Limiting probabilities of first order properties in sparse random graphs

Alberto Larrauri Marc Noy
Universitat Politècnica de Catalunya, Barcelona

Tobias Müller
University of Groningen

First order logic (FO) of graphs

Quantifiers: $\forall, \exists$
Variables: $x, y, z, \ldots$
Boolean connectives and equality: $\vee, \wedge, \neg, \rightarrow, =$
Predicates $P(x)$, $Q(x, y)$, ...

$E(x, y)$ adjacency relation written $x \sim y$
Assumed symmetric and antireflexive

Some examples

- Existence of a triangle: $\exists x \, \exists y \, \exists z \; (x \sim y) \wedge (y \sim z) \wedge (z \sim x)$
- Existence of fixed H as a subgraph
- There are at most a cycles of length at most k

FO cannot express connectivity, planarity, 3-colorability...

The $G(n, p)$ model of random graphs

$G(n, p)$ with $0 < p < 1$

Vertices: $V = \{1, 2, \ldots, n\}$
Each edge $\{i, j\}$ is in $G(n, p)$ independently with probability p

$$\mathbf{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$$

The expected number of edges is $p\binom{n}{2} \sim p\frac{n^2}{2}$

p constant — Dense graphs: $\Theta(n^2)$ edges

$p = \dfrac{c}{n}$ — Sparse graphs: $\Theta(n)$ edges

Sparse random graphs

$G_n = G(n, c/n)$

The phase transition Erdős-Rényi (1960)

- For $c < 1$, all components in G_n are either trees or have a unique cycle, and have size $O(\log n)$
- For $c > 1$ there is a unique component of size $\Theta(n)$

FO logic cannot capture the transition since it cannot express that a graph is acyclic or unicyclic

Limiting probabilities

$G_n = G(n, c/n)$

Given a FO property A we are interested in the limiting probability

$$\lim_{n\to\infty} \mathbf{P}[G_n \text{ satisfies } A]$$

Lynch 1992
For every FO property A

$$\lim_{n\to\infty} \mathbf{P}[G_n \text{ satisfies } A\,] = f_A(c)$$

and $f_A(c)$ is an expression in c using only $+$, $\times$, $\{\lambda\colon \lambda \in \Lambda(c)\}$ and exponentials, hence it is a C^∞ function

Remark Alberto Larrauri has recently generalized Lynch's result to sparse hypergraphs [Journal of Logic and Computation, to appear]

The set of limiting probabilities for sparse graphs

$L_c = \{\lim \mathbf{P}\left[G\left(n, \frac{c}{n}\right) \text{ satisfies } A\right] : \text{FO property } A\}$

L_c is a countable set, we consider its topological closure

$$\overline{L_c} \subseteq [0,1]$$

Theorem [Alberto Larrauri, Tobias Müller, M.N.]

Let $c_0 \approx 0.9368$ be the unique positive root of

$$e^{\frac{c}{2}+\frac{c^2}{4}}\sqrt{1-c} = \frac{1}{2}$$

- $\overline{L_c}$ is a finite union of intervals
- For $c \geq c_0$ we have $\overline{L_c} = [0,1]$
- For $0 < c < c_0$ there is at least one gap in $\overline{L_c}$, that is, an interval $[a,b] \subseteq [0,1]$ with $[a,b] \cap \overline{L_c} = \emptyset$

Previous work

$\mathcal{P}$ labelled planar graphs

Heinig, Müller, N., Taraz 2018

- With the uniform distribution on graphs in $\mathcal{P}$ with n vertices $\overline{L_c}$ is a finite union of intervals (in fact 108 intervals of length $\approx 10^{-6}$)
- For the class of random forests (acyclic graphs) $\overline{L_c} = [0, 0.170] \cup [0.223, 0.393] \cup [0.606, 0.776] \cup [0.830, 1]$
- For every minor-closed class of graphs which is addable (the forbidden minors are 2-connected) $\overline{L_c}$ is finite union of intervals and there is always a gap

Sketch of proof

1. No gap for $c \geq 1$
2. At least one gap for $c < c_0$
3. No gap for $c_0 \leq c < 1$

1. No gap for $c \geq 1$

X_k = number of k-cycles in $G(n, \frac{c}{n})$

$$X_k \xrightarrow[n\to\infty]{\text{d}} \text{Poisson}\left(\frac{c^k}{2k}\right)$$

and $X_3, \ldots, X_k$ are asymtotically independent for fixed k

$$X_{\leq k} := X_3 + \cdots + X_k \xrightarrow[n\to\infty]{\text{d}} \text{Poisson}\left(\mu_k = \sum_{i=3}^{k} \frac{c^k}{2k}\right)$$

- $\lim_{k\to\infty} \mu_k = \infty$ since $c \geq 1$
- The property $\{X_{\leq k} \leq a\}$ is FO expressible for fixed k

$$\mathbf{P}(\text{Poisson}(\mu) \leq \mu + x\sqrt{\mu}) \xrightarrow[\mu\to\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

Given $p \in (0,1)$ and $\epsilon > 0$ take x with $\Phi(x) = p$ and μ_0 such that

$$|\mathbf{P}(\text{Poisson}(\mu) \leq \mu + x\sqrt{\mu}) - p| < \epsilon, \qquad \text{for } \mu \geq \mu_0$$

