Logic and random graphs

Marc Noy
Universitat Politècnica de Catalunya
Barcelona

Outline of the lectures

1. First order logic and Ehrenfeucht-Fraïssé games
2. Logical limit laws: planar graphs an related classes

Partly based on joint work with

- Peter Heinig, Anusch Taraz (Munich/Hamburg), Tobias Müller (Utrecht) J. Combin. Theory Ser. B
- Albert Atserias (Barcelona), Stephan Kreutzer (Berlin) in preparation

First order logic (FO)

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

For a given class of structures we add relations of any given arity
Graphs: $E(x, y)$ adjacency relation, written $x \sim y$

Some examples in graphs

- Existence of an isolated vertex: $\exists x, \forall y \ \neg(x \sim y)$
- Existence of a triangle: $\exists x, \exists y, \exists z \ (x \sim y) \wedge (y \sim z) \wedge (z \sim x)$
- Existence of vertices with given degrees. Existence of fixed H as a subgraph (or induced subgraph)
- Existence of a connected component is isomorphic to H
- Connectivity?

A preview of things to come

- Inexpressability in FO
  Graph connectivity cannot be expressed in FO logic

- The classical Zero-One Law
  R_n random labelled graph on n vertices: $\mathbf{P}(R_n = G) = \frac{1}{2^{\binom{n}{2}}}$
  For every graph property $\mathcal{P}$ expressible in FO logic

$$\lim_{n\to\infty} \mathbf{P}(R_n \text{ satisfies } \mathcal{P}) \in \{0, 1\}$$

  Almost every graphs satisfies $\mathcal{P}$ or almost no graph satisfies $\mathcal{P}$

Graph connectivity

A graph (V, E) is connected if

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

Graph connectivity

A graph (V, E) is connected if

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

Not in FO

But diameter $\leq k$ (for fixed k) is in FO

Another attempt at expressing connectivity

$$\forall A \subset V, A \neq \emptyset, A \neq V \quad \exists x \in A, \exists y \notin A\ (x \sim y)$$

This is a second order formula: quantification over sets

Theorem Graph connectivity is not expressible in FO

First proof idea: analyze each FO formula and show it cannot express connectivity

$$\forall x \exists y \forall z\ ((x \sim z) \wedge \neg(y \sim z)) \vee \exists w((z \sim w)\ \vee \neg(x \sim w))$$

Theorem Graph connectivity is not expressible in FO

First proof idea: analyze each FO formula and show it cannot express connectivity

$$\forall x \exists y \forall z \ ((x \sim z) \wedge \neg (y \sim z)) \vee \exists w ((z \sim w) \ \vee \neg (x \sim w))$$

Theorem (Trakhtenbrot)
Given a FO formula ϕ it is undecidable whether there exists some finite graph satisfying ϕ

Winning idea: analyze simultaneously all formulas of given depth

Depth of formula ϕ = maximum number of nested quantifiers in ϕ

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

Logical equivalence of graphs

$G \equiv_k H$ if G and H satisfy exactly the same formulas of depth $\leq k$

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

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

Winning idea: analyze simultaneously all formulas of given depth

Depth of formula ϕ = maximum number of nested quantifiers in ϕ

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

Logical equivalence of graphs

$G \equiv_k H$ if G and H satisfy exactly the same formulas of depth $\leq k$

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 $k = \text{depth}(\phi)$

Contradiction!

Logical types

$\equiv_k$ is an equivalence relation in graphs
The equivalence classes are called $\equiv_k$ types

Theorem
For each k the number of $\equiv_k$ types is finite

But large: $2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$

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 any graph) and Duplicator picks a vertex from the other graph

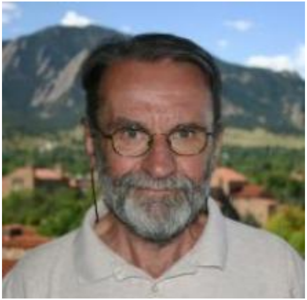
$(a_1, \ldots, a_k)$ vertices selected from G
$(b_1, \ldots, b_k)$ vertices selected from H

Duplicator wins if $(a_1, \ldots, a_k) \leftrightarrow (b_1, \ldots b_k)$ partial isomorphism (same adjacencies)

