Algorithms and Data Structures

Modelling with Flows

Bipartite Matching

Maximum Bipartite Matching or Maximum matching on a bipartite graph G.

Bipartite graphs

A graph G=(V,E) is bipartite *if any only if* it can be partitioned into sets A and B such that each edge has one endpoint in A and one endpoint in B.

Often, we write G=(L,R,E).



Bipartite Matching

Maximum Bipartite Matching or Maximum matching on a bipartite graph G.

Matching: A subset M of the edges E such that each node v of V appears in at most one edge e in E.

Maximum matching: A matching with maximum cardinality. (i.e., |M| is maximised).

Example

A maximum matching



A maximal matching





S







t









Claim: Assume that there is a matching M of size k on G. Then there is a flow f of value k in G^{f} .

Claim: Assume that there is a matching M of size k on G. Then there is a flow f of value k in G^{f} .

Consider the matching

 $M = \{(U_1, V_1), (U_2, V_2), \dots, (U_k, V_k)\}$

Claim: Assume that there is a matching M of size k on G. Then there is a flow f of value k in G^{f} .

Consider the matching

$$M = \{(U_1, V_1), (U_2, V_2), \dots, (U_k, V_k)\}$$

Consider the flow such that

 $f(s, u_i) = f(u_i, v_i) = f(v_i, t) = 1$ for all i = 1, ..., k

f(e) = 0, otherwise

Claim: Assume that there is a matching M of size k on G. Then there is a flow f of value k in G^{f} .

Consider the matching

$$M = \{(U_1, V_1), (U_2, V_2), \dots, (U_k, V_k)\}$$

Consider the flow such that

 $f(s, u_i) = f(u_i, v_i) = f(v_i, t) = 1$ for all i = 1, ..., k

f(e) = 0, otherwise

This is a feasible flow and obviously has value k.



Claim: Assume that there is a a flow f of value k in G^{f} . Then there is a matching M of size k on G.

Claim: Assume that there is a a flow f of value k in G^{f} . Then there is a matching M of size k on G.

For an edge e, f(e) is either 0 or 1. (why?)

Claim: Assume that there is a a flow f of value k in G^{f} . Then there is a matching M of size k on G.

For an edge e, f(e) is either 0 or 1. (why?)

Consider the set M' of edges (of the middle level) with f(e) = 1.

Consider the set M' of edges with f(e) = 1.

Consider the set M' of edges with f(e) = 1.

Claim: |M'| = k.

Consider the set M' of edges with f(e) = 1.

Claim: |M'| = k.



Consider the set M' of edges with f(e) = 1.

Claim: |M'| = k.



Consider the set M' of edges with f(e) = 1.

Claim: M' is a matching.



Maximum Flow and Maximum matching

The size of the maximum matching M in G is equal to the value of the maximum flow f in G^f.

The edges of M are the edges that carry flow from A to B in G^{f} .

Maximum Flow and Maximum matching

The size of the maximum matching M in G is equal to the value of the maximum flow f in G^f.

The edges of M are the edges that carry flow from A to B in G^{f} .

What was the crucial part, that allows us to establish this?

Maximum Flow and Maximum matching

The size of the maximum matching M in G is equal to the value of the maximum flow f in G^f.

The edges of M are the edges that carry flow from A to B in G^{f} .

What was the crucial part, that allows us to establish this?

The integrality theorem.

What is the running time of the algorithm?

By Edmonds - Karp, we get O(nm²).

What is the running time of the algorithm?

By Ford-Fulkerson, we get **O(mF).**

What is the running time of the algorithm?

By Ford-Fulkerson, we get **O(mF).**

How large is F here?

What is the running time of the algorithm?

By Ford-Fulkerson, we get **O(mF).**

How large is F here?

It is at most $max\{|L|, |R|\}$.

What is the running time of the algorithm?

By Ford-Fulkerson, we get **O(mF).**

How large is F here?

It is at most $max\{|L|, |R|\}$.

Running time **O(nm).**

Polynomial Time Reduction

Polynomial Time Reduction

We are given a problem A that we want to solve.
We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

does the transformation,

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.

We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.

Pictorially



We are given a problem A that we want to solve.

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.

We are given the maximum bipartite matching problem

We can *reduce* solving problem A to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.

We are given the maximum bipartite matching problem

We can *reduce* solving the MBP to solving some other problem B.

