Randomness is native to computer science

Serge Grigorieff
LIAFA, CNRS & Universite Paris 7
seg@liafa.jussieu.fr

Sommaire

Anyone who considers arithmetical methods of producing random reals is, of course, in a state of sin.

John von Neumann (1951)

Various techniques used in connection with random digits,

in *Monte Carlo Method*,

A.S. Householder, G.E. Forsythe, and H.H. Germond, eds.,

National Bureau of Standards Applied Mathematics Series, 12

(Washington, D.C.: U.S. Government Printing Office, 1951),

pages 36-38.

What is "true" randomness?

1 Two aspects of randomness

1.1 $\boxed{\textbf{Random objects and probabilities}}$

(reals, integers, strings,...)

random objects constitute the basic intuition for probabilities ... though they are not considered

- there is no formal notion ,of random object in probability theory
- *random variables* are formal objects which have nothing to do with random objects: they are just *measurable functions*

1.2 $\boxed{\textbf{Chaotic objects}}$

those of which one cannot know anything substantial

- *no regularity*
  a string is random if the sole way to describe it is to give its successive letters
- *"hereditary undecidability"*
  a random infinite sequence $a_0a_1a_2...$ has no computable subsequence $a_{k_0}a_{k_1}a_{k_2}...$ indexed by a computable infinite sequence $k_0 < k_1 < k_2 < ....$

2 Formalizing randomness: the case of finite objects

2.1 A paradoxical result

In probability theory,
if we toss an unbiaised coin 100 times then 100 heads are just as probable as any other outcome!

Who really believes that ?

> *The axioms of probability theory, as developped by Kolmogorov, do not solve all mysteries that they are somtimes supposed to.*
>
> *Peter Gàcs*

Probability theory, as axiomatized by Kolmogorov, 1933, **does not answer** the following questions:

- **what is a random string ?**
- **what is a random infinite sequence ?**

This quest for a formal notion of random object, in relation with this paradoxical result, goes back to Pierre Simon de Laplace (1749–1827).

2.2 Kolmogorov complexity

Even after his 1933 axiomatization of probability theory, Kolmogorov (1903–1987) never gave up the project to formalize the notion of random object.

> (By the way, it is really the unique probabilist – up to now – not to believe that Kolmogorov's axioms for probability theory do not constitute the last word about formalizing randomness...)

Descriptional complexity
built on the theory of computable functions

$$\text{independently by} \begin{cases} \text{Solomonoff} & (1962/1964) \\ \text{Kolmogorov} & (1963/1965) \\ \text{Chaitin} & (1964/1966) \end{cases}$$

Now called **Kolmogorov complexity**

Basic idea : $\boxed{\begin{array}{c} \textit{complexity of an objet} \\ = \\ \textit{length of the shortest descriptions} \end{array}}$

But, *description in which context?*
Care ! Berry's paradox (1908):
"the smallest integer which cannot be described by any sentence with less than twenty words"

$$\boxed{\begin{array}{c}\textit{complexity of an objet}\\ =\\ \textit{length of the shortest descriptions}\end{array}}$$

Key idea for Kolmogorov complexity:

> replace *"description"* by *"computation program"* in order to enter the formal framework of computability theory set up by Turing and Church.

Digression. Church's thesis insures that the notion of *computable function* has been formalized. However, the formalization yields a notion of *partially* (abusively called *partial computable function* or partial recursive function).

So, there are two theories :

- *the theory of computable functions*
- *the theory of partial computable functions*

The "right" theory, the one with a cornucopia of very nice results is that of partial computable functions.

Whence the more precise form if the key idea of Kolmogorov complexity

replace *"description"* by *"computation program" of a partial computable function*

Definition. *Let* $f : \{0,1\}^{<\omega} \to D$ *be a partial function (where* $D = \mathbb{N}$ *or* $D = \{0,1\}^{<\omega}$ *or...)*
Set, for $x \in D$,

$$K_f : D \to \mathbb{N} \quad , \quad K_f(x) = \min\{|p| : f(p) = x\}$$

$$\text{Intuition :} \begin{cases} p \text{ is a program} \\ f \text{ executes programs} \\ (f = \text{programming language} \\ \qquad + \text{compiler} + \text{computer}) \\ \text{f(p) is the output of program } p \end{cases}$$

for $x \in D$, $K_f(x)$ *is the length of shortest programs* p *with which* f *computes* x *(i.e.* $f(p) = x$*)*

Invariance theorem. *Among the* K_f*'s, where* f *varies in the family* PC^D *of partial computable function* $\{0,1\}^{<\omega} \to D$*, there is a smallest one, up to an additive constant (i.e. within some bounded interval) :*

$$\exists \varphi \in PC^D \, \forall f \in PC^D \, \exists c \, \forall x \in D \; K_\varphi(x) \leq K_f(x) + c$$

Such a φ is called *optimal.*

The proof is a simple application of the enumeration theorem for partial computable functions.

