Logic and Random Graphs

Marc Noy
Universitat Politècnica de Catalunya, Barcelona
Barcelona Graduate School of Mathematics

Limiting probabilities in classes of graphs

$\mathcal{G}$ class of labelled graphs $\quad V = \{1, 2, \ldots, n\}$
Probability distribution on $\mathcal{G}_n$ for each n

Given a graph property A consider

$$\lim_{n\to\infty} \mathbf{P}(G \text{ satisfies } A : \ G \in \mathcal{G}_n)$$

We are interested on properties expressible in First order logic (FO) or fragments of Second order logic (SO)

The classical Zero-one law
If $\mathcal{G}$ is the class of all labelled graphs, the limit exists for every FO property A and is equal to 0 or 1

Either A or $\neg A$ holds asymptotically almost surely

A holds with high probability (w.h.p.) if $\mathbf{P}(A) \to 1$ as $n \to \infty$

First order logic (FO)

Quantifiers: $\forall, \exists$
Variables: $x, y, z, \ldots$
Boolean connectives and equality: $\vee, \wedge, \neg, \rightarrow, =$

FO for simple undirected graphs
$E(x, y)$ 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

- Existence of an isolated vertex: $\exists x, \forall y\ \neg(x \sim y)$
- Existence of a connected component isomorphic to H

A sentence is a formula in which all variables are quantified

Limitations of FO logic

FO cannot express Graph Connectivity

$$\forall x \forall y \; \neg(x = y) \rightarrow \exists x_1, \ldots, x_k \text{ distinct from } x \text{ and } y$$
$$x \sim x_1, x_1 \sim x_2, \ldots, x_k \sim y$$

$$\bigvee_{k=1}^{\infty} (\exists x_1, \ldots, x_k \neq x, y \quad x \sim x_1, x_1 \sim x_2, \ldots, x_k \sim y)$$

FO logic is local in a technical sense (Gaifman's lemma)

Monadic Second Order Logic (MSO)

Graph connectivity is not expressible in FO, but it is in

Monadic Second Order (MSO) logic which is
FO + quantification over sets of vertices (unary relations)

$$\forall A \subset V \quad (\exists x \in A \ \wedge \ \exists x \notin A) \to \ \exists x \in A \ \exists y \notin A \ x \sim y$$

- Being connected, k-colorable, planar, are all in MSO
- Being Hamiltonian is not in MSO but it is in MSO_2: quantification over sets of vertices and sets of edges

Theorem
Graph connectivity is not expressible in FO

First proof idea: analyze each individual FO sentence and show that it cannot express connectivity

Theorem [Trakhtenbrot 1950]
Given a FO sentence ϕ, it is undecidable whether there exists a finite graph satisfying ϕ

G is a model of ϕ $\qquad G \models \phi$

Winning idea: analyze simultaneously all sentences of fixed depth
Depth of $\phi =$ maximum number of nested quantifiers in ϕ

- depth$(\phi) = 0$ if ϕ is quantifier free
- depth$(\psi) + 1$ if $\phi = \forall x \psi(x)$ or $\phi = \exists x \psi(x)$

Logical equivalence of graphs
$G \equiv_k H$ if G and H satisfy exactly same sentences of depth $\leq k$
The relation $\equiv_k$ has finitely many equivalence classes

Suppose for each $k \geq 1$ we find graphs G_k, H_k such that

- G_k is connected and H_k is not
- $G_k \equiv_k H_k$

Suppose ϕ expresses connectivity and let $k =$ depth(ϕ)
Contradiction!

Logic through combinatorial games

Ehrenfeucht-Fraïssé game $\mathrm{Ehr}_k(G, H)$

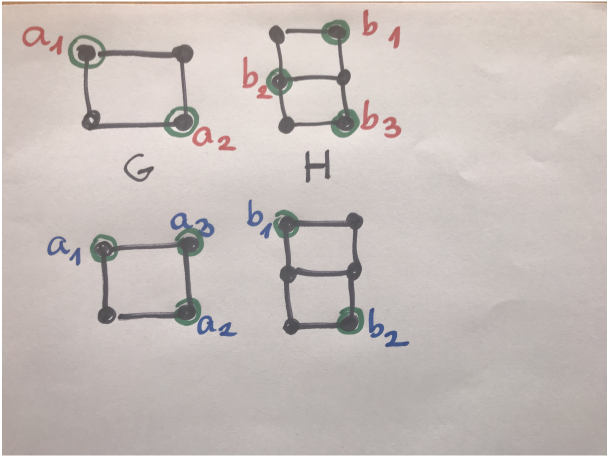
- Spoiler and Duplicator play k rounds on two graphs G, H
- At each round Spoiler picks a vertex (from either G or H) and Duplicator must pick a vertex from the other graph

