Negligibility, depth, and algorithmic randomness
Part 1

Christopher P. Porter
University of Florida

Joint work with Laurent Bienvenu

UF Logic Seminar
September 5, 2014

Introduction

The goal of this talk is to explain the way in which the theory of algorithmically random sequences can give us insight into certain limitations of probabilistic computation.

In particular, I will explain certain limitations in terms of two kinds of effectively closed classes (i.e. Π^0_1 classes):

1. negligible Π^0_1 classes;
2. deep Π^0_1 classes.

A brief history 1

Gödel's first incompleteness theorem tells us that there is no effective procedure for producing a consistent completion of Peano arithmetic (hereafter, PA).

In the early 1970's, Jockusch and Soare strengthened this result by proving (essentially) that the probability of producing a consistent completion of PA via a probabilistic procedure is zero.

In modern terminology, the set of consistent completions of PA is *negligible*.

A brief history 2

In the early 2000s, Levin strengthened the Jockusch/Soare result by proving that the probability of producing some initial segment of a consistent completion of PA goes to zero *quickly*.

This property is what we have isolated as the notion of *depth*.

Outline of today's talk

1. Some basics of algorithmic randomness
2. Probabilistic Turing computation
3. Negligible Π^0_1 classes

1. Some basics of algorithmic randomness

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 0000000000000000000000000000000000000000000000000000000

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

(3) 10100000110101000110101101000111110000111110100011

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

(3) 101000001101010001101011010001111100000111110100011

(4) 00100100001111110110101010001000100001011010001100

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 0000000000000000000000000000000000000000000000000000000000

(2) 0101010101010101010101010101010101010101010101010101010101

(3) 1010000011010100011001101011010001111101110000111110100011

(4) 0010010000111111011101101010100010001000100001011010001100

(5) 0100100101101011111111111101010100111100111111111110010

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

(3) 10100000110101000110101101000111110000111110100011

(4) 00100100001111110110101010001000100001011010001100

(5) 01001001011010111111110101010011110011111111110010

(3) List names of American states alphabetically: 0 = even # of letters, 1 = odd # of letters.

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

(3) 10100000110101000110101101000111110000111110100011

(4) 00100100001111110110101010001000100001011010001100

(5) 01001001011010111111101010100111100111111111110010

(3) List names of American states alphabetically: 0 = even # of letters, 1 = odd # of letters.

(4) First fifty digits of the binary expansion of π.

A motivating question

What does it mean for a sequence of 0s and 1s to be random?

Consider the following examples:

(1) 00000000000000000000000000000000000000000000000000

(2) 01010101010101010101010101010101010101010101010101

(3) 1010000001101010001101011010001111100000111110100011

(4) 00100100001111110110101010001000100001011010001100

(5) 01001001011010111111110101010011110011111111110010

(3) List names of American states alphabetically: $0 =$ even $\#$ of letters, $1 =$ odd $\#$ of letters.

(4) First fifty digits of the binary expansion of π.

(5) Fifty digits obtained from random.org (atmospheric noise?).

A rough definition of algorithmic randomness

Intuitively, a sequence is algorithmically random if it contains no "effectively definable regularities".

"effectively definable regularities" $\approx$ patterns definable in some computable way

In the absence of such regularities, algorithmically random sequences are not detected as non-random by any effective test for randomness.

However, if a sequence contains some "effectively definable regularity", there is some effective test for randomness that detects the sequence as non-random.

Towards a formal definition of algorithmic randomness

There are a number of ways one can formally characterize algorithmic randomness:

- in terms of betting strategies
- in terms of compressibility
- in terms of effectively definable null sets

Towards a formal definition of algorithmic randomness

There are a number of ways one can formally characterize algorithmic randomness:

- in terms of betting strategies
- in terms of compressibility
- in terms of effectively definable null sets $\Leftarrow$

Today I'll highlight a definition of randomness given in terms of statistical tests for randomness, where the statistical tests are effectively generated.

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?

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*:
$\alpha = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots, \frac{1}{2^n}, \dots$

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, $X\restriction k \notin U_i$.

$X \in 2^\omega$ 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

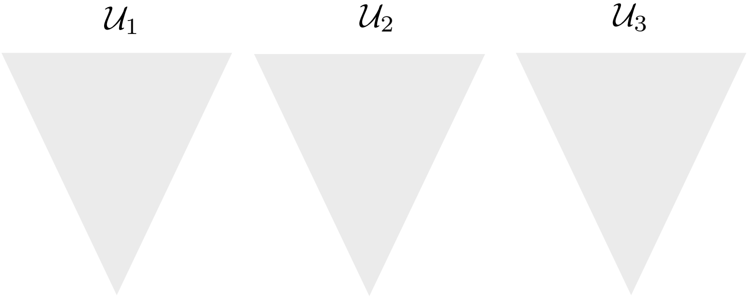
$$\lambda([\![\sigma]\!]) = 2^{-|\sigma|}.$$

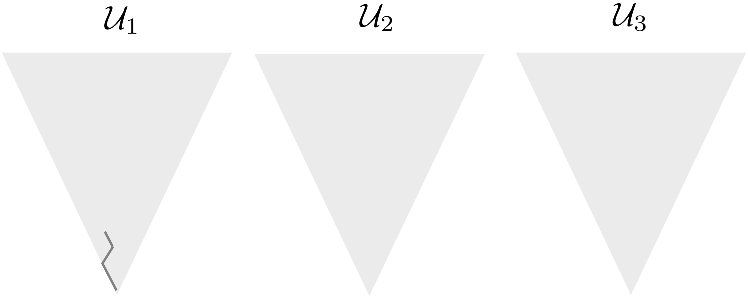
Thus we can consider a Martin-Löf test to be a collection $(\mathcal{U}_i)_{i\in\omega}$ of uniformly effectively open subsets of 2^ω such that

$$\lambda(\mathcal{U}_i) \leq 2^{-i}$$

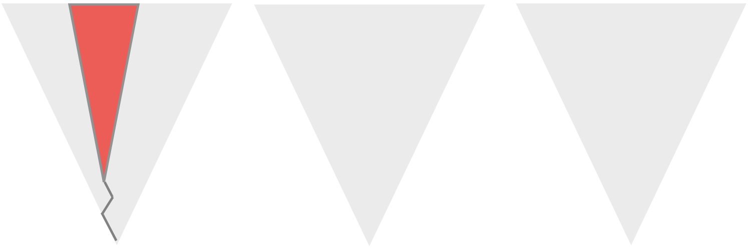
for every i.

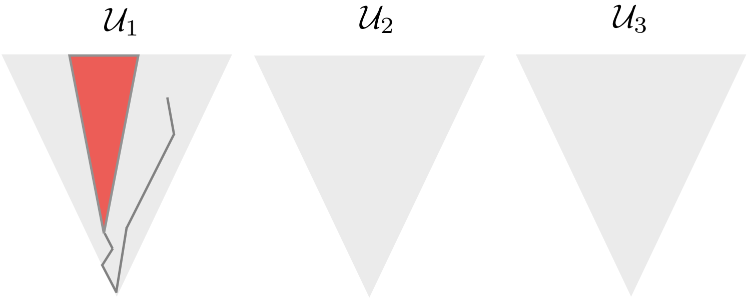
Moreover, X passes the test $(\mathcal{U}_i)_{i\in\omega}$ if $X \notin \bigcap_i \mathcal{U}_i$.

