Almost all graphs with $2.522n$ edges are not 3-colorable

Dimitris Achlioptas*
Department of Computer Science
University of Toronto

Michael Molloy*
Department of Computer Science
University of Toronto

Abstract

We show that if a random graph on n vertices has $m = (r + o(1))n$ edges, where r=2.522, then almost surely it is not 3-colorable. The previous best such value for r was 2.571[5]. Our result follows from applying ideas of [4] to the k-coloring problem. The development of the proof is for k-colorings, where k is an arbitrary constant, and we present similar improvements for small values of k.

1 Introduction

The term "almost all" in the title has the meaning introduced by Erdös and Rényi [3]. If $N(n, m, A)$ stands for the number of graphs with vertices $\{1, \ldots, n\}$, precisely m edges, and some property A, saying that almost all graphs with n vertices and $m(n)$ edges have property A, means that

$$\lim_{n \to \infty} \frac{N(n, m(n), A)}{\binom{\binom{n}{2}}{m(n)}} = 1 \ .$$

(In the terminology of Bollobás [1], we are using Model $\mathcal{G}_M$ of a random graph.)

2 Upper bounds for k-colorability

2.1 The first moment method

In the rest of the paper a random graph $G(V, E)$ will always have $m = m(n) = rn$ edges, for some constant r and it will be convenient to think of a k-coloring of G as a partition of V into k cells such that every $e \in E$ has its endpoints in distinct cells of the partition. The following "first moment" argument is due to Devroye (see [2]) and we include it for instructive purposes.

Let $P = V_1, \ldots, V_k$ be an arbitrary partition of V and let C_P denote the event that P is a k-coloring of G. If we let

$$T(P) = \sum_{i \neq j} |V_i||V_j| \leq \frac{k-1}{2k} n^2 \tag{1}$$

then the probability that P is a k-coloring of G is the probability that all m edges in E are enumerated by $T(P)$, i.e.,

$$\Pr[C_P] = \frac{\binom{T(P)}{m}}{\binom{\binom{n}{2}}{m}} \leq \left(\frac{T(P)}{\binom{n}{2}} \right)^m \leq O(1)(2\tau)^m, \text{ where } \tau = \frac{T(P)}{n^2} \ . \tag{2}$$

*Authors' address: Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 3G4. E-mail: {optas,molloy}@cs.toronto.edu

As there are exactly k^n k-partitions of V, and $\tau \leq \frac{k-1}{2k}$ the expected number of k-colorings of G is upper bounded by $O(1)\left(k\left(\frac{k-1}{k}\right)^r\right)^n$. Thus, if $m = (1+o(1))rn$, where $r = \frac{\ln k}{\ln k - \ln(k-1)}$, the expected number of k-colorings of G tends to zero as n tends to infinity and hence G is almost surely not k-colorable. For $k = 3$, this gives $r = 2.710$.

2.2 A refinement of the first moment method

The price of the simplicity of the first moment method is that, for a fixed r, the graphs with a very large number of colorings, although they may occur with very small probability, contribute substantially to the expectation. In other words, by substituting the number of k-colorings for the variable indicating the existence of a k-coloring (which is the essence of the first moment method) we might give away "a lot". The technique introduced in [4], as we apply it to coloring, amounts to taking the expectation of only those k-colorings that satisfy a certain "local maximality" condition. The underlying intuition is that for a random k-coloring of G, if we choose a node at random and assign it a different color, the probability that we will end up with a partition that is not a coloring is small. Consequently, colorings (when they exist) tend to appear in large "clusters" and instead of counting all the colorings in a cluster (as the first moment does) we can count only a "representative" coloring (locally maximum). The original application of this idea was to k-SAT [4] and it gave the best bound currently known.

Definitions and Notation. Recall that for a fixed partition $P = V_1, V_2, \ldots, V_k$ of V, C_P denotes the event that P is a k-coloring of G. We will say that a vertex $v \in V_i$ is *unmovable* in P if for every $j > i$ the partition resulting by moving v to V_j is not a coloring of G. We will say that P is a *rigid* coloring of G if every vertex is unmovable in P and we will denote this event by R_P. If colorings are interpreted as strings in $\{1, \ldots, k\}^n$ the lexicographically maximum coloring of G, if it exists, is rigid and hence we see that G has a rigid k-coloring iff G is k-colorable. Thus, in the following it suffices to establish a condition subject to which the expected number of rigid k-colorings of G tends to zero as n tends to infinity.

3 Probability Calculations

We will first bound the probability of R_P, for an arbitrary $P = V_1, V_2, \ldots, V_k$. By (2) it suffices to bound $\Pr[R_P | C_P]$. Let $\alpha_i = |V_i|/n$, for $1 \leq i \leq k$. The probability that a vertex v cannot be moved to V_j is the probability that for at least one $w \in V_j$, $\{v, w\} \in E$. As we have conditioned on C_P, all the edges of G are enumerated by T_P. Hence, each edge has probability $|V_j|/T_P = \frac{\alpha_j}{\tau n}$ to be a blocking edge of v. Since the edges of G are chosen uniformly and independently, the probability that v cannot be moved to V_j is $1 - (1 - \frac{\alpha_j}{\tau n})^{rn} = 1 - e^{-\alpha_j r/\tau} + o(1)$.

