computability theory are the computable (recursive) sets, i.e. sets whose membership can be decided by an algorithm.
- X is recursively enumerable (r.e.) iff it has a definition of the form $\exists y P(x, y)$, where P is a recursive predicate of natural numbers.

  Example: Diophantine sets $\{a \in \mathbb{N} \colon \exists \vec{x}\, p(a, x) = 0\}$,
  $p(a, \vec{x})$ a polynomial with integer coefficients.
  (In fact, every r.e. set can be represented this way (MDPR).)

- X is arithmetically definable iff there is a definition of X expressed solely in terms of addition, multiplication, and quantification ($\exists$, $\forall$) within the natural numbers.

Definability and Computability

These notions can be relativized. In particular, one can capture what it means a sequence X computes a sequence Y, written $Y \leqslant_T X$.

Three equivalent characterizations:

- There exists an algorithm that can query an oracle (think: USB-stick) such that if X is the oracle, the algorithm correctly decides membership in Y
- Both Y and $\mathbb{N} \setminus Y$ have a definition of the form $\exists y P(x, y)$, where P is an number-theoretic predicate that uses only bounded number quantifiers and expressions of the form '$\theta(\vec{v}) \in X$'.

  Example: $n \in Y$ iff $\exists x \forall z < 1942 (n \cdot (3x + z) \in X)$.

- There exists an effectively continuous mapping f from a G_δ subset of $2^{\mathbb{N}}$ to $2^{\mathbb{N}}$ such that $f(X) = Y$.

Definability and Computability

- There is a $\leqslant_T$-greatest r.e. subset of $\mathbb{N}$ denoted by $0'$ (the Halting Problem, the Turing jump).

  Similarly, for any X, X' is the $\leqslant_T$-greatest set which is recursively enumerable relative to X.

- The arithmetically definable sets are obtained by starting with the empty set, iterating relative existential definability (i.e. the map $X \mapsto X'$), and closing under relative computability.

- Beyond the arithmetically definable sets:

  hyperarithmetic – effectively Borel, continue jump operation through computable ordinals
  analytical, co-analytical and beyond – allow set/sequence quantifiers
  set-theoretical hierarchies – Gödel's constructible universe.

Martin-Löf Randomness

Every nullset is subset of a G_δ nullset.

A test for randomness is an effectively presented G_δ nullset.

Definition

- A Martin-Löf test for μ is a set $W \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ recursively enumerable relative to (a representation of) μ such that

$$\sum_{\sigma \in W_n} \mu[\![\sigma]\!] \leqslant 2^{-n},$$

  where $W_n = \{\sigma : (n, \sigma) \in W\}$
- A sequence $X = X_0 X_1 X_2 \ldots$ is μ-random if $X \notin \bigcap_n [\![W_n]\!]$ for every μ-test W.

The concept works more generally for premeasures, such Hausdorff premeasures.

Martin-Löf Randomness

Examples

- A computable sequence is not Martin-Löf random.

  For example, π is not random. (It fails the test of "being π").

- However, there is a recursively approximated ($\leqslant_T 0'$), but not recursive, sequence X such that X is Martin-Löf random.

- All commonly used statistical laws are effective in Martin-Löf's sense, so a Martin-Löf random sequence satisfies the law of large numbers, etc.

Martin-Löf Randomness

We can make tests more powerful by giving them access to an additional oracle Z.

μ-Z-test: W recursively enumerable relative to $\mu \oplus Z$.

n-randomness: random relative to $\mu^{(n-1)}$.

Summary

The set of μ-n-random sequences

- has μ-measure 1
  (there are only countably many r.e. sets in a given oracle, hence at most countably many tests)
- is decreasing in n
  (more computational power for tests, more non-randomness detected)

Kolmogorov Complexity

Let M be a Turing-machine. Define

$$C_M(\sigma) = \min\{|p| : p \in 2^{<\mathbb{N}}, M(p) = \sigma\},$$

i.e. $C_M(\sigma)$ is the length of the shortest program (for M) that outputs σ.

Kolmogorov's invariance theorem: There exists a machine U such that C_U is optimal (up to an additive constant), i.e. for all other machines M,

$$C_U(\sigma) \leqslant C_M(\sigma) + O(1)$$

Fix such a U and set $C(\sigma) = C_U(\sigma)$, the plain Kolmogorov complexity of σ.

A prefix-free Turing machine is a machine with prefix-free domain. The prefix-free version of C (use universal prefix free TM) is denoted by K.

