Homework 6

Submit your solutions in pdf format on Gradescope by 5pm on Friday, March 17. Solutions may be written either in \LaTeX (with either machine-drawn or hand-drawn diagrams) or legibly by hand. (The \LaTeX source for this homework is provided in case you want to use it as a template.) Please be sure to begin the solution for each problem on a new page, and to tag each of your solutions to the correct problem! Per course policy, no late solutions will be accepted. Take time to write clear and concise answers; confused and long-winded solutions may be penalized. You are encouraged to form small groups (two to four people) to work through the homework, but you must write up all your solutions on your own. Depending on grading resources, we reserve the right to grade a random subset of the problems and check off the rest; so you are advised to attempt all the problems.

1. [Adapted from MU, Exercise 5.17.]
   (a) Let $G$ be a random graph in the $G_{n,p}$ model. For each of the following structures, determine the value of $p$ (as a function of $n$) for which the expected number of the structures in $G$ is equal to 1. You should just give the asymptotic value of $p$ (i.e., ignore constants and lower order terms).
   
   (i) Cliques of size 5.
   
   (ii) Complete bipartite graphs $K_{5,5}$ (five vertices on each side). [Note that the edges “missing” from $K_{5,5}$ may or may not be included in $G$: i.e., each copy of $K_{5,5}$ need not be an induced subgraph of $G$.]
   (iii) Hamilton cycles.

   (b) It turns out that the values you identified in parts (i) and (ii) are in fact thresholds for the existence of the respective structures in $G$ (in the sense we discussed in class and in MU Section 6.5 for 4-cliques—and indeed this can be proved using the second moment method, exactly as we did for 4-cliques). We might think that the same holds for the value in part (iii). However, if you have done part (iii) correctly then you will have obtained a value that is different from the threshold for the existence of Hamilton cycles, which as we stated in class is $p = \frac{\ln n}{n}$. Explain carefully why this is not a contradiction.

2. In this problem we will see that the value $p = \frac{\ln n}{n}$ is a threshold for the property that a random graph in the $G_{n,p}$ model has an isolated vertex, i.e., a vertex with no adjacent edges. That is, we will prove that

   \[
   \Pr[G \text{ has an isolated vertex}] \xrightarrow{n \to \infty} \begin{cases} 
   0 & \text{if } p = \omega\left(\frac{\ln n}{n}\right); \\
   1 & \text{if } p = o\left(\frac{\ln n}{n}\right).
   \end{cases}
   \]

   (a) Let the r.v. $X$ denote the number of isolated vertices in $G$. Write down the expectation of $X$ as a function of $n$ and $p$.

   (b) Show that $E[X] \to 0$ for $p = \omega\left(\frac{\ln n}{n}\right)$, and that $E[X] \to \infty$ for $p = o\left(\frac{\ln n}{n}\right)$.

   (c) Deduce from part (b) that $\Pr[G \text{ has an isolated vertex}] \to 0$ for $p = \omega\left(\frac{\ln n}{n}\right)$.

   (d) Show that $\text{Var}[X] = n(1-p)^{n-1} + n(1-p)^{2n-3}(np - 1)$.

   (e) Deduce from parts (b) and (d) that $\Pr[G \text{ has an isolated vertex}] \to 1$ for $p = o\left(\frac{\ln n}{n}\right)$.

[Turn over for problems 3 & 4]
4. Recall that a graph (undirected, no self-loops) is \textit{2-colorable} if we can assign colors red and green to each vertex such that the endpoints of every edge are assigned different colors. Suppose we are told that a graph \( G = (V, E) \) is “locally 2-colorable”, in the sense that the induced subgraph\(^1\) on every subset of \( O(\log n) \) vertices is 2-colorable. Does this imply that \( G \) itself is 2-colorable? In this problem we will see that the answer is spectacularly “no”: namely, we will show that there exists a graph that is locally 2-colorable but is “very far away” from being 2-colorable, in the sense that we would have to remove a constant fraction of its edges in order to make it 2-colorable. We will prove the existence of this graph using the probabilistic method.

Throughout, set \( p = \frac{16}{n} \), and let \( G \) be a random graph from the model \( G_{n,p} \). The probabilities and expectations refer to the experiment of picking \( G \) at random.

(a) Write down the expected number of edges in \( G \).

(b) Apply the Chernoff bound to show that with probability \( 1 - 2^{-\Omega(n)} \), \( G \) has at most \( 10(n-1) \) edges. \[ \text{HINT: Use the Chernoff bound in the form } \Pr[X \geq (1 + \delta)\mu] \leq \exp(-\delta^2\mu/3) \text{ for } 0 < \delta < 1. \]

(c) Now fix an arbitrary (not random!) assignment of colors to the vertices. Show that the expected number of violated edges (i.e., edges with endpoints of the same color) in \( G \) is at least \( 4(n-2) \). Deduce by a Chernoff bound that the probability there are more than \( n-2 \) violated edges is at least \( 1 - \exp(-9(n-2)/8) \). \[ \text{HINTS: For the first part, think of the assignment of colors as being fixed before we choose the random edges of } G. \text{ What is the value for the number of red/green vertices that minimizes the expected number of violated edges? You may assume for simplicity that } n \text{ is even. Use the Chernoff bound in the form } \Pr[X \leq (1 - \delta)\mu] \leq \exp(-\delta^2\mu/2) \text{ for } 0 < \delta < 1. \]

(d) Show that for \( n \geq 6 \), with probability at least \( 3/4 \), \( G \) is not 2-colorable even if we delete any \( n-2 \) of its edges. \[ \text{HINT: Use the previous part and a union bound over colorings.} \]

(e) Show that the expected number of cycles of length exactly \( k \) in \( G \) is at most \( 16^k \). Deduce that the expected number of cycles of length at most \( \frac{1}{8} \log n \) is at most \( 16\sqrt{n} \). \[ \text{HINT: Use the fact that } \sum_{k=1}^{n} 16^k < 16^{n+1}. \]

(f) Use the previous part to deduce that, with probability at least \( 3/4 \), by deleting only \( O(\sqrt{n}) \) (suitably chosen) edges of \( G \) we can obtain a graph such that the induced subgraph on any subset of \( \frac{1}{8} \log n \) vertices is cycle-free (i.e., a forest – a collection of vertex-disjoint trees. Note that a forest is always 2-colorable.)

(g) Put parts (b), (d) and (f) together to deduce that, for every sufficiently large \( n \), there exists a graph \( G = G_n \) on \( n \) vertices such that:
   
   - The induced subgraph on any subset of \( \frac{1}{8} \log n \) vertices of \( G_n \) is 2-colorable; and
   - \( G_n \) is not 2-colorable, and remains not 2-colorable even after deleting any 0.05 fraction of its edges.

[\text{HINT: Remember to take into account the fact that, when we modify } G \text{ to remove cycles, we may also be deleting violated edges!}]

\(^1\)The \textit{induced} subgraph on a subset of vertices \( V' \subseteq V \) is the graph with vertex set \( V' \) and edge set consisting of all the edges of \( G \) both of whose endpoints are in \( V' \).