INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

Randomized Algorithm Design: Mid Semester Examination 2018-19

Date of Examination: 29 April 2019
Session (FN/AN): End Semester Examination
Duration: 3 hours
Full Marks: 100
Subject No: CS60029
Subject: Randomized Algorithm Design
Department/Center/School: COMPUTER SCIENCE AND ENGINEERING
Specific charts, graph paper, log book etc., required: NO
Special instruction (if any): NA

Answer question 6 and any 3 from rest (total 4).

1. (i) Suppose a bank is providing one ATM card per user to its n users. For security reason, the id every ATM card is picked uniformly randomly in the range of integers $[1, N]$ for some natural number $N \in \mathbb{N}_{\geqslant 1}$. For functional reason, the bank needs the id of every card to be distinct (that is, no two cards have the same id). Prove that with $N = \mathcal{O}(n^2)$, the above procedure (that is, giving ids uniformly randomly) provides unique ids to the cards with probability at least $\frac{9}{10}$.

[12.5 Marks]

(ii) Consider a Markov chain with state space being $\{i \in \mathbb{Z} : -n \leqslant i \leqslant n\}$ where n is any positive natural number. The transition probabilities are as follows.

$$p_{i,j} = 1/2 \text{ for every } i, j \in \{i \in \mathbb{Z} : -n \leqslant i \leqslant n\}, i \neq j, j = i + 1 \text{ or } j = i - 1$$

$$p_{n,n} = p_{-n,-n} = \frac{1}{2}$$

Let $\mathcal{S} = \{-n, n\}$. Find the expected number of steps to reach any state in $\mathcal{S}$ from state 0.

[12.5 Marks]

2. (i) Let $\mathcal{G}$ be a undirected graph on n vertices with vertex degrees $d_1, \ldots, d_n$. Then prove that the size of the largest independent set of $\mathcal{G}$ is at least $\sum_{i=1}^{n} \frac{1}{d_i+1}$. *[Hint: Consider a random permutation of vertices and use method of expectation.]*

[12.5 Marks]

(ii) A hyper graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a tuple of a set of vertices and a set of hyper edges. A hyper edge is a non-empty subset of $\mathcal{V}$. A hyper graph is called k uniform for some $k \in \mathbb{N}_{\geqslant 1}$ if every hyper edge in it is a subset of size k. A hyper graph is called 2 colorable if there is a coloring of its vertices with 2 colors such that every edge sees at least 2 colors. Let $\mathcal{H}$ be a k uniform hyper graph where every edge intersects at most d other edges. Prove that, if $e(d + 1) \leqslant 2^{k-1}$, then $\mathcal{H}$ is 2 colorable.

[12.5 Marks]

3. (i) Let C_n be a cycle on a set $\mathcal{V}$ of n vertices and $\mathcal{T}$ be a tree which is an embedding of C_n. That is, for every $x, y \in \mathcal{V}$, we have $d_{C_n}(x, y) \leqslant d_{\mathcal{T}}(x, y)$ where $d_{C_n}(x, y)$ and $d_{\mathcal{T}}(x, y)$ denote the distance between x and y in C_n and $\mathcal{T}$ respectively. Then prove that there exists an edge $\{u, v\} \in \mathcal{E}[C_n]$ such that the distance between u and v in $\mathcal{T}$ is $n - 1$.

[12.5 Marks]

(ii) In the *buy-at-bulk network design* problem, the input is an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where each edge $e \in \mathcal{E}$ has a length $\ell_e \in \mathbb{R}_+$. There are k source-destination pairs (s_i, t_i) with demand d_i for $i \in [k]$. We need to send d_i units of flow from s_i to t_i. Every edge has infinite capacity. However, if we wish to send c_e units of flow through an edge $e \in \mathcal{E}$, then we need to pay $f(c_e)\ell_e$ rupees. The function f is sub-additive: $f(0) = 0$ and $f(x + y) \leqslant f(x) + f(y)$ for every $x, y \in \mathbb{R}_+$. A solution is a set of paths $p_i, i \in [k]$ where p_i is a path from s_i to t_i. The cost of a solution is $\sum_{e \in \mathcal{E}} f(c_e)\ell_e$ where c_e is the total amount of flow passing through the edge e. That is, $c_e = \sum_{i \in [k], e \in p_i} d_i$. We wish to find a minimum cost solution. Design a $\mathcal{O}(\log n)$ factor approximation algorithm for the buy-at-bulk network design problem.

[12.5 Marks]

4. (i) Alice and Bob play each other in a checkers tournament, where the first player to win four games (in total) wins the match. The players are evenly matched, so the probability that each player wins each game is $\frac{1}{2}$, independent of all other games. The number of minutes for each game is uniformly distributed over the integers in the range $[30, 60]$, again independent of other games. What is the expected time they spend playing the match?

[12.5 Marks]

(ii) A gambler plays a sequence of games. In every game, the gambler earns INR 10 with probability $\frac{1}{2}$ and loses INR 10 with probability $\frac{1}{2}$. To begin with, the gambler has INR 1000. The gambler plays until he either earns INR 500 or loses all his money. What is the probability that the gambler earns INR 500?

[12.5 Marks]

5. (i) Let $\mathcal{S}$ be a set of n points drawn uniformly at random from (within) a two dimensional disk of diameter 1. Let $\mathcal{X}$ denote the length of a shortest cycle that covers all the points in $\mathcal{S}$ (that is the length of a shortest travelling salesman tour). Then prove the following for every $\varepsilon > 0$.

$$\Pr[|\mathcal{X} - \mathbb{E}[\mathcal{X}]| \geqslant 2\varepsilon\sqrt{n}] \leqslant 2\left\{-\frac{\varepsilon^2}{2}\right\}$$

[12.5 Marks]

(ii) For every $n \in \mathbb{N}, \ell \in \mathbb{R}_{>0}$, and $\varepsilon \in (0, 1)$, prove that there existsz a set $\mathcal{S}$ of n points in $m = \mathcal{O}\left(\frac{\log n}{\varepsilon^2}\right)$ dimensional Euclidean space (that is in $\mathbb{R}^m$) such that, for every $x, y \in \mathcal{S}, x \neq y$, we have $d(x, y) \in [(1 - \varepsilon)\ell, (1 + \varepsilon)\ell]$ where $d(x, y)$ is the Euclidean distance between x and y.

[12.5 Marks]

6. (i) Prove that, for any natural number $k \geqslant 2$ and any prime number p, the hash family $\mathcal{H} = \{h_{a_0, a_1, \ldots, a_{k-1}}(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} \pmod p : a_0, a_1, \ldots, a_{k-1} \in \mathbb{F}_p\}$ is k universal.

[12.5 Marks]

(ii) Prove that, for every $n \in \mathbb{N}_{\geqslant 1}$, there can be at most $n + 1$ points in $\mathbb{R}^n$ whose pairwise distance is exactly 1. *[Hint: Let $\{x_0, \ldots, x_k\}$ be a set of points with pairwise distance being 1. Assume w.l.o.g that x_0 is origin (how?). What can we say about $< x_i, x_j >$? Prove $x_1, \ldots, x_k$ are linearly independent.]*

[12.5 Marks]