Algorithmic randomness over general spaces

Kenshi Miyabe

Jun 10, 2010

Background

Motivation

Probability theory by Kolmogorov's measure-theoretic approach has been developed and has many applications.
His approach is also called an axiomatic approach.
Some people claim that math does not tell us meaning of probability and others say that this is the merit because we can use it freely.
Then interpretation of probability is a difficult problem.
Probability theory is the only one theory that has applications to unpredictable systems.
However in many cases it is not appropriate to use probability, for example, tomorrow's weather and quantum mechanics.
So we need a theory which

- gives interpretation of probability and
- has applications to unpredictable systems without probability.

I believe algorithmic randomness does!!

History of probability

The theory of probability began by Pascal (1623-1662).
One big question is why it is so late.
There are two plausible explanations.

- Unpredictable event had been used to ask for God.
- Aristotle's philosophy.

> That there is not a science of the accidental is manifest. *The Metaphysiscs, Aristotle*

To Aristotle, even accidental event has causes called "tyche" or "automaton" which cannot immediately be understod.

History of randomness

In 1919 von Mises tried to define random sequences to formalize probability theory, only to fail.
Kolmogorov was not satisfied with his approach because probability theory does not say whether an individual element is random.
To do that he founded algorithmic randomness and tried to use it to formalize probability theory.
But this trial does not succeed because of some reasons.
In 1990s Dawid and Vovk use the idea of algorithmic randomness to give another formalization of probability theory, whose basis is not algorithmic randomness but game theory.
Many people will agree with the fact that algorithmic randomness is a natural basis of probability theory.

Algorithmic randomness

Algorithmic randomness has more potential than probability theory in the following mean.

- It can consider individual random elements.
- It has a strong connection to computability theory.
- It gives us a natural philosophical meaning of probability.

Then it will have many applications like probability theory, i.e., statistics, machine learning, AI.
To do that, we need to overcome some difficulties.
One of them is the restriction of the space where we can consider randomness.
Then today's my talk is about the first step.

Martin-Löf randomness

We have three approaches to define randomness.
A *Cantor space* 2^ω is the set of all infinite binary sequences.
The topology is the one generated by the cylinder sets $[w] = \{A \in 2^\omega \;:\; w \preceq A\}$.
The measure μ is induced by $\mu([w]) = 2^{-|w|}$.
A open set W is *c.e.* if $W = \bigcup_{w \in V} [w]$ for some c.e.

Definition 1 (Martin-Löf 1966)**.** *A Martin-Löf test is a uniformly c.e. open set U_n with $\mu(U_n) \leq 2^{-n}$.*
A sequence A is Martin-Löf random if it passes all Martin-Löf test, that is, $A \notin \bigcap_n U_n$.

Theorem 2. *There is a universal Martin-Löf test.*
The class of Martin-Löf random sequences has measure 1.

Complexity and martingales

The *prefix-free Kolmogorov complexity* K of w is defined as $K(w) = \{|u| : U(u) = w\}$ where U is the universal Turing machine.

Theorem 3 (Schnorr 1971)**.** *A sequence A is Martin-Löf random iff $K(A \upharpoonright n) \geq n - O(1)$.*

A martingale is a function $d : 2^* \to \mathbb{R}^+$ satisfying $2d(w) = d(w0) + d(w1)$ for all $w \in 2^*$.

Theorem 4 (Schnorr 1971)**.** *A sequence A is Martin-Löf random iff no c.e. martingale succeeds on A, that is, $\sup_n d(A \upharpoonright n) < \infty$ for all d.*

This coincidence is one of the reasons of the fact that Martin-Löf randomness is considered a natural randomness.

Generalization of computability

Computability on $\mathbb{N}$ was defined by Church, Turing and others in 1930s.
Computable analysis studies computability on $\mathbb{R}$.
There are many approaches to this problem.

- Via representations by Hauck, Kreitz and Weihrauch.
- Via sequential computability and effective uniform continuity by Pour-El and Richards (and this is generalized to metric spaces by Yasugi).
- Ko's approach, Domain Theory, Markov's approach, etc.