$(a_1, \ldots, a_k)$ vertices selected from G
$(b_1, \ldots, b_k)$ vertices selected from H

Duplicator wins iff $(a_1, \ldots, a_k) \leftrightarrow (b_1, \ldots b_k)$ partial isomorphism

Theorem (Ehrenfeucht-Fraïssé)
$G \equiv_k H \iff$ Duplicator has a winning strategy for $\mathrm{Ehr}_k(G, H)$

Provides a purely combinatorial characterization of $\equiv_k$



$$\exists x\, \exists y\, \exists z\; (x \neq y \wedge x \neq z \wedge y \neq z) \wedge (x \not\sim y \wedge x \not\sim z \wedge y \not\sim z)$$

$$\forall x \forall y (x \neq y \wedge x \not\sim y) \rightarrow \exists z\, (z \neq x \wedge z \neq y \wedge z \sim x \wedge z \sim y)$$

Proofs of non-expressibility in FO

Connectivity

$G = C_{2^k}, \quad H = C_{2^k} \cup C_{2^k}$

Claim: $G \equiv_k H$

Proof is by induction on k

Being planar, acyclic, Hamiltonian, bipartite, 3-colorable ...
are not expressible in FO

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

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

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

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

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

- p constant dense graphs
- $p = \dfrac{c}{n}$ sparse graphs
- $p = \dfrac{\log n}{n}$ threshold for connectivity

Observation $G(n, \frac{1}{2})$ = uniform distribution on labelled graphs

Dense graphs: constant p

Theorem [Glebskiĭ, et al. 1969, Fagin 1976]
The zero-one law in FO holds for the class of all labelled graphs

The same holds in $G(n,p)$ with p constant

Extension Property E_k For fixed k and disjoint A, B with $|A| = |B| = k$ there exists $z \notin A \cup B$ with $z \sim A$ and $z \not\sim B$

$$\mathbf{P}(E_k \text{ does not hold }) \leq \binom{n}{k}\binom{n-k}{k}(1-p^k(1-p)^k)^{n-2k} \to 0,\ n \to \infty$$

Property E_k For fixed k and disjoint A, B with $|A| = |B| \leq k$ there exists $z \notin A \cup B$ with $z \sim A$ and $z \not\sim B$

Claim Duplicator wins $\mathrm{Ehr}_k(G, H)$ w.h.p. for large $G, H \sim G(n, p)$

At round i Spoiler plays a_i, say in G
$C = \{b_j \in H \colon a_j \sim a_i, j < i\}$, quad $D = \{b_j \in H \colon a_j \not\sim a_i, j < i\}$
Duplicator choses b_i such that $b_i \sim C$ and $b_i \not\sim D$

Hence $G \equiv_k H$

For every FO sentence ϕ either both G, H satisfy ϕ or both do not satisfy ϕ

The zero-one law holds n $G(n, p)$ for constant p $\square$

Sparse graphs

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

Erdős-Rényi 1960

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

FO logic does not capture this jump

Lynch 1992
For every FO sentence ϕ

$$\lim_{n\to\infty} \mathbf{P}\left[G\left(n, \frac{c}{n}\right) \models \phi\right] = f(c)$$

and $f(c)$ is analytic in c

The Shelah-Spencer theorem

No zero-one law in $G(n, n^{-2/3})$

$p = n^{-2/3}$ is a threshold for the appearance of K_4,
and the number of copies of K_4 is Poisson distributed

$$\lim_{n\to\infty} \mathbf{P}(G(n, n^{-2/3}) \text{ contains } K_4) \neq 0, 1$$

If $\alpha \in (0, 1)$ is rational, there exists a graph H such that $p = n^{-\alpha}$ is a threshold function for containing a copy of H

Shelah-Spencer 1989
For $p = n^{-\alpha}$, $\alpha \in (0, 1)$, the zero-one law holds iff α is irrational

Main idea: for $\alpha \notin \mathbb{Q}$ there is a (very subtle) strategy in the EF game based on rich extension properties on rooted graphs

Remark For $\alpha \in \mathbb{Q}$ there exist non-convergent sentences

Constrained classes of graphs

- Bipartite, Triangle-free, d-regular, Chordal (no chordless cycle)
- Acyclic, Planar, Graphs on surfaces

Uniform distribution in all cases

Proving 0-1 laws in these classes needs some kind of counting

- The 0-1 law holds for bipartite graphs, proof as for $G(n, \frac{1}{2})$
- Almost every triangle-free graph is bipartite, hence the 0-1 law holds for triangle-free graphs
- No 0-1 law for d-regular graphs, but there is a convergence law: every sentence has a limiting probability
- Almost all chordal graphs are Split: $V = K_n \cup \overline{K_m}$
  The 0-1 law holds, as for bipartite graphs

