WoLLIC 2005

THE STRANGE LOGIC OF RANDOM GRAPHS

Joel Spencer

Courant Institute

What is a Graph?

- Vertex Set V
- Edge Set E of $\{v, w\} \in V$
- Equivalently:

Areflexive Symmetric Relation

Write: $v \sim w$

Read: v, w are adjacent

Note to cognescenti: no loops, multiple edges

Usually: $n := |V|$, number of vertices

First Order Language

Relations $=$ (equality), $\sim$ (adjacency)

Usual Boolean $\wedge, \neg, \vee, \rightarrow, \ldots$

Universal $\forall_v$, $\exists_w$

NOTE: Quantification only over vertices!

There is a triangle:

$\exists_v \exists_w \exists_u (u \sim v) \wedge (v \sim w) \wedge (u \sim w)$

Diameter at most two:

$\forall_v \forall_w [v = w \vee v \sim w \vee \exists_u (v \sim u \wedge u \sim w)]$

Diameter at most k (k *fixed*)

Connectivity: NO!

The Random Graph $G(n,p)$

n vertices

$p =$ adjacency probability

Usually $p = p(n)$

$$p = \frac{3}{n};\ p = n^{-1/2};\ p = \frac{\ln n}{n} - \frac{5}{n}$$

Allow defined for n sufficiently large

Glebskii et. al. [1969]; Fagin [1976]:

Set $p = \frac{1}{2}$. Then for *all* first order sentences A

$$\lim_{n \to \infty} \Pr[G(n,p) \models A] = 0 \text{ or } 1$$

Extension Statement $A_{r,s}$: For all distinct $x_1, \ldots, x_r, y_1, \ldots, y_s$ there exists distinct z adjacent to all x_i and no y_j.

Probability:

$$\Pr[\neg A_{r,s}] \leq \binom{n}{r}\binom{n-r}{s}(1 - 2^{-r-s})^{n-r-s} \to 0$$

Combinatorics: There is a unique countable model satisfying all $A_{r,s}$.

Logic: Therefore $T = \{A_{r,s}\}$ is complete. If $T \models A$ then $\lim_n \Pr[A] = 1$ by compactness. Otherwise $T \models \neg A$ and $\lim_n \Pr[A] = 0$.

Given

- Distribution μ_n over n-point models
- Language

Possible Outcomes

- Zero-One Law: $\Pr[A] \to 0$ or 1 for all A
- Convergence: $\lim \Pr[A]$ exists for all A
- Slow Oscillation: $\Pr_{n+1}[A] - \Pr_n[A] \to 0$
- Nonseparability: No oracle separating A with

$\lim \Pr[A] = 1$ from A with $\lim \Pr[A] = 0$

Erdős-Rényi

On the *Evolution* of Random Graphs

Threshold Function

$r(n)$ is threshold function for A if

$p \ll r(n) \rightarrow \Pr[A] \rightarrow 0$

$p \gg r(n) \rightarrow \Pr[A] \rightarrow 1$

Existence of Triangle: n^{-1}

Existence of K_4: $n^{-2/3}$

Diameter Two: $n^{-1/2} \ln^{1/2} n$

Connectivity: $n^{-1} \ln n$

Shelah-Spencer [1988]: $\alpha \in (0,1)$, *irrational*,

$$\lim_{n\to\infty} \Pr[G(n, n^{-\alpha}) \models A] = 0 \text{ or } 1$$

Zero-One Law for $p(n) = n^{-\alpha}$

Interpretation: $n^{-\alpha}$ is *never* a threshold function for a first order A.

What happens in the evolution at $p = n^{-\pi/7}$?

NOTHING!

