Effective randomness for computable probability measures

Laurent Bienvenu\textsuperscript{1} Wolfgang Merkle\textsuperscript{2}

\textsuperscript{1}Laboratoire d’Informatique Fondamentale
Université de Provence
Marseille, France

\textsuperscript{2}Institut für Informatik
Ruprecht-Karls-Universität
Heidelberg, Germany
Outline

1. The Cantor space: topology, measures, computability

2. Our toolbox
   - Martingales
   - Effective randomness and Kolmogorov complexity

3. Main result
Outline

1. The Cantor space: topology, measures, computability

2. Our toolbox
   - Martingales
   - Effective randomness and Kolmogorov complexity

3. Main result
We work in the Cantor space $2^\omega$, which is the set of infinite binary sequences. We endow $2^\omega$ with the product topology, i.e. the topology generated by the sets

$$[w] = \{ \alpha \in 2^\omega : w \sqsupseteq \alpha \} \quad (\text{where } w \in 2^*)$$
We work in the Cantor space $2^{\omega}$, which is the set of infinite binary sequences. We endow $2^{\omega}$ with the product topology, i.e. the topology generated by the sets

$$[w] = \{ \alpha \in 2^{\omega} : w \sqsubseteq \alpha \} \quad (\text{where } w \in 2^*)$$

Every open set $\mathcal{U}$ is of the form

$$\mathcal{U} = \bigcup_{w \in A} [w] \quad (\text{where } A \subseteq 2^*)$$
We work in the Cantor space $2^\omega$, which is the set of infinite binary sequences. We endow $2^\omega$ with the product topology, i.e. the topology generated by the sets

$$[w] = \{ \alpha \in 2^\omega : w \sqsupseteq \alpha \} \quad (\text{where } w \in 2^*)$$

Every open set $\mathcal{U}$ is of the form

$$\mathcal{U} = \bigcup_{w \in A} [w] \quad (\text{where } A \subseteq 2^*)$$

If $A$ is a computably enumerable, $\mathcal{U}$ is said to be effectively open (or $\Sigma^0_1$).
We work in the Cantor space $2^\omega$, which is the set of infinite binary sequences. We endow $2^\omega$ with the product topology, i.e. the topology generated by the sets

$$[w] = \{ \alpha \in 2^\omega : w \sqsubseteq \alpha \} \quad \text{(where } w \in 2^* \text{)}$$

Every open set $U$ is of the form

$$U = \bigcup_{w \in A} [w] \quad \text{(where } A \subseteq 2^* \text{)}$$

If $A$ is a computably enumerable, $U$ is said to be effectively open (or $\Sigma^0_1$).

A set $C$ is **effectively closed** (or $\Pi^0_1$) if its complement is effectively open.
The arithmetic hierarchy.

We define inductively: a $\Sigma^0_{n+1}$ is an effective union of $\Pi^0_n$ sets and a $\Pi^0_{n+1}$ is an effective intersection of $\Sigma^0_n$ sets.
To specify a measure on $\mathbb{2}^\omega$, it suffices to specify the measure of $[w]$ for all $w \in 2^*$ (Caratheodory’s extension theorem).

Example of a measure
To specify a measure on $2^\omega$, it suffices to specify the measure of $[w]$ for all $w \in 2^*$ (Caratheodory’s extension theorem).
This allows us to define the notion of **computable measure**:

**Definition**

We say that $\mu$ is a **computable measure** if

$$w \mapsto \mu([w])$$

is a computable function.
A central notion in probability theory: the equivalence of two probability measures.

**Definition**

Two measures are *equivalent* if they have the same nullsets.

Classically, it is a useful notion (e.g. the Radon-Nikodym theorem).
The Cantor space: topology, measures, computability

Our toolbox

Main result

**Theorem**

The following are equivalent:

1. $\mu$ and $\nu$ are two equivalent measures
2. $\mu$ and $\nu$ have the same $G_\delta$ nullsets
3. $\mu$ and $\nu$ have the same closed nullsets

To what extent can this theorem be effectivized?

