The Logical Approach to Randomness

Christopher P. Porter
University of Florida

UConn Logic Colloquium
September 18, 2015

Introduction

The concept of randomness plays an important role in mathematical practice, particularly in areas such as

- probability theory,
- mathematical statistics,
- real analysis,
- dynamical systems,
- combinatorics, and
- number theory.

Although the forms that randomness takes can vary across mathematical disciplines, one standard definition of randomness is to take an object to be random if it is obtained as the result of randomly choosing an element from some fixed collection of objects.

An alternative approach

Recently, an alternative approach to defining random mathematical objects has garnered a considerable amount of attention from researchers in computability theory.

I refer to this approach as *the logical approach to randomness*; in the context of computability theory is instantiated as the theory of algorithmic randomness.

Motivating question

To what extent do these logical definitions of randomness capture anything that mathematicians take to be significant about the concept of randomness?

Main claim: Various logical definitions of randomness do illuminate certain uses of randomness in classical mathematics.

Outline

1. Logical definitions of randomness
2. Two problems for the logical approach
3. Towards a solution to the two problems

1. Logical definitions of randomness

The main ingredients of a logical definition

Each logical definition of randomness is formulated in terms of three ingredients:

- a collection $\mathscr{X}$ of objects;
- a probability measure μ on $\mathscr{X}$; and
- a collection of properties $\{\Phi_i(x)\}_{i\in\omega}$, expressible in some formal language $\mathcal{L}$, and satisfiable by objects in $\mathscr{X}$, such that for each $i \in \omega$,

$$\mu(\{x \in \mathscr{X} : \Phi_i(x)\}) = 1.$$

Hereafter, I will refer to the properties $\{\Phi_i\}_{i\in\omega}$ as *randomness properties*.

Putting the ingredients together

From a triple $(\mathscr{X}, \mu, \{\Phi_i\}_{i \in \omega})$ satisfying the conditions from the previous slide, we get a definition $\mathscr{D}$ of μ-randomness for objects in $\mathscr{X}$ by stipulating that

$$x \in \mathscr{X} \text{ is } \mathscr{D}\text{-random if and only if } \Phi_i(x) \text{ for every } i \in \omega.$$

One immediate consequence of this definitional framework is that, assuming that there is some $x \in \mathscr{X}$ and some $i \in \omega$ such that x does not satisfy Φ_i, we can partition $\mathscr{X}$ into

- a non-empty collection of $\mathscr{D}$-random objects, and
- a non-empty collection of non-$\mathscr{D}$-random objects.

Examples of logical definitions of randomness

Let $(\mathscr{X}, \mu)$ be a probability space that satisfies certain effectivity conditions, where μ is a probability measure whose values on basic open subsets of $\mathscr{X}$ can be effectively approximated. For example:

- $[0, 1]$ with the Lebesgue measure;
- 2^ω with a Bernoulli measure with parameter $p \in \mathbb{Q}$;
- $\mathcal{C}[0, 1]$ with the Wiener measure.

On such spaces, there are a number of non-equivalent definitions of algorithmic randomness.

The statistical definition of randomness (for $2^{<\omega}$)

Given a finite string $\sigma \in 2^{<\omega}$, we'd like to test whether it is random.

Null hypothesis: σ is random.

How do we test this hypothesis?

We employ a statistical test $\mathcal{T}$ that has a critical region U corresponding to the significance level α.

If our string is contained in the critical region U, we reject the hypothesis of randomness at level α (say, $\alpha = 0.05$ or $\alpha = 0.01$).

The statistical definition of randomness (for 2^ω)

Given an infinite sequence $X \in 2^\omega$, we'd like to test whether it is random.

Null hypothesis: X is random.

How do we test this hypothesis?

We test initial segments of X at *every level of signficance*:
$\alpha = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots, \frac{1}{2^n}, \ldots$

A test for 2^ω is now given by an infinite collection $(\mathcal{T}_i)_{i\in\omega}$ of tests for $2^{<\omega}$, where the critical region U_i of $\mathcal{T}_i$ corresponds to the significance level $\alpha = 2^{-i}$.

Formally. . .

A *Martin-Löf test* is a sequence $(U_i)_{i\in\omega}$ of uniformly computably enumerable sets of strings such that for each i,

$$\sum_{\sigma\in U_i} 2^{-|\sigma|} \leq 2^{-i}.$$

(Think of each U_i as the critical region for a statistical test $\mathcal{T}_i$ at significance level $\alpha = 2^{-i}$.)

A sequence $X \in 2^\omega$ *passes a Martin-Löf test* $(U_i)_{i\in\omega}$ if there is some i such that for every k, $X{\upharpoonright}k \notin U_i$.

$X \in 2^\omega$ is *Martin-Löf random*, denoted $X \in \mathrm{MLR}$, if X passes *every* Martin-Löf test.

