The Kolmogorov Complexity of Random Reals

Rod Downey

Victoria University

Wellington

New Zealand

- Joint work with Yu Liang (Nanjing and Victoria University)
  Ding Decheng (Nanjing)
  and some with Denis Hirschfeldt (Chicago)
  Andre Nies (Auckland)
  Evan Griffiths (Victoria)

The Basic Refs are van Lambalgen's Thesis, Solovay's unpublished notes, and Li-Vitanyi. The new material can be found in:

- The Kolmogorov Complexity of the Random Reals (Yu, Ding, Downey)
- The Initial Segment Complexity of Random Reals (Yu Liang and Ding Decheng)
- Density, randomness and computability, with D. Hirschfeldt and A. Nies, SIAM J. Comput. (Extended abstract appeared in *Symposium for Theoretical Aspects of Computer Science, STACS'01* January, 2001.)

- Some Computability-Theoretical Aspects of Reals and Randomness, to appear, in a *Lecture Notes in Logic* volume edited by Cholak. et. al. (to appear, 2010)
- *Algorithmic Randomness and Complexity* (Springer-Verlag) (with Hirschfeldt.
- On Ω^A with Griffiths, Hirschfeldt, Nies.

Motivation

- What is “random”?
- How can we calibrate levels randomness?
- How does this relate to classical computability notions, which calibrate levels of complexity?
- Von Mises, Church, Solomonoff, Chaitin, Kolmogorov, Shannon, etc.

Notation

- Real is a member of Cantor space 2^ω with topology with basic clopen sets $[\sigma] = \{\sigma\alpha : \alpha \in 2^\omega\}$ whose measure is $2^{-|\sigma|}$.
- for uniformity, a real is always nonrational.

Self-delimiting universal computers

- Computers have alphabet $\{0,1\}$.
- A computer M is *self-delimiting* if

$$(M(\sigma)\downarrow \ \wedge \ \sigma' \supsetneq \sigma) \Rightarrow M(\sigma')\uparrow .$$

- A self-delimiting M is *universal* if for every self-delimiting N there is a c s.t.

$$N(\sigma)\downarrow \rightarrow \exists\tau(|\tau| \leq |\sigma| + c \ \wedge$$

$$M(\tau)\downarrow = N(\sigma)).$$

- Fix a universal machine N, and a self-delimiting universal computer M.

Kolmogorov-Solomonoff-Levin-Chaitin randomness:

- The *Kolmogorov complexity* $C(\sigma)$ of $\sigma \in \{0,1\}^*$ is $|\tau|$ for the shortest τ s.t. $N(\tau)\downarrow= \sigma$.
- The *prefix-free complexity* $K(\sigma)$ of $\sigma \in \{0,1\}^*$ is $|\tau|$ for the shortest τ s.t. $M(\tau)\downarrow= \sigma$.
- α is *random* if there is a c s.t.

$$\forall n(K(\alpha \upharpoonright n) > n - c).$$

  This happens if there is a c such that for infinitely many n, $C(\alpha \upharpoonright n) > n - c$.

Martin-Löf Randomness

- This is equivalent to Martin-Löf randomness:
- A Martin-Löf test is a computable collection of (c.e.) open sets V_n such that $\mu(V_n) \leq 2^{-n}$.
- A real α is Martin-Löf random iff it *passes* all Martin-Löf tests, meaning $\alpha \notin \cap_n V_n$. (Schnorr)

Measures of Relative Randomness

- Initial segment complexity allows us to measure the relative randomness of reals.
- $\alpha \leq_Q \beta$ iff there is a c such that for all n,

  $$Q(\beta \upharpoonright n) + c \geq Q(\alpha \upharpoonright n).$$

  Here $Q \in \{C, K\}$.
- There are other (nonexplored) possibilities. e.g $\alpha \leq_{sm} \beta$ iff there is a c such that $cF(\alpha \upharpoonright n) \geq F(\beta \upharpoonright n)$ where F is the universal minimal supermartingale. (cf Lutz etc)

Computably enumerable reals:

- α is *c.e.* if there is a computable sequence of rationals

$$q_0 < q_1 < \ldots \longrightarrow \alpha.$$

  Equivalently, the lower cut of α is a c.e. set of rationals.

