Randomness, probability, and computation

Christopher P. Porter
Université Paris 7
LIAFA

Joint work with Laurent Bienvenu and Antoine Taveneaux

Midlands Logic Seminar
University of Birmingham
29 November 2013

Introduction

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

Outline of the talk

1. Some basic algorithmic randomness
2. Probabilistic Turing computation
3. Negligible Π^0_1 classes
4. Deep Π^0_1 classes

1. Some basic algorithmic randomness

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.

(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 specifiable regularities".

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

In other words, if a sequence contains some "effectively specifiable 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 effective unpredictability
- in terms of effective incompressibility
- in terms of effective typicality $\Leftarrow$

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?

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 uniform sequence $(U_i)_{i\in\omega}$ of 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

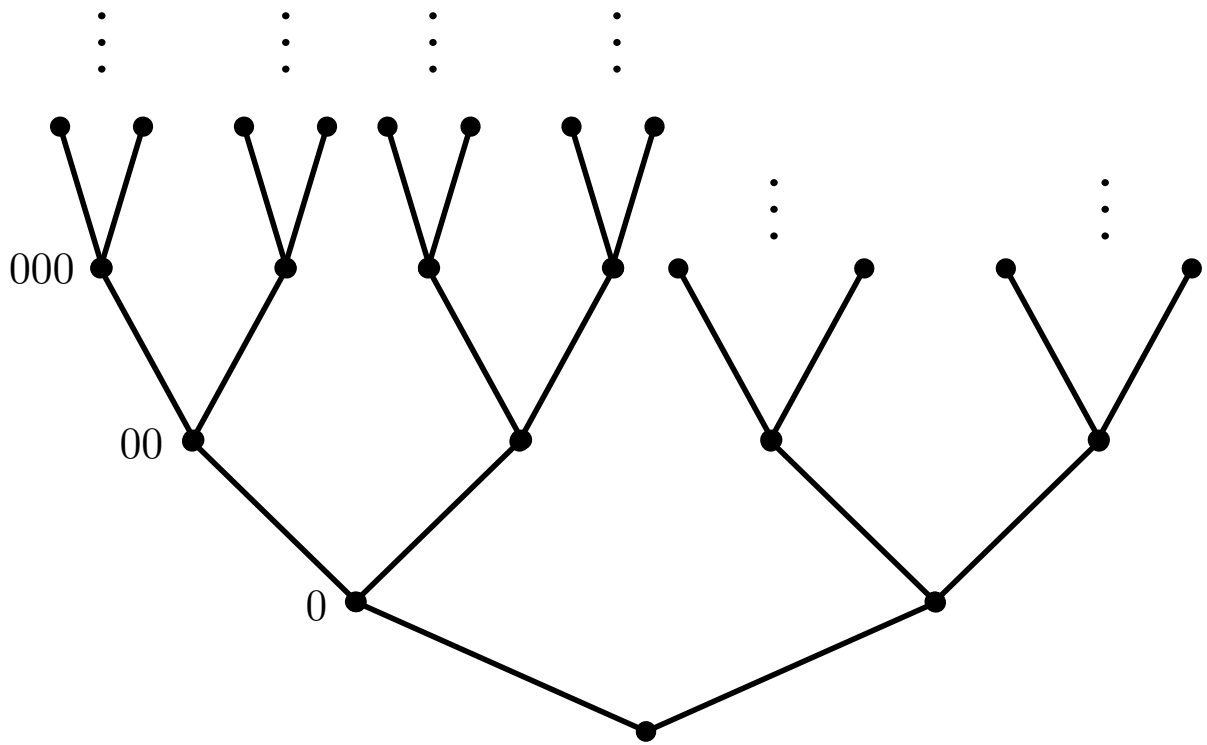
$$\lambda([\![\sigma]\!]) = 2^{-|\sigma|}.$$

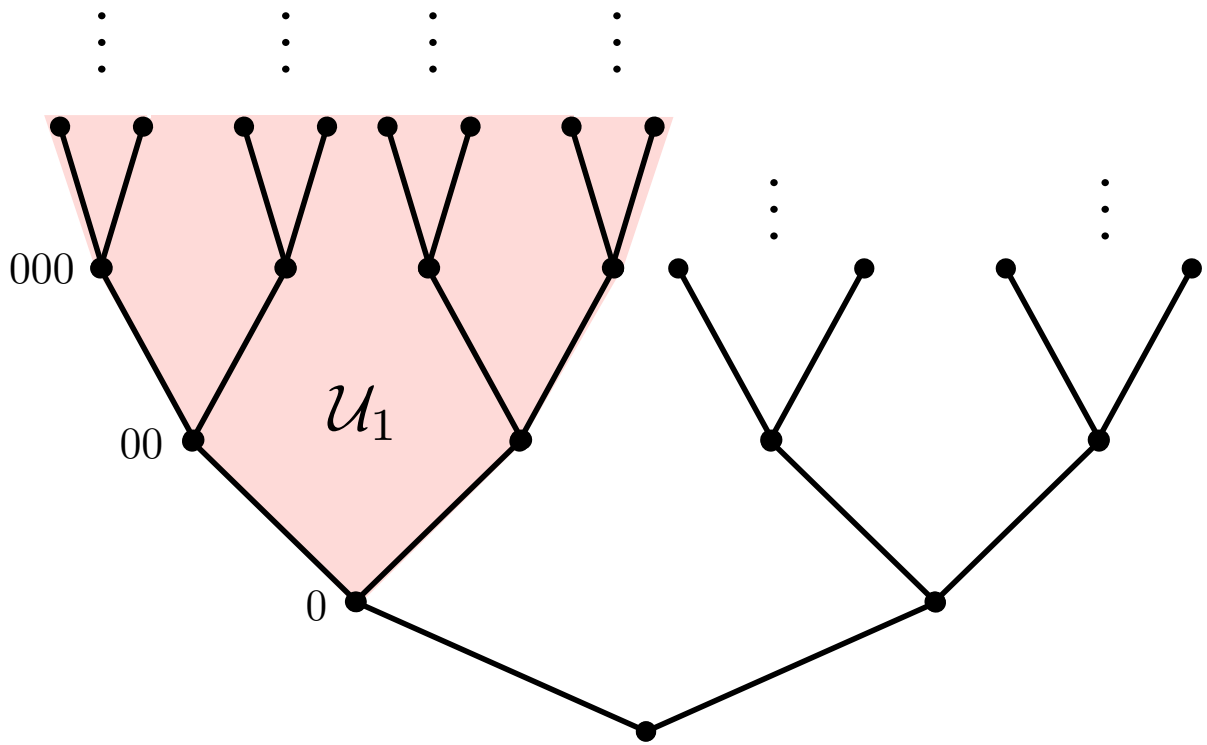
Thus we can consider a Martin-Löf test to be a collection $(\mathcal{U}_i)_{i\in\omega}$ of uniformly effectively open subsets of 2^ω such that

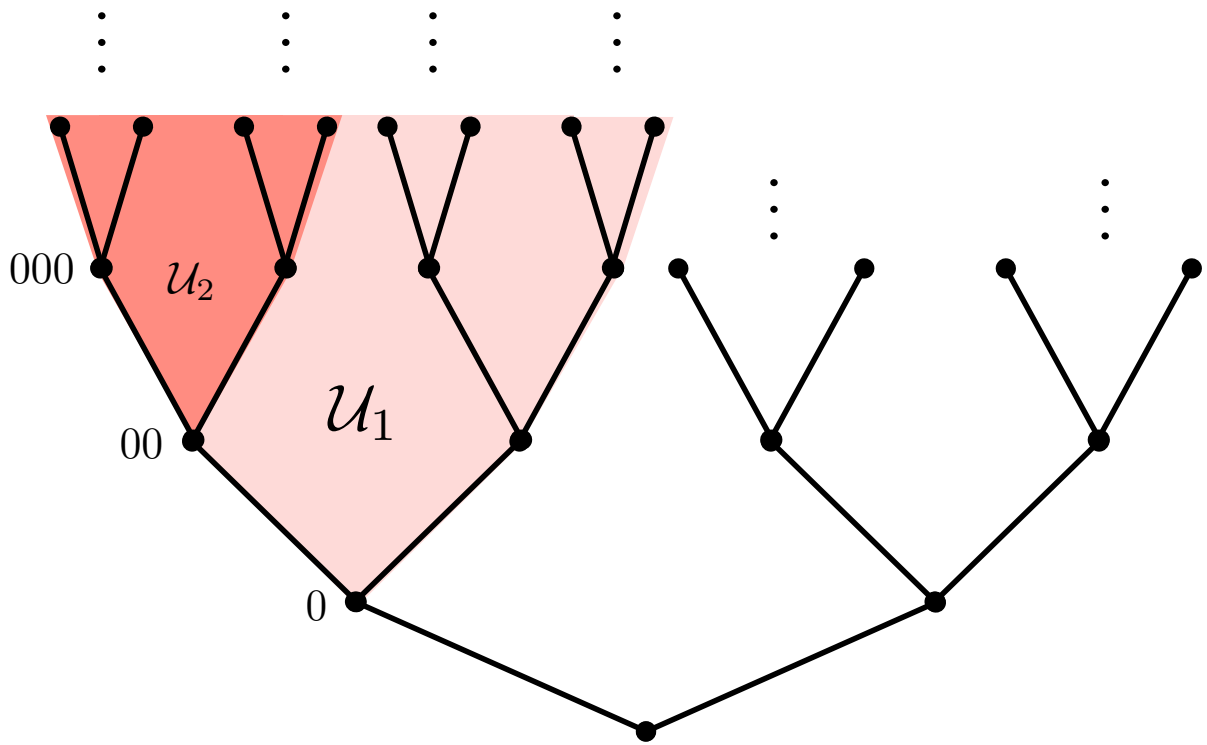
$$\lambda(\mathcal{U}_i) \leq 2^{-i}$$

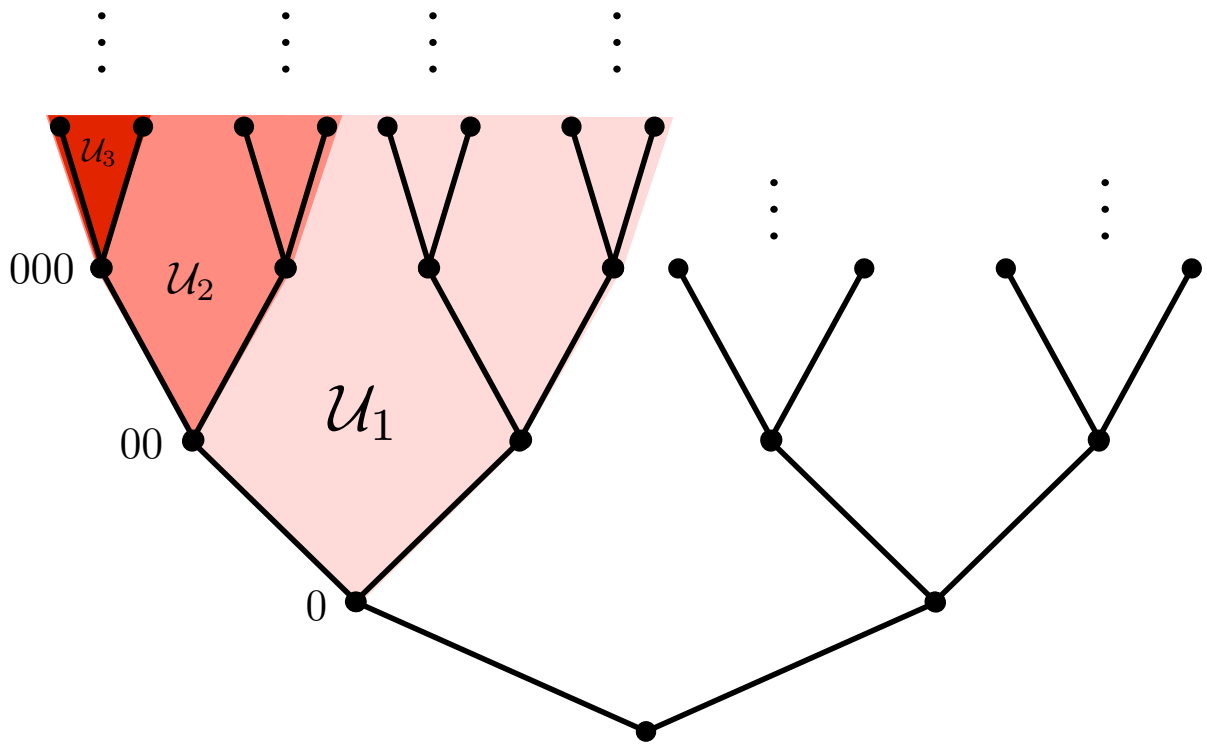
for every i.

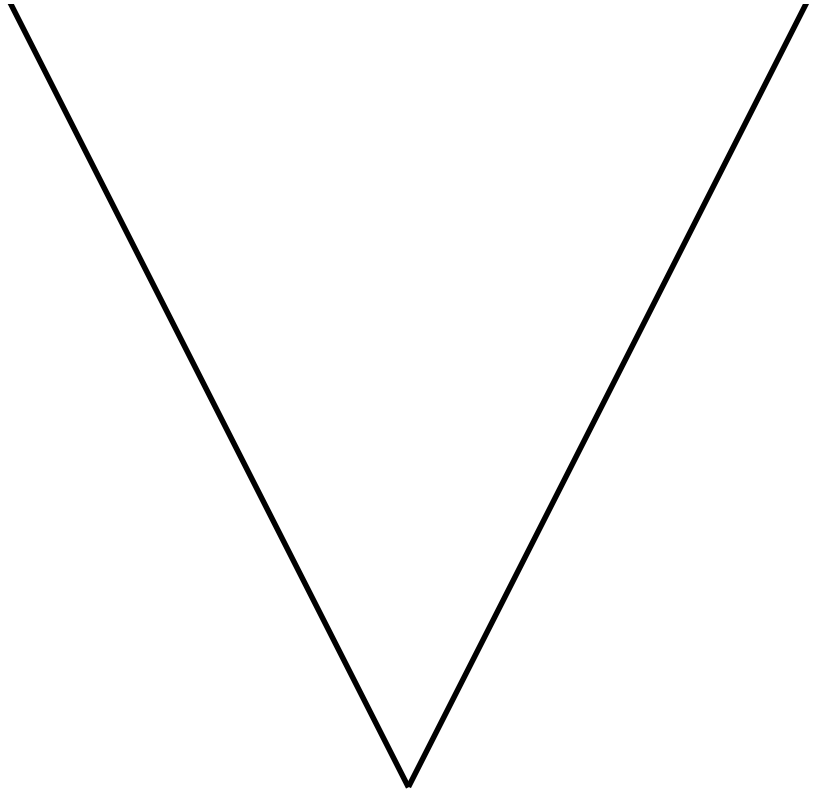
Moreover, X passes the test $(\mathcal{U}_i)_{i\in\omega}$ if $X \notin \bigcap_i \mathcal{U}_i$.

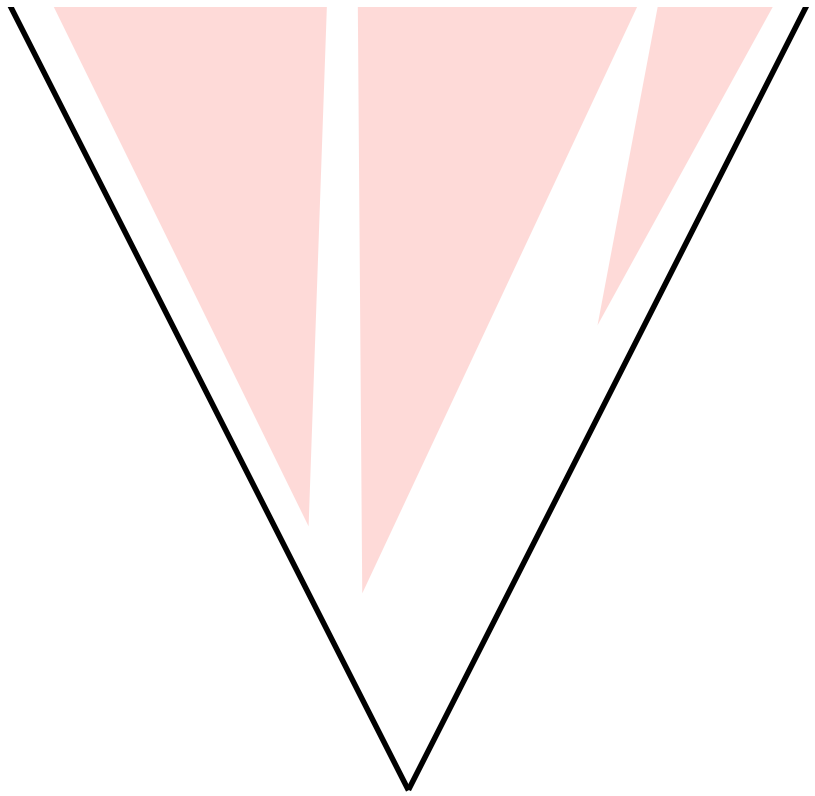

000
00
0

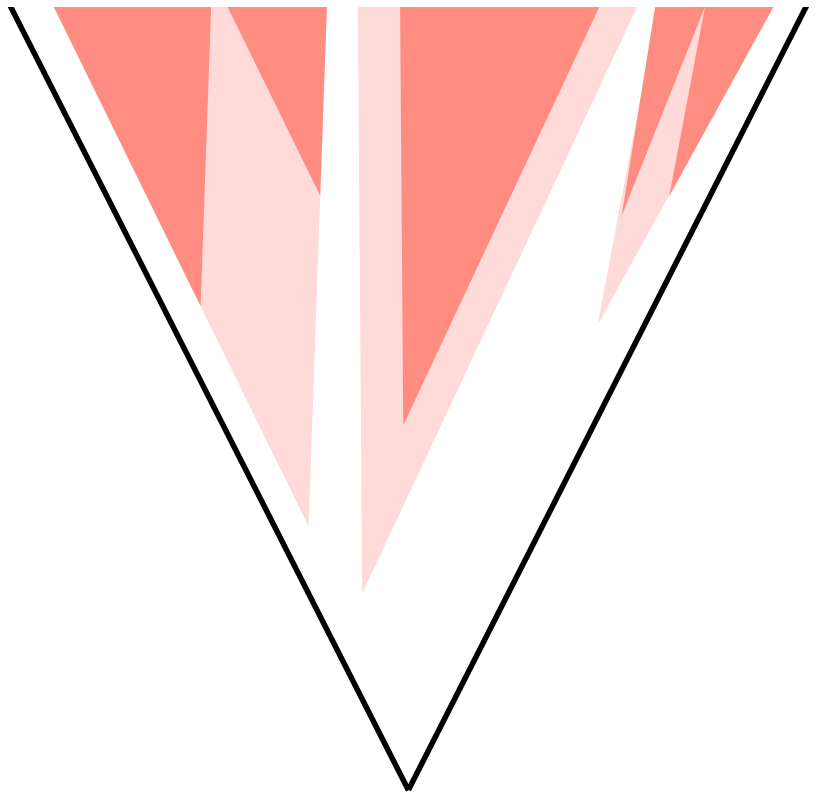

000
00
0
$\mathcal{U}_1$

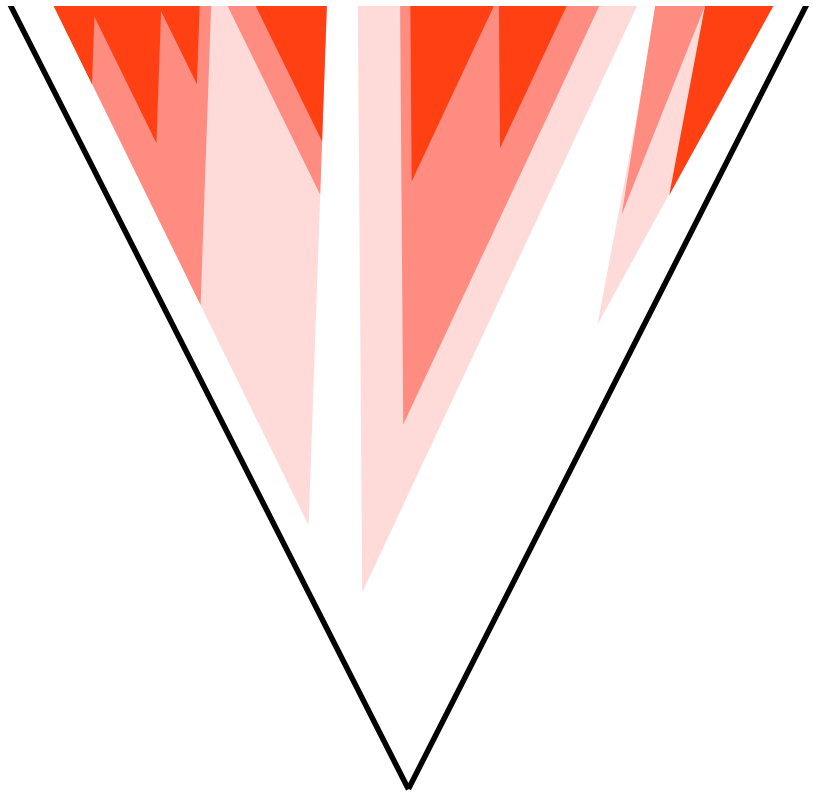

000
00
0
$\mathcal{U}_2$
$\mathcal{U}_1$

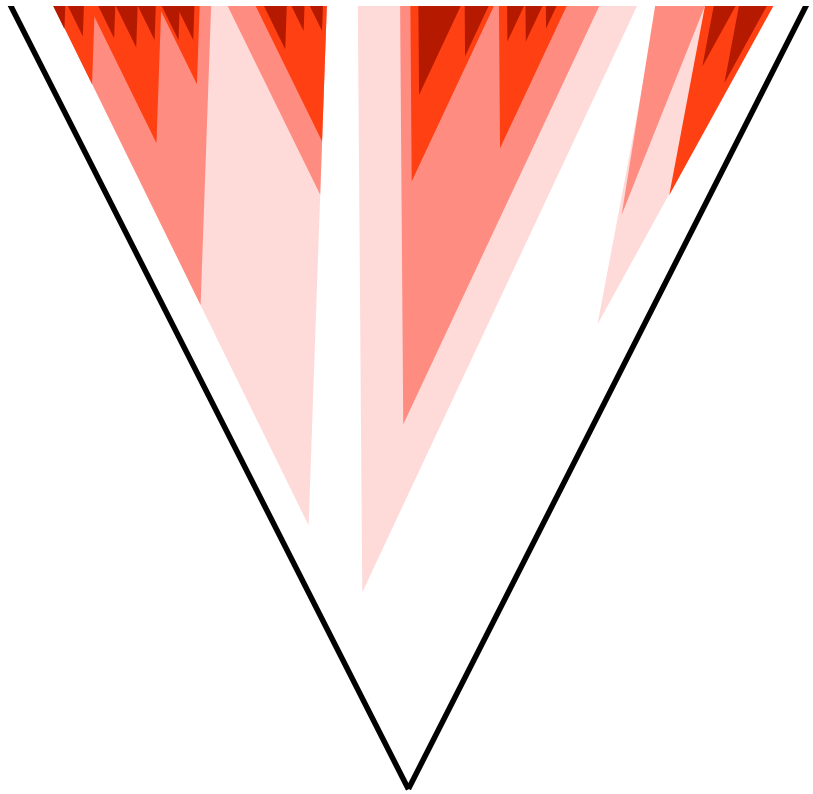

$\mathcal{U}_3$
$\mathcal{U}_2$
$\mathcal{U}_1$
000
00
0











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 $i \in \omega$.

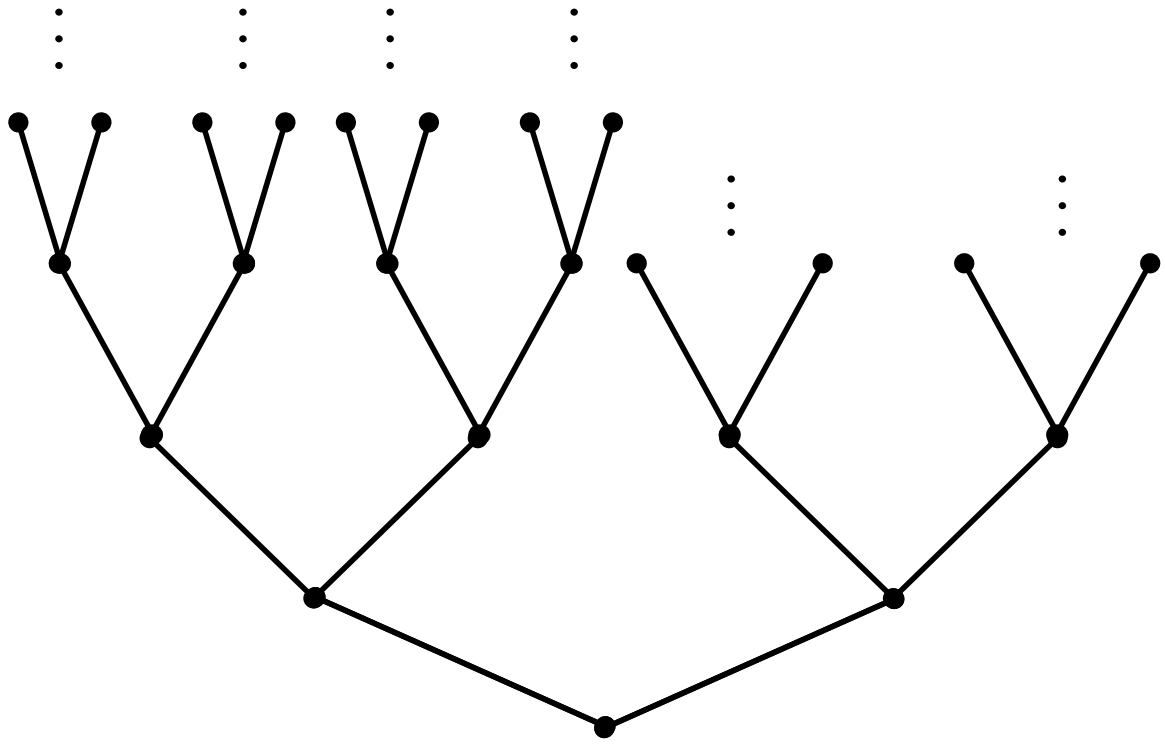
From now on we will write $\mu(\sigma)$ instead of $\mu([\![\sigma]\!])$.


LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
E PLURIBUS UNUM
QUARTER DOLLAR

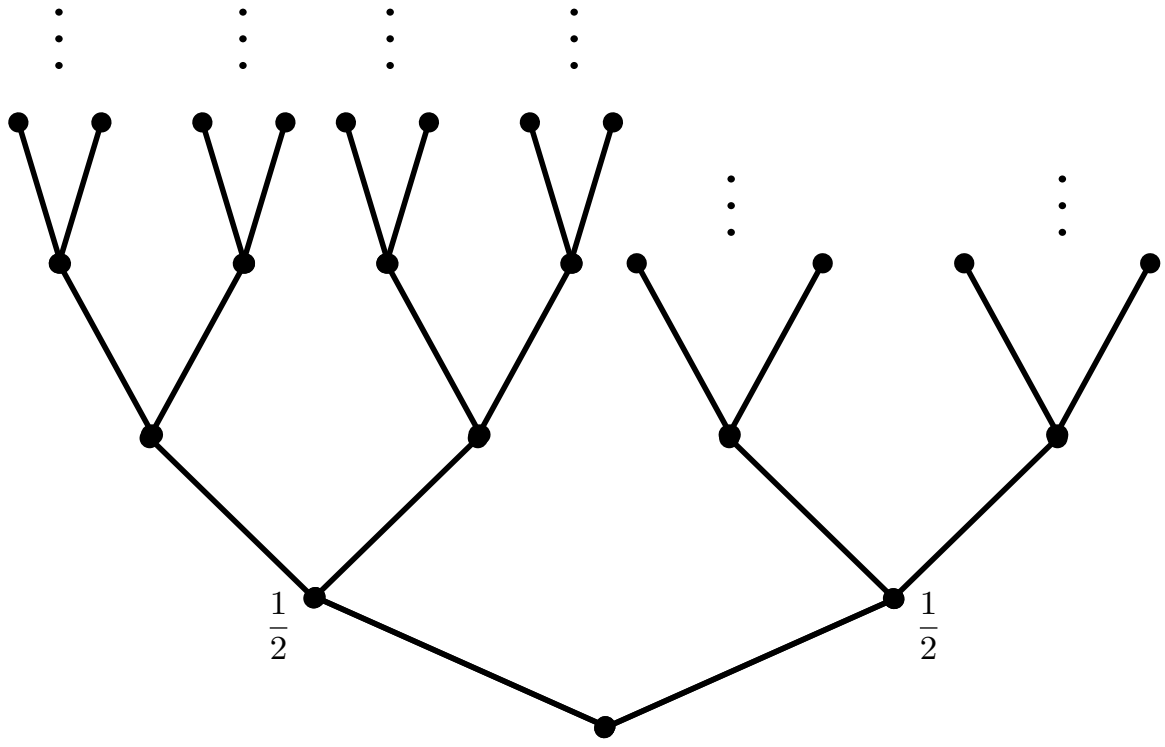

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
E PLURIBUS UNUM
QUARTER DOLLAR

$$\frac{1}{2}$$

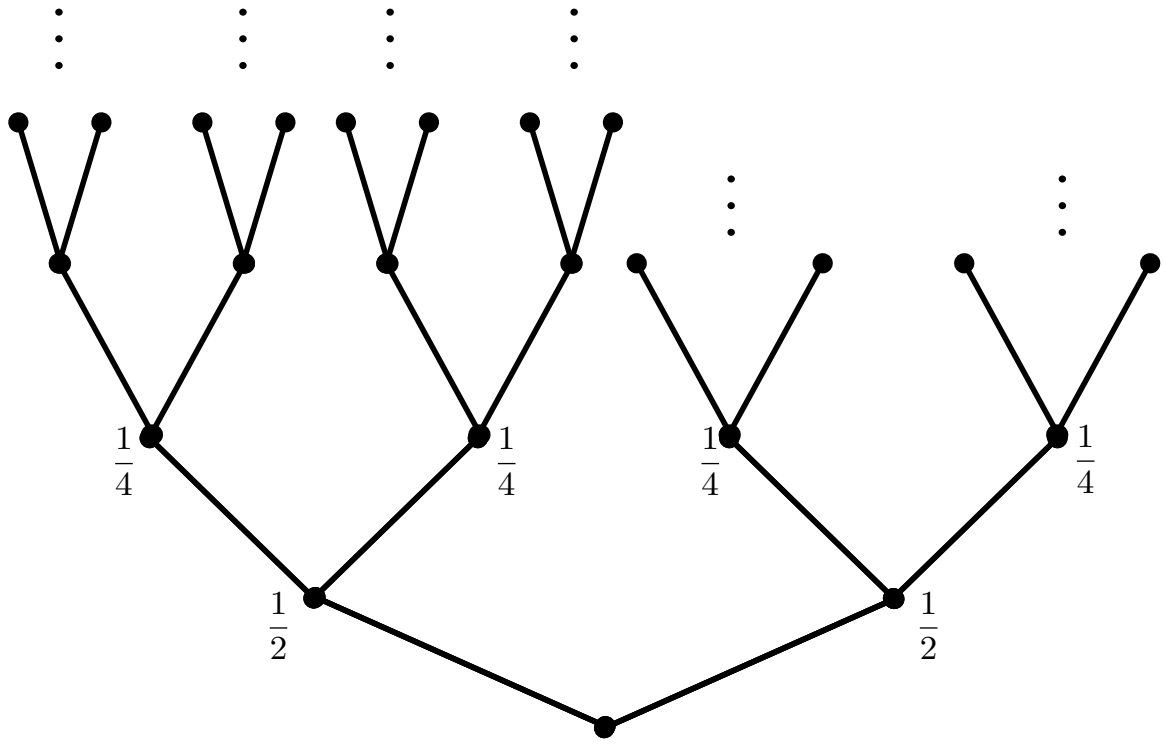
$\frac{1}{2}$ $\frac{1}{2}$



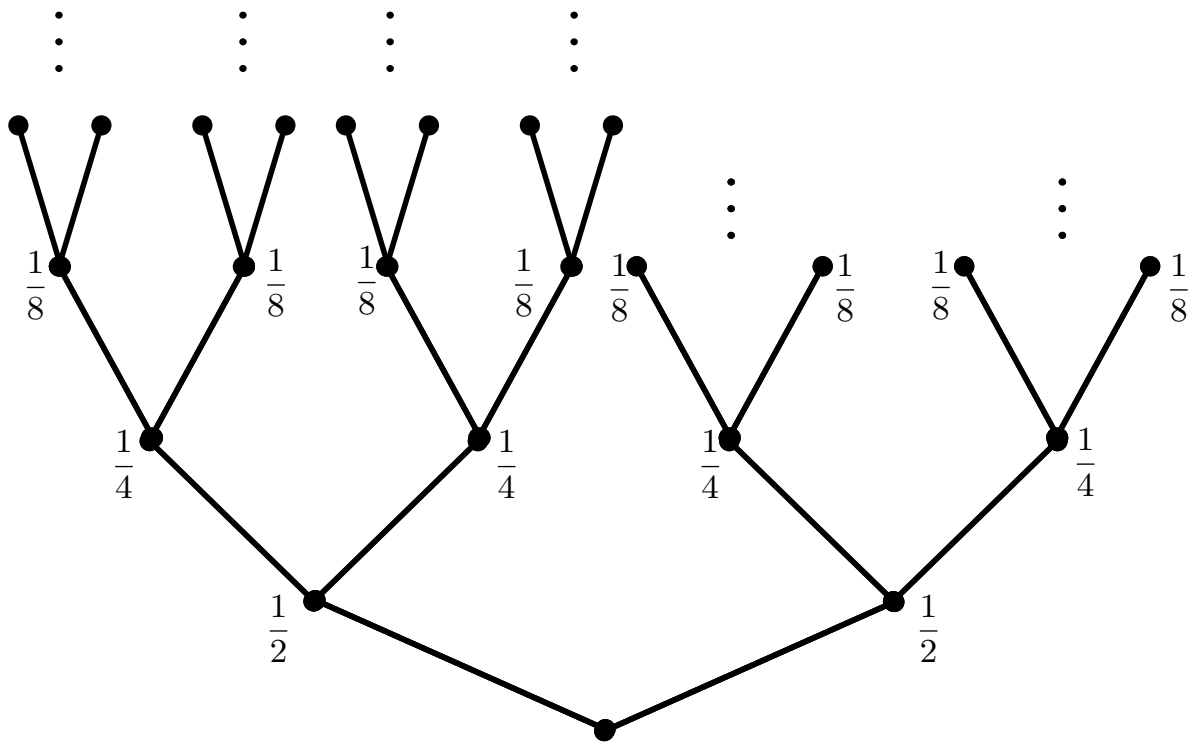
$\frac{1}{2}$ $\frac{1}{2}$



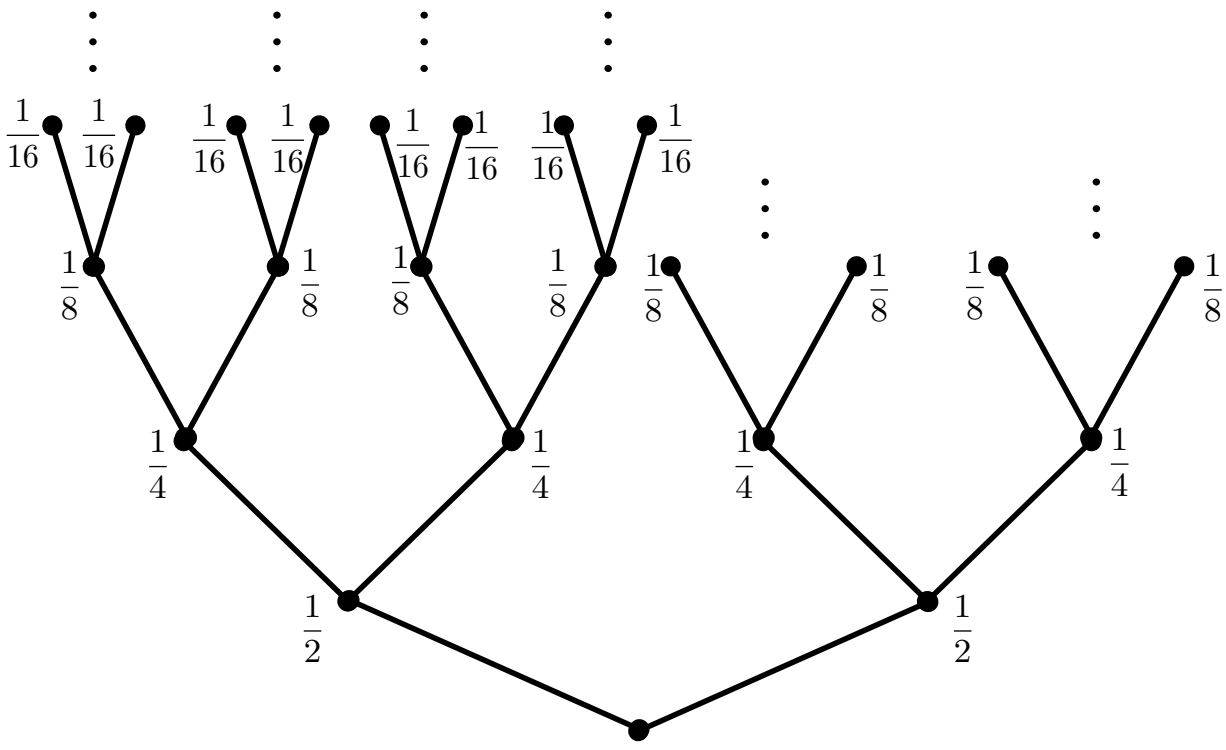
$\frac{1}{2}$ $\frac{1}{2}$



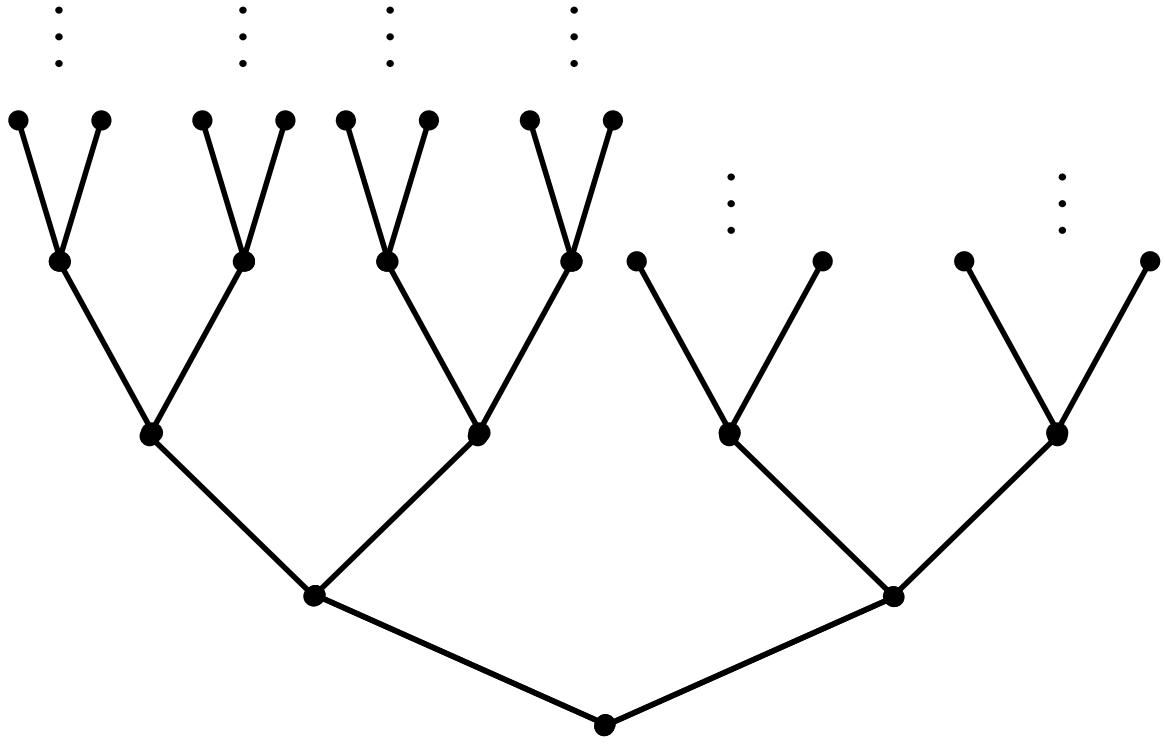
$\frac{1}{2}$ $\frac{1}{2}$



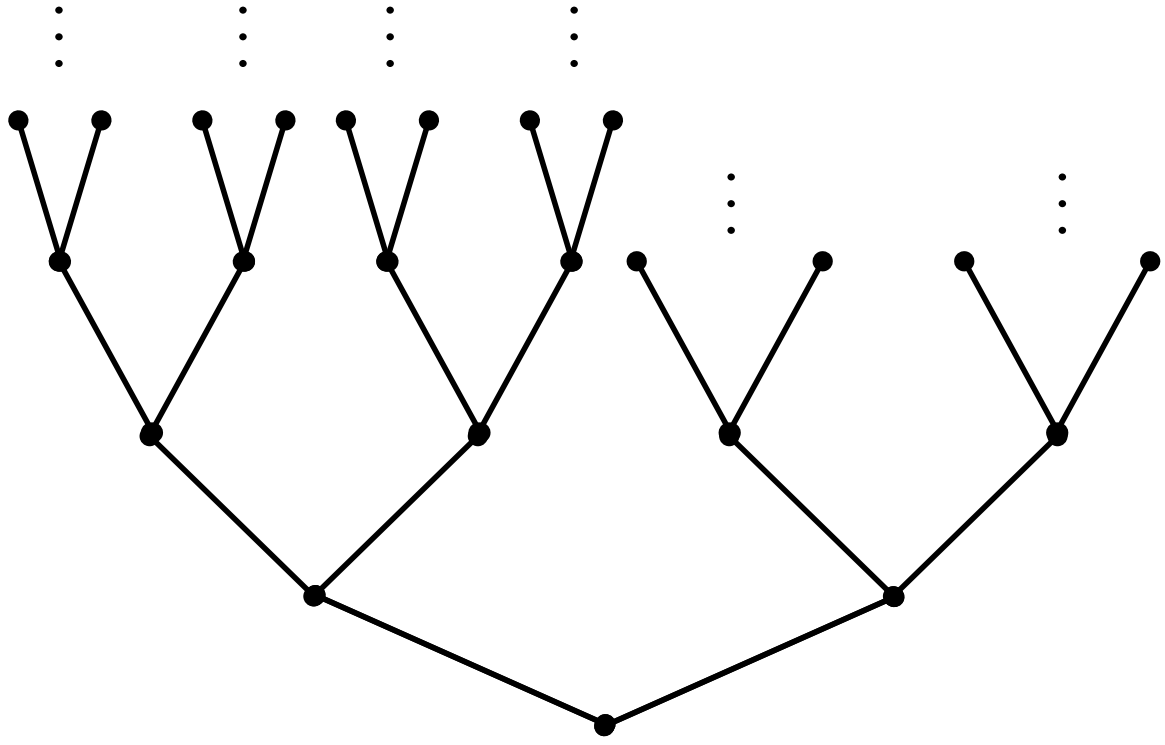
$\frac{1}{2}$ $\frac{1}{2}$



$\frac{1}{2}$ $\frac{1}{2}$

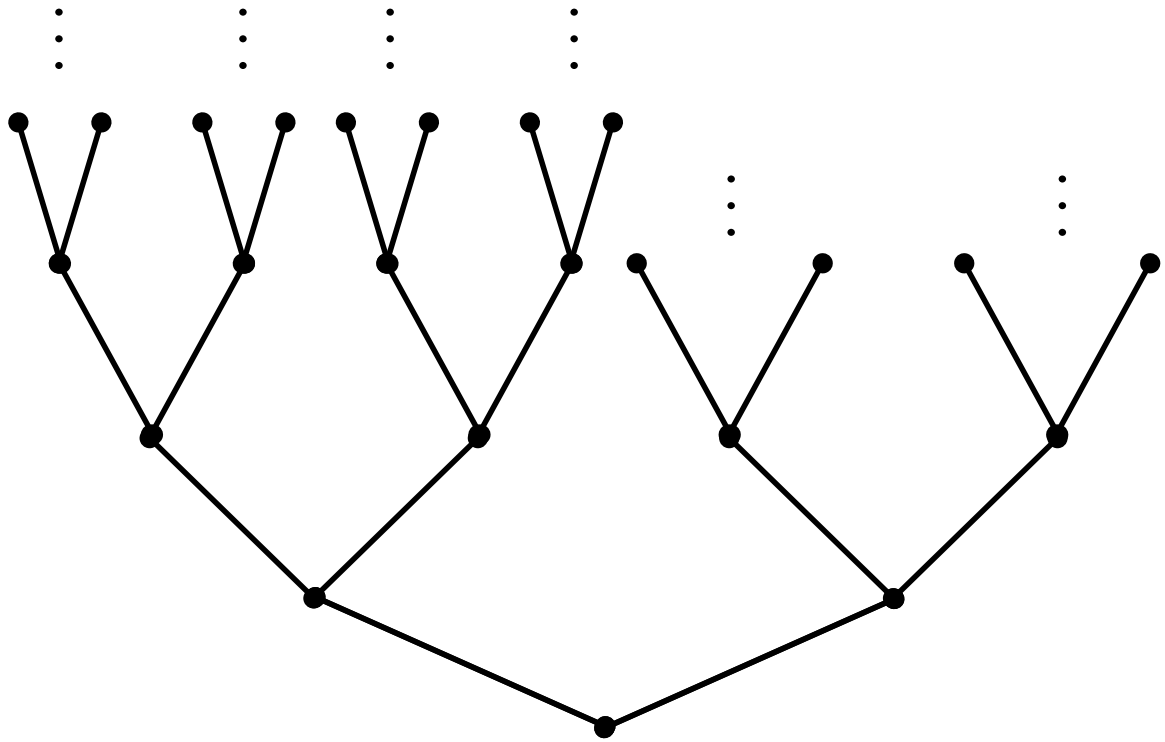



LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR

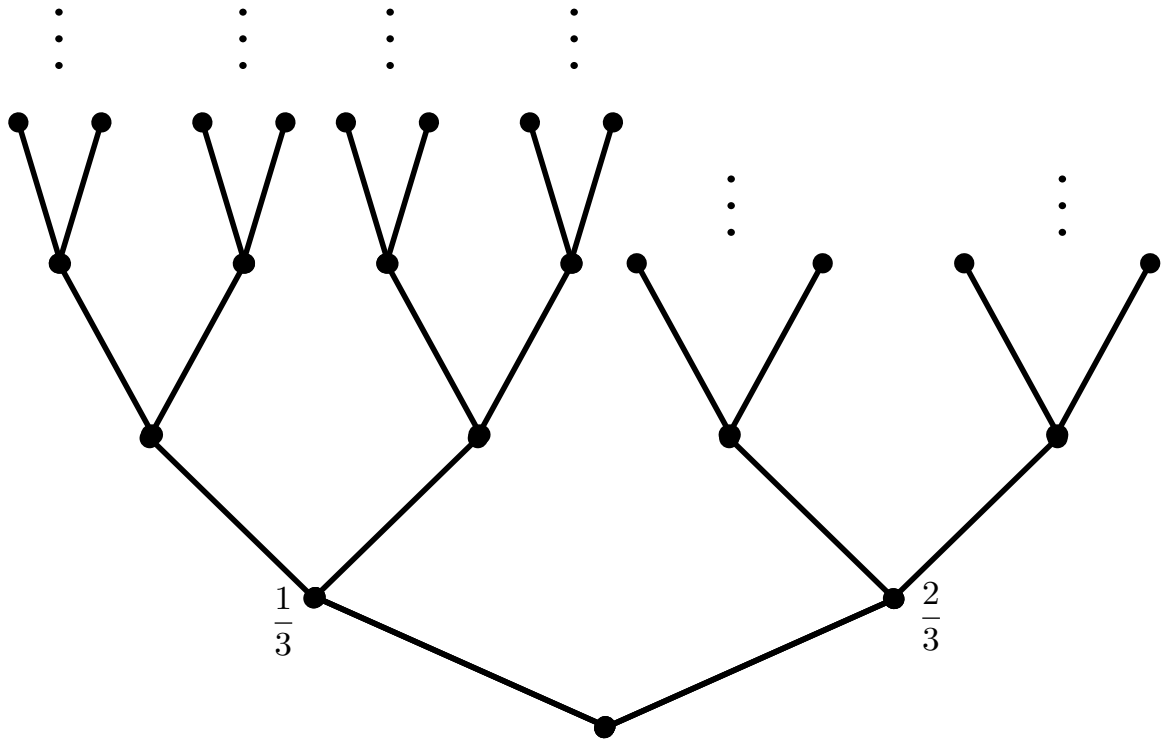



LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
E PLURIBUS UNUM
QUARTER DOLLAR