Randomness and Incompressibility

Schnorr-Levin Theorem

A sequence X is Martin-Löf random iff there exists a constant c such that

$$(\forall n)\ K(X\restriction_n) \geqslant n - c,$$

Proof: Short descriptions $\leftrightarrow$ open cover

Pointwise Shannon-McMillan-Breiman Theorem

If μ is a computable Bernoulli measure, then for any μ-random X

$$\lim_{n\to\infty} \frac{K(X\restriction_n)}{n} = h(\mu) = -[p \log p + (1-p)\log(1-p)].$$

[Levin,Brudno]

Randomness and Computability

Trivial Randomness

Obviously, every sequence X is trivially random with respect to μ if $\mu\{X\} > 0$, i.e. if X is an atom of μ.

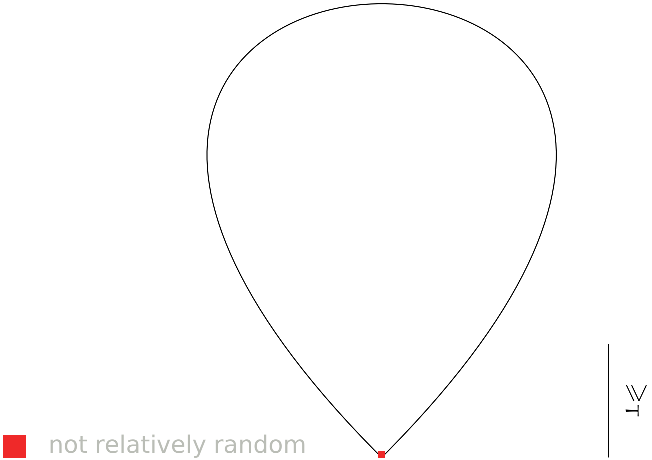
If we rule out trivial randomness, then being random means being non-computable.

Theorem [R. and Slaman]

For any sequence X, the following are equivalent.

- There exists a measure μ such that $\mu\{X\} = 0$ and X is μ-random.
- X is not computable.

$2^{\mathbb{N}}$ ordered by $\geqslant_T$



Non-trivial Randomness

Features of the proof

- Conservation of randomness.
  If Y is random for Lebesgue measure λ, and $f : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ is computable, then $f(Y)$ is random for λ_f, the image measure.

- A cone of λ-random reals.
  By the Kucera-Gacs Theorem, every sequence $\geqslant_T 0'$ is Turing equivalent to a λ-random real.

- Relativization using the Posner-Robinson Theorem.
  If X is not recursive, then $X \oplus G \geqslant_T G'$. ($X$ looks like a jump relative to G)

- A lowness argument for measures.

  Lowness is a property of low definability power. It is typically a consequence of compactness.

Randomness for Continuous Measures

In the proof we have little control over the measure that makes X random.

- In particular, atoms cannot be avoided (due to the use of Turing reducibilities).

Question

What if one admits only continuous (i.e. non-atomic) probability measures?.

Randomness for Continuous Measures

Using a technique based on Borel Determinacy, we obtain a cone of continuously n-random sequences.

For each n, there exists a sequence $X \in 2^{\mathbb{N}}$ such that for all $Y \geqslant_T X$, Y is random with respect to a continuous measure.

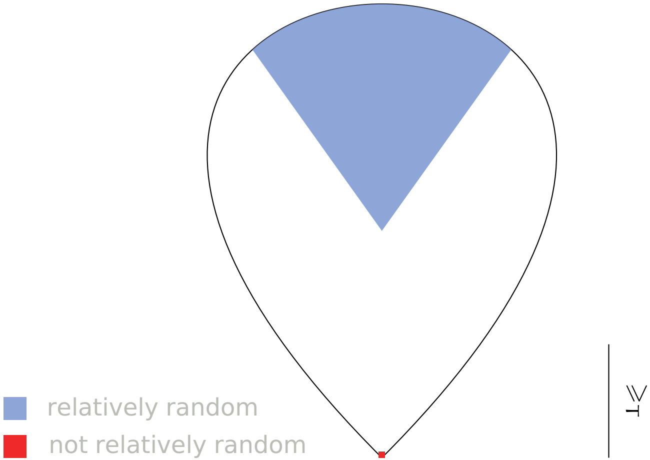
Borel Turing Determinacy:
If E is a Borel subset of $2^{\mathbb{N}}$ that is closed under $\equiv_T$, then either E or $2^{\mathbb{N}} \setminus E$ contains a $\geqslant_T$-cone.

This is a consequence of Borel Determinacy (Martin):
Two-player games with Borel winning sets are determined.