We will mostly use the approach via representations, which is called Type-2 Theory of Effectivity (TTE).
This is because this approach can be naturally adapted to randomness.

Generalization of randomness

History of generalization of randomness is the following.

	measure	complexity	martingales
$2^\omega, \mathbb{R}$	Martin-Löf 1966	Schnorr 1971	Schnorr 1971
compact CMS	-	Gács 2005	?
CMS	-	Hoyrup&Rojas 2009	?
CTS	Hertling&Weihrauch 1998	?	?

Here “CMS” means a computable metric space and “CTS” means a computable topological space.

Our goal is to define randomness by complexity and martingales on a computable topological space and to confirm the naturalness for randomness.

TTE

Representations

- $f : A \rightrightarrows B$ denotes a multi-function.
- $f :\subseteq A \to B$ denotes a partial function.
- $f : A \to B$ denotes a total function.

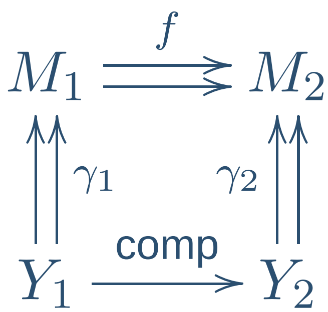
A function $f :\subseteq \Sigma^\omega \to \Sigma^\omega$ is *computable* if it is computed by Type-2 machine.

Informally, a *Type-2 machine* is a Turing machine, which reads from input tapes with finite or infinite inscription, operates on work tapes and write one-way to an output tape.

A *multi-representation* of a set M is a surjective function $\gamma : Y \rightrightarrows M$ where $Y \in \{\Sigma^*, \Sigma^\omega\}$.

A point is γ-*computable* if it has a computable representation by γ.

A function $f : M_1 \rightrightarrows M_2$ is (γ_1, γ_2)-*computable* if it has a computable realization.



Computable topological spaces

Definition 5 (Hertling and Weihrauch 2009)**.** *A computable topological space is a 4-tuple $\mathbf{X} = (X, \tau, \beta, \nu)$ such that*

- *(X, τ) is a topological T_0-space,*
- *$\nu :\subseteq \Sigma^* \to \beta$ is a notation of a base β of τ,*
- *$dom(\nu)$ is recursive and*
- *$\nu(u) \cap \nu(v) = \bigcup\{\nu(w) \ : \ (u, v, w) \in S\}$ for all $u, v \in dom(\nu)$ for some r.e. set $S \subseteq (dom(\nu))^3$.*

Remark 6. The definition has been revised many times.

Example 7. (i) *(real line) Define $\mathbf{R} = (\mathbb{R}, \tau_{\mathbb{R}}, \beta, \nu)$ such that $\tau_{\mathbb{R}}$ is the real line topology and ν is a canonical notation of the set of all open intervals with rational endpoints.*

(ii) *(lower unit interval) Define $\mathbf{I}_< = (\mathbb{I}, \tau_<, \beta_<, \nu_<)$ such that $\nu_<(w) = \{x \ : \ 0 \leq q < x \leq 1$ and $q \in \nu_{\mathbb{Q}}\}$. The representation δ for $\mathbf{I}_<$ is denoted by $\rho_<$.*

Representations for a computable topological space

To encode sequences of Σ^* in Σ^ω, we use the notation $u \ll p$ for $u \in \Sigma^*$ and $p \in \Sigma^\omega$.