L. Bienvenu, W. Merkle

Effective randomness for computable probability measures
Theorem

The following are equivalent:

1. $\mu$ and $\nu$ are two equivalent measures
2. $\mu$ and $\nu$ have the same $G_\delta$ nullsets
3. $\mu$ and $\nu$ have the same closed nullsets

To what extent can this theorem be effectivized?
... at least part of it can be!

**Theorem**

*Two computable measures are equivalent iff they have the same $\Pi^0_2$ nullsets.*
Proof. One direction is trivial. For the other one, suppose that \( \mu \) and \( \nu \) are not equivalent i.e. there exists \( X \subseteq 2^\omega \) such that, for example, \( \mu(X) \geq m > 0 \) and \( \nu(X) = 0 \).
**Proof.** One direction is trivial. For the other one, suppose that \( \mu \) and \( \nu \) are not equivalent i.e. there exists \( X \subseteq 2^\omega \) such that, for example, \( \mu(X) \geq m > 0 \) and \( \nu(X) = 0 \).

This means that for all \( n \) there exists an open set \( U_n \subseteq X \) such that \( \mu(U_n) \geq m \) and \( \nu(U_n) < 2^{-n} \).
**Proof.** One direction is trivial. For the other one, suppose that $\mu$ and $\nu$ are not equivalent i.e. there exists $X \subseteq 2^\omega$ such that, for example, $\mu(X) \geq m > 0$ and $\nu(X) = 0$.

This means that for all $n$ there exists an open set $U_n \subseteq X$ such that $\mu(U_n) \geq m$ and $\nu(U_n) < 2^{-n}$.

Hence, there exists $V_n$ finitely generated such that $\mu(V_n) \geq m - 2^{-n}$ and $\nu(V_n) < 2^{-n}$.
**Proof.** One direction is trivial. For the other one, suppose that $\mu$ and $\nu$ are not equivalent i.e. there exists $X \subseteq 2^\omega$ such that, for example, $\mu(X) \geq m > 0$ and $\nu(X) = 0$.

This means that for all $n$ there exists an open set $U_n \subseteq X$ such that $\mu(U_n) \geq m$ and $\nu(U_n) < 2^{-n}$.

Hence, there exists $V_n$ finitely generated such that $\mu(V_n) \geq m - 2^{-n}$ and $\nu(V_n) < 2^{-n}$. And such a $V_n$ can be found effectively.
**Proof.** One direction is trivial. For the other one, suppose that $\mu$ and $\nu$ are not equivalent i.e. there exists $X \subseteq 2^\omega$ such that, for example, $\mu(X) \geq m > 0$ and $\nu(X) = 0$.

This means that for all $n$ there exists an open set $U_n \subseteq X$ such that $\mu(U_n) \geq m$ and $\nu(U_n) < 2^{-n}$.

Hence, there exists $V_n$ finitely generated such that $\mu(V_n) \geq m - 2^{-n}$ and $\nu(V_n) < 2^{-n}$. And such a $V_n$ can be found effectively.

Then, consider the $\Pi^0_2$ set

$$G = \bigcap_{k} \bigcup_{n \geq k} V_n$$

One has $\mu(G) \geq m$ and $\nu(G) = 0$. 
This leaves the second part open:

**Question**

If two computable measures have the same $\Pi^0_1$ nullsets, are they necessarily equivalent?
Outline

1. The Cantor space: topology, measures, computability

2. Our toolbox
   - Martingales
   - Effective randomness and Kolmogorov complexity

3. Main result
We begin with the notion of martingale, inspired by the corresponding notion in classical probability theory:

**Definition**

A *μ-martingale* is a function $d : 2^* \rightarrow \mathbb{R}_+$ such that for all $w \in 2^*$:

$$
\mu(w)d(w) = \mu(w0)d(w0) + \mu(w1)d(w1)
$$

![Diagram of measure μ and μ-martingale](image-url)
Definition

A martingale succeeds on a sequence $\alpha \in 2^\omega$ if

$$\sup_n d(\alpha_0...\alpha_n) = +\infty$$
### Theorem (Ville’s inequality)

