Baseball Elimination

Given the standings in a sports league at some point during the season, determine which teams have been mathematically eliminated from winning their division.

The baseball elimination problem. In the baseball elimination problem, there is a league consisting of N teams. At some point during the season, team i has w[i] wins and g[i][j] games left to play against team j. A team is eliminated if it cannot possibly finish the season in first place or tied for first place. The goal is to determine exactly which teams are eliminated. The problem is not as easy as many sports writers would have you believe, in part because the answer depends not only on the number of games won and left to play, but also on the schedule of remaining games. To see the complication, consider the following scenario:

				Against			
Team	**Wins**	**Losses**	**Left**	**Atlanta**	**Philadelphia**	**New York**	**Montreal**
Atlanta	83	71	8		1	6	1
Philadelphia	80	79	3	1		0	2
New York	78	78	6	6	0		0
Montreal	77	82	3	1	2	0	

Montreal is mathematically eliminated since it can finish with at most 80 wins and Atlanta already has 83 wins. This is the simplest reason for elimination. However, there can be more complicated reasons. For example, Philadelphia is also mathematically eliminated. It can finish the season with as many as 83 wins, which appears to be enough to tie Atlanta. But this would require Atlanta to lose all of its remaining games, including the 6 against New York, in which case New York would finish with 84 wins. We note that New York is not yet mathematically eliminated despite the fact that it has fewer wins than Philadelphia.

It is sometimes not so easy for a sports writer to explain why a particular team is eliminated and newspapers occasionally get it wrong. Consider the following scenario from the American League East at the end of play on August 30, 1996. Note that a team's total number of remaining games does not necessarily equal the number of remaining games against divisional rivals since teams may play opponents outside of their own division.

					Against				
	Team	**Wins** (w[i])	**Losses** (l[i])	**Left** (r[i])	**New York**	**Baltimore**	**Boston**	**Toronto**	**Detroit**
0	New York	75	59	28		3	8	7	3
1	Baltimore	71	63	28	3		2	7	4
2	Boston	69	66	27	8	2		0	0
3	Toronto	63	72	27	7	7	0		0
4	Detroit	49	86	27	3	4	0	0	

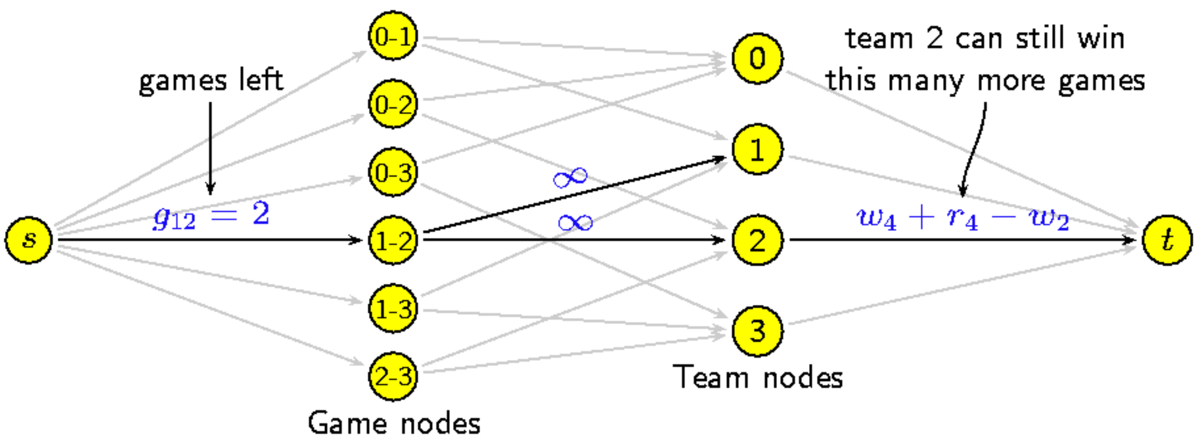
Here, g = $\begin{bmatrix} 0 & 3 & 8 & 7 & 3 \\ 3 & 0 & 2 & 7 & 4 \\ 8 & 2 & 0 & 0 & 0 \\ 7 & 7 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \end{bmatrix}$ (we assume that indices in rows and columns start at 0)

At first (or second) glance it might appear that Detroit still has a remote chance of catching New York and winning the division since Detroit can finish with as many as 76 wins. Try to convince yourself that this is *not* the case and that Detroit is already mathematically eliminated. We will provide an explanation below.