Definition 8. *Let $\mathbf{X} = (X, \tau, \beta, \nu)$ be an effective topological space.*
Define a representation $\delta :\subseteq \Sigma^\omega \to X$ of the points as

$$x = \delta(p) \iff (\forall w \in \Sigma^*)(w \ll p \iff x \in \nu(w))$$

and a representation $\theta :\subseteq \Sigma^\omega \to \tau$ of the set of open sets as

$$W = \theta(p) \iff \begin{cases} w \ll p \Rightarrow w \in \mathit{dom}(\nu) \\ W = \bigcup\{\nu(w) \ : \ w \ll p\} \end{cases} .$$

Generalization

Computability of measures

First we define computable probability measures.
For a topological space X, let $P(X)$ be the space of probability Borel measures on X with the usual topology of weak convergence.
It is known that

- X is countable iff $P(X)$ is countable,
- X is separable iff $P(X)$ is separable,
- X is compact iff $P(X)$ is compact,
- and other similar relations of the descriptive complexity.

Theorem 9 (Hoyrup and Rojas 2009)**.** *The similar relation holds for the property of a computable metric space.*

Theorem 10 (My result)**.** *So is a computable topological space.*

Definition 11. *A measure is computable if it is a computable point in the computable topological space.*

Definition by tests

Definition 12 (essentially Hertling and Weihrauch 1998)**.** *A measure test over $\mathbf{X}$ is a uniformly θ-computable sequence $\{U_n\}$ of open sets with $\mu(U_n) \le 2^{-n}$ for all n.*
A point x is measure μ-random over $\mathbf{X}$ if $x \notin \bigcap_n U_n$ for each measure test $\{U_n\}$.

Inspired by a uniform test by Gács and Levin, we give characterization by a function test.
Let $\overline{\rho}_<$ be the representation of lower real line with infinity.

Definition 13. *A function test over $\mathbf{X}$ is a $(\delta, \overline{\rho}_<)$-computable function $f : X \to \overline{\mathbb{R}}$ such that $\mu f = \int_X f d\mu \le 1$.*

Theorem 14. *A point x is measure μ-random iff $f(x) < \infty$ for each function test f.*

Definition by martingales

We use some terminologies from measure theory.
Let $(X, \mathcal{A}, \mu)$ be a measure space.
A *filtration* is a sequence of sub-σ-algebra $\{\mathcal{A}_n\}$ such that $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ for each n.
A sequence of $\mathcal{A}$-measurable functions $\{f_n\}$ is called *a martingale* if $\int f_n d\mu < \infty$ and $\int_A f_n d\mu = \int_A f_{n+1} d\mu$ for all $A \in \mathcal{A}_n$.

Theorem 15. *A point x is measure μ-random iff $\sup_n f_n(x) < \infty$ for each $([\nu_{\mathbb{N}}, \delta], \overline{\rho}_<)$-computable martingale $\{f_n\}$*

Proof idea. Let $U_{k,m} = \{y \ : \ \sup_{n \leq m} \sup f_n(y) > 2^k\}$ and use Doob's maximal inequality. □

Universal complexity of sequences

Let $f :\subseteq 2^* \to \Sigma^\omega$ be a prefix-free computable function.

$$K_f(p) = \min\{\sigma \;:\; f(\sigma) = p\}.$$

If p is not computable, $K_f(p) = \infty$ for all f.

Theorem 16. *There exists a prefix-free computable function $U :\subseteq 2^* \to \Sigma^\omega$ such that*

$$(\forall f)(\exists c)(\forall p)(\exists q)\theta(p) = \theta(q) \text{ and } K_U(q) \leq K_f(p) + c.$$

In the following we write K to mean K_U.

Definition by complexity

Let $\psi^{-}(p) = X \backslash \theta(p)$.

Definition 17. *A point x is complexity μ-random if*

$$x \in \psi^{-}(p) \Rightarrow K(p) \geq -\log \mu\psi^{-}(p) - O(1).$$

This definition is not a straightforward generalization but coincides with the definition on a Cantor space.

Definition 18. *The base $\beta = \{\nu(i)\}$ is complete if all equivalent bases are computably reducible to ν.*

There exists such a complete base.

Theorem 19. *A point x is complexity μ-random iff*

$$x \in \xi(u) \Rightarrow K(u) \geq -\log \mu\xi(u) - O(1)$$

where $\xi(u) = \nu(u)^{c}$.

When they coincide.

