Algorithmic Randomness and Complexity

Rod Downey
Victoria University
Wellington
New Zealand

Melbourne, 2011

- Lets begin by examining the title:

- Algorithmic
- Randomness
- and Complexity

Algorithmic

- Etymology : Al-Khwārizmī, Persian astronomer and mathematician, wrote a treatise in 825 AD, On Calculation with Hindu Numerals, together with an error in the Latin translation.

- What we intuitively mean
- From a set of basic instructions (ingredients) specify a mechanical method to obtain the desired result.

- Already you can see that I plan to be sloppy, but you should try to get the feel of the subject.

No. 1 – THE REVEREND JOHN MACFARLANE

(Reel)

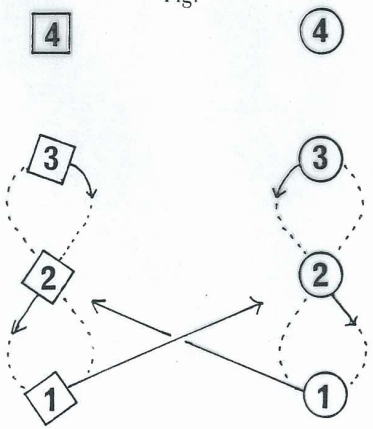
MUSIC *Bars*	DESCRIPTION
1– 8	1st woman dances a reel of three on the men's side with 2nd and 3rd men, while 1st man dances a reel of three on the women's side with 2nd and 3rd women. (Fig.)
	1st couple finish in partner's place.
9–12	1st couple dance a half figure of eight round 2nd couple.
13–16	1st couple, joining both hands, dance four slip steps down the middle to third place and then set with hands still joined. (1st man sets to the left and then to the right.) 2nd and 3rd couples step up and 4th couple step in to meet on bars 15–16.
17–24	1st and 4th couples poussette.
25–28	2nd couple with 3rd couple and 4th couple with 1st couple dance right hands across once round to places.
29–32	2nd, 3rd, 4th and 1st couples turn partners once round with the left hand.
	Repeat with a new top couple.

This dance commemorates the 150th anniversary of the founding in Wellington of New Zealand's first Scots Church, later known as St. Andrews.

The Rev. John Macfarlane, the first minister of Martyr's Memorial Church, Paisley, arrived in New Zealand on 20th February, 1840 and he held the first service on the beach at Petone on Sunday 23rd February.

Devised by Gary W. Morris (New Zealand Branch).

Fig.



Top

ice when it reaches the mushy stage and every 30 minutes after that until it is ready to serve, to insure smoothness. Garnish with pitted black cherries.

CREAM FRITTERS

READY TRAY — Serves 4 to 6

4 egg yolks
¼ cup sugar
½ cup flour
Salt to taste
4 cups milk, scalded
1 teaspoon grated orange or lemon rind
1 egg, beaten
Breadcrumbs
2 tablespoons oil
2 tablespoons butter
Powdered sugar
2 tablespoons brandy or rum

Beat egg yolks and sugar in top of double boiler. Cook over low heat, stirring with wooden spoon until slightly thickened. Mix in ¼ cup flour, salt and gradually add milk. Simmer, stirring, until very thick. At no time allow to boil. Blend in rind.

Rinse a square dish or pan with cold water and pour in mixture to a depth of 2 inches. Chill until firm. Cut into squares or rectangular pieces 2 inches long. Dip in remaining flour, in egg and then in breadcrumbs. Brown gently on both sides in hot oil and butter. Serve sprinkled with sugar, and flame with heated brandy or rum.

FRIED RICOTTA

READY TRAY — Serves 8

½ pound macaroons
1 pound ricotta cheese
Pinch cinnamon
3 eggs
Breadcrumbs
¼ pound butter
Powdered sugar
Brandy

Algorithmic

- From a set of basic instructions (ingredients) specify a mechanical method to obtain the desired result.

Greatest Common Divisors

- The greatest common divisor of two numbers x and y is the biggest number that is a factor of both.
- For instance, the greatest common divisor, gcd(4,8) is 4. gcd(6,10)=2; gcd(16,13)=1.
- Euclid, or perhaps Team Euclid, (around 300BC) devised what remains the "best" algorithm for determining the gcd of two numbers.

