Algorithmically random infinite structures

Bakh Khoussainov

Computer Science Department, The University of Auckland, New Zealand

September 4, 2017

I would like to thank Rod for his friendship

Plan

- Branching classes
- Martin-Löf randomness
- Computable structures and ML-randomness
- Algorithmically random c.e. and co-c.e. structures
- Degrees of ML-random structures
- Measures of varieties

Literature

1. B. Khoussainov. A quest for algorithmically random infinite structures. Proceedings of LICS-CSL conference. 2014.
2. B. Khoussainov. A quest for algorithmically random infinite structures, II. Proceedings of LFCS conference. 2016.
3. B. Khoussainov. Quantifire-free definability on infinite algebras. LICS proceedings. 2016.
4. B. Khoussainov and D. Turesky. Computability theoretic properties of algorithmically random structures. In preparation.

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 significant advances in the area of algorithmic randomness on strings.
- Many notions of randomness, various techniques, and ideas have been studied.

Randomness via effective null cover sets

Definition

String $\alpha \in \{0,1\}^\omega$ is algorithmically random if no effective measure 0 set contains α.

A set $V \subseteq 2^\omega$ has *effective measure* 0 if V is contained in 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 base open sets that form M_i,
- The measure of M_i is bounded by $1/2^i$.

The sequence $\{M_i\}_{i\in\omega}$ is called *Martin Löf (ML) test*.

So, the measure on the Cantor space plays the key role in introducing algorithmic randomness.

Motivating question

The question is the following:

What is an algorithmically random infinite algebraic structure?

To answer the question, we need to invent a meaningful measure in the classes of structures.

Expectations from algorithmically random structures

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

Questions:

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

Randomness by Erdos

On ω, for each pair $\{i, j\}$ put an edge between i and j at random. This determines an infinite graph.

Definition

Call the resulting graph random.

Theorem (Erdos and Spencer)

With probability 1 *any two random graphs are isomorphic.*

This theorem, as Erdos and Spencer write , "demolishes the theory of infinite random graphs".

Implications of string-randomness

Theorem

If G is a string-random, then G is isomorphic to the random 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 $\mathcal{K}$ be a class of finite structures. A computable function $h : \mathcal{K} \to \omega$ is a **height function** if each of the following is true:

1. We can compute the cardinality of $h^{-1}(i)$ for every i.
2. Each $\mathcal{A} \in \mathcal{K}$ of height i has 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 $\mathcal{K}$.

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. Trees of bounded degree $d > 1$. The height function is the height of the tree.

Example 2. Pointed connected graphs $(G, \bar{p})$ of bounded degree d. The height function is the max distance from $\bar{p}$ to vertices of G.

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

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. The class of δ-hyperbolic connected pointed graphs of bounded degree *d*.

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 $\mathcal{T}(K)$

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

1. The root is $\emptyset$. This is level -1.
2. The nodes of $T(K)$ 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 $\mathcal{T}(K)$:

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.

The class K_ω

Let K be a B-class. Set

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

Call this class K_ω a B-class.

Correspondence between K_ω and $[T(K)]$:

1. Each path $\eta = \mathcal{B}_0, \mathcal{B}_1, \ldots$ determines the limit structure $\mathcal{B}_\eta = \cup_i \mathcal{B}_i$ from the class K_ω.
2. The mapping $\eta \to \mathcal{B}_\eta$ is a bijection from $[\mathcal{T}(K)]$ to 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 and the continuum

Definition

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

Corollary (Randomness is a property of a collective)

The number of ML-random structures in K_ω is continuum.

Corollary

For all the examples of B-classes K we considered, the classes K_ω contains continuum ML-random structures.

ML-randomness is isomorphism invariant

All the definitions depend on constants $\bar{c}$ that we fixed at the start. In particular, the trees $T(\mathcal{K})$ and hence *ML*-randomness depend on the constants.

Theorem (Absoluteness)

For all the examples of B-classes, ML-randomness is independent on the choice of constants.

Summary of our set-up

ML randomness for structures, as we defined, depends on:

1. The class K (the context).
2. The height function h (approximation).
3. The measure μ or its refinements (measures).

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 $i \in \omega$.

Strict computability

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 v from G computes the number of vertices adjacent to 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

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

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

Computability in the Halting set

Theorem

Every B-class $\mathcal{K}_\omega$ contains ML-random structures computable in the halting set. □

Thus, we have the following corollary:

Corollary

All examples of B-classes $\mathcal{K}_\omega$ that we have considered contain ML-random structures computable in the halting set. □

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 (Computable structure theorem)

There exists a B-class $\mathcal{S}$ such that $\mathcal{S}_\omega$ 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.

We identify the tree $\{0,1\}^\star$ with $T(\mathcal{S})$ by the lemmas.

Further properties of $T(\mathcal{S})$

Lemma (Algebraic left-embedding 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 $|x| \leq |yz|$*

Consider Ω and take its left-c.e. limit $x_0 \preceq x_1 \preceq x_2 \preceq \ldots$. Because of the lemmas above, we have a computable sequence

$$A_{x_0} \subset A_{x_1} \subset A_{x_2} \subset \ldots.$$

The limit of this sequence is $\mathcal{A}_\Omega$. Hence, $\mathcal{A}_\omega$ is computable. □

Computably enumerable universal algebras

Let $\mathcal{A}$ be a universal finitely generated computable algebra. Then no *B*-class K_ω exists in which $\mathcal{A}$ is ML-random.

