Fall 2024, CS 598: Topics in Graph Algorithms
Homework 0
Due: Never

Instructions and Policy: This homework is to give you a sense of the prerequisite material and the nature of the course. It is not for submission. You should be able to solve at least four without help. If not, this course may not be appropriate for you at this point. I encourage all students to solve the problems informally even though they are not for credit. Some of the concepts and ideas will appear in lectures. And they may appear in other homeworks or on the short exams.

Problem 1 Let $G = (V, E)$ be a directed graph. Let $R \subseteq V$ be a set of reward nodes. Given $s, t \in V$ and an integer k, describe a linear-time algorithm to decide whether there is an s-t *walk* that visits at least k reward nodes.

Problem 2 Let $G = (V, E)$ be a directed graph and let $\ell : E \to \mathbb{R}$ be edge lengths (which can be negative). A directed cycle C in G is a negative length cycle if $\sum_{e \in C} \ell(e) < 0$. We say that a function $\pi : V \to \mathbb{R}$ is a valid potential if for each edge $(u, v) \in E$ we have $\pi(v) \leq \pi(u) + \ell(u, v)$.

- Prove that if G has a negative length cycle than there is no valid potential.
- Prove that if there is a valid potential then there is no negative length cycle in G.
- How can you use the Bellman-Ford algorithm to obtain a valid potential if G does not have a negative length cycle?
- Suppose you have computed a valid potential π in G. For each edge (u, v) define the reduced cost/length of (u, v) to be $\ell'(u, v) = \ell(u, v) + \pi(u) - \pi(v)$ which is non-negative. Suppose you want to find shortest path lengths from a node s to every other node. Show how you can compute shortest paths using the reduced costs via Dijkstra's algorithm and obtain the actual shortest paths in G. Thus, once we have potentials, finding single-source shortest paths can be reduced to the setting of non-negative edge lengths.

Problem 3 Let $G = (V, E)$ be an undirected graph. We use $\delta(A)$ to denote that set of edges with one end point in A and the other in $V - A$. For two disjoint sets of vertices $A, B \subset V$ we use $E(A, B)$ to denote the set of edges with one end point in A and another in B.

- Prove that for any $A, B \subseteq V$,

$$|\delta(A)| + |\delta(B)| = |\delta(A \cap B)| + |\delta(A \cup B)| + 2|E(A - B, B - A)|$$

- A real-valued set function $f : 2^V \to \mathbb{R}$ is called submodular iff $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$ for all $A, B \subseteq V$. Prove that for any graph $G = (V, E)$ the function $f(A) = |\delta(A)|$ is submodular. Note that f is also symmetric.

- Suppose $G = (V, E)$ be a directed graph and now consider the function $f(A) = |\delta^+(A)|$ where $|\delta^+(A)|$ is the number of edges that leave A (with the tail in A and head in $V - A$). Prove that f is submodular.

- A hypergraph $G = (V, E)$ consists of a finite vertex set V and a set of hyperedges E where each hyperedge $e \in E$ is a subset of V, that is $e \subseteq V$. Graphs are the special case when each hyperedge has two vertices. For a given hypergraph G and a set $A \subseteq V$ let $\delta(A)$ denote the set of hyperedges that contain (at least) one node in A and (at least) one in $V - A$. Prove that $f(A) = |\delta(A)|$ is submodular.

Problem 4 Let $G = (V, E)$ be a directed graph with integer edge capacities, and let s, t be distinct nodes. The Ford-Fulkerson augmenting path algorithm finds a maximum flow f and proof of the maxflow-mincut theorem is based on showing the following. Suppose there is no s-t path in the residual graph G_f. Let A be the set of nodes reachable from s in the residual graph G_f. The proof shows that in this case $|\delta^+(A)| = |f|$ and hence we have optimality of the flow and the cut. It is easy to observe that a graph can have many s-t mincuts. However one can prove that there is a ***unique*** minimal s-t cut and a unique maximal mincut. To be precise we say that a set S is a minimal mincut if there is no S' such that $S' \subset S$ and S' is also a mincut. Minimality does not necessarily imply uniqueness since there can be two mincuts S, S' such that neither is a subset of the other.

- Prove that if A and B are s-t mincuts then $A \cap B$ and $A \cup B$ are also s-t minimum cuts. You can use submodularity or prove it in other ways.

- Argue that if f is a maximum flow then the set of reachable nodes from s in G_f is the unique minimal s-t cut.

- How would you find the unique maximal s-t cut given a maximum flow f?

Problem 5 Let $G = (V, E)$ be an undirected graph. Let u, v, w be three distinct nodes. Describe a linear-time algorithm to check if there is a simple path from u to v that contains w. Note that the problem in directed graphs is NP-Complete!

Problem 6 For $0 < p \leq 1$, let $\mathsf{Geom}(p)$ denote the *geometric distribution with parameter* p, so that if $\mathsf{X} \sim \mathsf{Geom}(p)$ then for $i \geq 1$ we have $\Pr[\mathsf{X} = i] = p(1-p)^{i-1}$.

Let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be independent random variables distributed according to $\mathsf{Geom}(p)$, and let $\mathsf{X} = \sum_{i=1}^{n} \mathsf{X}_i$.

1. Compute $\mathrm{E}[\mathsf{X}]$ and $\mathrm{Var}(\mathsf{X})$.

2. For $c \geq 2$, use Chebyschev's inequality to give a bound for $\Pr[\mathsf{X} \geq c\,\mathrm{E}[\mathsf{X}]]$.

3. Let $\mathsf{Bern}(p)$ denote the *Bernoulli distribution with parameter* p, so that if $\mathsf{Y} \sim \mathsf{Bern}(p)$ then $\Pr[\mathsf{Y} = 1] = p = 1 - \Pr[\mathsf{Y} = 0]$. Let $\mathsf{Y}_1, \ldots, \mathsf{Y}_t$ be independent random variables distributed according to $\mathsf{Bern}(p)$, and let $\mathsf{Y} = \sum_{j=1}^{t} \mathsf{Y}_j$. Show that $\Pr[\mathsf{X} \geq t] \leq \Pr[\mathsf{Y} \leq n]$.

   Hint: Consider the following experiment. You are given a biased coin that yields heads with probability p, and tails with probability $1 - p$. Flip the coin until you both (a) have seen n heads, and (b) have flipped at least t times in total. Using only this experiment as a probability space, construct random variables identically distributed to the X_i and Y_j. First try proving that $\Pr[\mathsf{X} > t] = \Pr[\mathsf{Y} < n]$.

4. For $c \geq 2$, use Problem 1 to give a bound for $\Pr[\mathsf{X} \geq c\,\mathrm{E}[\mathsf{X}]]$.

Problem 7 Consider the following variant of quicksort. Given an array A of n distinct numbers the algorithm picks a pivot x uniformly at random from A and computes the rank of x. If the rank of x is between $n/4$ and $3n/4$ (a *balanced* pivot), the algorithm proceeds as with standard quicksort. However, if the pivot is not balanced, the algorithm selects a new random pivot.

1. Write a formal description of the algorithm.

2. Directly prove that the expected runtime of this algorithm is $O(n \log n)$.

Problem 8 Let $G = (V, E)$ be a d-regular bipartite graph.

- Prove that G has a perfect matching.

- Prove that G can be edge-colored with d colors. By an edge-coloring we mean a coloring of edges such that no two edges that intersect at a node have the same color.

- Prove that every bipartite graph with maximum degree d can be edge-colored with d colors.

- Give a simple example of a 2-regular non-bipartite graph that cannot be edge-colored with 2 colors.