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

Provides a purely combinatorial characterization of FO logic

Roland Fraïssé (1920-2008)
left (with Abraham Robinson)



Andrzej Ehrenfeucht

Roland Fraïssé [Wikipedia] Professeur à l'université de Provence où il a formé toute une génération de logiciens

An example

Let $(L_n, \leq)$ be a linear order on n elements

Lemma If $n, m \geq 2^k$ then $L_n \equiv_k L_m$

$(a_1, \dots, a_i)$ and $(b_1, \dots, b_i)$ selections up to move i

Guarantee that for $j, \ell \leq i$

1. $d(a_j, a_\ell) < 2^{k-i} \implies d(b_j, b_\ell) = d(a_j, a_\ell)$
2. $d(a_j, a_\ell) \geq 2^{k-i} \implies d(b_j, b_\ell) \geq 2^{k-i}$
3. $a_j \leq a_\ell \iff b_j \leq b_\ell$

Assume spoiler plays a_{i+1} with $a_j < a_{i+1} < a_\ell$

Choose b_{i+1} depending on whether

- $d(a_j, a_\ell) < 2^{k-i}$
- $d(a_j, a_\ell) \geq 2^{k-i}$

Proofs of non-expressability in FO

Connectivity

$G = C_{3^k}, \quad H = C_{3^k} \cup C_{3^k}$

Claim: $G \equiv_k H$

Proof by induction on k as before

- Aciclicity
- 3-colorability
- Hamiltonicity
- Eulerian
- Planarity
- Rigidity (no non-trivial automorphism)

Exercises

Zero-one laws

$\mathcal{G}$ class of (labelled) graphs
$\mathcal{G}_n$ graphs in $\mathcal{G}$ with n vertices
Probability distribution on $\mathcal{G}_n$ for each n

The zero-one law holds in $\mathcal{G}$ if for every formula ϕ in FO

$$\lim_{n\to\infty} \mathbf{P}(G \models \phi : \ G \in \mathcal{G}_n) \in \{0, 1\}$$

Whp every object satisfies ϕ or whp no object satisfies ϕ

Property A holds in $\mathcal{G}$ with high probability (whp) if
$\lim_{n\to\infty} \mathbf{P}(G \text{ satisfies } A \colon \ G \in \mathcal{G}_n) = 1$

The classical example

$\mathcal{G}$ class of all labelled graphs $\quad |\mathcal{G}_n| = 2^{\binom{n}{2}}$

Uniform distribution $\quad \mathbf{P}(G) = \frac{1}{2^{\binom{n}{2}}}, \qquad G \in \mathcal{G}_n$

Theorem Glebski, Kogan, Liagonkii, Talanov (1969) Fagin (1976)
The zero-one law holds for labelled graphs

The $G(n,p)$ model

- Class: Labelled graph with n vertices
- Each possible edge xy independently with probability p

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

$G(n,1/2)$ is the uniform distribution

The extension property E_r:
For all disjoint $A, B \subset \{1, \ldots, n\}$ with $|A| = |B| = r$

$$\exists z \notin A \cup B \qquad (\forall x \in A \quad z \sim x) \quad \wedge \quad (\forall y \in B \quad z \not\sim y)$$

Lemma $G(n,p)$ satisfies E_r whp for constant p

$$\mathbf{P}(G_n \not\models E_r) \le \binom{n}{r}\binom{n-r}{r}(1-p^r(1-p)^r)^{n-2r} \to 0, \qquad \text{as } n \to \infty$$

Theorem The 0-1 law holds in $G(n, p)$ for constant p

Assume $(a_1, \ldots, a_i) \leftrightarrow (b_1, \ldots, b_i)$ and Spoiler plays a_{i+1}
Let

$$A_1 = \{a_j | a_{i+1} \sim a_j, 1 \leq j \leq i\}$$

$$A_2 = \{a_j | a_{i+1} \not\sim a_j, 1 \leq j \leq i\}$$

Then Duplicator plays $b_{i+1} = z$ as in E_r for the sets A_1 and A_2
Hence Duplicator wins whp

It follows that for each k two random graphs are $\equiv_k$ equivalent
Hence they satisfy exactly the same same formulas of depth k

