Math 261B
Final Solutions

1. A treap is a binary tree whose nodes contain two values, a *key* x, and a *priority* p_x. Keys are drawn from a totally ordered set and the priorities are given by a random permutation of the keys. Without loss of generality, we assume that the set of elements is $X = \{1, 2, \ldots, n\}$. The tree is a heap according to the priorities (i.e., if x is a parent of y, then $p_x < p_y$). And the tree is a search tree according to the keys, (i.e., if a node has a key x, then its left subtree contains nodes with keys $< x$ and its right subtree contains nodes with keys $> x$). For example, if $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and $p = (3, 1, 5, 4, 8, 6, 2, 7)$, we have that $p_1 = 2, p_2 = 7, p_3 = 1$ and so on.

Tweaps allow for fast insertion, deletion and search of an element. The cost of these operations is proportional to the height of the treap. In what follows you will show that this quantity is $O(\log n)$ with high probability. The analysis boils down to the following problems on random permutations on $[n] = \{1, 2, \ldots, n\}$: Given a permutation $p : [n] \to [n]$ of the n elements, an element is *checked* if it is larger than all the elements appearing to its left in p. For instance, if

$$p = (\mathbf{3}, 1, \mathbf{5}, 4, \mathbf{8}, 6, 2, 7), \tag{1}$$

the elements that are checked are in bold. The problem is to show that the number of checks is concentrated around its expectation as described in the following subproblems:

(1a) Given a key x, let x_- be the set of elements that are smaller than or equal to x, We will use p_-^x to denote the permutation induced by p on x_-. For example, using p from (1), we have that $p_-^6 = (3, 1, 5, 4, 6, 2)$. Show that all elements of x_- that are checked in p_-^x appear along the path from the root to x in the tree.

(1b) Prove an analogous statement for the set x^+ of all elements $\geq x$ and use this to calculate exactly the number of elements from the root to x in the tree.

(1c) Denoting with X_n the number of elements that are checked for a random permutation $p : [n] \to [n]$, prove that

$$E[X_n] = 1 + \frac{1}{2} + \ldots + \frac{1}{n}.$$

(It is known that $H_n = \sum_{i=1}^{n} 1/i$, the nth *harmonic number*, is $\Theta(\log n)$.)

(1d) Let Y_i be an indicator random variable denoting whether the ith element of the permutation (starting from the left) is checked. Prove that

$$\Pr[Y_i = 1 \mid Y_{i+1} = y_{i+1}, \ldots, Y_n = y_n] = \frac{1}{i}$$

for any choice of the y's.

(1e) Show that for any index set S,

$$\Pr\left[\bigwedge_{i\in S}(Y_i = 1)\right] \leq \prod_{i\in S}\Pr[Y_i = 1]. \tag{2}$$

(1f) Prove that under the condition (2) the Chernoff bound holds for $Y = \sum_{i=1}^{n} Y_i$. Using this, give a concentration result for X_n.

Solution: (1a) Fix some x and consider p_x^-. Let c be some checked element of p_x^-. Clearly if $x = c$, then c is on the path from x to the root. Thus we may assume that $x \neq c$. In order to show that c is on the path from x to the root, it suffices to show that x is a descendant of c. Suppose not, then there is some proper ancestor c' of c, such that either $c' = x$ or such that x is a descendant of c' in the subtree not containing c. If $x = c'$, then the $p_x = p_{c'} < p_c$, but as x is the largest element in p_x^- this contradicts that c is checked. Thus x must be a proper descendant of c'. We have either $c < c' < x$ or $x < c' < c$. In the first case c' is in p_x^- and c' has higher priority than c, contradicting that c is checked. We can not have the second case, since neither c' or c is in p_x^-.

Solution: (1b) For p_x^+ , a similar argument as above shows that all checked elements of p_x^+ appear on the path from the root to x.

