Problem Set 6 CS265, Autumn 2022

Due: Friday November 11, 11:59pm on Gradescope.

Group members: Rylan Schaeffer, Kelly Brennan

Please follow the homework policies on the course website.

Note: In this homework you may find the following inequality useful. When $x \in (0,1)$:

$$\exp\left(-\frac{x}{1-x}\right) \leq 1 - x \leq \exp(-x).$$

A nice special case is that if $x \in (0, 1/2)$ then $1 - x \geq e^{-2x}$. Feel free to use these without proof!

1. **(10 pt.) [Threshold for isolation]**

   Recall that $G_{n,p}$ refers to a random graph with n vertices, where each of the $\binom{n}{2}$ possible edges is present independently with probability p.

   (a) **(2 pt.)** Suppose that $p = 1.01\frac{\ln n}{n}$. Show that $G_{n,p}$ has an isolated vertex with probability $o(1)$.

   (b) **(4 pt.)** Let $X_1, X_2, ..., X_n$ be (*not* necessarily independent, and *not* necessarily identically distributed) Bernoulli random variables[1], and let $X = \sum_{i=1}^{n} X_i$. Prove that

   $$\mathbb{E}[X^2] = \sum_{i=1}^{n} \Pr[X_i = 1] \cdot \mathbb{E}[X \mid X_i = 1].$$

   (c) **(4 pt.)** Suppose that $p = 0.99\frac{\ln n}{n}$. Show that $G_{n,p}$ has an isolated vertex with probability $1 - o(1)$.

   [**HINT:** *Consider using part (b) – it might make the math simpler.*]

SOLUTION:

(a) Let $p = 1.01\frac{\log n}{n}$. Define X as the number of isolated vertices. A vertex is isolated if the $n-1$ possible edges to other vertices do not exist. By linearity of expectation:

$$\mathbb{E}[X] = \sum_i \mathbb{E}[X_i \text{ is isolated}] = n(1-p)^{n-1} \leq n\exp(-1.01\frac{\log n}{n})^{n-1}$$

Simplifying:

$$n(\exp(-\log n))^{1.01(n-1)/n} = n^{1.01/n - 0.01} \leq \frac{2}{n^{0.01}} = o(1)$$

Because X is a non-negative random variable, we can apply Markov's:

$$\mathbb{P}[X \geq 1] \leq \mathbb{E}[X] \leq \frac{2}{n^{0.01}} = o(1)$$

[1] i.e., they only take on the values 0 and 1.

(b) (I prefer a slightly different notation for expectations that I feel is more clear.) Using linearity of expectation, the definition of X and the law of total expectation:

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}[\sum_{ij} X_i X_j] \\
&= \mathbb{E}[\sum_i X_i \sum_j X_j] \\
&= \sum_i \mathbb{E}[X_i X] \\
&= \sum_i \mathbb{E}_{X_i}[\mathbb{E}_{X|X_i}[X_i X]] \\
&= \sum_i \mathbb{E}_{X_i}[X_i[\mathbb{E}_{X|X_i}[X]] \\
&= \sum_i \mathbb{P}[X_i = 1]\, \mathbb{E}_{X|X_i}[X]
\end{aligned}$$

(c) Let X_i denote the indicator that the ith node is isolated, and let $X = \sum_i X_i$. We want to show that with $p = 0.99\frac{\log n}{n}$:

$$\mathbb{P}[X > 0] = 1 - \mathbb{P}[X = 0] = 1 - o(1)$$

Repeating 1a with our new probability, we find that:

$$\mathbb{E}[X] = \sum_i \mathbb{E}[X_i] = n(1-p)^{n-1}$$

From Lecture 11 Theorem 1, we know that:

$$\mathbb{P}[X = 0] \leq \frac{\mathbb{V}[X]}{\mathbb{E}[X]^2} = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mathbb{E}[X]^2} = \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2} - 1$$

Recalling that $X = \sum_i X_i$, denote $X_{-i} = \sum_{j \neq i} X_j$ and note that:

$$\mathbb{E}_{X|X_i=1}[X] = 1 + \mathbb{E}_{X_{-i}}[X_{-i}] = 1 + (n-1)(1-p)^{n-2}$$

because if $X_i = 1$, then the ith node is independent and X_{-i} refers to a disconnected graph we can think of as being independent. From 1b, we know that:

$$\mathbb{E}[X^2] = \sum_i \mathbb{P}[X_i = 1]\mathbb{E}_{X|X_i=1}[X] = n(1-p)^{n-1}(1 + (n-1)(1-p)^{n-2})$$

Combining the above, we have:

$$\begin{aligned}
\mathbb{P}[X=0] &\leq \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2} - 1 \\
&= \frac{n(1-p)^{n-1}(1+(n-1)(1-p)^{n-2})}{(n(1-p)^{n-1})^2} - 1 \\
&= \frac{1}{n(1-p)^{n-1}} + \frac{(n-1)(1-p)^{n-2}}{n(1-p)^{n-1}} - 1 \\
&\leq \frac{1}{n(1-p)^{n}} + \frac{(n-1)}{n(1-p)} - 1
\end{aligned}$$

Per Wolfram Alpha, as $n \to \infty$, the first term $\to 0$ and the middle term can be split into $1/(1-p)$ and $1/(n(1-p))$ $1/(1-p) \approx 1$, which annihilates with -1, whereas $1/(n(1-p) \approx n^{-1}$. Thus, $\mathbb{P}[X=0] \leq o(1)$ and therefore $\mathbb{P}[X>0] = 1 - o(1)$.