For each ϕ, almost all graphs satisfy ϕ or satisfy $\neg\phi$

The 0-1 law does not hold in $G\left(n, p = \frac{1}{n}\right)$
$p = 1/n$ is the threshold for the appearance of a triangle

The number of triangles in $G(n, p = 1/n)$ tends to Poisson$(1/6)$
The probability of having a triangle tends to $1 - \exp(-1/6)$

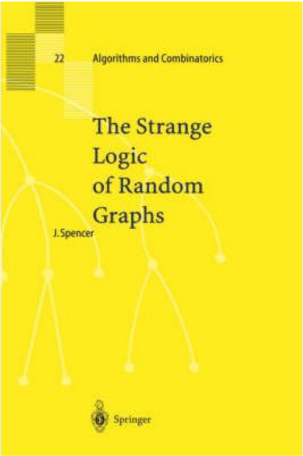
The threshold for the appearance of a balanced graph H is

$$n^{-v(H)/e(H)}$$

Shelah, Spencer 1988

- The 0-1 law holds in $G(n, p = n^{-\alpha})$ for $\alpha \in [0, 1]$ irrational
- For $\alpha \in [0, 1]$ rational there are non-convergent FO properties

Joel Spencer The strange logic of random graphs (Springer 2001)



Constrained classes of graphs

- H-free graphs
- d-regular graphs
- Trees
- Planar graphs

In all cases uniform distribution on labelled graphs with n vertices

The convergence law holds if $\mathcal{G}$ the limit

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

exists for each formula ϕ

Examples

- Triangle-free graphs

  Erdős, Kleitman, Rothschild (1976)
  Almost all triangle-free graphs are bipartite
  0-1 law as for $G(n,p)$ from extension axioms

- K_{t+1}-free graphs

  Kolaitis, Prömmel, Rothschild (1987)
  Almost all K_{t+1}-free are t-partite

- d-regular graphs
  - Lynch (2005)
    Convergence law for constant d using the configuration model
    Number of triangles $\to$ Poisson law
  - Haber, Krivelevich (2010)
    Zero-one law for $d \approx \delta n$ by comparison with $G(n,p)$

- Trees McColm (2002)

Random trees

$\mathcal{T}$ labelled trees $\quad |\mathcal{T}_n| = n^{n-2}$

Typical properties of a random tree

- Has $\sim e^{-1}n$ leaves
- Has αn pendant copies of any fixed tree

T has T' as a pendant copy if it has a rooted subtree isomorphic to T' joined to T by a single edge

Zero-one law for trees

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

Sketch of proof
Consider rooted trees for the game strategy (but the root is not part of the language)
$T_1, \ldots, T_m$ representatives of all $\equiv_k$ types of rooted trees
Construct a 'universal' tree U_k: take k copies of each T_i and glue them by identifying the roots

- A random tree contains a pendant copy of U_k w.h.p.
- If T, T' both contain a pendant copy of U_k then $T \equiv_k T'$

Duplicator wins $\mathrm{Ehr}_k(T, T')$ by playing in suitable subtrees of U_k

Hence T and T' satisfy the same formulas of depth $\leq k$ whp

What follows is joint work with

Tobias Müller

Peter Heinig

Anusch Taraz

- Extension to forests
- Extension to more general classes of graphs

Forests

There is no zero-one law in the class $\mathcal{F}$ of forests

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

Properties of random forests

- Is connected with probability $\to e^{-1/2} \approx 0.607$
- The largest component has expected size $n - O(1)$
- Fragment = complement of largest component
  H unlabelled forest, $\quad \mathbf{P}(\text{Fragment} \simeq H) \to \mu_H$

Theorem

A convergence law holds for forests

Sketch of proof

Type of the components determines type of the forest
Largest component has a.a. the same type (because of 0-1 law for trees).
Sum over fragments $\mathcal{A}(\phi)$ that make ϕ hold:

$$\lim_{n\to\infty} \mathbf{P}(F_n \models \phi) = \sum_{H \in \mathcal{A}(\phi)} \mu_H$$