The measure-theoretic formulation

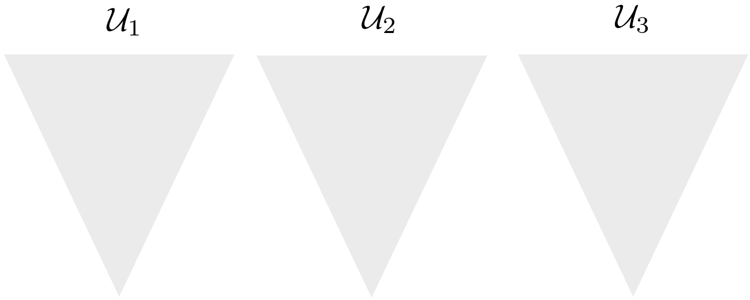
Given $\sigma \in 2^{<\omega}$,
$$[\![\sigma]\!] := \{X \in 2^\omega : \sigma \prec X\}.$$

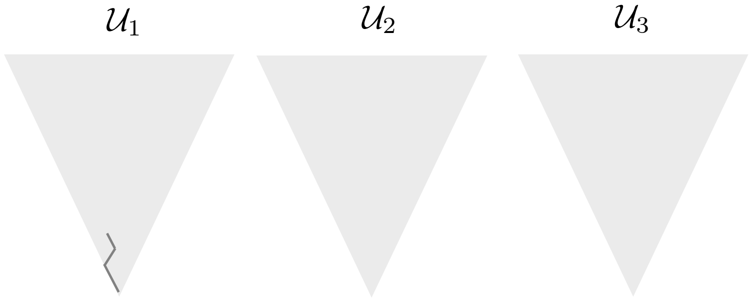
These are the basic open sets of 2^ω.

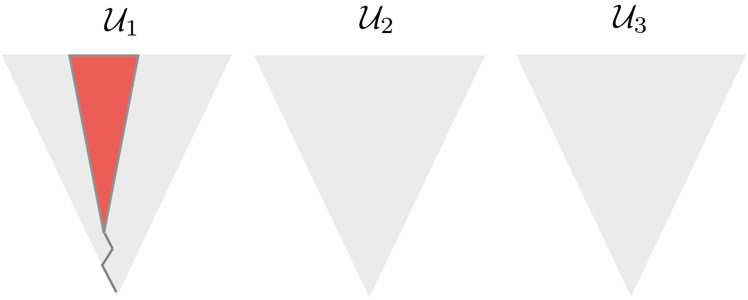
The Lebesgue measure on 2^ω is defined by
$$\lambda([\![\sigma]\!]) = 2^{-|\sigma|}.$$

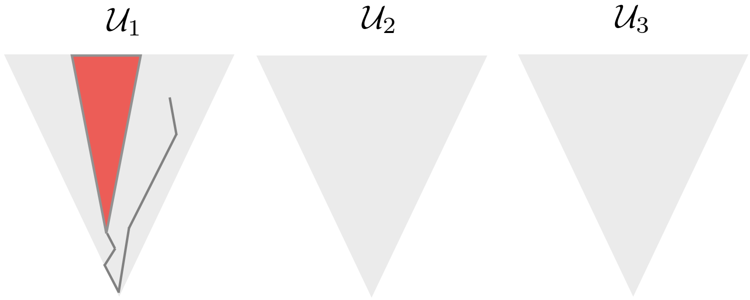
Thus we can consider a Martin-Löf test to be a collection $(\mathcal{U}_i)_{i\in\omega}$ of uniformly effectively open subsets of 2^ω such that
$$\lambda(\mathcal{U}_i) \leq 2^{-i}$$
for every i.

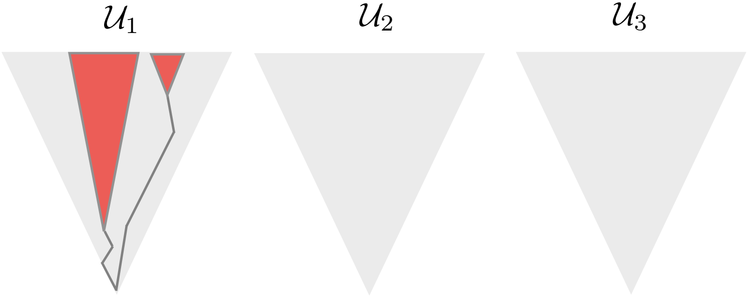
Moreover, X passes the test $(\mathcal{U}_i)_{i\in\omega}$ if $X \notin \bigcap_i \mathcal{U}_i$.