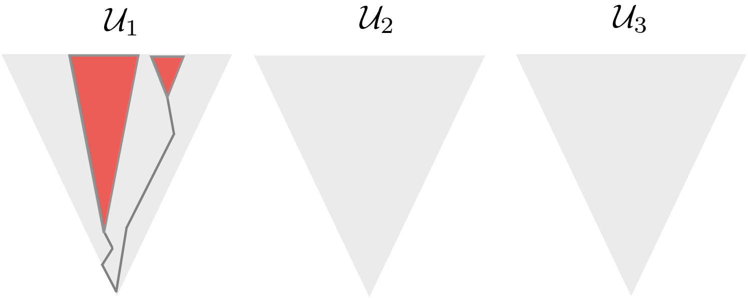

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

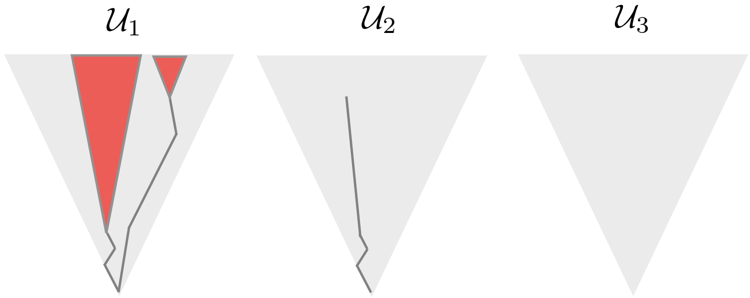

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

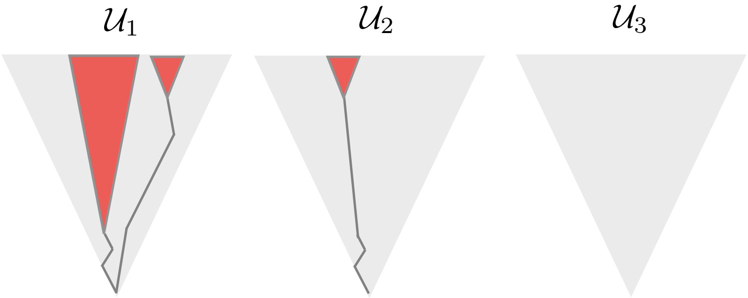
$\mathcal{U}_1$ $\mathcal{U}_2$ $\mathcal{U}_3$

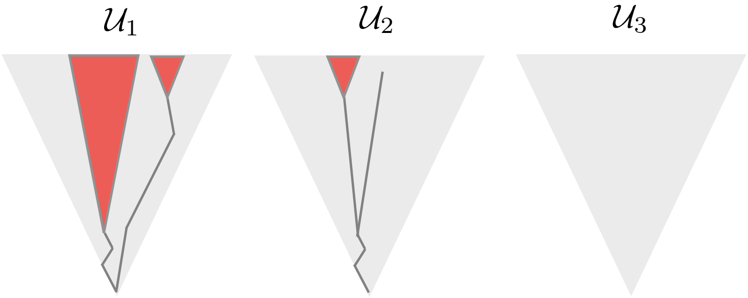


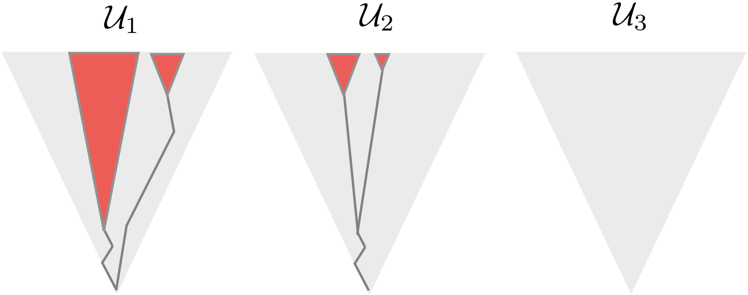

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

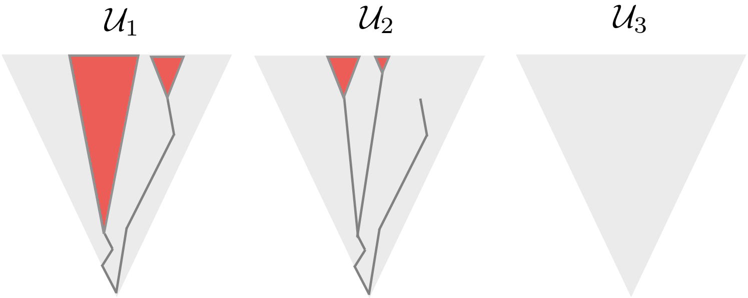

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

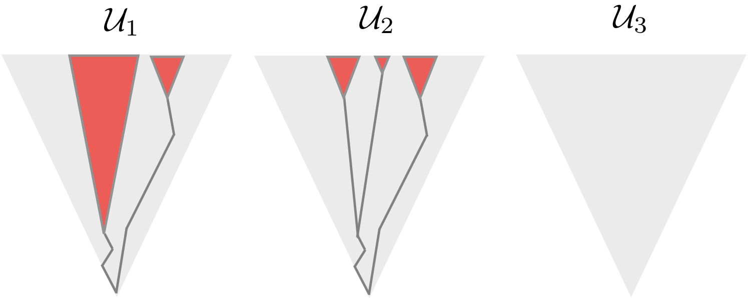

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

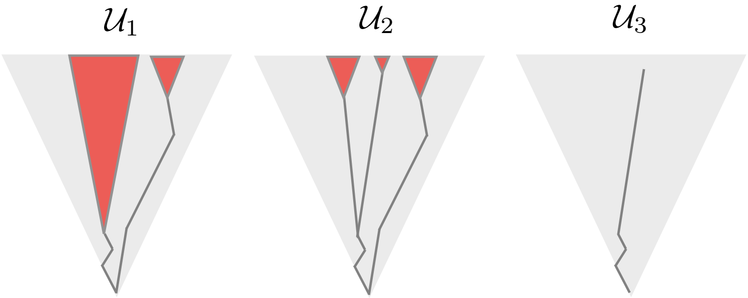

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


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

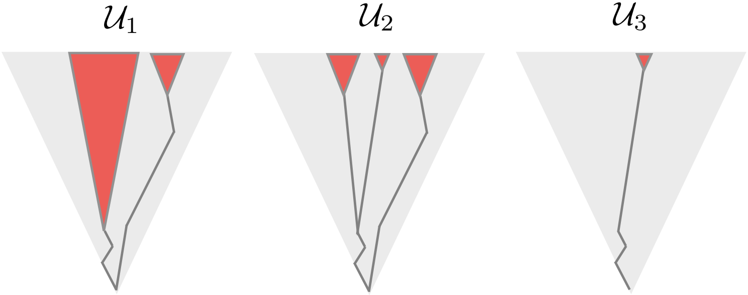

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

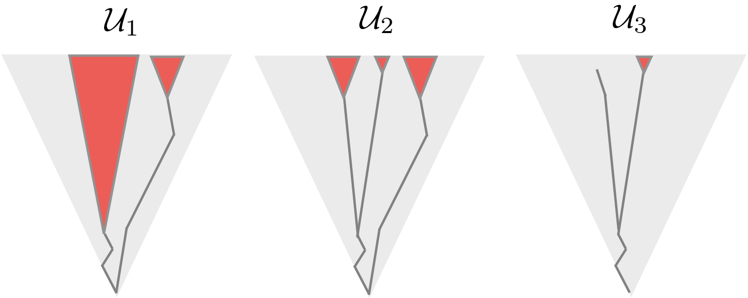

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

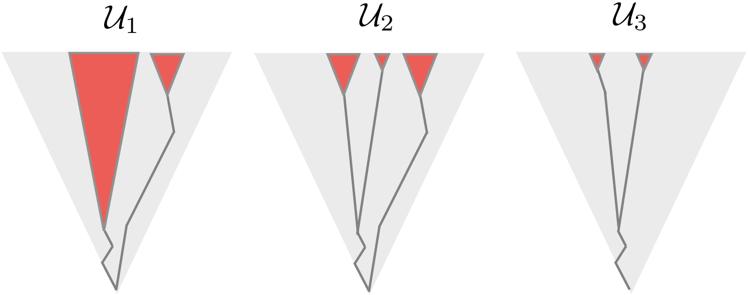

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$



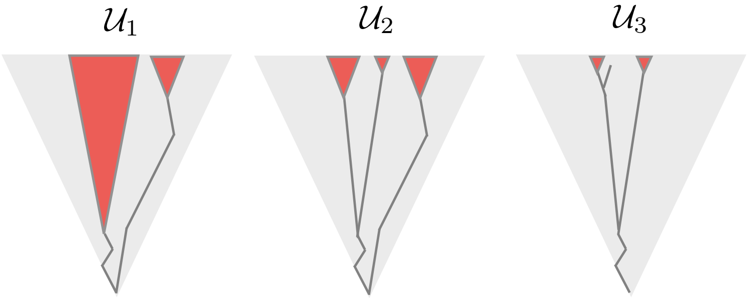
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