This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



This is a transformation of an instance of problem A to an instance of problem B.

Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



Assume that we had an algorithm ALG^B for solving problem B.

We can construct an algorithm ALG^A for solving problem A, which

- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



- does the transformation,
- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



- uses the algorithm ALG^B,
- transforms the solution back to a solution to problem A.



In the baseball league, there are 4 teams with the following number of wins:

New York92Baltimore91Toronto91Boston90

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

In the baseball league, there are 4 teams with the following number of wins:

Assume Boston wins all remaining games.

New York92Baltimore91Toronto91Boston90

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

In the baseball league, there are 4 teams with the following number of wins:

New York92Baltimore91Toronto91Boston90

Assume Boston wins all remaining games.

New York must lose all remaining games.

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

In the baseball league, there are 4 teams with the following number of wins:

New York 92 91 Baltimore 91 Toronto 90 Boston

remaining games.

New York must lose all remaining games.

Assume Boston wins all Baltimore and Toronto must win one game each.

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

In the baseball league, there are 4 teams with the following number of wins:

New York92Baltimore91Toronto91Boston90

Assume Boston wins all
remaining games.Baltimore and Toronto must
win one game each.New York must lose all
remaining games.Baltimore or Toronto must
win one more game (BLT vs TOR).

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

In the baseball league, there are 4 teams with the following number of wins:

New York92Baltimore91Toronto91Boston90

Assume Boston wins all remaining games.
New York must lose all remaining games.
Baltimore and Toronto must win one game each.
Baltimore or Toronto must win one more game (BLT vs TOR).

There are five games left in the season.

NY vs BLT, NY vs TOR, BLT vs TOR, BLT vs BOS, TOR vs BOS

Question: Can Boston finish (possibly tied for) first?

The answer is no.

In the baseball league, there are 4 teams with the following number of wins:`

New York90Baltimore88Toronto87Boston79

These are the games left in the season:

NY vs BLT

NY vs TOR 6 games

BLT vs TOR

BOS vs ANY 4 games (12 games total)

Generally:

Generally:

We have a set S of teams.

Generally:

We have a set S of teams.

For each team x in S, the current number of wins is w_x .

Generally:

We have a set S of teams.

For each team x in S, the current number of wins is w_x .

For teams x and y in S, they still have to play g_{xy} games against each other.

Generally:

We have a set S of teams.

For each team x in S, the current number of wins is w_x .

For teams x and y in S, they still have to play g_{xy} games against each other.

We are given a designated team z.

Generally:

We have a set S of teams.

For each team x in S, the current number of wins is w_x .

For teams x and y in S, they still have to play g_{xy} games against each other.

We are given a designated team z.

Can z win the tournament (possibly in a tie?)

From baseball to flows

From baseball to flows

Observation: If there is a way for z to be first, there is a way for z to be first when winning all remaining games.

From baseball to flows

Observation: If there is a way for z to be first, there is a way for z to be first when winning all remaining games.

Suppose that in the end, team z has m wins.
Observation: If there is a way for z to be first, there is a way for z to be first when winning all remaining games.

Suppose that in the end, team z has m wins.

What are we looking for?

Observation: If there is a way for z to be first, there is a way for z to be first when winning all remaining games.

Suppose that in the end, team z has m wins.

What are we looking for?

Is there an allocation of all the remaining g* games (between the other teams) such that no team ends up with more than m wins?



A pair of teams











Two edges if teams in p_j still have games to play between them.



























Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)



Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)

Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)

The following hold:

Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)

The following hold:

A pair (x, y) will play exactly g_{xy} games.

Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)

The following hold:

A pair (x, y) will play exactly g_{xy} games.

A team x will win at most $m-w_x$ games.

Assume that the algorithm says yes.

The value of the flow is equal to the number of remaining games. (why?)

The following hold:

A pair (x, y) will play exactly g_{xy} games.

A team x will win at most m-w_x games.

Team z can win.

Assume that the algorithm says no.

Assume that the algorithm says no.

The maximum flow has value $< g^*$.

Assume that the algorithm says no.

The maximum flow has value $< g^*$.

It is not possible to play all the remaining games without giving some team x more than $m - w_x$ points.

Assume that the algorithm says no.

The maximum flow has value $< g^*$.

It is not possible to play all the remaining games without giving some team x more than $m - w_x$ points.

