Discrete Math in CS	Problem Set 9
CS 280 Fall 2005	Due December 2, 2005

Reading: Rosen, Sections 8.1, 8.2, 8.4, 8.8, 9.4; and the handout on graphs.

---

(1) Given numbers $a \geq 1$ and $b \geq 1$, we define the directed acyclic graph $G(a, b)$ as follows. It has nodes

$$s, t, v_1, v_2, \ldots, v_a, w_1, w_2, \ldots, w_b,$$

and it has edges

$$(s, v_1), (s, w_1), (v_a, t), (w_b, t),$$

$$(v_i, v_{i+1}) \text{ for each } i = 1, 2, \ldots, a - 1, \text{ and}$$

$$(w_j, w_{j+1}) \text{ for each } j = 1, 2, \ldots, b - 1.$$

So intuitively, it consists of two paths from s to t: one passing through all nodes v_i in order, and the other passing through all nodes w_j in order. As an example, Figure 1 depicts the graph $G(3, 4)$.

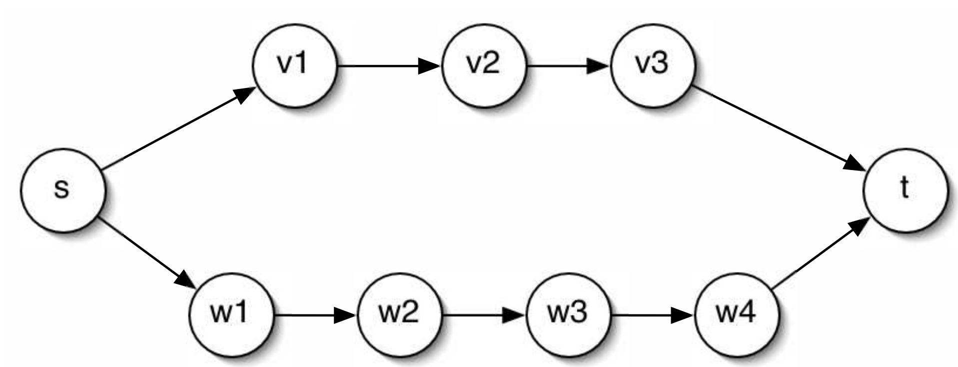

Figure 1: A drawing of the graph $G(3, 4)$.

For each choice of a and b, the directed graph $G(a, b)$ is acyclic, so it has a topological ordering. In fact, it has multiple topological orderings.

Give a formula in terms of a and b for the number of different topological orderings of $G(a, b)$, and give an explanation for your answer.

(2) Your friend Rui holds the title of Captain of Social Media for MyCrawl, and he comes to you with the following problem. The Social Media group is concerned that MyCrawl's user population is becoming too passive — or, to take his summary, that they're turning into a bunch of drones who just watch news headlines scroll across their computer screens.

They decide MyCrawl needs something to engage its users more directly, so they develop an interactive tool that lets people rank the top news stories of the day. At the end of each day, the tool offers n candidates for the top story, and each user puts these in order from most to least interesting. The analytical part of the tool then takes the ranking from each user and tries to determine a consensus on the "top" story.

The MyCrawl people had hoped it would always be clear how to synthesize all the rankings into a top choice, but in fact it's quite subtle. Actually, it's an example of the voting issues that we discussed way back in the first lecture, which you may recall are subtle enough that people have won Nobel Prizes in Economics for thinking about them.

Here's a way to view some of these problems using the language of directed graphs. We define a directed graph G with n nodes, one for each news story. Then, if a majority of users put story i ahead of story j in their ranking, we include a directed edge (i, j). We'll assume there are no ties — for each pair i, j either a majority of users put i ahead of j, or a majority put j ahead of i, so exactly one of the two edges (i, j) or (j, i) is present.

Directed graphs with this property — that for every pair of nodes i and j, exactly one of the two edges (i, j) or (j, i) is present — are called *tournaments*, since they can be viewed as expressing the outcome of an all-play-all round-robin tournament with a win-loss outcome for each pair.