Trees

Labelled trees with uniform distribution

$$\mathbf{P}(T) = \frac{1}{n^{n-2}}, \qquad T \in \mathcal{T}_n$$

Theorem McColm (2002)
The zero-one law holds for trees

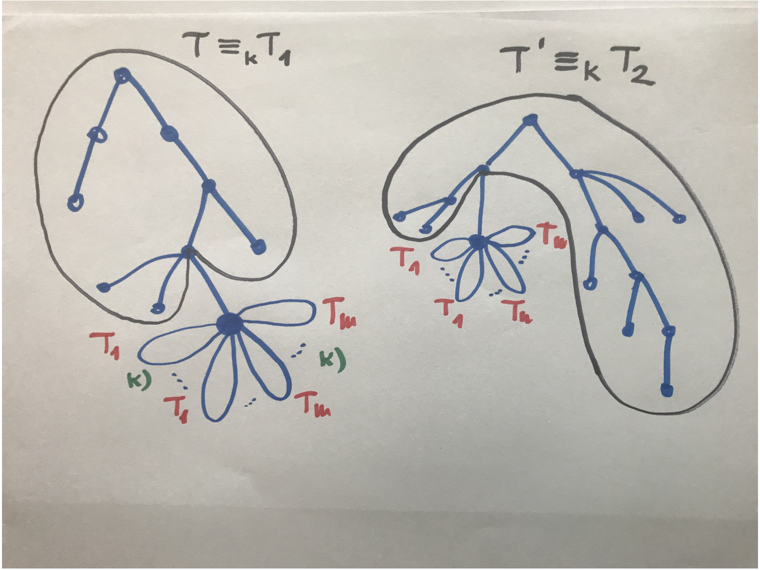
Proof via Ehrenfeucht-Fraïssé games

For each k there is a rooted tree U_k such that

- If T and T' contain a terminal copy of U_k, then Duplicator wins $\mathrm{Ehr}_k(T, T')$
- A random tree contains a terminal copy of U_k w.h.p.

U_k contains k terminal copies of each $\equiv_k$ equivalence class

$T_1, \ldots, T_m$ representatives of $\equiv_k$ equivalence classes



Forests (acyclic graphs)
There is no zero-one law in the class of forests

$$\mathbf{P}(\text{Random forest has an isolated vertex}) \to e^{-1}$$

Theorem Heinig, Müller, M.N., Taraz
The convergence law holds for forests: for every sentence ϕ

$$\lim_{n\to\infty} \mathbf{P}(G \models \phi : \ G \in \mathcal{G}_n) \quad \text{exists}$$

Planar graphs and graphs on surfaces

Theorem [HMNT]
The zero-one law holds for connected planar graphs
The convergence law holds for planar graphs

Holds also for classes closed under subgraphs and contractions

- Series-parallel and outerplanar
- Bounded tree-width

Uses results by McDiarmid and Giménez-M.N.

Theorem [HMNT]
The zero-one law holds in MSO for connected planar graphs
The convergence law holds in MSO for planar graphs

MSO = FO + quantification over sets of vertices

EF games for MSO. Vertex moves and set moves

$$\begin{array}{ll} a_1, \ldots, a_k, A_1, \ldots A_s & \text{selected in } G \\ b_1, \ldots, b_k, B_1, \ldots B_s & \text{selected in } H \end{array}$$

Duplicator wins if

- $(a_1, \ldots, a_k) \leftrightarrow (b_1, \ldots b_k)$ partial isomorphism
- $a_i \in A_j \iff b_i \in B_j$

U_k contains 2^k terminal copies of each $\equiv_k$ equivalence class

Duplicator uses pigeonhole to win the game

$\mathcal{G}^S$ class of graphs embeddable on a fixed surface S

The zero-one law in FO holds for connected graphs in $\mathcal{G}^S$
The convergence law in FO holds for graphs in $\mathcal{G}^S$

$B_r(x) = \{y \colon d(x,y) \leq r\}$

- Gaifman's locality Theorem
  Every FO formula is logically equivalent to a Boolean combination of elementary local formulas

$$\exists x_1, \ldots, \exists x_s \left(\bigwedge_{i \neq j} d(x_i, x_j) > 2r \right) \wedge \left(\bigwedge_i \psi_{B_r(x_i)}(x_i) \right)$$

- Random graphs in $\mathcal{G}^S$ are locally planar
  [Chapuy-Fusy-Giménez-Mohar-M.N. 2011]

  For fixed $r > 0$ the balls $B_r(x)$ are planar

$\mathcal{G}^S$ class of graphs embeddable on a fixed surface S

The zero-one law in FO holds for connected graphs in $\mathcal{G}^S$
The convergence law in FO holds for graphs in $\mathcal{G}^S$

$B_r(x) = \{y \colon d(x,y) \leq r\}$

- Gaifman's locality Theorem
  Every FO formula is logically equivalent to a Boolean combination of elementary local formulas

$$\exists x_1, \dots, \exists x_s \left(\bigwedge_{i \neq j} d(x_i, x_j) > 2r \right) \wedge \left(\bigwedge_i \psi_{B_r(x_i)}(x_i) \right)$$

- Random graphs in $\mathcal{G}^S$ are locally planar
  [Chapuy-Fusy-Giménez-Mohar-M.N. 2011]

  For fixed $r > 0$ the balls $B_r(x)$ are planar

Question
Does the convergence law in MSO holds for graphs in $\mathcal{G}^S$?

