The Metamathematics of Randomness

Jan Reimann

January 26, 2007

(Original) Motivation

Effective extraction of randomness

In my PhD-thesis I studied the computational power of reals effectively random for Hausdorff measures.

The Dimension Problem

Can we efficiently extract uniform randomness from such reals?

Examples:

- Von Neumann's trick
- randomness extractors in computational complexity
- factors of dynamical systems (Sinai, Ornstein)

This lead eventually to another question:

> Which reals are random with respect to
> a (continuous) probability measure?

The answers to this question took an unexpected turn.

Measures on Cantor Space

Outer measures from premeasures

Approximate sets from outside by open sets and weigh with a general measure function.

- A premeasure is a function $\rho : 2^{<\omega} \to \mathbb{R}_0^+ \cup \{\infty\}$.
- One can obtain an outer measure μ_ρ from ρ by letting

$$\mu_\rho(X) = \inf_{C \subseteq 2^{<\omega}} \left\{ \sum_{\sigma \in C} \rho(\sigma) : \bigcup_{\sigma \in C} N_\sigma \supseteq X \right\},$$

  where N_σ is the basic open set induced by σ.
  (Set $\mu_\rho(\emptyset) = 0$.)

The resulting $\mu = \mu_\rho$ is a countably subadditive, monotone set function, an outer measure.

Measures on Cantor Space

Types of measures

Probability measures: based on a premeasure ρ which satisfies

- $\rho(\emptyset) = 1$ and
- $\rho(\sigma) = \rho(\sigma^\frown 0) + \rho(\sigma^\frown 1)$.

For probability measures it holds that $\mu_\rho(N_\sigma) = \rho(\sigma)$.

Hausdorff measures: based on a premeasure ρ which satisfies

- If $|\sigma| = |\tau|$, then $\rho(\sigma) = \rho(\tau)$.
- $\rho(n)$ is nonincreasing.
- $\rho(n) \to 0$ as $n \to \infty$.
- For example: $\rho(\sigma) = 2^{-|\sigma|s}$, $s \geqslant 0$.

The actual definition of Hausdorff measures is more complicated, but we are only interested in nullsets.

Measures on Cantor Space

Nullsets

The way we constructed outer measures, $\mu(A) = 0$ is equivalent to the existence of a sequence $(W_n)_{n\in\omega}$, $W_n \subseteq 2^{<\omega}$, such that for all n,

$$A \subseteq \bigcup_{\sigma\in W_n} N_\sigma \quad \text{and} \quad \sum_{\sigma\in W_n} \rho(\sigma) \leqslant 2^{-n}.$$

Thus,

every nullset is contained in a G_δ nullset.

By requiring that the covering nullset is effectively G_δ, we obtain a notion of effective nullsets.

Definition

- A test relative to $z \in 2^\omega$ is a set $W \subseteq \mathbb{N} \times 2^{<\omega}$ which is c.e. in z.
- A real x passes a test W if $x \notin \bigcap_n N(W_n)$, where $W_n = \{\sigma : (n, \sigma) \in W\}$.

Hence a real passes a test W if it is not in the G_δ-set represented by W.

To test for randomness, we want to ensure that W actually describes a nullset.

Definition

Suppose μ is a measure on 2^ω. A test W is correct for μ if for all n,

$$\sum_{\sigma \in W_n} \mu(N_\sigma) \leqslant 2^{-n}.$$

Any test which is correct for μ will be called a test for μ.

An effective test for randomness should have access to the measure it is testing for.

- Therefore, represent it by an infinite binary sequence.
- Outer measures are determined by the underlying premeasure ρ. It seems reasonable to represent these values via approximation by rational intervals.

Definition

Given a premeasure ρ, define its rational representation r_ρ by letting, for all $\sigma \in 2^{<\omega}$, $q_1, q_2 \in \mathbb{Q}$,

$$\langle \sigma, q_1, q_2 \rangle \in r_\rho \iff q_1 < \rho(\sigma) < q_2.$$

The condition $q_1 < \rho(\sigma) < q_2$ induces a subbasis for the weak topology on the space of probability measures.

- More general, if a space X is Polish, so is the space $\mathcal{P}(X)$ of all probability measures on X (under the weak topology). Also, if X is compact metrizable, so is $\mathcal{P}(X)$.
- This yields various ways to represent a measure: Cauchy sequences, list of basic open balls it is contained in, etc.
- We can obtain a nice effective representation (e.g. by following the framework in Moschovakis' book).

Theorem

There is a recursive surjection $\pi : 2^\omega \to \mathcal{P}(2^\omega)$ and a Π^0_1 subset P of 2^ω such that $\pi\restriction_P$ is one-to-one and $\pi(P) = \mathcal{P}(2^\omega)$.

Definition

Suppose ρ is a premeasure on 2^ω and $z \in 2^\omega$. A real is μ_ρ-z-random if it passes all $r_\rho \oplus z$-tests which are correct for μ_ρ.

Hence, a real x is random with respect to an arbitrary measure μ_ρ if and only if it passes all tests which are enumerable in the representation r_ρ of the underlying premeasure ρ.

