

# Contents

<b>1</b>	<b>Preliminaries</b>	<b>4</b>
1.1	Models of Random Graphs . . . . .	4
1.2	First Order Logic . . . . .	4
1.3	Ehrenfeucht Fraisse Games and the Logic of Random Graphs . . .	4
<b>2</b>	<b>Probabilities of Sentences about Very Sparse Random Graphs</b>	<b>5</b>
2.1	Agreability Classes . . . . .	7

# Introduction

# Notation

# Chapter 1

## Preliminaries

### 1.1 Models of Random Graphs

### 1.2 First Order Logic

### 1.3 Ehrenfeucht Fraisse Games and the Logic of Random Graphs

## Chapter 2

# Probabilities of Sentences about Very Sparse Random Graphs

In this chapter we will review the results obtained in the paper with the same name by James F. Lynch [1]. In there, limit probabilities of sentences in the first order language of graphs  $\mathcal{L}$  are discussed for the binomial model  $G(n, p)$  in the cases  $p = \beta/n$  and  $p = \beta n^{-\alpha}$  with  $\alpha = (l + 1)/l$ .

More precisely, it is proven that in those cases the probability of every sentence converges and it is shown that for any of those sentences, its limit probability is among the values taken by some analytic formulas with parameter  $\beta$ .

We are interested in the case  $p = \frac{\beta}{n}$ , which is the one discussed more extensively in [1]. According to the author, the relevant theorems for the other case can be proven analogously. From now on we will only refer as random graphs to the ones in  $G(n, \beta/n)$

From now on we will denote by  $Poi_\lambda$  the probability function of the Poisson distribution with mean  $\lambda$ . That is, the one given by  $Poi_\lambda(n) = e^{-\lambda} \lambda^n / n!$  for any  $n \in \mathbb{N}$ . Also, we define  $Poi_\lambda(\leq n)$  and  $Poi_\lambda(> n)$  as  $\sum_{i=0}^n Poi_\lambda(n)$  and  $1 - Poi_\lambda(\leq n)$  respectively. Notice that for a fixed  $n$ , both  $Poi_\lambda(\leq n)$  and  $Poi_\lambda(> n)$  can be considered real functions of parameter  $\lambda$ .

We define the following sets of functions. Let  $\Lambda$  be the smallest set of expressions with parameter  $\beta$  such that:

- $1 \in \Lambda$ ,
- For any  $\lambda \in \Lambda$  and  $i \in \mathbb{N}$ , both  $Poi_{\beta\lambda}(n)$  and  $Poi_{\beta\lambda}(> n)$  are in  $\Lambda$ .
- For any  $\lambda_1, \lambda_2 \in \Lambda$ ,  $\lambda_1 \lambda_2$  belongs to  $\Lambda$  as well.

And let  $\Theta$  be the smallest set of functions with parameter  $\beta$  such that:

- For any  $\lambda \in \Lambda$  and  $n, a, i \in \mathbb{N}$ , with  $i \geq 3$ , both  $Poi_{\beta^i \lambda/a}(\leq n)$  and  $Poi_{\beta^i \lambda/a}(> n)$  are in  $\Theta$ .

The main result is the following:

**Theorem 2.1** (Lynch, 1992). *Let  $\phi$  be a sentence in the first order theory of graphs. Then the limit  $\lim_{n \rightarrow \infty} P(G(n, \beta/n) \models \phi)$  exists for all positive real numbers  $\beta$ , and it is a finite sum of expressions in  $\Theta$ .*

We show now an outline of the proof.

We show that for any quantifier rank  $k$  there are some classes of graphs  $C_1^k, \dots, C_{n_k}^k$  such that

- (1) a.a.s the rank  $k$  type of any two graphs in the same class coincide,
- (2) a.a.s. any random graph belongs to some of them, and
- (3) the limit probability of random graph belonging to any of them is an expression in  $\Theta$ .