Theorem Atserias, Kreutzer, M.N. 2018

- For every surface S different from the sphere there exists a sentence ϕ in MSO such that

$$\lim_{n\to\infty} \mathbf{P}(G \models \phi : \ G \in \mathcal{G}_n^S) \quad \text{does not exist}$$

- For every rational number $q \in [0,1]$ there exists a sentence ϕ in MSO such

$$\lim_{n\to\infty} \mathbf{P}(G \models \phi : \ G \in \mathcal{G}_n^S) = q$$

MSO for graphs on surfaces

$\mathcal{G}^S$ = graphs embeddable in S of genus $g > 0$

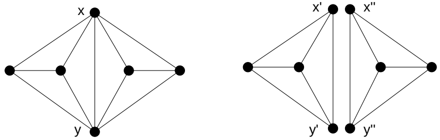
Random graphs in $\mathcal{G}^S$ satisfy

- There exists a unique non-planar component
- There exists a unique non-planar 2-connected component
- There exists a unique non-planar 3-connected component

[Chapuy-Fusy-Giménez-Mohar-M.N. 2011]

The following are definable in MSO:

- 3-connected components (Courcelle)
- Planarity: contractions to K_5 or $K_{3,3}$ are MSO expressible



$\mathbf{P}$(Unique non-planar 3-connected component has even size) $\to 1/2$

The proof uses a number of results

1. CFGMN 2011

   A random graph of genus $g > 0$ has w.h.p. a unique non-planar 3-connected component, which is MSO definable

2. Ellingham 1996

   A 3-connected graph of genus g has a spanning tree with maximum degree $\leq 4g$

3. Courcelle 2003

   For bounded genus MSO $\equiv$ MSO_2 (quantification over sets of vertices and edges)

4. Giménez-M.N.-Rué 2013

   Local limit law for $X_n = |\text{non-planar 3-connected component}|$

   Convergence to a stable law of parameter $3/2$

$\mathbf{P}(X_n \in \{0, 1, \dots, a-1\} \mod b) \to a/b$

$L = \{\lim \mathbf{P}(G_n \models \phi) : \phi \text{ MSO formula}\}$

$$\overline{L} = [0, 1]$$

Remark [HMNT]
For planar graphs L is a disjoint union of 108 intervals of length $\sim 10^{-6}$

Non-convergence for $g > 0$

We can produce an MSO formula ϕ such that
$\mathbf{P}(G_n \models \phi)$ does not converge for random graphs of genus $g > 0$

Claim The 3-connected component of genus g contains w.h.p. an MSO definable **grid** M with

$$\log \log n \leq |M| \leq n$$

Inspired on encoding Turing machine computations in a grid one can capture parity of the iterated logarithm $\log^* |M|$ and produce a formula without a limiting probability

For fixed $g \geq 0$ random graphs of genus g have the same global properties independently of g

$$\mathbf{P}(\text{being connected}) \sim 0.95$$
$$\mathbf{E}(\text{number of edges}) \sim 2.21n$$
$$\mathbf{E}(\text{size of largest 3-connected component}) \sim 0.73n$$

For planar planar the largest 3-connected component is indistinguishable in MSO from the other 3-connected components

For graphs of genus $g > 0$ the largest 3-connected component is non-planar, hence MSO definable (via contractions)

Anusch Taraz (Hamburg)

Tobias Müller (Groningen)

Albert Atserias (UPC)

Stephan Kreutzer (Berlin)

More non-convergence results

Second order (SO) logic: quantification over relations of any arity

Parity can be expressed as

$$\exists R(x,y) \text{ symmetric and antireflexive } \forall x \exists! y \ R(x,y)$$

Lemma

For every existential SO formula ϕ (all relation quantifiers are at the beginning of ϕ and are existential) there exists ϕ' in FO such that

$$\phi \text{ has a finite model} \iff \phi' \text{ has a finite model}$$

- Kuafmann-Shelah There exist non-convergent MSO properties for the class of all graphs (zero-one law in FO)
- Shelah-Spencer There exist non-convergent FO properties for $G(n, n^{-\alpha})$, when $\alpha \in (0,1)$ irrational
- Müller-M.N. (2018) There exist non-convergent FO properties for random perfect graphs (a dense class)
  Almost all prefect graphs are generalized split