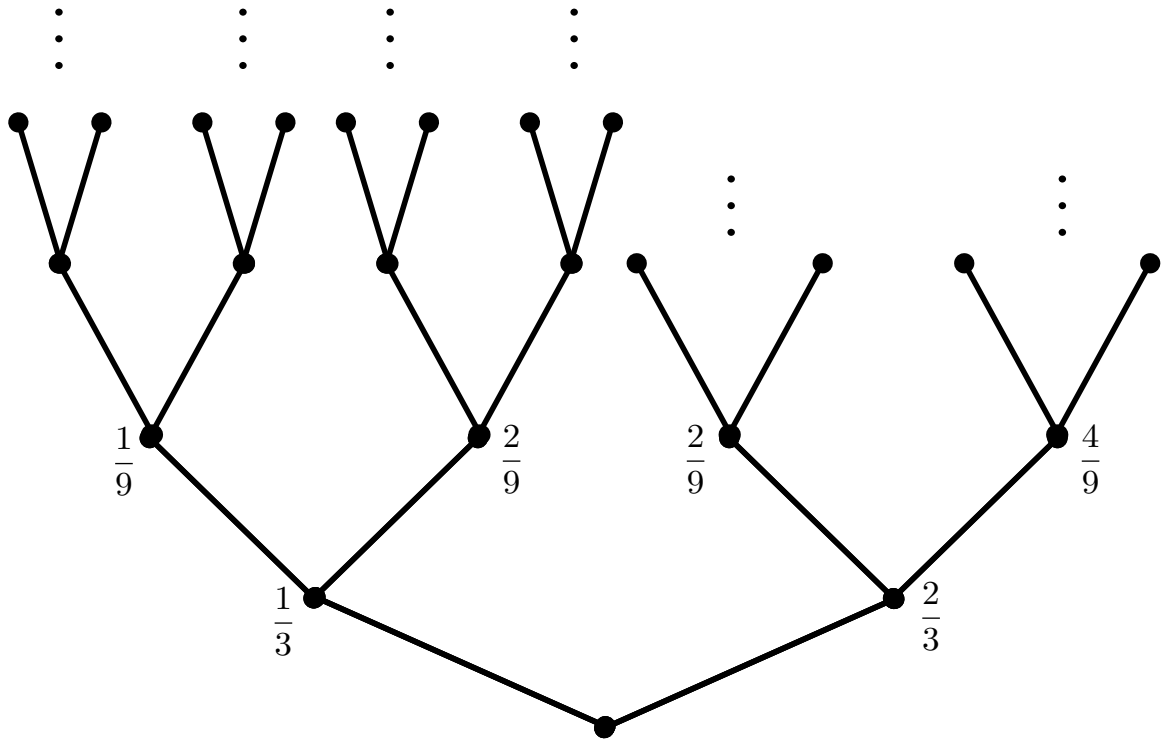
$$\frac{1}{3}$$



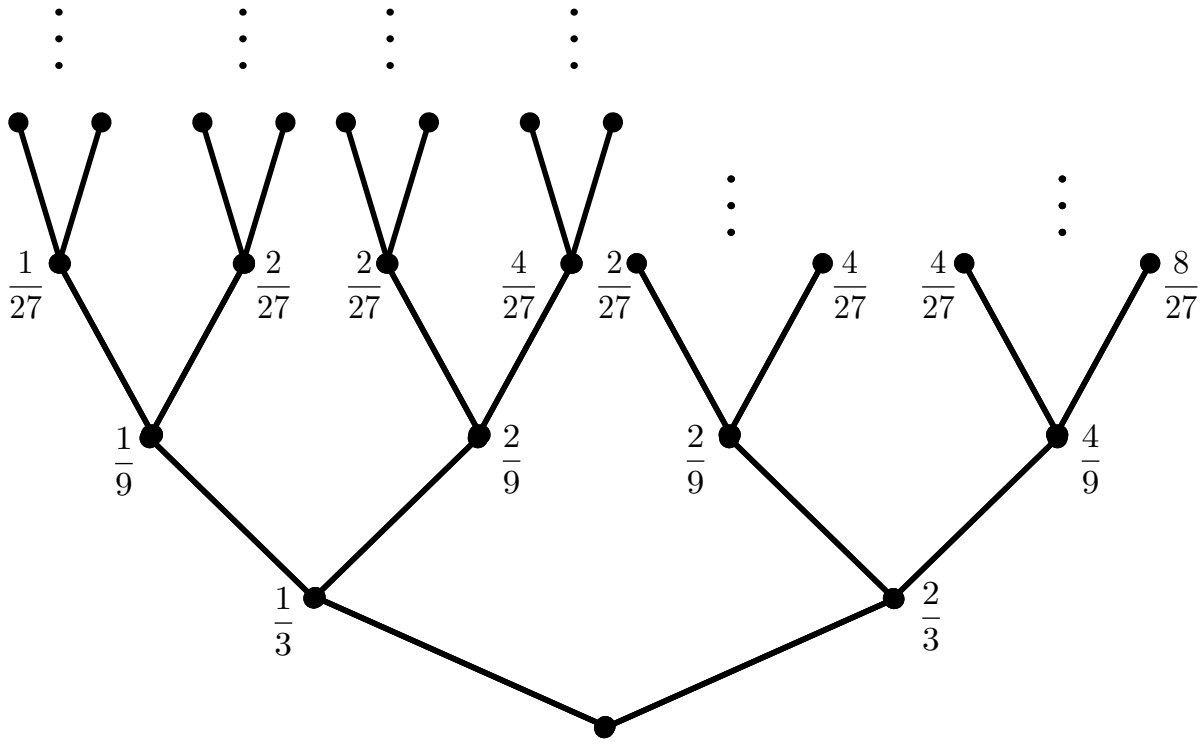
$\frac{1}{3}$ $\frac{2}{3}$



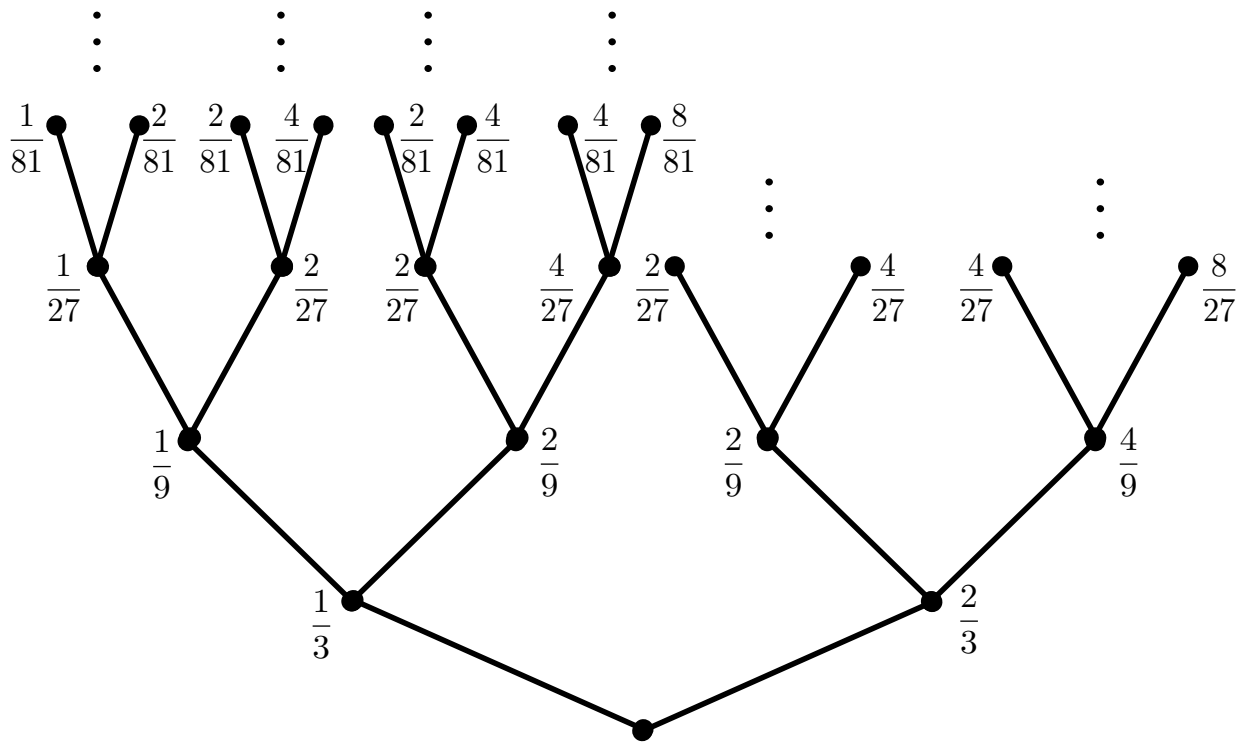
$\frac{1}{3}$ $\frac{2}{3}$



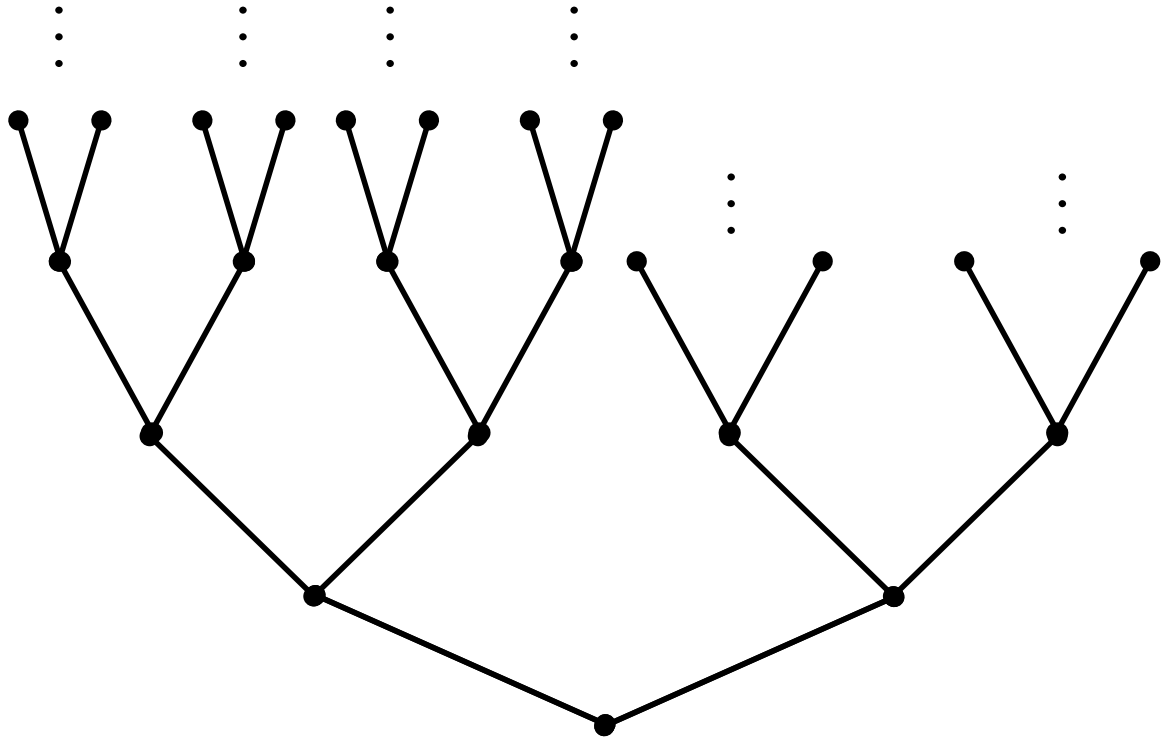
$\frac{1}{3}$ $\frac{2}{3}$



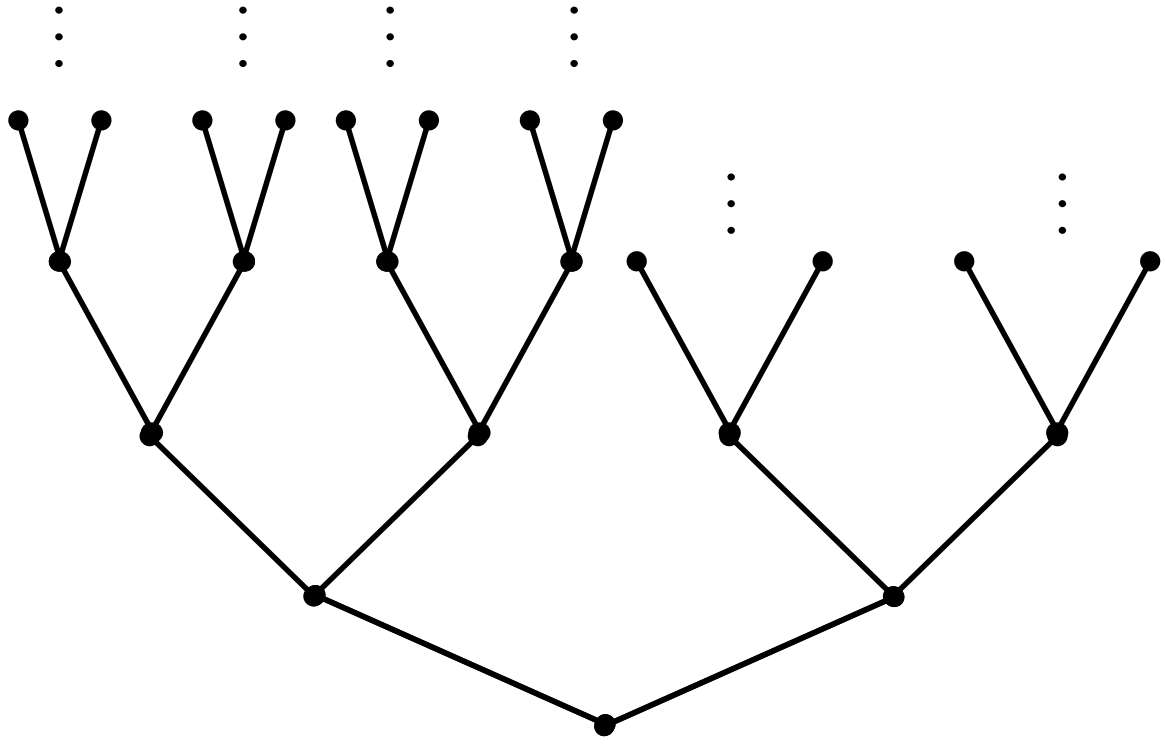
$\frac{1}{3}$ $\frac{2}{3}$



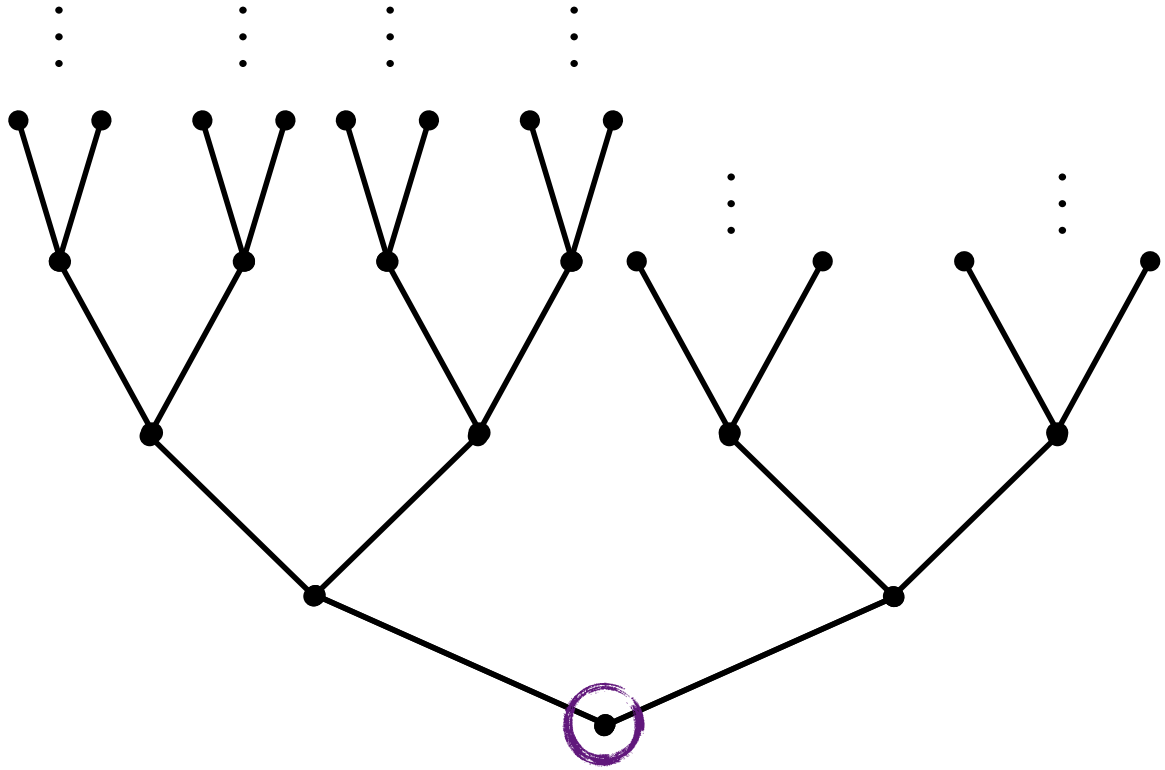
$\frac{1}{3}$ $\frac{2}{3}$



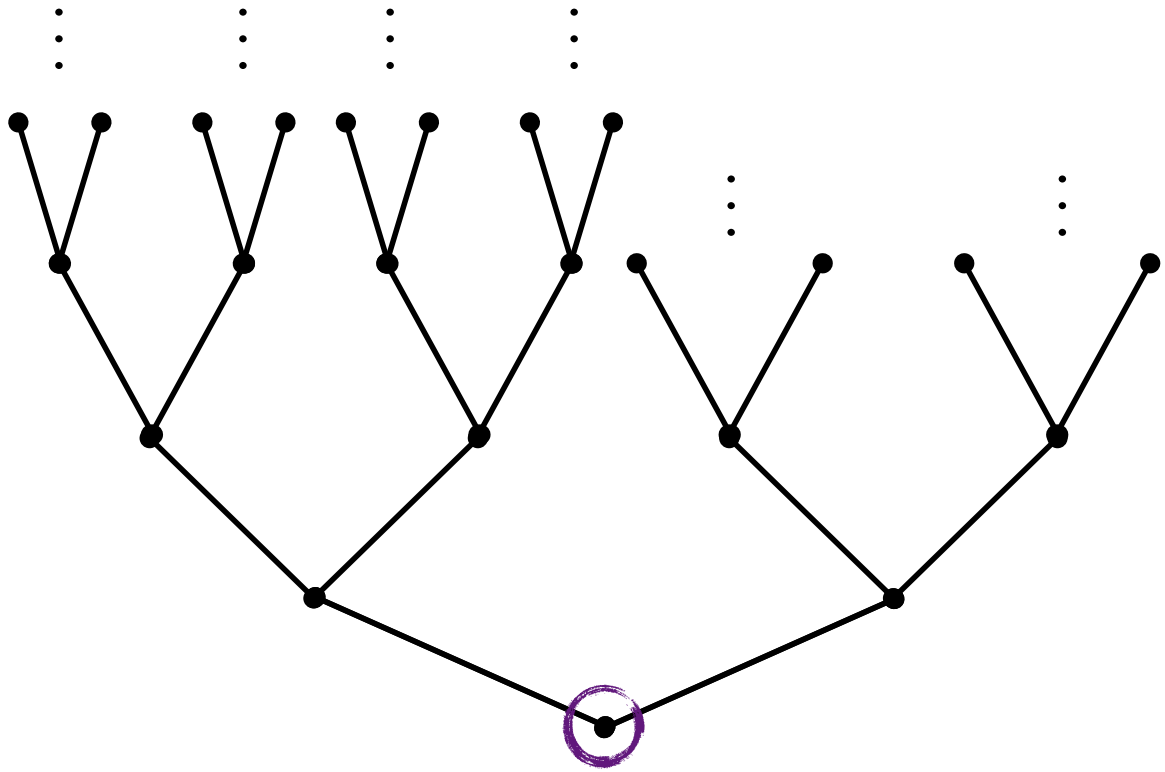

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



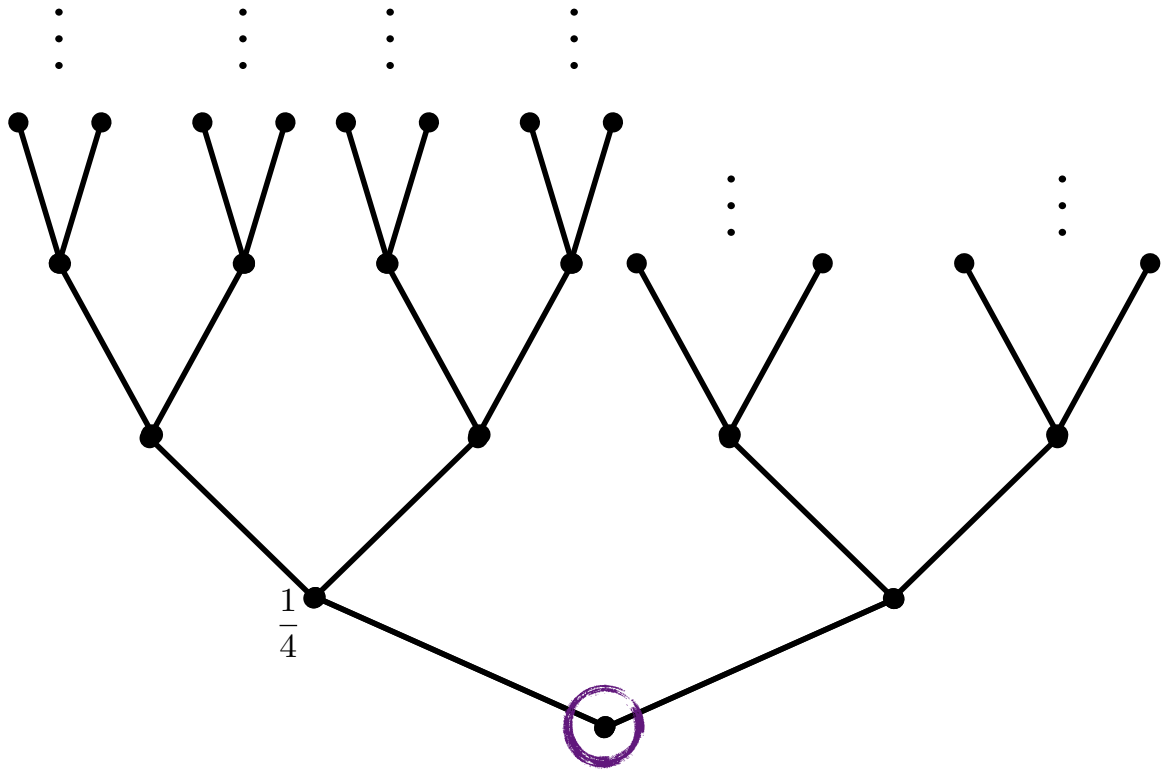

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



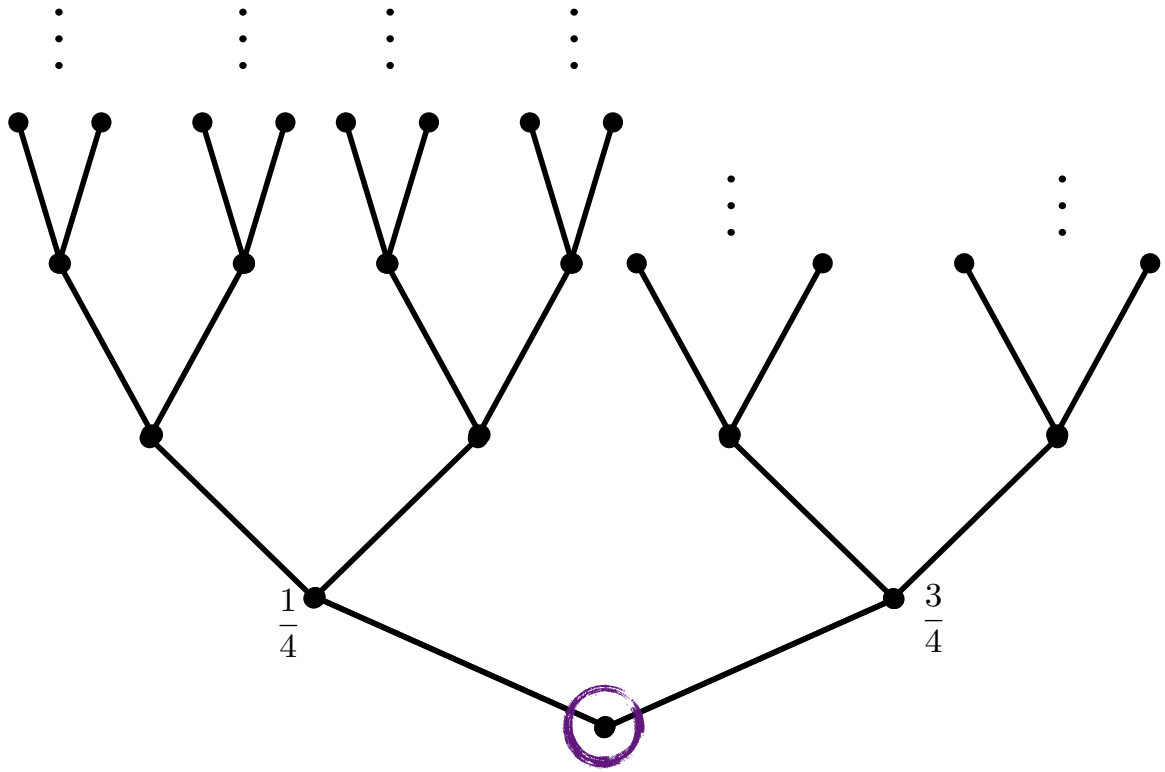

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



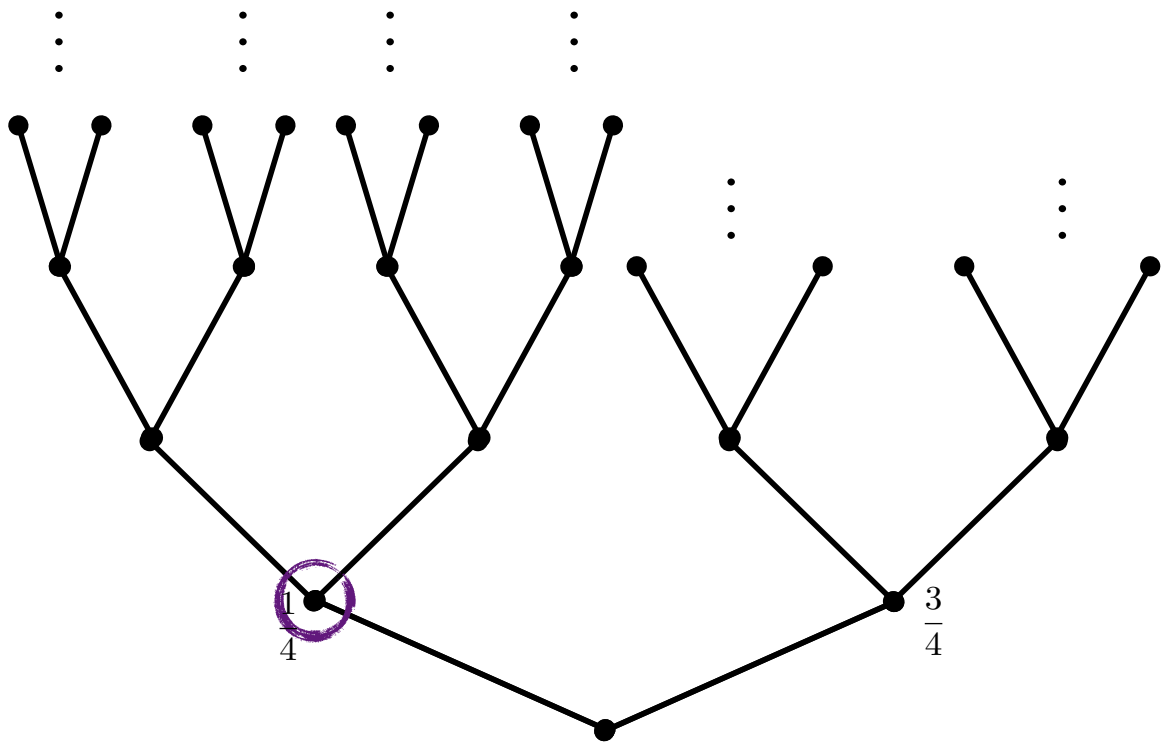
$\frac{1}{4}$ $\frac{3}{4}$



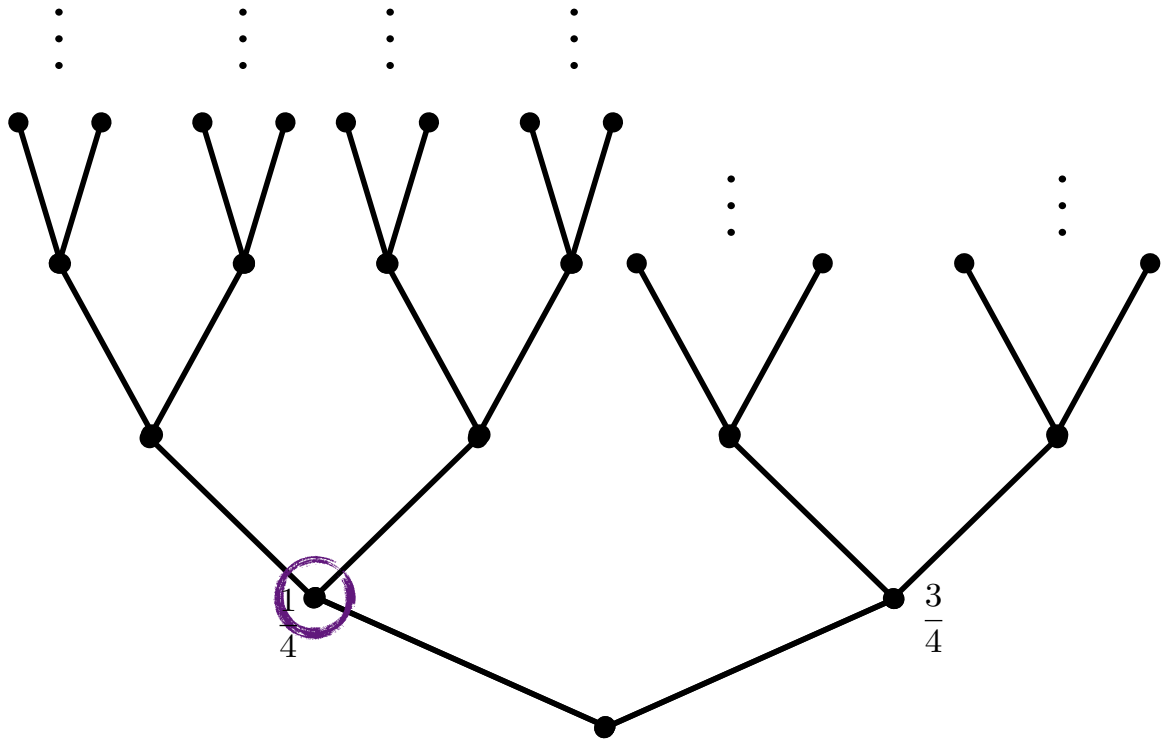
$\frac{1}{4}$ $\frac{3}{4}$



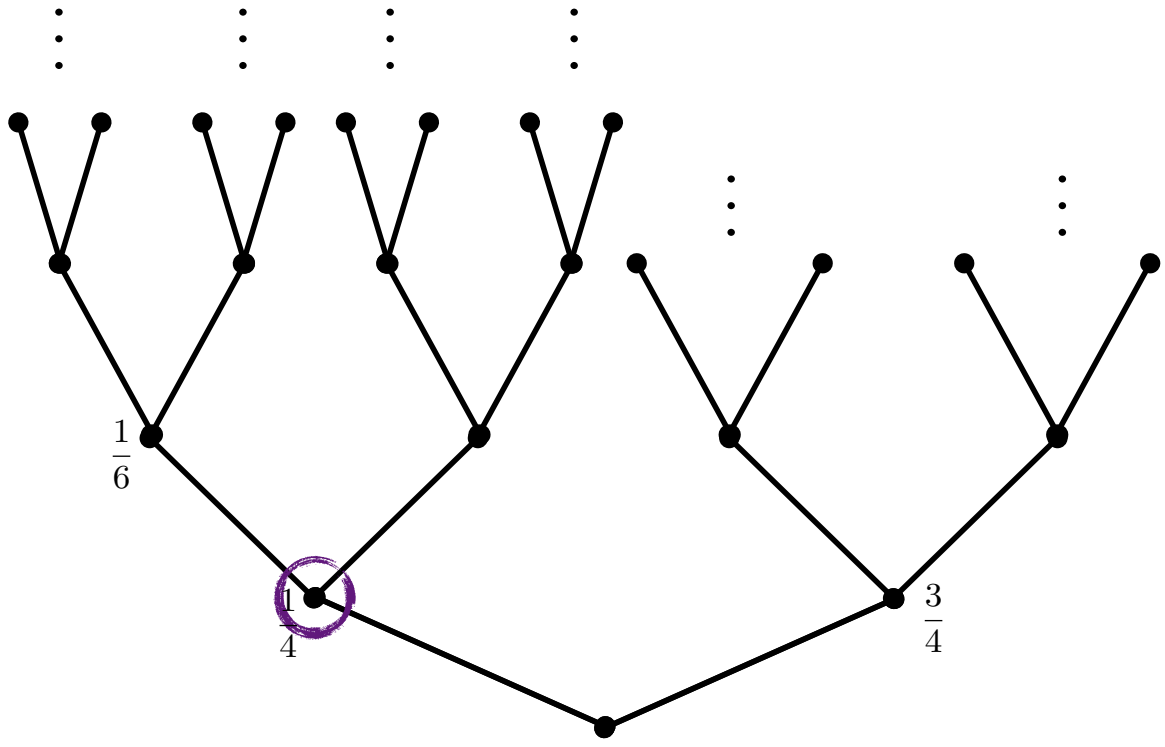
$\frac{1}{4}$ $\frac{3}{4}$


$\frac{1}{4}$
$\frac{3}{4}$

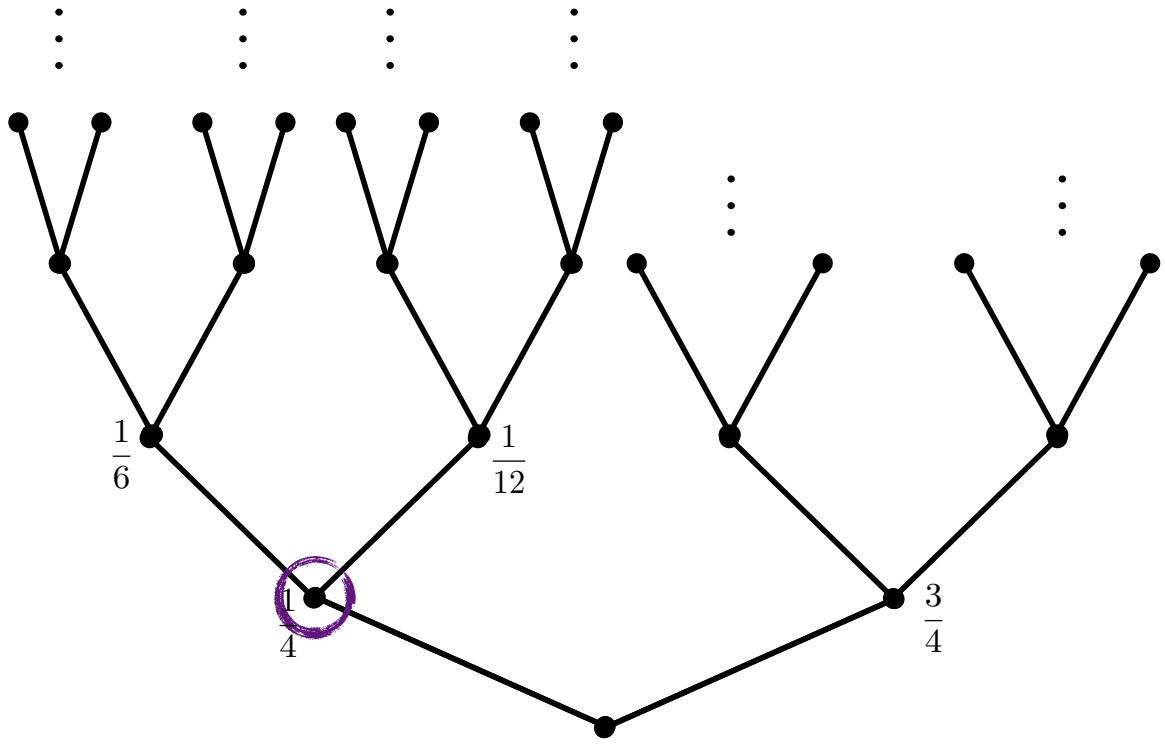

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



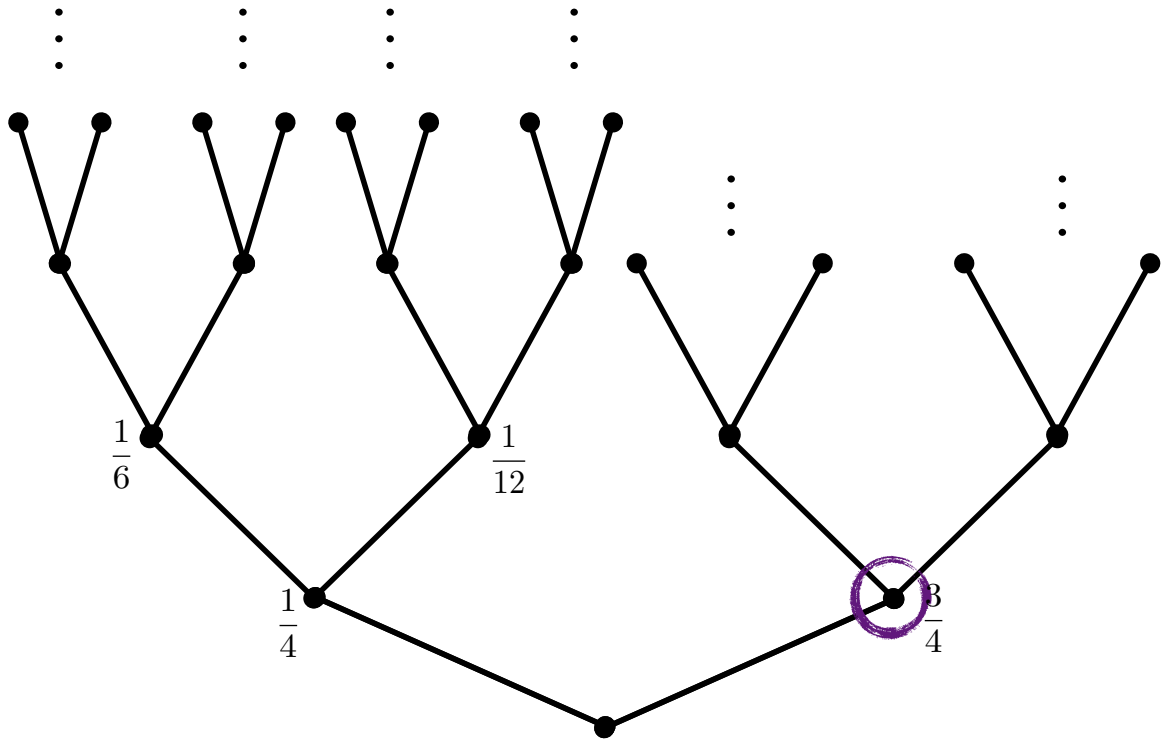
$\frac{2}{3}$ $\frac{1}{3}$



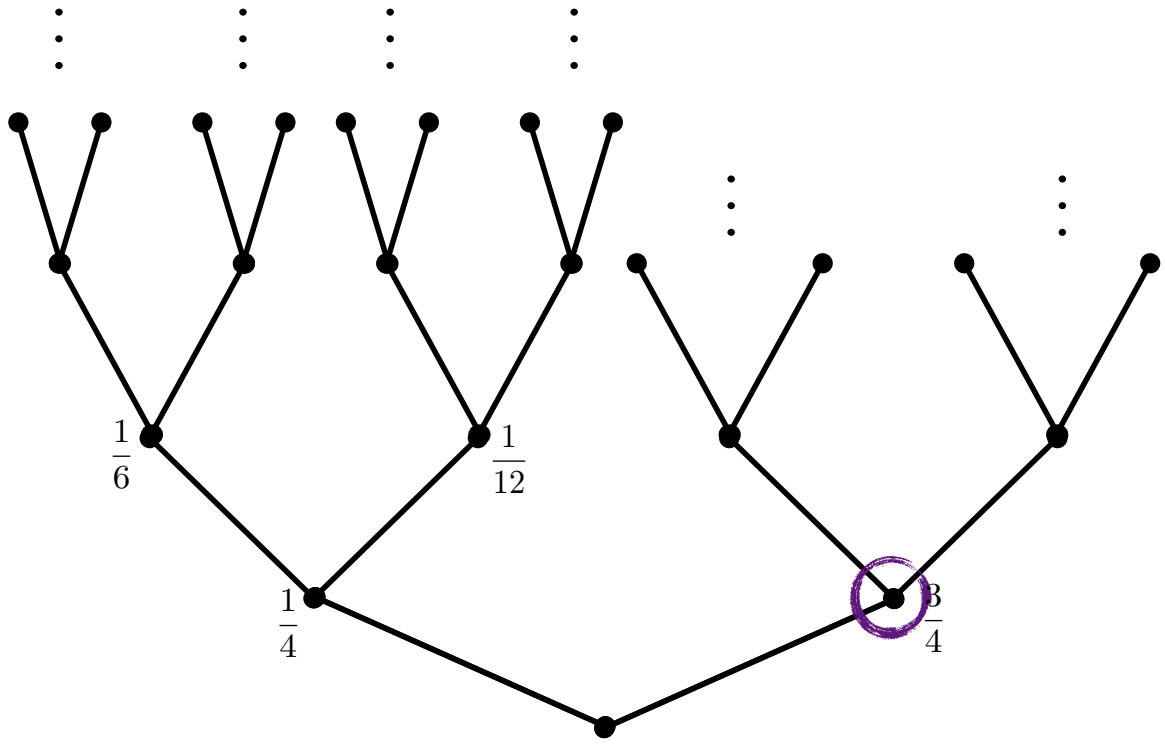
$\frac{2}{3}$ $\frac{1}{3}$



