First Order Logic of Sparse Random Hyper-Graphs

Lázaro Alberto Larrauri Borroto

July 8, 2019

Supervisor: Marc Noy Serrano

1. First Order Logic of Sparse Graphs

2. General relational structures

Contents

1 First Order Logic of Sparse Graphs

2 General relational structures

The first order language of graphs.

- Variables $x_1, \dots, x_n, \dots$
- Connectives $\wedge, \vee$, equality symbol $=$, and negation symbol $\neg$.
- Quantifiers $\forall, \exists$.
- A binary relation symbol R

- Vertices.
- "And", "or", "equals", "not".
- "For all", "there exists".
- Edges $x \sim y$.

$$\forall x_1, x_2\, R(x_1, x_2) \implies \exists x_3 (\neg(x_3 = x_1) \wedge \neg(x_3 = x_2) \wedge R(x_1, x_3))$$

The binomial model

The binomial model of random graphs $G(n,p)$ is a discrete probability space where we assign to each graph $G = ([n], E)$ the probability

$$\Pr(G) = p^{|E|} \cdot (1-p)^{\binom{n}{2} - |E|}.$$

Lynch's theorem

Theorem (Lynch, 1992)

Let φ be a sentence in the F.O. language of graphs. Then the map $\mathrm{F}_\varphi : [0,\infty) \to \mathbb{R}$ given by

$$\mathrm{F}_\varphi(\beta) = \lim_{n\to\infty} \mathrm{Pr}(\, G(n,\beta/n) \models \varphi \,)$$

is well defined and admits an analytic extension to $\mathbb{C}$.

Overview of the proof

Some properties of $G(n, \beta/n)$:

- The number of cycles of length $3, 4 \ldots, r$ are asymptotically distributed like independent Poisson variables.
- Small cycles are a.a.s far away.
- Fixed vertices are a.a.s far away.
- The ball of a given radius centered in fixed vertex is a.a.s a tree. Any tree occurs with a positive probability.

Overview of the proof

For each fixed quantifier rank k:

(1) It is given a finite classification of "small" uni-cycles.

(2) It is shown that the rank k type of random graph G in $G(n, \beta/n)$ a.a.s depends exclusively on the number of "small" uni-cycles belonging to each class.

(3) The asymptotic distribution of those quantities is obtained.

Contents

1. First Order Logic of Sparse Graphs

2. General relational structures

Edge sets

Definition

Fix

- natural numbers $n, a \in \mathbb{N}$, with $a \geq 2$
- a subgroup of the symmetric group $\Phi \leq S_a$
- and a subset of pairs $A \subset ([a] \times [a]) \setminus \Delta_a$.

The **total edge set** $\mathcal{H}_{(a,\Phi,A)}(n)$ of size a, symmetry group Φ and restrictions A, on n elements is the set

$$\mathcal{H}_{(a,\Phi,A)}(n) = ([n]^a/\Phi) \ \setminus R,$$

where

$$R = \{ [x_1, \dots, x_a] \in [n]^a/\Phi \mid x_i = x_j \text{ for some } (i,j) \in A\}$$

Graphs

Definition

An (hyper)-graph $([n], H_1, \ldots, H_c)$ with edge colors $1, \ldots, c$, sizes $a_1, \ldots, a_c$, symmetry groups $\Phi_1, \ldots, \Phi_c$ and restrictions $A_1, \ldots, A_c$ consists of

- The vertex set $[n]$ for some natural number n.
- For $i = 1, \ldots, c$, a "colored" edge set $H_i \subseteq \mathcal{H}_{(a_i, \Phi_i, A_i)}(n)$ whose elements have color i.

The first order language

Consider the first order purely relational language $\mathcal{L}$ whose signature consists of the relation symbols $R_1, \dots, R_c$ with arities $a_1, \dots, a_c$.

A graph $G = ([n], H_1, \dots, H_c)$ is a $\mathcal{L}$-structure in the following way:

- The universe of G is its vertex set, $[n]$.
- For each $1 \leq i \leq c$,

$$(x_1, \dots, x_{a_i}) \in R_i^G \iff [x_1, \dots, x_{a_i}] \in H_i.$$

The first order language

By definition, a graph $G = ([n], H_1, \ldots, H_n)$ satisfies, for each $1 \leq i \leq c$:

- Symmetry formulas:

$$S_g := \big(R_i(x_1 \ldots, x_{a_i}) \iff R_i(x_{g(1)} \ldots, x_{g(a_i)})\big),$$

  where g is an element from Φ_i.