Planar graphs

For each k there exists a planar graph U_k such that

- If G, G' planar contain a pedant copy of U_k then $G \equiv_k G'$
- W.h.p. a random planar graph contains a pendant copy of U_k
  McDiarmid, Steger, Welsh 2005 Giménez, N. 2009

Theorem

The zero-one law holds for connected planar graphs

The convergence law holds for arbitrary planar graphs

Minor-closed classes of graphs

H is a minor of G if it can be obtained from a subgraph of G by contracting edges

$\mathcal{G}$ is minor-closed if

$$G \in \mathcal{G}, \quad H \text{ minor of } G \Rightarrow H \in \mathcal{G}$$

Forests, Planar, Graphs embeddable in a fixed surface S
Outerplanar, Series-Parallel, Bounded tree-width

$\mathcal{G}$ addable if it is closed under disjoint unions and adding bridges between different components

Graphs on a fixed surface is **not** an addable class

Theorem (McDiarmid 2009)
$\mathcal{G}$ addable and minor-closed, H fixed graph in $\mathcal{G}$
A random graph in $\mathcal{G}$ contains a pendant copy of H w.h.p.

Theorem
The zero-one law holds for connected graphs in $\mathcal{G}$
The convergence law holds for arbitrary graphs in $\mathcal{G}$

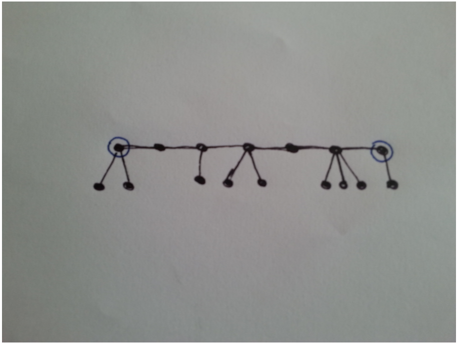
All these results hold in
Monadic Second Order (MSO) logic

MSO = FO + Quantification over sets of vertices

Connectivity

$$\forall A \subset V, A \neq \emptyset, A \neq V \quad \exists x \in A, \exists y \notin A \ (x \sim y)$$

No zero-one law for caterpillars (not addable)



$\mathbf{P}$(Endpoints of the spine of a caterpillar have given degrees)
$\to$ constant $\neq 0, 1$

The set of limiting probabilities

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

$L \subseteq [0,1]$ is countable and symmetric with respect to $1/2$

Theorem
If $\mathcal{G}$ addable minor-closed class
then $\overline{L}$ is a finite union of closed intervals

Forests

$$\overline{L} = [0, 0.1703] \cup [0.2231, 0.3935] \cup [0.6065, 0.7769] \cup [0.8297, 1]$$

$$0.6065\cdots = e^{-1/2} = \lim \mathbf{P}(\text{Random forest is connected})$$

ϕ a.s. true for trees $\Rightarrow \lim \mathbf{P}(\phi) \geq 0.6065$

ϕ a.s. false for trees $\Rightarrow \lim \mathbf{P}(\phi) \leq 1 - 0.6065 = 0.3935$

Lemma (Pólya)

$p_1 \geq p_2 \geq \cdots \geq p_n \cdots > 0$ and $\sum p_n < +\infty$

If $p_n \leq \sum_{k>n} p_k$ for $n \geq n_0$ then

$$\left\{ \sum_{i \in A} p_i \; : \; A \subset \mathbb{N} \right\}$$

is a finite union of closed intervals

In our case the p_i are the probabilities of the possible fragments

- Same $\overline{L}$ for FO and MSO
- At least two intervals since

$$\mathcal{G} \text{ addable} \implies \lim \mathbf{P}(\text{connectivity}) \geq e^{-1/2} \approx 0.06065$$

Conjecture (McDiarmid, Steger, Welsh) proved by Addario-Berry, McDiarmid, Reed (2012) and by Kang, Panagiotou (2013)
In a stronger form by Chapuy, Perarnau (2015)

Lemma (Pólya)