Definition. *$K^D : D \to \mathbb{N}$ is K_φ where φ is some fixed optimal function.*
K^D is therefore minimum among the K_f's, up to an additive constant.
K^D is defined up to an additive constant: if φ et ψ are both minimum then

$$\exists c \; \forall x \in D \; |K_\varphi(x) - K_\psi(x)| \leq c$$

An integer defined up to a constant...!
The constant is uniform in $x \in D$.
What Kolmogorov said about the constant:

> *Of course, one can avoid the indeterminacies associated with the [above] constants, by considering particular [...functions φ], but it is doubtful that this can be done without explicit arbitrariness.*
> *One must, however, suppose that the different "reasonable" [above optimal functions] will lead to "complexity estimates" that will converge on hundreds of bits instead of tens of thousands.*
> *Hence, such quantities as the "complexity" of the text of "War and Peace" can be assumed to be defined with what amounts to uniqueness.*

2.3 Measure of randomness of a string via Kolmogorov complexity

$K \leq K_I d + O(1)$ hence $K(x) \leq |x| + O(1)$ for every string x.

Word c-random if $K(x) \geq |x| - c$.

At most $2^{n-c} - 1$ programs of length $< n - c$ and 2^n strings of length n.

Thus, the proportion of strings of length n which are c-random is $\geq 2^{-c}$.

No absolute notion of random string but a measure of randomness.

$\boxed{\text{Random strings pass all statistical tests}}$

(Martin-Löf, 1965)

3 Formalizing randomness: the case of infinite sequences

3.1 von Mises top-down approach to random infinite sequences

Richard von Mises (1883–1953) pionneered the quest for a mathematical notion of random infinite sequences.

von Mises' idea. *A binary infinite sequence* $a_0 a_1 a_2 ...$ *is random if any "admissible" extracted subsequence has as many 0's than 1's, i.e.*

$$\lim_{n \to +\infty} \frac{a_{k_0} + a_{k_1} + a_{k_2} + ... + a_{k_n}}{n} = \frac{1}{2}$$

for any "admissible" extracted subsequence $a_{k_0} a_{k_1} a_{k_2} ...$ *(with* $k_0 < k_1 < k_2 < ...$*).*

Non clarified : what does "admissible" mean?

Church, 1939, proposes *admissible* = *computable.*

However, ***this is not the right notion:***

Theorem (Jean Ville, 1939). *there are sequences* $a_0a_1a_2...$ *which are von Mises & Church random and satisfy*

$$\forall n \quad a_0 + a_1 + a_2 + ... + a_n \geq \frac{n}{2}$$

Catastrophe : it is known that, with probability 1, an infinite sequence $\alpha_0\alpha_1\alpha_2...$ is such that

$$\alpha_0 + \alpha_1 + \alpha_2 + ... + \alpha_n - \frac{n}{2}$$

changes sign infinitely often.

3.2 Martin-Löf top-down approach to random infinite sequences

The naive approach on which it is based is as follows:

> $\alpha \in \{0,1\}^{\mathbb{N}}$ *is a random infinite sequence if it avoids every set of measure zero.*

Problem: any singleton set $\{\alpha\}$ has measure zero.

Martin-Löf's idea:

$\boxed{\textit{partially effectivize measure zero sets}}$

As we know,

- $\mathcal{X} \subseteq \{0,1\}^{\mathbb{N}}$ has measure 0 if and only if there exists a sequence of open sets $(\mathcal{V}_n)_{n\in\mathbb{N}}$ such that

$$\mathcal{X} \subseteq \mathcal{V}_n \subseteq \{0,1\}^{\mathbb{N}} \quad , \quad \lim_{n\to+\infty} \mu(\mathcal{V}_n) = 0$$

  i.e. $\mathcal{X} \subseteq \bigcap_{n\in\mathbb{N}} \mathcal{V}_n$ where $\bigcap_{n\in\mathbb{N}} \mathcal{V}_n$ is a measure zero G_δ set.

- $\mathcal{V}_n$ is open if and only if $\mathcal{V}_n = X_n\{0,1\}^{\mathbb{N}}$ for some $X_n \subseteq \{0,1\}^{<\omega}$

Martin-Löf defines *constructively measure zero G_δ set* as follows:

- $\{(n,u) \mid u \in X_n\}$ *is computably enumerable*
- $\mu(\mathcal{V}_n) \leq 2^{-n}$

Definition.
α ***is Martin-Löf random if it avoids all constructively measure zero G_δ set***

Clearly,

> *The set of Martin-Löf random sequences has measure* 1

Facts.

- **There exists a LARGEST constructively measure zero G_δ set**

  Thus,

  $$\boxed{\begin{array}{c}\textbf{Martin-Löf random}\\ =\\ \textbf{Out of this largest set}\end{array}}$$

- **Martin-Löf random sequences satisfy all usual probability laws**
  - law of large numbers,
  - law of iterated logarithm
  ...