- Anti-reflexivity formulas:

$$A_{i,(j,l)} := \big(R_i(x_1 \ldots, x_{a_i}) \implies \neg(x_j = x_l)\big),$$

  where $(j, l) \in A_i$.

The random model

The random model $HG(n, p_1, \ldots, p_c)$ is a discrete probability space where for each graph $G = ([n], H_1, \ldots, H_c)$,

$$\Pr(G) = \prod_{i=1}^{c} p_i^{|H_i|} \cdot (1 - p_i)^{|\mathcal{H}_{(a_i, \Phi_i, A_i)}(n)| - |H_i|}.$$

We consider the case where for each $1 \leq i \leq c$, $p_i(n) = \beta_i / n^{a_i - 1}$.

Let $\beta = (\beta_1, \ldots, \beta_c) \in [0, \infty)^c$. We abbreviate

$$HG(n, p(n, \beta)) := HG(n, \beta_1 / n^{a_1 - 1}, \ldots, \beta_c / n^{a_c - 1}).$$

The theorem

We want to prove the following

Theorem

Let φ be a first order sentence in $\mathcal{L}$. Then the map $\mathrm{F} : [0, \infty)^c \to \mathbb{R}$ given by

$$\mathrm{F}(\beta_1, \ldots, \beta_c) = \lim_{n \to \infty} \Pr(HG(n, p(n, \beta)) \models \varphi)$$

is well defined and admits an analytic extension to $\mathbb{C}^c$.

Distance and paths

Given any graph G, we define the following distance over its vertex-set:

$$d(x,y) = \min_{\substack{H \leq G \\ H \text{ connected} \\ x,y \in V(H)}} \big(|V(H)| - 1\big), \quad \text{or } \infty \text{ if } x,y \text{ are not connected.}$$

Definition

A path between two vertices x, y in a graph G is a connected subgraph $H \leq G$ containing both x, y whose number of vertices is minimum.

Likelihood, trees, cycles and clusters

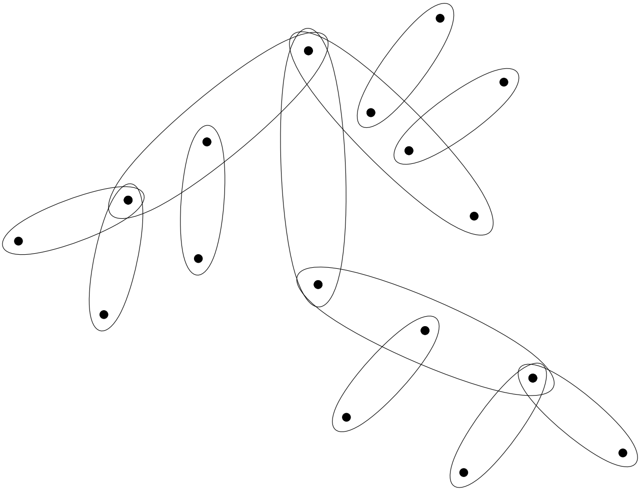
Definition

The likelihood $L(G)$ of a graph $G = (V, H_1, \ldots, H_c)$ is the number

$$|V(G)| - \sum_{i=1}^{c} |H_i|(a_i - 1).$$

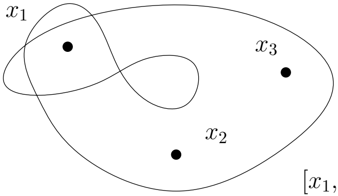
Likelihood, trees, cycles and clusters

- A tree is a connected graph with likelihood 1.

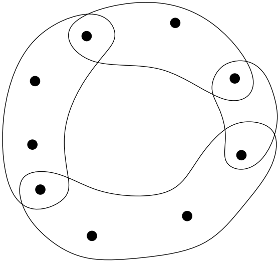


Likelihood, trees, cycles and clusters

- An unicycle is a connected graph with likelihood 0. A cycle is a minimal unicycle.

