Due: Wed, April 2, beginning of lecture

NOTE: Each problem set only counts 5% of your mark, but it is important to do your own work (but see below). Similar questions will appear on the first term test. You may consult with others concerning the general approach for solving problems on assignments, but you must write up all solutions entirely on your own. Anything else is *plagiarism*, and is subject to the University's Code of Behavior. You will receive 1/5 points for any (non bonus) question/subquestion for which you say "I do not know how to answer this question". You will receive .5/5 points if you just leave the question blank.
Advice: Do NOT spend an excessive amount of time on any question and especially not on any bonus questions (if given). If you wish to spend "free time" thinking about (say) bonus questions that is fine but you should not sacrifice the time needed for other courses.

1. Recall that by using Dijkstra's algorithm, we can efficiently compute the set of least costs paths from a given vertex s to all other vertices in a directed graph with non negative edge weights. Suppose we have a flow network $\mathcal{F} = (G, s, t, c)$ with integral capacities and suppose we are given an optimal integral flow f with $val(f) > 0$ in $\mathcal{F}$. Show how to efficiently determine (say in the time required for Dijkstra's algorithm) the least number of edges in E whose capacities can be increased (by 1) so as to increase the max flow to $val(f) + 1$.

[20 points]

2. Let $G = (V, E)$ be a directed graph with distinguished nodes $s, t \in V$. Describe a method for computing the maximum number of *node* disjoint directed paths from s to t in G.

[20 points]

3. Consider the following *2 from 4 frequency set cover problem*: We are given a collection of sets $\mathcal{S} = \{S_1, \ldots, S_m\}$ for $S_i \subseteq U$ with the property that every $u \in U$ occurs in exactly four different sets S_i in $\mathcal{S}$. There is a cost function $c : \mathcal{S} \rightarrow \mathbb{R}^{\geq 0}$ and we let c_i denote the cost of set S_i. A feasible solution is a sub-collection $\mathcal{S}' \subseteq \mathcal{S}$ such that every $u \in U$ occurs in at least two different sets S_i in $\mathcal{S}'$. The goal is to find a feasible solution $\mathcal{S}'$ so as to minimize the cost $c(\mathcal{S}') = \sum_{i:S_i \in \mathcal{S}'} c_i$.

   (a) Formulate the *2 from 4 frequency set cover problem* as a $\{0,1\}$ IP

   [10 points]

   (b) Show how to use LP relaxation + rounding to obtain a 3-approximation algorithm. Explain why your rounded solution is a feasible solution to the IP and why it provides a 3-approximation.

   [10 points]

4. Consider the following one machine TCSP problem (time constrained scheduling problem) when there are no release times (i.e. all release times are set to 0). That is, the input is a set of n jobs $J_1, \ldots J_n$ where each job is described by a triple (d_i, p_i, v_i) where d_i is the deadline, p_i is the processing time, and v_i is the value of job J_i. The goal is to maximize the sum of the values of scheduled jobs in a feasible schedule on one machine. Informally, in a feasible schedule

- all scheduled jobs must complete by their deadline
- jobs scheduled cannot overlap

[10 points] (a) Show that if a subset of jobs S can be feasibly scheduled then they can be scheduled in order of their deadlines.
Hint: Use an exchange argument to convert an arbitrary schedule to one satisfying the scheduling by deadlines constraint.

[10 points] (b) Show how to represent the "TCSP with no release times" problem by an IP. You may assume (from above) that the jobs are ordered so that $d_1 \leq d_2 \ldots \leq d_n$.