To show that every vertex on a path from x to the root r is checked in either p_x^+ or p_x^-, we first note that r is clearly checked in one of these, as it is the highest priority element. Now each of the subtrees are treaps with priorities given by p_{r-1}^- or p_{r+1}^+. Furthermore, other than r, the checked vertices are the same in $\left(p_{r-1}^-\right)_x^-$ and p_x^-, and similarly for $\left(p_{r+1}^+\right)_x^+$ and p_x^+. Thus by induction every vertex on the path from r to x is checked in p_x^+ or p_x^-. Thus the number of checked vertices is one more than total number of vertices on the path from x to r.

Solution: (1c) We proceed by induction. Clearly $X_1 = 1$ as there is only one permutation of one element. Now consider a random permutation τ on $\{2, 3, \ldots, n\}$ and insert 1 randomly in one of the n possible positions. The element 1 is checked if and only if it is inserted in the first position, which occurs with probability $\frac{1}{n}$. Thus the expected number of checked elements is $\frac{1}{n} + \mathbb{E}\left[X_{n-1}\right] = \sum_{i=1}^{n}\frac{1}{i}$.

Solution: (1d) Let $a_1 < a_2 < \cdots < a_i$ be the elements occurring in the first i positions. Independently of the configuration of the elements in the final $n-i$ positions, the position i is checked if and only if a_i is in position i. But the conditional distribution is independent of the choice of $a_1 < \cdots < a_i$, and thus the probability is $\frac{1}{i}$.

Solution: (1e) We proceed by induction. Clearly if $|S| = 1$, the bound holds.

Thus suppose that $|S| \geq 2$ and s is the least element of S. Then

$$\begin{aligned}
\mathbb{P}(\wedge_{i \in S} Y_i = 1) &= \mathbb{P}\big(Y_s = 1 \mid \wedge_{i \in S - \{s\}} Y_i = 1\big)\, \mathbb{P}\big(\wedge_{i \in S-\{s\}} Y_i = 1\big) \\
&\leq \mathbb{P}\big(Y_s = 1 \mid \wedge_{i \in S - \{s\}} Y_i = 1\big) \prod_{i \in S - \{s\}} \mathbb{P}(Y_i = 1) \\
&\qquad\qquad\qquad\qquad\qquad \text{by induction} \\
&= \frac{1}{s} \prod_{i \in S - \{s\}} \mathbb{P}(Y_i = 1) \\
&= \prod_{i \in S} \mathbb{P}(Y_i = 1)\,.
\end{aligned}$$

Solution: (1f) Note that from (1d), we have $\mathbb{P}(Y_i = 1) = \frac{1}{i}$ for all i. Thus $\mathbb{P}(Y_i = 0) = 1 - \frac{1}{i}$ for all i. As an extension of part (1e), we can derive in a similar way that $\mathbb{P}\big(\wedge_{i \in S} Y_i = 1, \wedge_{j \notin S} Y_j = 0\big) \leq \prod_{i \in S} \mathbb{P}(Y_i = 1) \prod_{j \notin S} \mathbb{P}(Y_j = 0)$.

Now for any $t > 0$, we have

$$\begin{aligned}
\mathbb{P}(Y \geq (1+\delta)\mathbb{E}[Y]) &= \mathbb{P}\Big(e^{tY} \geq e^{t(1+\delta)\mathbb{E}[Y]}\Big) \\
&\leq \frac{\mathbb{E}\left[e^{tY}\right]}{e^{t(1+\delta)\mathbb{E}[Y]}} \\
&= \frac{\sum_{S \subseteq [n]} \mathbb{P}\big(\wedge_{i \in S} Y_i = 1, \wedge_{j \notin S} Y_j = 0\big)\, e^{t|S|}}{e^{t(1+\delta)\mathbb{E}[Y]}} \\
&\leq \frac{\sum_{S \subseteq [n]} \prod_{j \notin S} \mathbb{P}(Y_j = 0) \prod_{i \in S} \big(\mathbb{P}(Y_i = 1)\, e^t\big)}{e^{t(1+\delta)\mathbb{E}[Y]}} \\
&= \frac{\prod_i \left(\mathbb{P}(Y_i = 0) + \mathbb{P}(Y_i = 1)\, e^t\right)}{e^{t(1+\delta)\mathbb{E}[Y]}} \\
&= \frac{\prod_i \mathbb{E}\left[e^{tY_i}\right]}{e^{t(1+\delta)\mathbb{E}[Y]}}.
\end{aligned}$$