After this is archived the theorem follows easily. Indeed, let  $\phi$  be a sentence in the first order language  $\mathcal{L}$  of graphs whose quantifier rank is  $k$ , and denote by  $G$  a random graph in  $G(n, \beta/n)$ . We define the events  $E_1, \dots, E_{n_k}$  as

$$E_i := (G \models \phi) \wedge (G \in C_i),$$

and the event  $F$  as

$$F := (G \models \phi) \bigwedge_{i=1}^{n_k} (G \notin C_i).$$

Then, for any  $n \in \mathbb{N}$

$$P(G \models \phi) = \sum_{i=1}^{n_k} P(E_i) + P(F), \quad (2.1)$$

as the events  $E_i$  together with  $F$  form a partition of all the cases where  $G$  satisfies  $\phi$ .

Fix and index  $i \in \{1, \dots, n_k\}$ . From the property (1) of the classes  $C_1, \dots, C_{n_k}$  it follows that if  $G$  and  $H$  are random graphs, then

$$\lim_{n \rightarrow \infty} P((G \models \phi) \wedge \neg(H \models \phi) \mid G \in C_i \wedge H \in C_i) = 0.$$

This is because  $G$  and  $H$  share a.a.s the same rank  $k$  type if they both belong to  $C_i$ . In consequence the limit

$$\lim_{n \rightarrow \infty} P(G \models \phi \mid G \in C_i)$$

is either zero or one, and

$$\lim_{n \rightarrow \infty} P(E_i) = \lim_{n \rightarrow \infty} P(G \in C_i) \cdot P(G \models \phi \mid G \in C_i) = \text{either } 0 \text{ or } \lim_{n \rightarrow \infty} P(G \in C_i). \quad (2.2)$$

Also, as a consequence of property (2) we obtain

$$\lim_{n \rightarrow \infty} P\left(\bigwedge_{i=1}^{n_k} G \notin C_i\right) = 0,$$

so

$$\lim_{n \rightarrow \infty} P(F) = \lim_{n \rightarrow \infty} P\left(\bigwedge_{i=1}^{n_k} G \notin C_i\right) \cdot P(G \models \phi \mid \bigwedge_{i=1}^{n_k} G \notin C_i) = 0. \quad (2.3)$$

Taking limits in equation 2.1 and using equations 2.2 and 2.3 we get

$$\lim_{n \rightarrow \infty} P(G \models \phi) = \sum_{C_i \in \mathcal{C}} \lim_{n \rightarrow \infty} P(G \in C_i),$$

where  $\mathcal{C}$  is a (possibly empty) subset of  $\{C_1, \dots, C_{n_k}\}$ . Finally, because of property (3) for each  $i$  the limit  $\lim_{n \rightarrow \infty} P(G \in C_i)$  is an expression in  $\Theta$ . Thus  $\lim_{n \rightarrow \infty} P(G \models \phi)$  is a finite sum of expressions in  $\Theta$  and the theorem follows.

The objective of next sections will be to define the classes  $C_1, \dots, C_{n_k}$  and to show that they satisfy properties (1), (2) and (3). Later we will prove a stronger result, so we will allow ourselves to just sketch some of the proofs during this chapter.

## 2.1 Agreability Classes

It is known that  $n^{-v/e}$  is the threshold probability for the appearance of a balanced graph of density  $v/e$ . In our case  $v/e = 1$ , so in consequence any connected graph  $H$  with  $e(H) < v(H)$  a.a.s will not appear as a subgraphs of  $G(n, \beta/n)$ . It can be easily shown that such graphs  $H$  are exactly the ones containing more than one cycle.

If  $H$  is a connected graph with  $v = e$ , then  $H$  is an uni-cyclic graph. In this case, the number  $X_H$  of copies of  $H$  in  $G(n, \beta/n)$  will asymptotically have non-zero bounded expectancy  $m$ . It does not take much work to prove, using Brun's sieve, that  $X_H$  converges in distribution to a Poisson random variable with mean  $m$  as  $n$  goes to infinity.

Finally, if  $H$  is a connected graph with  $v > e$  then it must be a tree. Here the expected number of copies of  $H$  in  $G(n, \beta/n)$  diverges asymptotically. Informally, trees of any kind will occur arbitrarily often.

This all means, in a sense, that a.a.s the only difference between large graphs in  $G(n, \beta/n)$  lies in their uni-cyclic subgraphs. More precisely, because of the “locality” of first order logic of quantifier rank  $k$  we will only be interested in the “small” neighborhoods of the “short” cycles. Thus, our goal will be to classify uni-cyclic graphs in a way that respects equivalence under first order logic of quantifier rank  $k$ .

