高级算法 (Fall 2019)/Problem Set 1

- 作业电子版于2019/10/8 23:59 之前提交到邮箱 njuadvalg@163.com
- 每道题目的解答都要有完整的解题过程。中英文不限。

Contents

Problem 1

Modify the Karger's Contraction algorithm so that it works for the *weighted min-cut problem*. Prove that the modified algorithm returns a weighted minimum cut with probability at least $\dfrac{2}{n(n-1)}$. The weighted min-cut problem is defined as follows.

- **Input**: an undirected weighted graph $G(V, E)$, where every edge $e \in E$ is associated with a positive real weight w_e;
- **Output**: a cut C in G such that $\sum_{e \in C} w_e$ is minimized.

Problem 2

Let $G = (V, E)$ be a graph, where $n = |V|$ and $m = |E|$. In class, we generated a random cut $\{S, T\}$ by sampling $X_v \in \{0, 1\}$ uniformly and independently for each $v \in V$ and constructing $S = \{v \in V \mid X_v = 1\}$ and $T = \{v \in V \mid X_v = 0\}$. Now, consider an alternative way to generate the random cut $\{S, T\}$. Suppose that n is an even number. Define a collection of subset as

$$\mathcal{F} = \{H \subseteq V : |H| = n/2\}.$$

We sample a random subset $S \in \mathcal{F}$ uniformly at random and construct $T = V \setminus S$.

- Give the expected size $|E(S, T)|$ of such random cut.
- Let $\mathcal{R}(\cdot)$ denote such a random source that given any $0 \le p \le 1$, $\mathcal{R}(p)$ returns an independent random sample $X \in \{0, 1\}$ such that $\Pr[X = 1] = p$. Give an algorithm that uses $\mathcal{R}(\cdot)$ as a subroutine to generate random subset $S \in \mathcal{F}$ uniformly at random. Prove the correctness of your algorithm. Analyze the number of times that the random sources is called by the algorithm.

Problem 3

Two *rooted* trees T_1 and T_2 are said to be **isomorphic** if there exists a bijection ϕ that maps vertices of T_1 to those of T_2 satisfying the following condition: for each *internal* vertex v of T_1 with children $u_1, u_2, \ldots, u_k$, the set of children of vertex $\phi(v)$ in T_2 is precisely $\{\phi(u_1), \phi(u_2), \ldots, \phi(u_k)\}$, no ordering among children assumed.

Given an efficient randomized algorithm with bounded one-side error (false positive), for testing isomorphism between rooted trees with n vertices. Analyze your algorithm.

Problem 4

Design a randomized algorithm to decide if an integer sequence $a_1, \ldots, a_n$ is a permutation of another integer sequence $b_1, \ldots, b_n$. Give upper bounds on the time complexity and the error probability.

Problem 5

Let $X_1, X_2, \ldots, X_n$ be n random variables, where each $X_i \in \{0, 1\}$ follows the distribution μ_i. For each $1 \le i \le n$, let $\rho_i = \mathbb{E}[X_i]$ and assume $\rho_i \ge \frac{1}{2}$. Consider the problem of estimating the value of

$$Z = \prod_{i=1}^{n} \rho_i.$$

For each $1 \le i \le n$, the algorithm draws s random samples $X_i^{(1)}, X_i^{(2)}, \ldots, X_i^{(s)}$ independently from the distribution μ_i, and computes

$$\hat{\rho}_i = \frac{1}{s} \sum_{j=1}^{s} X_i^{(j)}.$$

Finally, the algorithm outputs the product of all $\widehat{Z}_i$:

$$\widehat{Z} = \prod_{i=1}^{n} \hat{\rho}_i.$$

Express s as a function of n, ε, δ so that the output $\widehat{Z}$ satisfies

$$\Pr\left[e^{-\varepsilon} Z \le \widehat{Z} \le e^{\varepsilon} Z \right] \ge 1 - \delta.$$

Try to make s as small as possible.

Problem 6

In Balls-and-Bins model, we throw m balls independently and uniformly at random into n bins. We know that the maximum load is $\Theta\left(\dfrac{\log n}{\log \log n} \right)$ with high probability when $m = \Theta(n)$. The two-choice paradigm is another way to throw m balls into n bins: each ball is thrown into the least loaded of two bins chosen independently and uniformly at random(it could be the case that the two chosen bins are exactly the same, and then the ball will be thrown into that bin), and breaks the tie arbitrarily. When $m = \Theta(n)$, the maximum load of two-choice paradigm is known to be $\Theta(\log \log n)$ with high probability, which is exponentially less than the maxim load when there is only one random choice. This phenomenon is called *the power of two choices*.

Here are the questions:

- Consider the following paradigm: we throw n balls into n bins. The first $\dfrac{n}{2}$ balls are thrown into bins independently and uniformly at random. The remaining $\dfrac{n}{2}$ balls are thrown into bins using the two-choice paradigm. What is the maximum load with high probability? You need to give an asymptotically tight bound (in the form of $\Theta(\cdot)$).

- Replace the above paradigm to the following: the first $\dfrac{n}{2}$ balls are thrown into bins using the two-choice paradigm while the remaining $\dfrac{n}{2}$ balls are thrown into bins independently and uniformly at random. What is the maximum load with high probability in this case? You need to give an asymptotically tight bound.

- Replace the above paradigm to the following: assume all n balls are thrown in a sequence. For every $1 \le i \le n$, if i is odd, we throw i-th ball into bins independently and uniformly at random, otherwise, we throw it into bins using the two-choice paradigm. What is the maximum load with high probability in this case? You need to give an asymptotically tight bound.

-