(in our First Order universe)

Lynch [1992]: $p = cn^{-1}$: Convergence. $\lim \Pr[A]$ exists and is "nice" function of c.

$$\Pr[\text{no triangle}] \to e^{-c^3/6}$$

$$\Pr[\text{no isolated triangle}] \to e^{-c^3 e^{-3c}/6}$$

Spencer, Thoma [1999]: $p = \frac{\ln n}{n} + cn^{-1}$: Convergence. $\lim \Pr[A]$ exists and is "nice" function of c.

$$\Pr[\text{no isolated vertices}] \to e^{-e^{-c}}$$

Random *Ordered* Graph

$p = \frac{1}{2}$

Vertices $1, \ldots, n$

Relations $=, \sim, <$

Express $1 \sim 2$:

$\exists_x \exists_y (x < y) \wedge \forall_z (z < y \rightarrow z = x) \wedge (x \sim y)$

Convergence does *not* hold

Shelah: Slow Oscillation

Ehrenfeucht Game EHR[$G_1, G_2; k$]

Parameters G_1, G_2 (disjoint); $k =$ number rounds

Players: Duplicator and Spoiler

i-th Round

- Spoiler picks $x_i \in G_1$ *or* $y_i \in G_2$
- Duplicator then picks $y_i \in G_2$ or $x_i \in G_1$
- Duplicator wins if

$$x_i \sim x_j \leftrightarrow y_i \sim y_j \text{ and } x_i = x_j \leftrightarrow y_i = y_j$$

E.g.: G_1 has isolated point, G_2 does not. Spoiler wins EHR[$G_1, G_2; 2$]

Ehrenfeucht: Duplicator wins EHR[$G_1, G_2; k$] if and only if G_1, G_2 have same first order properties of quantifier depth at most k

Ehrenfeucht Classes

$G_1 \equiv_k G_2$ if Duplicator wins EHR$[G_1, G_2; k]$

Equivalence Relation

Finite number of equivalence classes

Very large (tower function!) number of equivalence classes

G_1: Cycle length n

G_2: Two disjoint Cycles length n

Thm: For all k if n sufficiently large Duplicator wins EHR$[G_1, G_2; k]$

Proof Idea: With s moves remaining Duplicator assures that 3^s-neighborhoods of points chosen are "isomorphic."

Corollary: Connectivity not first order

Ehrenfeucht and Zero-One Law

n-point random H_n

THM: Zero-One Law

if and only if

for all k

$$\lim_{m,n\to\infty} \Pr[\text{Dupl wins EHR}[H_m, H_n; k]] = 1$$

For arbitrary first order language Duplicator must preserve *all* relations

$p = \frac{1}{2}$ Zero-One Law

With Pr $\rightarrow 1$, H_m, H_n have all extension statements up to k points. Duplicator Strategy: Find point with proper adjacencies

With Pr $\rightarrow 1$ strategy succeeds

Why doesn't this always work??

$p = n^{-\alpha}$, $\frac{1}{2} < \alpha < 1$, $k = 3$

Some, not all v, w have common neighbor u

Spoiler picks $x_1, x_2 \in H_m$ with common neighbor

Duplicator needs *foresight* to pick $y_1, y_2 \in H_n$ with common neighbor

$(R, H)-$extensions

H on $a_1, \ldots, a_r, b_1, \ldots, b_v$ with designated roots $a_1, \ldots a_r$. Assume no edges between roots.

$Ext(R, H)$: For all $x_1, \ldots, x_r$ there exist $y_1, \ldots, y_s$ with the edges (maybe more) of H.

Every point in triangle

Every two points joined by path of length seven

Every two points x_1, x_2 in K_4 except maybe $\{x_1, x_2\}$

$v =$ number nonroots; $e =$ number of edges

Dense: $v - e\alpha < 0$

Sparse: $v - e\alpha > 0$ (*dichotomy!*)

Rigid: All (R, H') dense, $H' \subseteq H$

Safe: All (R', H) sparse, $R \subseteq R'$

. . . and $G(n, n^{-\alpha})$

Expected number of extensions of $x_1, \ldots, x_r$ is $\Theta(n^v p^e) = \Theta(n^{v-e\alpha})$

Dense. $v - e\alpha < 0$. *Most* $x_1, \ldots, x_r$ have no (R, H) extension.