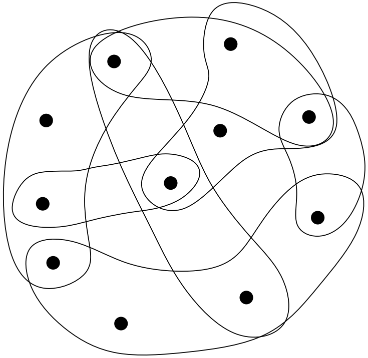


$[x_1, x_2, x_3, x_1]$



Likelihood, trees, cycles and clusters

- A cluster is a graph G with $L(G) \leq 0$ such that $L(H) > L(G)$ for any subgraph $H \leq G$.



The k-morphism relation over trees.

A rooted (T, x) is a tree T together with a distinguished vertex $x \in V(T)$.

$$Tree(y, T) = (T[X], y),$$
$$\text{where } X = \{z \in V(T) \mid d(x,z) = d(x,y) + d(y,z)\}.$$

The root of an edge $e \in H(T)$ in a rooted tree (T, x) is the vertex $y \in e$ such that $d(x, y) = d(x, e)$.

The radius of an edge $e \in H(T)$ is

$$\max_{\substack{y \in e \\ y \text{ not the root of } e}} r(Tree(y, T)).$$

The k-morphism relation over trees.

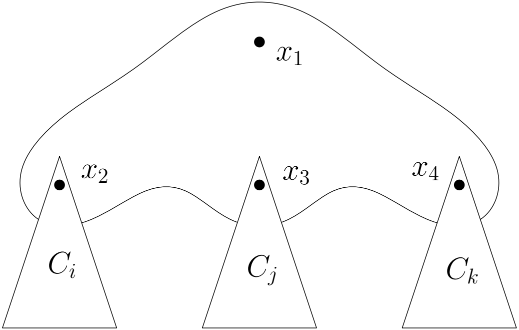
Fix $k \in \mathbb{N}$.
We define the k-morphism relation over rooted trees of the same radius inductively as follows:

- If $r(T_1) = r(T_2) = 0$ then $T_1 \overset{k}{\simeq} T_2$.

The k-morphism relation over trees.

For $r > 0$:

- First we define the type of a en edge with radius less than r:



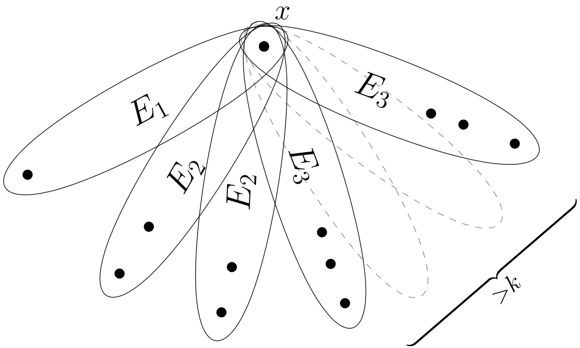
$[x_2, x_1, x_3, x_4]$
colored c

$\downarrow$

$[C_i, \tau, C_j, C_k]$
colored c

The k-morphism relation over trees.

- If $r(T_1) = r(T_2) = r$ we say that $T_1 \overset{K}{\simeq} T_2$ if for any edge k-type E of radius less than r either:
  - the number of initial edges in T_1 and T_2 of k-type E is the same, or
  - both T_1 and T_2 contain no less than $k+1$ initial edges of k-type E.



The k-morphism relation for graphs

To make the definitions suitable for proofs using E.F. games we need to work with "graphs with constants".

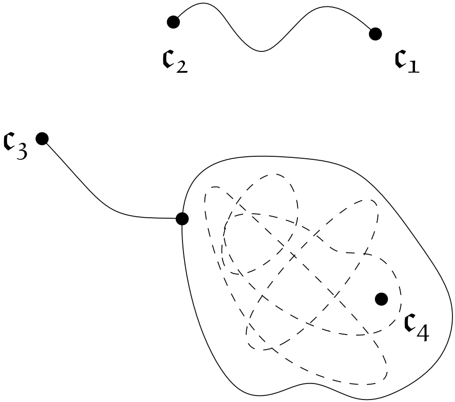
Definition

The center of a connected graph $G(\mathfrak{c}_1, \cdots, \mathfrak{c}_m)$ is its minimal connected subgraph containing all the constants and clusters.

