Algorithmically random infinite structures

Bakh Khoussainov

Computer Science Department, The University of Auckland, New Zealand

February 13, 2015

Partly based on and motivated by the author's LICS 2014 paper

Plan

- Motivation
- Branching classes
- Topology, metric, and measure.
- Martin-Löf randomness.
- Computable structures and ML-randomness.

A bit of motivation

- The modern history is fascinating; starts with the works of Kolmogorov, Martin-Löf, Chaitin, Schnorr and Levin.
- The last two decades have witnessed to significant advances in the area of algorithmic randomness on strings.
- Monographs by Downey and Hirschfeldt, and Nies.
- Many notions of randomness, various techniques, and ideas have been studied.

Algorithmic randomness of infinite strings

Consider the Cantor space $\{0,1\}^\omega$ and let $\alpha \in \{0,1\}^\omega$. There are several ways to define algorithmic randomness for α. The most traditional ones are based on the following ideas:

1. Incompressibilty.
2. Effective null cover.
3. Computable martingales.

Semi-informal explanation

1. The string $\alpha \in 2^\omega$ is algorithmically random if all sufficiently large segments of α are incompressible (up to a constant).
2. The string α is algorithmically random if no effective measure 0 set contains α.

A set $V \subseteq 2^\omega$ has *effective measure* 0 if V is the limit of embedded sets $M_0 \supset M_1 \supset M_2 \supset \ldots$ such that

- Each M_i is an open set,
- Given i we can compute the base open sets that form M_i,
- The measure of M_i is bounded by $1/2^i$.

Effective measure 0 sets are called *Martin Löf tests* (ML-tests).

From strings to algebraic structures

We want to investigate the following question:

What is an algorithmically random infinite tree, graph, monoid, or generally, a universal algebra?

In the case of strings, the measure on the Cantor space is fundamental in introducing algorithmic randomness.

Thus, the task is to invent a meaningful measure in the classes of structures.

What do we expect from an infinite algorithmically random structure?

- **Absoluteness**: Algorithmic randomness should be an isomorphism invariant property.
- **Continuum**: Random structures should be in abundance, the continuum. This is a property of a collective, the idea that goes back to Von Mises.
- **Selection**: There should be no effective way to describe the isomorphism type of an infinite part of the structure.
- **Axiomatization**: No set of simple (e.g. universal) axioms define the structure.

Questions:

- **Converting into strings**: Why don't we code structures as strings and transform algorithmic randomness for strings into structures?
- **Computability**: Can a computable structure be ML-random?
- **Immunity**: ML-random strings possess *immunity property*: No ML-random string has a computable subsequence. Does ML-random structure have immunity property?
- **Finite presentability**: Can a finitely presented structure, e.g. group, be ML-random?

Converting into strings

Let $G = (\omega; E)$ be a graph. Form the following string α_G:

$$\alpha_G(0)\alpha_G(1)\alpha_G(2)\ldots \in 2^\omega,$$

where $\alpha_G(i) = 1$ iff the i-th pair is an edge in G.

Definition

The graph G is *string-random* if the string α_G is ML-random.

Implications of string-randomness

Theorem

If G is a string-random, then G is Rado graph. Hence,

- *Any two string-random graphs are isomorphic.*
- *The first order theory of the graph is decidable.*
- *The string-random graph is axiomatised by extension axioms.*
- *Any countable infinite graph can be embedded into G.*

All of the above defy our intuition that we postulated for algorithmically random infinite structures.

Direct limits

Definition

An *embedded system* of structures is a sequence

$$(\mathcal{A}_0, f_0), (\mathcal{A}_1, f_1), \ldots, (\mathcal{A}_i, f_i), \ldots$$

such that (1) each $\mathcal{A}_i$ is a finite structure, and (2) each f_i is a *proper into embedding* from $\mathcal{A}_i$ into $\mathcal{A}_{i+1}$.