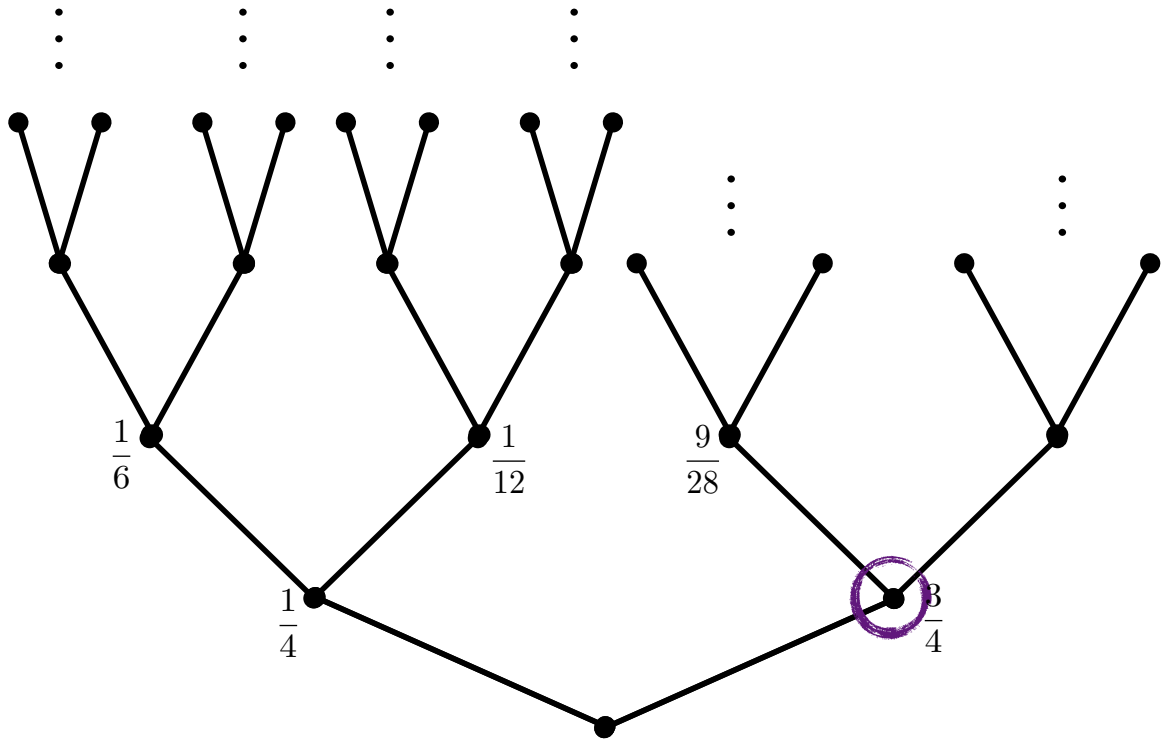
$\frac{2}{3}$ $\frac{1}{3}$


$\frac{1}{6}$
$\frac{1}{12}$
$\frac{1}{4}$
$\frac{3}{4}$

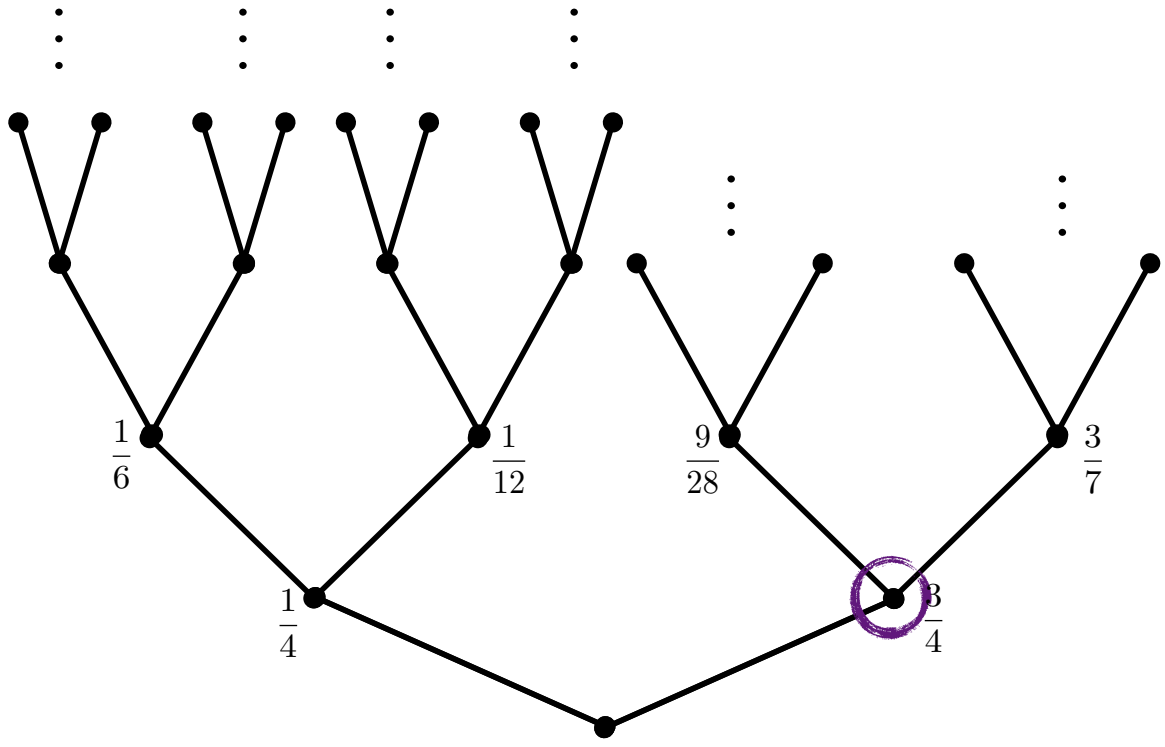

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



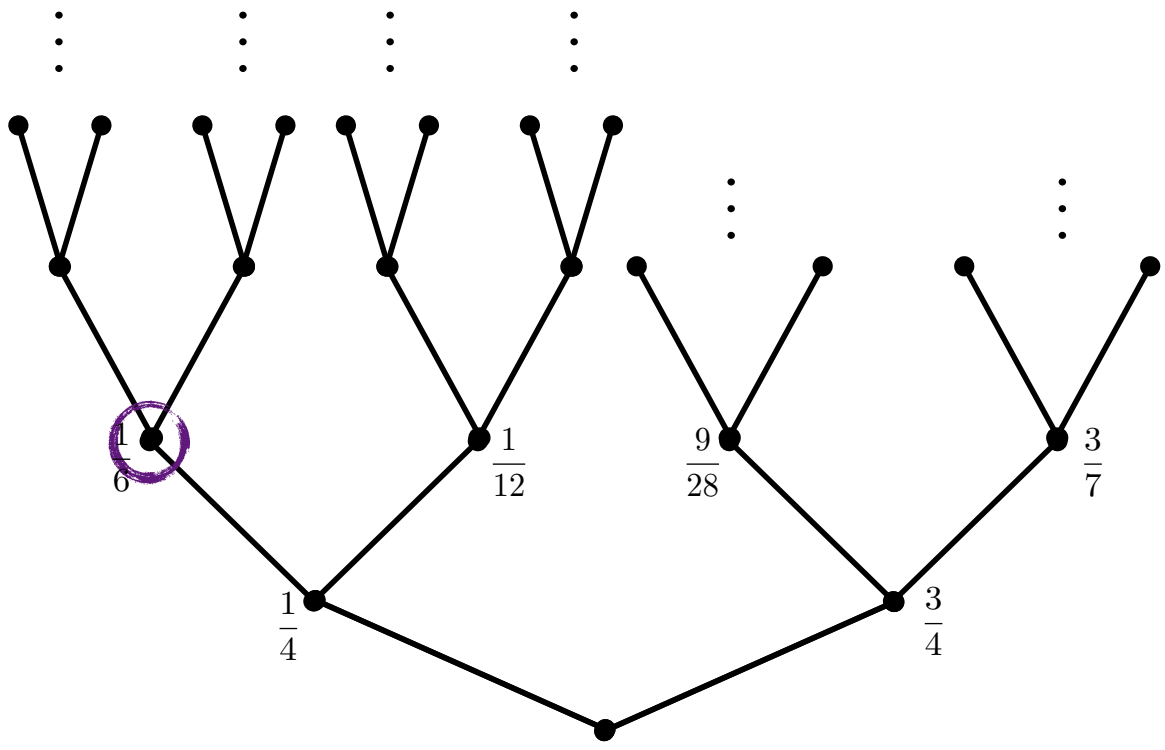
$\frac{3}{7}$ $\frac{4}{7}$



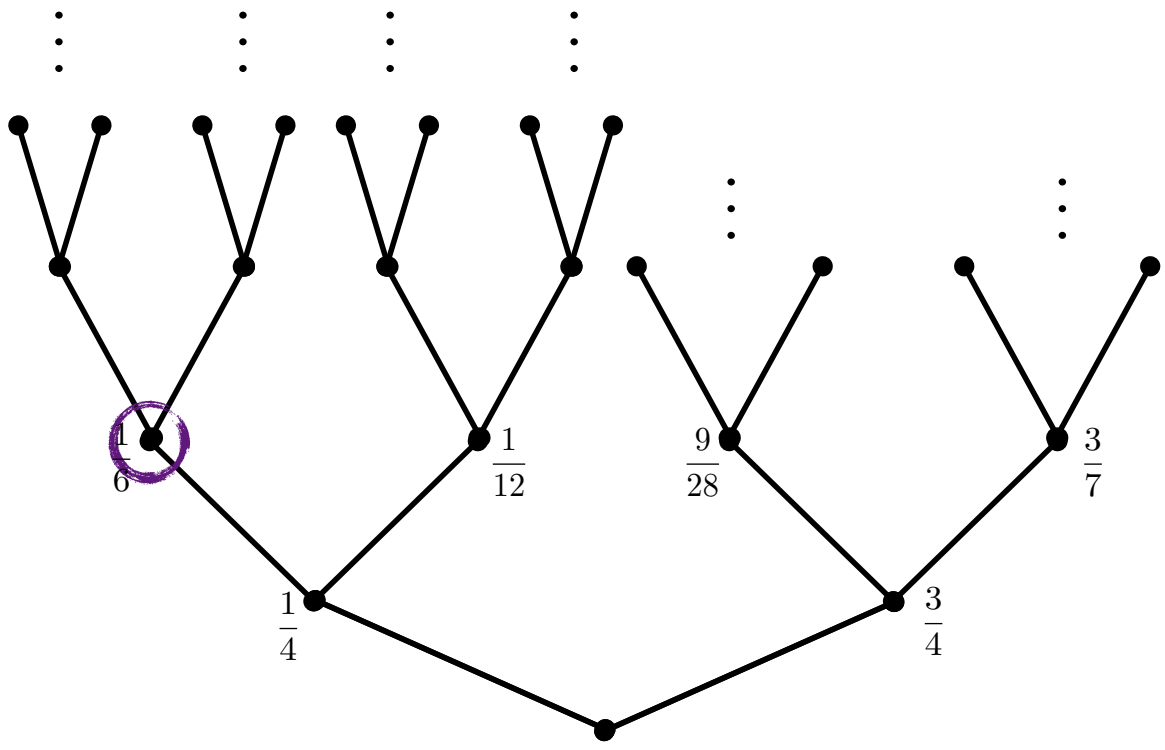
$\frac{3}{7}$ $\frac{4}{7}$



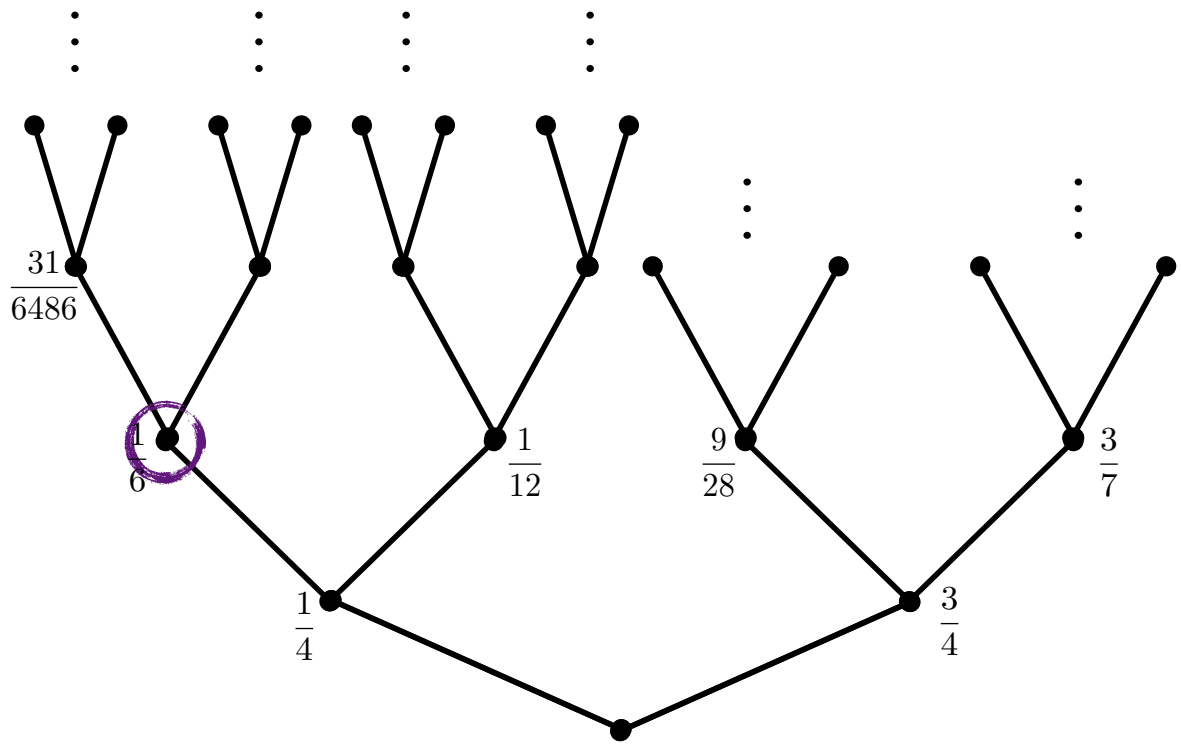
$\frac{3}{7}$ $\frac{4}{7}$


1/6
1/12
9/28
3/7
1/4
3/4

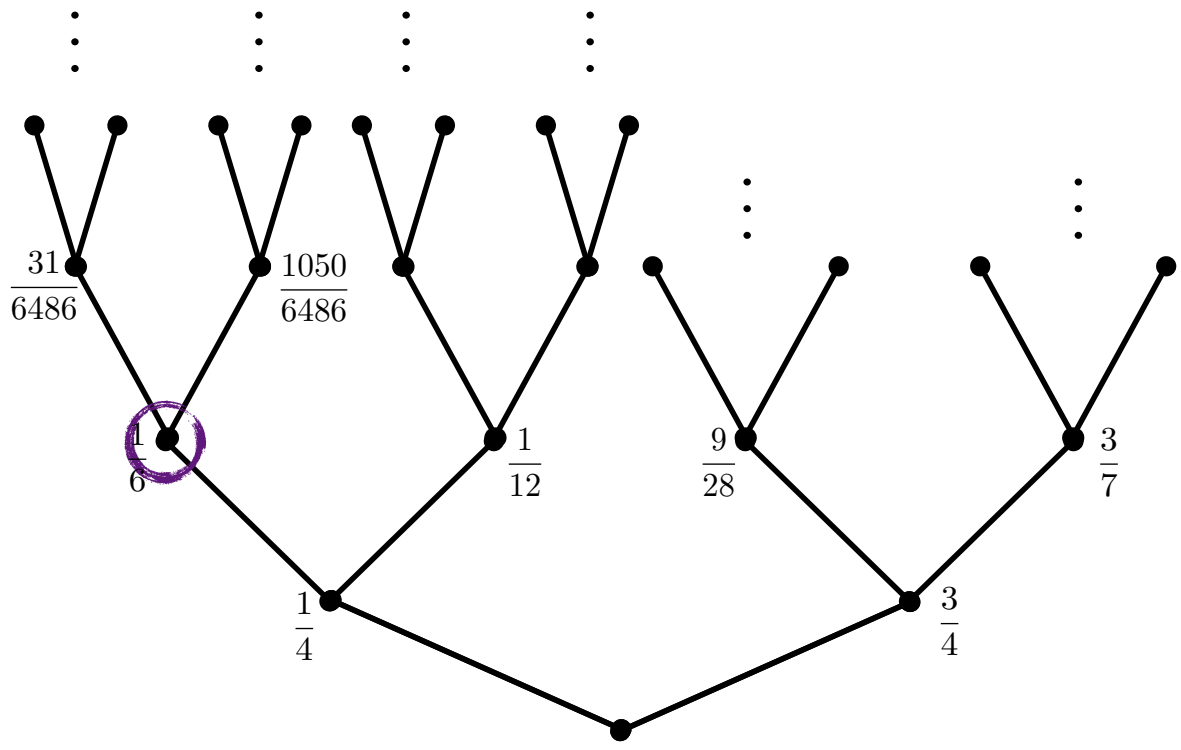

LIBERTY
IN GOD WE TRUST
1932
UNITED STATES OF AMERICA
QUARTER DOLLAR



$$\frac{31}{1081} \qquad \frac{1050}{1081}$$



$$\frac{31}{1081} \quad \frac{1050}{1081}$$



$$\frac{31}{1081} \quad \frac{1050}{1081}$$

Randomness with respect to computable measures

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.

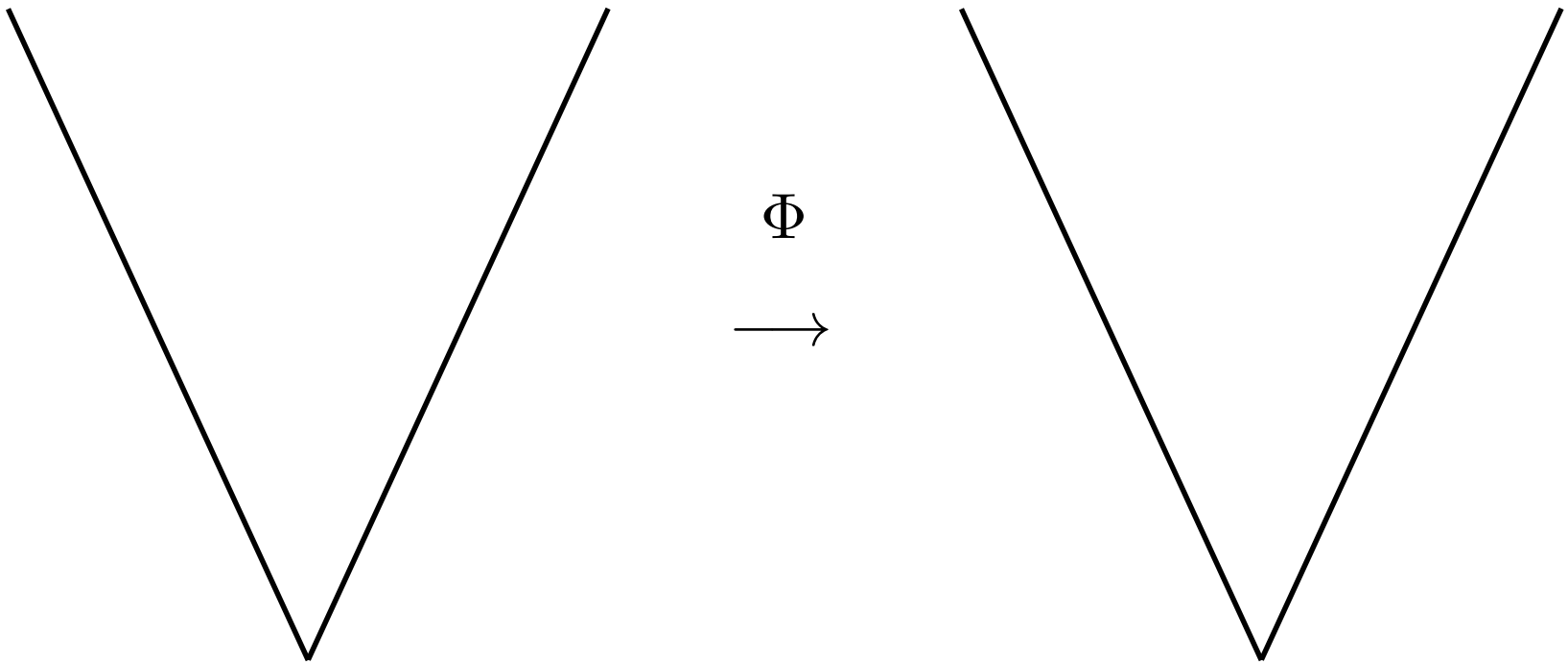
Alternatively, one can define a probabilistic machine to be an oracle Turing machine with some algorithmically random sequence as an oracle.

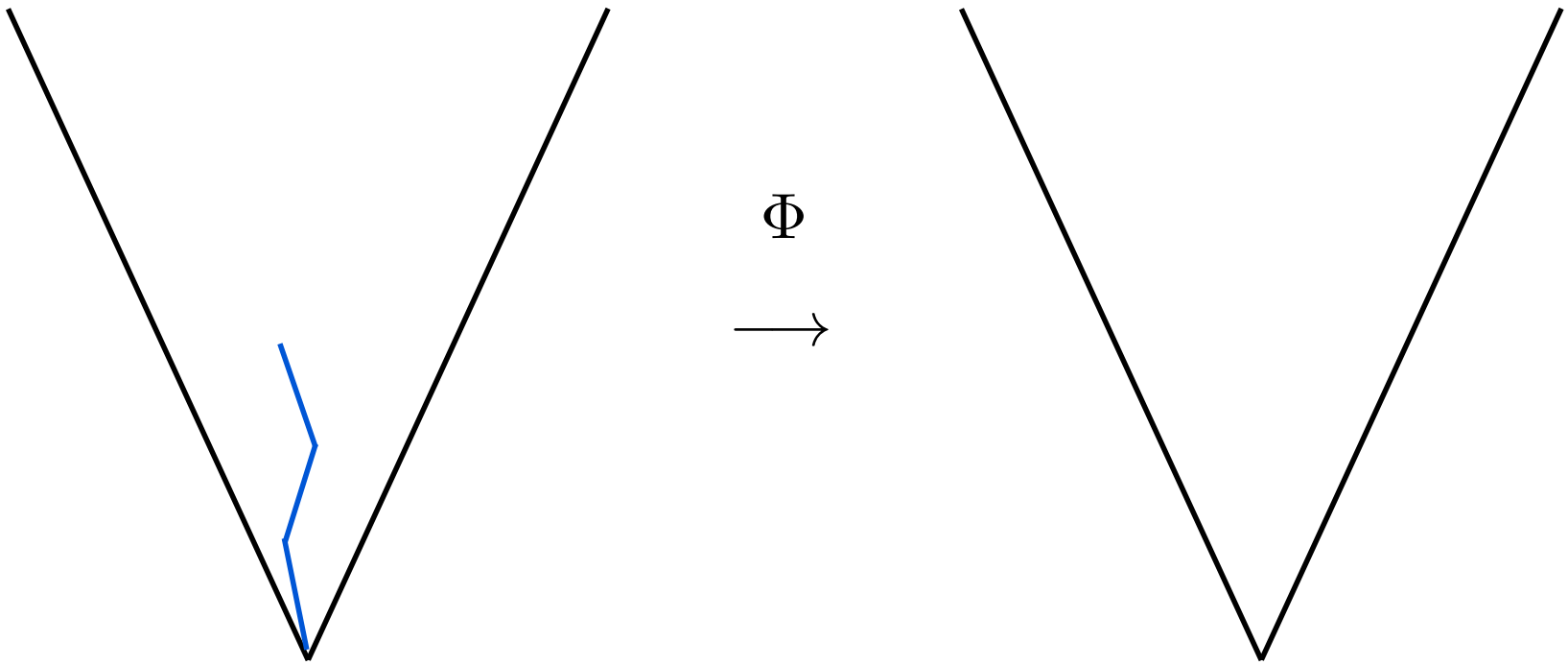
Key idea: For the purposes of computing a sequence or some sequence in a fixed class collection *with positive probability*, these two approaches are equivalent.

