

Randomness as Pointlessness

Greg Langmead

December 10, 2020

1 Introduction

Random real numbers satisfy the following hand-waving characterizations.

1. They cannot be obtained constructively.
2. They obey almost-everywhere properties quantified over 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 all three of these properties.

2 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, somewhat analogous to repeatedly folding dough. 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.” So, needlessly more violent than folding, but close. If we represent p and q in their binary expansion, we have

$$\phi(0.p_1p_2\dots, 0.q_1q_2\dots) = (0.p_2p_3\dots, 0.p_1q_1q_2\dots).$$

In the article 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 if you’re in a teasing mood, a 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 form a simple chain of inclusions, and the differences can be logically explained and outsourced to a family of choices or parameters that connect them. What I will present here is a consideration that is orthogonal to the flavor 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 beautiful punch line will be that the sublocale of randoms has no points! This is a wonderful result that feels tailor-made for discussing randomness. Locales are a generalization of

spaces with just where we can define continuity, computability (after passing to toposes of sheaves, which naturally live over locales).

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.

3 Frames and Locales

Locale theory is point-free topology. See [4] for an accessible survey, or Chapter IX of Mac Lane and Moerdijk [5] for a pleasantly breezy but more thorough introduction. 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 \perp , and where the infinite distributive law holds:

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

These are clearly snagged from the axioms for a topology. 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 σ **Frm** of σ -frames.

A continuous map between topological spaces induces a frame map on the frames of open sets, 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**^{op} and σ **Loc** = σ **Frm**^{op}. (Recall that the opposite

of a category is a category with all the same objects and morphisms, but with the direction of all the morphism arrows reversed.) We can define a functor $\mathcal{O} : \mathbf{Top} \rightarrow \mathbf{Loc}$ on objects by mapping a space X to its frame of open sets $\mathcal{O}(X)$, and on morphisms by mapping a function $f : X \rightarrow Y$ to the opposite arrow of the induced morphism on frames, so that's two inversions giving us an arrow $\mathcal{O}(f) : \mathcal{O}(X) \rightarrow \mathcal{O}(Y)$. 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 $\{\perp \leq \top\}$. To convert points into a derived concept then, we can note that a point in a space X is equivalent to a map $\text{pt} : \star \rightarrow X$, which at the level of frames is a map $\mathcal{O}(X) \rightarrow \{\top, \perp\}$. The inverse image of \top is the collection of all open sets that contain the point $\text{pt}(\star)$. 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 $S : \mathbf{Loc} \rightarrow \mathbf{Top}$ 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 so long as the original space is *sober*, which is a definition that was new to me, but is almost circularly defined to meet this condition. Hausdorff spaces are sober. This is the categorical situation whereby locales are generalized spaces, and being spatial, and having points, are derived special properties.

Note how compatible this is with a constructive point of view. To supply a point you must supply a function $\star \rightarrow X$. The points aren't given, you have to do work to produce one.

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 $[0, 1]$! To give a hint of how this may work, imagine deleting some of the open sets, so that the prime filters are not complete anymore. That's the approach we will take when we define certain *sublocales*.

4 Sublocales

There are three equivalent ways to define sublocales. Normally I'd expect two since we are bouncing between two opposite categories. But there is a super interesting third definition as well, and the interplay of all three sheds light on the application to randomness.

4.1 Definition 1: Embedded subspaces

To me the most intuitive definition is the one that lets one think of a sublocale as an *embedded subspace*, i.e. a monomorphic map with the extra requirement that the subspace is embedded. In the category of topological spaces and other concrete categories 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. In categorical terms the sublocales are *regular monomorphisms*.

4.2 Definition 2: Quotients of frames

In the category of frames the arrows are reversed and so a sublocale is in fact a quotient, i.e. an equivalence relation. Intuitively, two open sets are equivalent if they agree on the sublocale. In more detail, a sublocale $=_Y$ is an equivalence relation that preserves finite meets and arbitrary joins:

$$U =_Y V \text{ and } U' =_Y V' \implies U \cap U' =_Y V \cap V' \quad (1)$$

$$U_i =_Y V_i \text{ for all } i \geq 0 \implies \bigcup_{i \geq 0} U_i =_Y \bigcup_{i \geq 0} V_i. \quad (2)$$

4.3 Definition 3: Nuclei

The nucleus definition of a sublocale also takes place in the category of frames. A nucleus is a function $j : \mathcal{O}(X) \rightarrow \mathcal{O}(X)$ satisfying

$$U \leq jU \quad (3)$$

$$jjU = jU \quad (4)$$

$$j(U \wedge V) = j(U) \wedge j(V) \quad (5)$$

The intuition is that j takes an open set to some largest version of itself satisfying some property, which is why the first condition has the containment, and why the second condition is idempotence – doing the operation

twice yields the same largest open. The third condition is that forming the “largest version” preserves finite intersection.

Denote the fixed points of j by $\mathcal{O}_j(X)$ and the dual object, i.e. its locale, by X_j . The map j is a surjective frame map onto the set of fixed points, and so gives an embedding $i : X_j \hookrightarrow X$.

