

## 1. (20 points)

Consider the following variant of the knapsack problem. As in the standard knapsack problem, the input is  $\{(v_1, w_1), \dots, (v_n, w_n); W\}$  where  $v_i$  (respectively,  $w_i$ ) is the value (resp. size) of the  $i^{\text{th}}$  item and  $W$  is the total size limit of the knapsack. All input parameters are positive integers. Now (in this variant) a feasible solution is a non-negative integer vector  $(m_1, \dots, m_n)$  such that  $m_i \leq 3$  for all  $i$  and  $\sum_i m_i w_i \leq W$ . We want to maximize  $\sum_i m_i^2 v_i$  for a feasible solution. That is, items that are used more often increase the value of the solution. We wish to derive a polynomial time DP algorithm for computing the value of an optimal solution to this knapsack problem when the value parameters satisfy  $v_i \leq n^2$  for all  $i$ . That is,

- provide a semantic array for the problem and indicate how the desired answer is obtained from this array.

SOLUTION: Let  $W[i, v] = \min\{\sum_{1 \leq j \leq i} m_j \cdot w_j : \sum_{1 \leq j \leq i} m_j^2 v_j \geq v \text{ and } m_i \leq 3\}$  or  $\infty$  is no such  $(m_1, \dots, m_i)$ .

- provide an equivalent recursively defined computational array (including any base cases).
- briefly indicate why your computationally defined array is equivalent to the semantic array given.
- What is the time complexity of your algorithm assuming that  $(\forall i)v_i \leq n^2$ ?

## 2. (20 points)

Consider the following triangulation problem. We are given a set of  $n \geq 3$  points  $x_1, x_2, \dots, x_n$  in the plane which are the vertices (in clockwise order) of a convex polygon  $P$  in the plane. We wish to triangulate the polygon  $P$  so as to minimize the sum of the lengths of the  $n - 3$  interior edges used to form the triangulation. Provide a polynomial time dynamic programming algorithm with complexity polynomial in  $n$  for computing an optimal triangulation. Provide appropriate semantic and computational (i.e. recursively defined) arrays and justify why these are equivalent. What is the complexity of your algorithm assuming that one can calculate the length of a line in constant time?

Hint: You may find it useful to consider a dynamic program in the style used for the matrix chain problem.

SOLUTION: We discussed two ways to define an appropriate semantic array. Here I will discuss one such method/array. Every node must be a vertex of a triangle (in the triangulation) and hence must be adjacent to at least one internal edge (assuming there are at least 4 points). In particular,  $x_1$  must be connected by an internal edge to some  $x_i$  with  $3 \leq i \leq n - 2$ . Our dynamic programming solution will be based on finding the best choice of that point  $x_i$ . Once that  $x_i$  is chosen (with a cost of the length of the line  $(x_1, x_i)$ ), we are left with either one remaining problem (when  $i = 3$  or  $i = n - 2$ ) or with two smaller problems otherwise. In all cases the remaining triangulation problems can be represented by giving  $x_c$ , the immediate point clockwise, and  $x_{cc}$ , the immediate point counter-clockwise) from  $x_1$  with the

understanding that all points in between are consecutive. So the semantic array is say  $V[x_c, x_{cc}] = \min$  cost of internal edges in a triangulation of  $x_1, x_c, x_{c+1}, \dots, x_{cc}$ .

3. (10 points)

For each of the following statements about integral capacity flow networks, indicate if the statement is true or false. If true justify why and if false give a counter-example.

- There is always a *unique* minimum capacity cut.
- It is always possible to reduce a positive maximum flow by one unit (of flow) by decreasing the capacity of some edge by one unit.

4. (20 points) Suppose we are given a flow network  $\mathcal{F}$  with integer edge capacities and we are also given an integral max flow  $f$  in  $\mathcal{F}$ . Suppose we want to increase the max flow by one unit by increasing the capacity of certain edges. Show how to efficiently compute a smallest set of edges  $E'$  such that increasing the edge capacity for each edge  $e \in E'$  will increase the max flow by 1 unit.

