Yet More on Dimension

Rod Downey
Victoria University
Wellington
New Zealand

REFERENCES

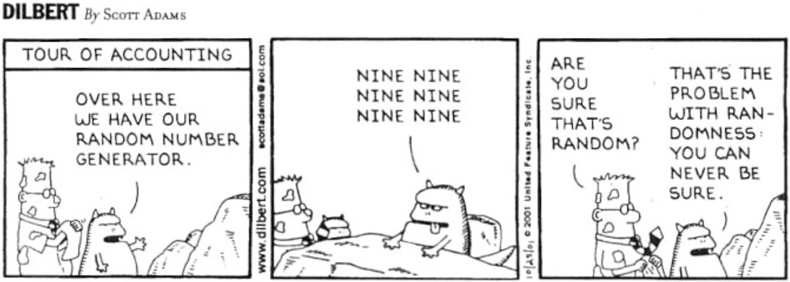
- Turing degrees of reals of positive effective packing dimension, (Downey and Greenberg) *Information Processing Letters*. Vol. 108 (2008), 298-303.
- PhD Thesis, Chris Conidis, University of Chicago, 2009.
- Effective Packing dimension and traceability (Downey and Ng), submitted.
- A Real of strictly positive packing dimension that does not compute a real of packing dimension 1, submitted, Chris Conidis.

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 $\mu([\sigma]) = 2^{-|\sigma|}$.
- Strings = members of $2^{<\omega} = \{0, 1\}^*$.

What the idea is

- Use tools from complexity theory and computability theory to understand the intuitive idea of randomness.
- How to reconcile the fact that one string looks more random than another but statistically they occur with the same probability.
- How to understand the idea that an individual sequence can be random; or “somewhat random” (and what does that mean, anyway?)
- Applications: complexity, crypto, quantum, Brownian motion, analysis, and others.


DILBERT By Scott Adams
TOUR OF ACCOUNTING
OVER HERE WE HAVE OUR RANDOM NUMBER GENERATOR.
NINE NINE NINE NINE NINE NINE
ARE YOU SURE THAT'S RANDOM?
THAT'S THE PROBLEM WITH RANDOMNESS: YOU CAN NEVER BE SURE.
www.dilbert.com
scottadams@aol.com
© 2001 United Feature Syndicate, Inc.

Which seem random?

000000000000000000000000000000000000000000000000000000000

00110100110100110100110100110100110100110100110100110100

0100011011000001010011100101110111000000010010001101000

0010011011011000100011110101001110110010011000000010110

0101011101101111011100100110101101110011011010000110111

0111011111001101100110100100001111100110110000001101110

00000110001011100010000000010100001011010100000001000000

0101001101111011011101010100000101111000000101011101010

Three Approaches to Randomness at an Intuitive Level

- The statistician's approach: Deal directly with rare patterns using measure theory. Random sequences should not have effectively rare properties. (von Mises, 1919, finally Martin-Löf 1966)
- Computably generated null sets represent effective statistical tests.
- The coder's approach: Rare patterns can be used to compress information. Random sequences should not be compressible (i.e., easily describable) (Kolmogorov, Levin, Chaitin 1960-1970's). Want $K(\alpha \upharpoonright n) \geq n$ for all n.
- Kolomogorov complexity; the complexity of σ is the length of the shortest description of σ.
- The gambler's approach: A betting strategy can exploit rare patterns. Random sequences should be unpredictable. (Solomonoff, 1961, Scnhorr, 1975, Levin 1970)

- No time to get into exact definitions.
- The plain Kolmogorov complexity $C(\sigma)$ is the *length* of the shortest description of σ via a universal transducer.
- Also use prefix-free complexity where the machine works like telephone numbers. (Levin, etc) This is denoted by K

Definition

A real α is 1-random iff $K(\alpha \upharpoonright n) \geq n - O(1)$. for all n.

