Randomness and Computability 1: Basic Facts

Rod Downey
Victoria University
Wellington
New Zealand

References

- van Lambalgen's Thesis, Solovay's unpublished notes, and Li-Vitanyi Also new book "to appear" by Downey and Hirschfeldt prelim version on my home page, and one by Nies available (maybe) if you ask him.
- Calibrating Randomness (with Hirschfeldt, Nies and Terwijn) for BSL.
- Five Lectures on Algorithm Randomness, to appear Proceedings Computational prospects of Infinity
- Some Computability-Theoretical Aspects of Reals and Randomness, in The Notre Dame Lectures

Motivation

- What is “random”?
- How can we calibrate levels randomness? Among randoms?, Among non-randoms?
- How does this relate to classical computability notions, which calibrate levels of computational complexity?

MOTIVATION

- What is “random”?
- How can we calibrate levels randomness? Among randoms?, Among non-randoms?
- How does this relate to classical computability notions, which calibrate levels of computational complexity?
- Von Mises, Church, Solomonoff, Levin, 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 $\mu([\sigma]) = 2^{-|\sigma|}$.
- for uniformity, a real is always nonrational.
- Strings = members of $2^{<\omega} = \{0, 1\}^*$.

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 $\mu([\sigma]) = 2^{-|\sigma|}$.
- for uniformity, a real is always nonrational.
- Strings = members of $2^{<\omega} = \{0, 1\}^*$.
- There are theories for more general spaces, notably by Gács, (see his web site), but this is still under development.

Plain Kolmogorov Complexity

- Capture the incompressibility paradigm. Random means hard to describe, incompressible: e.g. 1010101010.... (10000 times) would have a short program.
- A string σ is random iff the only way to describe it is by hardwiring it. (Formalizing the Berry paradox)

Plain Kolmogorov Complexity

- Capture the incompressibility paradigm. Random means hard to describe, incompressible: e.g. 1010101010.... (10000 times) would have a short program.
- A string σ is random iff the only way to describe it is by hardwiring it. (Formalizing the Berry paradox)
- For a fixed machine N, we can define
- The Kolmogorov complexity $C(\sigma)$ of $\sigma \in \{0,1\}^*$ with respect to N, is $|\tau|$ for the shortest τ s.t. $N(\tau)\downarrow = \sigma$. (Kolmogorov)

- A string σ is N-random iff $C_N(\sigma) \geq |\sigma|$.
- A machine U is called weakly universal iff for all N, there is a d such that for all σ, $C_U(\sigma) \leq C_N(\sigma) + d$.

- A string σ is N-random iff $C_N(\sigma) \geq |\sigma|$.
- A machine U is called weakly universal iff for all N, there is a d such that for all σ, $C_U(\sigma) \leq C_N(\sigma) + d$.
- Actually we will always use universal machines where the e-th machine is coded in a computable way.

- A string σ is N-random iff $C_N(\sigma) \geq |\sigma|$.
- A machine U is called weakly universal iff for all N, there is a d such that for all σ, $C_U(\sigma) \leq C_N(\sigma) + d$.
- Actually we will always use universal machines where the e-th machine is coded in a computable way.
- They exist (Kolmogorov). Hence there is a notion of Kolmogorov randomness for strings up to a constant.

- A string σ is N-random iff $C_N(\sigma) \geq |\sigma|$.
- A machine U is called weakly universal iff for all N, there is a d such that for all σ, $C_U(\sigma) \leq C_N(\sigma) + d$.
- Actually we will always use universal machines where the e-th machine is coded in a computable way.
- They exist (Kolmogorov). Hence there is a notion of Kolmogorov randomness for strings up to a constant.
- Proof: We can enumerate the Turing machines $\{M_e : e \in \mathbb{N}\}$. Define

$$U(1^e 0\sigma) = M_e(\sigma).$$

This particular coding gives $C(\tau) \leq M_e(\tau) + e + 1$.