Euclid's Algorithm

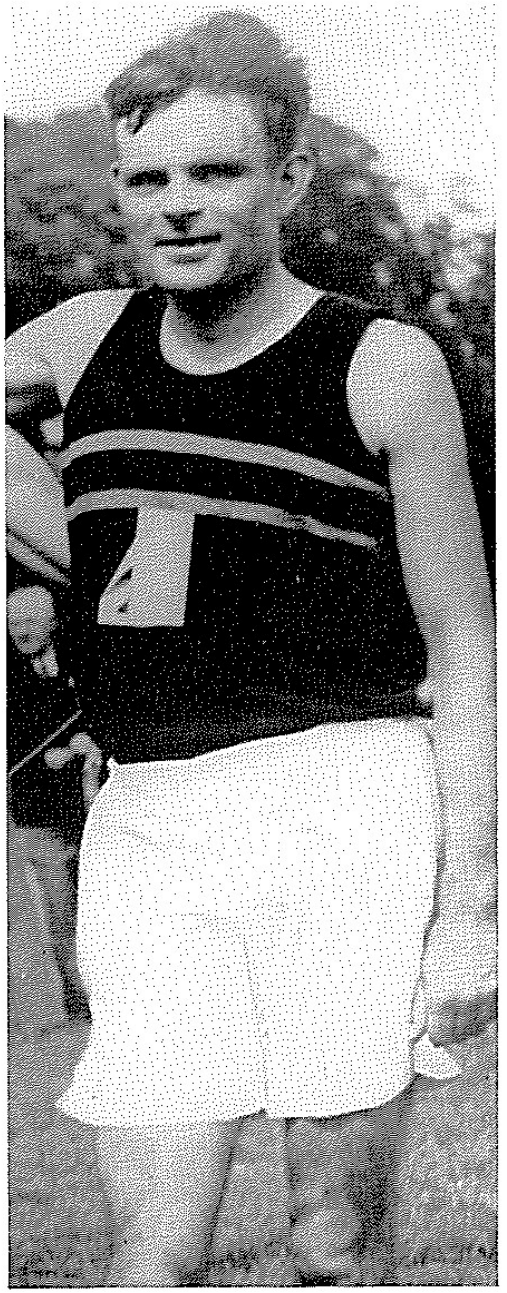
- To find gcd(1001,357).
- $1001 = 357 \cdot 2 + 287$
- $357 = 287 \cdot 1 + 70$
- $287 = 70 \cdot 4 + 7$
- $70 = 7 \cdot 10$
- 7=gcd(1001,357).

Computable functions and Church's Thesis

- The notion of a Computable Function can be made precise and was done in the 1930's by people like Church, Gödel, Turing and others.
- Became implemented by the work of Turing, von Neumann and others.

- Commonly accepted is Church's Thesis that the intuitively computable functions are the same as those defined by Turing machine (or your favourite programming language, such as JAVA, C++, etc.)
- Trickier when we talk about complexity theory. (feasible is a subset of polynomial time on a Turing Machine)

Aged 5



After a successful race. May, 1950

The Enigma Machine, employed by the Germans to encrypt classified and sensitive messages during World War II. (HultonArchive/Getty Images)