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

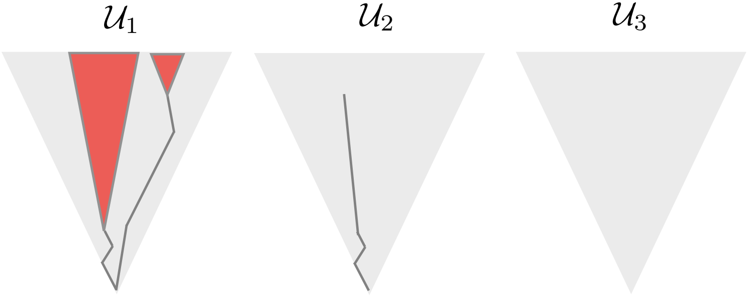

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

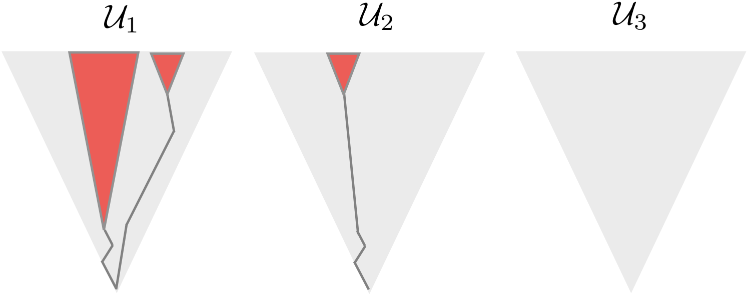

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

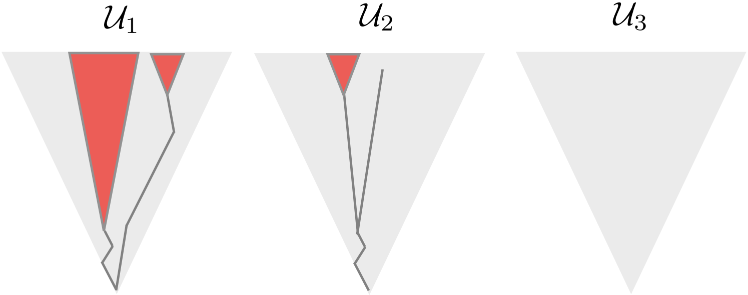

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

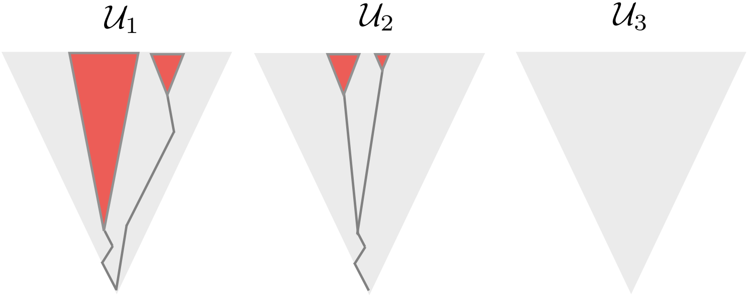


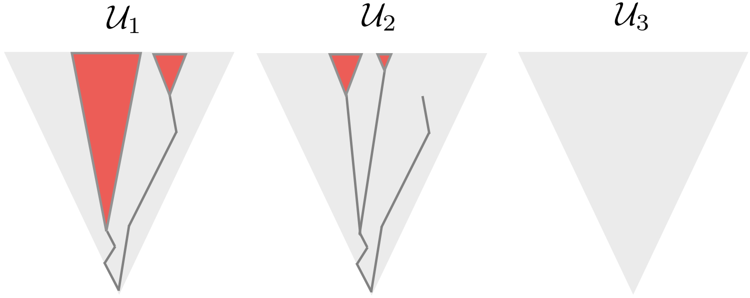
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

