The first order convergence law fails for random perfect graphs

Tobias Müller

Groningen University

Marc Noy

Universitat Politècnica de Catalunya
Barcelona Graduate School of Mathematics

Abstract

We consider first order expressible properties of random perfect graphs. That is, we pick a graph G_n uniformly at random from all (labelled) perfect graphs on n vertices and consider the probability that it satisfies some graph property that can be expressed in the first order language of graphs. We show that there exists such a first order expressible property for which the probability that G_n satisfies it does not converge as $n \to \infty$.

Keywords: Perfect graph, First Order Logic, Random graphs

1 Introduction

A graph is perfect if the chromatic number equals the clique number in each of its induced subgraphs. Perfect graphs are a central topic in graph theory and play an important role in combinatorial optimization. In this paper we will study the random graph chosen uniformly at random from all (labelled) perfect graphs on n vertices. The first thing one might want in order to prove results about this object is a mechanism for generating random perfect graphs that is more descriptive than "put all n-vertex perfect graphs in a bag and pick one uniformly at random". Such a mechanism has been introduced recently by McDiarmid and Yolov [6]. Before presenting it, let us discuss as a preparation a simpler subclass of perfect graphs.

A graph is chordal if every cycle of length four or more has a chord, that is, an edge joining non-consecutive vertices in the cycle. A graph is split if its vertex set can be partitioned into a clique and an independent set (with arbitrary edges across the partition). It is easy to see that a split graph is chordal, but not conversely. On the other hand, it is known that almost all chordal graphs are split [1], in the sense that the proportion of chordal graphs that are split tends to 1 as the number of vertices n tends to infinity. Thus we arrive at a very simple process for generating random chordal graphs: (randomly) partition the vertex set into a clique A and an independent set B, and add an arbitrary set of edges between A and B (chosen uniformly at random from all posible sets of edges between A and B). The distribution we obtain in this way is not uniform as a split graph may arise from different partitions into a clique and an independent set, but it can be seen that when the size of A is suitably sampled then its total variational distance to the uniform distribution tends to zero as n tends to infinity.

Now we turn to random perfect graphs. A graph G is *unipolar* if for some $k \geq 0$ its vertex set $V(G)$ can be partitioned into $k+1$ cliques $C_0, C_1, \ldots, C_k$, so that there are no edges between C_i and C_j for $1 \leq i < j \leq k$. A graph G is *co-unipolar* if the complement $\overline{G}$ is unipolar; and it is a *generalized split* graph if it is unipolar or co-unipolar. Notice that a graph can be both unipolar and co-unipolar, and that when the C_i for $i \geq 1$ are reduced to a single vertex, a generalized split graph is split. It can be shown that generalized split graphs are perfect, and it was proved in [7] that almost all perfect graphs are generalized split.

McDiarmid and Yolov [6] have devised the following process for generating random unipolar graphs. Choose an integer $k \in [n]$ according to a suitable distribution; choose a random k-subset $C_0 \subseteq [n]$; choose (uniformly) a random set partition $[n] \backslash C_0 = C_1 \cup \cdots \cup C_k$ of the complement and make all the C_i into cliques; finally add edges between C_0 and $[n] \backslash C_0$ independently with probability $1/2$, and no further edges. Again this scheme is not uniform but it is shown in [6] that it approximates the uniform distribution on unipolar graphs on n vertices up to total variational distance $o_n(1)$. This gives a useful scheme for random perfect graphs: pick a random unipolar graph G on n vertices according to the previous scheme, and flip a fair coin: if the coin turns up heads then output G, otherwise output its complement $\overline{G}$. Several properties of random perfect graphs are proved in [6] using this scheme. One notable such result is that for every fixed graph H the probability that the random perfect graph on n vertices has an induced subgraph isomorphic to H tends to a limit that is either $0, 1/2$ or 1.

In this paper we consider graph properties that can be expressed in the *first order language of graphs* (FO), on random perfect graphs. We say that a graph G is a model for the sentence $\varphi \in$ FO if G satisfies φ, and write $G \models \varphi$. Several restricted classes of graphs have been studied with respect to the limiting behaviour of FO properties, and usually one proves either a zero-one law (that is, every FO property has limiting probability $\in \{0, 1\}$) or a convergence law (that is, every FO property has a limiting probability). For instance, a zero-one law has been proved for trees [5] and for graphs not containing a clique of fixed size [3], while a convergence law has been proved for d-regular graphs for fixed d [4], and for forests and planar graphs [2].

In the light of the above mentioned result of McDiarmid and Yolov on the limiting probability of containing a fixed induced subgraph, one might expect the convergence law to hold for random perfect graphs, perhaps even with the limiting probabilities only taking the values $0, 1/2, 1$. The main result of this paper however states something rather different is the case.

