(CS1.406) Advanced Algorithms (Spring 2022)

Review Problem Set

1. Suppose we choose a permutation π of the ordered set $[n] = \{1, 2, \dots, n\}$ uniformly at random, from the space of all permutations of $[n]$. Let $L(\pi)$ denote the length of the longest increasing subsequence in permutation π. Then for large n and for some positive constant c, prove that $\mathbb{E}[L(\pi)] \geqslant c\sqrt{n}$.

2. Let Y be a non-negative integer-valued random variable with positive expectation. Then prove the following inequalities.

   (a)
   $$\mathbb{P}[Y = 0] \leqslant \frac{\mathbb{E}\left[Y^2\right] - (\mathbb{E}[Y])^2}{(\mathbb{E}[Y])^2}.$$

   (b)
   $$\frac{(\mathbb{E}[Y])^2}{\mathbb{E}[Y^2]} \leqslant \mathbb{P}[Y \neq 0] \leqslant \mathbb{E}[Y].$$

3. Suppose m balls are thrown into n bins. Give the best bound you can on m to ensure that the probability of there being a bin containing at least 2 balls is at least $1/2$.

4. Let a and b be chosen independently and uniformly at random from $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$, where n is a prime. Suppose we generate t pseudo-random numbers from $\mathbb{Z}_n$ by choosing $r_i = ai + b \mod n$ for $1 \leqslant i \leqslant t$. For any $\varepsilon \in [0, 1]$, show that there is a choice of the witness set $W \subset \mathbb{Z}_n$ such that $|W| \geqslant \varepsilon n$ and the probability that none of the r_i's lie in the set W is at least $\frac{(1-\varepsilon)^2}{4t}$.

5. *Lattice approximation problem:* We are given a $n \times n$ matrix A all of whose entries are 0 or 1. In addition, we are given a column vector p with n entries, all of which are in the interval $[0, 1]$. We wish to find a column vector q with n entries, all of which are entries in the set $\{0, 1\}$, so as to minimize $\|A(p - q)\|_\infty$. That is, minimize $\max_{i \in [n]}\{(A(p - q))_i\}$. In other words, the column vector q is an integer approximation to column vector p. Derive a bound on $\|A(p - q)\|_\infty$ assuming q were derived from p using randomized rounding.

6. Suppose we run Valiant's bit fixing scheme on an N node network in which every node is of degree d; each packet goes to the a random destination chosen uniformly from all the nodes and then to its final destination. Show that the expected number of steps for the completion of the first phase is
   $$\Omega\left(\frac{\log N}{d \log\log N} + \frac{\log N}{\log d}\right).$$

7. Let G be a graph on n vertices, with $\frac{nd}{2}$ edges. Consider the following probabilistic experiment for finding an independent set in G. Delete each vertex of G (together with its incident edges) independently with a probability of $1 - \frac{1}{d}$.

(a) Compute the expected number of vertices and edges that remain after the deletion process.

(b) From these, infer that there is an independent set with at least $\frac{n}{2d}$ vertices in any graph on n vertices with $\frac{nd}{2}$ edges.

8. Let X be a random variable with expectation 0 such that the moment generating function $\mathbb{E}[\exp(t|X|)]$ is finite for some $t \geqslant 0$. We can use the following two kinds of tail inequalities on X.

   - Chernoff bound:
   $$\mathbb{P}[|X| \geqslant \delta] \leqslant \min_{t \geqslant 0} \frac{\mathbb{E}\left[e^{t|X|}\right]}{e^{t\delta}}.$$
   - kth-moment bound:
   $$\mathbb{P}[|X| \geqslant \delta] \leqslant \frac{\mathbb{E}\left[|X^k|\right]}{\delta^k}.$$

   Show that for each δ, there exists a k such that the kth-moment bound is stronger than Chernoff bound. (**Hint:** Use taylor expansion on the moment generating function and use probabilistic method.)

9. A $n \times n$ matrix P is said to be stochastic if all entries are non-negative and for each row, its entries sum to 1. It is said to be doubly stochastic if for each column, the sum of entries is 1. If P is the transition probability matrix of a Markov chain, then show that the stationary distribution of this Markov chain is necessarily the uniform distribution.

10. Consider a random walk on an infinte line. At each step, the position of the particle is one of the integer points. At the next step, it moves to one of the neighbouring integral points with equal probability. Show that the expected distance of the particle from origin after n steps is $\Theta(\sqrt{n})$.

11. Two rooted trees T_1 and T_2 are said to be isomorphic if there exists a one-to-one onto mapping f from the vertices of T_1 to the vertices of T_2 satisfying the following condition: for each vertex v in T_1 with children $v_1, \ldots, v_k$, the vertex $f(v)$ has exactly the children $f(v_1), \ldots, f(v_k)$. Further, no ordering is assumed on the children of any of the internal nodes. Device an efficient randomized algorithm for testing the isomorphism of the rooted trees. (**Hint:** Associate a polynomial P_V with each vertex v in a tree T. The polynomials are defined recursively, the base case being the leaf verices all have $P = X_0$. An internal vertex v of height h with the children $v_1, \ldots, v_k$ has it polynomial defined to be $(x_h - P_{v_1})(x_h - P_{v_2}) \ldots (x_h - P_{v_k})$. Note that there is exactly one indeterminate for each level in the tree.)

12. Consider the problem of deciding whether two integer multisets S_1 and S_2 are identical in the sense that each integer occurs the same number of times in both sets. The problem can be solved by sorting the two sets in $O(n \log n)$ time where n is the cardinality of the multisets. Suggest a way of representing this as a problem involving verification of polynomial identity and thereby obtain an efficient randomized algorithm.

13. Let $Q(x_1, x_2, \ldots, x_n)$ be a multivariate and multilinear polynomial over $\mathbb{Z}_p$. That is, for all $1 \leqslant i \leqslant n$, the maximum exponent of x_i in Q is 1. Let $S_1, S_2, \ldots, S_n \subseteq \mathbb{Z}_p$ be arbitrary subsets each of size 2. If Q is a non-zero polynomial, show that there exists a point $(a_1, \ldots, a_n)$ in $S_1 \times S_2 \times \ldots \times S_n$ such that $Q(a_1, \ldots, a_n)$ is not zero.

14. Let $g(x) \equiv x \mod n$. For each $a \in \mathbb{Z}_p$, define the function $f_a(x) \equiv a \cdot x \mod p$, and $h_a(x) = g(f_a(x))$, and let $H = \{h_a \mid a \in \mathbb{Z}_p, a \neq 0\}$. Show that H is such that for all $x \neq y$,

$$\mathbb{P}\left[\text{There exists } h \in H \text{ such that } h(x) = h(y)\right] \leqslant \frac{2\,|H|}{n}.$$

15. A maximal matching in a graph is a matching that is not properly contained in any other matching. Use the parallel algorithm for Maximal Independent Set to device a parallel algorithm for finding a maximal matching in a graph.

16. Show that convolution of two n bit vectors $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ could be computed using Discrete Fourier Transform as follows.

$$\mathbf{a} * \mathbf{b} = F_{2n}^{-1}(F_{2n}(\mathbf{a}') \circledcirc F_{2n}(\mathbf{b}'))$$

where F_{2n} and F_{2n}^{-1} are DFT and Inverse DFT respectively on vectors of length $2n$, $\mathbf{a}'$ and $\mathbf{b}'$ are $2n$ length vectors obtained from $\mathbf{a}$ and $\mathbf{b}$ by appending n many zeros at the end, and $\circledcirc$ denoted pointwise multiplication of two vectors.

Good luck with your preparations!