The Turing-invariant set we consider is

$$\{X \colon X \equiv_T Z \oplus R, R \ (n+1)\text{-random for } \lambda \text{ relative to } Z\}$$

$2^{\mathbb{N}}$ ordered by $\geqslant_T$



Locating the Base of the Cone

The base of the randomness cone is given by the Turing degree of a winning strategy in a game given by Martin's Theorem.

How complex is the winning strategy?

It is definable but very complex.

Gödel's hierarchy of constructible sets L:

- $\mathrm{L}_0 = \emptyset$
- $\mathrm{L}_{\alpha+1} = \mathrm{Def}(\mathrm{L}_\alpha)$, the set of subsets of L_α which are first order definable in parameters over L_α.
- $\mathrm{L}_\lambda = \cup_{\alpha<\lambda} \mathrm{L}_\alpha$, λ limit ordinal.
- $\mathrm{L} = \bigcup_\alpha \mathrm{L}_\alpha$.

Locating the Base of the Cone

The winning strategy of a Borel game can be located in L.

- The more complicated the game is in the Borel hierarchy, the more iterates of the power set of the natural numbers are used in producing the winning strategy – trees, trees of trees, etc.
- The winning strategy (for Borel complexity n) is contained in $L_{\beta(n)}$, where β_n is the least ordinal such that

  $$L_{\beta(n)} \models \mathsf{ZF}_n^-,$$

  where ZF_n^- is Zermelo-Fraenkel set theory without the Power Set Axiom + "exist n many iterates of the power set of $\mathbb{N}$".
- Note that $L_{\beta(n)}$ is countable (Condensation Lemma).

The Co-Countability Theorem

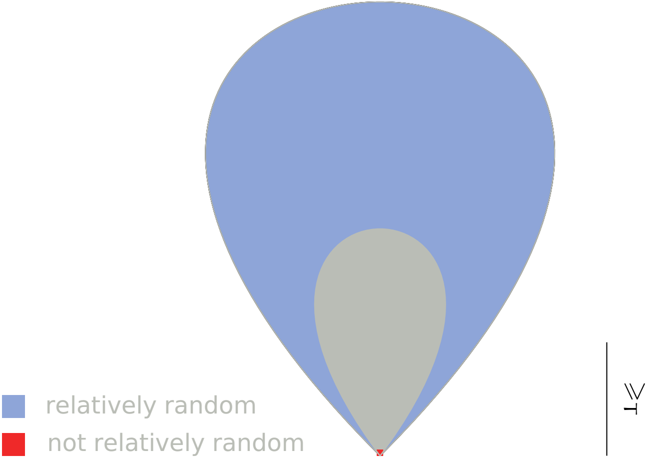
Relativizing the argument to work in other set-theoretic models (forcing extensions), we can extend the realm of ‘guaranteed’ randomness from a cone to co-countability many.

Co-Countability Theorem, R. and Slaman

For any n, all but countably many sequences are n-random with respect to a continuous measure.

> The realm of ‘guaranteed’ n-randomness is the set of sequences not in $L_{\beta(n)}$, i.e. sequences so complex that they cannot be defined in a model of a rather large fragment of set theory.

$2^{\mathbb{N}}$ ordered by $\geqslant_T$



Metamathematics necessary?

Question

Do we really need to deal with the existence of iterates of the power set of $\mathbb{N}$ (i.e. the set of real numbers, the set of all sets of real numners, the set of sets of sets of real numners, . . .) to prove that certain sets of real numbers (infinite sequences) are countable?

We make fundamental use of Borel determinacy; this suggests to analyze the metamathematics in this context.

- H. Friedman (with improvements from T. Martin) showed that infinitely many iterates of the power set of $\mathbb{R}$ are necessary to prove Borel Determinacy.

We can prove a similar fact concerning the Co-Countability Theorem: For any fixed k, ZF^{-}_{k} cannot prove the Co-countability Theorem.

Necessity of power sets

How do you prove such a thing?

- To show that the axioms of group theory do not prove that the group operation commutes, exhibit a nonabelian group.
- To show that the axioms of set theory with $\mathfrak{n}$-many iterates of the power set of $\mathbb{R}$ do not prove the Co-countability Theorem, exhibit a structure satisfying these axioms in which the Co-countability Theorem fails.

Iterates of the Power Set

A cofinal sequence of non-randoms

- Show that there is an n such that the set of non-n-randoms is cofinal in the Turing degrees of $L_{\beta(0)}$. (The approach does not change essentially for higher k.)
- The non-random witnesses will be canonical countings of the initial segments of $L_{\beta(0)}$, the so-called master codes.

Higher random reals and definability

The following is a key lemma.