Let S be the set containing all events of the type "v cannot be moved from V_i to V_j", where $v \in V$ and $i < j$. Since all blocking edges for v must have v as an endpoint (and $i < j$) it's clear that for any two distinct elements of S the corresponding sets of blocking edges are mutually disjoint. Hence, for any event $\langle v, i, j\rangle \in S$, if $\mathcal{C}$ is an arbitrary conjunction of events from $S - \langle v, i, j\rangle$ then $\Pr[\langle v, i, j\rangle] \geq \Pr[\langle v, i, j\rangle | \mathcal{C}]$ [1]. Using this fact we get,

$$\Pr[R_P | C_P] \quad = \quad \Pr[\bigwedge_{1 \leq i < j \leq k} \text{No vertex from } V_i \text{ can move to } V_j]$$

[1]Intuitively speaking, any event in $\mathcal{C}$ "exposes" only edges that do not participate in $\langle v, i, j\rangle$ and moreover it "exposes" the fact that some edges have been "used".

$$
\begin{aligned}
&\le \prod_{1\le i<j\le k} (1-e^{-\alpha_j r/\tau})^{\alpha_i n} \\
&= \left[\prod_{2\le j\le k} \left(1-e^{-\alpha_j r/\tau}\right)^{\sum_{i<j}\alpha_j}\right]^n \qquad (3)\\
&\equiv f(\alpha_1,\ldots,\alpha_k,n)
\end{aligned}
$$

For the expected number of rigid colorings of G, $E[\mathcal{G}]$, by using (3) above, we get

$$
\begin{aligned}
E[\mathcal{G}] &\le \sum_{\alpha_1+\cdots+\alpha_k=1} \binom{n}{\alpha_1 n,\ldots,\alpha_k n} f(\alpha_1,\ldots,\alpha_k,n) \\
&\le \max_{\alpha_1+\cdots+\alpha_k=1} \left\{ \binom{n}{\alpha_1 n,\ldots,\alpha_k n} f(\alpha_1,\ldots,\alpha_k,n) \right\} n^k \ , \qquad (4)
\end{aligned}
$$

since there are at most n^k ways to partition n into k integers. Moreover, for all sufficiently large n,

$$
\binom{n}{\alpha_1 n,\ldots,\alpha_k n} < \left(\frac{1}{\alpha_1^{\alpha_1}\cdots {\alpha_k}^{\alpha_k}}\right)^n, \text{ where } 0^0\equiv 1 \ . \qquad (5)
$$

From (1)...,(5) we get that $E[\mathcal{G}]$ is upper bounded by

$$
O(1)\left(\max_{\alpha_1+\alpha_2+\cdots+\alpha_k=1}\left\{\frac{1}{\alpha_1^{\alpha_1}\cdots {\alpha_k}^{\alpha_k}}\left(2\sum_{i\ne j}\alpha_i\alpha_j\right)^r \prod_{2\le j\le k}\left(1-e^{-\frac{\alpha_j r}{\sum_{i\ne j}\alpha_i\alpha_j}}\right)^{\sum_{i<j}\alpha_j}\right\}\right)^n n^k \ . \qquad (6)
$$

Let $g(\alpha_1,\cdots,\alpha_k,n)$ denote the term to be maximized in (6). The dependency of g on n comes from the definition $\alpha_i = |V_i|/n$ (implying that each a_i can only be a rational q/n, where $1\le q\le n$). By examining the underlying combinatorial structure, it's not hard to see that, subject to $\alpha_1+\cdots+\alpha_k=1$, g is unimodal. If we relax each each α_i to a real number in $[0,1]$ we get $g'(\alpha_1,\cdots,\alpha_k)$ which is continuous in $[0,1]^k$ and hence unimodal as well. Maximizing such a function can be done numerically with arbitrarily good precision and to simplify things we substitute α_k with $1-\alpha_1-\cdots-\alpha_{k-1}$ and treat g' as a function defined inside the $(k-1)$-dimensional unit sphere S^{k-1}. Thus, if for some value of r $\max_{\vec{\alpha}\in S^{k-1}} g'(\vec{\alpha})<1$ then $E[\mathcal{G}]$ tends to zero as n tends to infinity, giving us the desired condition. For $k=3$ g' is

$$
\frac{(2\phi)^r\left(1-e^{-\frac{\alpha_2 r}{\phi}}\right)^{\alpha_1}\left(1-e^{-\frac{(1-\alpha_1-\alpha_2)r}{\phi}}\right)^{\alpha_1+\alpha_2}}{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}(1-\alpha_1-\alpha_2)^{1-\alpha_1-\alpha_2}},
$$

where $\phi=\alpha_1+\alpha_2-\alpha_1\alpha_2-\alpha_1^2-\alpha_2^2$. For $r=2.5217$ it is maximized at $\alpha_1=0.30746, \alpha_2=0.33527$ and at that point it is less than 0.9999744. Similarly, we get

k	3	4	5	6	7
First moment method	2.7096	4.8188	7.2126	9.8275	12.6234
New bound	2.5217	4.5861	6.9479	9.5389	12.3160

Acknowledgements

We would like to thank the authors of [4] for providing us with an early draft of their paper and NSERC for financial support.

References

[1] B. Bollobás, *Random Graphs*, Academic Press, London, (1985).

[2] V. Chvátal, *Almost all graphs with 1.44n edges are 3-Colorable*, Random Structures and Algorithms, **2** (1991), 11–28.), 329–359.

[3] P. Erdös, A. Rényi, *On the evolution of random graphs*, Publication of the Mathematical Institute of the Hungarian Academy of Sciences, **5** (1960), 17–61.

[4] L. M. Kirousis, E. Kranakis, D. Krizanc, *Approximating the unsatisfiability threshold of random formulas*, Proceedings of the 4th European Symposium on Algorithms, (1992) 27–38.

[5] M. Molloy, *The chromatic number of sparse random graphs*, Masters Thesis, University of Waterloo, (1992).