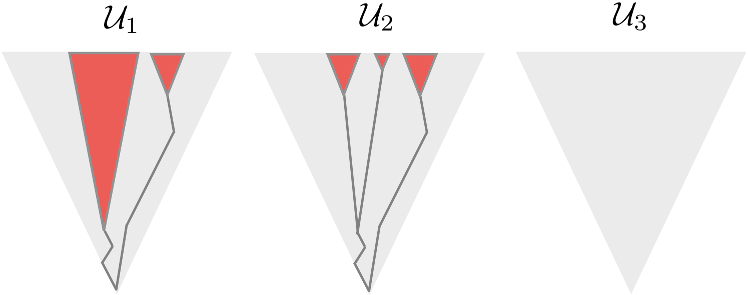

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

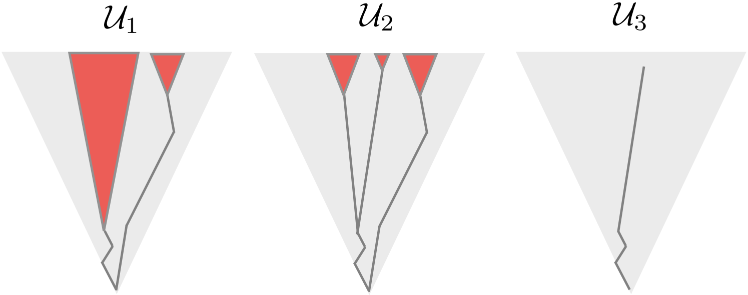

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

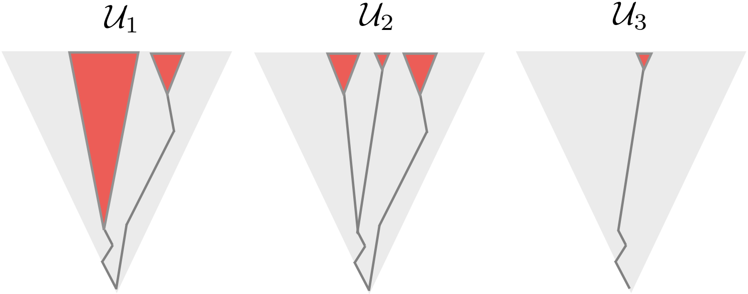

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

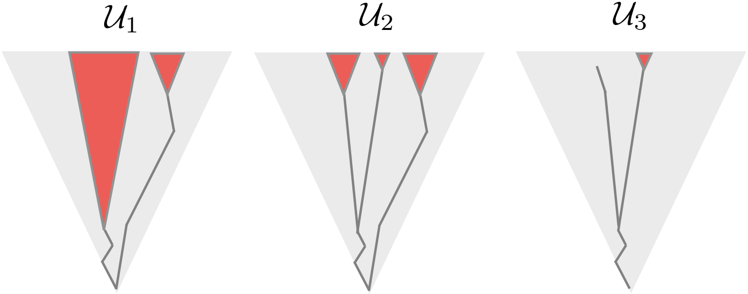

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

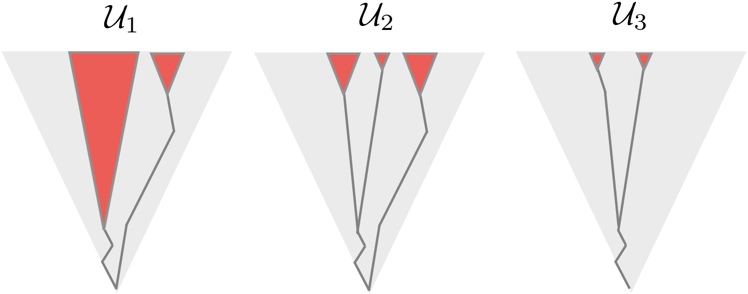

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

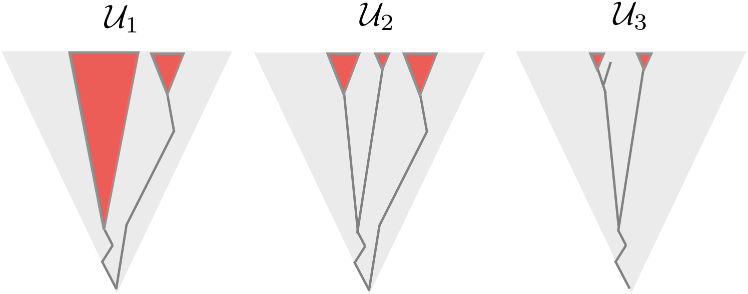

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

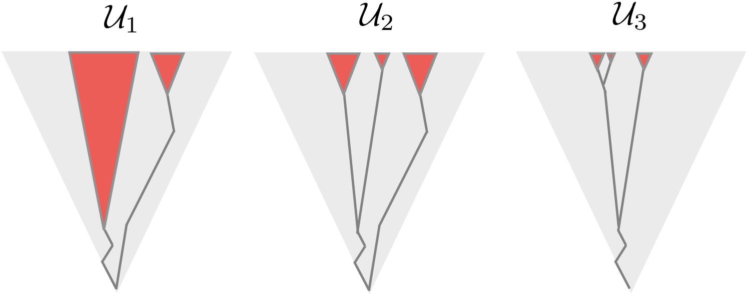

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$



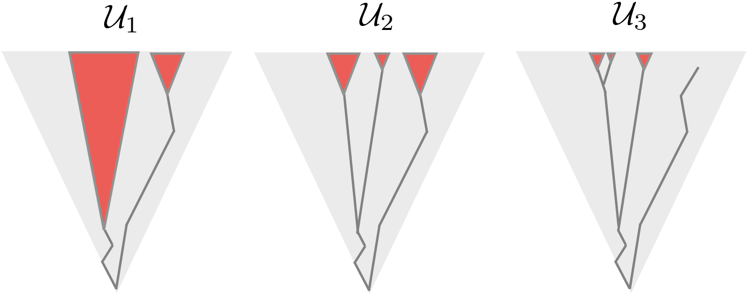
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

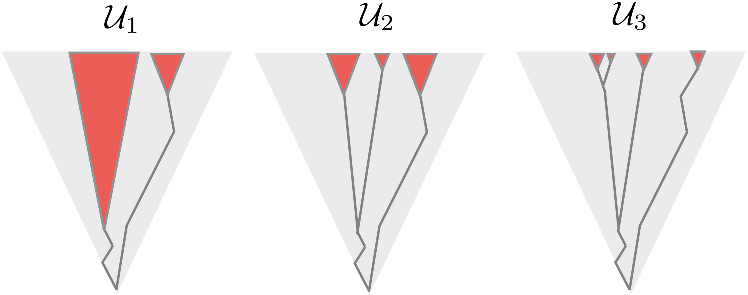


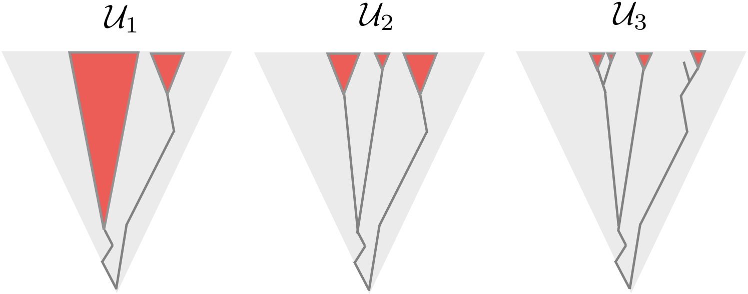
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