Team z cannot win.



Let's look at the final residual graph. Why do we have no augmenting paths?



Let's look at the final residual graph. Why do we have no augmenting paths?



Let's look at the final residual graph. Why do we have no augmenting paths?


Another way to think about it

Let's look at the final residual graph. Why do we have no augmenting paths?



Either all the games have been played, or some team cannot win any more games.

Example

In the baseball league, there are 4 teams with the following number of wins:`

New York90Baltimore88Toronto87Boston79

There are five games left in the season.

NY vs BLT

NY vs TOR 6 games

BLT vs TOR

BOS vs ANY 4 games (12 games total)

Question: Can Boston finish (possibly tied for) first?



m = 91

We extract blocks of earth from the surface, trying to find gold.

We extract blocks of earth from the surface, trying to find gold.

Each block z that we mine has

We extract blocks of earth from the surface, trying to find gold.

Each block z that we mine has

a value pz

We extract blocks of earth from the surface, trying to find gold.

Each block z that we mine has

a value p_z

a mining cost Cz

We extract blocks of earth from the surface, trying to find gold.

Each block z that we mine has

a value pz

a mining cost Cz

Constraint: We can not mine a block z unless we mine the two blocks x and y on top of it.

We extract blocks of earth from the surface, trying to find gold.

Each block z that we mine has

a value pz

a mining cost Cz

Constraint: We can not mine a block z unless we mine the two blocks x and y on top of it.

We want to earn as much money as possible.



From pits to flows



From pits to flows t Is pz - Cz > 0?













Consider an (S, T) cut C.

Consider an (S, T) cut C.

We will mine $S - \{s\}$.

Consider an (S, T) cut C.

We will mine S - {s}.

If C is minimum, we cannot have nodes that are connected with an *infinite capacity* edge in different sides of the cut.

Consider an (S, T) cut C.

We will mine S - {s}.

If C is minimum, we cannot have nodes that are connected with an *infinite capacity* edge in different sides of the cut.

If S contains z, it must contain x and y (needed to mine z).

Consider an (S, T) cut C.

We will mine S - {s}.

If C is minimum, we cannot have nodes that are connected with an *infinite capacity* edge in different sides of the cut.

If S contains z, it must contain x and y (needed to mine z).

Feasibility guaranteed by the above fact.

Consider an (S, T) cut C.

We will mine S - {s}.

If C is minimum, we cannot have nodes that are connected with an *infinite capacity* edge in different sides of the cut.

If S contains z, it must contain x and y (needed to mine z).

Feasibility guaranteed by the above fact.

Optimality?

$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$

From pits to cuts t **c**_x - **p**_x ∞ p_z - c_z **p**_y - **c**_y







$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$$



$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S: \ p_z - c_z < 0} (p_z - c_z)$$

Add and subtract this:

$$c(S,T) = \sum_{z \in S : p_z - c_z > 0} (p_z - c_z)$$

$$c(S,T) = \sum_{z \in T : p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S : p_z - c_z < 0} (c_z - p_z)$$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S: \ p_z - c_z < 0} (p_z - c_z)$$

Add and subtract this:
$$c(S,T) = \sum_{z \in S : p_z - c_z > 0} (p_z - c_z)$$

$$c(S,T) = \sum_{z \in V: \ p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S} (p_z - c_z)$$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S: \ p_z - c_z < 0} (p_z - c_z)$$

Add and subtract this:

$$c(S,T) = \sum_{z \in S : p_z - c_z > 0} (p_z - c_z)$$

$$c(S,T) = \sum_{z \in V : p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S} (p_z - c_z)$$

constant

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) + \sum_{z \in S: \ p_z - c_z < 0} (c_z - p_z)$$

$$c(S,T) = \sum_{z \in T: \ p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S: \ p_z - c_z < 0} (p_z - c_z)$$

Add and subtract this:

$$c(S,T) = \sum_{z \in S : p_z - c_z > 0} (p_z - c_z)$$

$$c(S,T) = \sum_{z \in V : p_z - c_z > 0} (p_z - c_z) - \sum_{z \in S} (p_z - c_z)$$

$$constant$$
Mining profit
Open-pit mining -Summarising

Construct the flow network.

Run Ford-Fulkerson to find a maximum flow.

Find a minimum cut using the final residual graph.

Mine the blocks in the S part of the cut.