HW 10 (due Wednesday, at noon, November 28, 2018)

CS 473: Algorithms, Fall 2018 Version: **1.22**

Submission guidelines and policies as in homework 1.

1 (100 PTS.) Diminishing returns.

You are running a company with n workers – let W be this set of worker. You have to do a new project, and you would like to assign as few people to do the project as possible. The project has $\tau \leq n$ tasks that needs to be completed (τ is a positive integer). Given a subset $S \subseteq W$, you have a function $f(S)$ that tells you the number of tasks in the project this group of people can complete on their own (i.e., $f(S)$ is an integer number). Note, that $f(W) = \tau$ (i.e., if you assign everybody to the project, then it can be done in full).

We are going to assume (maybe unrealistically) that the function is monotone, that is

$$\forall X \subseteq Y \subseteq W \qquad f(X) \leq f(Y).$$

Another assumption, known is economics as ***diminishing returns***, is that the "value" of a single worker decreases as we consider bigger sets. Formally, for any $w \in W$, we have

$$\forall X \subseteq Y \subseteq W \qquad f(X + w) - f(X) \geq f(Y + w) - f(Y),$$

where $X + w = X \cup \{w\}$.

Provide a $O(\log n)$ approximation algorithm to the smallest set $Z \subseteq W$, such that $f(Z) = \tau$. Prove the quality of approximation of your algorithm. You can assume that computing f on a single set, takes $O(n)$ Time. What is the running time of your algorithm?

2 (100 PTS.) Linear programming

2.A. (10 PTS.) Show the following problem in NP-HARD.

Integer Linear Programming

> **Instance**: A linear program in standard form, in which A and B contain only integers.
> **Question:** Is there a solution for the linear program, in which the x must take integer values?

2.B. (50 PTS.) A steel company must decide how to allocate next week's time on a rolling mill, which is a machine that takes unfinished slabs of steel as input and produce either of two semi-finished products: bands and coils. The mill's two products come off the rolling line at different rates:

Bands 200 tons/hr
Coils 140 tons/hr.

They also produce different profits:

Bands \$ 25/ton
Coils \$ 30/ton.

Based on current booked orders, the following upper bounds are placed on the amount of each product to produce:

Bands 6000 tons
Coils 4000 tons.

Given that there are 40 hours of production time available this week, the problem is to decide how many tons of bands and how many tons of coils should be produced to yield the greatest profit. Formulate this problem as a linear programming problem. Can you solve this problem by inspection?

2.C. (40 PTS.) A small airline, Ivy Air, flies between three cities: Ithaca (a small town in upstate New York), Newark (an eyesore in beautiful New Jersey), and Boston (a yuppie town in Massachusetts). They offer several flights but, for this problem, let us focus on the Friday afternoon flight that departs from Ithaca, stops in Newark, and continues to Boston. There are three types of passengers:

1. Those traveling from Ithaca to Newark (god only knows why).
2. Those traveling from Newark to Boston (a very good idea).
3. Those traveling from Ithaca to Boston (it depends on who you know).

The aircraft is a small commuter plane that seats 30 passengers. The airline offers three fare classes:

1. Y class: full coach.
2. B class: nonrefundable.
3. M class: nonrefundable, 3-week advanced purchase.

Ticket prices, which are largely determined by external influences (i.e., competitors), have been set and advertised as follows:

	Ithaca-Newark	Newark-Boston	Ithaca-Boston
Y	300	160	360
B	220	130	280
M	100	80	140

Based on past experience, demand forecasters at Ivy Air have determined the following upper bounds on the number of potential customers in each of the 9 possible origin-destination/fare-class combinations:

	Ithaca-Newark	Newark-Boston	Ithaca-Boston
Y	4	8	3
B	8	13	10
M	22	20	18

The goal is to decide how many tickets from each of the 9 origin/destination/fare-class combinations to sell. The constraints are that the place cannot be overbooked on either the two legs of the flight and that the number of tickets made available cannot exceed the forecasted maximum demand. The objective is to maximize the revenue. Formulate this problem as a linear programming problem.

