

# Randomness as Pointlessness

Greg Langmead

December 14, 2020

## 1 Introduction

Random real numbers satisfy the following hand-waving characterizations.

1. They cannot be obtained constructively.
2. They obey almost-everywhere properties.
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, and to have one is to provide its path through the process. 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 random real numbers has no points! This is a wonderful result that feels tailor-made for discussing randomness.

### 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, or Johnstone [6] chapter C1 for a detailed treatment. It is *synthetic* topology 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 isolates just the axioms we need to prove many theorems from topology. This 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 abstracted from the axioms for a topology. If we require only countable joins then it is called a  $\sigma$ -frame. The open sets of a topological space form a frame (or

a  $\sigma$ -frame). A frame map is a function that preserves finite meets and arbitrary joins. A  $\sigma$ -frame map preserves countable joins. We collect this data into the category  $\mathbf{Frm}$  of frames and  $\sigma\mathbf{Frm}$  of  $\sigma$ -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  $\mathbf{Loc}$  of locales to be simply  $\mathbf{Frm}^{\text{op}}$  and  $\sigma\mathbf{Loc} = \sigma\mathbf{Frm}^{\text{op}}$ . (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  $L : \mathbf{Top} \rightarrow \mathbf{Loc}$  on objects by mapping a space  $X$  to its frame of open sets  $\mathcal{O}(X)$  (so on objects,  $L = \mathcal{O}$ ), 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  $L(f) : L(X) \rightarrow L(Y)$ .

## 4 Points

For the topological space  $\star$  consisting of a single point, its topology is just the poset  $\{\perp \leq \top\}$ . 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  $L$ . If we go back and forth with  $S \circ L$ , 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 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.

## 5 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.

## 5.1 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*.

## 5.2 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:

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

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

## 5.3 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  $i : X_j \hookrightarrow X$ . This is basically performing the quotient from the previous definition.

## 5.4 The collection of sublocales

The collection of all sublocales of a fixed locale  $X$  is partially ordered by inclusion, just as we have in any category where we can define subobjects. It’s worth mentioning that even more is true: the collection of sublocales forms a *coframe*, which is a poset with finite joins and arbitrary meets, and where the joins distribute over the meets. This collection is nowhere near as straightforward to contemplate as the collection of subsets of a set, but we will keep thinking through the consequences to find how it connects with randomness.

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 frame of its open sets. Consider the double-negation operation  $\neg\neg : \mathcal{O}(X) \rightarrow \mathcal{O}(X)$ . What does this do? The negation of an open  $U$  is the largest open disjoint from  $U$ . In  $X$  this is the same as the interior of the complement.  $\neg\neg U$  is thus the largest open disjoint from the largest open disjoint from  $U$ . For any  $p \in X$  we have that  $\neg\neg(X - \{p\}) = \emptyset$ .

$\neg\neg$  is idempotent, and in fact defines a nucleus hence an embedding  $nn : X_{\neg\neg} \hookrightarrow X$ , where  $nn^{-1}U = \neg\neg U$ . 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 thus a point of  $X$ . Consider the open set given by deleting this point from  $X$ , i.e.  $X - \{nn(p(\star))\}$ . Taking the inverse image we have  $p^{-1}(\neg\neg(X - \{nn(p(\star))\})) = \perp$  because only opens that contain the image of  $\star$  give  $\top$ . But as we saw in the previous paragraph  $\neg\neg(X - \{nn(p(\star))\}) = X$ , so the above expression  $p^{-1}(\neg\neg(X - \{nn(p(\star))\}))$  becomes  $p^{-1}(X)$  which is  $\top$ , a contradiction.

Furthermore,  $\neg\neg X$  is dense in  $X$ . For a sublocale  $A \rightarrow X$  to be dense means that  $A$  intersects every open sublocale. (An open sublocale is one of the original open sets in  $X$ , promoted to a sublocale by embedding it in  $X$ .) This follows from being a nucleus (which we did not prove):  $\neg\neg A \supseteq A$ .

What makes the category of locales different from the category of spaces is that there is a *largest dense sublocale* of any locale  $X$  and it is exactly  $\neg\neg X$ . And by largest I mean in terms of the inclusion relation on sublocales. One consequence of having a largest dense sublocale is that any two dense sublocales must intersect! This is absolutely not the case with topological spaces such as  $\mathbb{R}$  which can be decomposed into disjoint dense subsets such as  $\mathbb{Q}$  and  $\mathbb{R} \setminus \mathbb{Q}$ . If we promote these two subsets to sublocales, they will intersect!