Definition

Let $\mathcal{A}$ be a finitely generated universal algebra.

1. Call $\mathcal{A}$ *computably enumerable* if the word problem for $\mathcal{A}$ is a computably enumerable set.
2. Call $\mathcal{A}$ *co-computably enumerable* if the word problem for $\mathcal{A}$ is a co-computably enumerable set.

Elements of finitely presented algebras are presented by terms. Hence, computability of operations is granted vacuously.

Questions

In the case of strings there are ML-random c.e. reals, e.g. the Ω number. Hence, natural questions arise:

1. Do there there exist computably enumerable algorithmically random universal algebras?
2. Do there there exist co-c.e. algorithmically random universal algebras?

ML-random co-c.e. algebras

Theorem (with D. Turetsky)

There exists a branching class $\mathcal{S}_\omega$ that contains co-computably enumerable ML-random universal algebra.

Proof (idea). Construct a class $\mathcal{S}$ of partial universal algebras such that the following properties hold:

1. For each n there are exactly 2^n algebras of height n from $\mathcal{S}$. So, the tree $T(\mathcal{S})$ is just the full binary tree $\{0,1\}^\star$.
2. Let $x \preceq y$ in $\{0,1\}^\star$.
   1. If $|x| = |y|$ then $\mathcal{A}_x$ is a homomorphic image of $\mathcal{A}_y$.
   2. If $|x| > |y|$ then $\mathcal{A}_x$ is a homomorphic image of $\mathcal{A}_{yz}$ for all z such that $|x| = |yz|$

Proof: continuation

Take Ω and its left-c.e. approximation $x_0 \preceq x_1 \preceq x_2 \preceq \ldots$. This corresponds to the sequence of partial algebras:

$$A_{x_0}, A_{x_1}, A_{x_2}, \ldots$$

Each A_{x_i} is a homomorphic image of $A_{x_{i+1}}$. Some equal elements in A_{x_i} are split to become non-equal elements in $A_{x_{i+1}}$. Non-equality is preserved.

The natural direct sub-sum of these algebras will be a total algebra in which equality is co-c.e. The direct sub-sum will be isomorphic to $\mathcal{A}_\Omega$.

ML-random c.e. algebras

Theorem (with D. Turetsky)

There exists a branching class $\mathcal{S}_\omega$ that contains computably enumerable ML-random universal algebra.

Proof (idea). Consider $1 - \Omega$. This is right c.e. real. Consider the right-c.e. approximation $\ldots \preceq x_2 \preceq x_1 \preceq x_1 \preceq x_0$. This corresponds to the sequence of partial algebras:

$$A_{x_0}, A_{x_1}, A_{x_2}, \ldots$$

Each $A_{x_{i+1}}$ is a homomorphic image of A_{x_i}. Once two elements in A_{x_i} are equal, they stay equal. The limit of this sequence converges to $\mathcal{A}_{1-\Omega}$ which is c.e. and ML-random.

Degrees of ML-random structures

Definition

A B-class K **jumpless** if for every path η through $T(K)$, every isomorphic copy of $\mathcal{A}_\eta$ computes η.

Theorem (with D. Turetsky)

If K is jumpless, then every structure in K_ω has degree, and the degrees of ML-random structures are precisely the Turing degrees which contain random binary strings.

Left-algebraic B-classes

Definition

A B-class K is **left-algebraic** if there is a computable ordering on the elements of each level of $T(K)$ such that for the induced lexicographic ordering $\leq$ we have:

1. For all $\eta \in [T(K)]$ and all $\eta_0 \leq \eta_1 \leq \eta_2 \leq \ldots$ with limit η, the sequence computes an isomorphic copy of $\mathcal{A}_\eta$.
2. For all $\eta \in [T(K)]$ and all isomorphic copies of $\mathcal{A}_\eta$, the copy computes a sequence $\eta_0 \leq \eta_1 \leq \eta_2 \leq \ldots$ with limit η.

Dichotomy theorem

Theorem (with D. Turetsky)

Let $\mathcal{A}$ be an ML-random structure in a left-algebraic branching class K_ω such that $\mathcal{A}$ has a degree. Then the degree of $\mathcal{A}$ is either $\mathbf{0}$ or $\mathbf{0}'$. Both degrees are realisable.

Proof (Idea). Consider Ω and $1 - \Omega$. The structure $\mathcal{A}_\Omega$ is computable and the structure $\mathcal{A}_{1-\Omega}$ computes the halting set. The rest requires forcing type of arguments.

Nontrivial varieties

Definition

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

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

An equation is $p(\bar{x}) = q(\bar{x})$ where p and q are terms.
The equation $p(\bar{x}) = q(\bar{x})$ is *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.

Open questions

1. Assume that a B-class K is neither strict nor left-algebraic. What degrees can be realised by ML-random structures?
2. Is the first order theory of ML-random graph with bounded degree decidable?
3. Are two ML-random graphs of the same bounded degree elementary equivalent?
4. Construct B-classes of finitely generated groups.
5. Are there computable ML-random graphs in the class of all connected graphs of bounded degree?
6. Is the class of the subgroups of $(Q; +)$ a branching class?