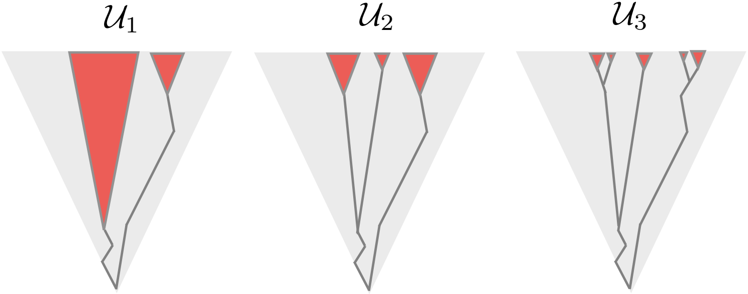


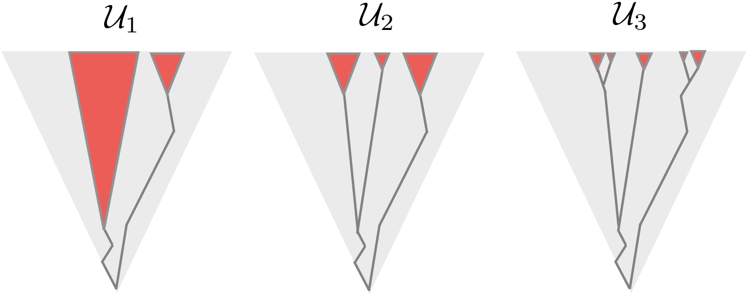
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

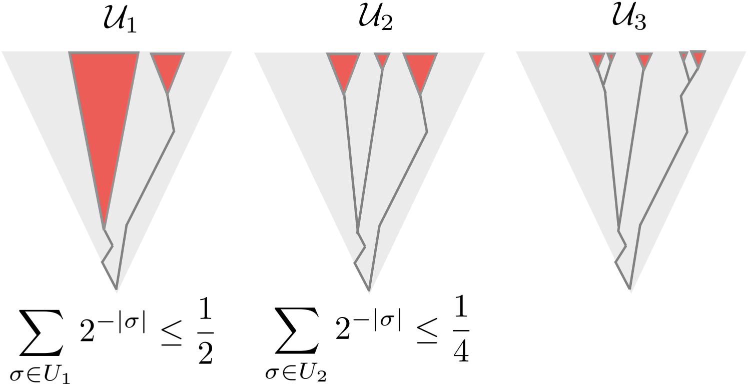

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

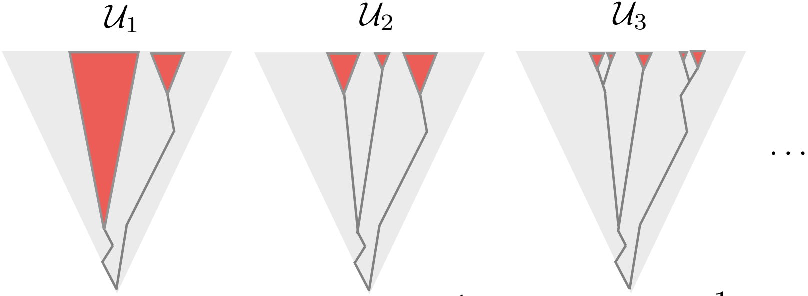

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$



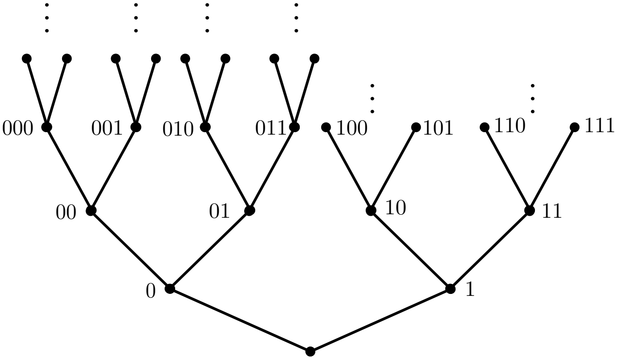
$$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2}$$

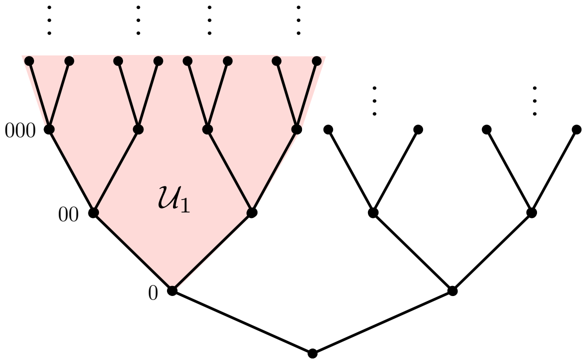


$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$
$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2}$
$\sum_{\sigma \in U_2} 2^{-|\sigma|} \leq \frac{1}{4}$