3 (100 PTS.) Linear programming.

Provide *detailed* solutions for the following problems, showing each pivoting stage separately.

3.A. (15 PTS.)
maximize $6x_1 + 3x_2 + 5x_3 + 9x_4$
subject to
$2x_1 + x_2 + x_3 + 3x_4 \leq 5$
$x_1 + 3x_2 + x_3 + 2x_4 \leq 3$
$x_1, x_2, x_3, x_4 \geq 0.$

3.B. (15 PTS.)
maximize $2x_1 + 4x_2$
subject to
$2x_1 + x_2 \leq 4$
$2x_1 + 3x_2 \leq 3$

$4x_1 + x_2 \leq 5$
$x_1 + 5x_2 \leq 1$
$x_1, x_2 \geq 0.$

Consider a directed graph G with source vertex s and target vertex t and associated costs $\mathrm{cost}(\cdot) \geq 0$ on the edges. Let $\mathcal{P}$ denote the set of all the directed (simple) paths from s to t in G.

Consider the following (very large) integer program:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{e \in E(\mathsf{G})} \mathrm{cost}(e) x_e \\
\text{subject to} \quad & x_e \in \{0,1\} \qquad \forall e \in E(\mathsf{G}) \\
& \sum_{e \in \pi} x_e \geq 1 \qquad \forall \pi \in \mathcal{P}.
\end{aligned}
$$

3.C. (10 PTS.) What does this IP computes?

3.D. (10 PTS.) Write down the relaxation of this IP into a linear program.

3.E. (10 PTS.) Write down the dual of the LP from (B). What is the interpretation of this new LP? What is it computing for the graph G (prove your answer)?

3.F. (10 PTS.) The strong duality theorem states the following.

> **Theorem 0.1.** *If the primal LP problem has an optimal solution $x^* = (x_1^*, \ldots, x_n^*)$ then the dual also has an optimal solution, $y^* = (y_1^*, \ldots, y_m^*)$, such that*
> $$\sum_j c_j x_j^* = \sum_i b_i y_i^*.$$

In the context of (C)-(E) what result is implied by this theorem if we apply it to the primal LP and its dual above? (For this, you can assume that the optimal solution to the LP of (B) is integral – which is not quite true – things are slightly more complicated than that.)

And now for something completely different.

3.G. (10 PTS.) Given a weighted, directed graph $G = (V, E)$, with weight function $w : E \to \mathcal{R}$ mapping edges to real-valued weights, a source vertex s, and a destination vertex t. Show how to compute the value $d[t]$, which is the weight of a shortest path from s to t, by linear programming.

3.H. (10 PTS.) Given a graph G as in **3.G.**, write a linear program to compute $d[v]$, which is the shortest-path weight from s to v, for each vertex $v \in V$.

3.I. (10 PTS.) In the *minimum-cost multicommodity-flow problem*, we are given a directed graph $G = (V, E)$, in which each edge $(u, v) \in E$ has a nonnegative capacity $c(u, v) \geq 0$ and a cost $\alpha(u, v)$. As in the multicommodity-flow problem (Chapter 29.2, CLRS), we are given k different commodities, K_1, K_2, ..., K_k, where commodity i is specified by the triple $K_i = (s_i, t_i, d_i)$. Here s_i is the source of commodity i, t_i is the sink of commodity i, and d_i is the demand, which is the desired flow value for commodity i from s_i to t_i. We define a flow for commodity i, denoted by f_i, (so that $f_i(u, v)$ is the flow of commodity i from vertex u to vertex v) to be a real-valued function that satisfies the flow-conservation, skew-symmetry, and capacity constraints. We now define $f(u, v)$, the *aggregate flow*, to be sum of the various commodity flows, so that $f(u, v) = \sum_{i=1}^{k} f_i(u, v)$. The aggregate flow on edge (u, v) must be no more than the capacity of edge (u, v).

The cost of a flow is $\sum_{u,v \in V} \alpha(u, v) f(u, v)$, and the goal is to find the feasible flow of minimum cost. Express this problem as a linear program.