Due on October 30, 2015

Problem 1 [15 points]
Consider the following problem. Given an undirected graph $G = (V, E)$ with a non-negative weight function $w(\cdot)$ on the vertices V, a set of terminals $T \subseteq V$ and a number K, decide whether there exists a subgraph G' of G such that in G', there exists a path between every pair of terminals and the total weight of nodes in G' is at most K. Show that this problem is NP-Complete.

(**Hint:** *Reduce from the set cover problem.*)

Problem 2 [5 + 10 points]
In the optimization version of the above problem, the objective is to find such a subgraph G' with minimum total node weight. First show that you can assume, without loss of generality, that all terminals are leaves of weight 0 in G.

We now describe a greedy algorithm for this problem. Define a spider to be a tree such that at most one vertex has degree greater than 2.

i. Initially, let G_0 be the set terminals T with no edge.

ii. At the i-th step, choose a subgraph S of G such that

- S is a spider
- all leaves of S are terminals
- S has the minimum cost-benefit ratio $\frac{C(S)}{B(S)}$. Here the cost $C(S)$ is the total node weight of S and the benefit $B(S) = \alpha(G_i) - \alpha(G_i \cup S)$, where $\alpha(G)$ denotes the number of connected components in graph G.

Let $G_{i+1} = G_i \cup S$. If G_{i+1} connects all terminals, stop the algorithm and output G_{i+1}; else go to ii.

Answer the following questions about this greedy algorithm.

(a) Give a polynomial time implementation of this greedy algorithm.

(b) Show that this algorithm returns an $O(\log \kappa)$-approximate solution, where $\kappa = |T|$.

(**Hint:** *Recall the analysis of the greedy algorithm for the set cover.*)

Problem 3 [5 + 5 points]
Consider the following problem. Given a set F of facilities and a set C of clients, we want to assign every client to one of the facilities. Let a_{ij} be the *connection cost* of assigning client c_i to facility f_j. Let b_j be the *opening cost* of facility f_j, that is, if f_j is assigned at least one client, this cost will be incurred (only once). The objective is to find an assignment of the clients to the facilities that minimizes the total cost.

(a) Show that the set cover problem is a special case of this problem.

(b) Show that the set cover problem is equivalent to this problem, by reduction.

Problem 4 [15 points]

Show that Kruskal's algorithm for the minimum spanning tree problem is a primal-dual algorithm. Write the primal and dual LP, then show the equivalence (step by step) between Kruskal's algorithm and the primal-dual method on this LP.

Problem 5 [5 + 15 points]

Consider the following problem. Given a set J of jobs and a set M of machines, let p_{ij} be the time needed to process job $j \in J$ on machine $i \in M$ and let c_i be the cost of opening machine i. Given a bound K, we want to open some subset of machines whose total cost is at most K such that the makespan is minimized.

(a) Show that both set cover and minimum makespan scheduling are special cases of this problem.

(b) Write an LP for this problem. Use randomized rounding to give a bi-criteria approximation of $(O(\log n), O(1))$, where n is the number of jobs.

Problem 6 [15 points]

Obtain an FPTAS for the following problem. Given n positive integers, $a_1 < \cdots < a_n$, find two disjoint nonempty subsets $S_1, S_2 \subseteq \{1, \ldots, n\}$ with $\sum_{i \in S_1} a_i \geq \sum_{i \in S_2} a_i$ such that the ratio

$$\frac{\sum_{i \in S_1} a_i}{\sum_{i \in S_2} a_i}$$

is minimized.

(**Hint:** *First, obtain a pseudo-polynomial time algorithm for this problem. Then scale and round appropriately.*)