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

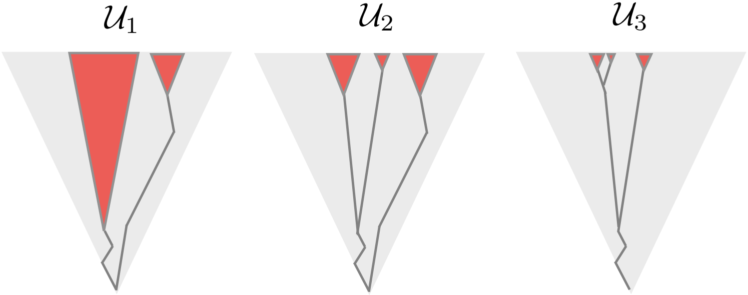

$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$



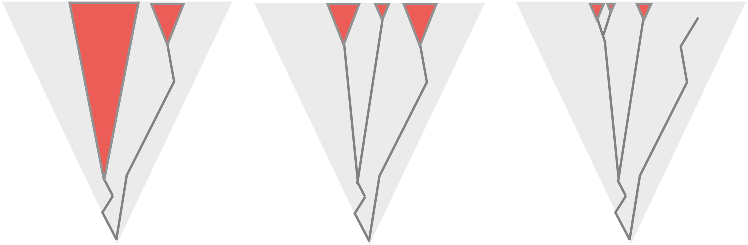
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$



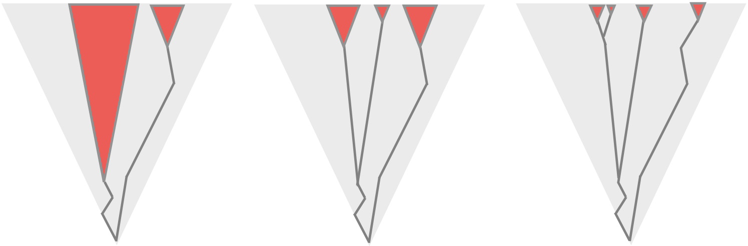
$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$


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

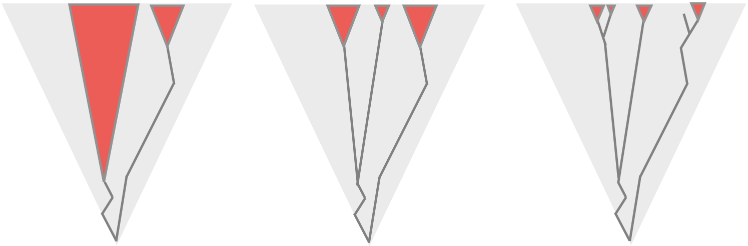
$\mathcal{U}_1$ $\mathcal{U}_2$ $\mathcal{U}_3$

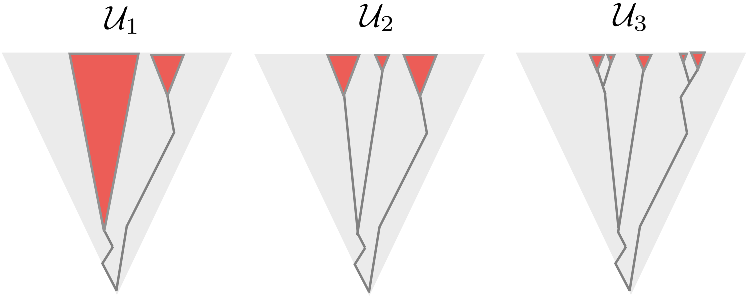


$\mathcal{U}_1$ $\qquad\qquad$ $\mathcal{U}_2$ $\qquad\qquad$ $\mathcal{U}_3$

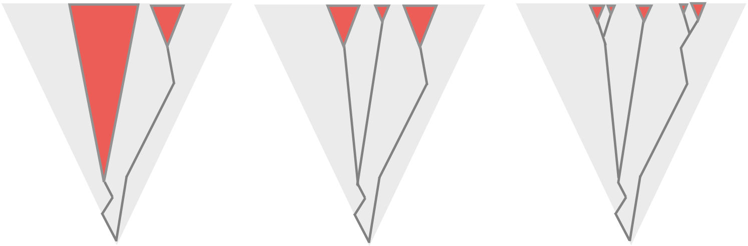


$\mathcal{U}_1$ $\mathcal{U}_2$ $\mathcal{U}_3$

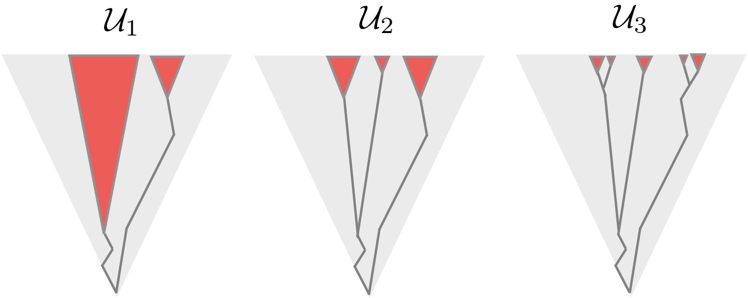



$\mathcal{U}_1$
$\mathcal{U}_2$
$\mathcal{U}_3$

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

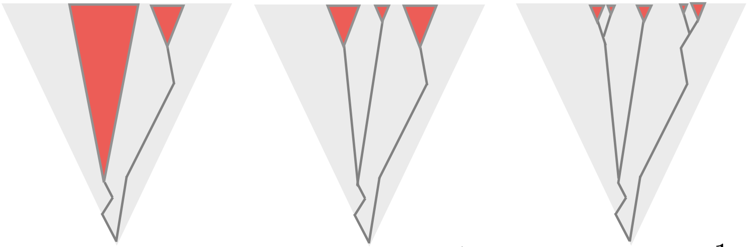


$$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2}$$

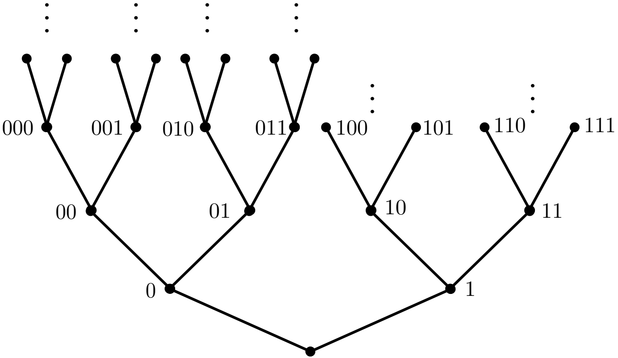


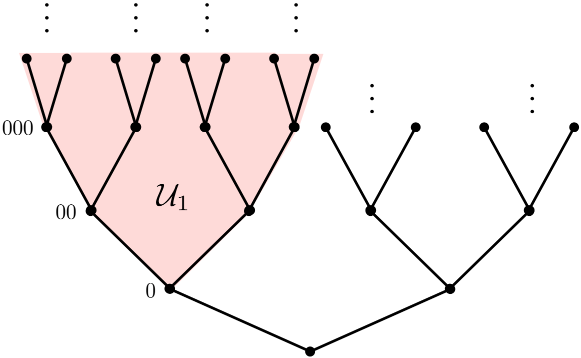
$$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2} \qquad \sum_{\sigma \in U_2} 2^{-|\sigma|} \leq \frac{1}{4}$$

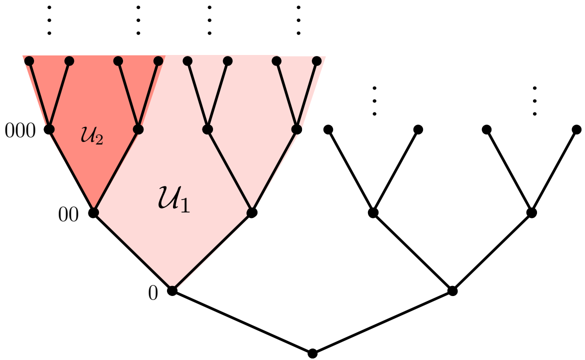
$\mathcal{U}_1$ $\qquad$ $\mathcal{U}_2$ $\qquad$ $\mathcal{U}_3$

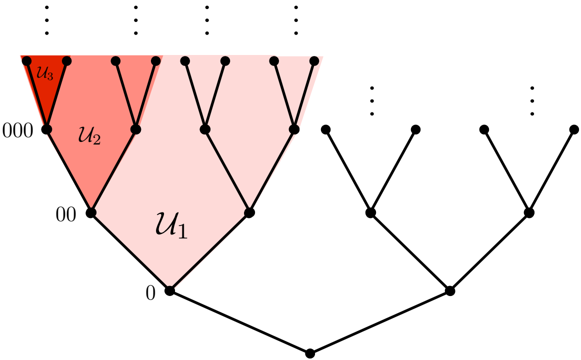


$$\sum_{\sigma \in U_1} 2^{-|\sigma|} \leq \frac{1}{2} \qquad \sum_{\sigma \in U_2} 2^{-|\sigma|} \leq \frac{1}{4} \qquad \sum_{\sigma \in U_3} 2^{-|\sigma|} \leq \frac{1}{8} \qquad \cdots$$


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


000
00
0
$\mathcal{U}_1$


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



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

Computable measures

We can also define Martin-Löf randomness with respect to any computable measure on 2^ω.