To make our classification suitable for proofs involving E.F. games we need to work graphs to which we “attach” labels. We define the set of symbols  $Const = \{c_i\}_{i \in \mathbb{N}}$  as the set of constants. Also, we will denote by  $Const_n$  the set  $\{c_1, \dots, c_n\}$ .

**Definition 2.1.** A **co-labeling** of a graph  $G = (V, E)$  is a map  $\sigma : D \rightarrow V$ , where  $D \subset C$  is a finite set of constant symbols. Given  $c_i \in D$ , we will say that the vertex  $\sigma(c_i)$  is labeled  $c_i$ . Equivalently, we can denote a labeling  $\sigma$  as a tuple  $(c_{i_1}[x_1], \dots, c_{i_m}[x_m])$  where each  $c_{i_j}$  is a constant symbol, and  $x_j$  is the vertex in  $V$  labeled  $c_{i_j}$ .

**Definition 2.2.** A **co-labeled graph**<sup>1</sup>  $G = (V, E, c_{i_1}[x_1], \dots, c_{i_m}[x_m])$  is a graph  $(V, E)$  together with a labeling  $(c_{i_1}[x_1], \dots, c_{i_m}[x_m])$ .

To keep our notation compact we will often drop the  $x_i$ 's and say  $G = (V, E, c_{i_1}, \dots, c_{i_m})$ .

**Definition 2.3.** Let  $G$  be a co-labeled graph. A co-labeled subgraph  $H$  of  $G$  is a co-labeled graph such that  $V(H) \subseteq V(G)$ ,  $E(H) \subseteq E(G)$  and all vertices in  $V(H)$  have the same labels in  $H$  and  $G$ .

An important abuse of notation we are going to make will be to identify the constants  $c_i$  with their labeled vertices  $\sigma(c_i)$ . This way things like  $c_i \sim c_j$  will make sense. In this context, notice that the expression  $c_i = c_j$  is ambiguous because the vertices labeled  $c_i$  and  $c_j$  may be the same for some  $i \neq j$ , but the constant symbols  $c_i$  and  $c_j$  will be equal only if  $i = j$ . We will make sure to leave no room for ambiguity in this situations.

**Proposition / Definition 2.1.** Let  $G = (V, E, c_{i_1}, \dots, c_{i_m})$  be a connected co-labeled graph. Then it has a unique minimal connected co-labeled subgraph  $H$  containing all its constants and cycles. We will call the **center** of  $G$  to such subgraph and denote it by  $Center(G)$ . If  $\bar{G}$  is an arbitrary co-labeled graph, then its center  $Center\bar{G}$  will be the union of the centers of its connected components.

*Proof.* TO DO □

For an arbitrary co-labeled graph we define the metric  $d(\cdot, \cdot)$  on  $V(G)$  as the one such that  $d(x, y)$  is the minimum length of a path connecting  $x$  and  $y$  in  $G$  or  $\infty$  if such path does not exist. For any vertex  $x \in V(G)$  and  $r \in \mathbb{N}$  we define the co-labeled subgraph  $N(x; r)$  as the ball of radius  $r$  centered at  $x$ . That is, the induced subgraph with vertex set

$$V(N(x; r)) = \{ y \in V \mid d(x, y) \leq r \}.$$

In a similar vein, given  $X \subseteq V(G)$  we define its neighborhood of radius  $r$  as the induced co-labeled subgraph  $N(X; r)$  whose vertex set is

$$V(N(X; r)) = \{ y \in V \mid \forall x \in X : d(x, y) \leq r \}.$$

Let  $G = (V, E)$ , and  $V' \subseteq V$ . Another important abuse of notation we will make is writing  $H = (V', E)$  for a subgraph  $H$  to mean that the edge set of  $E(H)$  is the one induced by  $E(G)$  on  $V'$ .

---

<sup>1</sup> Compare with [1], where they are called “rooted graphs”.



**Definition 2.4.** A **rooted tree**  $T = (V, E, x)$  is a tree  $(V, E)$  with distinguished vertex  $x \in V$  with we will call **root** of the tree.