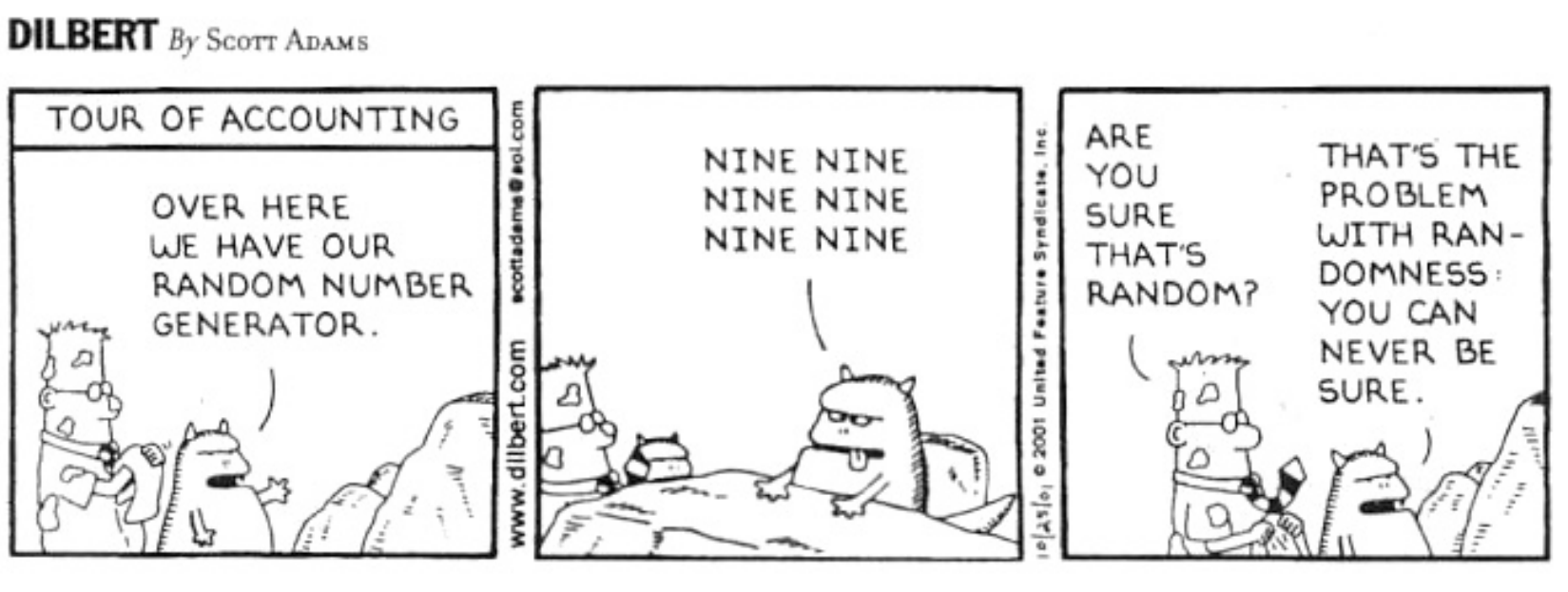
John von Neumann, Princeton, 1932

Randomness

"How dare we speak of the laws of chance?
Is not chance the antithesis of all law?"

— Joseph Bertrand, Calcul des Probabilités, 1889

Intuitive Randomness



Intuitive Randomness

Which of the following binary sequences seem random?

A 00000000000000000000000000000000000000000000000000

B 00110100110100110100110100110100110100110100110100

C 01000110110000010100111001011101110000000100100011010001010101

D 00100110110110001000111101010011101100100110000000101101010100

E 01010111011011110111001001101011011100110110100001101111011101

F 01110111110011011001101001000011111100110110000001101101010101

G 00000110001011100010000000010100001011010100000010000000100

H 01010011011110110111010101000001011110000001010111010101000101

Intuitive Randomness

Non-randomness: increasingly complex patterns.

A 00000000000000000000000000000000000000000000000000

B 00110100110100110100110100110100110100110100110101

C 01000110110000010100111001011101110000000100100011010001010101

D 00100110110110001000111101010011101100100110000000101101010100

E 01010111011011110111001001101011011100110110100001101111011101

F 01110111110011011001101001000011111100110110000001101101010101

G 00000110001011100010000000010100001011010100000010000000010000

H 01010011011110110111010101000001011110000001010111010101000101

Intuitive Randomness

Randomness: bits coming from atmospheric patterns.

A 00000000000000000000000000000000000000000000000000

B 00110100110100110100110100110100110100110100110100

C 01000110110000010100111001011101110000000100100011010001010101

D 00100110110110001000111101010011101100100110000000101101010100

E 01010111011011110111001001101011011100110110100001101111011101

F 01110111110011011001101001000011111100110110000001101101010101

G 00000110001011100010000000010100001011010100000010000000100

H 01010011011110110111010101000001011110000001010111010101000101

Intuitive Randomness

Partial Randomness: mixing random and nonrandom sequences.

A 00000000000000000000000000000000000000000000000000

B 00110100110100110100110100110100110100110100110101

C 01000110110000010100111001011101110000000100100011010001010 1

D 00100110110110001000111101010011101100100110000001011010100

E 01010111011011110111001001101011011100110110100001101111011 1

F 01110111110011011001101001000011111100110110000001101101010 1

G 00000110001011100010000000010100001011010100000100000000100

H 01010011011110110111010101000001011110000001010111010101000 1

Intuitive Randomness

Randomness relative to other measures: biased coins.