Definition

A measure μ on 2^ω is *computable* if $\sigma \mapsto \mu([\![\sigma]\!])$ is computable as a real-valued function.

In other words, μ is computable if there is a computable function $\hat{\mu} : 2^{<\omega} \times \omega \to \mathbb{Q}_2$ such that

$$|\mu([\![\sigma]\!]) - \hat{\mu}(\sigma, i)| \leq 2^{-i}$$

for every $\sigma \in 2^{<\omega}$ and $i \in \omega$.

From now on we will write $\mu(\sigma)$ instead of $\mu([\![\sigma]\!])$.

Randomness with respect to a computable measure

Definition

Let μ be a computable measure.

- A *μ-Martin-Löf test* is a sequence $(\mathcal{U}_i)_{i\in\omega}$ of uniformly effectively open subsets of 2^ω such that for each i,

$$\mu(\mathcal{U}_i) \leq 2^{-i}.$$

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

2. Probabilistic Turing computation

Two approaches to probabilistic computation

The standard definition of a probabilistic Turing machine is a non-deterministic Turing machine such that its transitions are chosen according to some probability distribution.

In the case of that this distribution is uniform, one can imagine that the machine is equipped with a fair coin that determines how it will transition from state to state.

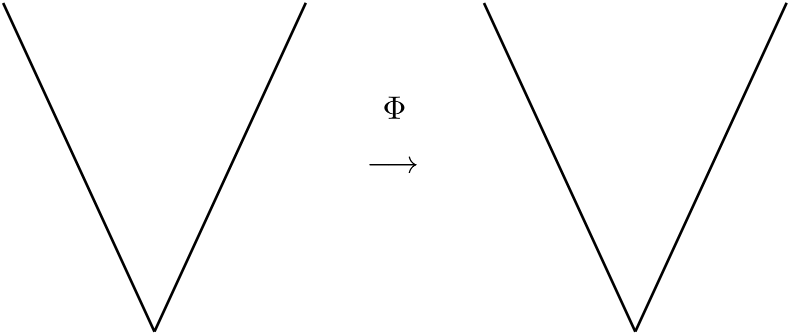
Alternatively, one can define a probabilistic machine to be an oracle Turing machine with some algorithmically random sequence as an oracle.

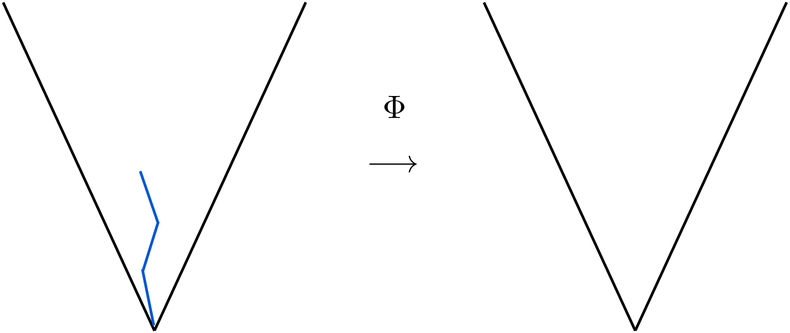
Key idea: For the purposes of computing a sequence or some sequence in a fixed collection *with positive probability*, these two approaches are equivalent.

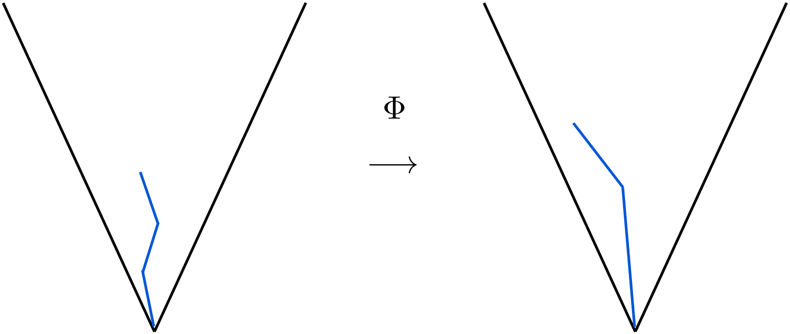
Turing functionals

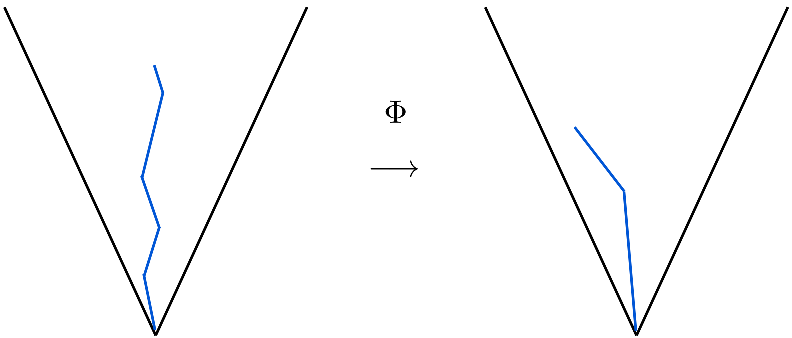
Definition

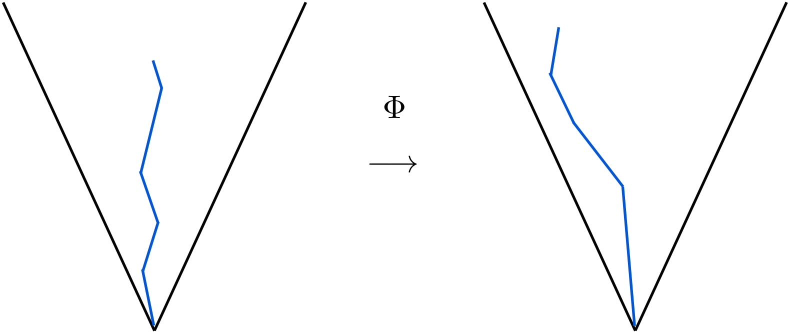
A *Turing functional* $\Phi : 2^\omega \to 2^\omega$ is a computably enumerable set S_Φ of pairs of strings (σ, τ) such that if $(\sigma, \tau), (\sigma', \tau') \in S_\Phi$ and $\sigma \preceq \sigma'$, then $\tau \preceq \tau'$ or $\tau' \preceq \tau$.