- A string σ is N-random iff $C_N(\sigma) \geq |\sigma|$.
- A machine U is called weakly universal iff for all N, there is a d such that for all σ, $C_U(\sigma) \leq C_N(\sigma) + d$.
- Actually we will always use universal machines where the e-th machine is coded in a computable way.
- They exist (Kolmogorov). Hence there is a notion of Kolmogorov randomness for strings up to a constant.
- Proof: We can enumerate the Turing machines $\{M_e : e \in \mathbb{N}\}$. Define

$$U(1^e 0\sigma) = M_e(\sigma).$$

  This particular coding gives $C(\tau) \leq M_e(\tau) + e + 1$.
- We will often write $=^+$, or $\leq^*$ where we mean $\pm O(1)$.

Definition

Thus we can define the plain Kolmogorov complexity of a string σ as $C(\sigma)$ for a fixed universal machine U.

- We can similarly do an oracle version of this and can define $C(x|y)$ as the Kolmogorov complexity of x given y.
- The unique string τ which first occurs of length $C(\sigma)$ is denoted by x^* (really x^*_C).

- Here are some basic facts about C-complexity:
  (I) $C(x, C(x)) =^* C(x^*)$.
  (I) $C(x|x^*) = O(1)$
  (III) $C(x, C(x)|x^*) =^* C(x^*|C(x), x) = O(1)$.
  (IV) $C(xy) \leq C(x, y) + O(1)$ where xy denotes the concatenation of x and y and $C(x, y)$ denotes $C(\langle x, y\rangle)$.

Plain Counting Theorem

- The following is the basic fact that makes the theory work.

Theorem (Plain Counting Theorem-Kolmogorov)

$|\{\tau : C(\tau) \leq |\tau| - d\}| \leq O(1)2^{|\tau|-d}$.

- Proof: pigeonhole principle.

Definition (Kolmogorov)

We say that σ is *C-random* iff $C(\sigma) \geq |\sigma|$.

Compression functions

- Thus plain complexity is a combinatorial fact

Definition (Nies, Stephan Terwijn)

We say that $F : \Sigma^* \mapsto \Sigma^*$ is a compression function if for all x $|F(x)| \leq C(x)$ and F is 1-1.

- Note that the counting theorem works for compression functions.
- Now we can form a Π^0_1 class of compression functions. We can apply then various basis Theorems, for instance, the Low Basis Theorem.
- There is a infinite low set of C-random strings.
- In some sense this is the best you could hope for. The collection of C-random strings is easily seen to be immune.

Compression functions

- Thus plain complexity is a combinatorial fact

Definition (Nies, Stephan Terwijn)

We say that $F : \Sigma^* \mapsto \Sigma^*$ is a compression function if for all x $|F(x)| \leq C(x)$ and F is 1-1.

- Note that the counting theorem works for compression functions.
- Now we can form a Π^0_1 class of compression functions. We can apply then various basis Theorems, for instance, the Low Basis Theorem.
- There is a infinite low set of C-random strings.
- In some sense this is the best you could hope for. The collection of C-random strings is easily seen to be immune.
- Proof: We can use the recursion theorem to play part of the universal machine, and lower the complexity of some string the opponent enumerates as part of a c.e. subset of the randoms.

C-overgraphs

- We can easily see that R_C, the collection of C-randoms is wtt complete.
- For each n, choose a length $f(n)$ and, at each stage s point at a string $\sigma(n, s)$ which is C_s-random.
- Should $\sigma(n, s)$ become nonrandom due to a play by our opponent choose the next string of this length. Should we see n enter $\emptyset'$ at s, we drops the complexity of $\sigma(n, s)$. (Here we use the recursion theorem)

Kummer's Theorem

- It was a question whether R_C could be tt-complete, so that the reduction above was non-adaptive.

Theorem (Kummer)

R_C and hence the overgraph $M_C = \{(x, y) : C(x) < y\}$ is tt-complete.

- The proof is tricky and nonuniform. It used blocks instead of the $\sigma(n, s)$ above and is a conjunctive tt-reduction. The nonuniformity comes from the combinatorics. A finite number of tries occur for these blocks, but this will be bounded and the number that occurs infinitely often is the one.