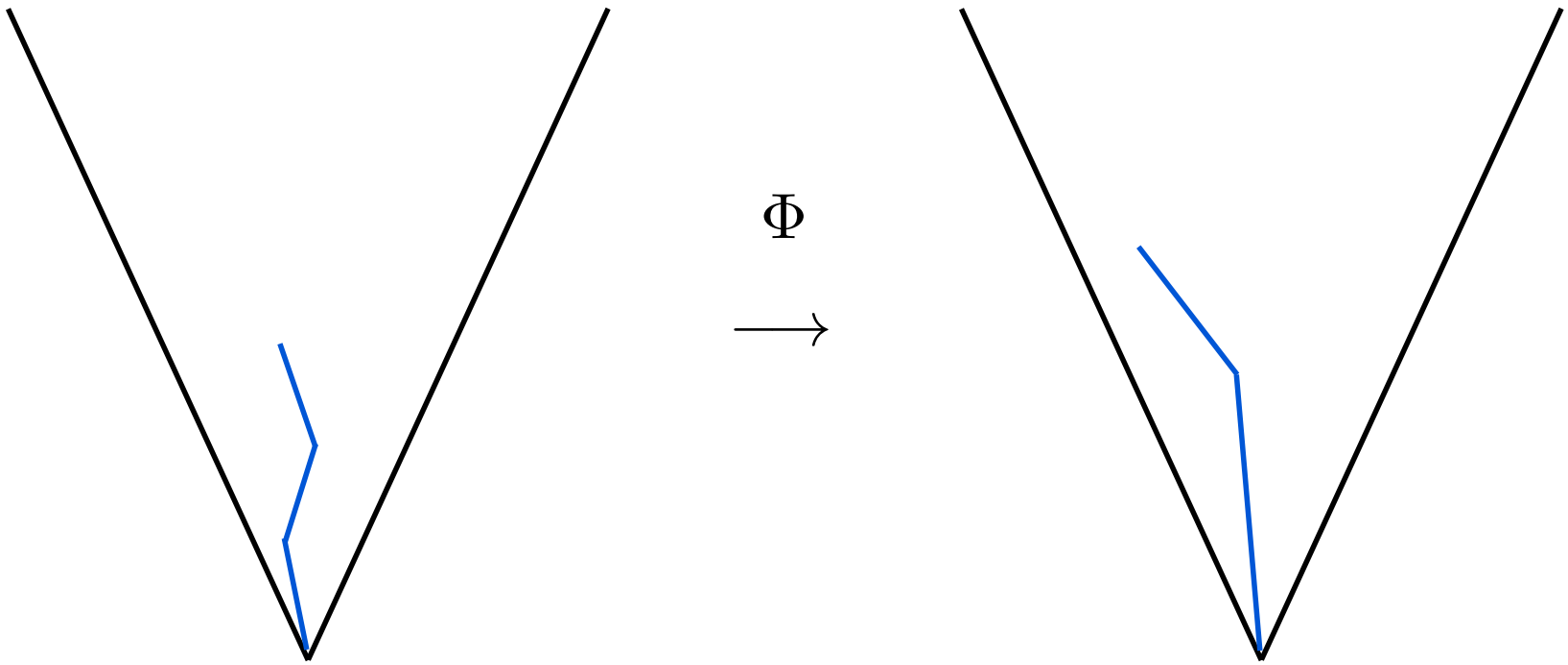
Turing functionals

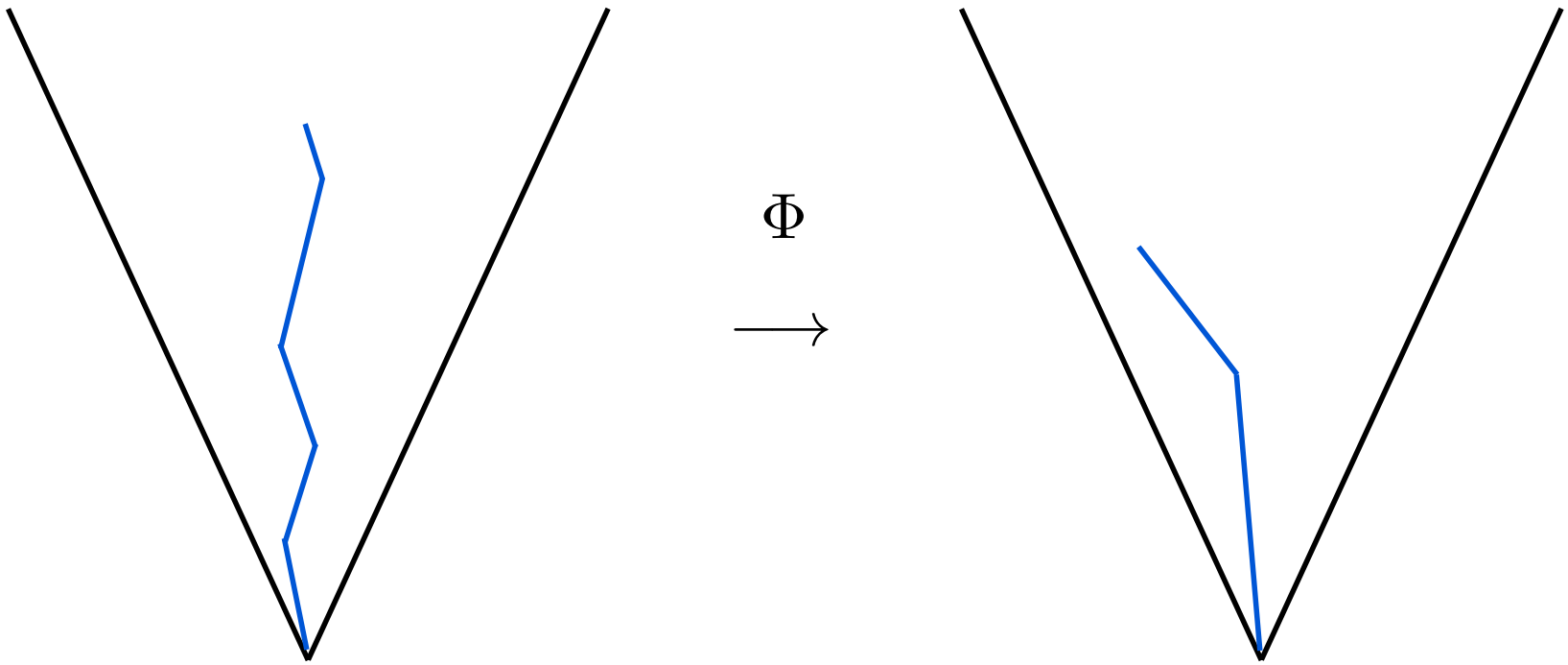
Definition

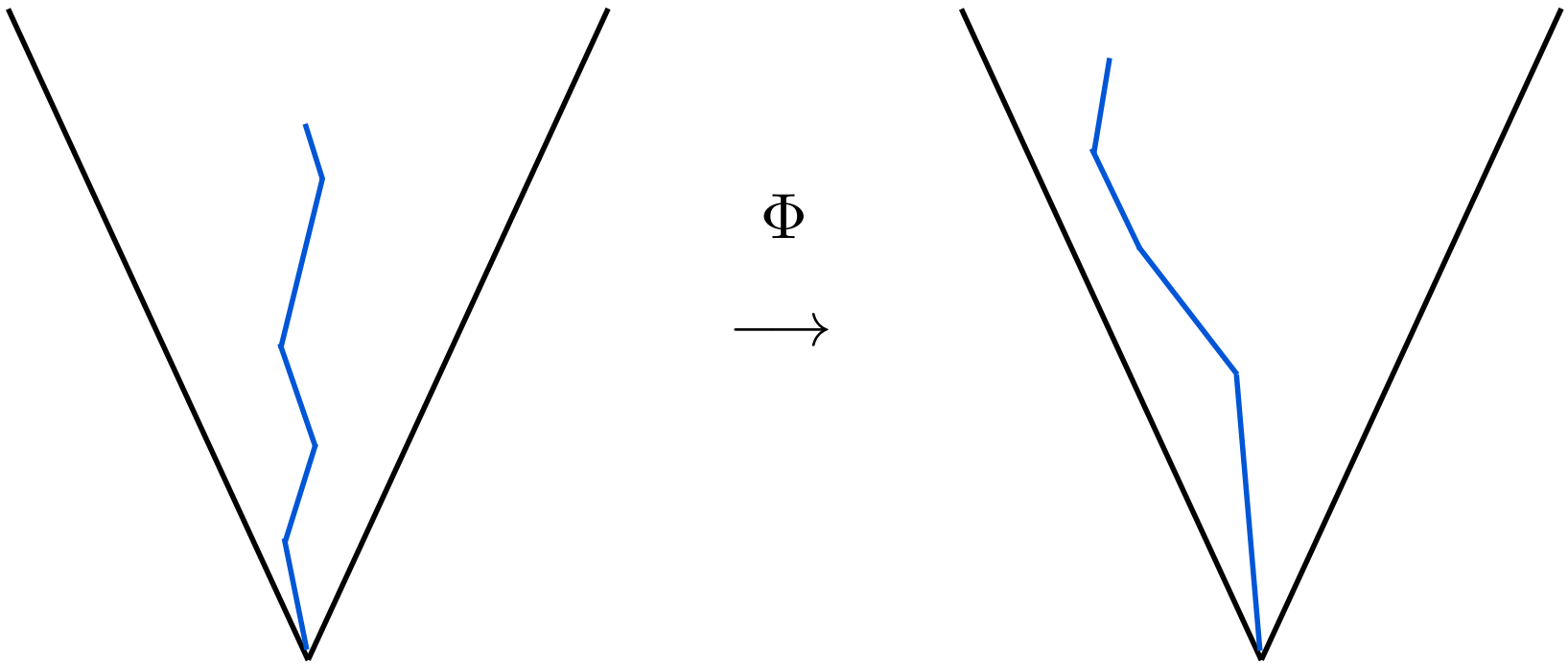
A *Turing functional* $\Phi : 2^\omega \to 2^\omega$ is a computably enumerable set of pairs of strings (σ, τ) such that if $(\sigma, \tau), (\sigma', \tau') \in \Phi$ and $\sigma \preceq \sigma'$, then $\tau \preceq \tau'$ or $\tau' \preceq \tau$.