Do measure randomness and complexity randomness coincide?
In general they are different.

Example 20. For lower unit interval $\mathbf{I}_<$ *and Lebesgue measure* μ*,*

- *the set of measure* μ*-random points is* $\{1\}$ *and*
- *the set of complexity* μ*-random points is* $\{0\}$*.*

However they coincide on a computable metric space with a computable measure, so on a Cantor space too.
We shall see the coditions on which they coincide.

Almost properties

Definition 21 (almost decidability, Hoyrup and Rojas 2009)**.** *A set A is said to be almost decidable if there are two θ-computable open sets U and V such that:*

$$U \subset A,\ V \subseteq A^c,\ U \cup V \text{ is dense and has measure one.}$$

In a computable metric space with a computable measure, there exists an equivalent basis that is uniformly almost decidable.

Definition 22 (almost disjointness)**.** *Suppose that the base β is complete. $(\mathbf{X}, \mu)$ has the property of almost disjointness if there exists a computable function $d :\subseteq \Sigma^* \times \mathbb{Q} \to \Sigma^*$ such that $\mu(\nu(u)) > 1 - q_1 \geq 1 - q_2$ implies $\nu(u) \cup \nu(d(u, q_i)) = X$, $\nu(d(u, q_1)) \subseteq \nu(d(u, q_2))$ and $\mu(\nu(d(u, q_i))) \leq q_i$ for $i = 1, 2$.*

This property also holds in a computable metric space with a computable measure.

Measure randomness and complexity randomness

Suppose μ is computable.This is an essential hypothesis.

Theorem 23. *With almost decidability, complexity μ-randomness implies measure μ-randomness.*

Each measure test is as a computably countable union of almost decidable base elements.
Then i-th element is covered by a complement of a base element.
By getting rid of the union of base elements until $i - 1$, we get a sequence of closed sets containing all non-random points.
The sum of measures of such sets is at most 1.

Theorem 24. *With almost disjointness, measure μ-randomness implies complexity μ-randomness.*

Each closed set can be covered by a base element with as little loss as you want.

Some natural properties

Theorem 25. *The set of complexity random points has measure 1.*

It is obvious for measure randomness, but it needs another proof for complexity randomness.
Recall that computable permutations preserves randomness.

Theorem 26. *Let $f : X_1 \to X_2$ s.t. $\mu_1(f^{-1}(V)) \leq C\mu_2(V)$ for all open $V \subseteq X_2$.*
If $x \in dom(f)$ is complexity μ_1-random then $f(x)$ is complexity μ_2-random.

Relative randomness

Definition

For $i = 1, 2$ let $\mathbf{X}_i = (X_i, \tau_i, \beta_i, \nu_i)$ be computable topological spaces with complete bases and μ_i be a measure on X_i.
Let $x_1 \in X_1$.

Definition 27. *A x_1-measure μ-test over $\mathbf{X_2}$ is a sequence $\{t_n\}$ of uniformly (δ, θ)-computable functions with $\mu(t_n(x_1)) \leq 2^{-n}$. $y \in X_2$ is x_1-measure μ-random if $x \notin \bigcap_n t_n(x_1)$ for each measure test.*

Definition 28. *A x is x_1-complexity μ-random if*

$$x \in \xi(u) \Rightarrow K_{f(x_1)}(u) \geq -\log(\mu(\xi(u))) - O(1)$$

for all (δ, η^{**})-computable functions $f :\subseteq X \to F^{**}$ such that $dom(f(x_1)) \subseteq 2^*$ and $f(x_1)$ is prefix-free.

Coincide

Note that if x_1 is computable, then x_1-randomness coincides with non-relativized randomness.
Almost all properties can be relativized.