A 00000000000000000000000000000000000000000000000000

B 00110100110100110100110100110100110100110100110101

C 01000110110000010100111001011101110000000100100011010001010101

D 00100110110110001000111101010011101100100110000000101101010100

E 01010111011011110111001001101011011100110110100001101111011101

F 01110111110011011001101001000011111100110110000001101101010101

G 00000110001011100010000000010100001011010100000010000000010000

H 01010011011110110111010101000001011110000001010111010101000101

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).
- 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 effective martingale (betting) can make an infinite amount betting of the bits.

The statisticians approach

- von Mises, 1919. A random sequence should have as many 0's as 1's. But what about 1010101010101010.....
- von Mises idea: If you select a subsequence $\{a_{f(1)}, a_{f(2)}, \dots\}$ (e.g. $f(1) = 3, f(2) = 10, f(3) = 29,000$, so the 3rd, the 10th, the 29,000 th etc) then the number of 0's and 1's divided by the number of elements selected should end to $\frac{1}{2}$. (Law of Large Numbers)
- But what selection functions should be allowed?
- Church: computable selctions.
- Ville, 1939 showed no countable selection possible. Essentially not enough statistical tests.

Ville's Theorem

Theorem (Ville)

Given any countable collection of selection functions, there is a real passing every member of the test yet the number of zero's less than or equal to n in the $A \upharpoonright n$ (the first n bits of the real A) is always less than or equal to the number of 1's.

Martin-Löf

- Martin-Löf, 1966 suggests using shrinking effective null sets as representing effective tests. Basis of modern effective randomness theory.
- 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)

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

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

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)

The Coder's Approach

- Have a Turing maching $U(\tau) = \sigma$ is a U-desription of σ. The length of the shortest τ is the Kolmogorov Complexity of σ relative to U. $C_U(\sigma)$.
- There are universal machines in the sense that for all M, $C_U(\sigma) \leq_C^+ (\sigma) =_{\text{def}} K_M(\sigma) + d_m$.

reals

- From this point of view we should have all the initial segments of a real to be random.
- First try α, a real, is random iff for all n, $C(\alpha \upharpoonright n) \geq n - d$.
- By complexity oscillations no such real can exist. The reason as is that C lacks the intentional meaning of Komogorov complexity. the bits of τ encode the information of the bits of σ. Because C really uses $\tau + |\tau|$ as we know it halts there.

Prefix free complexity

- K is the same except we use prefix-free complexity (Think telephone numbers.) i.e. $U(\tau)$ halts implies $U(\tau')$ does not for all τ comparible (but not equal to) τ.
- (Levin, later Schnorr and Chaitin) Now define α is K-random if there is a c s.t.

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

And...

- They all give the same class of randoms!

Theorem (Schnorr)

A is Martin-Löf random iff A is K-random.

- Similar ideas using martingales were you bet on the nest bit. A is random iff no "effective" marrtingale succeeds in achieving infinite winnings betting on the bits of A.
- $f(\sigma) = \frac{f(\sigma 0)+f(\sigma 1)}{2}$. (fairness)
- Many variations depending of sensitivity of the tests. Implementations approximate the truth: ZIP, GZIP, RAR and other text compression programmes.
- Notice no claims about randomness "in nature" But very intersting question as to e.g. how much is needed for physics etc.
- Interesting experiments can be done. E.g. ants. (or children) (Reznikova and Yu, 1986)

and Complexity

- How hard is it to compute the solution?
- How many steps does the algorithm take?
- Two examples.

CD's

- Algebraic coding: (Hamming, 1950)
  (something to be coded)$\mapsto$ (longer with redundancies for decoding)
- e.g. parity check 1010010$\mapsto$10100101. can decide if there is likely a single error. ISBN etc.
- More complex decoding uses algebra (specifically algebra following a long line for Fermat's Last Theorem (Kummer, 1840's), and non-solving the quintic (Galois 1830))
- Like the Yellow Pages, instead of letting your fingers do the walking, algebra talks the talk. (mixed metaphor)
- CD's first takes 00000000...11111111 and amplifies to something of length 256, so there are 2^{256} many possible codewords, which are decoded, and they are in real time.

Beer Delivery

- Take a big map and plan a tour to cost the least amount. (Beer delivery, or Travelling Salesman Problem)
- Yet Beer delivery (TSP) for 256 cities is computationally impossible. No way known except to try all possibilities! ($P =?NP$) (Cook, 1970, Karp, 1972, Levin-who knows?) Called computational intractability
- Sometimes intractablity is good e.g. RSA and credit cards. if factorization was easy, modern banking would break down!