The sequence $\mathcal{A}_0, \mathcal{A}_1, \ldots$ is *the base* of the system.

Each embedded system determines the limit structure.

Definition

An embedded system $\{(\mathcal{A}_i, f_i)\}_{i \in \omega}$ is *strict* if its direct limit is isomorphic to the direct limit of any embedded system with the same base.

Classes with height function

Let $h : \mathcal{K} \to \omega$ be a computable height function such that:

1. We can compute the cardinality of $h^{-1}(i)$ for every i.
2. For every $\mathcal{A} \in \mathcal{K}$ of height i there is a substructure $\mathcal{A}[i-1]$ of height $i-1$ such that all substructures of $\mathcal{A}$ of height $\leq i-1$ are contained in $\mathcal{A}[i-1]$.
3. For all $\mathcal{A} \in \mathcal{K}$ of height i and $C \subseteq A \setminus A[i-1]$, the height of the substructure $C \cup A[i-1]$, where $C \neq \emptyset$, is i in case the substructure belongs to $\mathcal{K}$.

Properties of classes with height function

Lemma

For all $\mathcal{A}, \mathcal{B} \in K$, the structures $\mathcal{A}$ and $\mathcal{B}$ are isomorphic iff $h(\mathcal{A}) = h(\mathcal{B})$ and $\mathcal{A}[j] = \mathcal{B}[j]$ for all $j \leq h(\mathcal{A})$. □

Lemma

Every embedded system of structures from the class $\mathcal{K}$ is strict. □

Definition

The class K is a *branching class*, or *B-class for short*, if for all $\mathcal{A} \in K$ of height i there exist distinct structures $\mathcal{B}, \mathcal{C} \in K$ such that $h(\mathcal{B}) = h(\mathcal{C}) > h(\mathcal{A})$ and $\mathcal{B}[i] = \mathcal{C}[i] = \mathcal{A}$.

Examples of B-classes

Example 1. Pointed connected graphs (G, p) of bounded degree. The height function is the max distance from p to vertices of G.

Example 2. Trees of bounded degree. The height function is the height of the tree.

Example 3. Relational structures whose Gaifman graph is a connected graph of a bounded degree.

Example 4. Partially ordered sets $(P; \leq, C, p)$, where p is the least element, $C(x, y)$ is the cover relation, and each x in P has at most d covers.

Examples of B-classes

Example 5. Two generated, by generators a and b, partial monoids. The height function is the max of the shortest words representing monoid elements.

Example 6. The class of binary rooted ordered trees.

Example 7. The class of n-generated universal partial algebras. The hight function is the max among the heights of the shortest terms representing the elements of the algebras.

Example 8. The class of (a, b)-sparse graphs. A connected pointed graph is (a, b)-sparce if every subgraph of G with m vertices has at most $am + b$ edges.

Definition of tree $T(K)$

Let K be a B-class. Define $T(K)$ as follows:

1. The root is $\emptyset$. This is level -1.
2. The nodes of at level $n \geq 0$ are structures of height n.
3. Let $\mathcal{B}$ be a structure of height n. Its successor is any structure $\mathcal{C}$ of height $n+1$ such that $\mathcal{B} = \mathcal{C}[n]$.

Properties of the tree $\mathcal{T}(K)$:

Let K be a B-class. Set

$$K_\omega = \{\mathcal{A} \mid \mathcal{A} \text{ is the direct limit of structures from } K\}.$$

1. Given any node x of the three $T(K)$, we can effectively compute the structure $\mathcal{B}_x$ associated with the node x.
2. Each x in $\mathcal{T}(K)$ has an immediate successor. We can compute the number of immediate successors of x.
3. Each path $\eta = \mathcal{B}_0, \mathcal{B}_1, \ldots$ determines the limit structure $\mathcal{B}_\eta = \cup_i \mathcal{B}_i \in K_\omega$.
4. The mapping $\eta \to \mathcal{B}_\eta$ is a bijection from $[\mathcal{T}(K)]$ to K_ω.