Muchnik's Theorem

- The following is easier and along the same lines.
- Theorem (An. A. Muchnik) The conditional overgraph $M = \{(x, y, n) : C(x|y) < n\}$ is creative

- The proof. We need $\emptyset' \leq_m M$.
- Parameter d known in advance.
- Construct possible g_x for $x \in [1, 2^d]$.
- Either we know $z \in \emptyset'$, or there is a unique y such that $g_x(z) = (x, y, d)$ and $x \in \emptyset'$ iff $g_x(z) \in M$.
- For some maximal x which enumerates elements infinitely often, g_x works.

- **Construction, stage** $s+1$ For each active $y \leq s$, find the least $q \in [1, 2^p]$ with

$$(q, y, d) \notin M_s.$$

  (Notice that such an x needs to exist since $\{q : (q, y, d) \in M\} < 2^d$.)
  If q is new, ie $(q', y, d) \in M_s$ for all $q' < q$, find the least z with $z \notin \emptyset'[s+1]$ and define

$$g_q(z) = (q, y, d).$$

- Now for any v, if v enters $\emptyset'[s+1]$, find the largest r, if any, with $g_r(v)$ defined. If one exists Find $\hat{y}$ with $g_r(v) = (r, \hat{y}, d)$. Declare that $\hat{y}$ is no longer active.

- Note that there must a largest $x \leq 2^d$ such that $\exists^\infty v(g_x(v) \in M)$. Call this x. We claim that g_x is the required m-reduction. Work in stages after which g_{x+1} enumerates nothing into M.

- Note that there must a largest $x \leq 2^d$ such that $\exists^\infty v(g_x(v) \in M)$. Call this x. We claim that g_x is the required m-reduction. Work in stages after which g_{x+1} enumerates nothing into M.
- Given z, since g_x is defined on infinitely many arguments and they are assigned in order,we can go to a stage s where either z has entered $\emptyset'[s]$, or $g_x(z)$ becomes defined, and $g_x(z) = (x, y, d)$ for some active y. $g_x(z)$ will be put into M should z enter $\emptyset'$ after s.

- There is a lot of very interesting work by Allender and others about what is efficiently reducible to R_C, and this (apparently) relates to standard classes like PSPACE, NP, etc. The point is that here the reductions are big.
- For instance, Allender, Buhrmann, Koucký look at the hypothesis

$$PSPACE = \cap_V P^{R_C^V}$$

(R_C^V is R_C for universal V.)

Complexity Oscillations

- Tempting but false $C(xy) \leq C(x) + C(y) + O(1)$. The false argument says : concatenate the machines

Complexity Oscillations

- Tempting but false $C(xy) \leq C(x) + C(y) + O(1)$. The false argument says : concatenate the machines
- The problem is where does x^* stop and y^* begin.

Complexity Oscillations

- Tempting but false $C(xy) \leq C(x) + C(y) + O(1)$. The false argument says : concatenate the machines
- The problem is where does x^* stop and y^* begin.
- Martin-Löf showed that the formula always fails for long enough srings and hence reals.

- Why? Take any α. Then, as a string $\alpha \restriction n$ corresponds to some number which we can interpret as a string using llex ordering: $\alpha \restriction n$ is the m-th string.

- Why? Take any α. Then, as a string $\alpha \upharpoonright n$ corresponds to some number which we can interpret as a string using llex ordering: $\alpha \upharpoonright n$ is the m-th string.
- Now consider the program that does the following. It takes a strings ν, interprets its length $m_\nu = |\nu|$ as a string, $\sigma = \sigma_m$ and outputs $\sigma\nu$.
- Apply this to the string τ whose length is m th code of $\alpha \upharpoonright n$.

- Why? Take any α. Then, as a string $\alpha \upharpoonright n$ corresponds to some number which we can interpret as a string using llex ordering: $\alpha \upharpoonright n$ is the m-th string.
- Now consider the program that does the following. It takes a strings ν, interprets its length $m_\nu = |\nu|$ as a string, $\sigma = \sigma_m$ and outputs $\sigma\nu$.
- Apply this to the string τ whose length is m th code of $\alpha \upharpoonright n$.
- The output would be much longer, and would be $\alpha \upharpoonright m + n$, with input having length m. Thus $C(\alpha \upharpoonright m + n) < m + n - O(1)$.

- This phenomenom is fundamental in our understanding of Kolmogorov complexity and is called complexity oscillations.
- There are several known ways to get round this problem to cause only to get the information provided by the bits of the strings.