We note that $\mathbb{E}\left[e^{tY_i}\right] = 1 - \frac{1}{i} + \frac{1}{i} e^t = \frac{1}{i}(e^t - 1) + 1 \leq e^{\frac{e^t - 1}{i}}$, where the last inequality comes from observe that $1 + x \leq e^x$ for all x. Thus

$$\mathbb{P}(Y \geq (1+\delta)\mathbb{E}[Y]) \leq \frac{\prod_i e^{\frac{e^t-1}{i}}}{e^{t(1+\delta)\mathbb{E}[Y]}} = \frac{e^{\left(e^t - 1\right)\sum_i \frac{1}{i}}}{e^{t(1+\delta)\mathbb{E}[Y]}} = e^{\left(e^t - (1+\delta)t - 1\right)\mathbb{E}[Y]}.$$

Letting $t = \ln(1+\delta)$, we have

$$\mathbb{P}(Y \geq (1+\delta)\mathbb{E}[Y]) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{\mathbb{E}[Y]}.$$

In the other direction, we have that for $t > 0$

$$\begin{aligned}
\mathbb{P}(Y \leq (1-\delta)\mathbb{E}[Y]) &= \mathbb{P}\Big(e^{-tY} \geq e^{-t(1-\delta)\mathbb{E}[Y]}\Big) \\
&\leq \frac{\mathbb{E}\left[e^{-tY}\right]}{e^{-t(1-\delta)\mathbb{E}[Y]}} \\
&= \frac{\sum_{S\subseteq[n]} \mathbb{P}\big(\wedge_{i\in S} Y_i = 1, \wedge_{j\notin S} Y_j = 0\big)\, e^{-t|S|}}{e^{-t(1-\delta)\mathbb{E}[Y]}} \\
&\leq \frac{\sum_{S\subseteq[n]} \prod_{j\notin S} \mathbb{P}(Y_j = 0) \prod_{i\in S} \left(\mathbb{P}(Y_i = 1)\, e^{-t}\right)}{e^{-t(1-\delta)\mathbb{E}[Y]}} \\
&= \frac{\prod_i \left(\mathbb{P}(Y_i = 0) + \mathbb{P}(Y_i = 1)\, e^{-t}\right)}{e^{-t(1-\delta)\mathbb{E}[Y]}} \\
&= \frac{\prod_i \mathbb{E}\left[e^{-tY_i}\right]}{e^{-t(1-\delta)\mathbb{E}[Y]}}.
\end{aligned}$$

Again, observing that $1 + x \leq e^x$, we have $\mathbb{E}\left[e^{-tY_i}\right] \leq e^{\frac{1}{i}(e^{-t}-1)}$, and thus

$$\mathbb{P}(Y \leq (1-\delta)\mathbb{E}[Y]) \leq e^{\mathbb{E}[Y]\left(e^{-t}-1+t(1-\delta)\right)} \leq e^{\mathbb{E}[Y]\left(1-t+\frac{t^2}{2}-1+t(1-\delta)\right)} = e^{\mathbb{E}[Y]\left(\frac{t^2}{2}-\delta t\right)}.$$

Choosing $t = \frac{\delta}{2}$, gives $\mathbb{P}(Y \leq (1-\delta)\mathbb{E}[Y]) \leq e^{-\frac{\delta^2 \mathbb{E}[Y]}{2}}$. As $Y = X_n$, this implies that there are constants $c_1, c_2, c_3 > 0$ such that $\mathbb{P}(c_1 \ln(n) \leq X_n \leq c_2 \ln(n)) \geq 1 - n^{-c_3}$.