- n-randomness: tests r.e. in $\emptyset^{(n-1)}$.
- Accordingly, arithmetical randomness.

Let μ be a probability measure and $f : 2^\omega \to 2^\omega$ be a continuous (Borel) function.

Define the image measure μ_f by setting

$$\mu_f(\sigma) = \mu(f^{-1}N_\sigma)$$

Conservation of randomness

If the transformation f is computable in z, then it should preserve randomness, i.e. it should map a μ-z-random real to a μ_f-z-random one.

Making Reals Random

Computable measures

Trivially, every atom of a measure is random with respect to it.

- For recursive reals, this is the only way to become random.

If μ is a computable measure, then an atom of μ is μ-random iff it is computable.

Theorem (Levin, Kautz)

If a real is noncomputable and random with respect to a computable probability measure, then it is Turing equivalent to a λ-random real.

- The proof works by showing that the distribution function can be computed effectively.

Let $h : \mathbb{N} \to \mathbb{R}^{\geqslant 0}$ be a computable, nondecreasing, unbounded function (effective dimension function).

- Interesting connection with Kolmogorov complexity: A real x is Hausdorff 2^{-h}-random if and only if for some c,

$$K(x\restriction_n) \geqslant h(n) - c \quad \text{for all } n.$$

The dimension problem

If x is Hausdorff 2^{-h}-random, does it compute a λ-random real?

Non-trivial Randomness

The basic question

Unfortunately, not every Hausdorff 2^{-h}-random real is random for some computable probability measure (for arbitrary h).

- Join a λ-random and a 1-generic with appropriate density.

Question

Are such reals at least non-trivially random with respect to some measure? What reals are in general?

Non-trivial Randomness

Too coarse

It turns out every non-recursive real is random.

Theorem (Reimann and Slaman)

For any real x, the following are equivalent.

(i) *There exists (a representation of) a measure μ such that $\mu(\{x\}) = 0$ and x is 1-random for μ.*

(ii) *x is not computable.*

Features of the proof:

- Conservation of randomness.
- Randomness of cones:
  - Kucera's coding argument shows that every degree above $\emptyset'$ contains a λ-random.
  - Relativize this using the Posner-Robinson Theorem.
  - Conclude that every non-recursive real x is Turing equivalent to some λ-z-random real for some real z.
- A basis theorem for relative randomness.

The Turing equivalence to a λ-random real translates into effectively closed consistency conditions for a probability measure.

- The following basis theorem (Downey, Hirschfeldt, Miller, and Nies; Reimann and Slaman) ensures that one of the measures will not affect the randomness of z.

Theorem

If $B \subseteq 2^\omega$ is nonempty and Π^0_1, then, for every y which is λ-random there is $z \in B$ such that y is λ-random relative to z.

- This argument seems to be applicable in more generality, proving existence of measures.

Randomness for Continuous Measures

In the proof there is no control over the measure that makes x random.

- Atoms cannot be avoided.
- Uses a special (though natural) representation of $M(2^\omega)$ as a particular Π^0_1 class.

Question

What if one admits only continuous probability measures?.

Theorem (Reimann and Slaman)

Let x be a real. For any $z \in 2^\omega$, the following are equivalent.

(i) *x is truth-table equivalent to a λ-z-random real.*

(ii) *x is random for a continuous (dyadic) measure recursive in z.*

(iii) *There exists a functional Φ recursive in z which is an order-preserving homeomorphism of 2^ω such that $\Phi(x)$ is λ-z-random.*

This is an effective version of the classical isomorphism theorem for continuous probability measures.

The Class NCR

Question

Which level of logical complexity guarantees continuous randomness?

Let NCR_n be the set of all reals which are not n-random relative to any continuous measure.

- Kjos-Hanssen and Montalban: Every member of a countable Π^0_1 class is contained in NCR_1. (It follows that elements of NCR_1 is cofinal in the hyperarithmetical Turing degrees.)
- Woodin: outside Δ^1_1 the Posner-Robinson theorem holds with tt-equivalence.
- Conclude that $\mathsf{NCR}_1 \subseteq \Delta^1_1$.

Theorem

Kleene's $\mathcal{O}$ is an element of NCR_3.

Based on this, one can use the theory of jump operators (Jockusch and Shore) to obtain a whole class of examples.

Proof:

- Tree representation
  $\mathcal{O} = \{e : \text{the } e\text{th recursive tree } T_e \subseteq \omega^{<\omega} \text{ is well-founded}\}$.
- Suppose $\mathcal{O}$ is 3-random for some μ.
- We want to use domination properties of random reals.

- Well-known (Kurtz and others): If X is n-random for μ, $n > 1$, then every function $f \leqslant_T X$ is dominated by a function recursive in μ'.
- Therefore, μ' computes a uniform family $\{g_e\}$ of functions dominating the leftmost infinite path of T_e.
- Infer: For every e, the following are equivalent.
  - (i) T_e is well-founded.
  - (ii) The subtree of T_e to the left of g_e is finite.
- The latter condition is $\Pi^0_1(\mu')$, hence $\mathcal{O}$ is $\Pi^0_2(\mu)$.
- But this is impossible if $\mathcal{O}$ is 3-random for μ.