2. **(6 pt.) [Echoing paths]**

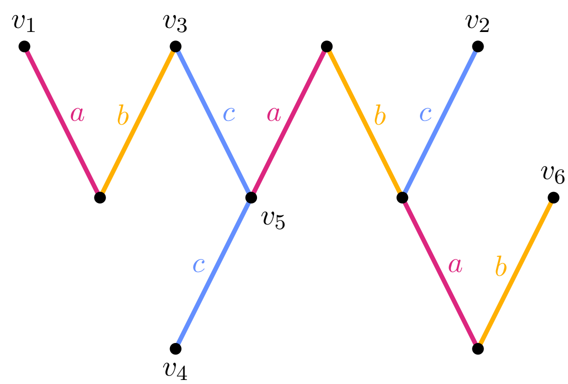

Figure 1: An edge coloring of a graph with some echoing paths.

An *edge coloring* of an (undirected) graph $G = (V, E)$ assigns exactly one color to each edge of the graph. We say that a colored path in the graph is *echoing* if the path has an even number of edges, and the second half of the path is colored identically to the first half of the path (i.e. the sequence of colors in the second half of the path is the same sequence as in the first half). For example, in Figure 1, the paths from v_1 and v_2, from v_3 to v_4, and from v_5 to v_6 are all echoing paths. Edges are colored and labeled a, b, or c corresponding to their color.

Throughout this problem, by "path" we refer only to simple paths—i.e. paths that do not re-use any edges.

(a) **(4 pt.)** Prove that for any graph whose maximum degree is d, there exists a coloring using $10 \cdot d^2$ colors such that there are no echoing paths of length 4 (i.e. no repeating paths consisting of 4 distinct edges).

[**HINT:** *Lovasz Local Lemma!*]

(b) **(2 pt.)** Given the setup in the previous part, give an algorithm that will find such a coloring in expected time polynomial in the size of the graph, and justify the runtime.

(c) **(0 pt.) [This problem is optional.]** Prove that there is some constant C such that for any graph whose maximum degree is d, there exists a coloring using $C \cdot d^2$ colors such that there are no echoing paths (of any length).

SOLUTION:

(a) Define A_i as the event that the ith simple path of length 4 echoes. If we consider that path, the first two edges can be arbitrarily colored, but the last two edges must have exactly matching colors. Therefore,

$$\mathbb{P}[A_i] = \left(\frac{1}{10d^2}\right)^2 = \frac{1}{100d^4} := p$$

For this ith simple path of length 4, how many other simple paths of length 4 are mutually dependent? The answer is: for a length 4 path, there are 4 ways to choose 1 overlapping edge, then $\leq d^3$ ways to extend that path, times 2 because the path could be extended from the "left" end or from the "right" end, for a total $\leq 8d^3$. There are 3 ways to choose 2 overlapping edges and 2 ways to choose 3 overlapping edges, but the number of ways to extend these paths is $O(d^2)$ and $O(d)$, and since $d \geq 1$, the cubic term dominates. Thus, $pd = \frac{1}{100d^4} 8n^3 \leq \frac{1}{4}$, and per the LLL, the probability that no length 4 simple path is echoing is greater than 0 and thus there must exist some coloring using $10d^2$ colors such that there are no echoing paths of length 4.

(b) Run Algorithm 2 from Lecture 12, using the edge colors as the independent random variables. Since $\mathbb{P}[A_i = 1] = \frac{1}{100d^4} \leq \frac{1}{e(d+1)}$, per Corollary 3, this algorithm will find an edge coloring such that no simple path of length 4 will be echoing.

3. **(0 pt.) [Tightness of the Lovasz Local Lemma]**

This whole problem is optional and will not be graded.

One version of the LLL that we saw asserts that for any set of events $A_1, \ldots, A_n$, such that for each i, A_i is mutually independent of all but at most d events, then as long as $\Pr[A_i] \leq \frac{1}{e(d+1)}$, then there is a nonzero chance of all events being simultaneously avoided.

(a) Define a set of events over a probability space such that each event is mutually independent of all but at most d other events, and $\Pr[A_i] \leq 1/(d+1)$ for all i, but the probability of simultaneously avoiding all events A_i is 0. This shows that the constant e in the statement of the LLL cannot be replaced by 1.

(b) **(Challenge!)** For some constant $c > 1$, prove that the constant e in the LLL cannot be replaced by c.

SOLUTION:

(a)