Random Graphs, First-Order Logic, and AC^0 Circuits

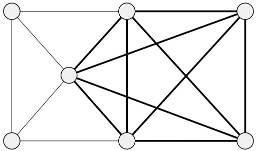
Yijia Chen
Fudan University

Introduction

The clique problem

Definition

Let G be a graph. A clique in G is a subgraph, wherein every two vertices are adjacent in G. A k-clique is a clique that contains k vertices.



A graph with a 5-clique.

The k-clique problems in first-order logic

1. For every fixed $k \in \mathbb{N}$

$$G \text{ has a } k\text{-clique} \iff G \models \exists x_1 \cdots \exists x_k \bigwedge_{1 \leqslant i < j \leqslant k} E x_i x_j.$$

2. There is no FO-sentence φ with $k - 1$ variables such that

$$G \text{ has a } k\text{-clique} \iff G \models \varphi.$$

The proof uses Ehrenfeucht-Fraïssé-game from model theory.

G with a built-in ordering, or even arithmetic

In computer science, G is always stored by some data structure, e.g., the adjacency matrix $A = A(G)$ where

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between the } i\text{-th and } j\text{-th vertices} \\ 0 & \text{otherwise.} \end{cases}$$

In particular, we assume an ordering on the vertices, hence the ordered graph $\langle G, < \rangle$.

We could even allow arithmetic on G, i.e.,

$$\langle G, <, +, \times \rangle.$$

Embarrassingly

Problem

Let $k \in \mathbb{N}$. Is there an FO*-sentence φ using only $k-1$ variables such that*

$$G \text{ has a } k\text{-clique} \quad \Longleftrightarrow \quad \langle G, < \rangle \models \varphi?$$

Or even

$$G \text{ has a } k\text{-clique} \quad \Longleftrightarrow \quad \langle G, <, +, \times \rangle \models \varphi?$$

Rossman's result

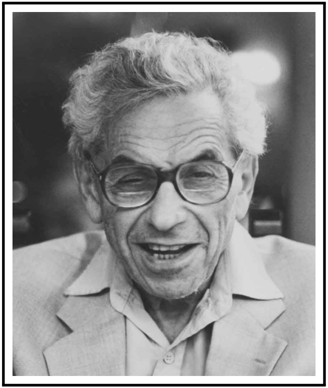
Theorem (Rossman, 2008)

Let $k \in \mathbb{N}$. There is no FO*-sentence φ using at most $k/4$ variables such that*

$$G \text{ has a } k\text{-clique} \quad \Longleftrightarrow \quad \langle G, <, +, \times \rangle \models \varphi.$$

Rossman's proof uses Erdős-Rényi random graphs and AC^0-circuits.

Random Graphs and FO



Erdős-Rényi random graphs

Definition

Let $n \in \mathbb{N}$ and $p \in \mathbb{R}$ with $0 \leqslant p \leqslant 1$. Then $G \in \mathrm{ER}(n, p)$ is the Erdős-Rényi random graph on vertex set $[n]$ constructed by adding every edge $e \in \binom{[n]}{2}$ independently with probability p.

Example

Let $n \in \mathbb{N}$ and $\mathscr{G}_n$ be the class of graphs with vertex set $[n]$. Then, $\mathrm{ER}(n, 1/2)$ is the uniform distribution on $\mathscr{G}_n$.

The 0-1 Law for FO

Theorem (Glebskii, Kogan, Liagonkii, and V.A. Talanov, 1969; Fagin 1976)

Let $\varphi \in \mathrm{FO}$. *Then*

$$\lim_{n\to\infty} \Pr_{G\in \mathrm{ER}(n,1/2)} \left[G \models \varphi\right] = \begin{cases} 1 & \text{if } G^\infty \models \varphi \\ 0 & \text{if } G^\infty \not\models \varphi, \end{cases}$$

where G^∞ *is the infinite Rado graph.*

A consequence of the 0-1 law

Definition

$$\textsc{Parity} = \{G \mid G \text{ has even number of vertices}\}.$$

The sequence

$$\frac{\left|\{G \in \mathscr{G}_n \mid G \text{ has even number of elements}\}\right|}{\left|\mathscr{G}_n\right|}$$

does not converge.

Corollary

$\textsc{Parity} \notin \mathsf{FO}$.

