CSE 521: Design and Analysis of Algorithms — Winter 2006
Problem Set #4 — Instructor: Venkatesan Guruswami
Due on **February 24, 2006** (Friday) 5pm (deliver to Ning's mailbox).

Instructions: Same as for Problem Set 1.

Readings: Vanderbei's Linear Programming Text: Chapter 2, Chapter 5 (till Section 5); Kleinberg and Tardos, Sections 11.1, 11.3, 11.4 and 11.6.

1. (12 points)

   (a) Solve the following linear program using the simplex method (show the intermediate steps):

   Maximize $3x_2 + x_3$ subject to

$$\begin{aligned} x_1 + 2x_2 + x_3 &\leq 2 \\ 2x_1 + x_2 - x_3 &\leq -1 \\ 3x_1 + 2x_2 + x_3 &\leq 3 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

   (b) Show that the variable that becomes nonbasic in one iteration of the simplex algorithm cannot become basic in the next iteration.

2. (27 points)

   (a) (8 points) In this problem, we will identify a very useful criterion for when a linear program has an integral optimum solution. Consider a LP in standard form:
   Maximize $c^T x$ subject to

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \end{aligned}$$

   where A is an integer matrix and b is an integer vector but the vector c need not be integral). Suppose that the determinant of every square submatrix of A is 0, 1, or -1. Prove that if the linear program is bounded and has an optimal solution, then it has an optimal solution x^* whose entries are integers.
   Hint: Use the fact that there is a basic optimal solution. If B, N are the basic and nonbasic variables in such a solution, then we have the equation $C_B \tilde{x}_B + C_N \tilde{x}_N = b$ where $C = [\ A \mid I\ ]$ and $\tilde{x}$ denotes the vector that includes the slack variables as well as x.

   (b) (3 points) Consider the problem of computing a maximum matching in a bipartite graph. Express this as an integer linear program using a 0/1-valued variable x_e to indicate whether an edge e is picked in the matching or not.

(c) (5 points) Consider the linear programming relaxation of the above integer linear program, obtained by relaxing the constraint $x_e \in \{0, 1\}$ to $x_e \geq 0$. Write down the dual linear program of this LP relaxation. What optimization problem is the dual a natural relaxation of?

(d) (8 points) Prove, using Part (a), that the linear program above for maximum matching on bipartite graphs as well as the its dual both have integral optimum solutions.

(e) (3 points) What "min-max theorem" concerning bipartite graphs can you conclude by using Part (d)?

3. (9 points) For a graph $G = (V, E)$, call a subset $E' \subseteq E$ to be an edge-dominating set of G if every edge $e \in E$ shares at least one its endpoints with an edge in E' (formally, if $e = (u, v)$, then either u or v has an edge in E' that is incident upon it). A minimum edge-dominating set is an edge-dominating set with a minimum number of edges.

   (a) Prove that every graph has a minimum edge-dominating set that is also a matching.

   (b) Using the above, give a polynomial time algorithm that, on every graph, finds an edge-dominating set with at most twice as many edges as the minimum edge-dominating set.

4. (12 points) Chapter 11, Problem 12 (facility location). Consider the following problem. There is a set U of n nodes, which we can think of as users (e.g., these are locations that need to access a service, such as a Web server). You would like to place servers at multiple locations. Suppose you are given a set S possible sites that would be willing to serve as locations for the servers. For each site $s \in S$ there is a fee $f_s \geq 0$ for placing a server at that location. Your goal will be to approximately minimize the cost while providing the service to each of the customers. So far this is very much like the set-cover problem: the places s are sets, the weight of set s is f_s, and we want to select a collection of sets that covers all users. There is one extra complication: users $u \in U$ can be served from multiple sites, but there is an associated cost d_{us} for serving user u from site s. When the value d_{us} is very high, we do not want to serve user u from site s; and in general the service cost d_{us} serves as an incentive to serve customers from "nearby" servers whenever possible.

   So here is the algorithmic problem: Given the sets U, and S, and costs f and d, you need to select a subset $A \subseteq S$ to activate (at the cost of $\sum_{s \in A} f_s$), and assign each user u to the active server where it is cheapest to be served $\min_{s \in A} d_{us}$. The goal is to minimize the overall cost $\sum_{s \in A} f_s + \sum_{u \in U} \min_{s \in A} d_{us}$. Give an $H(n)$-approximation for this problem.

   (Note that if all service costs d_{us} are 0 or infinity, than this problem is exactly the set cover problem: f_s is the cost of the set named s, and d_{us} is 0 if node u is in set s, and infinity otherwise.)

   Hint: Follow an approach similar to the greedy strategy for set cover. If you open a new facility, associate it with a subset of remaining users with minimum average cost, and divide this cost equally amongst the newly covered users. If you use an already open facility s to cover a remaining user u, assign the cost d_{us} to u. This ensures that the total cost is the sum of the costs associated with all the users. Now bound this sum from above in terms of the cost of an optimal solution using an argument similar to the one for set cover.