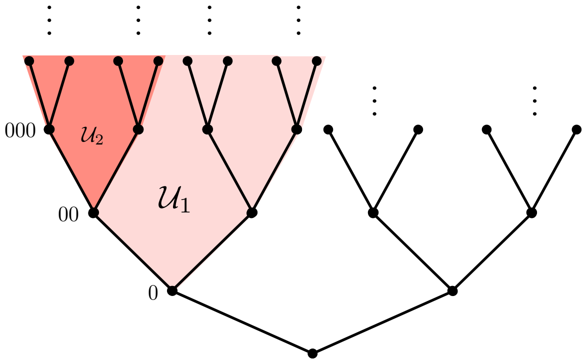


$$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2} \qquad \sum_{\sigma \in U_2} 2^{-|\sigma|} \leq \frac{1}{4} \qquad \sum_{\sigma \in U_3} 2^{-|\sigma|} \leq \frac{1}{8}$$

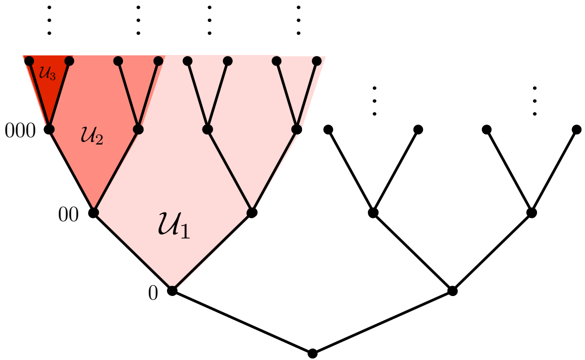

000
001
010
011
100
101
110
111
00
01
10
11
0
1



000
00
0
$\mathcal{U}_1$



000
$\mathcal{U}_2$
$\mathcal{U}_1$
00
0



$\mathcal{U}_3$
000
$\mathcal{U}_2$
00
$\mathcal{U}_1$
0

Additional logical definitions of randomness

Let μ be a computable probability measure on $\mathscr{X}$.

- Martin-Löf randomness:
  - A μ-Martin-Löf test $(\mathcal{U}_i)_{i\in\omega}$ is a sequence of uniformly effectively open subsets of $\mathscr{X}$ such that $\mu(\mathcal{U}_i) \leq 2^{-i}$ for every i.
  - Randomness properties $= \mathscr{X} \setminus \bigcap_{i\in\omega} \mathcal{U}_i$ for each Martin-Löf test $(\mathcal{U}_i)_{i\in\omega}$.
- Schnorr randomness:
  - A μ-Schnorr test $(\mathcal{U}_i)_{i\in\omega}$ is a Martin-Löf test that satisfies the property that $\mu(\mathcal{U}_i)$ is computable for every i.
  - Randomness properties $= \mathscr{X} \setminus \bigcap_{i\in\omega} \mathcal{U}_i$ for each Schnorr test $(\mathcal{U}_i)_{i\in\omega}$.
- Weak n-randomness:
  - Randomness properties $= \mathscr{X} \setminus \mathcal{S}$ for each Π^0_n $\mathcal{S} \subseteq \mathcal{X}$ satisfying $\mu(\mathcal{S}) = 0$.

In general, these definitions are not equivalent.

Valuative randomness

The logical approach to randomness differs radically from one frequently occurring notion of randomness in classical mathematics, which I refer to as *valuative randomness*.

Roughly speaking, the idea behind valuative randomness is this: to be random is to be the value of a random variable.

Recall that a random variable is a measurable function from a sample space Ω to some space, usually $\mathbb{R}$.

The usage of 'random' is not exact here; randomness is usually attributed to the function itself, but sometimes it is also attributed to individual outputs of the function.

ϕ-valued random variables

However, it is important to emphasize that in practice, the range of a random variable can be any collection of mathematical objects:

- complex numbers
- vectors
- matrices
- functions
- graphs
- closed sets
- measures
- and so on...

Let ϕ be a mathematical object such as one from any of the collections listed above.

Then a random ϕ is simply a ϕ-valued random variable.

Where exactly is the randomness?

It is common to think of a random variable as yielding the values of some random or chancy experiment (such as some measurement of some randomly selected individual).

Thus, a ϕ-valued random variable can be understood as yielding as output a randomly chosen ϕ from the relevant collection of objects.

Note that this random experiment/choice isn't technically part of the definition of a random variable, but in applications, such experiments or choices are often associated to random variables.

Almost sure events

Random variables can take values that we would not expect to arise as the result of some random experiment.

For instance, a real-valued random variable can take some value $q \in \mathbb{Q}$, or a graph-valued random variable can produce a complete graph as output.

However, there is a sense in which such outcomes are atypical.

In particular, one can associate a probability distribution to a random variable, and by means of this probability distribution, one can define events that happen almost surely (i.e. with probability one).

Thus, if some property Θ occurs almost surely with respect to the probability distribution associated to a ϕ-valued random variable, we say, "a random ϕ has Θ almost surely."

Comparing the logical and valuative approaches

The key distinction between the logical and valuative approaches is the former is *discriminative* while the latter is not.

That is, on the logical approach, one discriminates between the random and the non-random objects.

By contrast, on the valuative approach, *any* object in the relevant domain of objects can be the value of a random variable (and thus can be counted as random).

Moreover, on the valuative approach, one does not typically attribute non-randomness to any objects.

