Randomness as Pointlessness

Greg Langmead

December 7, 2020

Random points satisfy the following

1. They cannot be obtained constructively.

2. They obey almost-everywhere properties of classes of functions.

3. They appear tied to the foundations of mathematics.

For a survey of these and other properties, see Rute [1]. I will argue that defining random points in terms of locales is well suited to capturing or even explaining these properties.

1   Randomness vs determinism: a confused debate

In the Stanford Encyclopedia of Philosophy's article on Chance and Randomness [2] section 5.2 ("Chaotic Dynamics"), the authors describe an iterated transformation on the unit square. Given a starting point $(p, q), 0 \leq p, q \leq 1$, define a function

$$\phi(p, q) = \begin{cases} (2p, q/2), & \text{if } 0 \leq p \leq \frac{1}{2} \\ (2p - 1, (1 + q)/2), & \text{if } \frac{1}{2} \leq p \leq 1 \end{cases}$$

As the authors say, "This corresponds to transforming the unit square to a rectangle twice as wide and half the height, chopping off the right half, and stacking it back on top to fill the unit square again." If we represent p and q in their binary expansion, we have

$$\phi(0.p_1 p_2 \ldots, 0.q_1 q_2 \ldots) = (0.p_2 p_3 \ldots, 0.p_1 q_1 q_2 \ldots).$$

This example is intended to convey a paradox. On the one hand, the system is ergodic and behaves like a random process. Points will wander throughout the space without any pattern. But on the other hand the system is deterministic. The bit shift formula shows how it works: this is an engine that uses the next bit of the binary expansion to determine where the point moves next.

The resolution of the paradox, I believe, is to worry more about whether you can get your hands on p and q in the first place. Sure, once you have them, the machine behaves deterministically. But you can't have them! They require an infinite amount of information

to specify, so if you have them you have just recapitulated the future history of the machine. The problem with the paradox is therefore in one of its assumptions. The process is deterministic, but the randomness is in the input data.

Indeed, the most important property of random numbers appears to be their infinite nature, which prevents us from ever having one fully specified for us, either in a machine, in our minds, or in our theories. They are not computable, and they are not constructible.

However, there are several different definitions of algorithmic randomness. In fact there is a whole family (or less charitably, zoo) of them. They go by names such as Martin-Löf randomness, Schnorr randomness, weak-n-randomness, and computable randomness. They are not equivalent but they fit together in a comprehensible way (mostly through containment). But what their original mid-to-late 20th century definitions share is that they are dichotomies: each definition divides the interval $[0-1]$ into two sets: the randoms, and the non-randoms. What I will present here is a consideration that is orthogonal to the choice of randomness, and this is the *packaging* of the random numbers inside the reals. The main ingredient is a shift from *set theory* to *locale theory*. A locale is a generalization of a topological space, and I will argue that it is very well suited to a philosophical interpretation of randomness.

In [3], Alex Simpson has formulated a measure theory for locales and provided a new definition of randomness in this setting. We will describe this work, emphasizing the *point-free* interpretation. The sublocale of randoms has no points!

The beauty of the locale formulation of measure theory is that it can construct an environment for discussing randomness that highlights the inability to access individual random numbers. It does this because it *restricts* us from constructing certain measurable sets?? We will adopt the point of view of *pointless topology*. In a locale, we frame our theory not in terms of the points of a set, but in terms of the open subsets of a space. It is then extra data to produce a point of a locale. The central result of this paper is that the random sub-locale of a locale has no points.

2   What it means to have points

Locale theory is point-free topology. See [4] for an accessible survey. It is *synthetic* in that it takes open sets as basic, undefined objects, instead of defining them as collections of points. Instead it is points that will become derived. Topological spaces with their actual open sets of actual points are special cases of locales, but the enlarged category of locales has a few advantages. First of all it lets us argue constructively. Secondly it isolates just the axioms we need to prove many theorems from topology. Thirdly it will give us a unique point of view on randomness.

A *frame* is a partially ordered set with finite meets $\wedge$ and arbitrary joins $\vee$, with a maximal element $\top$ and minimal element $\bot$, and where the distributive law holds:

$$x \wedge (\bigvee_i y_i) = \bigvee_i (x \wedge y_i).$$

If there are only countable joins then it is called a σ-frame. The open sets of a topological space form a frame. A frame map is a function that preserves finite meets and arbitrary

joins. A σ-frame map preserves countable joins. We collect this data into the category Frm of frames and σFrm of σ-frames.