The breakdown of the 0-1 law on ordered graphs

Let

$$\varphi := \text{“there is an edge between the first and the second vertices.”}$$

Then

$$\lim_{n\to\infty} \frac{\left|\{G \in \mathscr{G}_n \mid \langle G, < \rangle \models \varphi\}\right|}{\left|\mathscr{G}_n\right|} = \frac{1}{2}.$$

Definition

$$\textsc{Parity}^* = \{G \mid G \text{ has even number of red vertices}\}.$$

Theorem (Furst, Saxe, and Sipser, 1981; Ajtai, 1983)

There is no FO*-sentence* φ *such that*

$$G \in \textsc{Parity}^* \quad \Longleftrightarrow \quad \langle G, <, +, \times \rangle \models \varphi.$$

Hence, there is no FO*-sentence* φ *such that*

$$G \in \textsc{Parity} \quad \Longleftrightarrow \quad \langle G, < \rangle \models \varphi.$$

The Ehrenfeucht-Fraïssé-game is extremely hard to play on $\langle G, < \rangle$, let alone $\langle G, <, +, \times \rangle$.

The actual theorem of Furst et. al and Ajtai

Theorem

$\textsc{Parity}^* \notin \mathsf{AC}^0$, *i.e., the parity problem has no* AC^0*-circuits.*

What is AC^0?

A sequence $(\mathsf{C}_n)_{n\in\mathbb{N}}$ of Boolean circuits is in AC^0 if there is a constant $d \in \mathbb{N}$ such that for every $n \in \mathbb{N}$

(i) the number of inputs of C_n is polynomially bounded in n;

(ii) the depth of C_n is bounded by d;

(iii) the size of C_n is polynomially bounded in n.

The k-clique problem can be computed by the following sequence of depth-2 circuits

$$\bigvee_{K \in \binom{[n]}{k}} \bigwedge_{\{i,j\} \in \binom{K}{2}} x_{\{i,j\}},$$

which are of size $n^{k+O(1)}$.

From FO to AC^0

Theorem

For every $\varphi \in \mathrm{FO}$ *there is a family of* AC^0*-circuits* $\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ *such that for every structure* $\mathcal{A}$ *with* n *elements*

$$\mathcal{A} \models \varphi \quad \Longleftrightarrow \quad \mathsf{C}_n(\mathcal{A}) = 1.$$

From AC^0 to FO

AC^0-circuits can decide uncomputable problems, hence $\mathsf{AC}^0 \not\subseteq \mathsf{FO}$.

Theorem

Let $\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ be a family of AC^0-circuits. Moreover, assume

$$1^n \mapsto C_n$$

can be computed by a deterministic logarithmic time Turing machine. Then, there is an FO-sentence φ such that every structure $\mathcal{A}$ with n elements we have

$$\mathsf{C}_n(\mathcal{A}) = 1 \quad \Longleftrightarrow \quad \langle \mathcal{A}, <, +, \times \rangle \models \varphi.$$

$\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ is said to be dlogtime-uniform.

Remark

The circuits $\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ is the previous theorem is in fact dlogtime-uniform.

Håstad's Switching Lemma

Lemma (Håstad, 1986)

Let f be expressible as a k-DNF, and ρ a random restriction that assigns random values to $t \geq n/2$ randomly selected input bits. Then for every $s \geq 2$,

$$\Pr_{\rho}\left[f\mid_{\rho} \textit{ is not expressible as } s\text{-}\mathrm{CNF}\right] \leqslant \left(\frac{(n-t)k^{10}}{n}\right)^{s/2}.$$

Håstad's Switching Lemma can be roughly viewed as probabilistic quantifier elimination.

The moral

By going from FO to AC^0, we can use many tools from combinatorics and probability that are very hard to apply to FO directly.

Our Results

Recall

Definition

Let $n \in \mathbb{N}$ and $p \in \mathbb{R}$ with $0 \leqslant p \leqslant 1$. Then $G \in \mathrm{ER}(n, p)$ is the Erdős-Rényi random (ordered) graph on vertex set $[n]$ constructed by adding every edge $e \in \binom{[n]}{2}$ independently with probability p.

Lemma

The expected size of a maximum clique in $G \in \mathrm{ER}(n, 1/2)$ *is approximately* $2 \log n$.