4.4 Warmup example: a sublocale with no points

Here’s something that warrants some reflection: let’s define a sublocale with no points. Let X be a Hausdorff topological space without isolated points and let $\mathcal{O}(X)$ be the locale of its open sets. Consider the double-negation operation $\neg\neg : \mathcal{O}(X) \rightarrow \mathcal{O}(X)$. This defines a nucleus hence an embedding $nn : X_{\neg\neg} \hookrightarrow X$.

This sublocale $\neg\neg X$ has no points! For suppose $p : \star \rightarrow X_{\neg\neg}$ is a point. The composition $nn \circ p$ is a point of X itself. Consider the open set given by deleting this point from X , i.e. $X - \{nn(p(\star))\}$. Mapping this open set backwards to $X_{\neg\neg}$ and then to \star , we have that $p^{-1}(\neg\neg(X - \{nn(p(\star))\})) = \perp$. It equals \perp because only opens that contain the image of \star give \top . But X has no isolated points and so $\neg\neg(X - \{nn(p(\star))\}) = X$, so the above expression $p^{-1}(\neg\neg(X - \{nn(p(\star))\}))$ becomes $p^{-1}(X)$ which is \top , a contradiction.

5 σ -versions

6 Measures thereon

The collection of all sublocales of a fixed locale X are ordered by inclusion and form a frame of their own which we denote by $\mathcal{S}(X)$. This is one level up so to speak, since X itself is (the opposite of) a frame of open sets, and the sublocales are formed from the whole of this frame by way of a quotient operation.

We will define a measure on a locale/frame and then extend it to the frame of sublocales of the frame.

A measure on a frame is a function $\mu : F \rightarrow [0, \infty]$ satisfying

$$\mu(\perp) = 0$$

$$x \leq y \implies \mu(x) \leq \mu(y)$$

$$\mu(x) + \mu(y) = \mu(x \vee y) + \mu(x \wedge y) \quad (\text{"modularity"})$$

$$(x_i)_{i \geq 0} \text{ ascending} \implies \mu\left(\bigvee_{i \geq 0} x_i\right) = \sup_{i \geq 0} \mu(x_i) \quad (\text{"continuity"}).$$

We say μ is a probability measure if $\mu(\top) = 1$. We will also call μ a measure on a locale if it is a measure on the corresponding frame of opens.

To bring the concept of measure to sublocales we will generalize the concept of outer measure. We say a locale is *fitted* if for every sublocale $Y \subseteq X$ it holds that $Y = \cap \mathcal{N}(Y)$ where $\mathcal{N}(Y) = \{U \in \mathcal{O}(X) | Y \subseteq U\}$. We extend the measure μ to a measure μ^* on any sublocale by $\mu^*(Y) = \inf_{U \in \mathcal{N}(Y)} \mu(U)$, i.e. the limit of the measure of the open sets that intersect to Y . There is work to show that spaces we care about are fitted, and that this definition really defines a measure (see section 5 of [3]), but I will skip all that.

7 The random sublocale

We are ready to define the random sublocale Ran_X of an arbitrary locale X equipped with a probability measure μ . We'll do it three times, once for each definition of sublocale.

The spatial definition of Ran_X is the largest sublocale of measure 1, or equivalently the union of all sublocales of measure 1. This tracks exactly with the definition of Kurtz randomness seen elsewhere in the literature.

The quotient definition of Ran_X is the equivalence relation $=_{\text{Ran}_X}$ given by $U =_{\text{Ran}_X} V \iff \mu(U) = \mu(U \cap V) = \mu(V)$. Opens are equivalent if they cover the same amount of measure.

The nucleus definition is that $j_{\text{Ran}_X}(U) = \bigcup_{V \supseteq U, \mu(U) = \mu(V)} V$. In other words, the maximal open that contains U and with the same measure as U . This is my educated guess, I haven't seen this in Simpson's writings or elsewhere as of yet.

This sublocale has no points.

8 Kurtz randoms are not so bad

How Kurtz randomness fails the LLN (see [6], Section 3.5).

The prospect for importing all the various randomness notions into this framework. The \mathcal{S} functor.

←
So
prove
it,
wimp.

9 But actually we can have the other kinds of randoms in locales too

This doesn't appear to be thought through carefully yet in published work but it appears in sketch form in slides by Simpson [?]. He defines the Borel hierarchy in the context of locales, giving him definitions for Σ_n^0 and Π_n^0 . Using these you can proceed with the definition of pointless sublocales of randoms. If every definition of randomness can be imported into this setting then we have a clean story where locales are a lens through which to look at the entire landscape of computable randomness. If, however, there are limitations to what can be expressed in locale theory, then the picture will get cloudy, but perhaps being compatible with locale theory can join the other popular metrics when assessing the tradeoffs between members of the randomness family.

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 \mathbb{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] Saunders MacLane and Ieke Moerdijk. *Sheaves in geometry and logic: a first introduction to topos theory*. Universitext. Springer, Berlin, 1992.
- [6] André Nies. *Computability and Randomness*. Oxford Logic Guides. Oxford University Press, Oxford, 2009.