Hint: Your method should take time proportional to Dijkstra's shortest path algorithm, say  $O(m \log n)$  where  $m$  is the number of edges in the network and  $n$  is the number of nodes.

5. (20 points)

Consider the  $m$ -machine makespan problem for the restricted machines model. In this model, a job  $J_i$  is specified by a pair  $(p_i, S_i)$  where  $p_i$  is the processing time of the job and  $S_i \subseteq \{1, \dots, m\}$  is the set of allowable machine for this job. That is, job  $J_i$  can only be scheduled on a machine whose index is in  $S_i$ . Recall that the goal of the makespan problem is to minimize the latest (over all machines) finishing time. Suppose that all jobs have unit processing time (i.e.  $p_i = 1$  for all  $i$ ). Show how to reduce the makespan problem for this special case to a flow problem.

Hint 1: Use a max flow algorithm to determine whether or not a makespan with a particular value can be achieved. Hint 2: This is somewhat similar to the max bipartite matching problem.

SOLUTION: Suppose there are  $n$  jobs and  $m$  machines. For each possible minimum makespan value  $T$  (with  $1 \leq T \leq n$ ) we set up a flow network as follows.

$\mathcal{F}(T) = (G, s, t, c_T)$  with  $G = (V, E)$  defined as :

$V = \{s, t\} \cup \{J_i : 1 \leq i \leq n\} \cup \{M_j : 1 \leq j \leq m\}$  and

$E = \{(s, J_i) : 1 \leq i \leq n\} \cup \{(M_j, t) : 1 \leq j \leq m\} \cup \{(J_i, M_j) : j \in S_i\}$ . The capacity  $c_T(e)$  for all edges  $e \in \{(M_j, t) : 1 \leq j \leq m\}$  is  $T$ . All other edges have capacity 1. The claim is that  $\mathcal{F}(T)$  has a max flow of  $n$  iff the jobs  $\{J_i = (1, S_i) : 1 \leq i \leq n\}$  can be scheduled with makespan at most  $T$ . Hence we say keep trying  $T = 1, T = 2, \dots$  until we find the first  $T$  that allows for a max flow of  $n$  in  $\mathcal{F}(T)$ . (We also use binary search to find the smallest  $T$  which reduces the complexity by a factor of  $n/\log n$ .)

6. (20 points)

Show how to polynomial time reduce the following optimization problems to the associated decision problem.

- (a) Reduce the max clique optimization problem (.e. given a graph  $G$ , find a maximum size clique in  $G$ ) to the decision problem CLIQUE (i.e. given  $(G, k)$  does  $G$  have a clique of size  $k$ ).
- (b) Reduce the min colouring optimization problem (i.e. given a graph  $G$ , find a minimum colouring of the graph) to the decision problem COLOUR (.e. given  $(G, k)$  does  $G$  have a colouring using  $k$  colours). Here you are to compute the actual colouring (i.e. the function mapping each node to a colour) and not just the value of the minimum colouring).

SOLUTION: We first use the decision problem to determine the minimum colouring number,  $\chi(G)$ , for graph  $G = (V, E)$  with say  $V = \{v_1, \dots, v_n\}$ . We then keep modifying  $G$  to a sequence of graphs  $G_0, G_1, G_2, \dots, G_n$  as follows:  $G_0$  is  $G$  with the disjoint addition of a complete graph  $K_{\chi(G)}$  on  $\chi(G)$  nodes; we will let  $u_1, \dots, u_{\chi(G)}$  be these clique nodes. The intended meaning is that  $u_j$  represents colour  $j$ . We will create the graphs  $G_i$  so that they can always be coloured with  $\chi(G)$  colours so that it follows that the same colouring restricted to the nodes of  $G$  are also a valid colouring. We create  $G_i$  from  $G_{i-1}$  by essentially trying possible colours for  $v_i$ ; that is, if the intended meaning is that  $v_i$  is to be coloured  $j$  then we add the edges  $(v_i, u_k)$  for all  $k \neq j$  to the graph  $G_{i-1}$  and test (using the decision procedure) whether or not the resulting graph can still be coloured with  $\chi(G)$  colours. There will be at least one such extension of  $G_{i-1}$  so as to create  $G_i$ .