2. The following type of geometric random graphs arises in the study of power control for wireless networks. We are given n points distributed uniformly at random within the unit square. Each point connects to the k-closest points. Let us denote the resulting (random) graph as G_k^n.

(2a) Show that there exists a contant α such that if $k \geq \alpha \log n$, then G_k^n is connected with probability at least $1 - 1/n$.

(2b) Show that there exists a constant β such that if $k \leq \beta \log n$, then G_k^n is not connected with positive probability.

Solution: (2a) Let T be an integer and consider partitioning the unit square into T^2 squares, each $\frac{1}{T}$ on a side. The probability that any particular point lands in a fixed square is $\frac{1}{T^2}$, thus the expected number of points in a square is $\frac{n}{T^2}$ and by Chernoff bounds the probability there is some square that has less than $\frac{n}{2T^2}$ points or more than $\frac{3n}{2T^2}$ points is at most $2T^2 e^{-\frac{n}{24T^2}} = 2e^{2\ln(T)-\frac{n}{24T^2}}$. If $T \leq \sqrt{\frac{n}{48\ln(2n)}}$, then this is at most $\frac{1}{n}$.

Now consider two adjacent squares. The maximum distance between any two points in these squares is $\frac{\sqrt{5}}{T}$. Now from any square S there are at most 48 other squares that contain a point of distance at most $\frac{\sqrt{5}}{T}$ from S. Thus if every square has at least one vertex and at most $\frac{k}{49}$ vertices, then for every square S, the vertices inside the square are connected to each other, and there is a vertex connected to a vertex in a neighboring square. That is, if every square has at most $\frac{k}{49}$ vertices and at least one vertex, then the graph is connected. But by the above each square has at most $\frac{3}{2}\frac{n}{T^2} \leq \frac{3}{2}48\ln(2n) = 72\ln(2n)$ vertices. Thus, letting $k \geq 3528\ln(2n)$ suffices.

Solution: (2b) The basic idea behind it is pretty straight forward but some very significant probabilistic analysis is necessary. See the paper https://netfiles.uiuc.edu/prkumar/www/ps_files/connect.pdf.

3. Consider the following parallel and distributed vertex-coloring algorithm. Every vertex u in a graph G initially has a list of colors $L_u = [\Delta(G) + 1]$ where $\Delta(G)$ denotes the maximum degree for vertices in G. The algorithm is in rounds. In each round the following happens.

- Every vertex not yet colored wakes up with probability 1/2.
- Every vertex that woke up picks a tentative color uniformly at random from its own list of colors.
- If t_u is the tentative color picked by u, and no neighbor of u picked t_u, then u colors itself with t_u.
- The color list of each uncolored vertex in the graph is updated by removing all colors successfully used by its neighbors.
- All uncolored vertices go back to sleep.

(3a) Show that the algorithm computes a legal coloring.

(3b) Show that in every round, each uncolored vertex colors itself with probability at least 1/4.

(3c) Show that within $O(\log|V(G)|)$ many rounds, the graph will be colored with high probability.

Solution: (3a) It is clear that at any stage no adjacent pair of vertices are colored the same color. Thus it suffices to show that after every step the partial coloring may be extended to a full coloring of the graph. To that end consider greedily coloring every vertex in an arbitrary order. For any vertex in the greedy coloring process, the set of forbidden colors is at most Δ (since the maximum degree is at most Δ), thus every vertex has at least one color choice available to it for any partial coloring.