Symmetry of Information

- The information content of a string y in a string x is defined as
$$I(x:y) = C(y) - C(y|x).$$
- (Levin-Kolmogorov)
$$\begin{aligned} I(x:y) &= I(y:x) \pm O(\log n) \\ &= I(y:x) \pm O(\log C(x,y)) \end{aligned}$$
where $n = \max\{|y|, |x|\}$.
- (restated) $C(x,y) = C(x) + C(y|x) + O(\log C(x,y))$

UNIVERSAL COMPUTERS

- Levin, Gaćs, Chaitin, Schnorr.
- Computers have alphabet $\{0, 1\}$.
- A computer M is prefix-free if

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

- A prefix-free machine is universal if every other one is coded in it.
- They exist, same proof.
- Building them uses what is now called Kraft-Chaitin.

KRAFT-CHAITIN

THEOREM (KRAFT, LEVIN, SCHNORR)

(I) *If A is prefix-free then $\sum_{n \in A} 2^{-|n|} \leq 1$.*

(II) (This part is now called Kraft-Chaitin, or Chaitin simulation) *Let $d_1, d_2, \cdots$ be a collection of lengths, possibly with repetitions, Then $\Sigma 2^{-d_i} \leq 1$ iff there is a prefix-free set A with members σ_i and σ_i has length d_i. Furthermore from the sequence d_i we can effectively compute the set A.*

- (Restatement) Suppose that we are effectively given a set of "requirements" $\langle n_k, \sigma_k \rangle$ for $k \in \omega$ with $\sum_k 2^{-n_k} \leq 1$. Then we can (primitive recursively) build a prefix-free machine M and a collection of strings τ_k with $|\tau_k| = n_k$ and $M(\tau_k) = \sigma_k$.

Prefix-free randomess

- Prefix freeness gets rid of the use of length as extra information: Machines concatenate!
- The prefix-free complexity $K(\sigma)$ of $\sigma \in \{0,1\}^*$ is $|\tau|$ for the shortest τ s.t. $M(\tau)\downarrow = \sigma$.
- Note now $K(\sigma) \le |\sigma| + K(|\sigma|) + d$, about $n + 2\log n$, for $\sigma| = n$.
- Build M, $M(z\sigma) = \sigma$ if $U(z) = |\sigma|$.

K-Counting Theorem

Theorem (Counting Theorem-Chaitin)

$$|\{\sigma : |\sigma| = n \wedge K(\sigma) \leq n + K(n) - c\}| \leq O(1)2^{n+K(n)-c}.$$

K-Counting Theorem

Theorem (Counting Theorem-Chaitin)

$|\{\sigma : |\sigma| = n \wedge K(\sigma) \leq n + K(n) - c\}| \leq O(1)2^{n+K(n)-c}$.

- The easiest proof uses semimeasures. A partial function $\widehat{K} : 2^{<\omega} \mapsto \mathbb{N}$ such that
  (I) $\sum_{\sigma \in 2^{<\omega}} 2^{-\widehat{K}(\sigma)} \leq 1$, and,
  (II) $\{\langle \sigma, k \rangle : \widehat{K}(\sigma) \leq k\}$ is c.e..

K-Counting Theorem

Theorem (Counting Theorem-Chaitin)

$|\{\sigma : |\sigma| = n \wedge K(\sigma) \leq n + K(n) - c\}| \leq O(1)2^{n+K(n)-c}$.

- The easiest proof uses semimeasures. A partial function $\widehat{K} : 2^{<\omega} \mapsto \mathbb{N}$ such that
  (I) $\sum_{\sigma \in 2^{<\omega}} 2^{-\widehat{K}(\sigma)} \leq 1$, and,
  (II) $\{\langle \sigma, k \rangle : \widehat{K}(\sigma) \leq k\}$ is c.e..
- There is a universal minimal one:

$$\widehat{K}(x) = \min_{k \geq 0} \{\widehat{K}_k(x) + k + 1\}.$$

- Using KC K is the same thing!
- Namely, at stage s, if we see $K_s(\sigma) = k$ and $K_{s+1}(\sigma) = k' < k$ enumerate a Kraft-Chaitin axiom $\langle 2^{-(k'+1)}, \sigma \rangle$ to describe M, and hence generate $\widehat{K} = K_M$.