- Recall that a martingale is a betting strategy $F : 2^{<\omega} \to \mathbb{R}^+ \cup \{0\}$ so that $F(\sigma) = \frac{F(\sigma 0)+F(\sigma 1)}{2}$. If $=$ is replaced by $\leq$ then this is a supermartingale.
- Succeeds if $\limsup_{n\to\infty} F(\alpha \upharpoonright n) = \infty$.
- For example, a random real should have long sequences of 0's else I could devise a (computable) martingale to succeeed on it.
- Recall α is 1-random iff no c.e. supermartingale succeeds on alpha. Here c.e. is computable from below. (Schnorr)

Orders

- This is concerned with the "speed" of success.
- Schnorr called a function h and order, if h is nondecreasing and $\lim_n h(n) = \infty$. Computable unless specified otherwise.
- If F is a martingale and h is an order the h-success set of F is the set:

$$S_h(F) = \{\alpha : \limsup_{n\to\infty} \frac{F(\alpha \upharpoonright n)}{h(n)} \to \infty\}.$$

- (Schnorr) A real α is Schnorr random iff for all computable orders h and all computable martingales F, $\alpha \notin S_h(F)$.

- 1895 Borel, Jordan
- Lebesgue 1904 measure
- In any n-dimensional Euclidean space, Carathéodory 1914

$$\mu^s(A) = \inf\{\sum_i |I_i|^s : A \subset \cup_i I_i\},$$

where each I_i is an interval in the space.
- 1919 Hausdorff s fractional; and refine measure 0.
- For $0 \leq s \leq 1$, the s-measure of a clopen set $[\sigma]$ is

$$\mu_s([\sigma]) = 2^{-s|\sigma|}.$$

- Mayordomo has the following characterization of effective Hausdorff dimension:
- (Lutz) An s-gale is a function $F : 2^{<\omega} \mapsto \mathbb{R}$ such that

$$F(\sigma) = 2^s(F(\sigma 0) + F(\sigma 1)).$$

Similarly we can define s-supergale, etc.

- Theorem (Mayordomo) For a class X the following are equivalent:
  (I) $\dim(X) = s$.
  (II) $s = \inf\{s \in \mathbb{Q} : X \subseteq S[d]$ for some s-supergale $F\}$.
- Lutz says the following:
  "Informally speaking, the above theorem says the the dimennsion of a set is the most hostile environment (i.e. most unfavorable payoff schedule, i.e. the infimum s) in which a single betting strategy can achieve infinite winnings on every element of the set."

- In (ii) we can replace supergale by gale because of the work of Hitchcock.
- This requires work. Essentially you show that for all $\epsilon > 0$ there is a $s + \epsilon$-martingale which is universal for all s-supermartingales.
- Open question: is there e.g. a multiplicatively optimal s-gale? Can one delete the ϵ from Hitchcock's Theorem?

- Theoerm (Mayordomo): The effective Hausdorff dimension of a real α is

$$\liminf_{n\to\infty} \frac{K(\alpha \restriction n)}{n} = (\liminf_{n\to\infty} \frac{C(\alpha \restriction n)}{n})$$

- (Schnorr) “To our opinion the important statistical laws correspond to null sets with fast growing orders. Here the exponentially growing orders are of special significance.”
- When asked at Dagstuhl he commented that he did not have Hausdorff dimension in mind.

Extracting randomness

- An easy example of something which has effective dimension $\frac{1}{2}$ is to take Ω and spread it out by inserting 0's every second bit. (Tadaki etc)
- Question: (Reimann, Terwijn) Can randomness always be extracted from positive dimension? What about dimension 1?
- Question (Reimann) Can dimension 1 always be extracted from positive dimension.

- Theorem (Miller) There is a Turing cone of dimension $\frac{1}{2}$.
- Theorem (Greenberg and Miller) There is a real of effective Hausdorff dimension 1 of minimal degree.
- Theorem (Zimand) Hausdorff dimension 1 can be extracted from two independent sources of positive dimension. (In fact $1-\epsilon$) can be extracted from independend sources where the initial segment (plain) complexity is eventually bigger than $c \log n$ for all c.)

