Due: Wed, November 28, 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 second 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 blankxa. Advice: Do NOT spend an excessive amount of time on any question and especially not on a bonus question. If you wish to spend "free time" thinking about (say) bonus questions that is fine but you should not sacrifice time needed for other courses.

1. Let $\mathcal{F} = (G, s, t, c)$ be a flow network with $G = (V, E)$ with integral capacites. Suppose you are given an integral max flow f. (You may assume that the positive flow edges do not contain a directed cycle.)

   (a) Suppose we are given an edge $e \in E$. Show how to efficiently determine (say in time $O(|E|)$ whether or not by reducing the capacity of e to $c'(e) = c(e) - 1$, the resulting network has max flow f' with $val(f') = val(f) - 1$.
   Hint: You may assume that breadth first serach takes time $O(|E|)$ to determine what nodes are reachable from a given node.

   [15 points]

   (b) Show how to determine in time $O(|E|)$ whether or not there is any single edge $e \in E$ such that by increasing the capacity of e to $c'(e) = c(e) + 1$, the resulting network has max flow f' with $val(f') = val(f) + 1$.

   [15 points]

2. (a) Recall the local search algorithm for the max cut problem. Complete the proof of the claim (12.6) on page 678. That is, show that by searching for a soultion $S' \in N(S)$ such that $w(S') \geq (1 + 2\epsilon/n)w(S)$, any local optimum is within a factor of $(2 + \epsilon)$ of a global optimum max cut.

   [10 points]

   (b) Use the probabilistic method to argue why there is always a cut (A, B) such that $w(A, B) \geq \frac{1}{2}W$ where $W = \sum_{e \in E} w_e$.

   [10 points]

   (c) Indicate how you would use the method of conditional expectations to de-randomize your algorithm. [10 points]

   (d) Consider the unweighted case where $w(e) = 1$ for all $e \in E$. Describe your de-randomized algorithm (in part c) as a simple efficient greedy algorithm in graph theoretic terms (i.e. without any discussion of expectations)? [10 points]

3. (a) Recall the local search algorithm for the Max2Sat problem using the potential function $\Phi(\tau) = \frac{3}{2}W(S_1)+2W(S_2)$. Complete the proof that any local optimum τ satisfies $w(\tau) \geq \frac{3}{4}W$ where $W = \sum_i w(C_i)$ is the sum of all clause weights.

10 points

Solution: As in the lecture we let τ be a truth assignment that achieves a local optimum with respect to the potential function above. To simplify the presentation, wlg rename variables (ie complementing some variables) so that such each x_i = true.

Let $P_{i,j}$ (resp $N_{i,j}$) denote the weight of clauses in S_i containing the literals x_j (resp $\bar{x}_j$).
Since τ is a local optimum, as we claimed in class the following must hold:

$$1/2N_{1,j} + 3/2N_{0,j} \leq 1/2P_{2,j} + 3/2P_{1,j}$$

summing over all j:

$$1/2\sum_j N_{1,j} + 3/2\sum_j N_{0,j} \leq 1/2\sum_j P_{2,j} + 3/2\sum_j P_{1,j}$$

$\sum P_{1,j} = \sum N_{1,j} = W(S_1)$
$\sum P_{2,j} = 2W(S_2)$
$\sum N_{0,j} = 2W(S_0)$

Therefore, $1/2W(S_1) + 3W(S_0) \leq W(S_2) + 3/2W(S_1)$

or $3W(S_0) \leq W(S_2) + W(S_1)$

The fraction of weight not satisfied is:

$W(S_0)/(W(S_0) + W(S_1) + W(S_2) \leq W(S_0)/4W(S_0) = 1/4$

(b) (Bonus question) Design and analyze a local search algorithm for exact Max3Sat with an approximation ratio as close to the $\frac{7}{8}$ bound guaranteed by the probabilistic method.

15 points

4. Consider the following one machine TCSP 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. Informally, in a feasible schedule all scheduled jobs must complete by their deadline and jobs cannot overlap.

   (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

   [10 points]

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

5. Consider the following *4 to 2 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 also a cost function $c : \mathcal{S} \rightarrow \mathbb{R}^{\geq \not{0}}$ 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 $c(\mathcal{S}') = \sum_{S_i \in \mathcal{S}'} c_i$.

(a) Formulate the *4 to 2 frequency set cover problem* as a $\{0,1\}$ IP [10 points]

Solution: The IP is as follows:

minimize $\sum_i c_i x_i$
subj to $\sum_{i:u \in S_i} x_i \geq 2$ for each $u \in U$
Note: There are exactly four variables in each of the above inequalities
$x_i \in \{0,1\}$ for $1 \leq i \leq m$

(b) Show how to use LP relaxation + rounding to obtain an 3-approximation algorithm. Explain why your rounded solution is a feasible solution to the IP and why it provides a 3-approximation. [10 points]

Solution: The LP changes the $\{0,1\}$ constraint to $0 \leq x_i \leq 1$

If $\{x_1^*, \ldots, x_m^*\}$ is an LP optimal solution we deterministically round to obtain $\{\bar{x}_1, \ldots, \bar{x}_m\}$ by setting $\bar{x}_i = 1$ if and only if $x_i^* \geq 1/3$.

We have to show that

- The rounded solution is an IP solution. This follows since if at most 1 of the 4 variables in any of the inequalities is $\geq 1/3$ (i.e. at least 3 of the 4 variables are $< 1/3$ and one variable is at most 1, and hence the sum is $< 1 + 1$), then the inequality was not satisfied by the fractional optimal.
- The cost of this solution is within a factor of 3 of the fractional optimum cost. This is obvious as every fractional variable is being multipled by at most a factor of 3 to get an integral solution.