- Many proofs exploit the minimality of K.
- Strictly speaking, A discrete semimeasure is function $m : 2^{<\omega} \mapsto \mathbb{R}^+ \cup \{0\}$ such that

$$\sum_{\sigma \in 2^{<\omega}} m(\sigma) \leq 1.$$

- NB Discrete Lebesgue measure is $\lambda(\sigma) = 2^{-2|\sigma|-1}$.
- Let m denote the minimal universal discrete semimeasure. Then
- $K(\sigma) = -\log m(\sigma) + O(1)$.

The Coding Theorem

- Let $Q_D(\sigma) = \mu(D^{-1}(\sigma))$, the probability tht σ is output.

Theorem (Levin)

$-\log m(\sigma) = -\log Q(\sigma) + O(1) = K(\sigma) + O(1).$

- (Proof) $Q(\sigma) \geq 2^{-K(\sigma)} = 2^{-|\sigma^*|}$, since $D(\sigma^*) = \sigma$.

- (Proof) $Q(\sigma) \geq 2^{-K(\sigma)} = 2^{-|\sigma^*|}$, since $D(\sigma^*) = \sigma$.
- So $-\log Q(\sigma) \leq K(\sigma)$.

- (Proof) $Q(\sigma) \geq 2^{-K(\sigma)} = 2^{-|\sigma^*|}$, since $D(\sigma^*) = \sigma$.
- So $-\log Q(\sigma) \leq K(\sigma)$.
- But: $\sum 2^{-logQ(\sigma)} \leq \sum_\sigma Q(\sigma) \leq 1$.
- Now use minimality of K.

- (Proof) $Q(\sigma) \geq 2^{-K(\sigma)} = 2^{-|\sigma^*|}$, since $D(\sigma^*) = \sigma$.
- So $-\log Q(\sigma) \leq K(\sigma)$.
- But: $\sum 2^{-logQ(\sigma)} \leq \sum_\sigma Q(\sigma) \leq 1$.
- Now use minimality of K.
- (Remark) It is not hard to show that for any σ $Q(\sigma)$ is random.

An Application

- One nice applications shows that within a fixed diameter there are relatively few descriptions.

Theorem (Levin, Chaitin)

There is a constant d such that for all c and all σ,

$$|\{\nu : U(\nu) = \sigma \wedge |\nu| \leq K(\sigma) + c\}| \leq d2^c.$$

An Application

- One nice applications shows that within a fixed diameter there are relatively few descriptions.

Theorem (Levin, Chaitin)

There is a constant d such that for all c and all σ,

$$|\{\nu : U(\nu) = \sigma \wedge |\nu| \leq K(\sigma) + c\}| \leq d2^c.$$

- The point here is that d is independent of $|\nu|$ and depends only on the Recursion Theorem, and c

Symmetry of Information

- $K(xy) \leq K(x) + K(y) + O(1)$.
- Define $I(x : y) = K(y) - K(y|x)$.
- Levin and Gács, Chaitin
  $I(\langle x, K(x)\rangle : y) = I(\langle y, K(y)\rangle : x) + O(1)$.
- (restated)
  $K(x, y) = K(x) + K(y|x^*) = K(x) + K(x|x, K(x))$.
- The proof uses KC again. And the Coding Theorem.

Prefix free randomness

- Levin-Chaitin random $K(x) \geq |x| + O(1)$.
- Strongly $K(x) \geq |x| + K(|x|) + O(1)$.
- Strongly K-random implies C-random implies K-random.
- NO reversals (the first is nontrivial and due to Solovay)

- As with life, relationships here are complex (Solovay)

$$K(x) = C(x) + C^{(2)}(x) + \mathcal{O}(C^{(3)}(x)).$$

and

$$C(x) = K(x) - K^{(2)}(x) + \mathcal{O}(K^{(3)}(x)).$$

- As with life, relationships here are complex (Solovay)

$$K(x) = C(x) + C^{(2)}(x) + \mathcal{O}(C^{(3)}(x)).$$

and

$$C(x) = K(x) - K^{(2)}(x) + \mathcal{O}(K^{(3)}(x)).$$

- These 3's are sharp (Solovay) That is, for example, $K = C + C^2 + C^3 + O(C^4)$ is NOT true.