**Proposition / Definition 2.2.** Let  $G = (V, E, c_{i_1}, \dots, c_{i_m})$  be a connected graph and  $x \in V$ . Then define  $Tree(x, G)$  as the rooted tree

$$Tree(x, G) = (V_x, E, x),$$

where

$$V_x = \{y \in V \mid d(Center(G), y) = d(Center(G), x) + d(x, y)\}.$$

*Proof.* TO DO □

The radius  $r(T)$  of a rooted tree  $T = (V, E, x)$  is the maximum distance between its root  $x$  and any other of its vertices. The branches of  $T$  are the rooted trees of the form  $Tree(y, T)$ , where  $y \sim x$ . We will denote by  $Br(T)$  the set of branches of  $T$ .

We begin by defining an equivalence relation between rooted trees for each quantifier rank  $k$ .

**Definition 2.5.** Let  $k \in \mathbb{N}$  with  $k \geq 1$ . The **k-morphism** equivalence relation  $\overset{k}{\simeq}$  between co-labeled graphs is the one inductively defined as follows:

- If  $T_1, T_2$  are rooted trees of radius 0 -i.e., they consist only of their roots- they are  $k$ -morphic.
- Let  $T_1, T_2$  be rooted trees of radius  $r$  whose roots have the same label. Then  $T_1 \overset{k}{\simeq} T_2$  if for any  $k$ -morphism class  $C$  of trees with radius less than  $r$  and root either

“ $T_1$  and  $T_2$  have the same number of branches of type  $C$ ”

$$|Br(T_1) \cap C| = |Br(T_2) \cap C|,$$

or

“ $T_1$  and  $T_2$  have both more than  $k$  branches of type  $C$ ”

$$|Br(T_i) \cap C| \geq k + 1 \text{ for } i = 1, 2.$$

It follows from the definition that  $k$ -morphic trees have the same radius. It is also easy to check that the  $k$ -morphism relation is indeed an equivalence one.

**Proposition 2.1.** For all  $k, r \in \mathbb{N}$  and with  $k \geq 1$ , the set of classes of  $k$ -morphic trees with radius lesser or equal than  $r$  is finite.

*Proof.* TO DO □

We define now the  $k$ -morphism relation for arbitrary co-labeled graphs.

**Definition 2.6.** Let  $G^1 = (V^1, E^1, c_{i_1}[x_1^1], \dots, c_{i_m}[x_m^1])$  and  $G^2 = (V^2, E^2, c_{i_1}[x_1^2], \dots, c_{i_m}[x_m^2])$  be co-labeled graphs with the same constant symbols. We will say that they are  $k$ -morphic (denoted by  $G^1 \stackrel{k}{\simeq} G^2$ ) if there is a bijection  $f : V(\text{Center}(G^1)) \rightarrow V(\text{Center}(G^2))$  such that

- “ $f$  preserves edges”

$$\forall x, y \in V(\text{Center}(G^1)) : \quad x \sim y \iff f(x) \sim f(y).$$

- “ $f$  preserves labels”

$$\forall j \in \{1, \dots, m\} : \quad f(x_j^1) = x_j^2.$$

- “ $f$  preserves  $k$ -morphism classes of trees”

$$\forall x \in V(\text{Center}(G^1)) : \quad \text{Tree}(x, G^1) \stackrel{k}{\simeq} \text{Tree}(f(x), G^2).$$

In this case we will say that  $G^1 \stackrel{k}{\simeq} G^2$  via  $f$ .

We are going to show that the rank  $k$  type of a random graph a.a.s only depends on the neighborhoods of its small cycles. In consequence the following definition is motivated:

**Definition 2.7.** Let  $G = (V, E, c_{i_1}, \dots, c_{i_m})$  be a co-labeled graph. Then its core of radius  $r$ ,  $\text{Core}(G, r)$  is the co-labeled subgraph  $N(X; r)$ , where  $X$  is the union of the (vertex sets of the) cycles in  $G$  with size at most  $2r + 1$  and all of the labeled vertices in  $G$ .

# Bibliography

- [1] James F Lynch. Probabilities of sentences about very sparse random graphs.  
*Random Structures & Algorithms*, 3(1):33–53, 1992.