Topology on K_ω

Definition (Topology)

Let $\mathcal{B}$ be a structure of height n. The **cone** of $\mathcal{B}$ is:

$$Cone(\mathcal{B}) = \{\mathcal{A} \mid \mathcal{A} \in K_\omega, \text{ and } \mathcal{A}[n] = \mathcal{B} \text{ for all } n\}.$$

Declare the cones $Cone(\mathcal{B})$ to be the *base open sets* of the topology on K_ω. We refer to $\mathcal{B}$ as the *base of the cone*.

Measure

Definition (Measure)

- The measure of the cone based at the root is 1.
- Assume that the measure $\mu(Cone(\mathcal{B}_x))$ has been defined. Let e_x be the number of immediate successors of x. Then for any immediate successor y of x the measure of $Cone(\mathcal{B}_y)$ is

$$\mu(Cone(\mathcal{B}_y)) = \frac{\mu(Cone(\mathcal{B}_x))}{e_x}.$$

Metric

Definition (Metric)

For $\mathcal{A}, \mathcal{B} \in K_\omega$, let n be the maximal level at which $\mathcal{A}[n] = \mathcal{B}[n]$. The distance $d(\mathcal{A}, \mathcal{B})$ is then: $d(\mathcal{A}, \mathcal{B}) = \mu_m(Cone(\mathcal{A}[n]))$.

Lemma

The function d is a metric in the space K_ω .

Basic properties of the space K_ω

Fact

1. *K_ω is compact.*
2. *The set K is countable and dense in K_ω.*
3. *Finite unions of cones form clo-open sets in the topology.*
4. *The set of all μ-measurable sets is a σ-algebra.* □

ML-random structures

Definition

A structure $\mathcal{A} \in K_\omega$ is *ML-random* if it passes every ML-test.

Corollary

Let $\mathcal{K}$ be a B-class. The number of ML-random structures in the class $\mathcal{K}_\omega$ is continuum. In particular, for all the examples of B-classes $\mathcal{K}$ we considered, each of the classes $\mathcal{K}_\omega$ contains continuum ML-random structures.

Computable structures

Definition

An infinite structure $\mathcal{A}$ is *computable* if it is isomorphic to a structure with domain ω such that all atomic operations and relations of the structure are computable.

Definition

A computable structure $\mathcal{A}$ from $\mathcal{K}_\omega$ is *strictly computable* if the size of the substructure $\mathcal{A}[i]$ can be computed for all $i \in \omega$.

The following are true:

1. Every computable finitely generated algebra is strictly computable.
2. A computable pointed graph G of bounded degree is strictly computable iff there is an algorithm that given a vertex v of G computes $degree(v)$.
3. A computable rooted tree T of bounded degree is strictly computable iff there is an algorithm that given a node $v \in T$ computes the number of immediate successors of v.
4. A computable d-bounded partial order with the least element is strictly computable iff there is an algorithm that for every v of the partial order computes all covers of v.

Strict computability and randomness

Theorem

If $\mathcal{A}$ is strictly computable then $\mathcal{A}$ is not ML-random.

Corollary

Let $\mathcal{A}$ be either an infinite pointed graph or tree or partial order of bounded degree. If $\mathcal{A}$ is computable and its $\exists$-diagram, that is the set

$$\{\phi(\bar{a}) \mid \bar{a} \in A \text{ and } \mathcal{A} \models \phi(\bar{a}) \text{ and } \phi(\bar{x}) \text{ is an existential first-order formula}\},$$

is decidable then $\mathcal{A}$ is not ML-random.

