Homework 3
Linear Program Rounding

Problem 3-1. Consider the facility location problem. Suppose we were interested in optimizing the sum of facility costs and the *squares* of the distances. The linear relaxation (where y_i denotes opening a facility at i and x_{ij} denotes if demand j is sent to facility i) is:

$$\text{Minimize} \sum_i y_i + \sum_{i,j} d(i,j)^2 x_{ij}$$

$$\begin{array}{rcll} x_{ij} & \leq & y_i & \forall i,j \\ \sum_i x_{ij} & \geq & 1 & \forall j \end{array}$$

Design an approximation algorithm for this problem. First, show that filtering works the same way as filtering for facility location. Next, show that for any setof points (i,j,k) in a metric space, $d(i,k)^2 \leq 2(d(i,j)^2 + d(j,k)^2)$. What is the approximation ratio you obtain?

Problem 3-2. This problem is based on the facility location algorithm discussed in class. In *capacitated* facility location, we have the additional constraint that each open facility can serve no more than u amount of demand. In addition, the facilities have different costs f_i. So, we have the constraints:

$$\text{Minimize} \sum_i f_i y_i + \sum_{i,j} d(i,j) x_{ij}$$

$$\begin{array}{rcll} x_{ij} & \leq & y_i & \forall i,j \\ \sum_i x_{ij} & \geq & 1 & \forall j \\ \sum_j x_{ij} & \leq & u y_i & \forall i \end{array}$$

Construct a problem instance with two possible facilities (one with cost 0 and the other with cost 1) and $u+1$ demands, such that the fractional cost is $\frac{1}{u+1}$ and the integer cost is 1. This "gap" shows that there is no hope of approximating this problem via LP rounding if we were to respect the capacity constraints. However, show that if we are allowed to use $2u$ amount of capacity at the facilities we open (while restricting the optimal fractional solution to use only u), then the optimal integer solution has cost 0, and the "gap" disappears!

Problem 3-3. Given a graph $G(V, E)$ with costs $c(e)$ on the edges, we wish to find the minimum cost set of edges to delete to make the graph bipartite (or two-colorable). A bipartite graph has the property that all cycles are of even length.

We will show how to get a $O(\log |V|)$ approximation to this problem. Construct a new graph $G'(V', E')$ as follows. For each node $v \in V$, we have vertices v_1 and v_2 in V'. For each edge $(v, w) \in E$ of cost c_{vw}, we have edges (v_1, w_2) and (v_2, w_1) in E', each with cost c_{vw}.

Consider the multicut in G' that separates the pairs (v_1, v_2) for all $v \in V$. Show that this multicut gives a feasible set of edges to delete in G so that G becomes bipartite. Show that the cost of this multicut is within a factor of 2 of the cost of the corresponding edges deleted in G. Hence, show an approximation algorithm for this problem. You may use the approximation algorithm for multicuts discussed in class as a subroutine.

Problem 3-4. The Densest Subgraph Problem

We are given a graph $G(V, E)$. The goal is to find the subgraph $H(V', E')$, where $V' \subseteq V$, and $E' \subseteq E$ is the set of edges induced by V', such that $\frac{|E'|}{|V'|}$ is maximized.

Show that the following linear program is a valid relaxation for this problem. Here $x(v)$ is a variable for each vertex, and $y(e)$ is the variable for each edge. In other words, show that for every subgraph, there exists a setting of the values of the variables so that the objective function is precisely the density of the subgraph.

$$\text{Maximize: } \sum_e y(e)$$

$$\begin{array}{rcll} y(e) & \leq & x(v_1) & \forall e = (v_1, v_2) \\ y(e) & \leq & x(v_2) & \forall e = (v_1, v_2) \\ \sum_v x(v) & = & 1 & \\ y(e) & \in & [0, 1] & \forall e \\ x(v) & \in & [0, 1] & \forall v \end{array}$$

1. For some $r > 0$, let $V(r) = \{v | x(v) \geq r\}$ and $E(r) = \{e | y(e) \geq r\}$. Show that the $E(r)$ is precisely the set of edges induced by the set of vertices $V(r)$. For this r, the density of the subgraph induced by $V(r)$ is therefore $\frac{|E(r)|}{|V(r)|}$.
2. Show that $\int_{r=0}^{1} |V(r)| dr = 1$.
3. Show that $\int_{r=0}^{1} |E(r)| dr = \sum_e y(e)$.
4. Show that this implies there exists a $p \in [0, 1]$ such that $\frac{|E(p)|}{|V(p)|} \geq \sum_e y(e)$.
5. Therefore, show that there is a subgraph whose density is at least the optimal solution value of the linear program. This implies the densest subgraph problem has an optimal algorithm by rounding the linear program, which in turn means it is polynomial time solvable!