E0249 Approximation Algorithms

Midterm

27th Feb, 2015. 2pm to 5pm.

There are 3 pages to this exam and 100 points in all. There are choices in Part 1 and Part 2. If you attempt more than what is advised, please **explicitly mark** which ones you want graded. We will not grade all and then choose max. If you do not mark, we will grade the first 4 in part 1 and first 2 in part 2.

Good luck!

Part 0 (2 points). Write your name. Write both your instructors' names.

Part 1 (12 points each) Attempt any <u>four</u>.

1. Show that the greedy algorithm for set cover is an H_K-approximation, where K is the largest size of a set in the instance.

2. In the max-cut problem, we are given an undirected graph $G = (V, E)$ with capacities c_e on edges. The goal is to output a set $S \subseteq V$ so as to **maximize** $\sum_{e \in \delta(S)} c_e$ where $\delta(S)$ is the subset of edges with exactly one endpoint in S. Design a 2-approximation for this problem.

3. $(1, 2)$-*TSP* is TSP on n points and an added restriction that $c(i, j) \in \{1, 2\}$ for all pairs. For any $\varepsilon > 0$, find an instance of the $(1, 2)$-TSP where the Christofides algorithm done in class returns a tour of cost $\geq (3/2 - \varepsilon)OPT$.

4. In the k-median problem, we are given a set C of n points and costs c_{ij} between any pair of points. The objective is to select a subset H of these points such that $|H| \leq k$, and connect every point in $C \setminus H$ to one point in H such that the connection costs are minimized. Connecting point j to point i costs c_{ij}.
   (a) Write an LP relaxation for this problem. **(4 points)**
   (b) Write the dual of this LP relaxation. **(4 points)**
   (c) Give an example to show that the integrality gap is $\geq \rho$ for some constant $\rho > 1$. **(4 points)**

5. In the set cover problem, suppose every element is in at most f sets. Prove that the following algorithm is an f-factor approximation for the set cover problem: let x be **any** optimal solution to the natural LP relaxation for the set cover problem, and pick **all** sets with $x_i > 0$. Caution: in class, we only picked sets with $x_i \geq 1/f$. Hint: complementary slackness.

6. We are given a collection of n sets $S_1, \ldots, S_n$ where each S_i is a subset of $U = \{1, 2, \ldots, n\}$. A 2-coloring is an assignment $c : U \to \{-1, +1\}$ to the elements of U. The *discrepancy* of a 2-coloring is $\max_{j=1}^{n} \sum_{i \in S_j} c(i)$. What is the expected discrepancy of the *random* coloring which independently assigns each element $+1$ or -1 with probability $1/2$? Caution: note the **max** in the definition.

7. Suppose you are given a graph G for which you can query the degree of an arbitrary vertex in unit time. Its average degree d is unknown. For any $\varepsilon > 0$, show an algorithm to compute d^* in expected time $O(d/\varepsilon)$ such that with probability at least $3/4$, at most εn vertices have degree more than d^*.

Part 2 (25 marks each). Attempt any <u>two</u>.

1. In the makespan minimization problem, we are given n jobs J and m machines M. A job j takes time p_{ij} to run on machine i. The goal is to assign every job to some machine such that the maximum load on any machine is minimized. Formally, we have to find an assignment $\sigma : J \to M$ so as to minimize $\max_{i \in M} \sum_{j:\sigma(j)=i} p_{ij}$. In this problem you need to design a 2-approximation.

   (a) **(5 points)** Write an LP relaxation for the problem. Hint: Assume you know the optimum *value* M.

   (b) **(20 points)** Use the LP solution to get a 2-approximation. Hint: Try using the techniques done in class to get a 2-approximation for GAP.

2. In the KNAPSACK problem, we are given a bound B which is an integer, a set of n items with each item j having an integer weight w_j and an integer profit p_j. The goal is to choose a subset S of items such that $\sum_{j \in S} w_j \leq B$ and $\sum_{j \in S} p_j$ is maximized.

   (a) **(9 points)** Design a polynomial time algorithm for this problem if $p_j \leq n$ for all j. (Note B can be large).

   (b) **(16 points)** Design a polynomial time approximation scheme (PTAS) for the knapsack problem. Hint: Scale each p_j by a factor such that the new p'_js are not too large **and** the optimum value of the scaled instance isn't much smaller than that of the original instance.

3. Recall the facility location problem: F is a set of facilities with f_i the opening cost for facility i, C is the set of clients and c_{ij} is the cost of connecting client j to facility i. We need to open a set of facilities and connect clients to open facilities such that the sum of facility opening and connection costs are minimized. In this problem you have to design an $O(\log n)$-approximation (we are not assuming c_{ij}s form a metric any more). You could (but don't necessarily have to) do the following.

   (a) Write the LP relaxation for the problem. **(3 points)**

   (b) Perform the filtering step. **(6 points)**

   (c) Then try using randomized rounding. **(16 points)**

4. Consider a linear program of the following form:

$$\max\{ \quad \sum_{i=1}^{n} p_i x_i \quad \text{such that} \quad Ax \leq b, \quad 0 \leq x_i \leq 1, \forall i\}$$

where A is an $m \times n$ matrix with non-negative entries and b is an m-dimensional vector with non-negative entries. Let the sum of all the entries of any column of A be at most Δ. Let LP be the value of the above LP.

Design an algorithm that returns an integral vector $z \in \{0,1\}^n$ such that $\sum_{i=1}^{n} p_i z_i \geq LP$ **and** $Az \leq b + \mathbf{\Delta}$, where $\mathbf{\Delta}$ is an m-dimensional vector with all entries equal to Δ.

Hint: What can you say if $m < n$? If $m \geq n$, can you see how to "remove" a row without violating the corresponding constraint by much?

5. (a) **(10 points)** A matrix A is said to be *totally unimodular* if the determinant of any square sub-matrix is in $\{-1, 0, +1\}$. Prove that if A is totally unimodular, there is always an integral optimum solution to the LP for any cost vector c.

$$\min\{ \, c^\top x : \quad Ax \geq b, \quad x \geq 0\}$$

Hint: consider a bfs for the LP.

(b) **(15 points)** Consider a matrix A where for each row i there exists columns $\ell_i \leq u_i$ such that $A[i,j] = 1$ if $\ell_i \leq j \leq u_i$ and $A[i,j] = 0$ otherwise. That is, each row has a contiguous string of 1s and rest 0s. Show that A is totally unimodular.

Hint: Recall some facts about determinants – swapping columns or rows only swaps the sign. Subtracting one row from another doesn't change the determinant.