Solution: (3b) When considering whether a vertex colors itself we may assume that the waking status and tentative color of all other vertices have been decided. Thus, suppose that a vertex v has $0 \leq d \leq \Delta$ uncolored neighbors and w of these neighbors have woken up. Then the probability that v colors itself is at least $\frac{d+1-w}{d+1}$. Hence the probability a vertex colors itself is at least

$$\frac{1}{2}\sum_{w=0}^{d}\binom{d}{w}2^{-d}\left(1-\frac{w}{d+1}\right) = \frac{1}{2}\left(1-\sum_{w=0}^{d}\binom{d}{w}2^{-d}\frac{w}{d+1}\right) = \frac{1}{2}\left(1-\frac{d}{2(d+1)}\right) = \frac{1}{4}.$$

Solution: (3c) Let $n = |V(G)|$. We use the stochastic recurrence relation method of Karp. Let $T(n)$ be the total number of rounds necessary to color n vertices and let $H(n)$ be the random variable representing the number of uncolored vertices from a round where there were n vertices. $T(n)$ satisfies the recurrence $T(n) = 1 + T(H(n))$. Observing that by part (b), $\mathbb{E}[H(n)] \leq \frac{3}{4}n$, we consider the deterministic recurrence $u(n) = 1 + u(3/4n)$ with $u(x) = 0$ if $x < 1$. Thus $u(n) = \left\lceil \frac{\ln(n)}{\ln(\frac{4}{3})} \right\rceil$. Thus the probability that it takes more than $\left\lceil \frac{\ln(n)}{\ln(\frac{4}{3})} \right\rceil + t$ rounds is at most $\frac{3}{4}^t$ for each node, and so with high probability it the algorithm terminates in $\mathcal{O}(\ln(n))$ rounds.

4. Let $f(x_1, \ldots, x_n)$ be a Lipschitz function with constant c. Namely, changing any coordinate changes the value of f by at most c. Let σ be a permutation of $[n]$ chosen uniformly at random. Show a strong concentration for $f(\sigma(1), \ldots, \sigma(n))$.
Solution: (4) We first consider the form of $\mathbb{E}[f \mid X_1, \ldots, X_{i-1}, X_i = a_i]$. To that end let $A = [n] - \{X_1, \ldots, X_{i-1}\} \cup \{a_i\}$ and note

$$\mathbb{E}[f \mid X_1, \ldots, X_{i-1}, X_i = a_i] = \sum_{\tau \in S_A} \frac{1}{(n-i)!} f(X_1, \ldots, X_{i-1}, a_i, \tau_1, \ldots, \tau_{n-i}),$$

where S_A is the set of permutations of A. Similarly,

$$\mathbb{E}[f \mid X_1, \ldots, X_{i-1}, X_i = b_i] = \sum_{\tau \in S_B} \frac{1}{(n-i)!} f(X_1, \ldots, X_{i-1}, b_i, \tau_1, \ldots, \tau_{n-i}).$$

But $S_B = (a_i, b_i)S_A$ and thus

$$\begin{aligned}
&|\mathbb{E}[f \mid X_1, \ldots, X_{i-1}, X_i = a_i] - \mathbb{E}[f \mid X_1, \ldots, X_{i-1}, X_i = b_i]| \\
&= \left| \sum_{\tau \in S_A} \frac{1}{(n-i)!} f(X_1, \ldots, X_{i-1}, a_i, \tau_1, \ldots, \tau_{n-i}) - \sum_{\tau \in S_B} \frac{1}{(n-i)!} f(X_1, \ldots, X_{i-1}, b_i, \tau_1, \ldots, \tau_{n-i}) \right| \\
&= \left| \frac{1}{(n-i)!} \sum_{\tau \in S_A} f(X_1, \ldots, X_{i-1}, a_i, \tau_1, \ldots, \tau_{n-i}) - f(X_1, \ldots, X_{i-1}, b_i, (a_i, b_i)\tau_1, \ldots, (a_i, b_i)\tau_{n-i}) \right| \\
&\leq \sum_{\tau \in S_A} \frac{1}{(n-i)!} 2c \\
&= 2c,
\end{aligned}$$

where the inequality comes from the c-Lipschitz property of f and the fact that the two evaluations of f differ in two coordinates. But then by the bounded difference version of Azuma-Hoeffding, $\mathbb{P}(|\mathbb{E}[f] - f| \geq t) \leq 2e^{\frac{-t^2}{4c^2 n}}$.