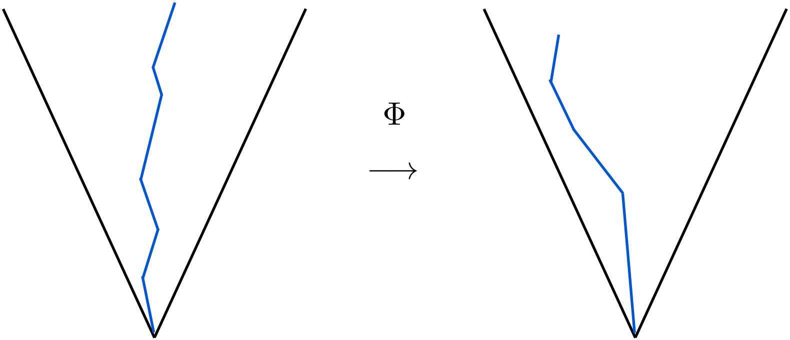

Φ

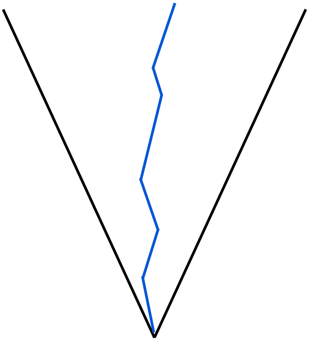

Φ


Φ

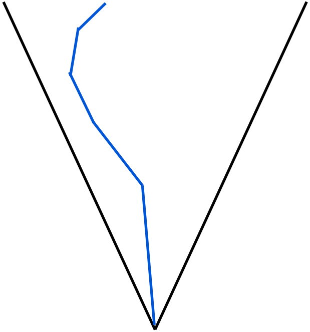

$\xrightarrow{\Phi}$

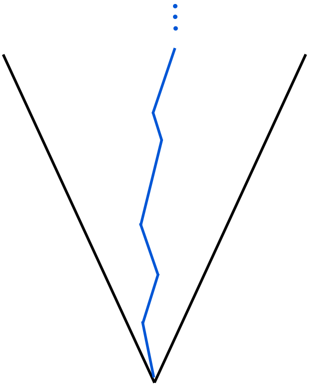

Φ
$\longrightarrow$


Φ

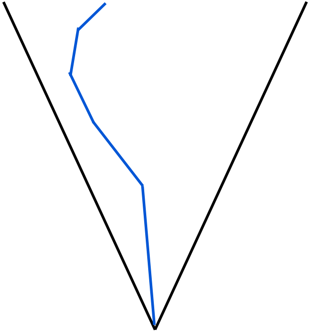


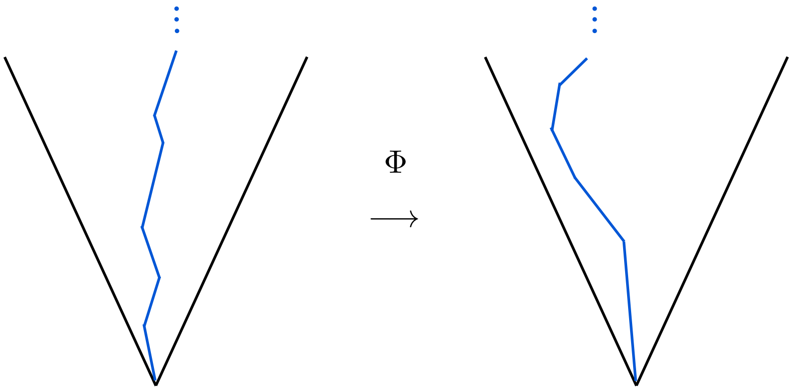
$$\xrightarrow{\Phi}$$

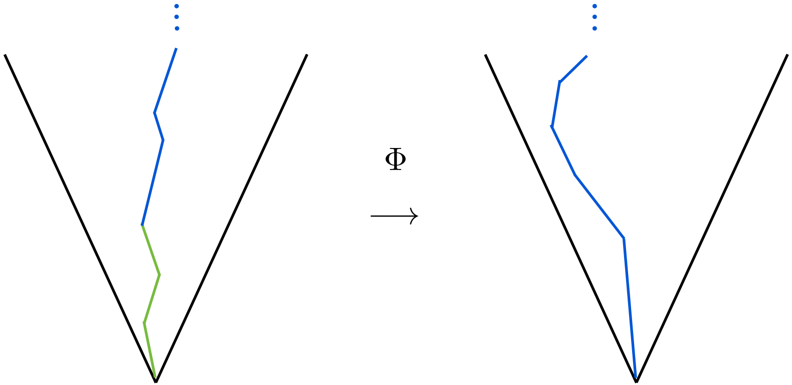


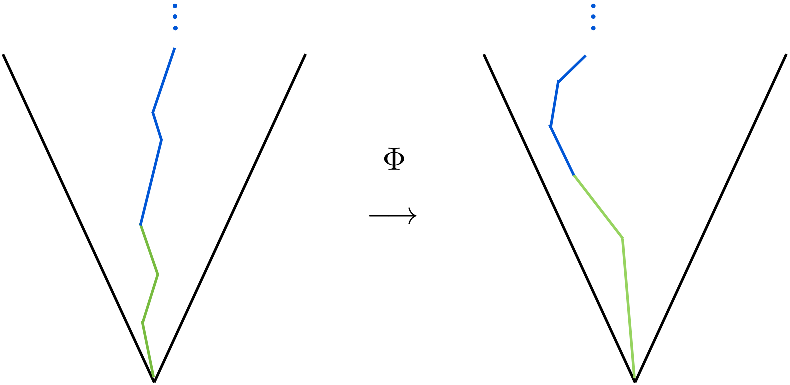


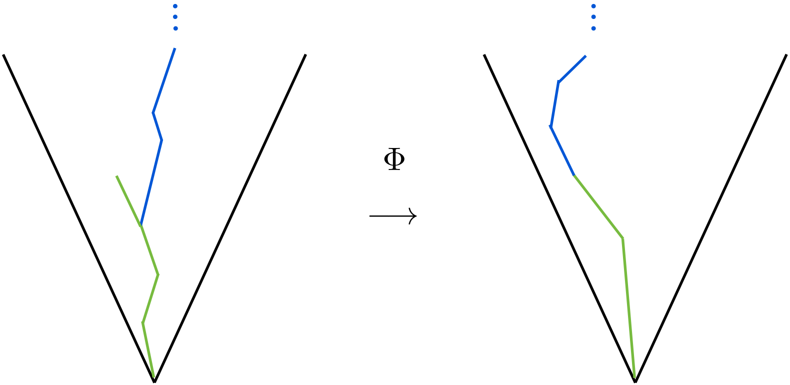
$$\xrightarrow{\Phi}$$

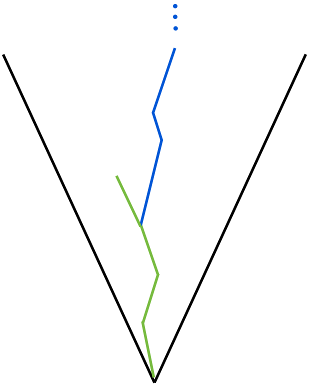



Φ

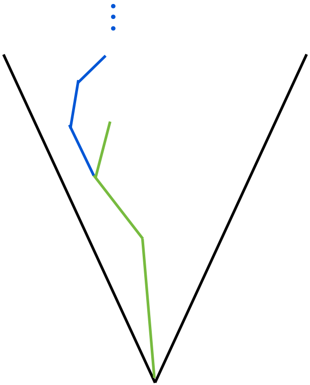

$\xrightarrow{\Phi}$

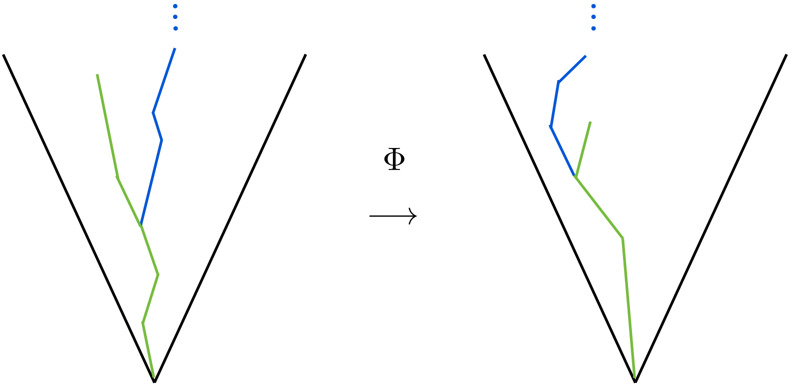

Φ

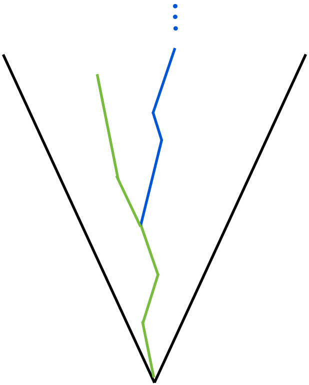

Φ



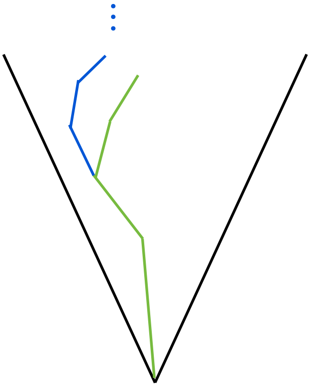
$$\xrightarrow{\Phi}$$

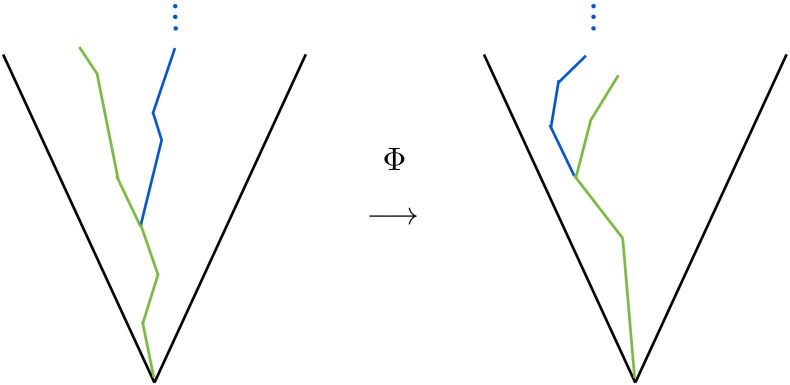


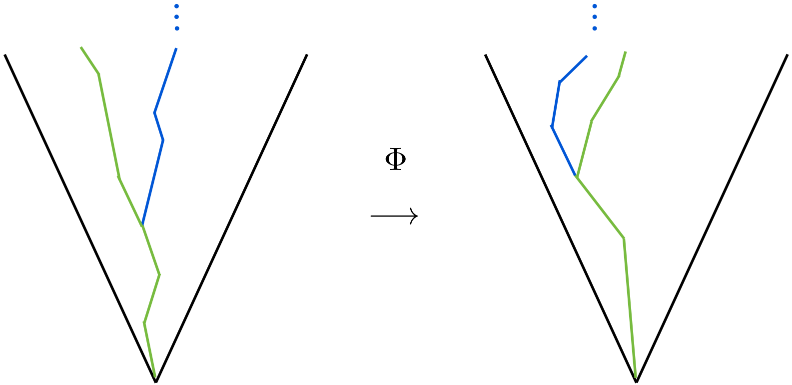


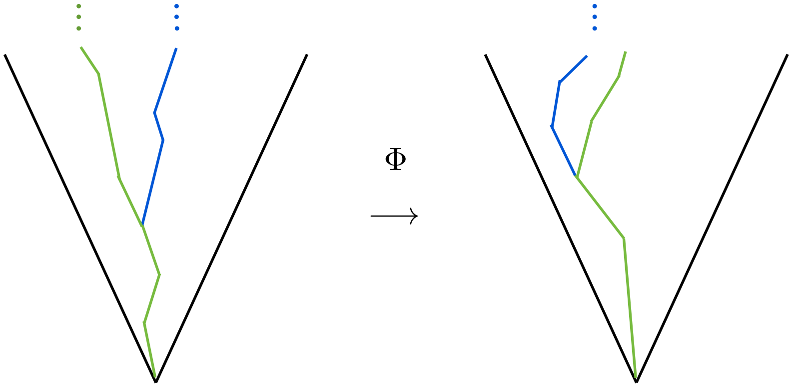


$$\xrightarrow{\Phi}$$

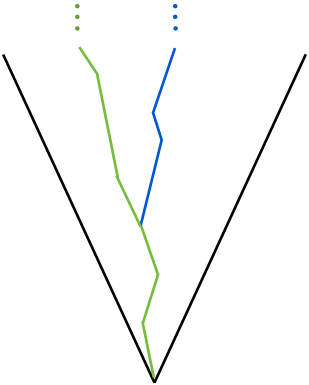



Φ

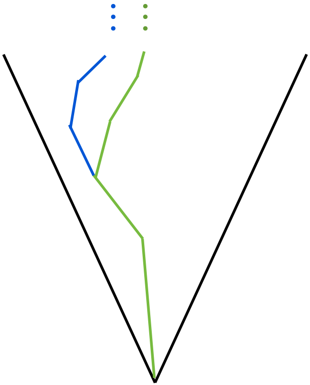

Φ

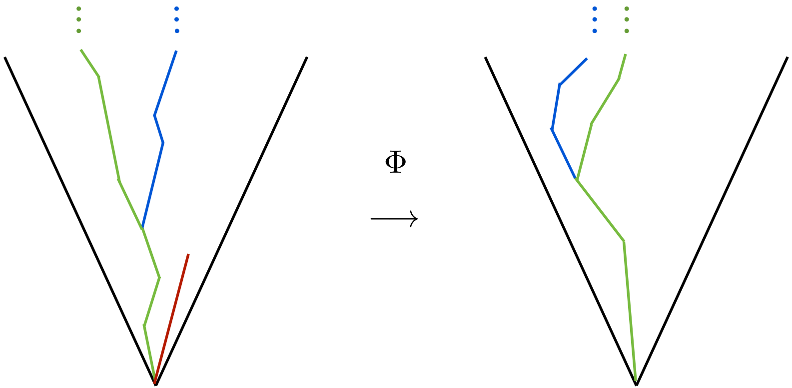


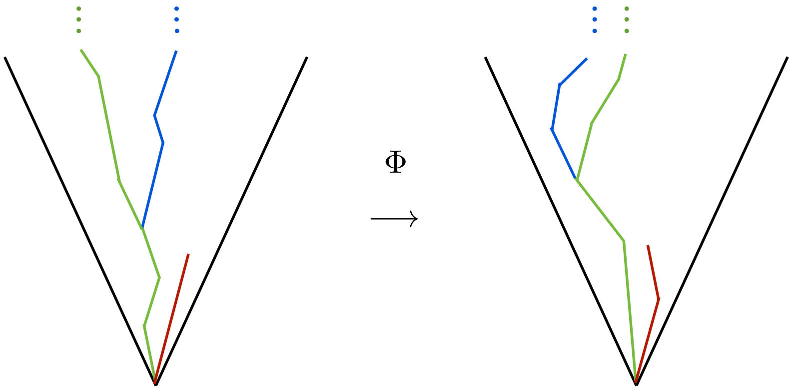
Φ



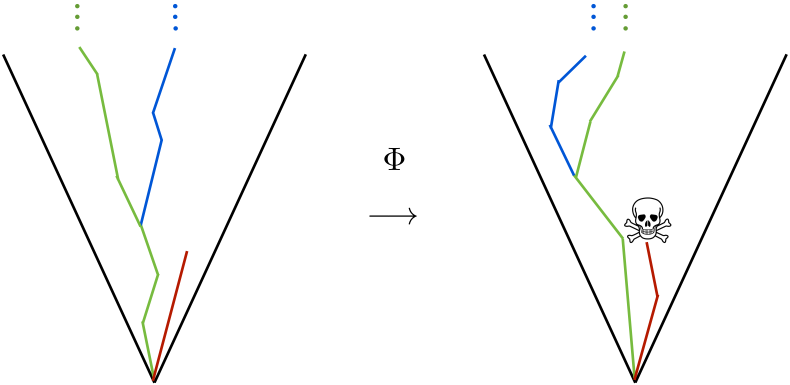
$$\xrightarrow{\Phi}$$

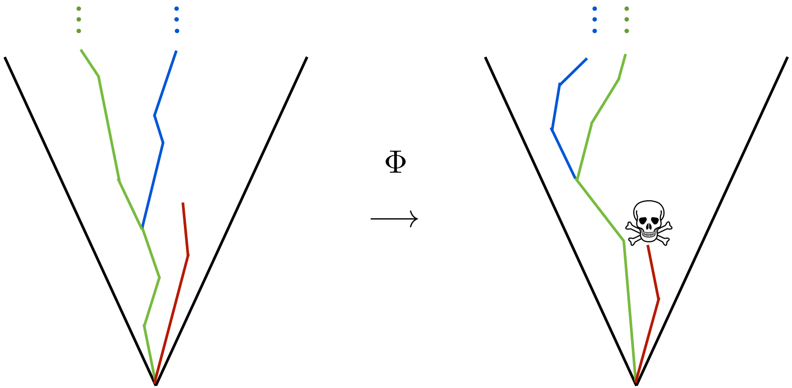


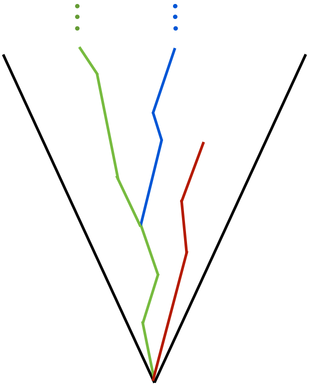

$\xrightarrow{\Phi}$



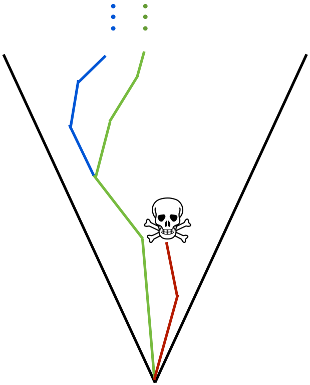
$\xrightarrow{\Phi}$

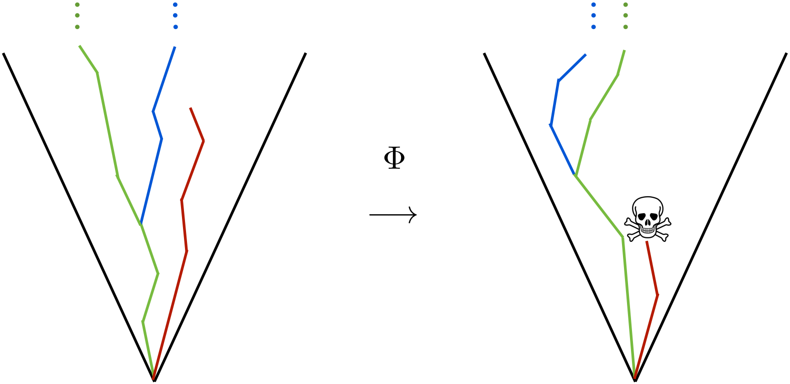

Φ

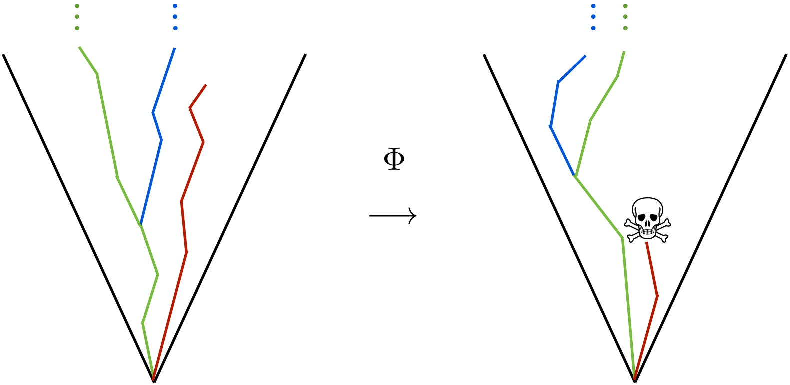

Φ

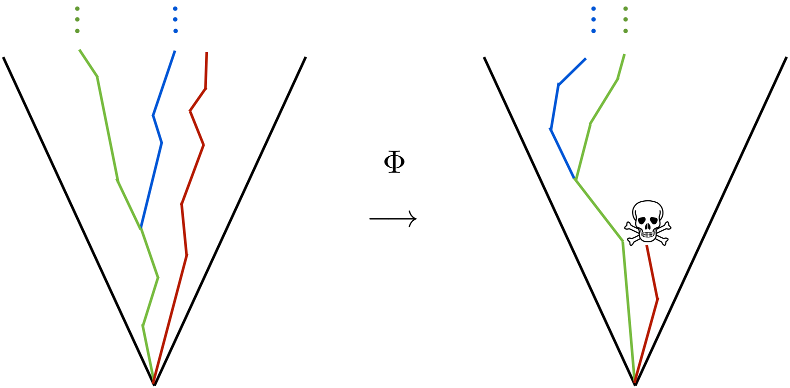


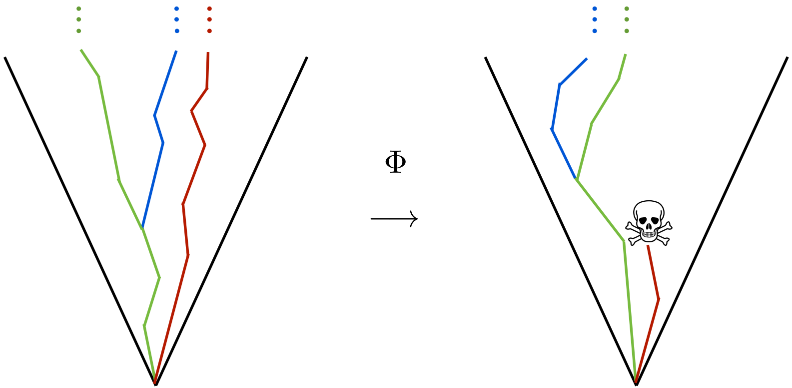
$$\xrightarrow{\Phi}$$

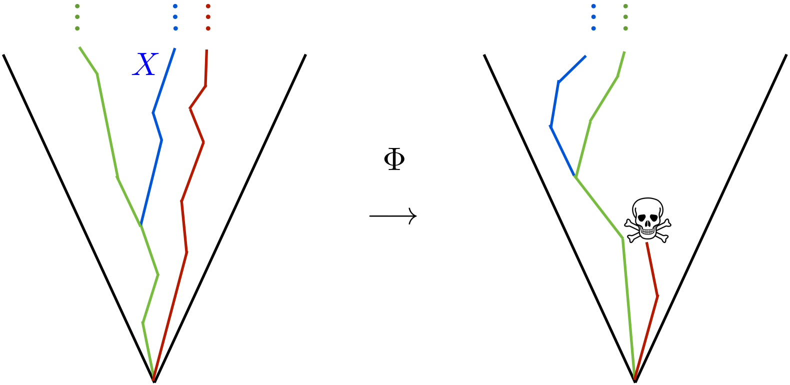


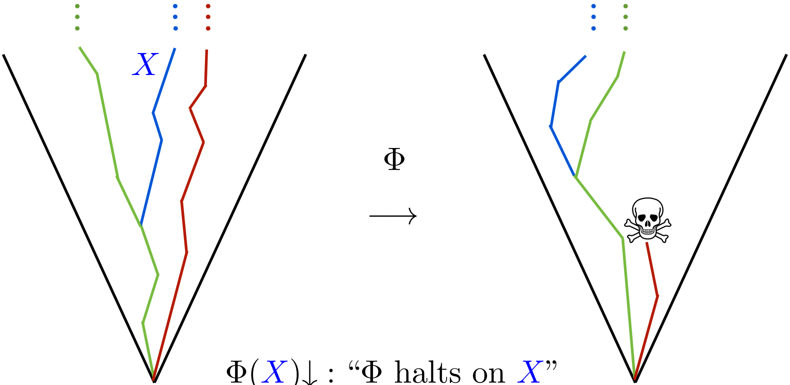

Φ


Φ

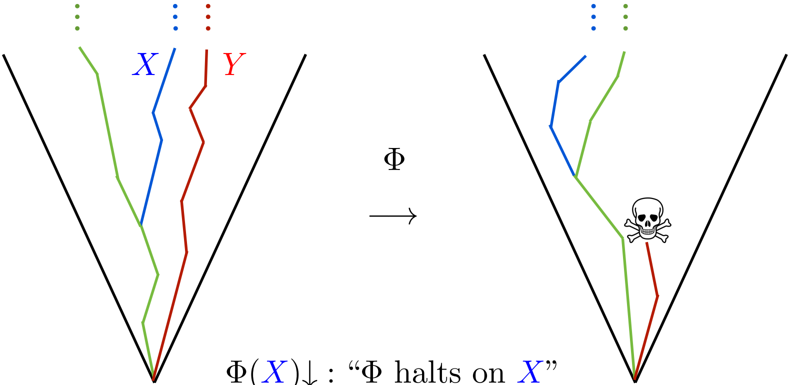

Φ


Φ

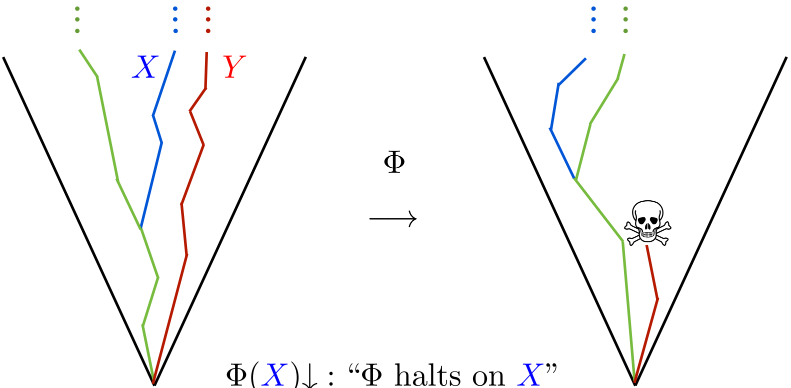

X
Φ



$$\Phi(X)\!\downarrow : \text{“}\Phi \text{ halts on } X\text{”}$$



$\Phi(X)\downarrow$: "Φ halts on X"



$\Phi(X)\downarrow$: "Φ halts on X"

$\Phi(Y)\uparrow$: "Φ diverges on Y"

Turing reducibility

If Φ is a Turing functional and $\Phi(B)\downarrow = A$, then we say that A is *Turing reducible* to B, denoted $A \leq_T B$.

