Class 25

April 19, 2023

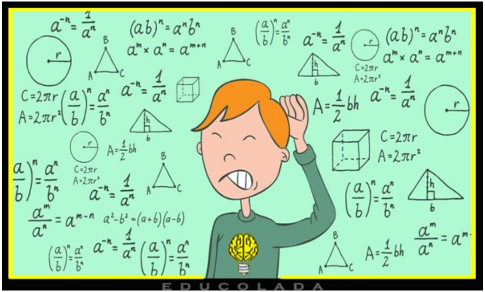
Max Flow and Baseball Elimination

Introduction to Dynamic Programming

An Application to Baseball

Announcements

Exam 2: Next Monday Evening



7 PM – ?

Baseball Elimination Via Max Flow

"See that thing in the paper last week about Einstein? ...
Some reporter asked him to figure out the mathematics
of the pennant race.
You know, one team wins so many of their remaining games,
the other teams win this number or that number.
What are the myriad possibilities?
Who's got the edge?"
"The hell does he know?"
"Apparently not much.
He picked the Dodgers to eliminate the Giants last Friday."
Don DeLillo, *Underworld*

Team	Wins	Games To Play	Against Atlanta	Against Phily	Against NY	Against Miami
Atlanta	83	8	-	1	6	1
Philadelphia	80	3	1	-	0	2
New York	78	6	6	0	-	0
Miami	77	3	1	2	0	-

Which teams have a chance of finishing the season with the most wins?

Miami is eliminated: it can finish with at most 80 wins, but Atlanta already has 83.
Sportswriters use *The Magic Number*.
Magic Number

Another Example: Can Boston finish with at least as many wins as every other team in the division?

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	92	2	-	1	1	0
Baltimore	91	3	1	-	1	1
Toronto	91	3	1	1	-	1
Boston	90	2	0	1	1	-

First Glance: Yes. Boston wins both its remaining games, Baltimore and Toronto win exactly 1, and New York loses both its games.

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	92	2	-	1	1	0
Baltimore	91	3	1	-	1	1
Toronto	91	3	1	1	-	1
Boston	90	2	0	1	1	-

Second Thought: No. If New York loses both of its games, then Baltimore and Toronto each pick up a win; the winner of the Baltimore -Toronto game finishes with 93 victories.

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	92	2	-	1	1	0
Baltimore	91	3	1	-	1	1
Toronto	91	3	1	1	-	1
Boston	90	2	0	1	1	-

A Different Analysis: Boston can win at most 92 games. The other three teams have a cumulative total of $92 + 91 + 91 = 274$ wins. Their three games against each other will produce an additional 3 wins for a final total of 277 wins. One of the teams must end up with more than 92 wins since the average number of wins is $277/3 > 92$.

- Is there an efficient algorithm to determine whether a given team has been eliminated from first place?
- When a team has been eliminated, is there an averaging argument that proves it?

A Mathematical Formulation

We have a set S of teams.

For each team x in S, let w_x = its current number of wins.

For each pair of teams x, y, let g_{xy} = the number of games they will play against each other.

Let z represent the team in S whose fate we wish to examine.

if T is a subset of the set of teams, then $|T|$ denotes the number of teams in T.

Theorem *Characterization Theorem for Baseball Elimination*):
Suppose team z has been eliminated. Then there exists a proof of this fact of the following form:

- z can finish with at most m wins.
- There is a subset T of S teams such that

$$\sum_{x \in T} w_x + \sum_{x,y \in T} g_{xy} > m|T|$$

Another Example

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	90	7	-	1	6	-
Baltimore	88	2	1	-	1	-
Toronto	87	7	6	1	-	-
Boston	79	12	-	-	-	-

Claim: Boston has been eliminated.
Boston can finish with at most $m = 79 + 12 = 91$ wins.
Let $T = \{\text{New York, Toronto}\}$. Then

$$\sum_{x \in T} w_x + \sum_{x,y \in T} g_{xy} = 90 + 87 + 6 = 183 > 91 \cdot 2 = 182.$$

One of New York or Toronto will finish with at least 92 wins.

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	90	7	-	1	6	-
Baltimore	88	2	1	-	1	-
Toronto	87	7	6	1	-	-
Boston	79	12	-	-	-	-

Note: The set $T =$ New York, Toronto, Baltimore doesn't work: Here

$$\sum_{x \in T} w_x + \sum_{x,y \in T} g_{xy} = 265 + 8 = 273$$

with average $273/3 = 91$.

Designing and Analyzing the Algorithm

Suppose there's a way for z to end up in first place with m wins. We now want to allocate wins for all remaining games so no other team finishes with more than m wins.

We'll allocate wins using a maximum flow computation.
We have a source s from which all wins emanate.
The ith win can pass through one of the two teams involved in the ith game.
We will then impose a capacity constraint: at most $m - w_x$ wins can pass through team x.