If $G(\mathfrak{c}_1, \cdots, \mathfrak{c}_m)$ is not connected its center is the union of the centers of its connected components.

The k-morphism relation for graphs

A sketch of a center:



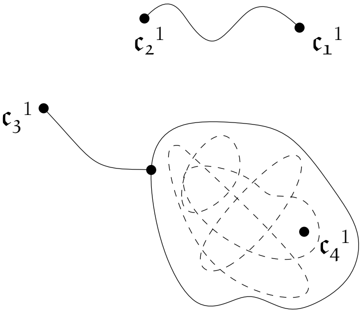
The k-morphism relation for graphs

Definition

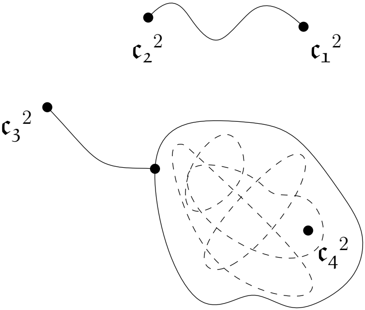
We say that $G_1(\mathfrak{c}_1, \cdots, \mathfrak{c}_m) \overset{k}{\simeq} G_2(\mathfrak{c}_1, \cdots, \mathfrak{c}_m)$, if there is an isomorphism $f : Center(G_1) \to Center(G_2)$ s.t.

- $f(c_i^1) = c_i^2$ for all constants, and
- $Tree(x, G_1) \overset{k}{\simeq} Tree(f(x), G_2)$ for all $x \in V(Center(G_1))$.

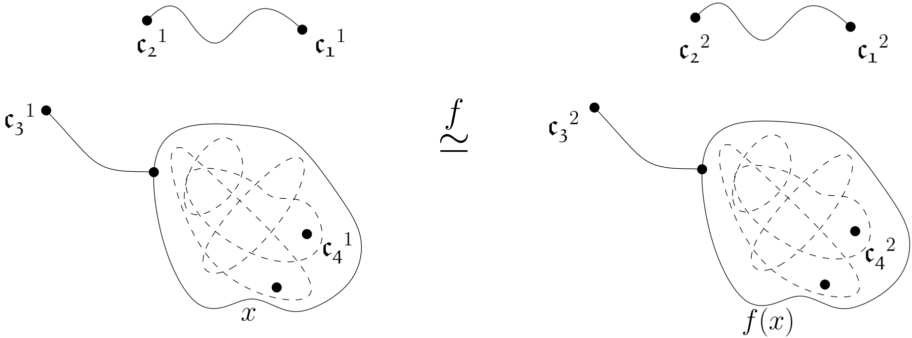
The k-morphism relation for graphs



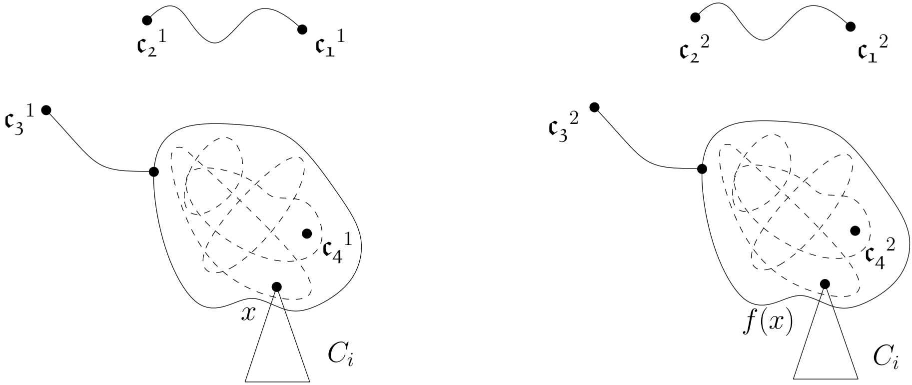
$\overset{f}{\simeq}$



The k-morphism relation for graphs



The k-morphism relation for graphs



The k-morphism relation for graphs

Connected k-morphic graphs G, H with $L(G), L(H) \leq 0$ have the same first order rank k-type.

The tool for proving this are Ehrenfeucht Fraisse games.

The k-morphism relation gives, for any r, a finite classification of the cycles of diameter at most r with trees of radii at most r hanging from their vertices.