Theorem 1.1 *There exists a sentence $\varphi \in FO$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left[G_n \models \varphi\right] \text{ does not exist,}$$

where G_n is chosen uniformly at random from all (labelled) perfect graphs on n vertices.

This is in strong contrast with random chordal graphs. The scheme we discussed above based on random split graphs is in fact very similar to the binomial bipartite random graph with independent edge probabilities equal to $1/2$. A standard argument shows that in fact a zero-one law holds in this case, that is, the limiting probability that a FO property is satisfied tends either to 0 or 1 as $n \to \infty$ [9].

Our proof of Theorem 1.1 draws on the techniques introduced in the proof of the celebrated Shelah-Spencer result of non-convergence in the classical $G(n, p)$ model when $p = n^{-\alpha}$ and $\alpha \in (0, 1)$ is a rational number [8] (see also [9]). In fact, it is the richness of unipolar graphs together with the properties of random set partitions that allow us to produce a non-convergent first order sentence.

2 Sketch of the proof

A simple argument shows that it is enough to prove the main result for unipolar graphs, and then the result extends to perfect graphs. The following properties, referred to random unipolar graphs, are proved using the following properties of the scheme for generating random unipolar graphs of McDiarmid and Yolov.

(i) $|C_0| = \frac{n}{2}(1 + o(1))$.

(ii) Let r be the unique root of $re^r = n - |C_0|$. For $t = 1, \ldots, (e - \varepsilon) \ln n$, with $\varepsilon > 0$ arbitrary but fixed, we have

$$|\{j : |C_j| = t\}| = \Omega\left(r^t/t!\right). \tag{1}$$

We note that with high probability we have $r = \ln n - (1 + o(1)) \ln \ln n$.

For $S \subseteq [n]$, let $N(S) = \bigcap_{v \in S} N(v)$ the set of common neighbours of S, and for $S, T \subseteq [n]$ we let $H(S, T)$ denote the graph with vertex set S and an edge between $a, b \in S$ if and only if there is a $v \in T$ that is adjacent to both a and b.

Lemma 2.1 (i) *There exists an FO-formula CN with two free variables such that, with high probability, CN(x, y) holds if and only if $x \in C_0, y \in C_i$ for some $i > 0$, and $x \in N(C_i)$.*

(ii) *The exists an FO-formula Hedge with three free variables such that, with high probability, Hedge(x, y, z) holds if and only if $x, y \in C_0$, $x \neq y$, $z \in C_i$ for some $i > 0$, and xy is an edge of $H(C_0, C_i)$.*

(iii) *For every $\varphi \in$ FO there exists an FO-formula $\Phi(x, y)$ with two free variables such that, with high probability, $\Phi(x, y)$ holds if and only if $x \in C_i, y \in C_j$ for some $i, j > 0$, and $H(N(C_i), C_j) \models \varphi$.*

Let us write $\ell = \lceil \ln \ln \ln n \rceil$. Then we have:

Lemma 2.2 (i) *With high probability, for every $0 \leq \ell' \leq \ell$, there exist $n^{\Omega(1)}$ indices $i > 0$ with $|N(C_i)| = \ell'$.*

(ii) *With high probability, the following holds. For every $i, j > 0$ such that $|N(C_i) \cup N(C_{j'})| \leq 2\ell$, and for every (labelled) graph G with $V(G) = N(C_i) \cup N(C_j)$, there is a $k > 0$ such that $H(N(C_i) \cup N(C_j), C_k) = G$.*

Now comes the key ingredient in the proof of our main theorem.

Lemma 2.3 *There exists an FO-formula Bigger with two free variables such that, with high probability:*

- *If Bigger(x, y) holds then there exist $i, j > 0$ such that $x \in C_i, y \in C_j$ and $|N(C_i)| > |N(C_j)|$;*
- *If $x \in C_i, y \in C_j$ for some $i, j > 0$ with $|N(C_j)| < |N(C_i)| \leq \ell$ then Bigger(x, y) holds.*

Proof of the main result. We need the following lemma, which is a straightforward adaptation of a construction of Shelah and Spencer [8]. Given a formula φ, the spectrum $\mathrm{spec}(\varphi)$ is the set of all n such that there exists a graph on n vertices that satisfies φ. The function $\log^* n$ is the classical iterated logarithm.