- Is there a infinite low collection of strongly K-random strings? Joe Miller showed that the set is not co-c.e..

THEOREM (AN A MUCHNIK)

There exist universal prefix-free machines V and U such that

(I) M_K^V *is* tt*-complete.*

(II) M_K^U *(and hence* $\overline{R}_K^U$*) is not* tt*-complete.*

- The proof of (ii) is very interesting, using strategies for finite games do diagonalize against *tt*-reductions.

- Thus, the overgraph may or may not be tt-complete depending on the universal machine. Open for monotone complexity, open for the nonrandoms.

Monotone Complexity

- Levin's original idea here was to try to assign a complexity to the real itself. That is, think of the complexity of the real as the shortest machine that outputs the real. Hence now we are thinking of machines that take a program σ and might perhaps output a real α. (Nonsense unless α is computable)

Monotone Complexity

- Levin's original idea here was to try to assign a complexity to the real itself. That is, think of the complexity of the real as the shortest machine that outputs the real. Hence now we are thinking of machines that take a program σ and might perhaps output a real α. (Nonsense unless α is computable)
- The following definition can be applied to Turing machines with potentially infinite output, and to discrete ones mapping strings to strings. In this definition, we regard $M(\sigma)\downarrow$ to mean that at some stage s, $M(\sigma)\downarrow [s]$.

- We say that a machine M is monotone if its action is continuous. That is, for all $\sigma \preceq \tau$, if $M(\sigma)\downarrow$ and $M(\tau)\downarrow$ then

$$M(\sigma) \preceq M(\tau).$$

- Levin's (standard) monotone complexity Km is defined as follows. Fix a universal monotone machine U.

$$Km(\sigma) = \min\{|\tau| : \sigma \preceq U(\tau)\}.$$

- We say that a machine M is monotone if its action is continuous. That is, for all $\sigma \preceq \tau$, if $M(\sigma)\downarrow$ and $M(\tau)\downarrow$ then

$$M(\sigma) \preceq M(\tau).$$

- Levin's (standard) monotone complexity Km is defined as follows. Fix a universal monotone machine U.

$$Km(\sigma) = \min\{|\tau| : \sigma \preceq U(\tau)\}.$$

- If we only look at $2^{<\omega}$ then we get to Schnorr's process complexity.

Continuous Semimeasures

- The coding theorem relates K to discrete semimeasures. Here we would like an analog.
- Continuous semimeasures.
- A continuous semimeasure is a function $\delta : [2^{<\omega}] \mapsto \mathbb{R}^+ \cup \{0\}$ satisfying
  (i) $\delta([\lambda]) \leq 1$, and
  (ii) $\delta([\sigma]) \geq \delta([\sigma 0]) + \delta([\sigma 1])$.

- There is a minimal optimal continuous semimeasure δ. (Actually $\delta([\sigma]) = 2^{-|\sigma|}F(\sigma)$ where F is the optimal supermartingale, for those who know.)
- $KM(\sigma) = -\log \delta([\sigma])$.

- There is a minimal optimal continuous semimeasure δ. (Actually $\delta([\sigma]) = 2^{-|\sigma|}F(\sigma)$ where F is the optimal supermartingale, for those who know.)
- $KM(\sigma) = -\log \delta([\sigma])$.
- The analog of the Coding Theorem would state $KM = Km$. That is the probability that a string is output (KM) is the same as its Kolmogorov complexity (Km). Note $2^{-Km(\sigma)}$ is a semimeasure.

Gács Theorem

Theorem (Gács)

(i) *There exists a function f with $\lim_s f(s) = \infty$, such that for infinitely many σ,*

$$Km(\sigma) - KM(\sigma) \geq f(|\sigma|).$$

(ii) *Indeed, we may choose f to be the inverse of Ackkermann's function.*

- This shows $\leq_{Km}$ is not the same as $\leq_{KM}$. (Miller observation). Is this true for c.e. reals?
- Find a reasonable proof of Gács Theorem. (Here reasonable=one I can understand)

We turn from strings to looking at randomness for reals.

