UC Berkeley—CS 170 | Problem Set 7
Lecturer: Michael Jordan | Due on October 27 at 3:30 p.m.

Problem Set 7 for CS 170

Problem 1 [Graph Coloring]

Suppose that we are given a graph $G = (V, E)$, and we want to color each node with one of three colors, even if we aren't necessarily able to give different colors to every pair of adjacent nodes. Rather, we say that an edge (u, v) is *satisfied* if the colors assigned to u and v are different.

Consider a 3-coloring that maximizes the number of satisfied edges, and let c^* denote this number. Give a randomized polynomial-time algorithm for which the expected number of edges it satisfies should be at least $\frac{2}{3}c^*$.

Problem 2 [Faster Min-Cut]

The randomized algorithm for min-cut that we gave in class ran in $O(n^4)$ time. In this problem, we will develop another min-cut algorithm that is based on the same ideas but has better asymptotic running time.

Consider the following algorithm:

`F`(G):
1: do $|V|/2$ times:
2:     pick $(u, v) \in E$ with probability proportional to $\mathrm{wt}(u, v)$
3:     contract (u, v), updating V and E
4: return `Trial`(G)

`Trial`(G):
1: if $|V| < 16$:
2:     *use the* $O(n^4)$ *algorithm given in class*
3: else:
4:     for $i := 1, 2, 3, 4$:
5:         $C_i :=$ `F`(G)
6:     return the C_i of least weight

(a) Write a recurrence relation for the worst-case running time of `Trial` in terms of n, where $n = |V|$.

(b) Solve your recurrence relation from (a) to get the asymptotic worst-case running time of `Trial`. Use $\Theta()$ notation.

(c) Let C be a min-cut of $G = (V, E)$. Consider the probability that steps 1–3 of `F`(G) never do a contraction on an edge crossing between C and $V \setminus C$. Show that this probability is at least $\frac{n-2}{4n-4}$.

(d) Let $p(n)$ denote the probability that `Trial`(G) finds the min-cut of G with n vertices. Show how to derive this recurrence relation:

$$p(n) \geq 1 - \left(1 - \frac{n-2}{4n-4} \cdot p(n/2)\right)^4.$$

(e) It turns out that the solution to the recurrence relation in (d) is $p(n) = \Omega(\frac{1}{\lg n})$. (You do *not* need to show this!)

Using this fact, describe an efficient algorithm that returns a min-cut of G with probability at least $1 - \frac{1}{2^{100}}$. What is its running time?

Problem 3 [Most Likely Partition]

Let $G = (V, E)$ be an undirected graph and $p : E \to [0, 1]$ be a weight function on the edges. G represents a communication network. For $(u, v) \in E$, $p(u, v)$ is the probability that the link between u and v fails on May 11, 2003. (May 11 is Mother's Day—the busiest day of the year for the public phone network!) Assume that the links fail independently.

We say that a set $S \subseteq E$ of edges *partitions* G if its removal disconnects G (i.e., if $G' = (V, E \setminus S)$ is disconnected). If $S \subseteq E$, let $p(S)$ be the probability that all edges in S fail on Mother's Day.

(a) Let $f(n)$ be the maximum number of partitions for a graph on n vertices. Is there some constant $c > 0$ such that $f(n) = O(n^c)$? Justify your answer informally.

(b) We say that a set $S \subseteq E$ that partitions G is a *most likely partition* for G if $p(S)$ is maximal among all sets that partition G. Give an $O(n^4)$ algorithm that computes a most likely partition for G.

Your algorithm may have a small chance of providing an erroneous answer, as long as the probability of erring is negligibly small.

Problem 4 [2-Universal Hash Functions]

(a) Consider the following family H of hash functions mapping $\{1, 2, 3, 4\}$ into $\{0, 1\}$. The family H contains the three functions h_1, h_2, and h_3 defined below. Is H a family of 2-universal hash functions? Justify your answer.

$$\begin{array}{lll} h_1(1) = 0, & h_2(1) = 1, & h_3(1) = 1 \\ h_1(2) = 1, & h_2(2) = 0, & h_3(2) = 1 \\ h_1(3) = 1, & h_2(3) = 1, & h_3(3) = 0 \\ h_1(4) = 0, & h_2(4) = 0, & h_3(4) = 0 \end{array}$$

(b) Suppose m is a prime and consider hash functions $h(x, y)$ that map a pair of keys each in $[0, m-1]$ to a value in $[0, m-1]$. Let H be the family of m^2 functions $\{h(x, y) = ax + by \ (mod\ m) : a, b \in [0, m-1]\}$ parameterized by a and b. Is H a family of 2-universal hash functions? Justify your answer.

(c) Now suppose m is 2^k for some $k > 1$. With this choice of m is the H of (4b) above a 2-universal family of hash functions? Justify your answer.

Problem 5 [Collisions]

Let H be a family of 2-universal hash functions whose output is the set $\{1, 2, \ldots, m\}$. Suppose we hash m different inputs, and let N_i denote the number of inputs that hash to i. Prove that $\mathbb{E}\left[N_1^2 + \ldots + N_m^2\right] \le 10m$. (Here $\mathbb{E}[X]$ denotes the expected value of X.)