Lemma 2.4 *There exist $\varphi_0, \varphi_1 \in FO$ such that*

$$\begin{aligned} \log^* n \mod 100 \in \{2, \dots 49\} &\Rightarrow n \in \mathit{spec}(\varphi_0) \setminus \mathit{spec}(\varphi_1), \\ \log^* n \mod 100 \in \{52, \dots 99\} &\Rightarrow n \in \mathit{spec}(\varphi_1) \setminus \mathit{spec}(\varphi_0). \end{aligned}$$

Let Φ_i denote the formula that Lemma 2.1 produces when applied to the sentence φ_i from Lemma 2.4. We define the following FO-sentence:

$$\varphi = \exists x, y : \Phi_1(x, y) \wedge \neg(\exists x', y' : \mathrm{Bigger}(x', x) \wedge \Phi_0(x', y')).$$

Up to error probability $o(1)$, φ will hold if and only if $H(N(C_i), C_j) \models \varphi_1$ for some $i, j > 0$, and moreover if $H(N(C_{i'}), C_{j'}) \not\models \varphi_0$ for some $i', j' > 0$ then $\mathrm{Bigger}(x, x')$ does not hold for any $x \in C_i, x' \in C_{i'}$. We briefly explain how this implies that φ does not have a limiting probability.

First we consider an increasing subsequence $(n_k)_k$ of the natural numbers for which $\log^* n \mod 100 = 75$. Observe that

$$\log^* n - 10 \leq \log^* \ell \leq \log^* n. \tag{2}$$

With high probability there are lots of C_i for which $|N(C_i)| = \ell$ by Lemma 2.2, and for each of these there is a j such that $H(N(C_i), C_j) \models \varphi_1$ (since $\ell \in \mathrm{spec}(\varphi_1)$ as $\log^* \ell \mathrm{mod} 100 \in \{65, \dots, 75\}$ by (2) and the choice of n). So there are (lots of) pairs of vertices x, y such that $\Phi_1(x, y)$ holds and $x \in C_i$ for some $i > 0$ with $|N(C_i)| = \ell$. On the other hand, with high probability, for any x' such that $\mathrm{Bigger}(x', x)$ it must hold that $x' \in C_{i'}$ for some $i' > 0$ with $\ell = |N(C_i)| < |N(C_{i'})| \leq n$. So in particular $\log^*(|N(C_{i'})|) \in \{65, \dots, 75\}$. Thus $|N(C_{i'})| \notin \mathrm{spec}(\varphi_0)$, which shows that $H(N(C_{i'}), C_{j'}) \not\models \varphi_0$ for any $j' > 0$. In other words, if $\mathrm{Bigger}(x', x)$ holds then there cannot be any y' such that $\Phi_0(x', y')$ holds. This shows that

$$\lim_{\substack{n\to\infty,\\ \log^* n \bmod 100 = 75}} \mathbb{P}(G_n \models \varphi) = 1.$$

Next, let us consider an increasing subsequence $(n_k)_k$ of the natural numbers for which $\log^* n \bmod 100 = 25$. In this case $\log^* \ell \bmod 100 \in \{15, \ldots, 25\}$. In particular $\ell, \ldots, n \notin \operatorname{spec}(\varphi_1)$. So, with high probability, if there is pair x, y such that $\Phi_1(x, y)$ holds then we must have $x \in C_i$ for some $i > 0$ with $|N(C_i)|$ strictly smaller than ℓ. But then we can again apply Lemma 2.2 to find that, with high probability, there exist x', y' with $x' \in C_{i'}, y \in C_{j'}$ for some $i', j' > 0$ such that $|N(C_{i'})| = \ell$ and $H(N(C_{i'}), C_{j'}) \models \varphi_0$. Since $\ell = |N(C_{i'})| > |N(C_i)|$, with high probability, $\mathrm{Bigger}(x', x)$ will hold by Lemma 2.3. This shows that

$$\lim_{\substack{n\to\infty,\\ \log^* n \bmod 100 = 25}} \mathbb{P}(G_n \models \varphi) = 0.$$

References

[1] E. A. Bender, L. B. Richmond, N. C. Wormald. Almost all chordal graphs split. *J. Austral. Math. Soc. Ser. A* 38 (1985), 214-221.

[2] P. Heinig, T. Müller, M. Noy, A. Taraz. Logical limit laws for minor-closed classes of graphs, to appear in *J. Combin. Th. Ser. B*

[3] Ph. G. Kolaitis, H. J. Prömel, B. L. Rothschild. $K_{\ell+1}$-free graphs: asymptotic structure and a 0-1 law. *Trans. Amer. Math. Soc.* 303 (1987), 637–671.

[4] J. F. Lynch. Probabilities of sentences about very sparse random graphs. *Random Structures and Algorithms* 3 (1992), 33–53.

[5] G. L. McColm. MSO zero-one laws on random labelled acyclic graphs. *Discrete Math.* 254 (2002), no. 1-3, 331–347.

[6] C. McDiarmid, N. Yolov. Random prefect graphs, arXiv:1604.00890.

[7] H. J. Prömel, A. Steger. Almost all Berge graphs are perfect. *Combin. Probab. Comput.*1 (1992), 53–79.

[8] S. Shelah and J. Spencer. Zero-one laws for sparse random graphs. *J. Amer. Math. Soc.* 1 (1988), 97-115.

[9] J. Spencer. *The Strange Logic of Random Graphs*. Springer (2001).