CS 170 HW 13 (Solve any 3 out of the 5 problems)

Due **2020-27-04, at 10:00 pm**

1 Study Group

List the names and SIDs of the members in your study group. If you have no collaborators, you must explicitly write none.

2 Opting for releasing your solutions

We are considering releasing a subset of homework submissions written by students for students to see what a full score submission looks like. If your homework solutions are well written, we may consider releasing your solution. If you wish that your solutions not be released, please respond to this question with a "No, do not release any submission to any problems". Otherwise, say "Yes, you may release any of my submissions to any problems".

3 QuickSelect

Let $\mathbf{X}_{i,j}$ be an indicator random variable for the event that the i-th smallest number is ever compared with the j-th smallest in $\mathsf{QuickSelect}(A, k)$.

(a) Write an exact expression for $\mathbf{E}[\mathbf{X}_{i,j}]$.

(b) Show that the expected runtime of $\mathsf{QuickSelect}(A, k)$ is $O(n)$.

4 $\sqrt{n}$ alloys

(a) Let G be a graph of maximum degree Δ. Show that G is $(\Delta + 1)$-colorable.

(b) Suppose G is a 3-colorable graph. Let v be any vertex in G. Show that the graph induced on the neighborhood of v is 2-colorable. *Clarification: the graph induced on the neighborhood of v refers to the subgraph of G obtained from the vertex set V' comprising vertices adjacent to v and edge set comprising all edges of G with both endpoints in V'.*

(c) Give a polynomial time algorithm that takes in a 3-colorable n-vertex graph G as input and outputs a valid coloring of its vertices using $O(\sqrt{n})$ colors. Prove that your algorithm is correct and also analyze its runtime.
Hint: think of an algorithm that first colors "high-degree" vertices and their neighborhoods, and then colors the rest of the graph. The previous two parts might be useful.

5 Cuts from electric charges

Given a graph $G = (V, E)$ on n vertices and m edges, and a vector $x \in \{\pm 1\}^n$, we say

$$\mathsf{Cut}(G, x) := \frac{1}{m} \sum_{\{i,j\} \in E} \left(\frac{x_i - x_j}{2} \right)^2$$

and define

$$\mathsf{MaxCut}(G) := \max_{x \in \{\pm 1\}^n} \mathsf{Cut}(G, x).$$

For every algorithmic question below, please analyze your runtime and prove that your algorithm is correct.

(a) (Warmup; ungraded) Let G be any graph. Prove that

$$\mathsf{MaxCut}(G) \geq \frac{1}{2}$$

always, and give a polynomial time randomized algorithm that outputs $\boldsymbol{x} \in \{\pm 1\}^n$ satisfying $\mathbf{E}[\mathsf{Cut}(G, \boldsymbol{x})] \geq \frac{1}{2}$.
What if each x_i is chosen to be $+1$ or -1 uniformly independently?

(b) Let G be any graph. Prove that

$$\mathsf{MaxCut}(G) \geq \frac{1}{2} + \Omega\left(\frac{1}{n}\right)$$

always, and give a polynomial time randomized algorithm that outputs $\boldsymbol{x} \in \{\pm 1\}^n$ satisfying $\mathbf{E}[\mathsf{Cut}(G, \boldsymbol{x})] \geq \frac{1}{2} + \Omega\left(\frac{1}{n}\right)$.
Hint: try to construct a simple distribution over ± 1 vectors such that for $\boldsymbol{x}$ drawn from this distribution any $i, j \in [n]$, $\mathbf{E}[\boldsymbol{x}_i \boldsymbol{x}_j] = -\frac{c}{n}$ for some absolute constant $c > 0$.

(c) Let G be any k-colorable graph. Prove that

$$\mathsf{MaxCut}(G) \geq \frac{1}{2} + \Omega\left(\frac{1}{k}\right).$$

Note that we are not asking you to give an algorithm to find such a cut, but instead just asking you to prove existence. Try to reduce this problem to the previous part.

(d) Let G be any 3-colorable graph. Give a polynomial time randomized algorithm to find a $\boldsymbol{x} \in \{\pm 1\}^n$ satisfying:

$$\mathbf{E}[\mathsf{Cut}(G, \boldsymbol{x})] \geq \frac{1}{2} + \Omega\left(\frac{1}{\sqrt{n}}\right).$$

Hint: Part (b) of Question 4 may be useful.

(e) Let G be any graph with maximum degree Δ. Prove that

$$\mathsf{MaxCut}(G) \geq \frac{1}{2} + \Omega\left(\frac{1}{\Delta}\right).$$

Give a polynomial time randomized algorithm to find a $\boldsymbol{x} \in \{\pm 1\}^n$ satisfying:

$$\mathbf{E}[\mathsf{Cut}(G, \boldsymbol{x})] \geq \frac{1}{2} + \Omega\left(\frac{1}{\Delta}\right).$$

Hint: Part (a) of Question 4 might be useful here.

6 Fixing Blemishes

We use $G(n,p)$ to denote the distribution of graphs obtained by taking n vertices and for each pair of vertices i,j placing edge $\{i,j\}$ independently with probability p.

(a) Compute the expected number of edges in $G(n,p)$?

(b) Compute the expected number of 4-cycles in $G(n,p)$?

(c) Give a polynomial time randomized algorithm that takes in n as input and in poly(n)-time outputs a graph G such that G has no 4-cycles and the expected number of edges in G is $\Omega(n^{4/3})$.

7 Porcupine Trio

Let L be a vector of integers in $[-M,M]^n$ given to us as input. For any other vector x, we will use $\langle L,x\rangle$ to denote $\sum_{i=1}^n L_i x_i$. In this problem we will assume $M < 2^n/(4n)$.

(a) Prove that there exists distinct $x_1, x_2 \in \{\pm 1\}^n$ such that $\langle L, x_1\rangle = \langle L, x_2\rangle$.
Hint: try to prove this using the pigeonhole principle.

(b) Let $\boldsymbol{x}$ be sampled uniformly at random from $\{\pm 1\}^n$. Use Chebyshev's inequality to prove that:

$$\Pr[|\langle L, \boldsymbol{x}\rangle| > 10M\sqrt{n}] \leq \frac{1}{100}.$$

(c) Let $\boldsymbol{Y}$ be a random variable that is equal to 1 with probability $\frac{1}{k}$ and 0 otherwise. Let $Y_1, \ldots, Y_{2k\ln r}$ be $2k\ln r$ independent copies of $\boldsymbol{Y}$. Prove:

$$\Pr[\boldsymbol{Y}_1 + \cdots + \boldsymbol{Y}_{2k\ln r} < 2] \leq \frac{2}{r}.$$

Hint: you may use the fact that $\left(1 - \frac{1}{k}\right)^k \leq \frac{1}{e}$ *without proof.*

(d) Give a randomized algorithm that runs in time $O(Mn^{3/2}\log(n))$ and with probability $1 - o_n(1)$ outputs distinct $x_1, x_2 \in \{\pm 1\}^n$ such that $\langle L, x_1\rangle = \langle L, x_2\rangle$.

(e) (Fun bonus question worth no points) Can you improve the runtime in the previous part to $O(\sqrt{M}n^{5/4}\log n)$?