Due: Wed, January 30, 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. This question concerns Dijkstra's shortest (i.e. least cost) paths algorithm. (See section 4.4).

   (a) Construct an example of an edge weighted directed graph having no negative cycles (but having some negative weight edges) for which Dijkstra's algorithm will not compute the optimal set of paths. [10 points]

   (b) Consider the following alternative definition of the "cost" of a path π. Define $cost(\pi) = max_{e \in \pi} \ell_e$. Show how to modify Dijkstra's algorithm so as to compute the least cost paths from s to all other vertices. Briefly justify why your algorithm correctly computes the optimal paths. (It is sufficient, for example, to set up an appropriate induction.) [10 points]

   (c) Consider the (usual) least cost paths and MST problems. Construct an example of an undirected graph with positive edge weights where the optimal solutions for these problems differ. [10 points]

2. Consider the following one machine flexible intervals scheduling problem. An interval I_j is now a triple s_j, f_j, k_j where s_j and f_j are (as before) the starting and finishing times and k_j is the "client" to which the interval belongs. A feasible schedule now is one where (as before) intervals do not intersect and for every client, at most one interval is scheduled.

   (a) Describe a greedy algorithm ALG for the one machine flexible interval scheduling problem which is a $\frac{1}{2}$ approximation algorithm. That is, for every input set $\mathcal{I}$, $|ALG(\mathcal{I})| \geq \frac{1}{2}|OPT(\mathcal{I})|$. [10 points]

   (b) Sketch a charging argument to prove that your algorithm achieves a $\frac{1}{2}$ approximation.

   Hint: The charging argument should consider *clients* in OPT-ALG. [10 points]

3. This question concerns Huffman coding.

(a) Let $p_a = p_b = p_c = p_d = 1/8; p_e = p_f = 1/4$. Construct an optimal prefix code. [5 points]

(b) Let $p_a = p_b = 1/6; p_c = p_d = 1/3$. Construct an optimal prefix code. [5 points]

(c) Suppose we have a set of symbols $a_1, \ldots, a_n$ where $p(a_i) < 1/3$ for all i and $\sum_{1 \leq i \leq n} p(a_i) = 1$. Sketch an argument to show that the Huffman tree cannot have a code word of length 1. [10 points]

4. Modify the argument in section 11.1 to show that the makespan greedy algorithm achieves an approximation ratio of $2 - \frac{1}{m}$ where m is the number of machines. That is show for every input set $\mathcal{I}, GREEDY(\mathcal{I}) \leq (2 - \frac{1}{m}) \cdot OPT(\mathcal{I})$. [10 points]