## 5.5 Locale morphisms

Embedding a locale as a sublocale of another locale is a more exotic operation than set injection. We can directly compare these morphisms by starting with an embedding of topological spaces  $f : X \hookrightarrow Y$  and asking what the corresponding sublocale morphism is. The answer is given by a frame map called the push-forward  $f_* : \mathcal{O}(X) \rightarrow \mathcal{O}(Y)$ , which is right adjoint to the frame map  $f^{-1} : \mathcal{O}(Y) \rightarrow \mathcal{O}(X)$ . It is defined as follows:

$$f_*(U) = \bigcup_{f^{-1}(V) \subseteq U} V.$$

In other words, an open set  $U \subseteq X$  maps under  $f_*$  to the *largest* open set in  $Y$  that pulls back to some subset of  $U$ .

This is an unusual operation to contemplate. Here's an example that helps me: imagine  $f : [0, 1] \rightarrow \mathbb{R}^2$  is some smooth curve in the plane. Given an open interval  $U \subseteq [0, 1]$ , its image in the plane is a one-dimensional subset, so won't itself be open. Instead to compute  $f_*(U)$  we have to find the largest open set in the plane whose inverse image is  $U$ , which will look like *all of*  $\mathbb{R}^2$  minus two skinny pieces of the curve that aren't in the image of  $U$ .

Let's consider the inclusion  $i : \mathbb{Q} \hookrightarrow \mathbb{R}$ . The push-forward map  $i_*$  sends the open set "all of  $\mathbb{Q}$ " to the largest open in  $\mathbb{R}$  whose inverse image is contained in  $\mathbb{Q}$ . But  $\mathbb{Q}$  is dense so this is all of  $\mathbb{R}$ , there's no smaller open that contains the rationals! The requirement that

this push-forward be an open set is forbidding us from carving out these pathological dense subsets in the category of locales.

The requirement that  $f_*$  give the *largest* open set such that some condition holds should remind us of a nucleus. In fact  $f_*$  is a nucleus that defines the sublocale  $f : \text{Loc}(X) \rightarrow \text{Loc}(Y)$ , the “promotion” of the subspace inclusion to a sublocale.

## 6 Measures thereon

We have prepared all the abstract background. We have a new setting for topology called locale theory, which has broken away from the concept of a point. We have seen that the monomorphisms are different than in the category of topological spaces, and we have examples of sublocales that have no points. Let’s bring measure theory into the mix and define randomness in this setting.

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

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

$$\begin{aligned}\mu(\perp) &= 0 \\ x \leq y &\implies \mu(x) \leq \mu(y) \\ \mu(x) + \mu(y) &= \mu(x \vee y) + \mu(x \wedge y) \quad (\text{“modularity”}) \\ (x_i)_{i \geq 0} \text{ ascending} &\implies \mu\left(\bigvee_{i \geq 0} x_i\right) = \sup_{i \geq 0} \mu(x_i) \quad (\text{“continuity”}).\end{aligned}$$

We say  $\mu$  is a probability measure if  $\mu(\top) = 1$ . We will also call  $\mu$  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  $\mathcal{N}(Y) = \{U \in \mathcal{O}(X) \mid Y \subseteq U\}$ . We extend the measure  $\mu$  to a measure  $\mu^*$  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.

Before we get to randomness let’s stop to mention a huge payoff we can now reap. Simpson [3] proves in Example 4.7 that the standard measure on open subsets of  $\mathbb{R}^n$ , which extends to a measure on sublocales in the way we just defined, is preserved under Euclidean translations. This measure agrees with Lebesgue measure whenever the sublocale agrees with a measurable set. In standard measure theory there is no translation invariant measure on arbitrary subsets of  $\mathbb{R}^n$  due to pathological examples that are automatically avoided in locale theory. We have traded the initial awkwardness of working with locales and sublocales for a rather compelling cleanup of measure theory.

### 6.1 The random sublocale

We are ready to define the random sublocale  $\text{Ran}_X$  of an arbitrary locale  $X$  equipped with a probability measure  $\mu$ . We’ll do it three times, once for each definition of sublocale.

The spatial definition of  $\text{Ran}_X$  is the smallest sublocale of measure 1, or equivalently the intersection of all sublocales of measure 1. This tracks exactly with the definition of Kurtz randomness seen elsewhere in the literature. It's morally similar to  $\neg\neg X$  but making use of measure.  $\neg\neg X$  is the smallest dense sublocale of  $X$ , whereas  $\text{Ran}_X$  is the smallest measure-1 sublocale.

The quotient definition of  $\text{Ran}_X$  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 sublocale has no points. Given a point  $p \in X$ , the set  $X - \{p\}$  has measure 1 and so  $\text{Ran}_X \subset X - \{p\}$  and so  $p$  is not present in the random sublocale.

## 7 Kurtz randoms are not so bad

How Kurtz randomness fails the LLN (see [7], Section 3.5).

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

## 8 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  $\Sigma_n^0$  and  $\Pi_n^0$ . 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.

## 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] Peter T Johnstone. *Sketches of an elephant: a Topos theory compendium*. Oxford logic guides. Oxford Univ. Press, New York, NY, 2002.
- [7] André Nies. *Computability and Randomness*. Oxford Logic Guides. Oxford University Press, Oxford, 2009.