- (GM) There is a real of effective Hausdorff dimension 1 which does not bound a random real. (Earlier claimed was of minimal degree but that is open as it relied on Kumabe-Lewis.)
- The proof idea. First generalize the notion of s-measure to functions (orders).
- Observe that if the order is sufficiently slowly growing then the resultant set has effective Hausdorff dimension 1.
- Now force with bushy (Kumabe) trees in something like "computably bounded" Baire space. This is a kind of miniature Prikry forcing.

Zimand's Theorem

- "small" is $O(\log)$ for our purposes.
- Independence: want to express the fact that X and Y have little common information.
- X and Y are C-independent iff for all n, m, $C(X \upharpoonright nY \upharpoonright m) \geq C(X \upharpoonright n) + C(Y \upharpoonright m) - O(\log n + \log m)$.
- A stronger form is noted by Calude and Zimand $C^X(y \upharpoonright n) \geq C(Y \upharpoonright n) - O(\log n)$ and $C^Y(X \upharpoonright n) \geq C(X \upharpoonright n) - O(\log n)$.
- Now suppose we have independent sources X and Y of positive dimension.
- Break the X and Y into blocks $X_1 X_2 \ldots, Y_1 Y_2 \ldots$ suitably chosen so that the conditional complexity of X_{i+1} is reasonably high relative to $X_1 \ldots X_i$.

- Done using: If q are rational, and that for almost all n, $C(X \upharpoonright n) > qn$ and $C(Y \upharpoonright n) > qn$. Let $0 < r < q$. For any n_0 sufficiently large, if we take $0 < r' < q - r$, and then $n_1 = \lceil \frac{1-r}{r'} \rceil n_0$. Then: $C(X \upharpoonright_{n_0+1}^{n_1} | X \upharpoonright n_0) > r(n_1 - n_0)$.
- Thus if $b = \lceil \frac{1-r}{r'} \rceil$.
- Let $t_0 = 0$ and $t_1 = b(t_0)$ with $t_{i+1} = b(t_0 + \cdots + t_i)$. For $i \geq 1$ define $X_i = X \upharpoonright_{t_{i-1}}^{t_i}$.
- $|X_i| = |Y_i| = n_0 b^2 (1 + b)^{i-3}$ for $i \geq 3$.

The combinatorial heart

- Compress the pair $E_i(X_i, Y_i) \mapsto Z_i$. We get a truth table reduction generated by the sequence $E_1, E_2, \ldots$..
- $Z = Z_1 Z_2 \ldots$ is the desired real.
- We say that a function $E : 2^n \times 2^n \to 2^m$ is $(r, 2)$-regular iff for every $k_1, k_2 \geq rn$, and any subsets $B_i \subseteq 2^n$ with $|B_i| = k_i$ for $i = 1, 2$, then for any $\sigma \in 2^m$,

$$|E^{-1}(\sigma) \cap (B_1 \times B_2)| \leq \frac{2}{2^m} |B_1 \times B_2|.$$

- Here $m = m_i = i^2$. The idea is that any target string z has essentially the same number of pre-images in $B_1 \times B_2$, and hence $E^{-1}(z) \cap B \times B$ can be enumerated effectively, so that if z has low complexity, then it becomes too easy to describe the pair. (Devil in details)

The extractor idea

- Independent strings x and y of length n with $C(x), C(y) = qn$ for positive rational q.
- $E : 2^n \times 2^n \to 2^m$ for each suitably large enough rectangle $B_1 \times B_2$ E maps about the same number of pairs to each $\tau \in 2^m$.
- $B \times B \in 2^{qn} \times 2^{qn}$, any $A \subseteq 2^m$, $|E^{-1}(A)| \approx \frac{|B \times B|}{2^m}|A|$.
- $z = E(x, y)$, the C-complexity of z must be large.
- If $C(z) < (1 - \epsilon)m$, then we note that
  (I) The set $B = \{\sigma \in 2^n \mid C(\sigma) = qn\}$ has size aproximately 2^{qn}.
  (II) The set $A = \{\tau \in 2^m \mid C(\tau) < (1 - \epsilon)m\}$ has size $< 2^{(1-\epsilon)m}$.
  (III) $(x, y) \in E^{-1}(A) \cap B \times B$.
- $|E^{-1}(A) \cap B \times B| \leq \frac{(2^{qn})^2}{2^{\epsilon m}}$.
- Hence $C(x, y) \leq 2qn - \epsilon m$, by c.e. listing.
- But, x and y are C-independent and hence $C(xy) \approx C(x) + C(y) = 2qn$, a contradiction.