Let \( d \) be a \( \mu \)-martingale. For all \( k > 0 \):

\[
\mu \left( \left\{ \alpha \in \mathbb{2}^\omega : \sup_n d(\alpha_0...\alpha_n) \geq k \right\} \right) \leq \frac{1}{k}
\]

### Corollary

Let \( d \) be a \( \mu \)-martingale.

\[
\mu \left( \left\{ \alpha \in \mathbb{2}^\omega : \text{d succeeds on } \alpha \right\} \right) = 0
\]
There exists a very close correspondence between measures and martingales:

**Theorem**

For all (computable) measures $\mu$ and $\nu$, $\frac{\nu}{\mu}$ is a (computable) $\mu$-martingale.

Conversely, every (computable) $\mu$-martingale $d$ can be written as $d = \frac{\nu}{\mu}$ for some (computable) measure $\nu$. 
There exists a very close correspondence between measures and martingales:

**Theorem**

For all (computable) measures $\mu$ and $\nu$, $\frac{\nu}{\mu}$ is a (computable) $\mu$-martingale.

Conversely, every (computable) $\mu$-martingale $d$ can be written as $d = \frac{\nu}{\mu}$ for some (computable) measure $\nu$.

**Theorem**

Two measures $\mu$ and $\nu$ are equivalent iff:

$$\mu\left(\{\alpha : \frac{\mu}{\nu}\text{ succeeds on }\alpha\}\right) = 0$$

$$\nu\left(\{\alpha : \frac{\nu}{\mu}\text{ succeeds on }\alpha\}\right) = 0$$
Outline

1. The Cantor space: topology, measures, computability

2. Our toolbox
   - Martingales
   - Effective randomness and Kolmogorov complexity

3. Main result
The goal of algorithmic randomness (or effective randomness) is to define what it means for an individual sequence to be random.
The goal of algorithmic randomness (or effective randomness) is to define what it means for an *individual* sequence to be random.

The theory started in 1919, with R. von Mises and his notion of *kollektiv*, which turned out to be too weak a notion of randomness.
The goal of algorithmic randomness (or effective randomness) is to define what is means for an **individual** sequence to be random.

The theory started in 1919, with R. von Mises and his notion of *kollektiv*, which turned out to be too weak a notion of randomness.

The first satisfactory (and the best up till now) approach was given by Kolmogorov-Chaitin-Solomonov for finite objects, and by Martin-Löf for infinite ones.
A \( \Pi^0_2 \) is a \( \mu \)-Martin-Löf nullset if it is the effective intersection of \( \Sigma^0_1 \) sets \( \{U_n : n \in \mathbb{N}\} \) such that \( \mu(U_n) \leq 2^{-n} \).

A sequence \( \alpha \in 2^\omega \) is \( \mu \)-Martin-Löf random if it belongs to no \( \mu \)-Martin-Löf nullset.
A weaker notion of randomness....

**Definition (Kurtz)**

A sequence $\alpha \in 2^\omega$ is $\mu$-weakly random if it belongs to no $\Pi^0_1$ set of $\mu$-measure 0.

**Proposition**

*Two computable measures $\mu$ and $\nu$ have the same $\Pi^0_1$ nullsets iff they have the same weakly random sequences.*
The following theorem characterizes weak randomness by means of martingales:

**Theorem (Wang)**

A sequence $\alpha$ is not $\mu$-weakly random if there exists a $\mu$-martingale $d$ and a computable order $h$ such that $d(\alpha_0...\alpha_n) \geq h(n)$ for all $n$. 
Definition

Let $w \in 2^*$. The **Kolmogorov complexity** of $w$, denoted by $C(w)$ is the length of the shortest program which outputs $w$.

Notice that, up to a fixed additive constant, $C(w) \leq |w|$.