A maximum flow formulation. We now solve the baseball elimination problem by reducing it to the maximum flow problem. To check whether or not one particular team x is eliminated, we create a network and solve a maximum flow problem in it. In the network, feasible integral flows correspond to outcomes of the remaining schedule. There are vertices corresponding to teams (other than team x) and to remaining divisional games not involving team x. Intuitively, each unit of flow in the network corresponds to a remaining game. As it flows through the network, it passes from a game node, say between teams i and j, then through one of the team nodes i or j, classifying this game as being won by that team.

More precisely, our graph includes the following edges and capacities. We connect an artificial source s to each game node i-j and set its capacity to g[i][j]. If a flow uses all g[i][j] units of capacity on this edge, then we think of all these games as being played, and the wins are distributed between team nodes i and j. We connect each game node i-j with the two opposing teams to ensure that one of the two teams earns a win. We do not need to restrict the amount of flow on such edges. Finally, we connect each team node to an artificial sink t. Team x can win as many as w[x] + r[x] games, so to prevent team i from winning more than that many games in total, we include an edge from team node i to the sink with capacity w[x] + r[x] - w[i].

If the max flow in the network saturates all arcs leaving s (i.e. flow through each of these arcs is equal to the corresponding capacity), then this corresponds to assigning winners to all of the remaining games in such a way that no team wins more games than x. If the max flow does not saturate all arcs leaving s, then there is no scenario in which team x wins the division, so it gets eliminated.



What the min cut tells us. By solving a max flow problem, we can determine which teams are mathematically eliminated. It would be nice if we could also explain the reason for a team's elimination to a friend without resorting to network flow theory. Here's a more convincing and succinct argument for Detroit's elimination in the American League East example above. With the best possible luck, Detroit could finish the season with as many as 49 + 27 = 76 wins. Consider the subset of teams R = { NY, Bal, Bos, Tor}. Collectively, they have 75 + 71 + 69 + 63 = 278 wins among them. There are 3 + 8 + 7 + 2 + 7 = 27 remaining games among them (consider the upper triangular matrix of these four teams in g), so collectively, these four teams must win at least 27 more games, resulting in a total win of 278 + 27 = 305 games. On average, the teams in R win at least 305 / 4 = 76.25 games. Regardless of the outcome, one team in R will win at least $\lceil 76.25 \rceil$ = 77 games (as the number of wins is always an integer), thereby eliminating Detroit.

In fact, when a team is mathematically eliminated there always exists such a convincing *certificate of elimination*, although the subset of teams R may not be the rest of the division as it is above. However, you can always find such a subset R by choosing the team vertices on the source side of a min s-t cut in the baseball elimination network. Note that although we use max flows and min cuts to find the subset R, once we have it, the argument for a team's elimination does not involve any sophisticated mathematics.

Your assignment. Write a program that reads in a sports league and prints out a list of all of the teams that are mathematically eliminated. (**Bonus and optional**: For each team, give a convincing reason of the form described above). For example, for the input given below

```
4
Atlanta       83 71  8  0 1 6 1
```

```
Philadelphia  80 79  3  1 0 0 2
New_York      78 78  6  6 0 0 0
Montreal      77 82  3  1 2 0 0
```

your program should output something like

Philadelphia is eliminated.
They can win at most 80 + 3 = 83 games.
Atlanta and New York have won a total of 161 games.
They play each other 6 times.
So on average, each of the teams wins 167/2 = 83.5 games.

Montreal is eliminated.
They can win at most 77 + 3 = 80 games.
Atlanta has won a total of 83 games.
They play each other 0 times.
So on average, each of the teams in this group wins 83/1 = 83 games.

(The blue colored lines are optional and for bonus marks.)

Simplifying assumptions. Assume that no games end in a tie (as is the case in Major League Baseball). Also assume that there are no rainouts, i.e., every scheduled game is played. Ignore wildcard possibilities, i.e., when a team can make the playoffs without finishing first in its division. Finally, assume that there are no whitespace characters in the name of a team.

Implement Edmonds-Karp algorithm to find the maximum flow.

Submission guideline: Copy all the files in a folder named after your 7 digit student id and then compress the folder in a .zip file. Submit that .zip file.

Copy-checking will be done using all your codes and the codes available online, so don't copy.

Deadline: 12 January 2022 3:00 am