Effective randomness for computable probability measures

Laurent Bienvenu[1] Wolfgang Merkle[2]

[1]Laboratoire d'Informatique Fondamentale
Université de Provence
Marseille, France

[2]Institut für Informatik
Ruprecht-Karls-Universität
Heidelberg, Germany

Outline

Outline

We work in the Cantor space 2^ω, which is the set of infinite binary sequences. We endow 2^ω with the product topology, i.e. the topology generated by the sets

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

We work in the Cantor space 2^ω, which is the set of infinite binary sequences. We endow 2^ω with the product topology, i.e. the topology generated by the sets

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

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

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

We work in the Cantor space 2^ω, which is the set of infinite binary sequences. We endow 2^ω with the product topology, i.e. the topology generated by the sets

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

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

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

If A is a computably enumerable, $\mathcal{U}$ is said to be **effectively open** (or Σ^0_1).

We work in the Cantor space 2^ω, which is the set of infinite binary sequences. We endow 2^ω with the product topology, i.e. the topology generated by the sets

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

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

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

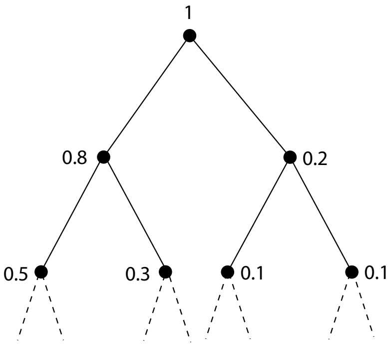
If A is a computably enumerable, $\mathcal{U}$ is said to be **effectively open** (or Σ^0_1).

A set $\mathcal{C}$ is **effectively closed** (or Π^0_1) if its complement is effectively open.

The arithmetic hierarchy.

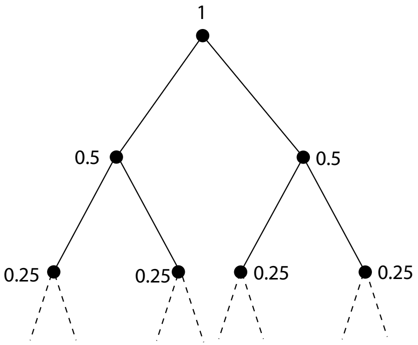
We define inductively: a Σ^0_{n+1} is an effective union of Π^0_n sets and a Π^0_{n+1} is an effective intersection of Σ^0_n sets.

To specify a measure on 2^ω, 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^ω, it suffices to specify the measure of $[w]$ for all $w \in 2^*$ (Caratheodory's extension theorem).



Lebesgue measure λ

This allows us to define the notion of **computable measure**:

Definition

We say that μ 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 usefull notion (e.g. the Radon-Nikodym theorem).

Theorem

The following are equivalent:

1. μ *and* ν *are two equivalent measures*
2. μ *and* ν *have the same* G_δ *nullsets*
3. μ *and* ν *have the same closed nullsets*

Theorem

The following are equivalent:

1. μ *and* ν *are two equivalent measures*
2. μ *and* ν *have the same* G_δ *nullsets*
3. μ *and* ν *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 Π^0_2 nullsets.

Proof. One direction is trivial. For the other one, suppose that μ and ν 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 μ and ν 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 $\mathcal{U}_n \subseteq X$ such that $\mu(\mathcal{U}_n) \geq m$ and $\nu(\mathcal{U}_n) < 2^{-n}$.

Proof. One direction is trivial. For the other one, suppose that μ and ν 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 $\mathcal{U}_n \subseteq X$ such that $\mu(\mathcal{U}_n) \geq m$ and $\nu(\mathcal{U}_n) < 2^{-n}$.

Hence, there exists $\mathcal{V}_n$ finitely generated such that $\mu(\mathcal{V}_n) \geq m - 2^{-n}$ and $\nu(\mathcal{V}_n) < 2^{-n}$.

Proof. One direction is trivial. For the other one, suppose that μ and ν 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 $\mathcal{U}_n \subseteq X$ such that $\mu(\mathcal{U}_n) \geq m$ and $\nu(\mathcal{U}_n) < 2^{-n}$.

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

Proof. One direction is trivial. For the other one, suppose that μ and ν 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 $\mathcal{U}_n \subseteq X$ such that $\mu(\mathcal{U}_n) \geq m$ and $\nu(\mathcal{U}_n) < 2^{-n}$.

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

Then, consider the Π^0_2 set

$$G = \bigcap_k \bigcup_{n \geq k} \mathcal{V}_n$$

One has $\mu(G) \geq m$ and $\nu(G) = 0$.

This leaves the second part open:

Question

If two computable measures have the same Π^0_1 nullsets, are they necessarily equivalent?

Outline

1. The Cantor space: topology, measures, computability

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

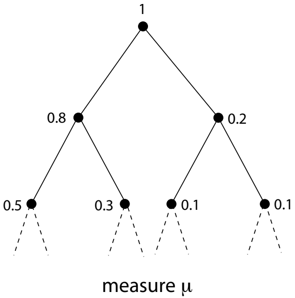
3. Main result

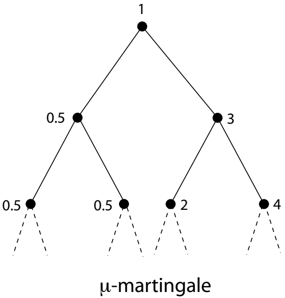
We begin with the notion of martingale, inspired by the corresponding notion in classical probability theory:

Definition

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

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





Definition

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

$$sup_n \; d(\alpha_0 \ldots \alpha_n) = +\infty$$

Theorem (Ville's inequality)

Let d be a μ-martingale. For all $k > 0$:

$$\mu\Big(\{\alpha \in 2^\omega : \sup_n d(\alpha_0 \ldots \alpha_n) \geq k\}\Big) \leq 1/k$$

Corollary

Let d be a μ-martingale.

$$\mu\Big(\{\alpha \in 2^\omega : d \text{ succeeds on } \alpha\}\Big) = 0$$

There exists a very close correspondence between measures and martingales:

Theorem

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

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

There exists a very close correspondence between measures and martingales:

Theorem

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

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

Theorem

Two measures μ and ν 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

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

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 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.