So the MyCrawl people take the rankings and (because we are assuming no ties), they build a tournament G as above. Now, suppose there were a node i in G so that there was an edge (i, j) for every other node j — that is, an edge from i to each other node. Then it would be very natural to declare i to be the top story — for every other story j, a majority of people put i ahead of j. (In the language of voting theory (as in the first lecture), i is a *Condorcet winner*.)

Unfortunately, there may not be such a node i. The goal in this problem, however, is to prove that there's always something which is not much worse.

Specifically, prove that in every tournament there is a node k such that k has a directed path of length at most 2 to every other node. (That is, for every other node $j \neq k$, there is either an edge (k, j), or there exists a node j' and edges (k, j') and (j', j).)

(3) In class we discussed how graphs can be used to model social networks, capturing the relationships within a group of people. One area of study within social network analysis is *social balance theory*, which tries to identify a sense in which some patterns of friendship and antagonism are more "balanced" or "stable" than others.

Here's one of the most basic social balance models. We take an undirected graph G on n nodes, with an edge between each pair of nodes. (We call such an undirected graph a *complete graph*, since it contains all possible edges.) Each node represents a person. We then *label* each edge $\{i, j\}$ with either $+$ or $-$; a $+$ label indicates that i and j are friends, while a $-$ label indicates that i and j are enemies. (Note that since there's an edge connecting each pair, we assume that each pair of people are either friends or enemies — no two people are indifferent to one another.)

We say that the labeling is *balanced* if it has the following property:

($*$) For *every* set of three nodes in G, if we consider the three edges connecting

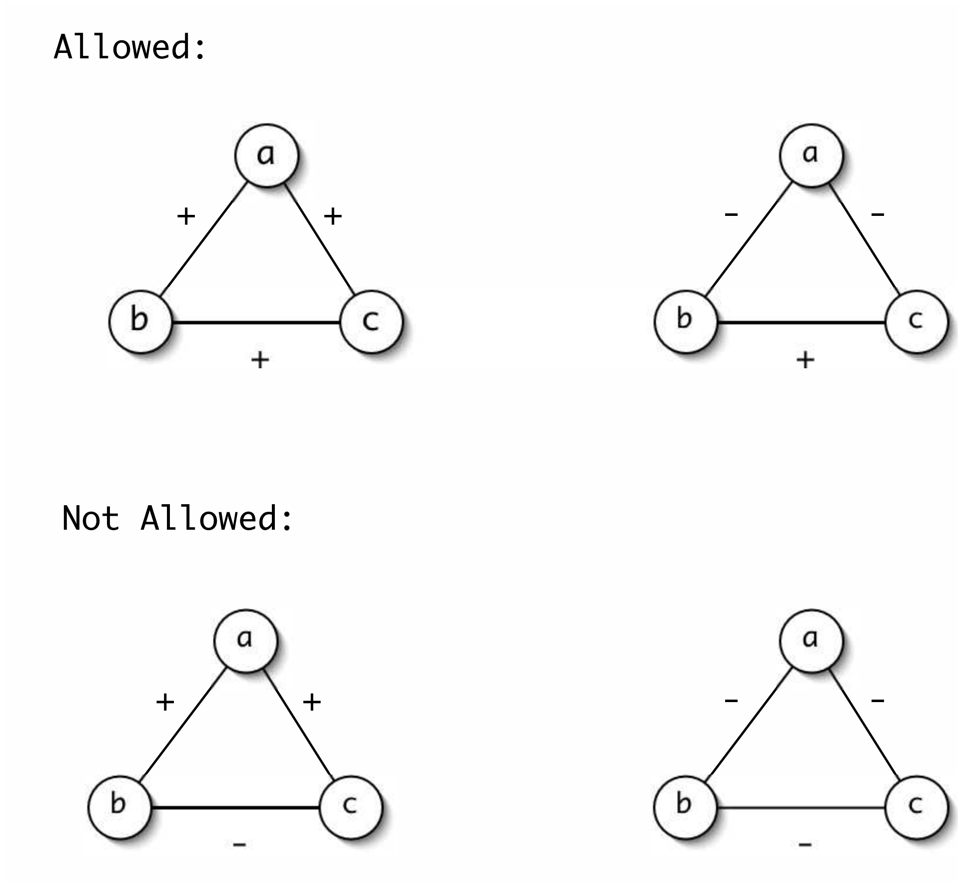

Figure 2: Social balance: Each labeled triangle must have 0 or 2 negative edges.

them, either all three of these edges are labeled +, or else exactly one of them is labeled +.

