Computation and Logic on Dynamic Random Graphs

Wesley Calvert

Southern Illinois University

ASL North American Annual Meeting
Berkeley, California
March 26, 2011

Theorem (0-1 Law)

Every sentence in the language of graphs is true for either almost all finite graphs or almost none of them.

Theorem (0-1 Law)

Every sentence in the language of graphs is true for either almost all finite graphs or almost none of them.

Definition

The theory of the random graph is the set of all sentences in the language of graphs which are true for almost all finite graphs.

Theorem

The theory of the random graph is decidable.

Theorem

The theory of the random graph is decidable.

Theorem

The theory of the random graph is $\aleph_0$-categorical.

Theorem

The theory of the random graph is decidable.

Theorem

The theory of the random graph is $\aleph_0$-categorical.

Proof.

Back-and-forth □

Theorem

The theory of the random graph is decidable.

Theorem

The theory of the random graph is $\aleph_0$-categorical.

Proof.

Back-and-forth □

Theorem

The theory of the random graph is computably categorical.

Theorem

The theory of the random graph is decidable.

Theorem

The theory of the random graph is $\aleph_0$-categorical.

Proof.

Back-and-forth □

Theorem

The theory of the random graph is computably categorical.

Theorem

The theory of the random graph is properly simple.

When I showed this to some colleagues in my department, they didn't like it much.

When I showed this to some colleagues in my department, they didn't like it much.

- "There's nothing random in it!"

When I showed this to some colleagues in my department, they didn't like it much.

- "There's nothing random in it!"
- "A model of that theory isn't a random graph!"

When I showed this to some colleagues in my department, they didn't like it much.

- "There's nothing random in it!"
- "A model of that theory isn't a random graph!"
- "You mean your random graphs can be infinite?"

When I showed this to some colleagues in my department, they didn't like it much.

- "There's nothing random in it!"
- "A model of that theory isn't a random graph!"
- "You mean your random graphs can be infinite?"
- "Maybe a better name would be 'random theory of graphs.' "

Here's what they're used to as a random graph:

Definition

The Erdős-Renyi random graph $G_p(n)$ is constructed by taking n vertices and connecting each pair independently with probability p.

Here's what they're used to as a random graph:

Definition

The Erdős-Renyi random graph $G_p(n)$ is constructed by taking n vertices and connecting each pair independently with probability p.

Proposition

The theory of the random graph is the almost-sure theory of $G_p(\omega)$ if $p \in (0, 1)$.

Definition

Continuous first-order logic is a logic taking truth values on $[0,1]$, and having all continuous functions on $[0,1]$ for its Boolean connectives and sup and inf for its quantifiers. We typically also include a metric d in the signature.

Definition

Continuous first-order logic is a logic taking truth values on $[0,1]$, and having all continuous functions on $[0,1]$ for its Boolean connectives and sup and inf for its quantifiers. We typically also include a metric d in the signature.

Proposition (Ben Yaacov-Berenstein-Henson-Usvyatsov)

Any continuous function $[0,1]^n \to [0,1]$ *can be approximated by* $(\dot{-}, \neg = x \mapsto 1 - x, \frac{1}{2})$.

Definition

A continuous structure $\mathcal{M}$ is a metric space (M, d)

Definition

A continuous structure $\mathcal{M}$ is a metric space (M, d) with

- Some distinguished uniformly continuous functions $f : M^n \to M$

Definition

A continuous structure $\mathcal{M}$ is a metric space (M, d) with

- Some distinguished uniformly continuous functions $f : M^n \to M$, and
- Some distinguished uniformly continuous predicates $R : M^n \to [0, 1]$.

Example

Consider the set $\{0, 1, \ldots, n\}$, with the discrete metric.

Example

Consider the set $\{0, 1, \ldots, n\}$, with the discrete metric. Define a predicate R such that

- $R(x, y) = 1 - p$ if $x \neq y$

Example

Consider the set $\{0, 1, \ldots, n\}$, with the discrete metric. Define a predicate R such that

- $R(x, y) = 1 - p$ if $x \neq y$
- $R(x, x) = 1$.

Example

Consider the set $\{0, 1, \ldots, n\}$, with the discrete metric. Define a predicate R such that

- $R(x, y) = 1 - p$ if $x \neq y$
- $R(x, x) = 1$.

Note that this is a continuous structure, and "is" the Erdős-Renyi random graph $G_p(n)$.