2. Two problems for the logical approach

The thrust of the two problems

The logical approach to defining randomness faces two serious problems, which I refer to as

- the randomness property problem; and
- the underlying measure problem.

The general thrust of these problems is that each logical definition of randomness depends on the choice of specific parameters, which, if not chosen on some principled basis, threaten to trivialize the logical approach to randomness.

Motivating the randomness property problem

For each object $x \in \mathscr{X}$, if we let the formula $\phi_x(y)$ be

$$y \neq x,$$

then assuming that $\mu(\{x\}) = 0$, we will have

$$\mu\big(\{y \in \mathscr{X} : \phi_x(y)\}\big) = 1.$$

Moreover, if μ is continuous (i.e., $\mu\big(\{y\}\big) = 0$ for every $y \in \mathscr{X}$), then each of the formulas in $\{\phi_x\}_{x \in \mathscr{X}}$ defines a set of μ-measure one.

The randomness property problem (1)

The example on the previous slide shows that one cannot require random objects to satisfy *every* measure one property, for otherwise the resulting definition of randomness would have an empty extension.

But note that for any given object $x \in \mathscr{X}$, we can always include the property ϕ_x among the collection $\{\Phi_i\}_{i \in \omega}$ of randomness properties.

That is, for any $x \in \mathscr{X}$ there is always some choice of randomness properties that excludes x as non-random.

In light of this problem, for nearly 45 years, one central question in the development of algorithmic randomness was: Which properties should we count as *the* randomness properties?

The randomness property problem (2)

The answer to this question about a choice of randomness properties depends on the role we want a logical definition of randomness to play.

We can thus cast the randomness property problem relative to some aims or purposes:

RPP: For a given set of purposes, is there a principled choice of randomness properties that yields a notion of randomness that successfully fulfills these purposes?

The RPP in context (1)

The prototype for the logical definitions that are studied today was first given by von Mises in 1919.

On von Mises' approach, a sequence is random if

(i) the limiting relative frequency of each element in the sequence exists, and
(ii) this limiting relative frequency is invariant under selecting subsequences from the original sequence.

Note that we cannot require invariance under the selection of all possible subsequences, as the only sequences that would be counted as random are those that are nearly constant.

However, von Mises did not initially specify which selection rules were to be used in his definition.

Aware of this problem, von Mises' contemporaries objected that his definition was defective.

The RPP in context (2)

Subsequently, Wald proved that any *countable* collection of selection rules yields a definition of randomness satisfied by continuum many sequences.

Doob proved that invariance under a single selection rule is a measure one property (with respect to the relevant measure).

Which countable collection of selection rules should be used to define randomness?

- Wald: Those rules definable in some "logic."
- Church: Those rules that are effectively computable.
- Kruse: Studied selection rules definable in various set theories.
- Agafonov: Rules given by a finite-state automaton yield the normal sequences as random.

The RPP in context (3)

Ville proved that no matter which countable collection of selection rules is chosen, there is some measure one property that fails to be satisfied by the resulting notion of randomness.

In response to this problem, Ville developed his own notion of randomness in terms of certain betting strategies he called *martingales*.

Ville proved that the collection of sequences on which a martingale fails to win unbounded capital has measure one (again, with respect to the relevant measure).

Which martingales should be used to define randomness?

The RPP in context (4)

Martin-Löf: Martin-Löf tests capture all randomness properties that one will encounter in "present or future use in statistics."

Schnorr argued that Martin-Löf tests yield too many randomness properties and thus fail to capture "the true concept of randomness."

In Schnorr's view, only measure one properties defined by Martin-Löf tests that are "visualizable" should be counted as randomness properties.

Such properties correspond precisely to the collection of Schnorr tests.

Has the RPP been adequately addressed?

Despite these latter developments, over the 40+ years since the contributions of Martin-Löf and Schnorr, there has yet to be a clear articulation of what these definitions are intended to capture.

That being the case, no one has offered a systematic account as to why any of the currently available definitions of algorithmic randomness adequately address the randomness property problem.

Motivating the underlying measure problem

For a given probability space $\mathscr{X}$, a formula ϕ that defines a set of measure one with respect to one measure μ may define a set of measure zero with respect to another measure ν (so that the formula $\neg\phi$ defines a set of ν-measure one).

Consequently, the extension of any logical definition of randomness with underlying measure μ that counts the property ϕ as a randomness property will be disjoint from the extension of any logical definition with underlying measure ν that counts the property $\neg\phi$ as a randomness property.

To complicate matters...

For any object $x \in \mathscr{X}$, there is a measure μ_x on $\mathscr{X}$ such that $\mu_x(\{x\}) = 1$ (i.e., the *Dirac measure* concentrated on x).

Then the only μ-random element of $\mathscr{X}$ is x.

Randomness with respect to non-computable measures

One need not appeal to Dirac measures to find such peculiar measures.

If we consider, say, Martin-Löf randomness with respect to non-computable measures on 2^ω, one can prove the following:

Theorem (Reimann-Slaman)

For every sequence $X \in 2^\omega$, X is non-computable if and only if there is some measure μ such that

(i) $\mu(\{X\}) = 0$ *and*

(ii) *X is Martin-Löf random with respect to μ.*

Surprisingly, this fact can be witnessed by a *single* measure!