Definition

A Π^0_2 is a μ-Martin-Löf nullset if it is the effective intersection of Σ^0_1 sets $\{\mathcal{U}_n : n \in \mathbb{N}\}$ such that $\mu(\mathcal{U}_n) \leq 2^{-n}$.

Definition

A sequence $\alpha \in 2^\omega$ is μ-Martin-Löf random if it belongs to no μ-Martin-Löf nullset.

A weaker notion of randomness....

Definition (Kurtz)

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

Proposition

Two computable measures μ and ν have the same Π^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 α is not μ-weakly random if there exists a μ-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.

Theorem (Miller-Yu)

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

$$C(\alpha_0 \ldots \alpha_n) \geq n - f(n) + O(1)$$

Theorem (Miller-Nies-Stephan-Terwijn)

A sequence α *is* $\emptyset'$*-Martin-Löf random iff*

$$\exists^\infty n \; C(\alpha_0 \ldots \alpha_n) \geq n - O(1)$$

Theorem (Miller-Yu)

A sequence α is λ-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 α is $\emptyset'$-λ-Martin-Löf random iff

$$\exists^\infty n \; C^*(\alpha_0 ... \alpha_n) \geq n - O(1)$$

Outline

Theorem

Two measures μ and ν are equivalent if:

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

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

Proposition

Two computable measures μ and ν have the same Π^0_1 nullsets iff they have the same weakly random sequences.

Theorem

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

The idea is to construct a measure μ such that

The idea is to construct a measure μ such that

- $\frac{\lambda}{\mu}$ succeeds on a set S such that $\lambda(S) > 0$

The idea is to construct a measure μ 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

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

Definition

... equivalently, α 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)

$$\lambda\Big(\{\alpha \in 2^\omega : \alpha \textit{ has hyperimmune degree } \}\Big) = 1$$

Good news...

Theorem (Martin)

$$\lambda\Big(\{\alpha \in 2^\omega : \alpha \textit{ has hyperimmune degree }\}\Big) = 1$$

Bad news...

Theorem (Kurtz)

There exists no operator $\mathfrak{h}$ such that $\mathfrak{h}^\alpha$ is a slow (in the sense above) order for almost all α.

We need a more constructive version:

Theorem (Nies-Stephan-Terwijn)

Every λ-$\emptyset'$-Martin-Löf random sequence α 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 μ with the desired properties.

We are ready for the construction of a computable measure μ with the desired properties.

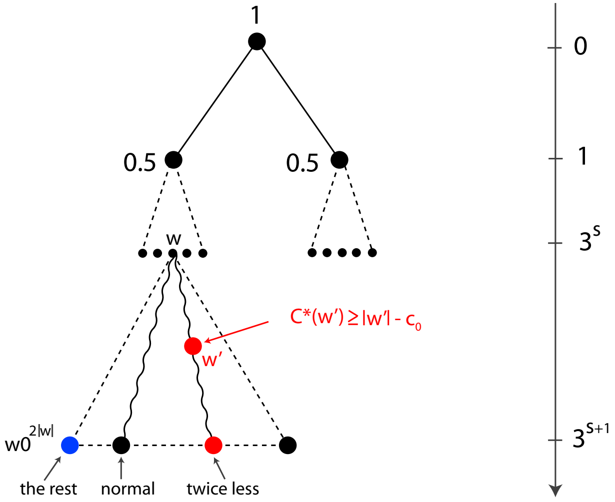
Recall that

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

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!


1
0.5
0.5
w
C*(w') ≥ |w'| - c0
w'
w0^2|w|
the rest
normal
twice less
0
1
3^s
3^s+1

Verification.

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

Verification.

In S, all the elements are λ-$\emptyset'$-MLR, hence λ-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 λ-$\emptyset'$-MLR, hence λ-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 μ-martingale against λ-random elements. Hence, every other computable μ-martingale succeeds as slowly as it on elements of S.

Verification.

In S, all the elements are λ-$\emptyset'$-MLR, hence λ-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 μ-martingale against λ-random elements. Hence, every other computable μ-martingale succeeds as slowly as it on elements of S.

Hence, all elements of S are μ-WR.

Verification.

In S, all the elements are λ-$\emptyset'$-MLR, hence λ-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 μ-martingale against λ-random elements. Hence, every other computable μ-martingale succeeds as slowly as it on elements of S.

Hence, all elements of S are μ-WR.

Outside S, μ and λ are equal up to a multiplicative constant, hence for all $\alpha \in S$: $\alpha \in \lambda$-$WR \Leftrightarrow \alpha \in \mu$-$WR$.

We have seen how the theory of algorithmic randomness can be used to solve which have *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:

$$\mu \equiv \nu \quad \textit{iff} \quad \mu - MLR = \nu - MLR$$

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 *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:

$$\mu \equiv \nu \quad iff \quad \mu - MLR = \nu - MLR$$

How does this new relation compare to the classical one?

This question is adressed in: *L. Bienvenu, W. Merkle. Effective randomness for computable probability measures. Electronic Notes in Theoretical Computer Science 167, pp 117-130 (2007).*

THANK YOU