Theorem 29. *With almost disjointness and almost decidability, measure randomness and complexity randomness coincide.*

Remark 30. *Universality of complexity randomness many not hold because of multi-functions.*

Van Lambalgen's Theorem

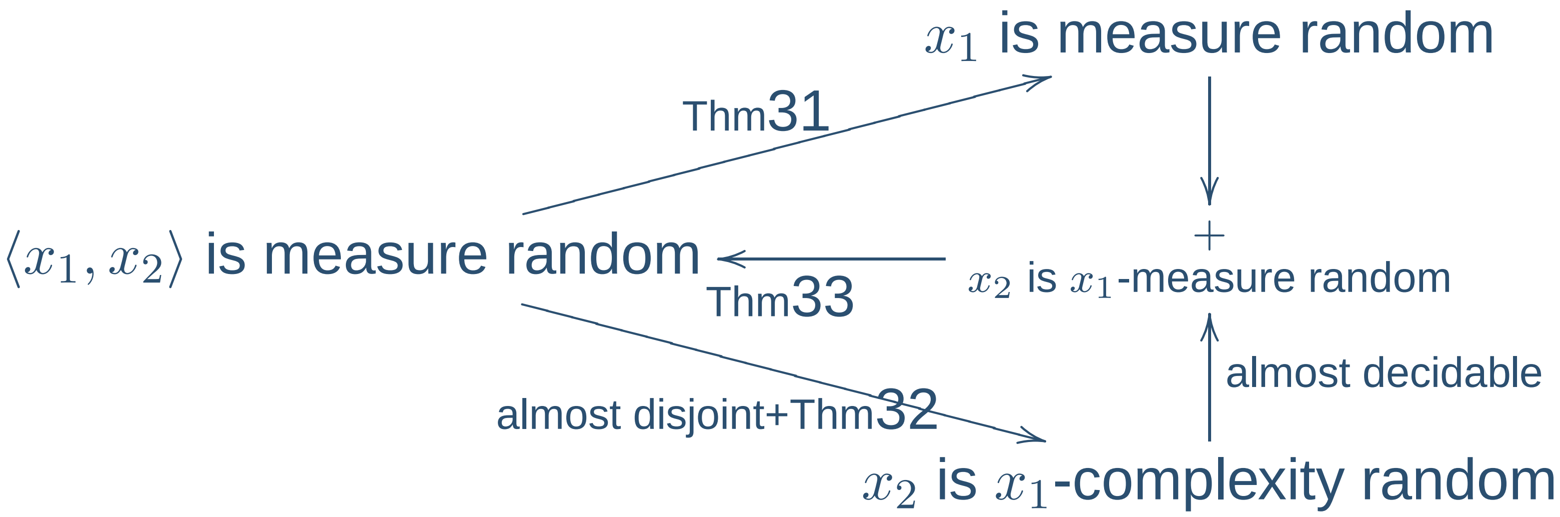
Theorem 31. *If $\langle x_1, x_2 \rangle \in \overline{X}$ is measure $\overline{\mu}$-random, then x_1 is measure μ_1-random.*

Theorem 32. *If X_2 with μ_2 has almost disjointness and $\langle x_1, x_2 \rangle \in \overline{X}$ is measure $\overline{\mu}$-random, then x_2 is x_1-complexity μ_2-random.*

Theorem 33. *If x_1 is measure μ_1-random and x_2 is x_1-measure μ_2-random, then $\langle x_1, x_2 \rangle \in \overline{X}$ is measure $\overline{\mu}$-random.*

Theorem 34. *With almost disjointness and almost decidability, van Lambalgen's Theorem holds.*

Situation



For some reason, it is a difficult problem to find whether van Lambalgen's Theorem holds with no conditions.
Recall that van Lambalgen's theorem is a criterion of natural randomness.
So the pair of the conditions is a sufficient condition for natural randomness.

Summary

Discussion

- We gave the characterizations of measure randomness by martingales.
- We defined complexity randomness and proved it has some natural properties.
- We proposed two conditions which hold on a computable metric space with a computable measure.
- With the conditions, measure randomness and complexity randomness coincide.
- With the conditions, van Lambalgen's Theorem holds.
- The pair of the conditions is a sufficient condition of the space where natural randomness can be defined.

End

Thank you!