Three views of effective randomness

1. Measure-Theoretical:
   - Random means no distinguishing features. (Think of a statistical test as generating a set of tests: considered as open sets.)
   - In effective terms:
     - Avoids all effective sets of measure 0.

2 Algorithmic:

- Random means hard to describe, incompressible: e.g. 1010101010.... (10000 times) would have a short program.
- In effective terms:
- Initial segments have high Kolmogorov complexity.

3 Other views: e.g. random means unpredictable.

- No effective betting strategy succeeds on α.

Richard von Mises:

- Actually, the first attempt to "define" randomness was by von Mises 1919.
- Stochastic approach: $\alpha = a_1 a_2 \ldots$, "select" some subsequence assuming "acceptable" selection rules,
- say positions $f(1) < f(2) \ldots$, then $n \to \infty$, the number of $a_{f(i)} = 1$ divided by those with $a_{f(i)} = 0$ for $i \leq n$ should be 1.
- generalization of the law of large numbers.
- What are acceptable selection rules?

Richard von Mises:

- Actually, the first attempt to "define" randomness was by von Mises 1919.
- Stochastic approach: $\alpha = a_1 a_2 \ldots$, "select" some subsequence assuming "acceptable" selection rules,
- say positions $f(1) < f(2) \ldots$, then $n \to \infty$, the number of $a_{f(i)} = 1$ divided by those with $a_{f(i)} = 0$ for $i \leq n$ should be 1.
- generalization of the law of large numbers.
- What are acceptable selection rules?
- Some problems (later). Solved by Martin-Löf who said we should view effective statistical tests as effective null sets.

Martin-Löf randomness:

- A c.e. open set is one of the form $\bigcup_i (q_i, r_i)$ where $\{q_i : i \in \omega\}$ and $\{r_i : i \in \omega\}$ are c.e.. $U = \{[\sigma] : \sigma \in W\}$.
- A Martin-Löf test is a uniformly c.e. sequence $U_1, U_2, \ldots$ of c.e. open sets s.t.

$$\forall i(\mu(U_i) \leq 2^{-i}).$$

(Computably shrinking to measure 0)

Definition

α is Martin-Löf random if for every Martin-Löf test,

$$\alpha \notin \bigcap_{i>0} U_i.$$

Martin-Löf randomness:

- A c.e. open set is one of the form $\bigcup_i (q_i, r_i)$ where $\{q_i : i \in \omega\}$ and $\{r_i : i \in \omega\}$ are c.e.. $U = \{[\sigma] : \sigma \in W\}$.
- A Martin-Löf test is a uniformly c.e. sequence $U_1, U_2, \ldots$ of c.e. open sets s.t.

$$\forall i(\mu(U_i) \leq 2^{-i}).$$

(Computably shrinking to measure 0)

Definition

α is Martin-Löf random if for every Martin-Löf test,

$$\alpha \notin \bigcap_{i>0} U_i.$$

- (Solovay) same as for all c.e. sets of open intervals $\{I_n : n \in \omega\}$, with $\sum_n |I_n| < \infty$, $\alpha \in I_n$ for at most finitely many n.

UNIVERSAL TESTS

- Enumerate all c.e. tests, $\{W_{e,j,s} : e, j, s \in \mathbb{N}\}$, stopping should one threated to exceed its bound.
- $U_n = \cup_{e \in \mathbb{N}} W_{e,n+e+1}$.

Universal Tests

- Enumerate all c.e. tests, $\{W_{e,j,s} : e, j, s \in \mathbb{N}\}$, stopping should one threated to exceed its bound.
- $U_n = \cup_{e \in \mathbb{N}} W_{e,n+e+1}$.
- A passes this test iff it passes all tests. It is a universal martin-Löf test. (Martin-Löf)

Universal Tests

- Enumerate all c.e. tests, $\{W_{e,j,s} : e, j, s \in \mathbb{N}\}$, stopping should one threated to exceed its bound.
- $U_n = \cup_{e \in \mathbb{N}} W_{e,n+e+1}$.
- A passes this test iff it passes all tests. It is a universal martin-Löf test. (Martin-Löf)
- There are other clever constructions we may need later. (Kučera)

Kolmogorov Complexity, again

- From this point of view we should have all the initial segments of a real to be random.
- (Can also use selected places and factor in the complexity of the selection.)