Construct a flow network G.
Let $S' = S - \{z\}$ (The set of other teams).
Let $g^* = \sum_{x,y \in S'} g_{xy}$ (total number of games left between all pairs.

Nodes

- s a source and t a sink.
- a node v for each team in S'.
- a node u_{xy} for each pair of teams in S' with a nonzero number of games left to play against each other.

Edges

- (s, u_{xy}): wins emanate from s.
- (v_x, t): wins are absorbed at t.
- (u_{xy}, v_x) and (u_{xy}, v_y): only x or y can win a game that they play against each other.

Capacities:

- Capacity of (s, u_{xy}) should be g_{xy}.
- Capacity of (v_x, t) should be $m - w_x$ (Ensure that team x cannot win more than $m - w_x$ games).
- Capacity of each (u_{xy}, v_x) will be infinite.

Team	Wins	Games To Play	Against New York	Against Baltimore	Against Toronto	Against Boston
New York	90	7	-	1	6	-
Baltimore	88	2	1	-	1	-
Toronto	87	7	6	1	-	-
Boston	79	12	-	-	-	-

Is Boston eliminated?

$m = 91$

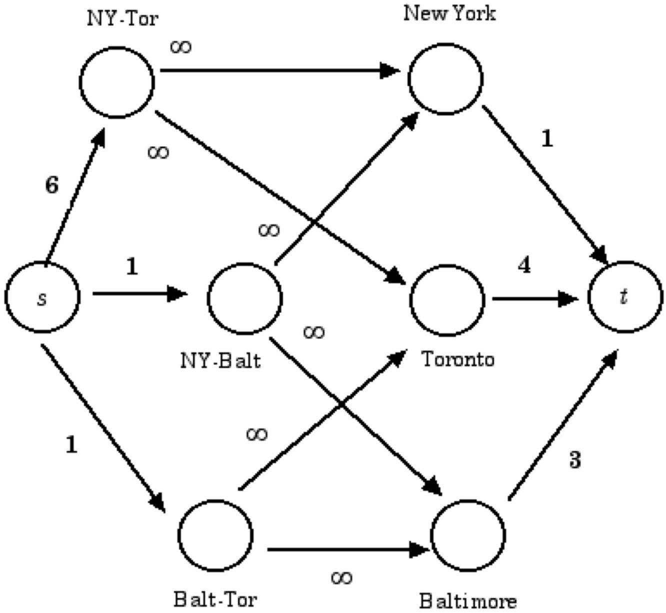
$S' = S - \{z\} = \{$ New York, Baltimore, Toronto$\}$

$g^* = 8$ games left

Capacity of $(v_{NewYork}, t) = m - w_{NewYork} = 91 - 90 = 1$

Capacity of $(v_{Baltmore}, t) = m - w_{Baltimore} = 91 - 88 = 3$

Capacity of $(v_{Toronto}, t) = m - w_{Toronto} = 91 - 87 = 4$

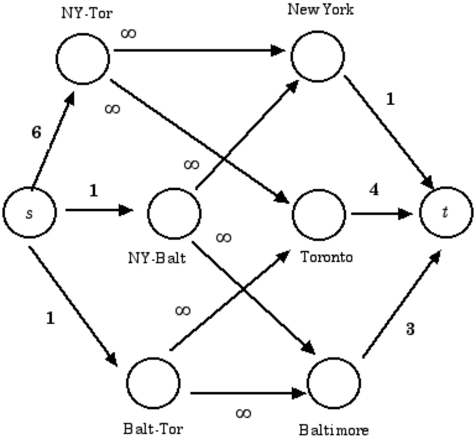

NY-Tor
New York
∞
∞
1
6
∞
1
4
s
t
∞
NY-Balt
Toronto
∞
1
3
∞
Balt-Tor
Baltimore

If there is a flow of value g^*, then it is possible for the outcomes of all remaining games to yield a situation where no team has more than m wins.

Hence if z wins all its remaining games, it can still achieve at least a tie for first place.

Conversely, if there are outcomes for the remaining games in which z does achieve at least a tie, we can use these outcomes to define a flow of value g^*.

Boston has a chance if and only if the maximum flow in the network is at least g* = 8 .



The maximum flow in this network is only 7

A Minimum cut is { $s \to$ Bal-Tor, $s \to$ NY-Bal, Toronto $\to t$, New York $\to t$. }

We have shown: **Team z has been eliminated if and only if the maximum flow in G has value strictly less than g^*.**

Characterizing When a Team is Eliminated

Theorem (*Characterization Theorem for Baseball Elimination*): Suppose team z has been eliminated. Then there exists a proof of this fact of the following form:

- z can finish with at most m wins.
- There is a subset T of S such that

$$\sum_{x \in T} w_x + \sum_{x,y \in T} g_{xy} > m|T|$$

Proof of Theorem: I. Suppose z has been eliminated.
The maximum $s - t$ flow in G has value $g' < g^*$

There is an $s - t$ minimum cut (A, B) of capacity g'

Let T be the set of teams x for which v_x is in A.