The domination property of higher randomness implies that random reals are not helpful when adding them as oracles/parameters.

Lemma

Suppose that $n \geqslant 2$, $y \in 2^\omega$, and R is λ-n-random relative to μ. If $i < n$, y is recursive in $(R \oplus \mu)$ and recursive in $\mu^{(i)}$, then y is recursive in μ.

Corollary: For all k, $\emptyset^{(k)}$ is not n-random relative to any μ, $n \geqslant 2$.

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

Upper Bounds for Continuous Randomness

In general, can we give a distinct bound on NCR_n like in the case $n = 1$?

- There is some evidence that NCR_n grows very quickly with n.
- Can we give an upper bound?

Theorem (Reimann and Slaman)

For all n, NCR_n *is countable.*

NCR_n is Countable

Main Features of the Proof

Show that the complement of NCR_n contains an upper Turing cone.

- Show that the complement of NCR_n contains a Turing invariant and cofinal (in the Turing degrees) Borel set.
- We can use the set of all y that are Turing equivalent to some $z \oplus R$, where R is $(n+1)$-random relative to a given z.
- These y will be n-random relative to some continuous measure and are T-above z.
- Use Martin's result on Borel Turing sets to infer that the complement of NCR_n contains a cone.
- The cone is given by the Turing degree of a winning strategy in the corresponding game.

NCR_n is Countable

Main Features of the Proof

Go on to show that the elements of NCR_n show up at a rather low level of the constructible universe.

- $NCR_n \subseteq L_{\beta_n}$, where β_n is the least ordinal such that

  $$L_{\beta_n} \models ZFC^- + \text{there exist } n \text{ many iterates of the power set of } \omega,$$

  where ZFC^- is Zermelo-Fraenkel set theory without the Power Set Axiom.

NCR_n is Countable

Main Features of the Proof

Given $x \notin L_{\beta_n}$, construct a set G such that

(i) $L_{\beta_n}[G]$ is a model of ZFC_n^-.

(ii) For all $y \in L_{\beta_n}[G] \cap 2^\omega$, $y \leqslant_T x \oplus G$.

G is constructed by Kumabe-Slaman forcing.

The existence of G allows to conclude:

- If x is not in L_{β_n}, it will belong to every cone with base in $L_{\beta_n}[G]$.
- In particular, it will belong to the cone given by Martin's argument (relativized to G – use absoluteness), i.e. the cone avoiding NCR_n.
- Hence x is random relative to G for some continuous μ, hence in particular μ-random.

NCR_n is Countable

Is the metamathematics necessary?

Question

Do we need to use metamathematical methods to prove the countability of NCR_n?

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

The necessity of iterates of the power set is known from a result by Friedman.

- Martin's proof of Borel determinacy starts with a description of a Borel game and produces a winning strategy for one of the players.
- The more complicated the game is in the Borel hierarchy, the more iterates of the power set of the continuum are used in producing the strategy.

Theorem (Friedman)

$\mathrm{ZFC}^- \nvdash$ *All $\mathbf{\Sigma}^0_5$-games on countable trees are determined.*

Martin improved this to $\mathbf{\Sigma}^0_4$.

Inductively one can infer from Friedman's result that in order to prove full Borel determinacy, a result about sets of reals, one needs infinitely many iterates of the power set of the continuum.

- The proof works by showing that there is a model of ZFC^- for which $\mathbf{\Sigma}^0_4$-determinacy does not hold.
- This model is L_{β_0}.

NCR and Iterates of the Power Set

We can work along similar lines to obtain a result concerning the countability of NCR_n.

Theorem

For every k, the statement

For every n, NCR_n is countable.

cannot be proven in

$\mathsf{ZFC}^- +$ *there exists k many iterates of the power set of ω.*

NCR and Iterates of the Power Set

Features of the proof

The proof (for $k = 0$) shows that there is an n such that NCR_n is cofinal in the Turing degrees of L_{β_0}. Hence, NCR_n is not countable in L_{β_0}.

- The witnesses for NCR_n are Jensen's master codes of models L_α for limit ordinals $\alpha < \beta_0$.

We choose n large enough to capture recognition and comparison (of well-foundedness) of models they code.

NCR and Iterates of the Power Set

Features of the proof

Suppose some M_λ, $\lambda < \beta_0$, were n-random relative to μ.

- Let $\mathfrak{M}$ be the sequence of possible master codes which are recursive in μ (satisfying some arithmetical formula).
  - The well-founded part of $\mathfrak{M}$ is of the form $\mathfrak{M}_{<\gamma} = (M_\alpha : \alpha < \gamma)$ for some $\gamma \leqslant \lambda$.
  - $\mathfrak{M}_{<\gamma}$ is uniformly arithmetically definable from M_λ and hence from μ.
- M_γ is obtained by iterating uniformly arithmetically definable operations on $\mathfrak{M}_{<\gamma}$.
- The results at each step and M_γ itself are recursive in M_λ.
- The results at each step and M_γ itself are recursive in μ, by the non-helpfulness lemma.
- M_γ is in the well-founded part of $\mathfrak{M}$. Contradiction.