A continuous map between topological spaces induces a frame map but in the opposite direction because the *inverse* image of an open is open. This motivates the definition of the category Loc of locales to be simply Frm$^{\mathsf{op}}$ and σLoc $= \sigma$Frm$^{\mathsf{op}}$. We can define a functor $\mathcal{O}$: Top $\to$ Loc on objects by mapping a space X to its frame of open sets $\mathcal{O}(X)$, and on morphisms by mapping a function $f : X \to Y$ to the opposite arrow of the induced morphism on frames, so that's two inversions giving us an arrow $\mathcal{O}(f) : o(X) \to \mathcal{O}(Y)$. This functor has an important right adjoint which we will discuss next.

For the topological space $\star$ consisting of a single point, its topology is just $\{\top \le \bot\}$. To convert points into a derived concept then, we can note that a point in a space X is equivalent to a map pt : $\star \to X$, which at the level of frames is a map $\mathcal{O}(X) \to \{\top, \bot\}$. The inverse image of $\top$ is the collection of all open sets that contain the point pt($\star$). It forms what is known as a *completely prime filter* in the frame, which is a collection P of opens such that if $x \vee y \in P$ then $x \in P$ or $y \in P$. (This is analogous to a prime ideal, or the ideal generated by a prime number.) We can then move forward with locales and completely prime filters instead of spaces and points.

Given a locale L, the set of completely prime filters constitutes the *spatial part* of L, and this set inherits a topology from the structure of L. Locale maps create an association between the completely prime filters, and so we obtain a set function between the spatial parts of the locales. This is a functor S : Loc $\to$ Top and is right adjoint to $\mathcal{O}$. If we go back and forth with $S \circ \mathcal{O}$, starting from an honest topological space, then we get back an equivalent space with an equivalent topology. This is the categorical situation whereby locales are generalized spaces.

So given an arbitrary locale, it may or may not have points. And given some underlying set such as the real interval $[0, 1]$, there may be a locale structure other than the standard topology, and this locale structure may have fewer points than the set $[0, 1]$! We will go there now, by defining sublocales.

There are a few equivalent ways to define sublocales. To me the most intuitive is to think of them as *embedded subspaces*, i.e. monomorphic maps with the extra requirement that the subspace is embedded. In the category of topological spaces this means the domain is homeomorphic to its image. For example, for a map from $[0, 1]$ into a space X, to be an embedding the curve must not self-intersect or do anything pathological. To say this cleanly in terms of open sets, we say that $f : X \to Y$ is an embedding if $f^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$ is surjective. In categorical terms these are *regular monomorphisms*. That's the spatial definition. But in the category of frames the arrows are reversed and so a sublocale is in fact a quotient, i.e. an equivalence relation. Simpson focuses on this definition (Definition 3.6 in [3]). The equivalence relation is given by a subcollection $Y \subset \mathcal{O}(X)$ of opens, together with intersecting

3 σ-locales and randomness

Define σ-locale.

Define σ-sublocale.

Define measures on these.

The payoff: there is a way to bring measure theory into the pointless paradigm such that the random part has no points.

How Kurtz randomness fails the LLN (see [5], Section 3.5).

The prospect for importing all the various randomness notions into this framework.

aper: locale theory - Ran, the sublocale of random points - presumably we can say some class of functions are smooth on this sublocale? - Porter provides some justification for the locale approach since it's "all measure 1 properties" - are there countable bounds on the collections, in the locale theory? - given some notion of randomness, the interval [0,1] can be given the corresponding locale structure - does this mean that each notion of randomness has its own topology? or are the topologies the same just with different numbers of points, and different Ran? (must be a different topology, right? because once you have a point you have an ideal of open sets, else not) - what corresponds to the different amounts of computability and/or martingale bounds? - will the punch line be that the locale approach is agnostic to which definition of randomness we work with, but simply takes any of the notions and removes random points from the space? - how does the existence of a universal ML test play out in locales? - the a.e. results like "if f is nondecreasing then it is smooth at every computable random" feel like statements about cohesion, and this feels natural in the locale interpretation - she is saying that Birkhoff Ergodicity theorem, suitably effectivized, holds at ML randoms, i.e. the same set that functions of bounded variation are smooth at – making them "equally strong". But in another sense these are different topologies on R and so we can also say that given this topology these are the classes of functions. - Church's thesis that there are no non-computable reals – has a natural home here.

References

[1] Jason Rute. On the close interaction between algorithmic randomness and constructive/computable measure theory, 2019.

[2] Antony Eagle. Chance versus Randomness. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

[3] Alex Simpson. Measure, randomness and sublocales. *Annals of Pure and Applied Logic*, 163(11):1642 – 1659, 2012. Kurt Goedel Research Prize Fellowships 2010.

[4] Peter T. Johnstone. The point of pointless topology. *Bull. Amer. Math. Soc. (N.S.)*, 8(1):41–53, 01 1983.

[5] André Nies. *Computability and Randomness*. Oxford Logic Guides. Oxford University Press, Oxford, 2009.