Claim: We can use T in the "averaging argument."

First, consider the node u_{xy} and suppose one of x or y is not in T, but u_{xy} is in A. Then the edge (u_{xy}, v_x) would cross from A to B, and hence the cut (A, B) would have infinite capacity.

But this contradicts the assumption that (A, B) is a minimum cut of capacity less than g^*.

Thus, if x or y is not in T, then u_{xy} is in B.

On the other hand, suppose x and y both belong to T, but u_{xy} is in B. Consider the cut (A', B') obtained by adding u_{xy} to the set A and deleting it from the set B.

The capacity of (A', B') is simply the capacity of (A, B) minus the capacity g_{xy} of the edge (s, u_{xy}) for this edge (s, u_{xy}) used to cross from A to B, but now does not cross from A' to B'.

Since $g_{xy} > 0$, the cut (A', B') has smaller capacity than the cut (A, B) , contradicting the minimality assumption on (A, B).

Thus, if x and y belong to T, then u_{xy} is in A.

We have established: u_{xy} is in A if and only if both x and y are in T.

Now let's determine the cut-value $c(A, B)$. The edges crossing from A to B have two possible forms:

- edges of the form (v_x, t) where x is in T, and
- edges of the form (s, u_{xy}) where at least one of x or y does not belong to T. (ie, $\{x, y\}$ is not a subset of T.)

$$\text{Thus } c(A, B) = \sum_{x \in T} (m - w_x) + \sum_{x,y \in T} g_{xy}$$

$$= m|T| - \sum_{x \in T} w_x + (g^* - \sum_{x,y \in T} g_{xy})$$

but we know that $c(A, B) = g' < g^*$ so

$$m|T| - \sum_{x \in T} w_x + (g^* - \sum_{x,y \in T} g_{xy}) < g^*$$

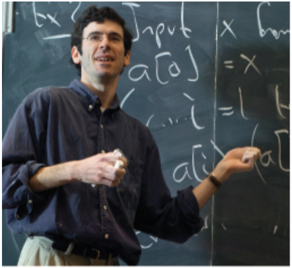
$$\text{implying } m|T| - \sum_{x \in T} w_x - \sum_{x,y \in T} g_{xy} < 0$$

or

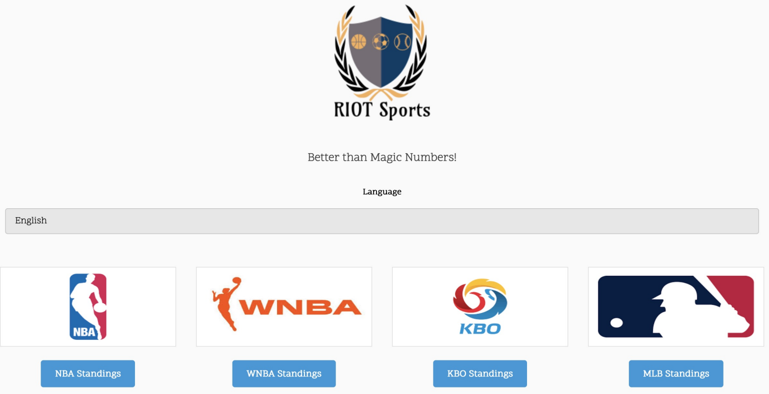
$$\sum_{x \in T} w_x + \sum_{x,y \in T} g_{xy} > m|T|$$

Adapted from Jon Kleinberg and Éva Tardos, *Algorithm Design*, Boston: Pearson Addison-Wesley, 2006; Second Edition, 2022.

Éva Tardos
October 1, 1957
Tardos Home Page



Jon Kleinberg
October 16, 1971
Kleinberg Home Page



Link to RIOT Site

DYNAMIC PROGRAMMING

Richard Ernest Bellman

Born: August 26, 1920 in Brooklyn

Died: March 19, 1984 in Los Angeles

Dynamic Programming

Dynamic programming is a useful technique that we can use to solve many optimization problems by breaking up large problems into a sequence of smaller, more tractable problems and then **Working Backward** from the end of the problem toward the beginning of the problem.

Dynamic Programming

The Advantage of Working Backwards

Example: Variation of The Game of Nim

There are 30 pennies on a table.

I begin by picking up 1, 2 or 3 pennies.

You then remove 1, 2 or 3.

We continue until the last penny is removed.

The player who picks up the last penny loses.

How should I play this game?