Also, $C$ is approximable from above but non-computable.
The Cantor space: topology, measures, computability

Our toolbox

Main result

Martingales

Effective randomness and Kolmogorov complexity

Theorem (Miller-Yu)

A sequence \( \alpha \) is Martin-Löf random iff for every computable function \( f : \mathbb{N} \to \mathbb{N} \) such that \( \sum 2^{-f(n)} < +\infty \)

\[
C(\alpha_0...\alpha_n) \geq n - f(n) + O(1)
\]

Theorem (Miller-Nies-Stephan-Terwijn)

A sequence \( \alpha \) is \( \emptyset' \)-Martin-Löf random iff

\[
\exists \infty n \ C(\alpha_0...\alpha_n) \geq n - O(1)
\]
The Cantor space: topology, measures, computability

Our toolbox

Main result

Martingales

Effective randomness and Kolmogorov complexity

Theorem (Miller-Yu)

A sequence $\alpha$ is $\lambda$-Martin-Löf random iff for every computable function $f : \mathbb{N} \to \mathbb{N}$ such that $\sum 2^{-f(n)} < +\infty$

$$C^*(\alpha_0...\alpha_n) \geq n - f(n) + O(1)$$

Theorem (Miller-Nies-Stephan-Terwijn)

A sequence $\alpha$ is $\emptyset'\text{-}\lambda$-Martin-Löf random iff

$$\exists \infty n \quad C^*(\alpha_0...\alpha_n) \geq n - O(1)$$
The Cantor space: topology, measures, computability

Our toolbox
- Martingales
- Effective randomness and Kolmogorov complexity

Main result

L. Bienvenu, W. Merkle
Effective randomness for computable probability measures
The Cantor space: topology, measures, computability

Our toolbox

Main result

**Theorem**

Two measures $\mu$ and $\nu$ are equivalent if:

$$\mu \left( \left\{ \alpha : \frac{\mu}{\nu} \text{ succeeds on } \alpha \right\} \right) = 0$$

$$\nu \left( \left\{ \alpha : \frac{\nu}{\mu} \text{ succeeds on } \alpha \right\} \right) = 0$$

**Proposition**

Two computable measures $\mu$ and $\nu$ have the same $\Pi^0_1$ nullsets iff they have the same weakly random sequences.

**Theorem**

A sequence $\alpha$ is not $\mu$-weakly random if there exists a $\mu$-martingale $d$ and a computable order $h$ such that $d(\alpha_0...\alpha_n) \geq h(n)$ for all $n$. 

L. Bienvenu, W. Merkle

Effective randomness for computable probability measures
The idea is to construct a measure $\mu$ such that
The idea is to construct a measure $\mu$ such that

$$\frac{\lambda}{\mu}$$ succeeds on a set $S$ such that $\lambda(S) > 0$
The idea is to construct a measure $\mu$ such that

- $\frac{\lambda}{\mu}$ succeeds on a set $S$ such that $\lambda(S) > 0$
- $\frac{\lambda}{\mu}$ succeeds slowly on sequences in $S$ (i.e. not faster than any computable order)
This brings us to the notion of hyperimmunity:

### Definition
A sequence $\alpha \in 2^\omega$ has hyperimmune degree if it Turing-computes a function $f : \mathbb{N} \to \mathbb{N}$ such that

$$\forall g \text{ computable } f \not\leq g$$

... equivalently, $\alpha$ has hyperimmune degree if it computes an order $h : \mathbb{N} \to \mathbb{N}$ such that

$$\forall g \text{ computable order } g \not\leq h$$
Good news...

**Theorem (Martin)**

\[ \lambda \left( \{ \alpha \in 2^\omega : \alpha \text{ has hyperimmune degree} \} \right) = 1 \]
Good news...

Theorem (Martin)

\[\lambda\left(\{\alpha \in 2^\omega : \alpha \text{ has hyperimmune degree}\}\right) = 1\]

Bad news...

Theorem (Kurtz)

*There exists no operator \(\mathcal{H}\) such that \(\mathcal{H}^\alpha\) is a slow (in the sense above) order for almost all \(\alpha\).*
We need a more constructive version:

**Theorem (Nies-Stephan-Terwijn)**

Every $\lambda$-\$\emptyset'\$-Martin-Löf random sequence $\alpha$ has hyperimmune degree.