Higher randomness has little definability power

Suppose that $n \geqslant 2$, $Y \in 2^{\mathbb{N}}$, and X is n-random for μ. Then, for $i < n$,

$$Y \leqslant_T X \oplus \mu \text{ and } Y \leqslant_T \mu^{(i)} \quad \text{implies} \quad Y \leqslant_T \mu.$$

Master codes, on the other hand, have extremely high definability power, hence cannot be n-random for n sufficiently large.

The "Stairmaster" Method

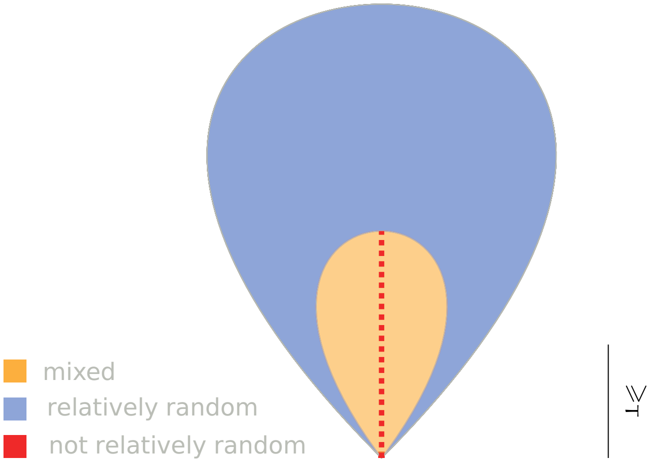
Proposition

For all k, $0^{(k)}$ is not 3-random for any μ.

Proof.

- Suppose $0^{(k)}$ is 3-random relative to μ.
- $0'$ is recursively enumerable relative to μ and recursive in the supposedly 3-random $0^{(k)}$. Hence, $0'$ is recursive in μ and so $0''$ is enumerable relative to μ.
- Use induction to conclude $0^{(k)}$ is recursive in μ, a contradiction.

$2^{\mathbb{N}}$ **ordered by** $\geqslant_T$



A Different Application

Basic principle of the previous result

random sequences $+$ Turing reductions $=$ existence of measures

Application: Frostman's Lemma

Sets of positive Hausdorff dimension support a "nice" probability measure.

Hausdorff Dimension

Hausdorff measures and dimension

Given a real $s \geqslant 0$, let $\mathcal{H}^s$ denote the premeasure given by

$$\mathcal{H}^s[\![\sigma]\!] = 2^{-|\sigma|s}.$$

(Note that we are only interested in nullsets.)

The Hausdorff dimension of a set $E \subseteq 2^{\mathbb{N}}$ is given by

$$\dim_H E = \inf\{s : E \text{ is } \mathcal{H}^s\text{-null}\}.$$

Since Martin-Löf's approach works for arbitrary premeasures, we can define the effective Hausdorff dimension $\dim_H^1$ of a sequence as

$$\dim_H^1 X = \inf\{s \in \mathbb{Q}^+ : X \text{ is not } \mathcal{H}^s\text{-random}\}$$

Effective Dimension

Dimension and Kolmogorov complexity

$$\dim_H^1 X = \liminf_n \frac{K(X\upharpoonright_n)}{n}$$

(Billingsley, Ryabko, Mayordomo)

This way effective Hausdorff dimension can be interpreted as a pointwise dimension, taken with respect to Levin's optimal enumerable semimeasure.

Example: If X is λ-random, then

$$\dim_H^1 (X_0\,0\,X_1\,0\,X_2\,0\dots) = 1/2.$$

Pointwise Frostman Lemma

Theorem
If for $X \in 2^{\mathbb{N}}$ $\dim^1_H X > s$, then X is random with respect to a probability measure μ such that

$$(\forall \sigma)\ \mu[\![\sigma]\!] \leqslant c2^{-|\sigma|s}. \qquad (*)$$

In particular, sequences of positive dimension are random with respect to a continuous measure.

This implies the classical Frostman Lemma:

If $\dim_H E > s$, $E \subseteq 2^{\mathbb{N}}$ Borel, then there exists a probability measure μ satisfying (*) such that

$$\text{supp}(\mu) \subseteq E.$$

Pointwise Frostman Lemma

However, the proof is of an effective nature.

- By the Kucera-Gacs Theorem, there exists a λ-random real R such that $R \geqslant_{\text{wtt}} X$ via some reduction Φ.

- The effective process transforming R into X induces a "defective" probability measure on $2^{\mathbb{N}}$, a semimeasure.

- Using a recursion theoretic lowness argument,

  Every effectively closed set contains an element that has low definability power ("almost recursive").

  one can show that among the possible completions of this semimeasure into a probability measure, there must exist one that makes X random and satisfies (*).

Ende