Algorithms
Computer Science 140 & Mathematics 168
Spring 2007
Homework 9b
Due Tuesday, April 3

This assignment is on network flow, with one problem at the end on linear programming. Keep in mind that network flow can be used to find maximum flows and also minimum cuts. Sometimes, it's the minimum cut that we're interested in! Also keep in mind that some problems may require that you solve several network flow problems.

1. **[15 Points] Scheduling System Administrators at Giigle.** You've been hired by the web search company, Giigle. Giigle has n system administrators ("sysadmins") and must have one sysadmin on site every day of the year. The sysadmins have a regular work schedule, but dealing with holidays has not been resolved. Giigle has asked you to develop an algorithm to efficiently schedule sysadmin coverage of holidays for the next several years.

   There are k vacation periods (e.g. Thanksgiving 2008, Giigleweek 2008, Memorial Day weekend 2008, winter break 2008, Thanksgiving 2009, etc.) where vacation period i comprises d_i consecutive days, $1 \leq i \leq k$. The total number of vacation days is $D = \sum_{i=1}^{k} d_i$.

   Each sysadmin has indicated the specific vacation days that she or he is willing to work. For example, a sysadmin might be able to work on Thursday or Saturday of Thanksgiving 2007, Monday, Wednesday or Friday during the winter holiday, Sunday of Thanksgiving 2008, etc.

   Giigle would like to assign sysadmins to vacation days subject to the following constraints:

   - Every vacation day is covered by a single sysadmin.
   - Every sysadmin is only asked to work on days that she or he has indicated willingness to work.
   - A sysadmin should not be asked to work more than one day during any particular vacation period.

Your objective is to develop an algorithm to find a solution that is as equitable as possible. The *equity* of a solution is the smallest value t such that all the constraints above are met and no sysadmin works more than t vacation days total. Thus, a solution that satisfies these constraints and has each sysadmin working at most 1 vacation day is more equitable than one in which some sysadmins work 2 vacation days. etc. If no solution exists that satisfies the constraints, the algorithm should report this.

(a) Carefully describe an algorithm for solving this problem.

(b) Explain why your algorithm is correct.

(c) Derive the running time of your algorithm. The running time may be polynomial in the various parameters in the problem (e.g. number of sysadmins, number of vacation days, etc.).

2. **[15 Points] Graph Partitioning!** Given an *undirected* graph G, a *partition* of the graph is a division of the vertices into two sets A and B such that every vertex is in exactly one of A and B. The only constraint on A and B is that neither set can be empty. The *crossing number* of the partition is the number of edges with one endpoint in A and one endpoint in B. Your job is to find an algorithm that finds a partition with the smallest possible crossing number. (Fast graph partitioning algorithms are used in a variety of applications ranging from data clustering to circuit design.)

   (a) Describe your algorithm. (You may use existing algorithms to help!)

   (b) Prove that your algorithm is correct.

   (c) Derive the running time of your algorithm. It must be polynomial in the number of vertices and edges in the graph.

3. **[20 Points] Space Shuttle Profit Optimization!**

   After successful careers at Millisoft, Hurts Car Rental, the brokerage firm of Weil, Proffet, and Howe, and Giigle, you've been hired by NASA to help optimize their Space Shuttle program.

   On a given mission, NASA will consider a set of experiments that industrial sponsors would like to conduct (for example, "Does zero-gravity compromise the integrity of Spam?"). Let $E = \{E_1, E_2, \ldots, E_m\}$ denote the set of experiments under consideration. Let p_j denote the amount of money the sponsor will pay to conduct experiment E_j. The experiments use a set $I = \{I_1, I_2, \ldots, I_n\}$ of instruments to conduct the experiments. For each experiment E_j, let R_j be the subset of I that contains all of the instruments needed to conduct experiment E_j. Notice that this allows a single instrument to be used in multiple experiments. The instruments are potentially large and heavy and the cost of taking instrument I_k is c_k dollars. Your job is to determine which experiments should be performed in order to maximize the net revenue, which is the total income from the performed experiments minus the total cost of all instruments carried. Amazingly, this problem can be solved using **network flow**!