E.g.: $\alpha = \pi/7 \sim 0.448$. Most pairs have no common neighbor

Safe. $v - e\alpha > 0$ and no "dense parts"

Thm: All $x_1, \ldots, x_r$ have $\Theta(n^{v-e\alpha})$ extensions.

E.g.: $\alpha = \pi/7$, all pairs joined by $\Theta(n^{2-3\alpha})$ paths of length three.

t-closure $\mathrm{cl}_t(X)$ in G

For any $1 \leq u \leq t$

and any rigid (R, H) extension with u roots

and any $x_1, \ldots, x_u \in X$ with

(R, H) extension to $y_1, \ldots, y_v$

Add $y_1, \ldots, y_v$ to X

Iterate

E.g: $\alpha = \pi/6 \sim 0.523$. $t = 1$. $X = \{x_1, x_2\}$

Add common neighbors to any pair of X.

Iterate

Bounded Closure Size

E.g.: $|\mathsf{cl}_1(X)| \leq 44$ for all $|X| = 2$

n^2 choices of X

Bounded number of pictures

$np^2 = n^{-0.017\cdots}$ factor for each extension

$n^2(np^2)^{42} = o(1)$

Duplicator Look-Ahead Strategy

Constants $0 = a_0 < a_1 = 1 < \ldots < a_k$

Select so $|\mathrm{cl}_{a_i}(x_1, \ldots, x_{k-i})| < (k-i) + a_{i+1}$

After i rounds Duplicator assures that

x's and y's have "same" a_{k-i}-closures.

$a = a_i$, $b = a_{i+1}$, $X = (x_1, \ldots, x_i)$, $Y = (y_1, \ldots, y_i)$,

$x = x_{i+1}, y = y_{i+1}$

Need: If $\mathrm{cl}_a(X) \cong \mathrm{cl}_a(Y)$ then after one round

Duplicator can assure $\mathrm{cl}_b(X, x) \cong \mathrm{cl}_b(Y, y)$

Assume $\mathsf{cl}_a(X) \cong \mathsf{cl}_a(Y)$

WLOG Spoiler picks $x \in G_1$

`Inside:` $x \in \mathsf{cl}_a(X)$.

Duplicator picks "isomorphic" $y \in \mathsf{cl}_a(Y)$

`Outside:` Not Inside

$H = \mathsf{cl}_b(X, x)$, $R = cl_b(X, x) \cap cl_a(X)$

$x \in H$, $x \notin R$, (R, H) safe

Safe extensions always exist, find y

Zero-One Law

$\Rightarrow$ Complete Theory

$\Rightarrow$ Countable Model(s)

$p = n^{-\alpha}$, $0 < \alpha < 1$ irrational.

Countable list of safe (R, H)

Countable list of "witness requests"

E.g.: $\exists_{y_1,y_2} 842 \sim y_1 \wedge y_1 \sim y_2 \wedge y_2 \sim 3712$

Use "new" vertices to satisfy each witness request minimally. Get countable G

Thm: G is Countable Model

Thm: G independent of order of requests

Thm: Theory *not* $\aleph_0$-categorical

The Very Sparse Cases

$p << n^{-2}$: No Edge!