Definition

Let 2^ω be the set of infinite binary sequences, with the usual Lebesgue probability measure μ.

Definition

Let 2^ω be the set of infinite binary sequences, with the usual Lebesgue probability measure μ.

1. A *randomized Turing machine* is a Turing machine equipped with an oracle for an element of 2^ω, called the *random bits*, with output in $\{0, 1\}$.

Definition

Let 2^ω be the set of infinite binary sequences, with the usual Lebesgue probability measure μ.

1. A *randomized Turing machine* is a Turing machine equipped with an oracle for an element of 2^ω, called the *random bits*, with output in $\{0, 1\}$.
2. We say that a randomized Turing machine M *accepts* n *with probability* p if and only if $\mu\{x \in 2^\omega : M^x(n) \downarrow = 0\} = p$.

Definition

Let 2^ω be the set of infinite binary sequences, with the usual Lebesgue probability measure μ.

1. A *randomized Turing machine* is a Turing machine equipped with an oracle for an element of 2^ω, called the *random bits*, with output in $\{0, 1\}$.
2. We say that a randomized Turing machine M *accepts* n *with probability* p if and only if $\mu\{x \in 2^\omega : M^x(n) \downarrow = 0\} = p$.
3. We say that a randomized Turing machine M *rejects* n *with probability* p if and only if $\mu\{x \in 2^\omega : M^x(n) \downarrow = 1\} = p$.

Definition

The continuous atomic diagram $D(\mathcal{M})$ of a continuous structure $\mathcal{M}$ is the set of pairs (φ, p), where φ is an atomic CFO formula (in $\mathcal{M}$ with unary distance) and the truth value of φ in $\mathcal{M}$ is equal to p.

Definition

The continuous atomic diagram $D(\mathcal{M})$ of a continuous structure $\mathcal{M}$ is the set of pairs (φ, p), where φ is an atomic CFO formula (in $\mathcal{M}$ with unary distance) and the truth value of φ in $\mathcal{M}$ is equal to p.

Definition

A probabilistically computable structure $\mathcal{M}$ is a continuous structure equipped with a randomized Turing machine which, for any pair $(\varphi, p) \in D(\mathcal{M})$, accepts φ with probability equal to p.

Proposition

Every classically computable structure, with the discrete metric, is a probabilistically computable structure.

Proposition

Every classically computable structure, with the discrete metric, is a probabilistically computable structure.

Theorem

Classically computable structures are to RCA_0 *as probabilistically computable structures are to* ACA_0*.*

Note that the Erdős-Renyi random graph we described earlier is probabilistically computable.

But, the Erdős-Renyi graphs are too homogeneous.

But, the Erdős-Renyi graphs are too homogeneous.

Example

The FaceBook friend graph.

But, the Erdős-Renyi graphs are too homogeneous.

Example

The FaceBook friend graph.

Example

The web graph.

But, the Erdős-Renyi graphs are too homogeneous.

Example

The FaceBook friend graph.

Example

The web graph.

Example

The collaboration graph.

But, the Erdős-Renyi graphs are too homogeneous.

Example

The FaceBook friend graph.

Example

The web graph.

Example

The collaboration graph.

Example

Consider the reactions catalyzed by *E. Coli*. Make a node for each collection of reactants and products, and a connection for any two that share a metabolite (intermediate compound).

Major features we want to capture:

1. Development over time

Major features we want to capture:

1. Development over time
2. Non-homogeneity [number of vertices of degree k proportional to $\frac{1}{k^\beta}$ for fixed β]

The language:

- Infinitely many predicates $(V_i)_{i \in \omega}$, which will be interpreted as the vertex set at stage i.

The language:

- Infinitely many predicates $(V_i)_{i\in\omega}$, which will be interpreted as the vertex set at stage i.
- Infinitely many predicates $(E_i)_{i\in\omega}$, interpreted as the edge set at stage i.

Definition

Let G_0 be a finite graph with vertex set V_* and edge set E_*. The preferential attachment graph $G(p, G_0)$ is the random graph process with $V_0 = V_*$, with $E_0 = E_*$, and with the following property:

Definition

Let G_0 be a finite graph with vertex set V_* and edge set E_*. The preferential attachment graph $G(p, G_0)$ is the random graph process with $V_0 = V_*$, with $E_0 = E_*$, and with the following property:

- For each $t > 0$, we form G_t by independently
  - With probability p let V_t consist of V_{t-1} plus a single new element v, and independently choose a vertex u in proportion to its degree in G_{t-1}, setting $E_t = E_{t+1} \cup \{(u, v)\}$.