We'll construct a network flow problem as follows: The network contains a source vertex, s, vertices $I_1, I_2, \ldots, I_n$, vertices $E_1, E_2, \ldots, E_m$, and a sink vertex t. For each instrument I_k there is a directed edge from s to vertex I_k with capacity c_k (the cost of taking this instrument). For each experiment E_j there is an edge from vertex E_j to t with capacity p_j (the payment for this experiment). Finally, if instrument I_k is in set R_j (meaning that instrument I_k is needed for experiment E_j) then there is a directed edge from vertex I_k to vertex E_j with **infinite** capacity.

(a) First, try this. Consider a situation in which there are three experiments E_1, E_2, E_3 which will bring in 10, 6, and 6 dollars, respectively. Also there are four instruments I_1, I_2, I_3, and I_4 which cost 3, 2, 5, and 7 dollars, respectively to take on the shuttle. Experiment E_1 requires instruments I_1 and I_2, experiment E_2 requires instruments I_1 and I_3, and experiment E_3 requires instruments I_3 and I_4. Using brute-force, enumerate all seven possible combinations of experiments that could be taken and determine which combination is the most profitable. What is this combination and what is the net revenue?

(b) For the example problem above, construct the corresponding network flow problem. Find the maximum flow in the network (show your residual graphs at each step) and then find the corresponding cut whose capacity is equal to this flow.

(c) Let's let S denote the vertices that are on the same side of the above cut as vertex s and let T denote the vertices that are on the same side of the cut as t. What do you notice about the instruments and experiments that are in the set T?

(d) Now, let τ be the sum of the payments that NASA would receive for taking all of the possible experiments. In this case, $\tau = 10 + 6 + 6 = 22$. From τ, subtract the capacity of the cut you found above. Surprise! What is this number and how does it appear to relate to this problem?

(e) Finally, we're ready to generalize all of this into an efficient algorithm for solving the profit maximization problem in general. Assume that we've set up a network flow problem corresponding to a given set of experiments and instruments. Show that for any cut with finite total capacity, if an experiment E_j is in T (the side of the cut containing vertex t), then all of the instruments used in this experiment must **also** be in T.

(f) Now argue that the maximum net revenue that can be achieved is simply the total sum τ of the payments that would be received for taking all of the experiments minus the capacity of the minimum cut.

(g) Now just summarize all of this by describing the algorithm, step-by-step, for finding the maximum net revenue, given a set of experiments, instruments, and the corresponding payments and costs. Give a careful derivation of the worst-case running time of the algorithm assuming there are m experiments and n instruments.

4. **[15 Points] A Small Linear Program.** Consider the following optimization problem.

Minimize $x_1 - 2x_2$ subject to the following constraints:

$$\begin{aligned} 3x_1 - x_2 &\geq -1 \\ x_1 + x_2 &\leq 7 \\ x_1 &\leq 5 \\ x_2 + 5 &\leq 9 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

(a) Convert this into standard linear programming form.

(b) Draw the feasible region. Shade it in so that it is clear what is inside and what is outside the feasible region. Make sure that your coordinates are clearly labeled and that your drawing is fairly precise.

(c) Draw and label the dotted lines indicating where the standard linear program's optimization function is equal to 0, 2, and 4.

(d) Using your drawing, which values of x_1 and x_2 give the optimal solution? Explain briefly.

(e) Now solve this linear program using the Simplex Algorithm. Show all of your work. In particular, be sure to show each dictionary. Show how you found the optimal solution from the last dictionary.

(f) Finally, verify your solution by running the applet at

`www-fp.mcs.anl.gov/otc/Guide/CaseStudies/simplex/applet/SimplexTool.html`

on this linear program. What does the purple box in the applet window indicate?