"B computes A": $A \leq_T B$

One limitation: computing individual sequences

A sequence $A \in 2^\omega$ is *computable with positive probability* if

$$\lambda(\{X \in 2^\omega : A \leq_T X\}) > 0.$$

Theorem (Sacks)

A sequence is computable with positive probability if and only if it is computable.

Computing members of effectively closed classes

We cannot probabilistically compute any individual sequence that is not already Turing computable.

However, the situation is more interesting when we consider certain *collections* of sequences, namely effectively closed classes, also known as Π^0_1 classes.

Π^0_1 classes

- $\mathcal{P} \subseteq 2^\omega$ is a Π^0_1 class if its complement is effectively open, i.e., the complement is given by a computable enumeration of basic open sets.
- Equivalently, $\mathcal{P} \subseteq 2^\omega$ is a Π^0_1 if it is the collection of infinite paths through through a computable tree (a subset of $2^{<\omega}$ that is closed downwards under $\preceq$).
- We can also define a Π^0_1 class to be the collection of infinite paths through a tree whose complement is computably enumerable.

Computationally powerful random sequences

It is worth noting that some Martin-Löf random sequences can compute a member of every Π^0_1 class.

- $X \in 2^\omega$ has *PA degree* if X computes a consistent completion of Peano arithmetic.
- Every sequence of PA degree computes a member of every Π^0_1 class.
- Some Martin-Löf random sequences have PA degree.

Dichotomy: A Martin-Löf random sequence has PA degree if and only if it computes the halting set $K = \{e : \phi_e(e)\downarrow\}$.

However, by Sack's theorem, only measure zero many Martin-Löf random sequences have this property.

3. Negligible Π^0_1 classes

When probabilistic computation fails

$\mathcal{A} \subseteq 2^\omega$ is *negligible* if we cannot compute some member of $\mathcal{A}$ with positive probability.

That is,

$$\lambda\big(\{X \in 2^\omega : (\exists Y \in \mathcal{A})[Y \leq_T X]\}\big) = 0.$$