Planting a clique of size $5 \log n$

Theorem

$$\lim_{n \to \infty} \Pr_{G \in \mathrm{ER}(n,1/2)} \left[G \textit{ contains a clique of size } 5 \log n \right] = 0.$$

We consider a "planted clique" distribution $G + K$ with $K \in P(n, 5 \log n)$:

Definition

Let $n, k \in \mathbb{N}$. Then $P(n, k)$ is the uniform distribution over all cliques of size k on the vertex set $[n]$.

The planted clique conjecture

Conjecture (Jerrum, 1992; Kucera, 1995)

For every polynomial time algorithm $\mathbb{A}$ *there is a polynomial* $p \in \mathbb{N}[X]$ *such that for all sufficiently large* $n \in \mathbb{N}$

$$\Pr_{\substack{G \in \mathrm{ER}(n,1/2) \\ K \in P(n, 5 \log n)}} \left[\mathbb{A}(G + K) \neq K \right] \geq \frac{1}{p(n)}.$$

That is, $\mathbb{A}$ *fails to find the planted clique with non-negligible probability.*

Assuming the planted clique conjecture

1. The classical clique problem is hard on average.
2. $(G, K) \mapsto G + K$ is a one-way function [Juels and Peinado, 2000].
3. Nash equilibrium is hard to approximate [Minder and Vilenchik, 2009; Hazan and Krauthgamer, 2011].
4. It is hard to decide whether a graph is Ramsey [Santhanam, 2010].

Our main technical result

Theorem

Assume:

(i) AC^0 *circuits,*

(ii) $\alpha : \mathbb{N} \to \mathbb{R}^+$ *with* $\lim_{n\to\infty} \alpha(n) = 0$,

(iii) $k : \mathbb{N} \to \mathbb{N}$ *with* $k(n) \leqslant n^{1-\varepsilon}$ *for* $1 \geq \varepsilon > 0$.

Then

$$\lim_{n\to\infty} \Pr_{\substack{G \in \mathrm{ER}(n, n^{-\alpha(n)}),\\ K \in P(n, k(n))}} \Big[\mathsf{C}_n(G) = \mathsf{C}_n(G+K) \Big] = 1.$$

Our main technical result (cont'd)

Theorem

Assume . . . . Then

$$\lim_{n\to\infty} \Pr_{\substack{G\in \mathrm{ER}(n,n^{-\alpha(n)}),\\ K\in P(n,k(n))}} \Big[\mathrm{C}_n(G) = \mathrm{C}_n(G+K) \Big] = 1.$$

By taking $\alpha(n) = 1/\log n$ *and* $k(n) = 5\log n$

$$\lim_{n\to\infty} \Pr_{\substack{G\in \mathrm{ER}(n,1/2),\\ K\in P(n,5\log n)}} \Big[\mathrm{C}_n(G) = \mathrm{C}_n(G+K) \Big] = 1.$$

Thus,

$$\lim_{n\to\infty} \left| \Pr_{G\in \mathrm{ER}(n,1/2)} \big[\mathrm{C}_n(G) = 1\big] - \Pr_{\substack{G\in \mathrm{ER}(n,1/2),\\ K\in P(n,5\log n)}} \big[\mathrm{C}_n(G+K) = 1\big] \right| = 0.$$

Our main technical result (cont'd)

If we take

$$\alpha(n) = \frac{1}{\log\log\log n} \quad \text{and} \quad k(n) = n^{0.999},$$

then with high probability $G \in \mathrm{ER}(n, n^{-\alpha(n)})$ has no clique of size

$$2\log\log\log n,$$

but still we have

$$\lim_{n\to\infty} \Pr_{\substack{G\in\mathrm{ER}(n,n^{-\alpha(n)}),\\ K\in P(n,n^{0.999})}} \Big[\mathrm{C}_n(G) = \mathrm{C}_n(G+K)\Big] = 1.$$

An application in logic

Theorem

For every $m \in \mathbb{N}$

$$\lim_{n\to\infty} \Pr_{\substack{G\in \mathrm{ER}(n,1/2),\\ K\in P(n,5\log n)}} \Big[\langle G,<,+,\times\rangle \equiv_m \big\langle (G+K),<,+,\times\big\rangle\Big] = 1,$$

where $\langle G,<,+,\times\rangle \equiv_m \big\langle (G+K),<,+,\times\big\rangle$ means that for every FO*-sentence φ with quantifier rank at most m*

$$\langle G,<,+,\times\rangle \models \varphi \quad\Longleftrightarrow\quad \big\langle (G+K),<,+,\times\big\rangle \models \varphi.$$

Thank You!