The underlying measure problem

UMP: How can we countenance notions of randomness with respect to different probability measures without potentially counting every object as random?

More concisely, which measures yield “legitimate” notions of randomness?

3. Towards a solution of the two problems

One possible strategy

One strategy for responding to these problems is to identify a definition of randomness, given by one collection of randomness properties and one underlying measure, and argue that this is the "correct" definition of randomness.

Just as the notion of Turing computable function captures the intuitive conception of effectively calculable function, we could hope to isolate a single definition of randomness that captures the intuitive conception of randomness.

A worry about this strategy

Although some have held that there is such a single correct definition of randomness, this view has always been articulated for definitions of random sequence with respect to the Lebesgue measure.

For instance, both Martin-Löf randomness and Schnorr randomness have been held to capture the intuitive conception of randomness.

But what about definitions of randomness for other objects, and with respect to different measures?

Should we hope for one general definition of randomness that is correct for each choice of objects and each choice of underlying measure?

A two-pronged solution

1. Seek to ground the various choices of randomness properties in classical results involving valuative randomness.

2. Concede that logical definitions of randomness give rise to artifacts, i.e. unintended consequences that result when we apply the tools of logic to the task of defining randomness.
   - Classify the various kinds of artifacts that arise.
   - Diagnose the sources of these artifacts.

Almost sure behavior

As we saw in our discussion of valuative randomness, in classical mathematics one commonly finds theorems of the form

- "the random ϕ has property Θ almost surely."

Recently, there has been a number of theorems in algorithmic randomness of the form:

- "the algorithmically random ϕ has Θ"

More significantly, we have a number of stronger results of the form:

ϕ has Θ *if and only if* ϕ is algorithmically random.

A theorem involving almost sure behavior

Consider the following example:

Theorem: For every real-valued function $f : [0, 1] \to \mathbb{R}$ of bounded variation, f is differentiable almost everywhere.

A few observations:

- The function quantifier in this theorem ranges over sets of size $2^{\mathfrak{c}}$, the size of the power set of the continuum.
- The properties "being a point of differentiability of some real-valued function of bounded variation" and "being a point of non-differentiability of some real-valued function of bounded variation" are satisfied by every point in [0,1].

A restricted version of the theorem

Now consider:

For every *computable* $f : [0, 1] \to \mathbb{R}$ of bounded variation, f is differentiable almost everywhere.

A few observations:

- The function quantifier in this theorem now ranges over countably many functions.
- Thus the property "being a point of differentiability of every computable real-valued function of bounded variation" is the intersection of countably many sets of Lebesgue measure one, which is itself a set of Lebesgue measure one.

The connection to randomness

Theorem (Brattka, Miller, Nies)

$z \in [0,1]$ is Martin-Löf random if and only if every computable, real-valued function $f : [0,1] \to \mathbb{R}$ of bounded variation is differentiable at z.

That is, Martin-Löf randomness is necessary and sufficient for this particular instance of almost sure behavior.

Almost sure behavior in classical analysis

Such results hold for a number of definitions of algorithmic randomness:

$x \in \mathrm{MLR}$	$\Leftrightarrow$	Every computable real-valued function of bounded variation is differentiable at x.
$x \in \mathrm{SR}$	$\Leftrightarrow$	For every L_1-computable real-valued function f, the Lebesgue differentiation theorem holds for f at x.
$x \in \mathrm{W2R}$	$\Leftrightarrow$	Every computable real-valued a.e.-differentiable function is differentiable at x.

There are a number of other examples, some involving definitions of randomness that we have not considered here.

More examples

There are other promising developments along similar lines:

- results in ergodic theory with respect to shift-invariant measures;
- Martin-Löf random closed sets;
- Martin-Löf random Brownian motion;
- effective notions of Hausdorff and packing dimension.

What do these examples tell us?

Defintions such as MLR, SR, and W2R correspond to effective versions of almost sure behavior that are of independent interest to mathematicians.

For the purposes of classifying the effective content of almost sure behavior in classical mathematics, these definitions thus prove to be extremely useful.

The different choices of randomness properties that yield these definitions are thus vindicated by these examples.

Artifacts of logical definitions

Although the previous results indicate that certain choices of randomness properties and underlying measures yield interesting and informative definitions of randomness, we still have to account for the pathological behavior that logical definitions can yield.

The challenge is to determine which features of our logical definitions are artifacts and which are not.

The Reimann-Slaman example

The measures in Reimann-Slaman theorem are admittedly exotic (for instance, it is necessary that they give *some* points positive measure, i.e. they are necessarily discontinuous).

A case can be made that the Reimann-Slaman theorem and related results are artifacts of the computational framework used to define randomness (particularly when we consider non-computable measures).

But on what grounds can we rule out these definitions as illegitimate?

Even if there is at best a piecemeal answer to this question, as we have seen, there are choices of underlying measure that yield mathematically significant notions of randomness.

In conclusion

By grounding the choice of randomness properties and underlying measures in results concerning almost sure behavior in classical mathematics, we vindicate these choices.

However, much work remains to be done in accounting for which aspects of our logical definitions legitimately reflect features of mathematical randomness, and which are merely artifacts.