Definition

Let G_0 be a finite graph with vertex set V_* and edge set E_*. The preferential attachment graph $G(p, G_0)$ is the random graph process with $V_0 = V_*$, with $E_0 = E_*$, and with the following property:

- For each $t > 0$, we form G_t by independently
  - With probability p let V_t consist of V_{t-1} plus a single new element v, and independently choose a vertex u in proportion to its degree in G_{t-1}, setting $E_t = E_{t+1} \cup \{(u, v)\}$.
  - Otherwise, independently choose two vertices u, v with probability proportional to their degrees in G_{t-1}, and set $V_t = V_{t-1}$ and $E_t = E_{t-1} \cup \{(u, v)\}$.

Proposition

There is an effective procedure which, given $\sigma \in 2^{\omega}$, will produce an index for the sample path corresponding to σ; that is, for the uniformly computable sequence of (classical) relations $(V_i^{\sigma}, E_i^{\sigma})_{i \in \omega}$.

Lemma

For any $p \in [0,1]$, there exists an infinite uniformly computable sequence of subsets $(S_i^p)_{i \in \omega}$ of 2^ω which are independent, such that S_i^p has probability p.

Lemma

For any $p \in [0,1]$, there exists an infinite uniformly computable sequence of subsets $(S_i^p)_{i\in\omega}$ of 2^ω which are independent, such that S_i^p has probability p.

Lemma

For any positive integer n, there exist infinite independent uniformly computable sequences of tuples $(Q_t^n)_{t\in\omega}$ and $(P_t^n)_{t\in\omega}$, where each Q_t^n and each P_t^n has the form $(Q_{t,k}^n)_{k<n}$ and the $Q_{t,k}^n$ are disjoint subsets of 2^ω, each with measure $\frac{1}{n}$.

Proof of Proposition.

Start with G_0.

Proof of Proposition.

Start with G_0. Check σ against S^p to decide whether to do a vertex step or an edge step.

Proof of Proposition.

Start with G_0. Check σ against S^p to decide whether to do a vertex step or an edge step. If a vertex step, check with the appropriate Q to figure out where the new vertex should connect.

Proof of Proposition.

Start with G_0. Check σ against S^p to decide whether to do a vertex step or an edge step. If a vertex step, check with the appropriate Q to figure out where the new vertex should connect. If an edge step, check with the appropriate P. □

Theorem

For any $p \in [0,1]$ and any finite graph G_0, there is a probabilistically computable random graph process of form $G(p, G_0)$.

Theorem

For any $p \in [0,1]$ and any finite graph G_0, there is a probabilistically computable random graph process of form $G(p, G_0)$.

Proof.

Take the Turing machine we just built, and let it work over all oracles. □

Problem

What is $Th(G(p, G_0))$?

Proposition

The axioms of preferential attachment are neither complete nor categorical.

Proposition

The axioms of preferential attachment are neither complete nor categorical.

Proof.

Do the same construction as before, but arrange that individual numbers have small probability of coming in as vertices at stage t. □

The metabolic pathway graphs tend to look different.

The metabolic pathway graphs tend to look different.
Recall: $n_k \propto \frac{1}{k^\beta}$.

The metabolic pathway graphs tend to look different.
Recall: $n_k \propto \frac{1}{k^\beta}$.

Preferential attachment $\beta \geq 2$ (generally)

Metabolic Networks $\beta \approx 1.5$.

Definition

Let G_0 be a finite graph with vertex set V_* and edge set E_*. The duplication graph $B(p, G_0)$ is the random graph process with $V_0 = V_*$, with $E_0 = E_*$, and with the following property: For each $t > 0$, we form G_t by

- Independently selecting a vertex v_t from V_{t-1} uniformly at random, add a new vertex y_t, and
- For each neighbor u connected to v_t at stage $t - 1$, we independently attach u to y_t with probability p.

Theorem

For any $p \in [0,1]$ and any finite graph G_0, there is a probabilistically computable random graph process of the form $B(p, G_0)$.

Theorem

For any $p \in [0,1]$ and any finite graph G_0, there is a probabilistically computable random graph process of the form $B(p, G_0)$.

And it has all the same issues with completeness.

Computation and Logic on Dynamic Random Graphs

Wesley Calvert

Southern Illinois University

ASL North American Annual Meeting
Berkeley, California
March 26, 2011