The Construction

- Step 1. Split $X = X_1X_2\ldots$ and $Y = Y_1Y_2\ldots$ as above, using the parameters $r = \frac{q}{2}$ and $r' = \frac{q}{4}$..
- We remark that for each i

$$C(X_i|\overline{X}_{i-1}) > rn_i \text{ and } C(Y_i|\overline{Y}_{i-1}) > rn_i.$$

- Step 2. For the parameter $m_i = i^2$, find a $(\frac{r}{2}, 2)$-regular function. Define $Z_i = E_i(X_i, Y_i)$.
- Step 3. Define $Z = Z_1Z_2\ldots$.

Packing dimension

- Idea is to replace outer measure by inner measure.
- We use the Athreya, Hitchcock, Lutz, Mayordomo characterization. The packing dimension of a real α is of a real α is

$$\limsup_{n\to\infty} \frac{K(\alpha \upharpoonright n)}{n} = (\limsup_{n\to\infty} \frac{C(\alpha \upharpoonright n)}{n})$$

- Interesting as 2-generics have high efffective packing dimension, measure meets category.

- What Turing degrees contain reals of high packing dimension?
- Fortnow,Hitchcock,Aduri,Vinochandran, Wang have proven that if a real has packing dimension above > 0, then there is one of the same weak truth table degree of packing dimension $1 - \epsilon$.
- hence for degrees a 0-1 Law for effective packing dimension.
- (Open Question) is there a real of effective packing dimension 1 inside each degree of packing dimension 1?

THE PROOF

- This proof is due to Bienvenu.
- Have $K(X \upharpoonright n) \geq tn$ some t. Break X into intervals of size $[m^k, m^{k+1})$ a large number. Then for any $t' < \frac{t}{m}$ $\exists^\infty k C(X \upharpoonright m^k) \geq t' m^k$. (Kolmogorov computations)
- Now let $s = \limsup_k \frac{C(X \upharpoonright m^k)}{m^k}$.
- Now we have rationals $s_1 < s < s_2$ and when we see τ_k with $\tau_k \mapsto X \upharpoonright m_k$ $|\tau_k| \geq s_1 m_k$, $|\tau_k| < s_2 m^k$ we output $Z_k = \tau_k$. Then $Z = Z_1 Z_2 \ldots$ works by easy calculations.
- The original proof was a bit different, but also nonuniform, and actually gave polynomial time reductions using complex multisource extractors of Impagliazzo and Widgerson.

The following lemma is implicit in, e.g. Conidis

LEMMA

There is a computable mapping $(\sigma, \epsilon) \mapsto n_\epsilon(\sigma)$ which maps a finite binary string $\sigma \in 2^{<\omega}$ and a positive rational ϵ to a natural number n such that there is some binary string τ of length n such that

$$\frac{K(\sigma\tau)}{|\sigma\tau|} \geq 1 - \epsilon.$$

- We prove this theorem of Downey and Greenberg.
- We force with clumpy trees. These are clumps generated by the n_ϵ above and separated by long stretches.
- The Lemma allows us to make sure that we only have the branches of the perfect clumpy trees at the clumps in a Spector style forcing.

- The same kind of idea can be used to construct a rank one c.e. real of packing dimension 1. (Conidis)
- Have a clump, move only left, with long stretches of zeroes extending.
- Imteresting as this is not possible for Hausdorff dimension.