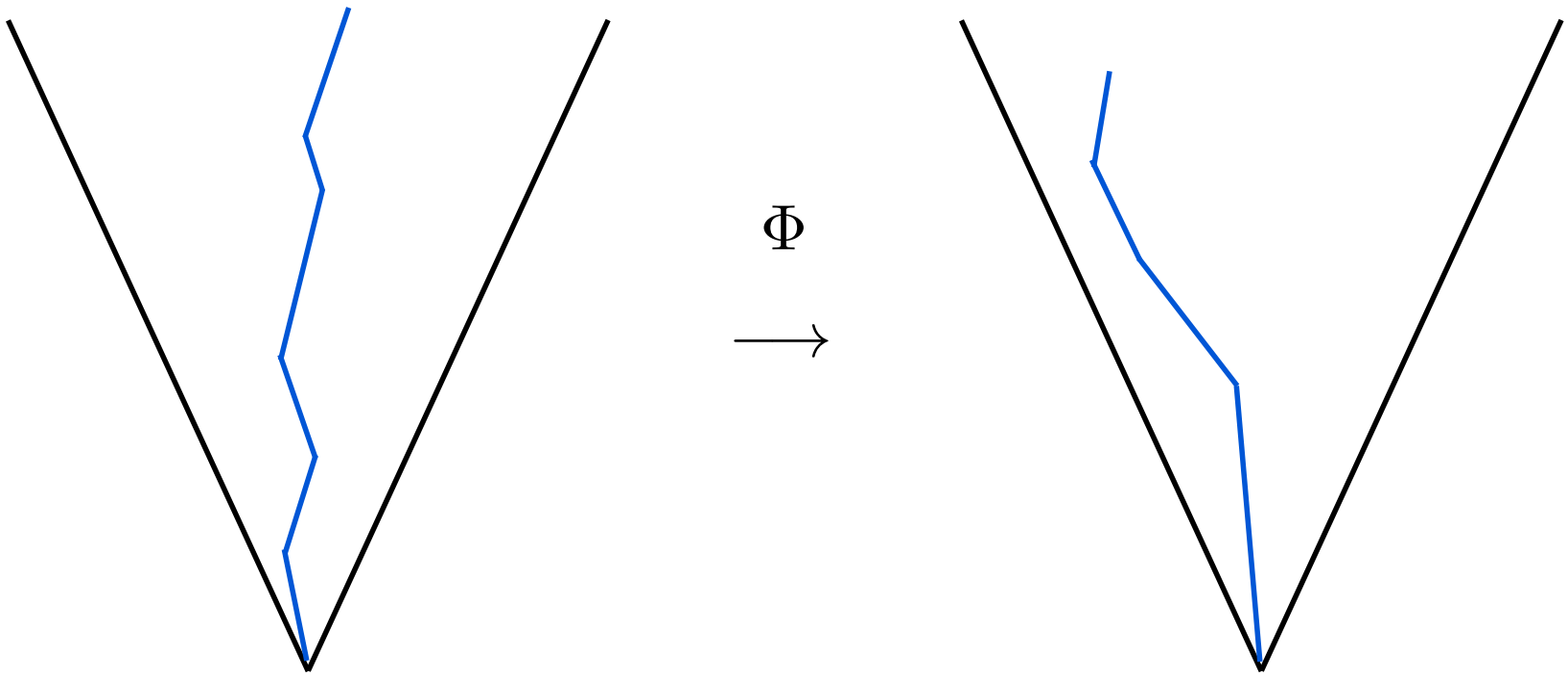

Φ

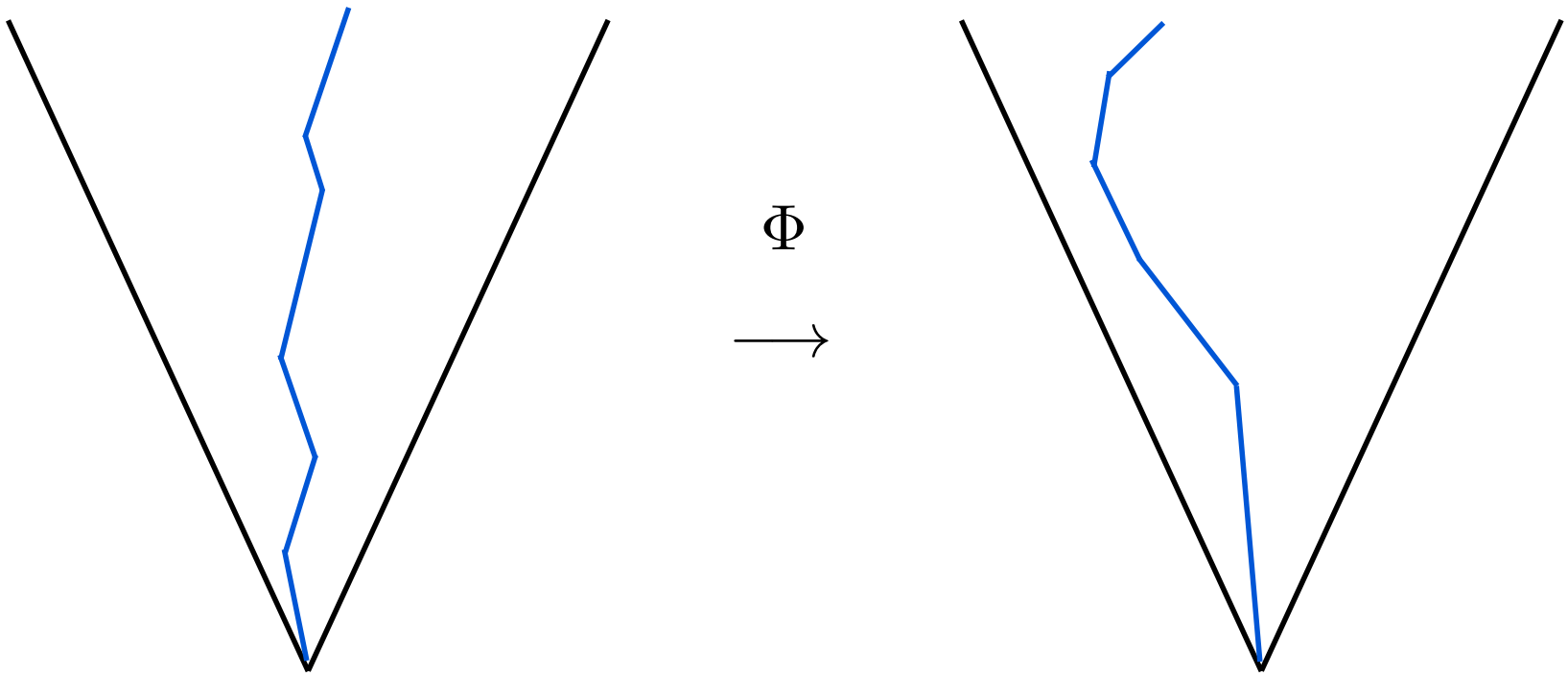

Φ

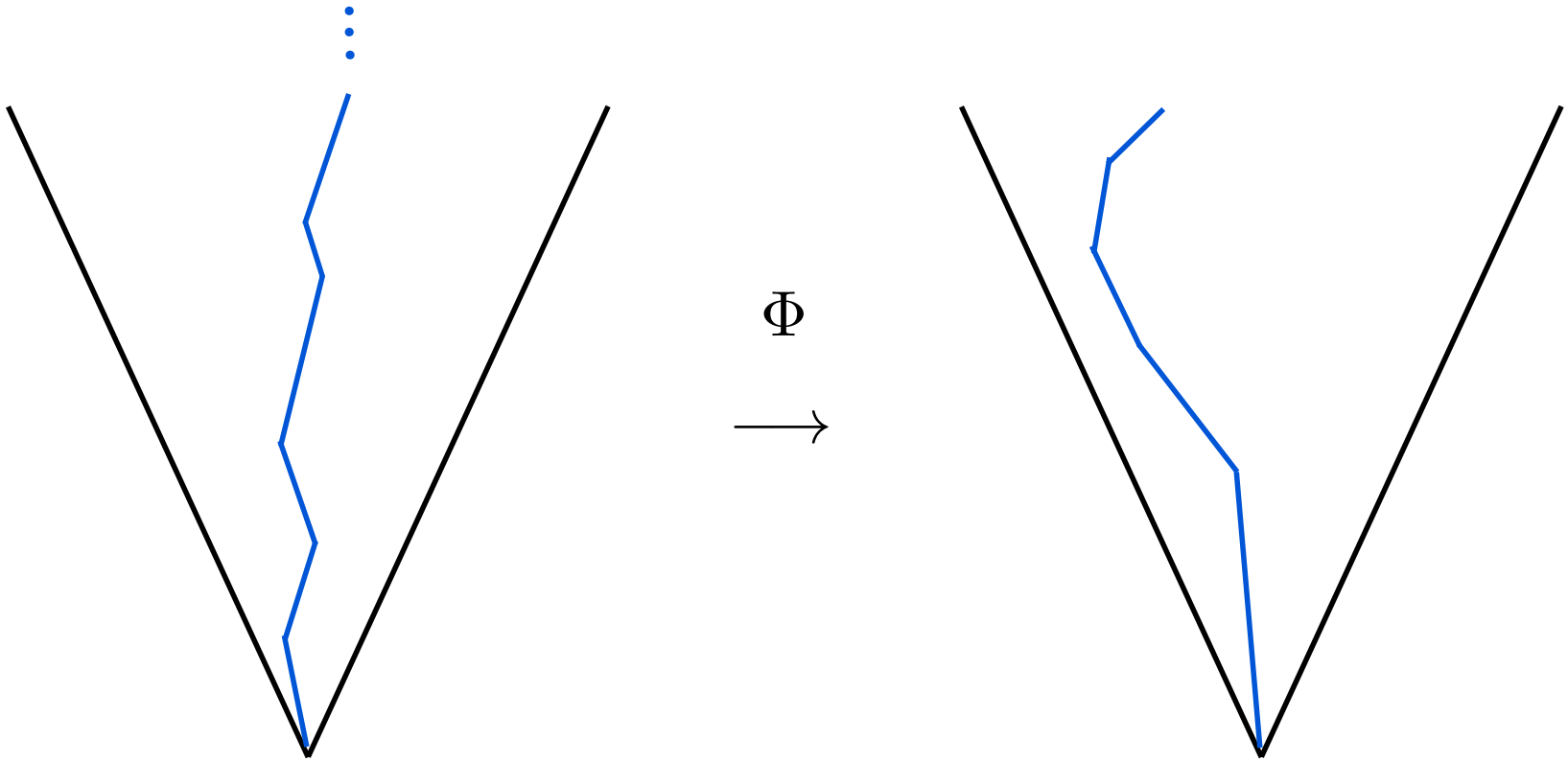

Φ

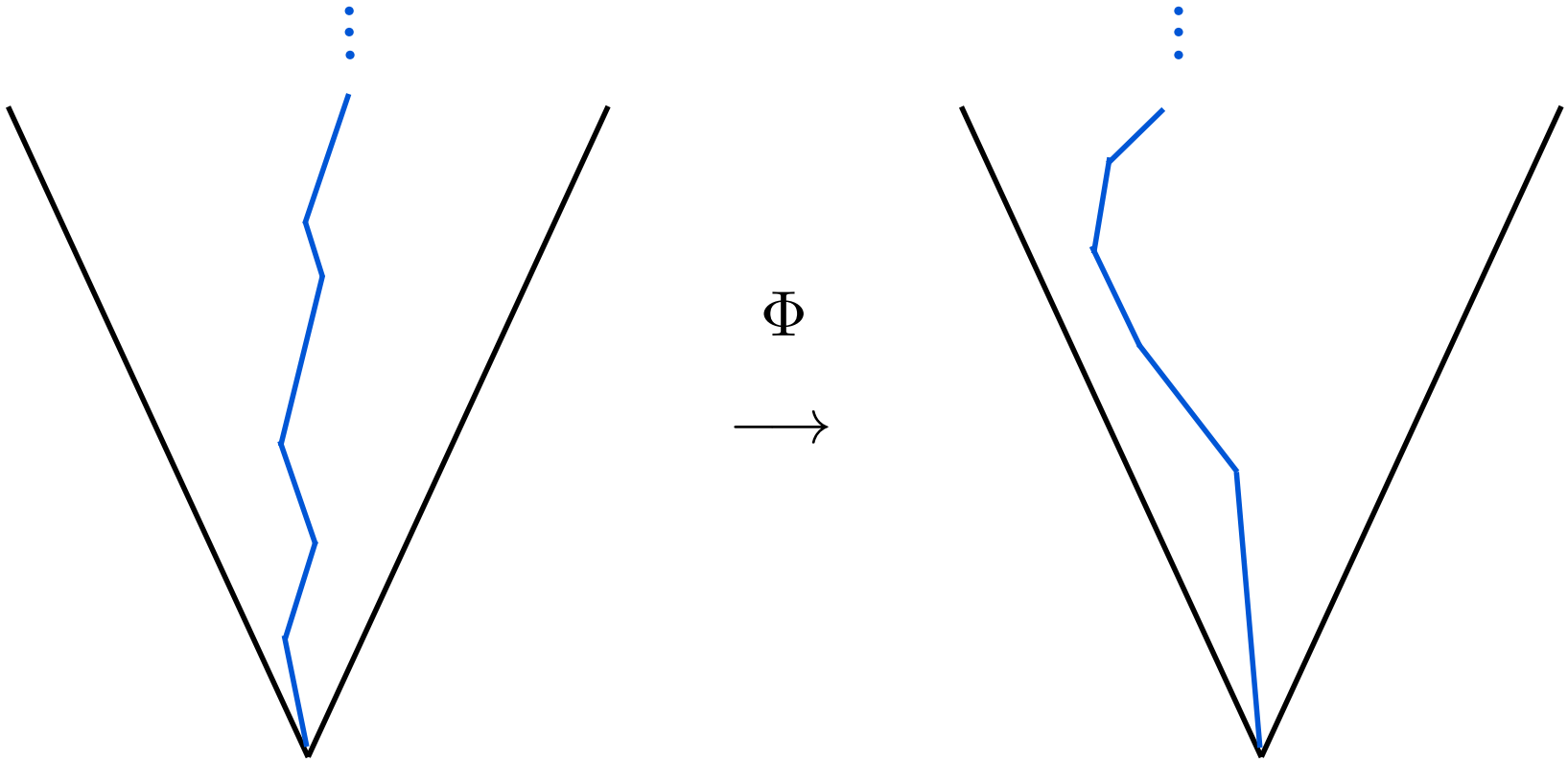

Φ

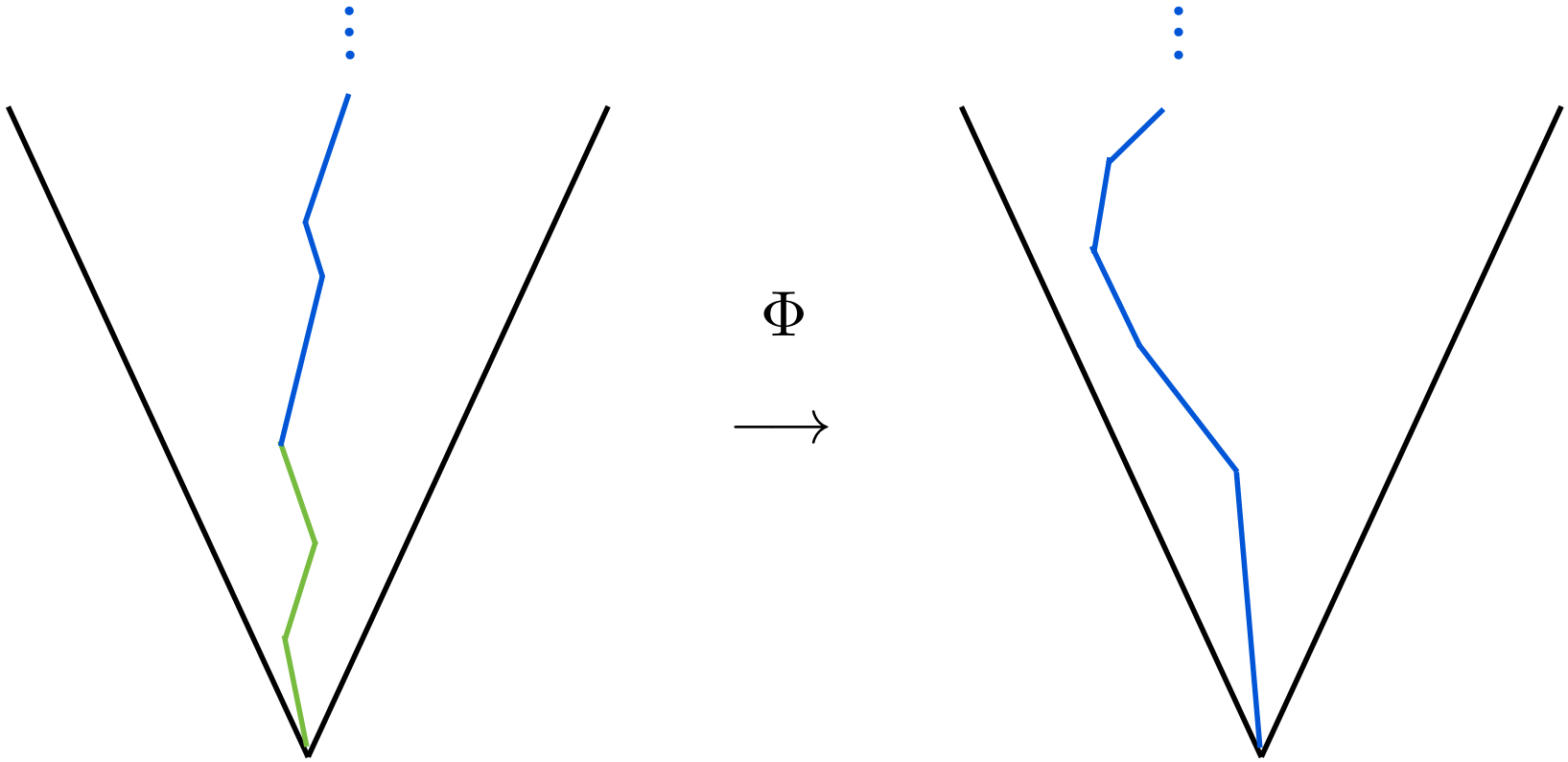

Φ

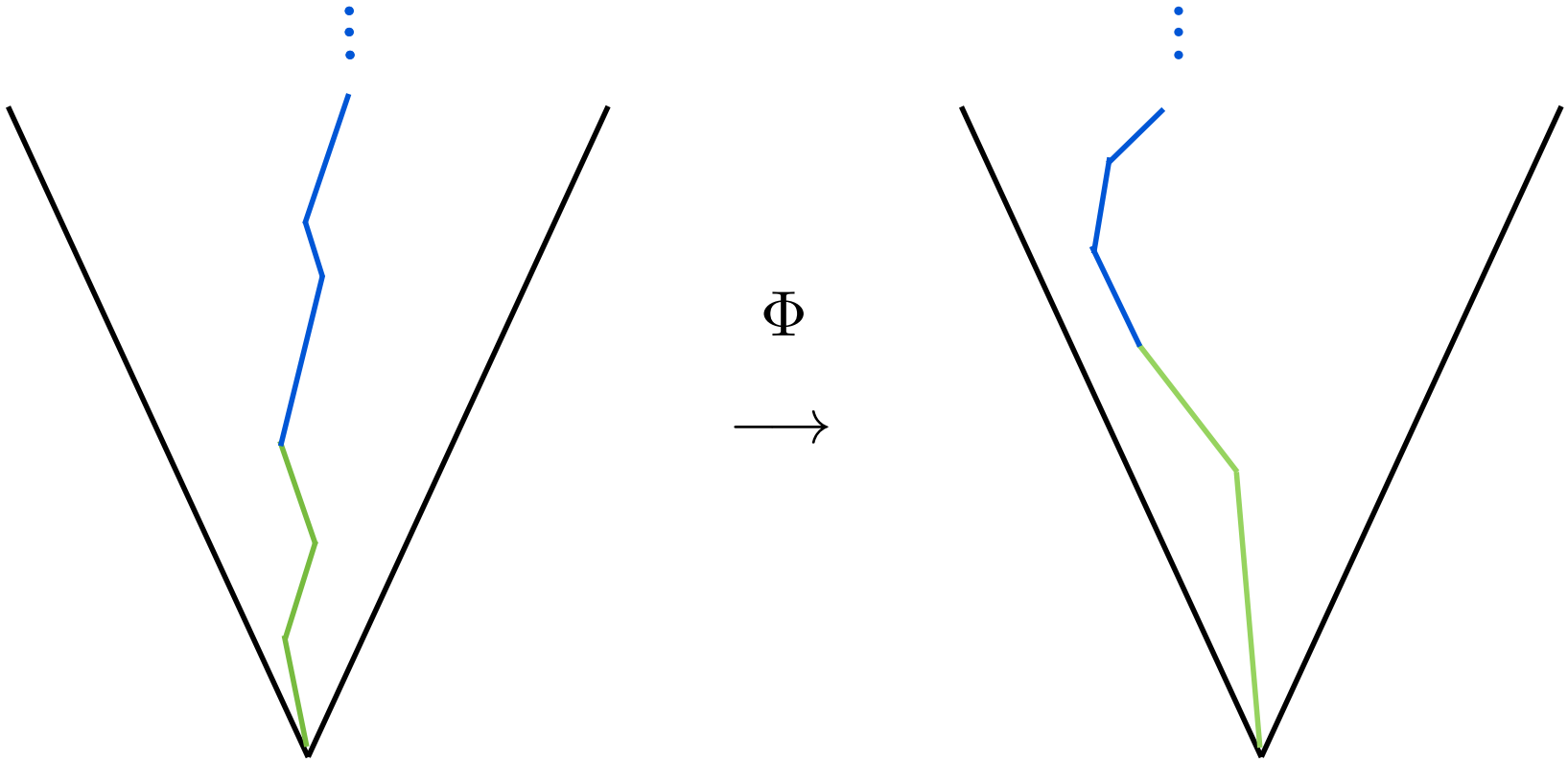

Φ

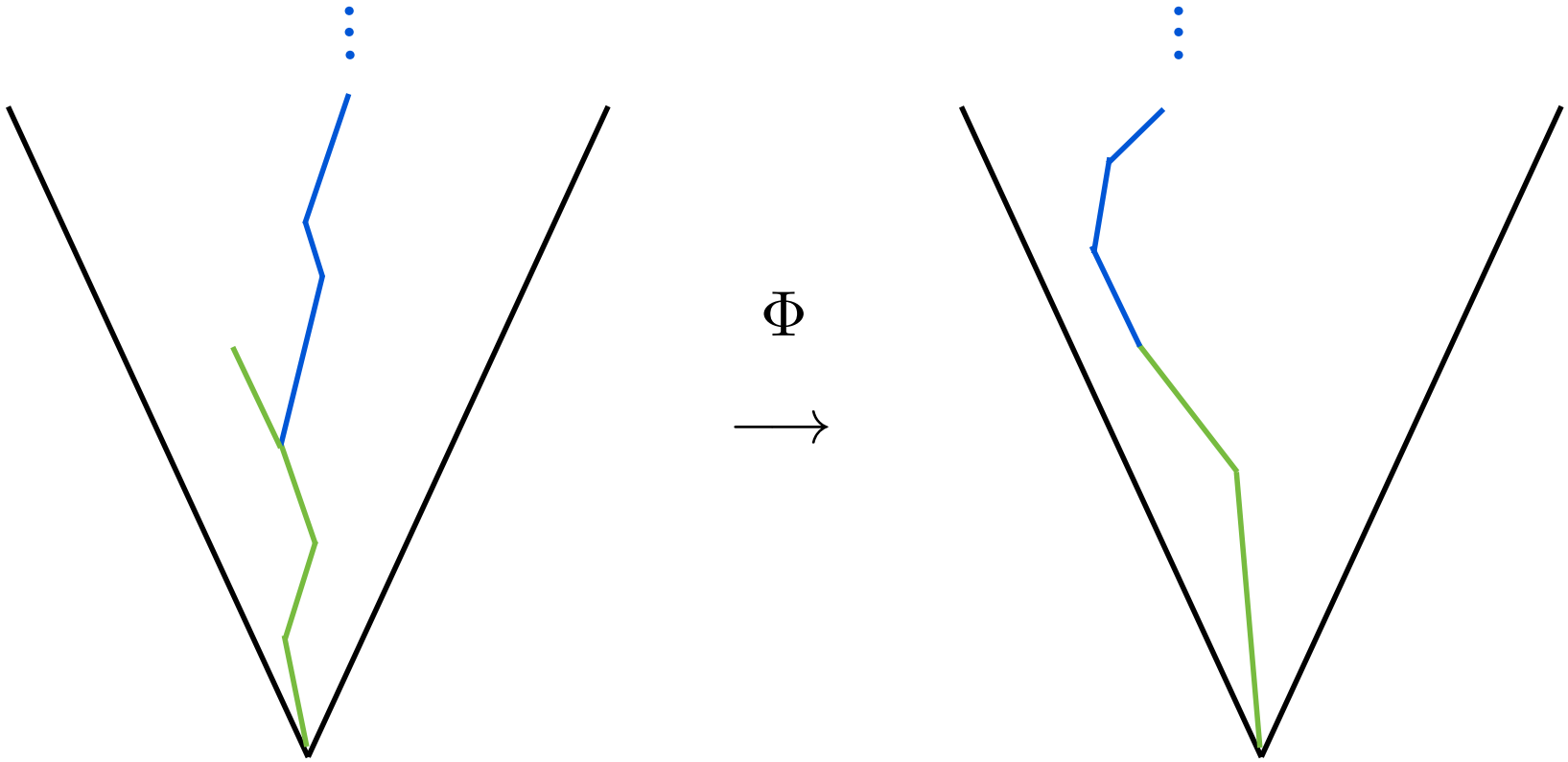

Φ

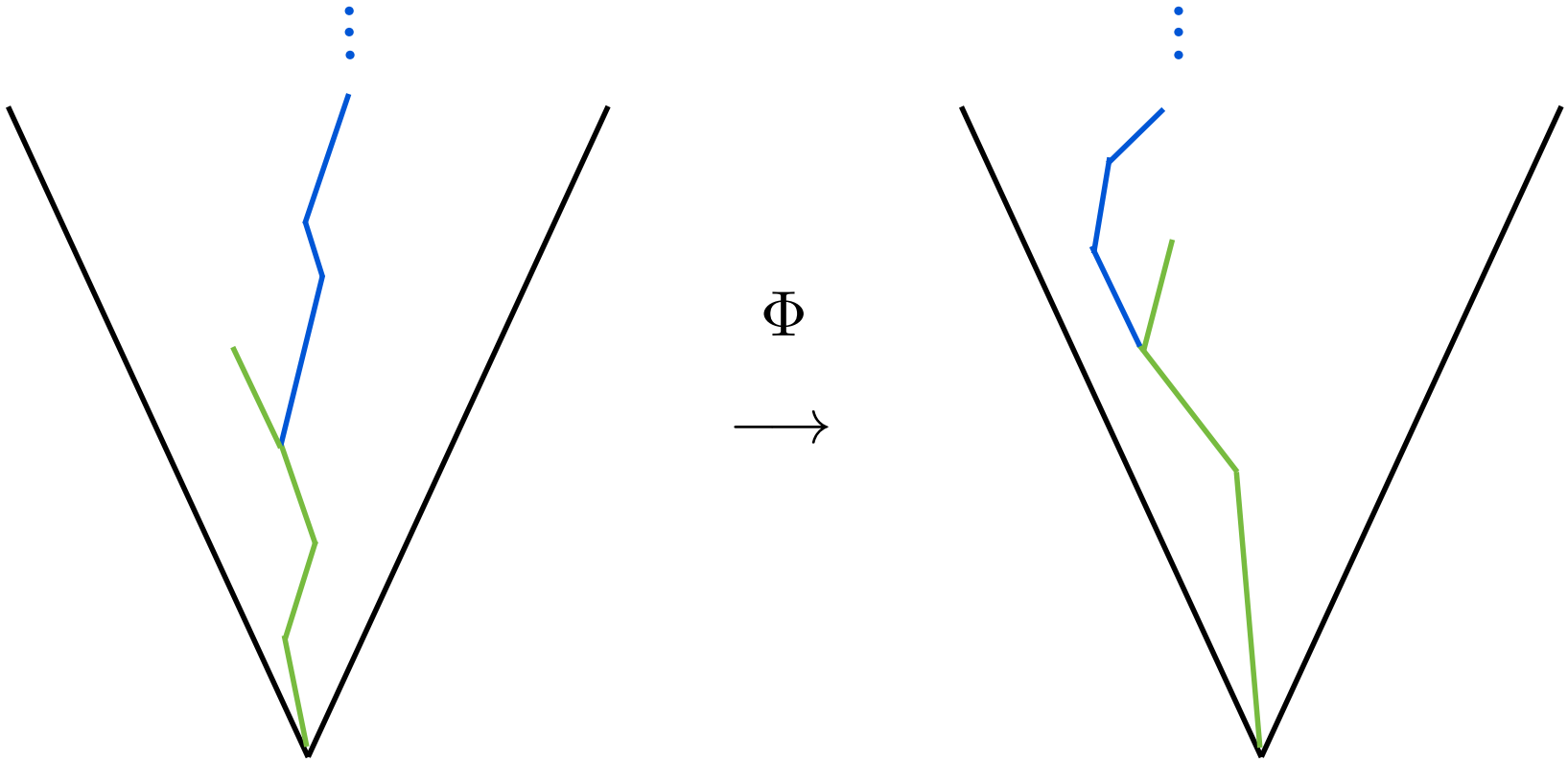

Φ

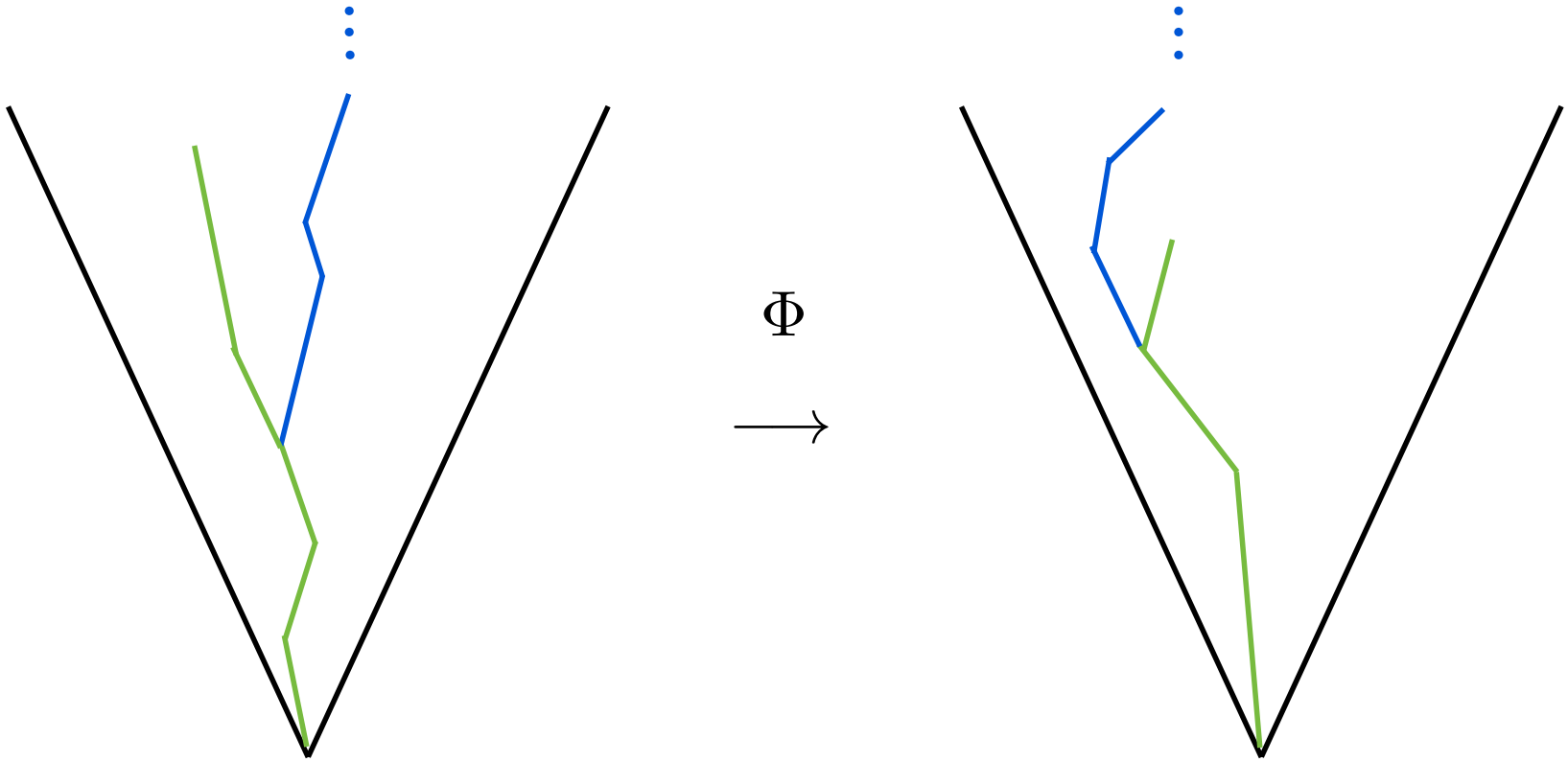

Φ

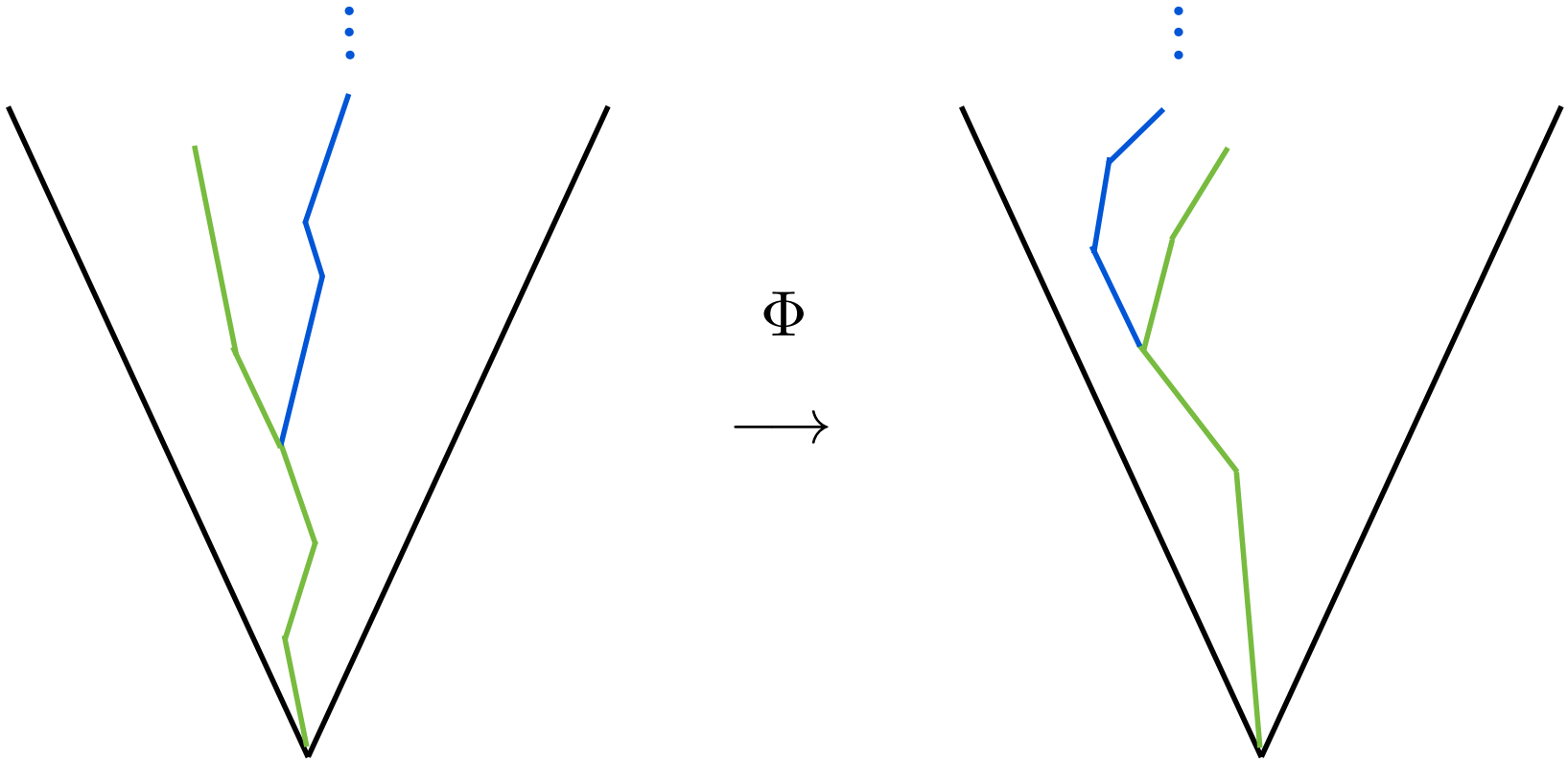

Φ

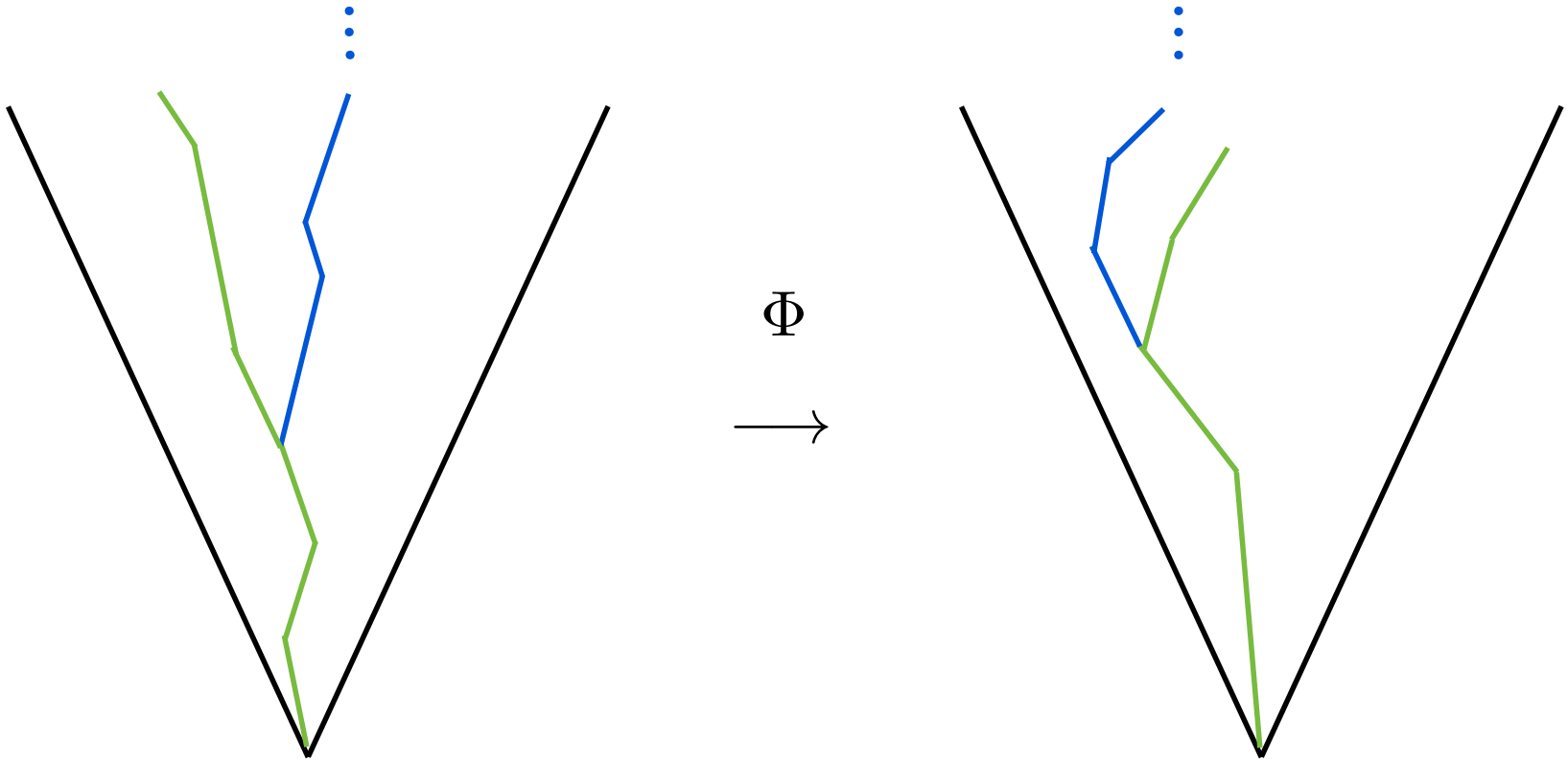

Φ

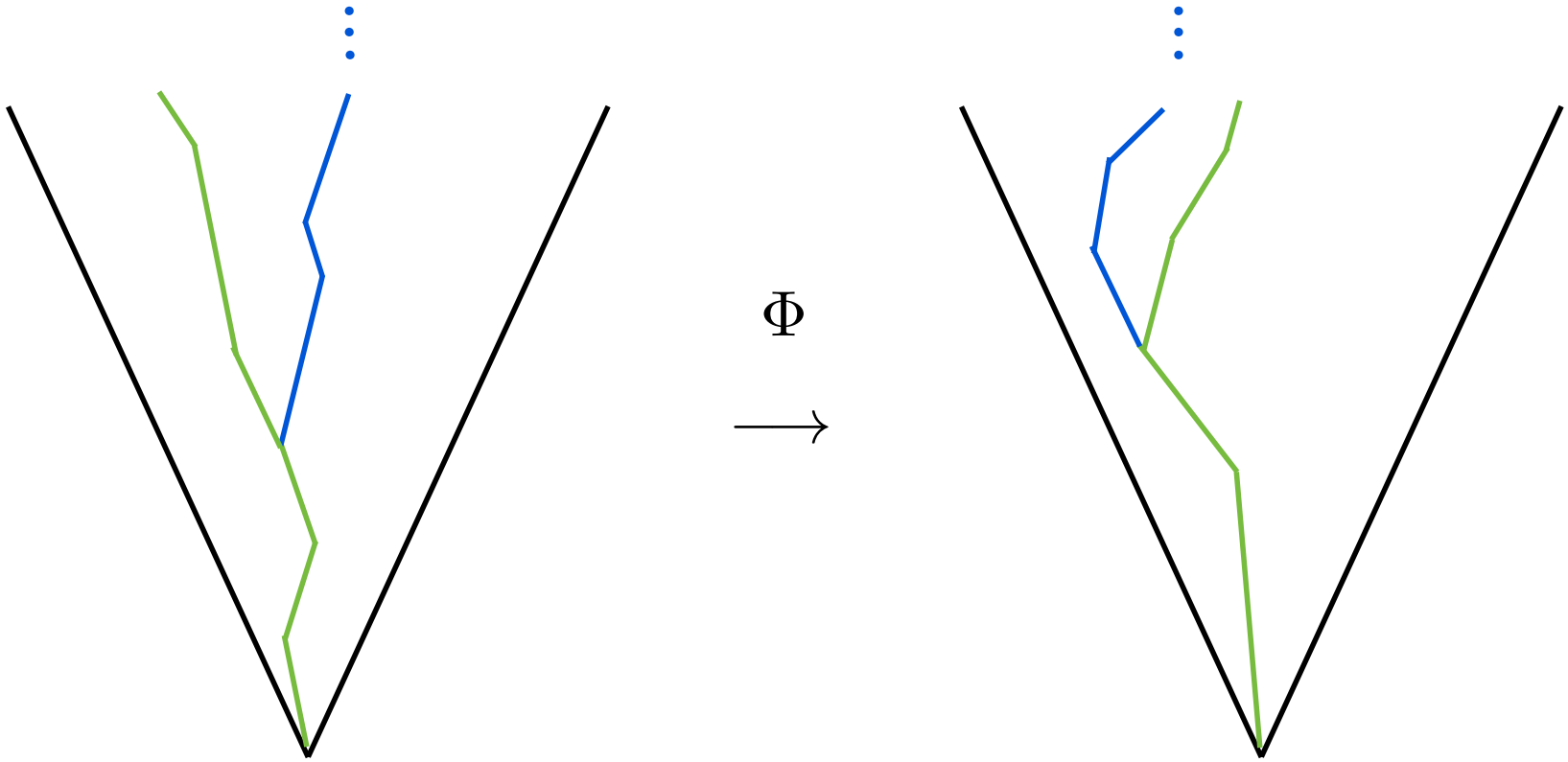

Φ

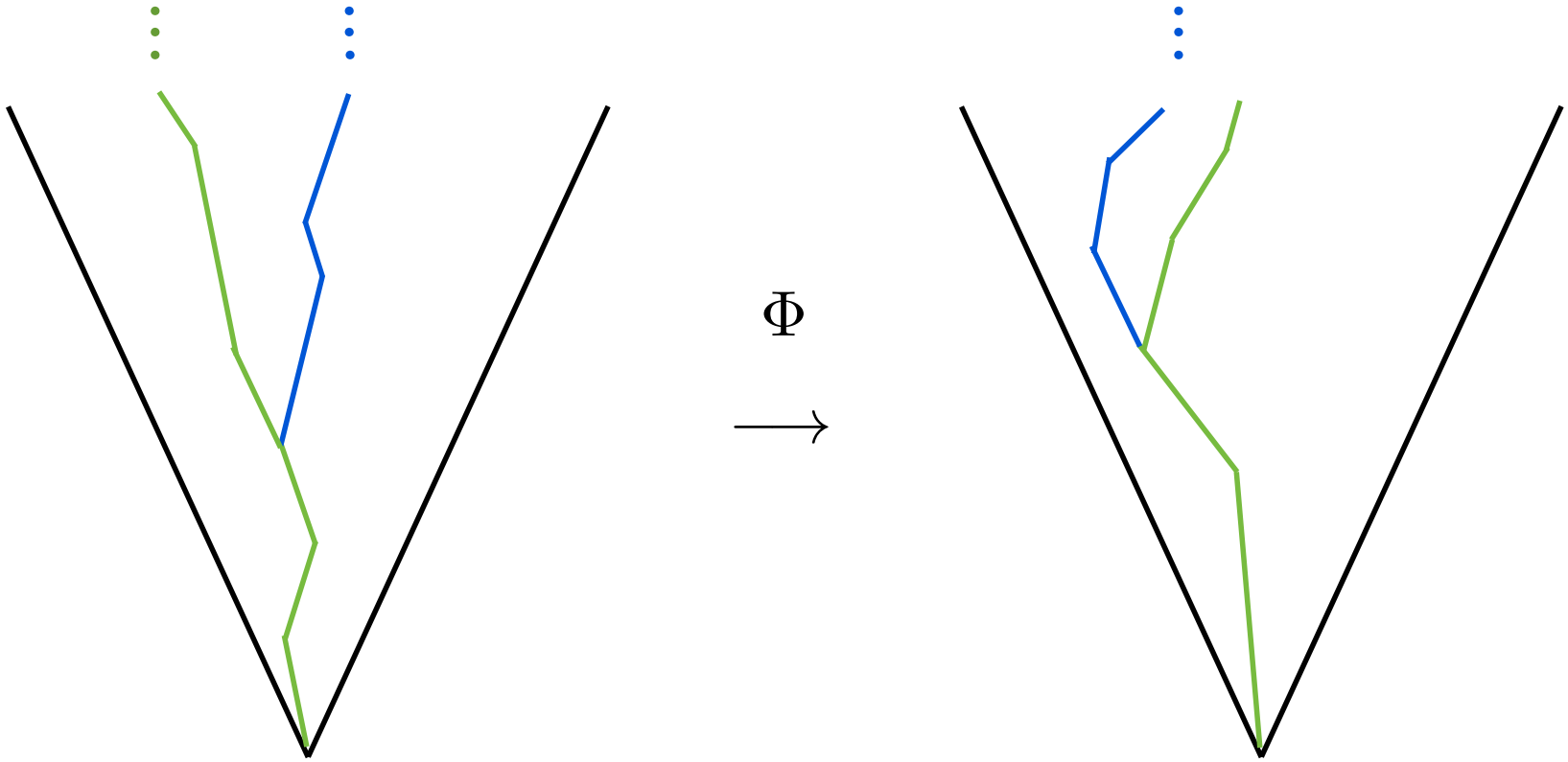

Φ

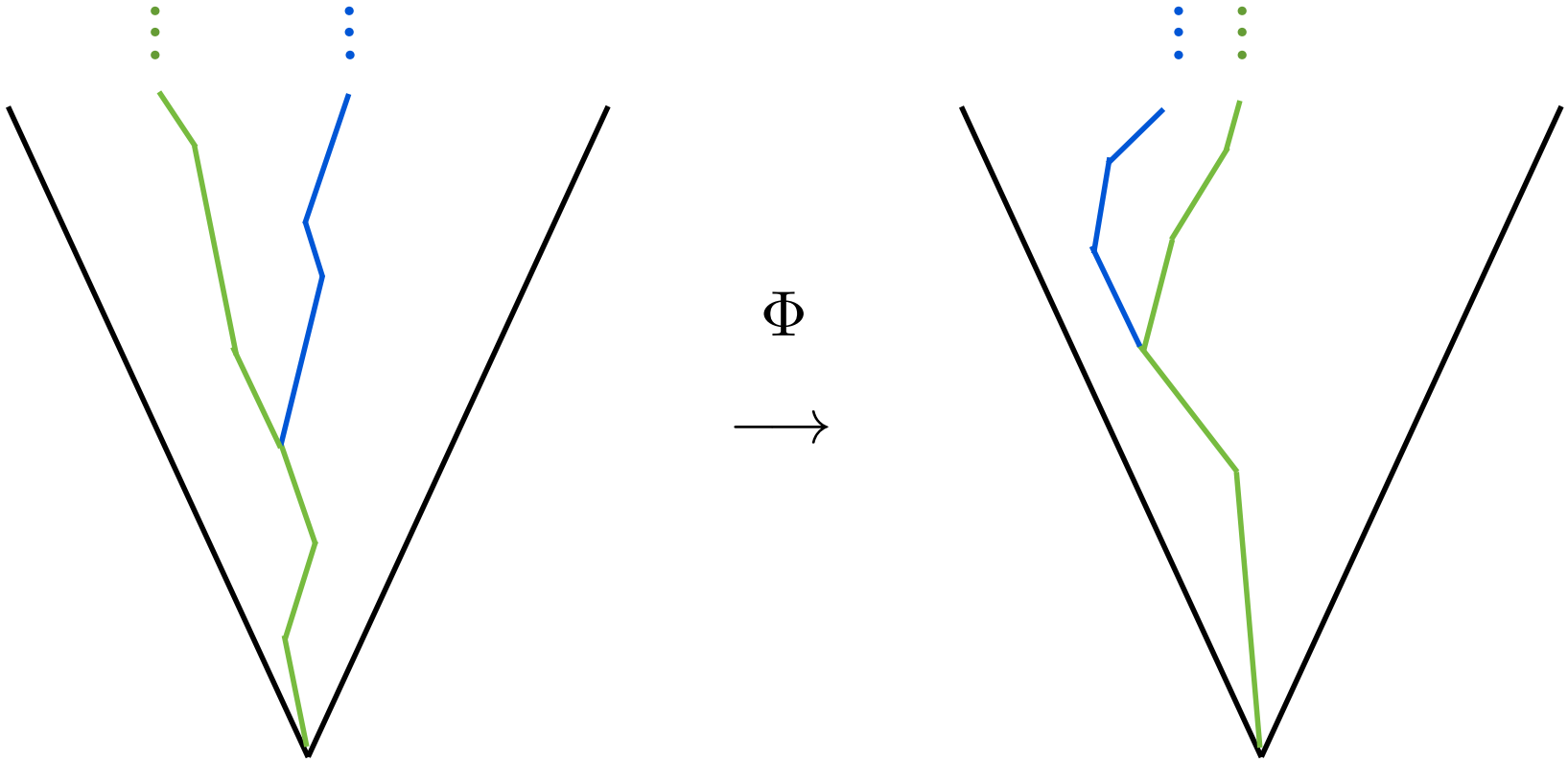

Φ

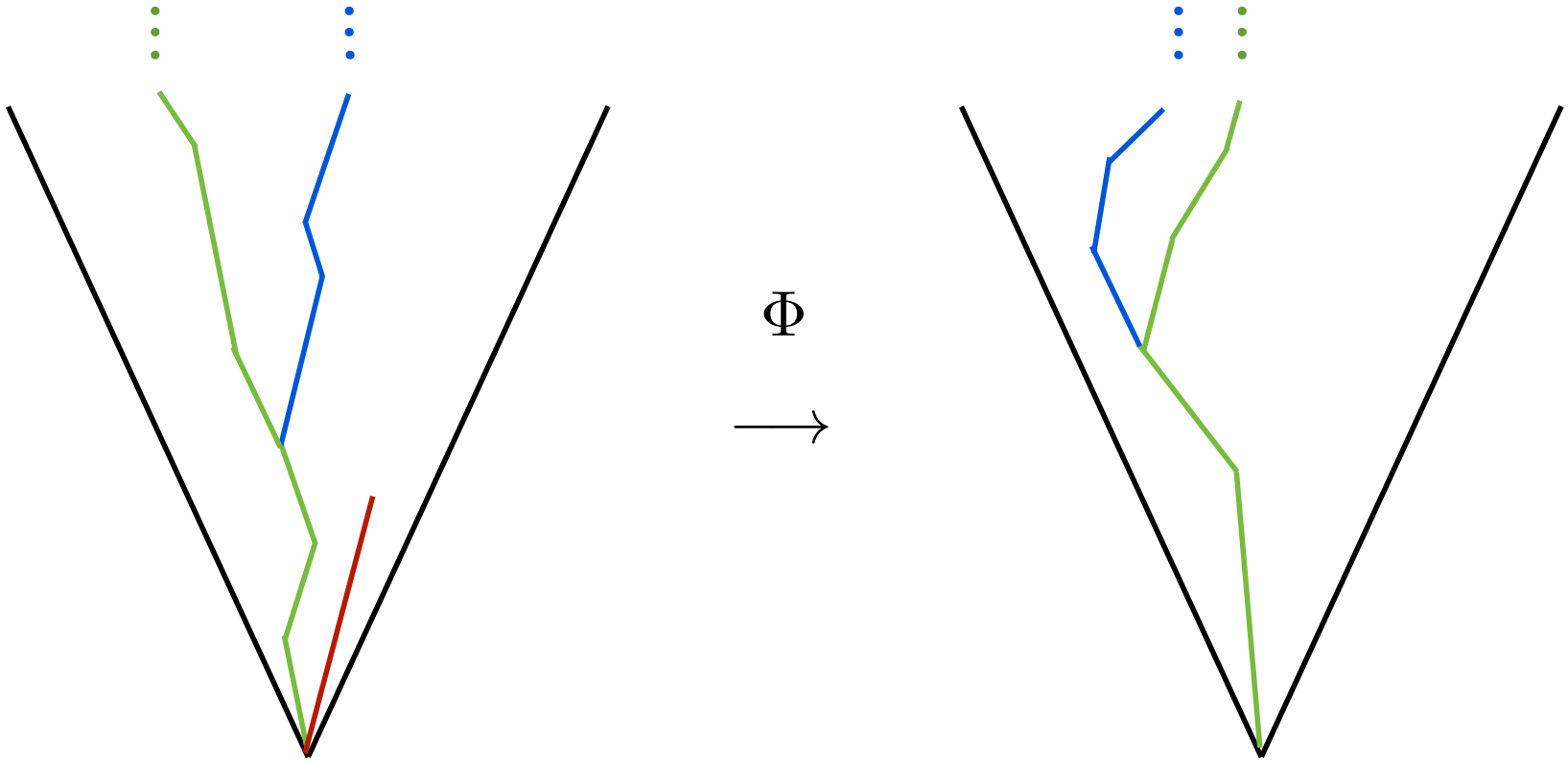

Φ

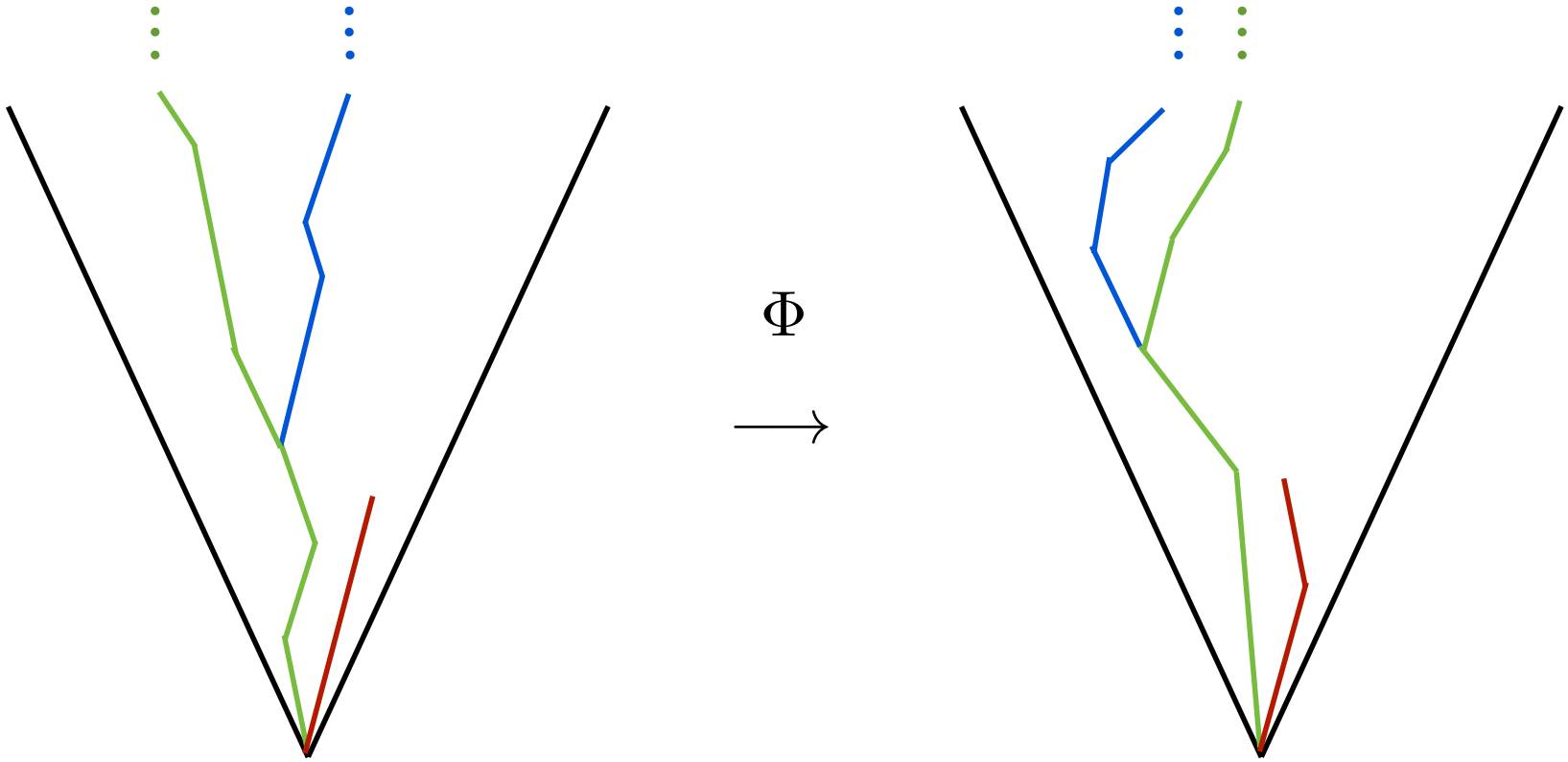

Φ

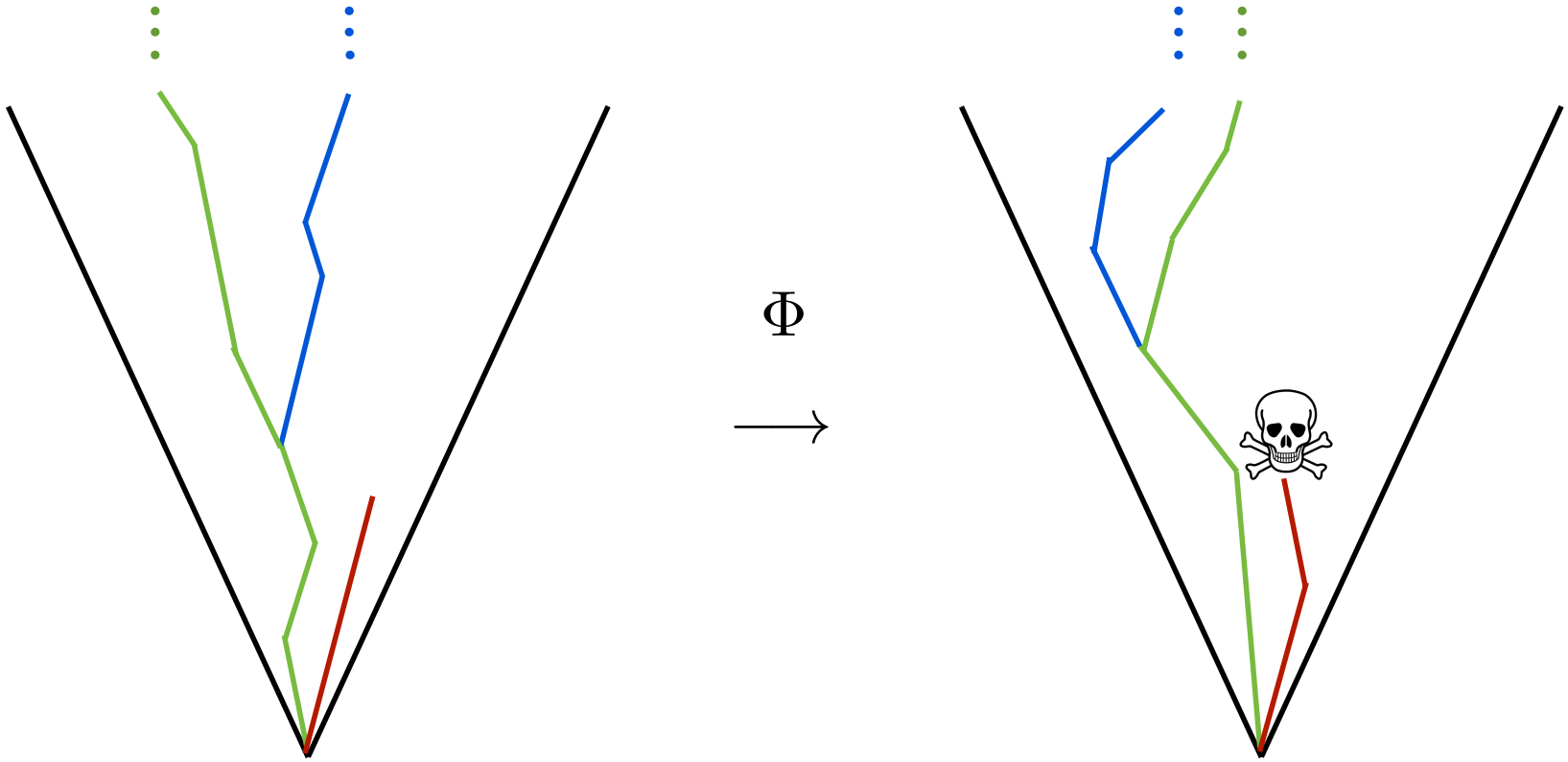

Φ

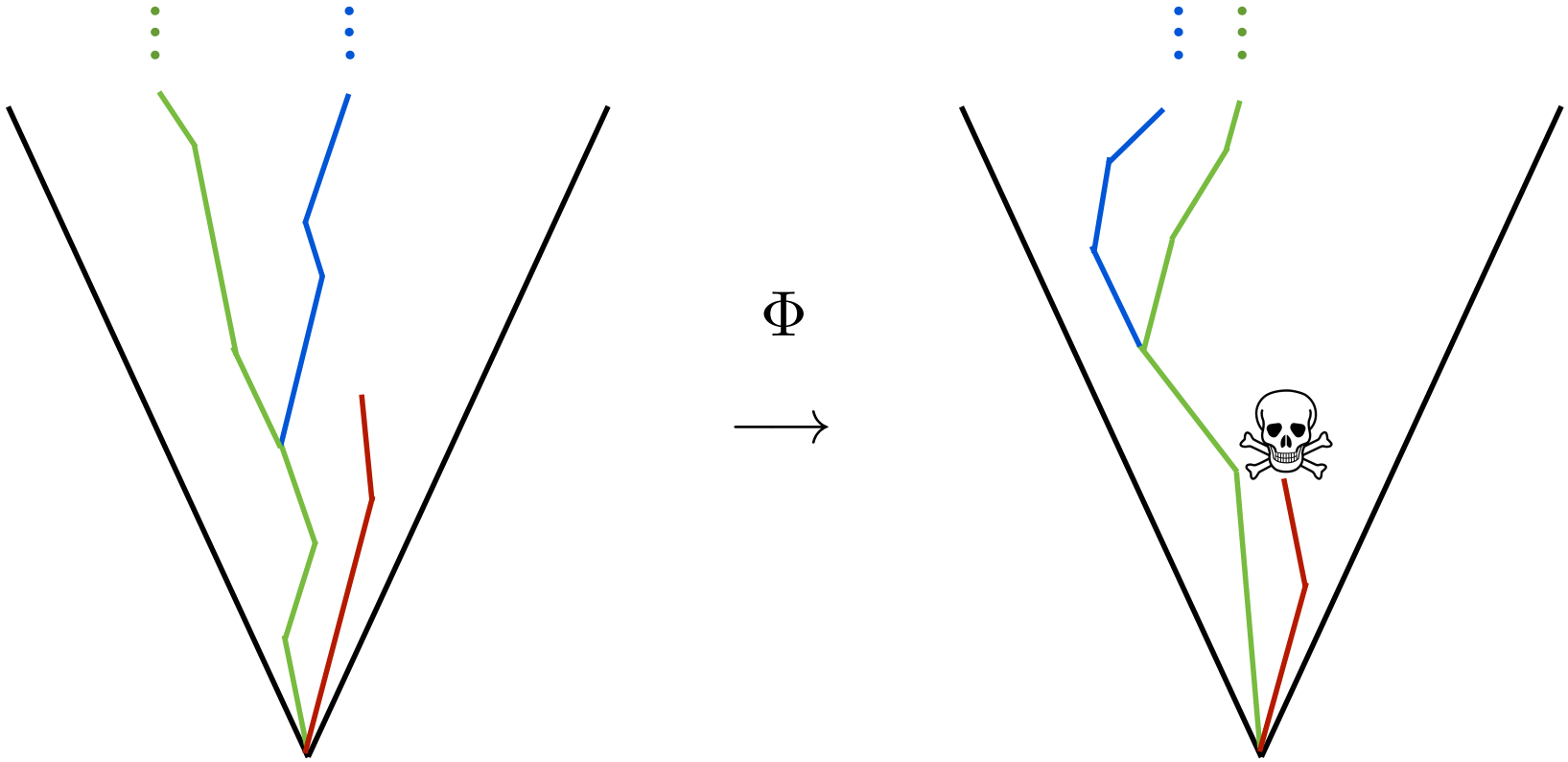

Φ

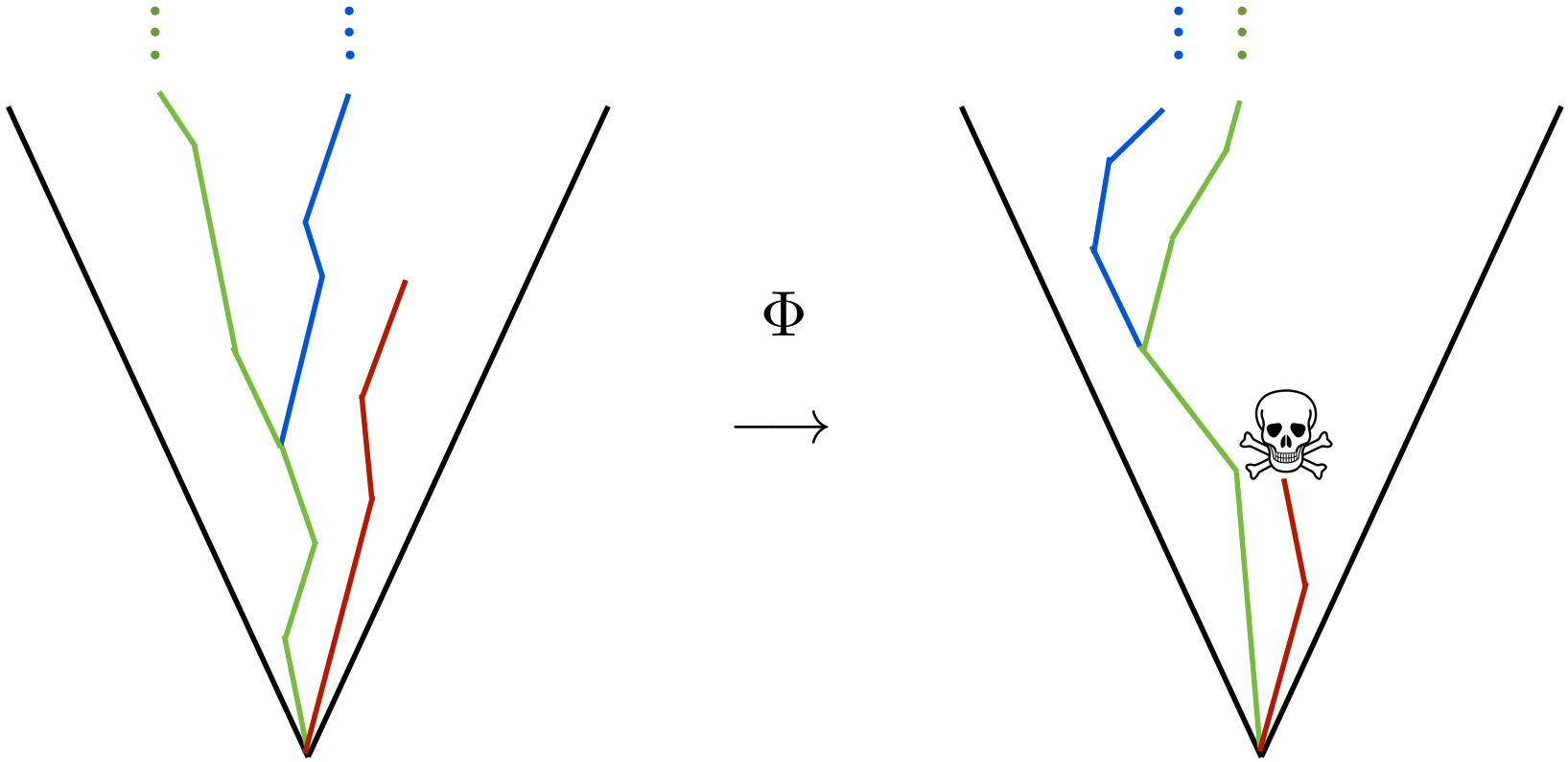

Φ

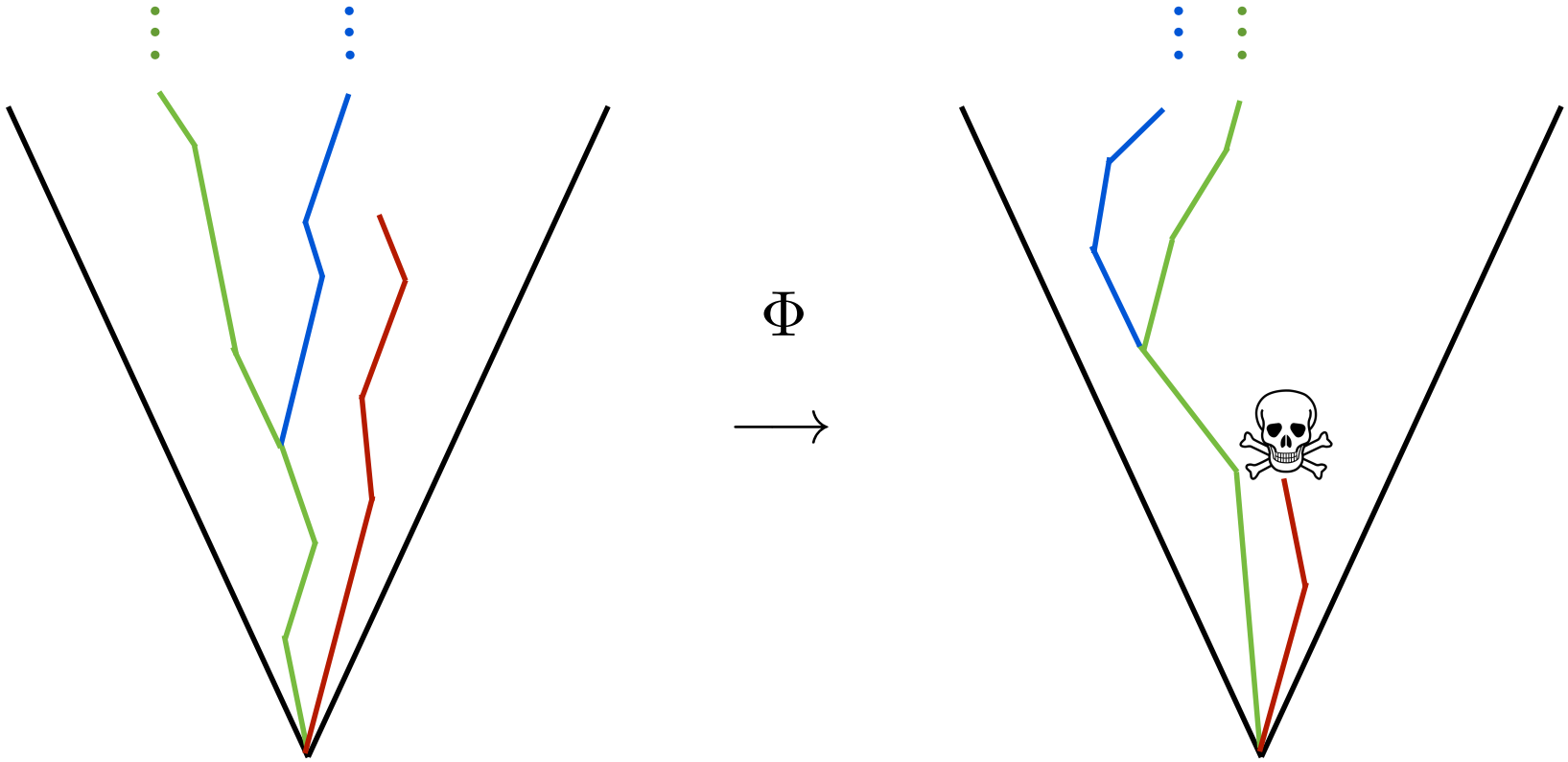

Φ

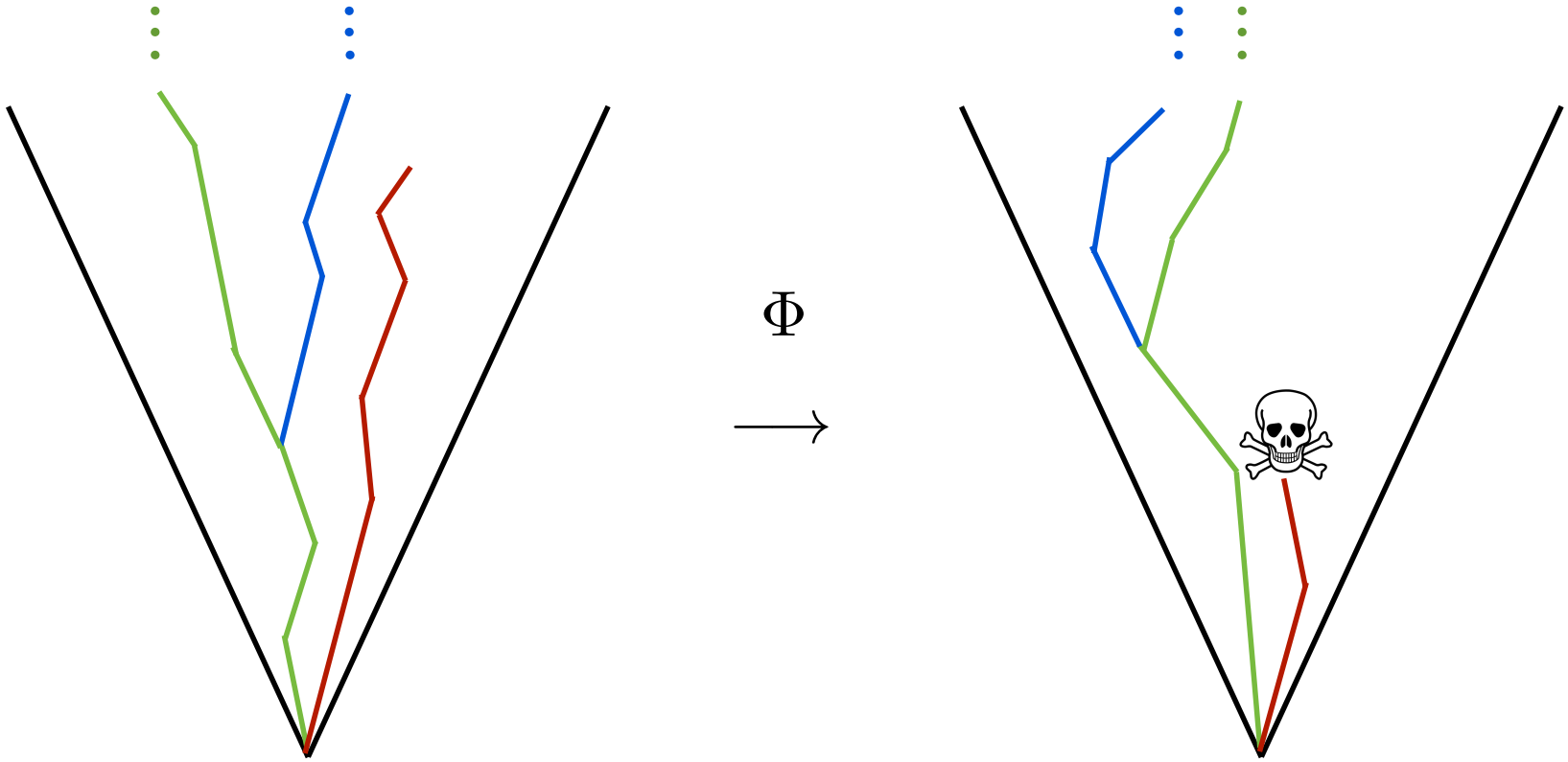

Φ

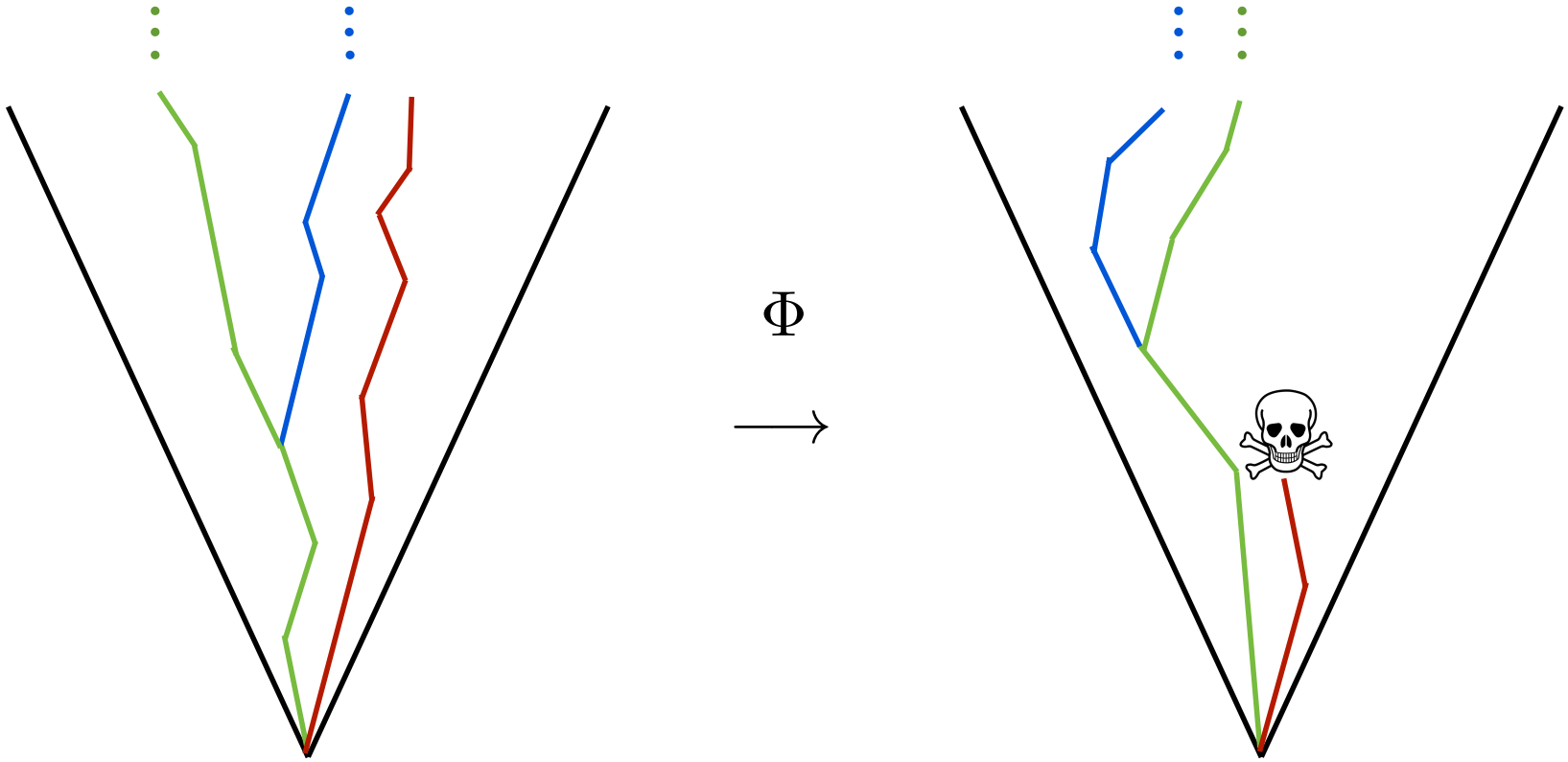

Φ

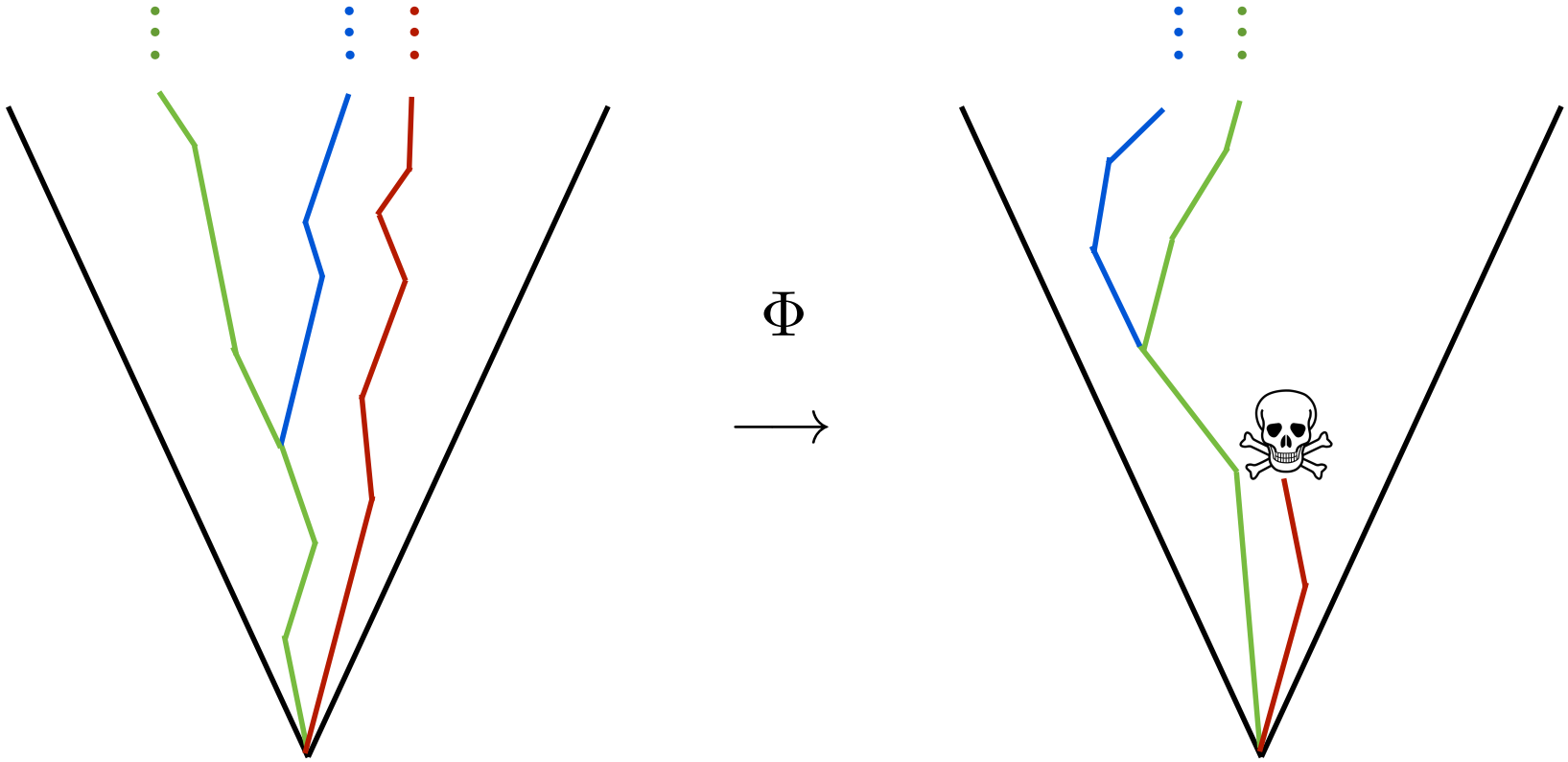

Φ

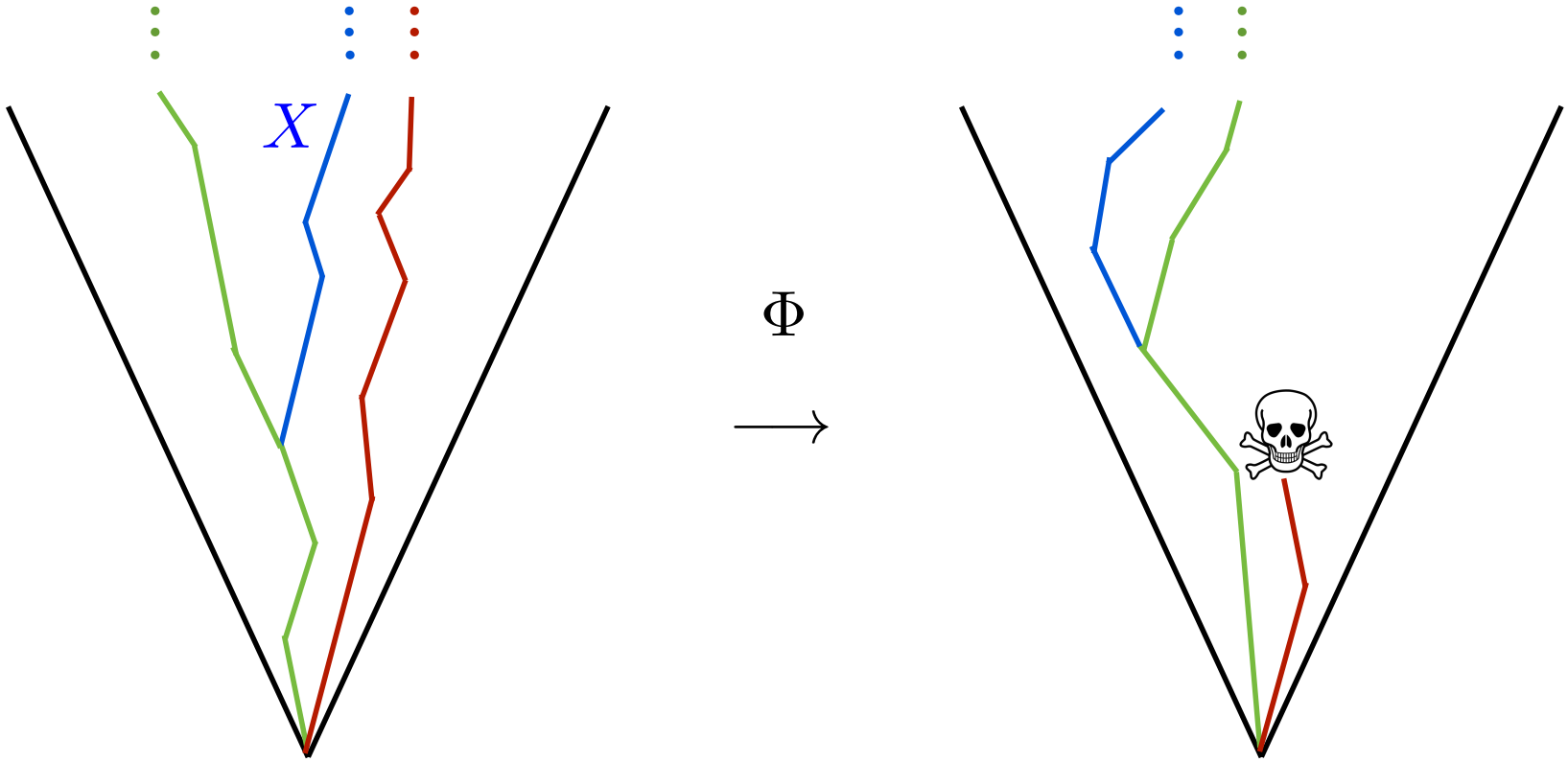

Φ

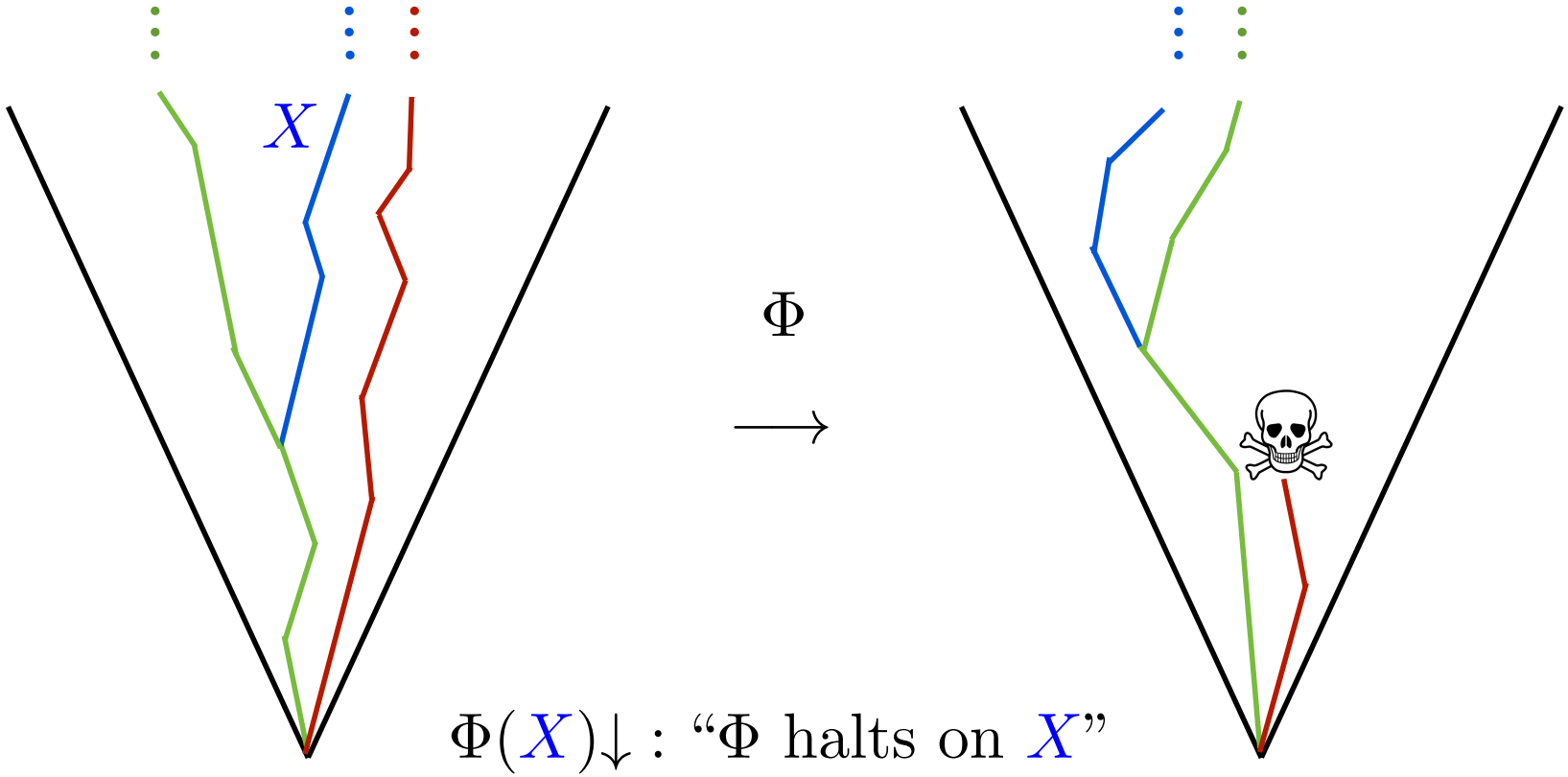

Φ


Φ

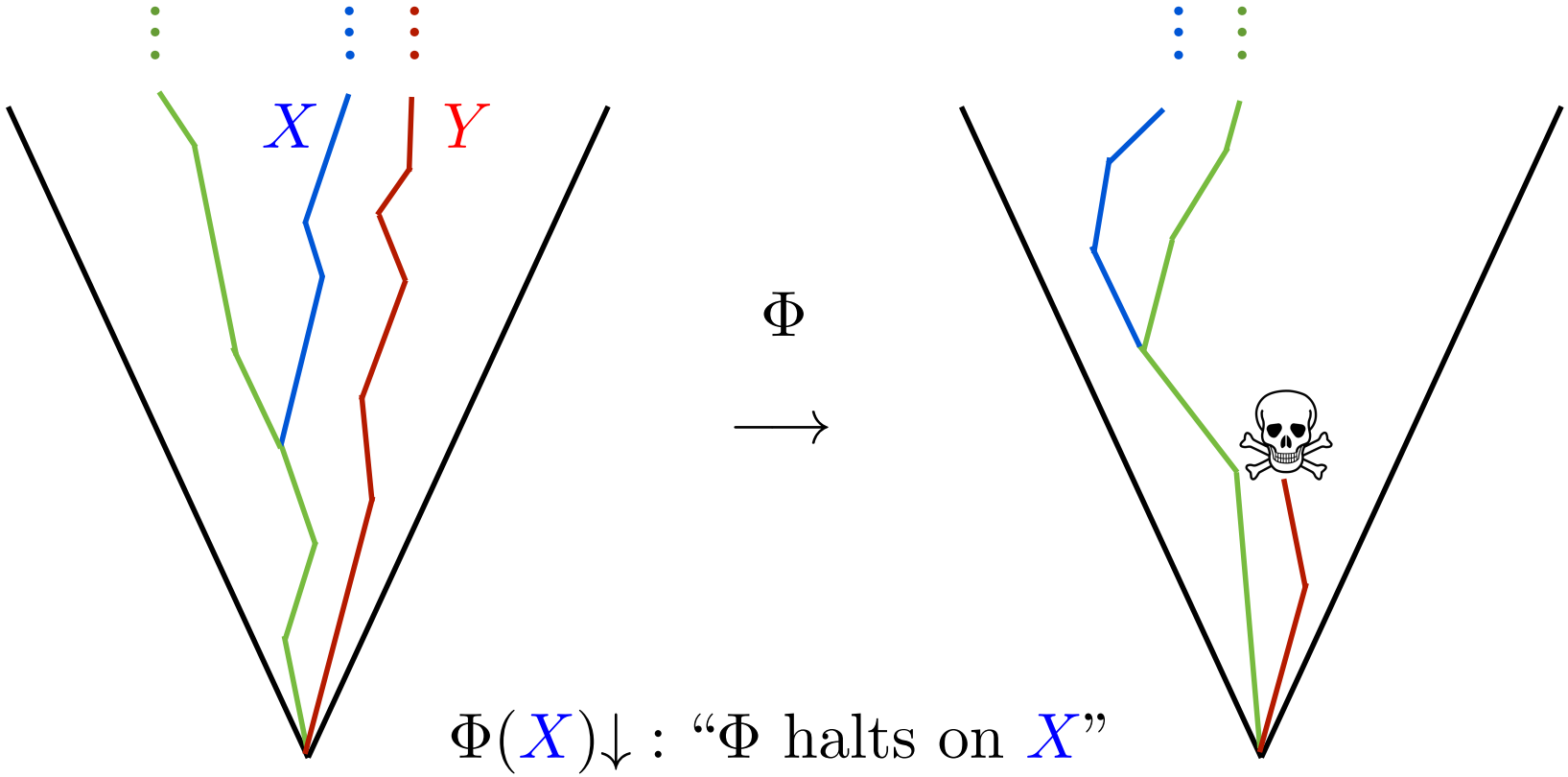

Φ


Φ

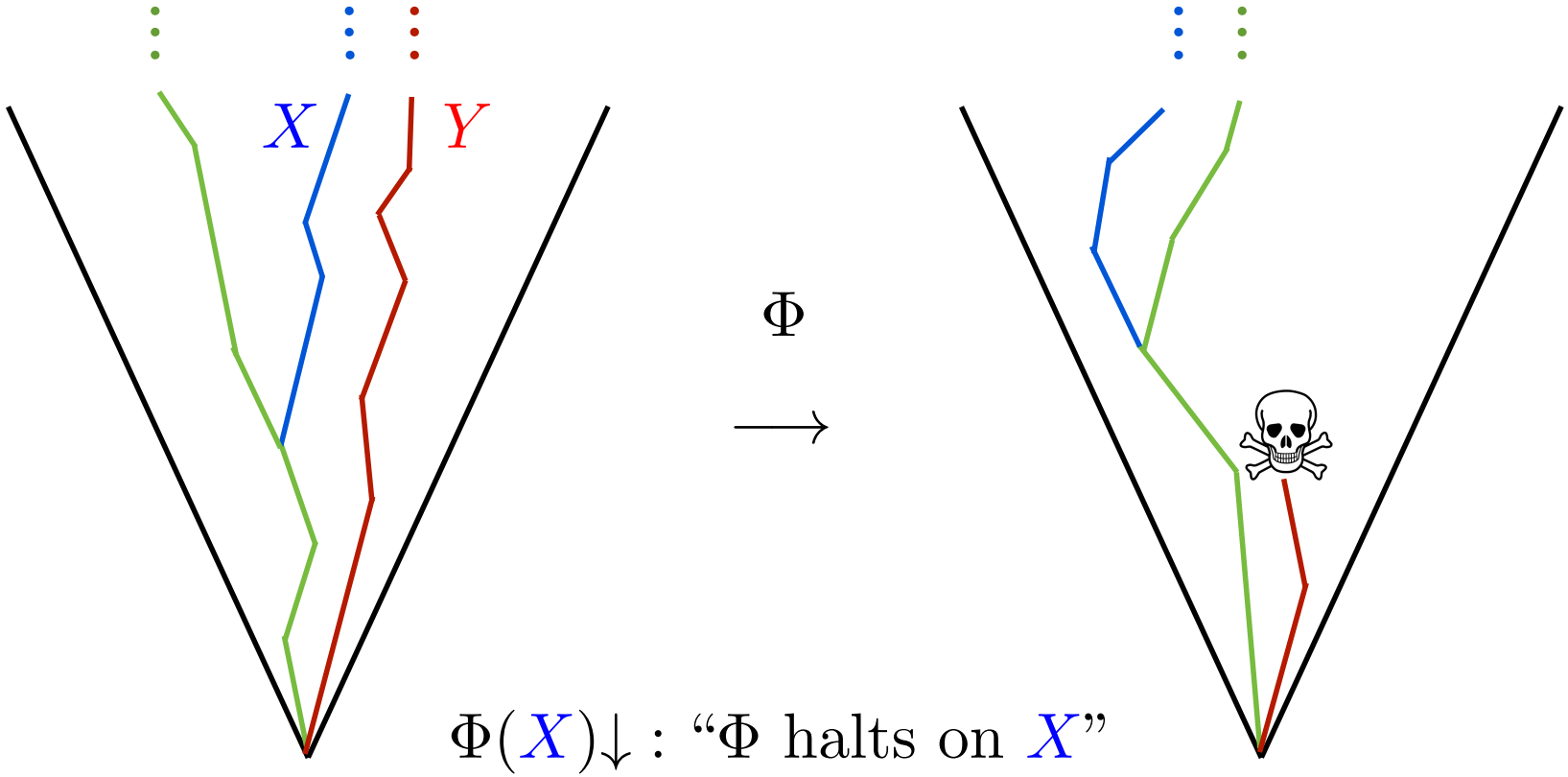

X
Φ



$\Phi(X)\downarrow$: "Φ halts on X"



$\Phi(X)\downarrow$: "Φ halts on X"



$\Phi(X)\downarrow$: "Φ halts on X"

$\Phi(Y)\uparrow$: "Φ diverges on Y"

Turing reducibility

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

One limitation

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 Π^0_1 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.

Let $\mathcal{P} \subseteq 2^\omega$ be a Π^0_1 class (or equivalently, the set of infinite paths through a computable tree).

We investigate whether we can compute some member of $\mathcal{P}$ with positive probability.

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 $K = \{e : \phi_e(e)\downarrow\}$.

3. Negligible Π^0_1 classes

When probabilistic computation fails

Negligible Π^0_1 classes are precisely those classes such that we cannot compute some member with positive probability.

To give a precise definition of negligibility, we need to define what is known as a universal left-c.e. semimeasure.

Left-c.e. semi-measures

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

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

We will be particularly interested in *left-c.e.* semi-measures.

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

Proposition

(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-measures

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

The definition of a negligible classe

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.

Definition

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

The intuition behind negligibility

Let $\mathcal{P}$ be a negligible Π^0_1 class.

$\overline{M}(\mathcal{P}) = 0$ means that the probability of producing some member of $\mathcal{P}$ by means of any Turing functional equipped with any sufficiently random oracle is 0:

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

where $(\Phi_i)_{i\in\omega}$ is an effective enumeration of all Turing functionals.

In particular, for each Φ_i, $\lambda(\{X \in \mathrm{MLR} : \Phi_i(X) \in \mathcal{P}\}) = 0$.

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, Taveneaux)

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

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? No.

Theorem (Bienvenu, Porter, Taveneaux)

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.

4. Deep Π^0_1 classes

Deep classes: the idea

Depth is a property that is stronger than negligibility for Π^0_1 classes.

Instead of considering how difficult it is to produce a path through a Π^0_1 class $\mathcal{P}$, we can consider how difficult it is to produce an *initial segment* of some path through $\mathcal{P}$, level by level.

Deep classes are the "most difficult" of Π^0_1 classes in this respect.

A few more definitions

Let $\mathcal{P} \subseteq 2^\omega$ be a Π^0_1 class.

Let $T^{ext} \subseteq 2^{<\omega}$ be the set of extendible nodes of $\mathcal{P}$,

$$T^{ext} = \{\sigma \in 2^{<\omega} : [\![\sigma]\!] \cap \mathcal{P} \neq \emptyset\}.$$

Thus T^{ext} is the canonical co-c.e. tree such that $\mathcal{P} = [T^{ext}]$ (the set of infinite paths through T^{ext}).

For each $n \in \omega$, T^{ext}_n consists of all strings in T^{ext} of length n.

(I will write T instead of T^{ext} hereafter.)

Deep classes: the definition

Let $\mathcal{P}$ be a Π^0_1 class and let T be the canonical co-c.e. tree such that $\mathcal{P} = [T]$.

$\mathcal{P}$ is a *deep class* if there is some computable, non-decreasing, unbounded function $h : \omega \to \omega$ such that

$$M(T_n) \leq 2^{-h(n)},$$

where $M(T_n) = \sum_{\sigma \in T_n} M(\sigma)$.

That is, the probability of producing some initial segment of a path through $\mathcal{P}$ is effectively bounded above.

Depth vs. negligibility

It's clear that every deep class is negligible.

Is every negligible class deep? No.

Theorem (Bienvenu, Porter, Taveneaux)

There is a negligible class $\mathcal{P}$ that is not deep.

Why use the co-c.e. tree in the definition of depth?

For every Π^0_1 class $\mathcal{P}$ there is a computable tree $T \subseteq 2^{<\omega}$ such that $\mathcal{P} = [T]$.

Why can't we use this computable tree T in the definition of depth?

In general, T will contain non-extendible nodes, so even if we can compute some element in T_n, we still may fail to compute an initial segment of a member of $\mathcal{P}$.

Theorem (Bienvenu, Porter, Taveneaux)

Let T be a computable tree. Then there is no computable order h such that $M(T_n) \leq 2^{-h(n)}$ for every $n \in \omega$.

Examples of deep classes

There are a number of deep classes that naturally arise in computability theory.

We don't, however, have any "natural" examples of negligible classes that aren't deep.