$n^{-2} \ll p(n) \ll n^{-3/2}$

No tree (or more) on 3 vertices

(for all r)

- r (or more) isolated vertices
- r (or more) isolated edges

$\aleph_0$-Categorical

$n^{-(k+1)/k} \ll p(n) \ll n^{-(k+2)/(k+1)}$

No trees (or more) on $k+2$ vertices

All trees on $\leq k+1$ vertices

$\aleph_0$-Categorical

$p = n^{-1+o(1)}$ and $p \ll n^{-1}$

All finite trees. No cycles

Not $\aleph_0$-Categorical: *May* have infinite trees!

$$p = \frac{c}{n}$$

Theory of A with Pr$[A] \rightarrow 1$:

- All trees as components
- No bicyclic (or more) subgraphs

Open: Cycles and their Neighborhoods

Countable Models:

All tree components infinitely often

Maybe infinite trees

Maybe unicyclic graphs

Binary Strings

Models $\{0,1\}^* =$ finite strings

Set $\{1,\ldots,n\}$; unary predicates U_0, U_1

$U_\alpha(x)$: x-th position α

$=$;$<$; $U_\alpha, \alpha = 0,1$

There exist two consecutive ones:

$\exists_x \exists_y [U_1(x) \wedge U_1(y) \wedge (x < y) \wedge \neg \exists_z (x < z \wedge z < y)]$

Random String $U(n,p)$: $\Pr[U_1(x) = p]$

Ehrenfeucht Semigroup

$\sigma \equiv_k \tau$: Duplicator wins EHR$[\sigma, \tau; k]$

Equivalence Relation

$E =$ set of equivalence classes

E *finite*, though very large!

$\sigma \equiv_k \sigma'$, $\tau \equiv_k \tau'$ implies $\sigma + \tau \equiv_k \sigma' + \tau'$

E forms Semigroup under concatenation

$e =$ empty string

$m\sigma = n\sigma$ if $m, n \geq 3^k$

Convergence for $U(n,p)$

Ehrenfeucht: $\lim \Pr[A]$ exists

$k =$ quantifier depth, $E =$ equivalence classes

Markov Chain!

Initial State $e =$ empty chain

$\Pr[x \to x1] = p$; $\Pr[x \to x0] = 1 - p$

NonPeriodic

Therefore: Stationary Distribution on E

$\lim \Pr[A] = \sum \lim \Pr[x]$ over x with A.

Persistent Strings

Following equivalent for $x \in E_k$:

- $\forall_y \exists_z x + y + z = x$
- $\forall_y \exists_z z + y + x = x$
- $\exists_p \exists_s \forall_y p + y + s = x$
- x persistent in Markov Chain

x called persistent.

There exist (many) persistent x (very long!)

Persistency not dependent on edge effects

x persistent implies $p + x + s$ persistent

$\lim_n \Pr[\text{persistent}] = 1$

Circular Strings

Over Z_n with $C(x,y,z) =$ "clockwise"

No edge effects

Zero-One Law for p constant

Thm (Shelah/JS):

Zero-One Law if $n^{-1/k} \ll p(n) \ll n^{-1/(k+1)}$

Countable Models (StJohn/JS):

$p \ll n^{-1}$ Line Z All 0

$n^{-1} \ll n^{-1} \ll n^{-1/2}$ Page Z^2

One 1 on each "line"

$n^{-1/2} \ll p(n) \ll n^{-1/3}$ Book Z^3

Each page with one line with two 1's

Volume, Library,. . .

Coming Attractions

Thursday, 1:30

Analytic Questions

Given Zero-One Law

A with quantifier depth k

Asymptotics of $n(k)$ so that

$n \geq n(k) \Rightarrow \Pr[A] < 0.01$ or $\Pr[A] > 0.99$

$G(n,p)$ with $p = \frac{1}{2}$:

$\binom{n}{k} 2^k (1 - 2^{-k})^{n-k} \to 0$

$n = \Theta(2^k k^2)$

Succint Definitions

General First Order Structure

Def: $D(G) =$ smallest *quantifier depth*

of A that defines G

What is $D(G)$ for random n-element model?

Kim/Pikhurko/Verbitsky/JS

$G(n, \frac{1}{2})$: $\Theta(\ln n)$

StJohn/JS:

$G_{<}(n, \frac{1}{2})$: $\Theta(\ln^* n)$

BitString $U(n, \frac{1}{2})$: $\Theta(\ln \ln n)$