$p_1 \geq p_2 \geq \cdots \geq p_n \cdots > 0$ and $\sum p_n < +\infty$

If $p_n \leq \sum_{k>n} p_k$ for $n \geq n_0$ then

$$\left\{ \sum_{i \in A} p_i \, : \; A \subset \mathbb{N} \right\}$$

is a finite union of closed intervals

In our case the p_i are the probabilities of the possible fragments

- Same $\overline{L}$ for FO and MSO
- At least two intervals since

$$\mathcal{G} \text{ addable } \implies \lim \mathbf{P}(\text{connectivity}) \geq e^{-1/2} \approx 0.06065$$

Conjecture (McDiarmid, Steger, Welsh) proved by Addario-Berry, McDiarmid, Reed (2012) and by Kang, Panagiotou (2013)
In a stronger form by Chapuy, Perarnau (2015)

For planar graphs $\overline{L}$ = union of 108 intervals of length $\approx 10^{-6}$

Graphs on surfaces

$\mathcal{G}_S$ class of graphs embeddable in S
Minor-closed but not addable: K_5 embeds in the torus not $K_5 \cup K_5$

$$B(x,r) = \{y : d(x,y) \leq r\}$$

A random graph in $\mathcal{G}_S$ satisfies w.h.p.

- All balls $B(x,R)$ are planar for fixed $R > 0$
  Chapuy-Fusy-Giménez-Mohar-N., Bender-Gao 2011
- Contains a pendant copy of any fixed connected planar graph
  McDiarmid 2008 CFGMN

Gaifman's locality theorem

Every FO formula is equivalent to a Boolean combination of basic local sentences of the form

$$\exists x_1 \cdots \exists x_s \quad \left(\bigwedge_{i \neq j} d(x_i, x_j) > 2r \right) \wedge \left(\bigwedge \psi^{\mathrm{Ball}_r(x_i)}(x_i) \right)$$

Theorem

A zero-one FO law holds for connected graphs in $\mathcal{G}_S$

A convergence FO law holds for arbitrary graphs in $\mathcal{G}_S$

$$p(\phi) = \lim \mathbf{P}(G_n \models \phi) \text{ independent of } S$$

Gaifman's locality theorem

Every FO formula is equivalent to a Boolean combination of basic local sentences of the form

$$\exists x_1 \cdots \exists x_s \quad \left(\bigwedge_{i \neq j} d(x_i, x_j) > 2r \right) \wedge \left(\bigwedge \psi^{\mathrm{Ball}_r(x_i)}(x_i) \right)$$

Theorem

A zero-one FO law holds for connected graphs in $\mathcal{G}_S$
A convergence FO law holds for arbitrary graphs in $\mathcal{G}_S$

$$p(\phi) = \lim \mathbf{P}(G_n \models \phi) \text{ independent of } S$$

We conjectured the same results hold in Monadic Second Order logic

What follows is joint work with

Albert Atserias

Stephan Kreutzer

Our results

- No Zero-One MSO law for connected graphs of genus $g > 0$
- No convergence MSO law for graphs of genus $g > 0$

Proofs use several facts

1. CFGMN 2011

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

   - 3-connected components are MSO definable
   - Minors are MSO definable, hence planarity too

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 vertices and edges)

4. Giménez-Noy-Rué 2013

   Local limit law for $X_n = |\text{3-connected component of genus } g|$

$$\mathbf{P}(X_n = \alpha n + x n^{2/3}) \sim n^{-2/3} f(x)$$

$f(x)$ density of an Airy distribution

Theorem

The probability that X_n is even is MSO expressible and

$$\mathbf{P}(X_n \text{ even}) \to 1/2$$

Sketch of proof

Because of spanning tree of bounded degree, parity is MSO expressible

Because of local limit law for X_n, $\mathbf{P}(X_n \text{ even}) \to 1/2$

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

Hence every rational number in $[0,1]$ is the limiting probability of some MSO formula

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

HMNT For planar graphs $\overline{L}$ is a finite union of disjoint intervals

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 large grid M

$$|M| \geq \log \log n$$

We use the fact that the unique non-planar 3-connected component has face-width $\Omega(\log n)$

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

For fixed $g \geq 0$ random graphs of genus g share many 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 fixed $g \geq 0$ random graphs of genus g share many 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 minors)

Non convergence typically comes from structures where one can capture parity of some substructure

Theorem Tobias Müller, MN

There exist non-convergent FO formulas in the class of perfect graphs