For the sake of brevity, I will only consider one example here.

Consistent completions of Peano arithmetic

The following is implicit in work of Levin and Stephan.

Theorem
The Π^0_1 class of consistent completions of PA is a deep class.

What exactly does this tell us?

Not only can we not probabilistically compute some consistent completion of *PA* with positive probability, but we cannot even hope to produce longer and longer initial segments of a consistent completion of *PA* with sufficiently high probability.

Proving depth

The technique for showing that the class of consistent completions of PA is deep is what we refer to as a *wait and kill* argument.

We need to work with some object that we have control over in some way.

We define a partial computable $\{0,1\}$-valued function ϕ using the recursion theorem.

We *wait* to see a sufficiently large collection of oracles compute some possible extension of ϕ (at some place at which ϕ is currently undefined).

We then define ϕ at this place in such a way as to *kill* off each of these oracles.

Thank you for your attention!

Completions of PA proof sketch 1

Theorem
The Π^0_1 class of consistent completions of PA is a deep class.

Equivalently, we can consider the class $\mathcal{P}$ of total extensions of a universal partial computable $\{0,1\}$-valued function.

Let $u(\langle e, x\rangle) = \phi_e(x)$, where $(\phi_e)_{e\in\omega}$ is an effective enumeration of all partial computable $\{0,1\}$-valued functions.

We will define a partial computable $\{0,1\}$-valued function ϕ_e (where we know e in advance by the recursion theorem), and this will allow us to show that $\mathcal{P}$ is deep.

Completions of PA proof sketch 2

Since we are defining ϕ_e, we have control of the values $u(\langle e, x \rangle)$ for every $x \in \omega$.

Let $(I_k)_{k \in \omega}$ be an effective collection of intervals forming a partition of ω, where we have control of 2^{k+1} values of u inside of I_k for each $k \in \omega$.

Step 1: For each k, we consider the sets

$$E_{k,s} = \{\sigma \in 2^{<\omega} : \sigma \restriction I_k \text{ extends } u_s \restriction I_k\},$$

and wait for a stage s such that

$$M(E_{k,s}) \geq 2^{-k}.$$

Completions of PA proof sketch 3

Step 2: Pick some $y \in I_k$ on which we have yet to define u.

Consider the sets

$$E^0_{k,s}(y) = \{\sigma \in E_{k,s} : \sigma(y) = 0\}$$

and

$$E^1_{k,s}(y) = \{\sigma \in E_{k,s} : \sigma(y) = 1\}.$$

Then $M(E^i_{k,s}(y)) \geq 2^{-(k+1)}$ for $i = 0$ or 1 (or both).

If this holds for $i = 0$, we set $u(y) = 1$; otherwise we set $u(y) = 0$.

Completions of PA proof sketch 4

We repeat the process, going back to Step 1.

We can repeat the process at most 2^{k+1} times (since we have enough values to work with in I_k).

Eventually, we will get stuck at Step 1.

Setting $f(k) = \max(I_k)$, we will have

$$M(\{\sigma : \sigma \restriction f(k) \text{ extends } u\}) \leq 2^{-k}.$$

That is,

$$M(T_{f(k)}) \leq 2^{-k}.$$