**Proof.** Suppose that $C^*(\alpha_0...\alpha_n) \geq n - c$ for infinitely many $n$’s. Then the function

$$h : n \mapsto \#\{k \leq n : C^*(\alpha_0...\alpha_k) \geq k - c\}$$

is a slow order.
We are ready for the construction of a computable measure $\mu$ with the desired properties.
We are ready for the construction of a computable measure $\mu$ with the desired properties.

Recall that

$$\lambda - \emptyset' - MLR = \{ \alpha : \exists c \exists \infty n \ C^*(\alpha_0...\alpha_n) \geq n - c \}$$

Thus, there must be a $c_0 > 0$ such that

$$\lambda\{ \alpha : \exists \infty n \ C^*(\alpha_0...\alpha_n) \geq n - c_0 \} > 0$$

This will be our set $S$!
The Cantor space: topology, measures, computability

Our toolbox

Main result

L. Bienvenu, W. Merkle

Effective randomness for computable probability measures
Verification.

In $S$, all the elements are $\lambda$-$\emptyset'$-$MLR$, hence $\lambda$-$WR$. 
Verification.

In $S$, all the elements are $\lambda$-$\emptyset'$-MLR, hence $\lambda$-WR.

On these elements, $\frac{\mu}{\lambda}$ decreases slowly (and tends to 0) hence $\frac{\lambda}{\mu}$ increases slowly and tends to $+\infty$. 
Verification.

In $S$, all the elements are $\lambda$-$\emptyset'$-$MLR$, hence $\lambda$-$WR$.

On these elements, $\frac{\mu}{\lambda}$ decreases slowly (and tends to 0) hence $\frac{\lambda}{\mu}$ increases slowly and tends to $+\infty$.

$\frac{\lambda}{\mu}$ is basically the best computable $\mu$-martingale against $\lambda$-random elements. Hence, every other computable $\mu$-martingale succeeds as slowly as it on elements of $S$. 
Verification.

In $S$, all the elements are $\lambda$-$\emptyset'$-MLR, hence $\lambda$-WR.

On these elements, $\frac{\lambda}{\mu}$ decreases slowly (and tends to 0) hence $\frac{\lambda}{\mu}$ increases slowly and tends to $+\infty$.

$\frac{\lambda}{\mu}$ is basically the best computable $\mu$-martingale against $\lambda$-random elements. Hence, every other computable $\mu$-martingale succeeds as slowly as it on elements of $S$.

Hence, all elements of $S$ are $\mu$-WR.
Verification.

In $S$, all the elements are $\lambda$-$\emptyset'$-$MLR$, hence $\lambda$-$WR$.

On these elements, $\frac{\mu}{\lambda}$ decreases slowly (and tends to 0) hence $\frac{\lambda}{\mu}$ increases slowly and tends to $+\infty$.

$\frac{\lambda}{\mu}$ is basically the best computable $\mu$-martingale against $\lambda$-random elements. Hence, every other computable $\mu$-martingale succeeds as slowly as it on elements of $S$.

Hence, all elements of $S$ are $\mu$-$WR$.

Outside $S$, $\mu$ and $\lambda$ are equal up to a multiplicative constant, hence for all $\alpha \in S$: $\alpha \in \lambda$-$WR \iff \alpha \in \mu$-$WR$. 
We have seen how the theory of algorithmic randomness can be used to solve which have \textit{a priori} nothing to do with it. Other questions arise from what we have seen. For example:

**Question**

One can define a new equivalence relation between computable measures:

\[ \mu \equiv \nu \text{ iff } \mu - MLR = \nu - MLR \]

How does this new relation compare to the classical one?
We have seen how the theory of algorithmic randomness can be used to solve which have \textit{a priori} nothing to do with it. Other questions arise from what we have seen. For example:

**Question**

One can define a new equivalence relation between computable measures:

\[ \mu \equiv \nu \text{ iff } \mu - \text{MLR} = \nu - \text{MLR} \]

How does this new relation compare to the classical one?

THANK YOU