DEGREES

- (Downey, Jockusch, Stob) Recall that **a** is array noncomputable iff for all $f \leq_{wtt} \emptyset'$ there is a function $g \leq_T \mathbf{a}$ such that
$$\exists^\infty n(g(n) > f(n)).$$
- Array computability is stronger than being totally ω-c.e. (DG) where *bfb* is this iff all functions $g \leq_T \mathbf{b}$ are ω-c.e..
- These latter ones crop up in randomness via e.g. computable finite randomness (Brodhead,D,Ng). Also in the cL-degrees (Barmpalias,D, Greenberg) These c.e. degrees are definable (D, Greenberg, Weber)

THEOREM

(DG) A c.e. degree contains a real of effective packing dimension 1 iff it is array noncomputable.

- One direction. First notice that for c.e. sets, array computable is the same as traceable. (Ismukhametov)
- That is for any computable order h, and all functions $g \leq_T A$, there is a weak array $W_{q(n)} : n \in \omega$, such that $|W_{q(n)}| < h(n)$ and $g(n) \in W_{h(n)}$.
- Think $g(n) = A \upharpoonright n$.
- If the trace is very slow growing, then we can describe with very few bits of information, an idea of Kummer.

- The harder direction. To make the real complex, at some clump we need to be able to move left often enough lift the dimension.
- Then you could use the classical version of anc.
- c.e. set A is anc iff for all very strong arrays $D_{k(n)} : n \in \omega$ (ie $|D_{k(n+1)}| > |D_{k(n)}|$), for all e there is a n with $W_e \cap D_{k(n)} = A \cap D_{k(n)}$. This is a kind of "multiple permitting".
- Actually works for pb-generic. so outside of the c.e. degrees.

- Kummer's gap. We know that a c.e. set can have maximal complexity $C(A \upharpoonright n)$ as $2 \log n$. Solovay showed that it is impossible to have that almost always.
- (Kummer) Either a c.e. degree is array computable and all initial segments are within $(1+\epsilon)\log n + O(1)$. or the degree contains a set which is infinitely often $2\log n - O(1)$.

- First guess: packing dimension 1 iff anc.
- False superlow randoms are ac, and similarly hypermmune-free randoms.
- Second guess: packing dimension 1 iff non-c.e. traceable. Reasonable since random reals are all non-c.e. traceable.
- Theorem (Downey and Ng) There is a Δ^0_3 real A which is of hyperimmune-free degree and not c.e. traceable, such that every real $\alpha \leq_T A$ has effective packing dimension 0.
- Maybe this has something to do with lowness like Schnorr, Kurtz etc:
- Theorem (Downey and Ng) There is a real $A \leq_T \emptyset'$ which is not c.e. traceable, such that every real $\alpha \leq_T A$ has effective packing d imension 0.

- The proofs again use this notion of highly branching trees instead of Cantor space, (a finite extension argument+) over a Spector-style forcing. within the sequence of conditions, for $A \in [T_e]$ we need to kill off $\Phi_j^A(x \upharpoonright n) \leq \frac{x}{2}$ for almost all x. The fatness of the tree will be enough to make sure that there is enough of the condition left to perform the construction.
- This is an external function describing the splits of the tree. Diagonalization is possible as the tracing must be arbitrarily slow.
- Leaving enough of a tree relies on a certain level by level "majority vote" argument. This relies on the fact we only need to describe *sets* below the trees rather than *functions*.
- In some sense this gives implicit descriptions on the survivors of the tree, and hence allows us to keep the complexity down with long intervals and clumps.

CONIDIS' THEOREM

- There is a degree **a** of packing dimension 1 that does not contain a real of packing dimension 1.
- The proof uses the Downey-Ng ideas, with slowly growing trees.
- Question : Is there a degree of Hausdorff dimension 1 that does not compute a real of (packing) (Hausdorff) dimension 1?

Other interesting things

- Shift complex sets for all n, m there is a d such that the $C(A \restriction_m^{m+n} \geq (1 - \epsilon)n - d$.
- The exist (Levin) can be used for aperiodic tilings.
- Miniaturizing things e.g. Automatic dimension (Schnorr, Lutz) Lempl-Ziv, etc
- Poly classes and separations. E.g. can there be a polynomial reduction of $\emptyset'$ to the collection of non-random strings. (Allender etc)
- Many, many more: see my recent open question paper (home page)

Thank you