Figure 2 illustrates this definition: we allow "triangles" where the number of + labels is 1 or 3, but none where the number of + labels is 0 or 2.

For example, consider the two labeled four-node complete graphs in Figure 3. The one on the left is balanced, since we can check that each set of three nodes satisfies condition $(*)$ above. On the other hand, the one on the right is not balanced, since the sets $\{a, b, c\}$ and $\{b, c, d\}$ each have exactly two edges labeled +, in violation of condition $(*)$.

Intuitively, why do social network theorists think of such a labeling as "balanced"? Essentially, both of the configurations in the top row of Figure 2 correspond to natural social situations: three mutual friends, or two friends united against a common enemy. On the other hand, the bottom row of Figure 2 corresponds to less natural situations: a person a who is friends with each of two enemies b and c; or three mutual enemies. Cognitive psychologists have argued that people generally find the latter two kinds of situations more "dissonant" and strive to minimize them in their personal relationships. (For example, if you were friends with two mutual enemies, you might try to influence them to become more

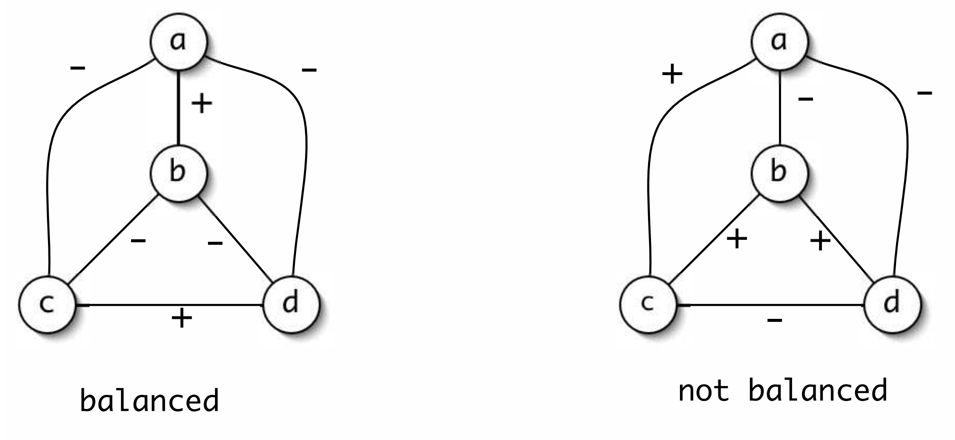

Figure 3: The labeling of the four-node complete graph on the left is balanced; the one on the right is not.

friendly to one another, or you might find yourself gradually turning against one of the two; if you were an enemy of each of two mutual enemies, you might end up allying yourself with one of the two of them. (Despite this "dynamic" intuition, though, note that in this problem we're taking the labeling of each edge as fixed and given, not as something that is changing.))

What's interesting is that these apparently simple, local rules determine fairly specific global properties of the labeled complete graph G. Given a balanced labeling of G, let's define G^+ to be the graph on the same set of nodes, but consisting only of the edges labeled $+$. (This can be viewed as the graph with edges connecting all the pairs of friends in the network.)

(a) Prove that for any balanced labeling of a complete graph G, the graph G^+ has at most two connected components.

(b) Prove that if v and w are any two nodes that belong to the same connected component of G^+, then in fact there is an edge of G^+ connecting v and w. (Note that this is certainly not true in an arbitrary undirected graph; in general, one can easily have two nodes in a connected component that are not directly connected by an edge, but only have some longer path between them. So in proving this, you need to make use of the fact that G^+ is derived from a balanced labeling of a complete graph G.)