Computability in the Halting set

Theorem

For every B-class $\mathcal{K}$ there exist an $\mathcal{H}$-computable ML-random structure in class $\mathcal{K}_\omega$. □

Thus, we have the following corollary:

Corollary

For all the examples of B-classes $\mathcal{K}$ that we have considered each of the classes $\mathcal{K}_\omega$ contains ML-random $\mathcal{H}$-computable structures. □

The mapping from $[T(K)]$ to K_ω

Construction of $\mathcal{A}_\eta$ from η is computable in η. Hence, if η is computable then so is $\mathcal{A}_\eta$.

How about the opposite:

How complex is that to compute η from $\mathcal{A}_\eta$?

Answer:

To compute η, we need to compute $\mathcal{A}_\eta[i]$ for each i. Computing $\mathcal{A}_\eta[i]$ requires the jump of the open diagram of $\mathcal{A}_\eta$.

ML-randomness and computability

Theorem

There exists a B-class K such that K_ω contains an ML-random yet a computable structure.

Proof (idea). A binary ordered tree $\mathcal{B}$ belongs to $\mathcal{S}$ if:

1. All leaves of $\mathcal{B}$ are of the same height,
2. If v in $\mathcal{B}$ has the right child then all nodes left of v on the v's level-order including v have both children,
3. At each level i there is at most one node such that it is the left child of its parent that does not have a right child.

The structure of $T(\mathcal{S})$

Lemma

If $\mathcal{B}$ belongs to $\mathcal{S}$ and has height n then there are exactly two non-isomorphic extensions of $\mathcal{B}$ of height $n+1$ both in $\mathcal{S}$. Hence, the tree $T(\mathcal{S})$ is isomorphic to the infinite binary tree.

Lemma

For every $n \geq 0$, the set of all trees in $\mathcal{S}$ of height n form a chain of embedded structures.

Embedding Lemma

So, we naturally identify the infinite binary tree $\{0,1\}^\star$ with $T(\mathcal{S})$ through the lemmas above.

Lemma

Let $x \preceq y$, where $\preceq$ is the lexicographical order on binary strings. Then:

1. *If $|x| \leq |y|$ then $\mathcal{A}_x$ is embedded into $\mathcal{A}_y$.*
2. *If $|x| > |y|$ then $\mathcal{A}_x$ is embedded into $\mathcal{A}_{yz}$ for all z such that $|x| \leq |yz|$*

Computable structure lemma

Lemma

The B-class $\mathcal{S}$ possesses a computable ML-random structure.

Proof (idea). Consider the tree $T(\mathcal{S})$. Let η be the ML-random path obtained as the leftmost path that avoids the universal ML-test. Because of the lemmas above, the structure $\mathcal{A}_\eta$ can be constructed computably. □

Nontrivial varieties

Definition

A class of universal algebras is a *variety* if its closed under sub-algebras, homomorphisms, and ultra-products.

A class of algebras is variety if and only if is axiomatised by a set E of universally quantified equitations.

An equation is $p(\bar{x}) = q(\bar{x})$ where p and q are terms.
The equation $p(\bar{x}) = q(\bar{x})$ *non-trivial* if at least one of the terms contains a variable and $p \neq q$ syntactically.

If E contains at least one non-trivial equation then we call the variety of algebras satisfying E a *non-trivial variety*.

The measure of nontrivial varieties

Theorem

The class of all infinite n-generated algebras that belong to a non-trivial variety has an effective measure zero. Hence, no finitely presented algebra of a non-trivial variety is ML-random.

Corollary

No finitely generated ML-random algebra exists that satisfies a nontrivial set of equations. Hence, no ML-random group, monoid, or lattice exist. □

Corollary

A finitely axiomatised variety V has either an effective measure 0 or its measure is a rational number > 0. The latter case occurs iff the variety is axiomatised by a trivial set of equations. Moreover, $\mu(V) > 0$ iff V is a finite union of cones.

Open questions

1. Are there ML-random c.e. or co-c.e. universal algebras?
2. Build ML-random finitely generated groups and rings.
3. Is ML-randomness quasi-isometry invariant in the class of graphs?