Some applications

- Using Chaos and randomness enable us to treat dynamical systems like the weather.
- Replace statistical tools by computational ones.
- Speeding up algorithms. E.g. supplying primes for things like RSA. (of course open if BPP=P)
- Phylogeny and language etc evolution (something of a dream).
- Understanding how levels of randomness relate to performance, etc.
- Differential geometry, reverse mathematics, Brownian motion, sampling randoms, etc. (AND misuses such as creationists!)

My work

- What is "random"? What level of randomness is necessary for applications.
- Suppose I have a source of weak randomness, how can I amplify this to get better randomness?
- 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? If a real is random does it have strong or weak computational power?

Randoms should be computationally weak

- We now know that there are two kinds of randoms, those which resemble Chaitin's $\Omega = \sum_\sigma 2^{-K(\sigma)}$ and more typical ones. (Specifically a theorem of Stephan in 2002.)

- There has been a lot of popular press about the "number of knowledge" etc, which is random, but has high computational power.

- We would theorize randoms would be stupid: computationally weak.

One example-from music

- Stupidity Tests
- There are two ways to convince someone you are stupid:
- The first are random as they pass the stupidity test as they are so smart that they know how to be stupid, the second really are stupid.
- That is, with sufficient randomness, randomness begins to resemble order. This is kind of remarkable. We are still tryintg to understand it.
- One of the following music examples is aleatoric (or chance) and the other is totally serial (based on a pattern). Which is which?

How Chaos Resembles Order

Highly random objects can resemble highly patterned ones.

A musical example.

Excerpt A: from *Music of Changes* by John Cage

Excerpt B: from *Structures for Two Pianos* by Pierre Boulez

Cage's piece is an example of aleatory music.

Boulez's piece is an example of total serialism.

Theorem (Stephan)

A random real can compute a DNC function (we say the real has PA degree) iff A computes the halting problem.

- f is DNC iff for all x, $f(x) \neq \varphi_x(x)$, and $f(x) \in \{0, 1\}$.
- If we remove the 0,1 restriction then the f is called fixed point free and any random can compute one.

Theorem (Barmpalias, Lewis, Ng)

Every PA degee is the join of two random segrees.

Halting probabilities

- One would think therefore that Ω has nothing to do with most randoms, but:

Theorem (Downey, Hirschfeldt, Miller, Nies)

Almost every random A is Ω^B for some B.

Theorem (Kurtz)

Almost every random A is computably enumerable relative to some $B <_T A$.

K-triviality

Theorem (Chaitin)

If $C(A \upharpoonright n) \leq^+ C(n)$ for all n, then A is computable.

- This is proven using the fact that a Π^0_1 class with a finite number of paths has computable paths, combined with the Counting Theorem $\{\sigma : C(\sigma) \leq C(n) + d \wedge |\sigma| = n\} \leq A2^d$. (The Loveland Technique)

K-triviality

Theorem (Chaitin)

If $C(A \upharpoonright n) \leq^+ C(n)$ for all n, then A is computable.

- This is proven using the fact that a Π^0_1 class with a finite number of paths has computable paths, combined with the Counting Theorem $\{\sigma : C(\sigma) \leq C(n) + d \wedge |\sigma| = n\} \leq A2^d$. (The Loveland Technique)
- What is $K(A \upharpoonright n) \leq^+ K(n)$ for all n? We call such reals K-trivial. Does A K-trivial imply A computable?

K-triviality

Theorem (Solovay)

There are noncomputable K-trivial reals.

- $A = \{\langle e, n\rangle : \exists s\,(W_{e,s} \cap A_s = \emptyset \wedge \langle e, n\rangle \in W_{e,s} and \sum_{\langle e,n\rangle \leq j \leq s} 2^{-K(j)[s]} < 2^{-(e+2)})\}$.

Theorem (Downey, Hirschfeldt, Nies, Stephan)

A is K-trivial and noncomputable implies $\emptyset <_T A <_T \emptyset'$, and hence they solve Post's problem.

Want to know more?

- My homepage : just type Rod Downey into google.
- and I am the one who is not the author of gay Lolita.
- Buy that wonderful book, $57 from Amazon at present.
- Buy some for your friends.

Thank You