- First try α, a real, is random iff for all *n*, $C(\alpha \upharpoonright n) \geq n - d$.
- By complexity oscillations (Martin-Löf) no such real can exist. The reason as we have seen is that *C* lacks the intentional meaning of Komogorov complexity.

K-RANDOMNESS

- Recall from earlier prefix freeness gets rid of the use of length as extra information:
- α is K- 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, $C(\alpha \upharpoonright n) > n - c$.

Schnorr's Theorem

Theorem (Schnorr)

K-random $\iff$ Martin-Löf random.

Kraft-Chaitin

- Recall from KC:
  Suppose that we are effectively given a set of "requirements" $\langle n_k, \sigma_k \rangle$ for $k \in \omega$ with $\sum_k 2^{-n_k} \leq 1$. Then we can (primitive recursively) build a prefix-free machine M and a collection of strings τ_k with $|\tau_k| = n_k$ and $M(\tau_k) = \sigma_k$.

Proof of Schnorr's Theorem

- $\Longrightarrow$ Suppose that α is not Martin-Löf random and $\alpha \in \cap_i U_i$, with $\mu(U_i) \leq 2^{-i}$.
- We use Kraft-Chaitin.
- Let $n \geq 3$. For all strings σ in U_{n^2}, enumerate the pair $|\sigma| - n, \sigma$ into B.
- By prefix-freeness, note that $\sum_B 2^{-n} \leq \sum_{n \geq 3} 2^{-n}(\mu(U_{n^2}) \leq \sum_{n \geq 3} 2^{n-n^2} \leq 1.$
- Thus by Kraft-Chaitin there is a machine M and strings $\tau_n \in \mathrm{dom} M$ with $M(\tau_n) = \sigma_n$ and $|\tau_n| = |\sigma| - n$. Since $\alpha \in \cap_n U_{n^2}$, this means that α is not Chaitin random.
- $\Longleftarrow$ Suppose that α is Martin-Löf random. Consider

$$U_k = \{\beta : \exists n (K(\beta \upharpoonright n) \leq n - k\}.$$

Then $\mu(U_k) \leq 2^{-k}$ (as the domain of M is prefix-free) and hence, as $\alpha \notin \cap_K U_k$, we are done.

K AND C

- Recall weakly Chaitin random string : $K(x) > |x|$.

COROLLARY (TO SCHNORR'S THEOREM)

For all c, there are infinitely many weakly K random strings σ with $C(\sigma) < |\sigma| - c$.

- (Proof) Consider the initial segments of a random real and C-oscillations.
- Actually with a more refied analysis of the complexity oscillations, you can have $C(x) \leq n - \log n$.

Lots of random reals

- $\mu\{A : A \text{ random }\} = 1$.
- Consider the Σ^0_2 class $\{A : \exists k \forall n K(A \restriction n > n - k\}$ contains all random reals.
- Hence there are ones of low Turing degree (low basis theorem) and hyperimmune free degree. (Kučera)
- There are ones of all jumps and even Δ^0_2 ones of all jumps (Kučera, Downey-Miller)

Levin and monotone complexity

- Recall from that for a a universal monotone machine U.

$$Km(\sigma) = \min\{|\tau| : \sigma \preceq U(\tau)\}.$$

Theorem (Levin's Theorem)
A is Martin-Löf random iff $Km(A \upharpoonright n) > n - O(1)$.

- (One direction holds since every prefix-free machine is monotone, the other we again put $[\sigma]$ into U_k iff $Km_M(\sigma) \leq |\sigma| - k$. where M is a universal monotone machine, and

$$\mu(U_k) = \sum\{2^{-|\sigma|} : K_M(\sigma) \leq |\sigma| - k \wedge$$

$$\forall \tau \prec \sigma (K_M(\tau) > |\tau| - k)\} \leq 2^{-k}.$$

- In fact A is Martin-Löf random iff $Km(A \upharpoonright n) = n - O(1)$.