We can also provide a useful equivalent formulation of negligibility in terms of left-c.e. semi-measures.

Left-c.e. semi-measures

A *semi-measure* $\rho : 2^{<\omega} \to [0,1]$ satisfies

- $\rho(\epsilon) = 1$ and
- $\rho(\sigma) \geq \rho(\sigma 0) + \rho(\sigma 1)$ for every $\sigma \in 2^{<\omega}$.

A semi-measure ρ is *left-c.e.* if each value $\rho(\sigma)$ is the limit of a non-decreasing computable sequence of rationals, uniformly in σ.

Semi-measures and Turing functionals

For $\sigma \in 2^{<\omega}$, we define $\Phi^{-1}(\sigma) := \{X \in 2^\omega : \exists n\, (X\restriction n, \sigma) \in S_\Phi\}$.

Proposition (Levin)

(i) *If Φ is a Turing functional, then λ_Φ, defined by*

$$\lambda_\Phi(\sigma) = \lambda(\Phi^{-1}(\sigma))$$

for every $\sigma \in 2^{<\omega}$, is a left-c.e. semi-measure.

(ii) *For every left c.e. semi-measure ρ, there is a Turing functional Φ such that $\rho = \lambda_\Phi$.*

A universal semi-measure

Levin proved the existence of a universal left-c.e. semi-measure.

A left-c.e. semi-measure M is *universal* if for every left-c.e. semi-measure ρ, there is some $c \in \omega$ such that

$$\rho(\sigma) \leq c \cdot M(\sigma)$$

for every $\sigma \in 2^{<\omega}$.

Defining negligibility in terms of semi-measures

Let M be a universal left-c.e. semi-measure.

Let $\overline{M}$ be the largest measure such that $\overline{M} \leq M$, which can be seen as a universal measure.

Proposition

$\mathcal{S} \subseteq 2^\omega$ is negligible if and only if $\overline{M}(\mathcal{S}) = 0$.

Proof idea: Use the correspondence between Turing functionals and left-c.e. semi-measures and the fact that M multiplicatively dominates all left-c.e. semi-measures to show

$$\overline{M}(\mathcal{S}) = 0 \text{ if and only if } \lambda\Big(\bigcup_{i \in \omega} \Phi_i^{-1}(\mathcal{S})\Big) = 0$$

for every $\mathcal{S} \subseteq 2^\omega$ (where $(\Phi_i)_{i \in \omega}$ is an effective enumeration of all Turing functionals).

Members of negligible classes