Finally take k such that $\mu_k \geq \mu_0$

Hence $\overline{L_c} = [0,1]$

2. At least one one gap for $c < c_0$

Structure of $G_n = G(n, c/n)$ for $c < 1$
Erdős-Rènyi (1960)

- $G_n = F_n \cup H_n$

  F_n is a **forest**, H_n is a collection of **unicyclic** graphs

- For every fixed tree T and $m \geq 1$,
  F_n contains at least m copies of T

Zero-one law in F_n For every FO property A

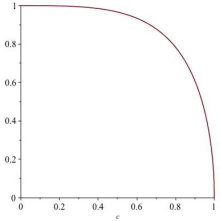
$$\lim_{n\to\infty} \mathbf{P}[F_n \text{ satisfies } A] \in \{0, 1\}$$

Idea F_n contains arbitrarily many copies of each tree, hence two random instances of F_n cannot be distinguished by FO properties

This can be proved for instance using combinatorial games (Ehrenfeucht-Fraïssé)

Let F be the property of G_n being acyclic (a forest)

$$\lim_{n\to\infty} \mathbf{P}(F) = \prod_{k\geq 3} e^{-c^k/(2k)} = e^{-\sum_{k\geq 3} \frac{c^k}{2k}} = \sqrt{1-c}\,e^{c/2+c^2/4} = f(c)$$



We have $\lim_{n\to\infty} \mathbf{P}(F) > 1/2$ for $c < c_0$
Given a FO property A, we have in terms of limiting probabilities

$$\mathbf{P}(A) = \mathbf{P}(A|F)f(c) + \mathbf{P}(A|\neg F)(1-f(c))$$

If $\mathbf{P}(A|F) = 1$ then $\mathbf{P}(A) \geq f(c) > 1/2$

If $\mathbf{P}(A|F) = 0$ then $\mathbf{P}(A) \leq 1 - f(c) < 1/2$

Hence $[1-f(c), f(c)]$ is a gap

3. No gap for $c_0 \leq c \leq 1$

- $G_n = F_n \cup H_n$
  F_n is a **forest**, H_n is a collection of **unicyclic** graphs
- $\mathbf{E}(|H_n|)$ is bounded

Restricting to F_n we have the FO zero-one law

Hence whether G_n satisfies a FO property depends solely on the fragment H_n

$$\mathbf{P}[H_n \cong H] \to p_H = f(c)\frac{(ce^{-c})^{|V(H)|}}{\mathrm{aut}(H)}$$

$$\mathbf{P}[G_n \text{ satisfies } A] \to \sum_{H \in \mathcal{H}_A} p_H$$

It follows that $\overline{L_c}$ is the collection of subsums of the series

$$\sum_{H \text{ fragment}} p_H = 1$$

Nymann-Sáenz 2000 (Kakeya 1915)

Let $\sum_{n\geq 0} p_n < +\infty$ with $p_n \geq 0$

If $p_i \leq p_{i+1} + p_{i+2} + \cdots$ (term $\leq$ tail) for all $i \geq 0$ then

$$\left\{ \sum_{i \in A} p_i : A \subset \mathbb{N} \right\} = \left[0, \sum_{n=0}^{\infty} p_n \right]$$

Order the limiting probabilities of the fragments H

$$p_0 \geq p_1 \geq p_2 \geq \cdots$$

Considering

$$\sum_{|V(H)|=k} p_H = f(c)(ce^{-c})^k \sum_{|V(H)|=k} \frac{1}{\mathrm{aut(H)}}$$

we show that $p_i \leq \sum_{j>i} p_j$ for fragments of size k ≥ 4

We complete the argument for size 3 (a triangle)

$p_0 =$ probability of being acyclic (empty fragment)

$p_0 \leq 1 - p_0$ means $p_0 \geq 1/2$ which holds because $c \geq c_0$

Sparse hypergraphs

$G^d(n,p)$ random d-hypergraph in which each d-edge has independent probability p

Take $p = \dfrac{c}{n^{d-1}}$

Sparse sin the expected number of edges is $p\binom{n}{d} = \Theta(n)$

Theorem

Let c_0 be the unique positive root of

$$\exp\left(\frac{c}{2(d-2)!}\right)\sqrt{1-\frac{c}{2(d-2)!}} = \frac{1}{2}. \qquad (1)$$

- $\overline{L_c}$ is a finite union of intervals
- $\overline{L_c} = [0,1]$ for $c \geq c_0$
- At least one gap for $0 < c < c_0$

Graphs with given degree sequence

(ongoing project with Alberto Larrauri and Guillem Perarnau)

Consider random graphs (uniform distribution) with degree sequence $\mathcal{D} = d_0(n), d_1(n), \ldots$

- $d_i(n)$ = number of vertices of degree i
- $\sum_i d_i(n) = n$
- $\lim_{n\to\infty} d_i(n)/n = \lambda_i$
- $\mu_1 = \lim_{n\to\infty} \sum_i i\lambda_i$ finite
- $\mu_2 = \lim_{n\to\infty} \sum_i i^2\lambda_i$ finite

Molloy-Reed 1995
There exists a component of size $\Theta(n)$ iff $\mu_2 - 2\mu_1 > 0$

$\overline{L_{\mathcal{D}}}$ the closure of the set of limiting probabilities of FO properties for random graphs with degree sequence $\mathcal{D}$

When is $\overline{L_{\mathcal{D}}} = [0, 1]$?

Partial results so far indicate a behavior generalizing what we have found for $G(n, p)$