  Why?. Classical proofs of these results are effective in the following sense:
  for all n one can get a sequence of words $(u_{n,i})_{i\in\mathbb{N}}$ such that

  1. *the sequence $(u_{n,i})_{n\in\mathbb{N},i\in\mathbb{N}}$ is computable*
  2. *the set of exceptions to the law is included in $\bigcup_{i\in\mathbb{N}} u_{n,i}\{0,1\}^{\mathbb{N}}$*
  3. $\mu(\bigcup_{i\in\mathbb{N}} u_{n,i}\{0,1\}^{\mathbb{N}}) \leq 2^{-n}$

All the theory relativizes :

- **1-randomness** : Martin-Löf randomness
- **2-randomness** : use oracle $\emptyset'$
- **3-randomness** : use oracle $\emptyset''$
- ...
- **and worse...**

Digression.

$\boxed{\textbf{Computable analysis}}$

versus

$\boxed{\textbf{Computably enumerable analysis}}$

An alien world
rich of "newlook" very nice results

- **There exists a LARGEST constructively measure zero G_δ set**

Another example (Levin, 1973) :

- **Among series of real numbers $(r_n)_{n\in\mathbb{N}}$ which are computably enumerable from below (i.e. $\{(n,q) \in \mathbb{N} \times \mathbb{Q} \mid q < r_n\}$ is computably enumerable) there exists a BIGGEST one, up to a multiplicative factor.**

3.3 The bottom-up approach to random infinite sequences

Get randomness on infinite strings from randomness on finite strings

This was Kolmogorov first approach: ask that $K(\alpha \restriction n) = n + O(1)$ where $\alpha \restriction n = \alpha(0)\alpha(1)...\alpha(n-1)$

Alas...

Large oscillations theorem. [Martin-Löf]
For any α there exists infinitely many n's such that $K(\alpha \restriction n) \leq n - \log(n)$

In fact, if $\sum_{n\in\mathbb{N}} 2^{-f(n)} = +\infty$ then there exists infinitely many n's such that $K(\alpha \restriction n) \leq n - f(n)$

However, this bottom-up approach is possible with modified versions of Kolmogorov complexity.

Claus Peter **Schnorr's process complexity**

consider partial computable $f : \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ which are *monotone*:

$$p \leq_{prefix} q \wedge p, q \in domain(f) \ \Rightarrow \ f(p) \leq_{prefix} f(q)$$

The invariance theorem still holds: among the K_f's, where f varies over monotone partial computable functions, there is a minimum one.
(Recall $K_f(x) = \min\{|p| : f(p) = x\}$)
Let's denote by S the smallest such K_f

Theorem. [Schnorr] *α is Martin-Löf random if and only if $\exists c\ \forall n\ |S(\alpha \upharpoonright n) - n| \leq c$.*
In other words, $S(\alpha \upharpoonright n) = n + O(1)$.

Levin, Chaitin prefix-free complexity

Consider partial computable $f : \{0,1\}^{<\omega} \to \{0,1\}^{<\omega}$ which have *prefix-free domains*:
The invariance theorem still holds: among the K_f's, for such f's, there is a minimum one.
Let's denote by H the smallest such K_f

Theorem. [Schnorr] *α is Martin-Löf random if and only if $\exists c\ \forall n\ H(\alpha \upharpoonright n) \geq n - c$*
if and only if $\exists c\ \forall n\ \lim_{n\to+\infty} H(\alpha \upharpoonright n) - n = +\infty$.

3.4 Chaotic sequences

uncomputable set = complex subset of a simple set

But this subset may contain subsubsets which are computable

(Recall : every infinite computably enumerable set does contain an infinite computable set)

However there exists complement of computable sets which contain no infinite computable set.

Get worse : $X \subset \mathbb{N}$ such that if A is an infinite coinfinite computable subset of X then $X \cap A$ is not computable with oracle $X \setminus A$.

Random sequences are like this.

Theorem (van Lambalgen). *If* $(a_0, a_1, a_2, ...)$ *est random then* $(a_0, a_2, a_4, ...)$ *is random relative to oracle* $(a_1, a_3, a_5, ...)$.

3.5 Randomness is quite fragile

No stability under computable transformation.

If $a_0a_1a_2...$ is random then $0a_00a_10a_20...$ IS NOT random

4 Explicit random sequences

Intuition : start with known non computable results.

Up to now, all random sequences come as digits of a real which is a probability

Sole known way to diffuse non computability

4.1 Chaitin's real Ω

An avatar of the halting problem.

Consider a universal Turing machine $\mathcal{U}$ and let it run on an infinite input string.

Ω is the probability that $\mathcal{U}$ on input $\alpha \in \{0,1\}^{\mathbb{N}}$ halts in finite time

Theorem (Chaitin, 1975).
Ω is Martin-Löf random.