A few observations:

- If a Π^0_1 class contains a computable member, it cannot be negligible.
- Moreover, if a Π^0_1 class contains a Martin-Löf random member, it cannot be negligible, since any Π^0_1 class with a random member must have positive Lebesgue measure.

These two facts are subsumed by the following result:

Proposition (Bienvenu, Porter)

Let $\mathcal{P}$ be a negligible Π^0_1 class. Then for every computable measure μ, $\mathcal{P}$ contains no $X \in \mathsf{MLR}_\mu$.

Does the converse hold?

Suppose that $\mathcal{P}$ is a Π^0_1 class such that $\mathcal{P} \cap \mathsf{MLR}_\mu = \emptyset$ for every computable measure μ.

Does it follow that $\mathcal{P}$ is negligible?

Does the converse hold?

Suppose that $\mathcal{P}$ is a Π^0_1 class such that $\mathcal{P} \cap \mathsf{MLR}_\mu = \emptyset$ for every computable measure μ.

Does it follow that $\mathcal{P}$ is negligible? No.

Does the converse hold?

Suppose that $\mathcal{P}$ is a Π^0_1 class such that $\mathcal{P} \cap \mathsf{MLR}_\mu = \emptyset$ for every computable measure μ.

Does it follow that $\mathcal{P}$ is negligible? No.

Theorem (Bienvenu, Porter)

There is a non-negligible Π^0_1 class $\mathcal{P}$ such that $\mathcal{P} \cap \mathsf{MLR}_\mu = \emptyset$ for every computable measure μ.

Computing members of negligible Π^0_1 classes

As mentioned above, some Martin-Löf random sequences can compute a member of every Π^0_1 class (but only measure zero many random sequences have this property).

If we consider a slightly stronger notion of randomness known as *weak 2-randomness*, we get a stronger result.

Weak 2-randomness is the notion of randomness that results from replacing Martin-Löf tests with *generalized Martin-Löf tests*: a collection $(\mathcal{U}_i)_{i\in\omega}$ of uniformly effectively open subsets of 2^ω such that $\lambda(\mathcal{U}_i) \to 0$.

Theorem (Bienvenu, Porter)

If $X \in 2^\omega$ is weakly 2-random, then X cannot compute any member of a negligible Π^0_1 class.