- α is *computable* if there is a computable f s.t.

$$\forall n(\alpha - q_{f(n)} < 2^{-n}).$$

- α is *strongly c.e.* if $\alpha = 0.\chi_A$ for a c.e. A.

- Soare was the first to observe that there are c.e. reals that are not strongly c.e..

- *ce reals* are to this area what *ce sets* are in classical computability as they are the measures of domians of prefix free machines.

Chaitin's Ω-numbers

- $$\Omega_M = \sum_{M(\sigma)\downarrow} 2^{-|\sigma|}$$
- Ω_M is the halting probability of M.
- Ω_M is c.e. and random.
- Such numbers are called Ω-*numbers*.
- One talks about *the* Chaitin number Ω, but it is apparently Machine dependent.

Domination, Solovay reducibility:

- $\beta \leq_S \alpha$ if there are a partial computable $\varphi : \mathbb{Q} \to \mathbb{Q}$ and a c s.t.

  $$\forall q < \alpha(\varphi(q)\downarrow < \beta \ \wedge \ \beta - \varphi(q) \leq c(\alpha - q)).$$

- The idea is that however I approximate β I can also approximate α. Can be thought of as the relevant version of m-reducibility for randomness.

- Solovay's Lemma: *Let c be given. Then there is a constant d depending on c alone, such that, for $Q = C$ or K, for all n, if $|\sigma| = |\tau| = n$ and*

$|\sigma - \tau| \leq 2^{c-n}$, *then*

$$Q(\sigma) = Q(\tau) + d.$$

- Solovay Reducibility provides a measure of relative randomness for c.e. reals:

$$\alpha \leq_S \beta \text{ implies}$$

$$\exists c \forall n (Q(\alpha \upharpoonright n) \leq Q(\beta \upharpoonright n) + c.$$

- The proof is to apply Solovay's Lemma on $\alpha \upharpoonright n$ and $\varphi(\beta \upharpoonright n) \upharpoonright n$.

Ω-like reals:

- A c.e. real is Ω-*like* if it dominates all c.e. reals.
- By the Solovay's theorem, any Ω-like real is random.
- **Theorem (Calude, Hertling, Khoussainov, and Wang).** If a c.e. real is Ω-like then it is an Ω-number. (Proof is Kraft-Chaitin)
- **Theorem** [Kucera-Slaman, SICOMP, 2002] If a c.e. real is random then it is Ω-like.

- Chaitin observed that $K(\sigma) \leq |\sigma| + K(|\sigma|) + c$.
- To "qualify" as being random, all α has to have is $K(\alpha \upharpoonright n) \geq n - c$.
- One would expect that there ought to be lots of differing complexity random reals.
- The conclusion is that all random c.e. reals are essentially the same. Their initial segment complexity oscillates at exactly the same rate.

How many K-degrees are there?

- Here we are not caring about being c.e. or even below $\emptyset'$.
- The question is definitely nontrivial in view of the following.
- (Yu, Ding Downey) Suppose that α is random. There are $2^{\aleph_0}$ many reals which are K-reducible to α.

- Proof: Define $\mathcal{A} = \mathcal{P}(\{2^n : n \in \mathbb{N}\})$.

- For every set A, define $B(A) = \{n \mid 2^n \in A\}$.

- Then for every $A \in \mathcal{A}$ and n, $K(A \upharpoonright n) \leq K(\log n) + K(B(A) \upharpoonright \log n) + c$
$\leq 2 \log n + 4 \log \log n + c'$
$\leq n + c''$

- Thus $A \leq_K \alpha$ for every $A \in \mathcal{A}$.

Solovay's Theorems:

- Solovay, and van Lambalgen, amongst others noted that Δ^0_2 sets differed from general sets.
- Call real α is *Kolmogorov random* iff $\exists^\infty n(C(\alpha \upharpoonright n) \geq n - \mathcal{O}(1))$.
- (YDD-probably implicit in van Lambalgen) No Δ^0_2 real (and in particular Ω) is Kolmogorov random
- (Proof)
- Suppose that $\alpha \leq_T \emptyset'$. $\alpha \upharpoonright n = \lim_s f(n,s)$.
- Let g be any computable function, such as $g(n) = 2^{2^n}$, say. Let s_n be

sufficiently large that
$f(g(n), s) = f(g(n), s_n)$ for all $s \geq s_n$.

- Then for $k \geq n + s_n$, the following is a short K-program for $\alpha \restriction k$:

- The input is n, γ where γ is the part of α of lengths between $g(n)$ and k, which we write as $\gamma = \alpha[g(n), k]$. This input has length $2 \log n$ plus $k - g(n)$. Then on this input, we first scan the length (say t) then calculate $f(g(n), t)$ and output $f(g(n), t)\gamma$, which will equal $\alpha \restriction k$. The length of this program is bounded away from $n - c$ for any c.

n-randomness

- (Kurtz) Can a real n-random iff it avoids all Σ^0_n Martin-Löf tests.
- Suppose that α is 3-random. Then α is Kolmogorov random. Indeed, $\exists^\infty n(K(\alpha \upharpoonright n) = n + K(n) + \mathcal{O}(1))$. This implies that α is Kolmogorov random via Solovay.

Nies' Theorem

- Nies used the YDD method to prove :
- Suppose that α is Kolmogorov random. Then α is already 2-random.
- (Proof) There is an n with $K'(\alpha \upharpoonright n) < n - k$. Take σ with $U^{\emptyset'}(\sigma) = \alpha \upharpoonright n$. Now consider $\sigma\alpha \upharpoonright_{n+1}^{s}$ for sufficiently long s.
- Nies also observed that 2-random = Kolmogorov random with a computable time bound. (Miller later)

More degrees

- A consequence of the above is that there are least two H-degrees. This is a result first proven by Solovay.
- Solovay's method:
- $\alpha(n) = \min\{K(j) : j \geq n\}$.
- 2-random is 1-random relative to any Turing complete set. Use Ω, and then we can define Ω', α' etc. Note that $K'(\Omega' \upharpoonright n) \geq n - c$, but this tells us little about K.
- Solovay 1. If α is 2-random then $K(\alpha \upharpoonright n) \geq n + \alpha(n) + O(\log \alpha(n)$.

- Solovay 2. For infinitely many n, $K(\Omega \upharpoonright n) \leq n + \alpha'(n) + O(\log \alpha'(n)$.
- Note that $\alpha'(n)$ grows slower than any function computable from $\emptyset'$, such as $\alpha(\alpha(n))$.
- We don't know how to ramp this up. But (later) we can show that n-randoms always have differing H degrees.
- is $K(\Omega' \upharpoonright n) \geq K(\Omega \upharpoonright n)$?

There are uncountably many K-degrees

- (YDD) gives a method of showing that $\mu(\{x : x \leq_K y\} = 0$, giving the result.
- Lemma (Solovay) If $\sum_{m \in \mathbb{N}} 2^{-f(m)} = \infty$ for any computable function f, then $K(\Omega \upharpoonright m \geq m + f(m) - \mathcal{O}(1)$ infinitely often.
- Prototype $f(m) = \log m$
- Two improvements on this lemma:

- Joe Miller's Theorem :

  A real α is 1-random iff $\sum 2^{n-K(\alpha\upharpoonright n)} < \infty$.

- Corollary

  Suppose that f is *any* function with $\sum_{m\in\mathbb{N}} 2^{-f(m)} = \infty$. Then if α is random, $\exists^i nftym(K(\alpha \upharpoonright m) > m + f(m) + \mathcal{O}(1)$.

- The other generalization (YDD) Given any computable f, as above, there are low c.e. reals α such that $\exists^\infty n(K(\alpha \upharpoonright n) > n + f(n) + \mathcal{O}(1)$.

The Idea

- The idea for proving that there are *many* K-degrees of random reals.

- We show that every so often the initial segment complexity of any given real β must oscillate downwards, say below $n + \log n + \mathcal{O}(1)$.

- Now we would like *our* real α to be random and have its complexity *high* at these very places where the complexity of β oscillates downwards.

- To do this we need to build a (suitable) Martin-Löf test to avoid the places where β's complexity oscillates downwards. This involves having some reasonable estimate of the measure of such places. The idea is to tie the low n's to m's where β's complexity is high enough to be able to estimate the sizes enough to be able to sparseify and create a ML-test.

- The places will be the m's from Miller's Theorem.

- Lemma: For every A and every m, define $n(m) = n$ if $A \upharpoonright m$ is the nth string under the standard length/lexicographic order. Then $K(A \upharpoonright n(m)) \leq n(m) + \log(n(m)) + c$ for some constant c.

- Define

  $\mathcal{D} = \{n \mid$there is a m so that $\Omega \upharpoonright m$ is the nth string under the standard length/lexicographic order and $K(\Omega \upharpoonright m) > m + \log m + \log\log m\}$. (Here $\log = \lfloor \log \rfloor$.)

- $K(\Omega \upharpoonright n) \leq n + \log n + c$.

- Note $K(n(m)) = K(\Omega \upharpoonright m) - c >$ $m + \log m + \log\log m - c =$ $\log n + \log\log n + \log\log\log n - c'$

- (Lemma) There is an infinite set $\mathcal{D}' \subseteq \mathcal{D}$ so that $\sum_{n \in \mathcal{D}'} 2^{-\log\log\log n} \leq \frac{1}{2}$.
- The decisive action is to use this to define tests:

  $U_k = \{y : (\exists n \in \mathcal{D}')[K(y \upharpoonright n) \leq n + \log n + \log\log n - k]\}$. Clearly, $U_k \supseteq U_{k+1}$ for every $k \in \mathbb{N}$. These turn out to have measure $\leq 2^{-k}C$.
- Thus there is a 2-random real α so that $K(\alpha \upharpoonright n) > n + \log n + \log\log n - k$ for every $n \in \mathcal{D}'$ and a fixed k. But for every $n \in \mathcal{D}'$, $K(\Omega \upharpoonright n) \leq n + \log n$ and $|\mathcal{D}'| = \infty$.

Ramping this up

- Use $\mathcal{D}_i = \{n \mid$there is a m so that $\Omega \upharpoonright m$ the nth string under the standard length/ lexicographic order and $K(\Omega \upharpoonright m) > \sum_{0 \le j \le i} \log^{(j)} m\}$ $(i \geq 2)$. etc.

- The we can show that this gives a way of avoiding countable sets. And showing that $\mu(\{x : x \leq_K y\} = 0$.

- Subsequently, Yu and Ding used this process and a clever argument to show that there are $2^{\aleph_0}$ many H-degrees.

- Note that the first argument we can replace Ω by $\Omega^{\emptyset^{(p)}}$ and get that $\Omega^p \not\leq_K \Omega^q$ if $q > p$. (Indeed $\limsup K(\Omega^q \upharpoonright n) - K(\Omega^p \upharpoonright n) = \infty$.)
- Unknown if there are infinitely many K -degrees of random Δ_2^0 randoms.

Joe Miller's Theorems

- 2-random= Kolmogorov random (!)
- if $p \neq q > 0$, $K(\Omega^p) \not\leq K(\Omega^q)$.
- if $q > 0$ $\mu(\{x : \Omega^q < x\} = 0$.
- He conjectures that there should be *maximal* K-degrees.
- Question: are there comparible K-degrees?

On Ω^A

- While on the subject of n-randomness this is of interest.
- What does it really mean? (Answer if V^A is a real ce relative to A and $A-$random, then it is a halting probability relative to a "genuine" machine: a U such that U^X is a universal machine for all X.
- There are non-delta 2 A with Ω^A delta 2.
- Ω^A is never low.
- If A is delta 2, then Ω^A is high.
- Invariant? lots of questions.

- Thank you.