Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem

CS 401/MCS 401 Lecture 16
Computer Algorithms I
Jan Verschelde, 25 July 2018
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
In a bipartite graph $G = (V, E)$, we have:
1. $V = A \cup B$, $A \cap B = \emptyset$, $A \neq \emptyset$, $B \neq \emptyset$.
2. for any edge $(a, b) \in E$: $a \in A$ and $b \in B$.

A matching $M$ in a bipartite graph $G = (V, E)$:
1. $M \subseteq E$, is a subset of the set of edges,
2. every vertex appears at most once in $M$.

Given a bipartite graph $G$, the bipartite matching problem asks to find the matching $M$ in $G$ of the largest size.
the flow network for a bipartite graph

Given $G = (V, E)$, with $V = A \cup B$, add a source $s$ and sink $t$, $V' = V \cup \{s, t\}$, and construct the flow network $G' = (V', E')$ as:

1. For all $e \in E$: $e \in E'$, all edges of $G'$ are directed.
2. For all $a \in A$: $(s, a) \in E'$, connect the source to $A$.
3. For all $b \in B$: $(b, t) \in E'$, connect the sink to $B$.
4. For all edges $e \in E'$: $c_e = 1$, all capacities are one.
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
matchings and flows

To a matching $M$ in $G$ corresponds a flow $f$ in $G'$:

- Let $M = \{ (a_1, b_1), (a_2, b_2), \ldots, (a_m, b_m) \}$.
- For all $(a, b) \in M$: $f(s, a) = 1$, $f(a, b) = 1$, and $f(b, t) = 1$.

The value of the flow is $m$, $\nu(f) = m$.

Exercise 1:
Verify that the $f$ corresponding to a matching $M$ is an $s$-$t$ flow:

1. Show that the capacity conditions on $f$ in $G'$ hold.
2. Show that the conservation conditions on $f$ in $G'$ hold.
flows and matchings

If all capacities are integers, then there is an integer valued flow. If we have an integer valued flow \( f \) in \( G' = (V, E) \), then

\[
M' = \{ (a, b) \in E' \mid f(a, b) = 1 \}.
\]

**Lemma**

*If the value of the flow \( f \) is \( m \), then \( \#M' = m \).*

**Proof.** \( G' \) is the flow network for the bipartite graph \( G = (V, E) \), with \( V = A \cup B \). Consider the cut which separates the source \( s \) from \( A \). The value of the flow is \( m = f^{\text{in}}(A) \). By the conservation conditions, \( f^{\text{out}}(A) = f^{\text{in}}(A) \), the flow into \( A \) equals the flow out of \( A \).

- The edges out of \( A \) that carry flow are the edges of \( M' \).
- The flow at each edge equals one.

The number of edges with flow equals the size of \( M' \). Q.E.D.
the flow and its corresponding matching

If we have an integer valued flow $f$ in $G' = (V, E)$, then for

$$M' = \{ (a, b) \in E' \mid f(a, b) = 1 \},$$

we have to show that $M'$ leads to a matching in $G = (V, E)$:

**Lemma**

*Every $v \in V$ appears at most once in $M'$.*

**Proof.** By contradiction, suppose $v$ appears twice in $M'$. If $v \in A$, then

This picture violates the conservation conditions at $v$, which contradicts that $f$ is a flow in $G'$.

Thus $v$ cannot appear twice in $M'$.

If $v \in B$, then a similar picture will contradict the conservation conditions of the flow.

Q.E.D.
maximum flow and maximum matching

**Theorem**

Given a flow $f$ in a flow network $G'$, constructed for a bipartite graph $G = (V, E)$, consider

$$M = \{ (a, b) \in E \mid f(a, b) = 1 \}.$$  

Then, $v(f) = \#M$ and $M$ is a matching in $G$.

**Corollary**

*If the value of the flow $f$ in $G'$ is maximal over all flows, then the matching $M$ corresponding to $f$ is maximal as well.*

**Exercise 2:**
Write a proof for the corollary.
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
finding an alternating path
the maximum bipartite matching
When the Ford-Fulkerson algorithm is applied to the bipartite matching problem, the augmenting paths are called *alternating paths*, because those paths alternate between forward and backward edges when rerouting flow in the bipartite graph.

**Exercise 3:**
Consider the pictures on slides 11 and 12.
Describe the transitions between the top and the bottom graph, using the proper terminology introduced in lecture 15.
In particular, describe the construction of the residual graph with forward and backward edges in sufficient detail.
The Ford-Fulkerson algorithm runs in time $O(mC)$, $C = \sum_{e \text{ out of } s} c_e$.

For the flow network $G'$ constructed from the bipartite graph $G = (V, E)$, $V = A \cup B$, $C \leq \#A \leq n$, where $n = \#V$.

Therefore, we arrive at the following theorem:

**Theorem**

Consider a bipartite graph $G = (V, E)$, with $n = \#V$ and $m = \#E$. The maximum matching problem is solved by the Ford-Fulkerson algorithm in $O(mn)$ time.
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
perfect matchings

**Definition**

Given a bipartite graph \( G = (V, E) \) and a matching \( M \). \( M \) is a *perfect* matching if every \( v \in V \) appears in \( M \).

What if no perfect matching is found?

Consider a bipartite graph \( G = (V, E) \), \( V = A \cup B \).
If \( a_1, a_2 \in A \):

- each incident to only one edge, and
- the same \( b \) is that end of their incident edge.

Either \( a_1 \) or \( a_2 \) is excluded from the matching.
a certificate for no perfect matching?

How can we be sure no perfect matching exists?
Is there a certificate?

We generalize the example on the previous slide.

**Definition**

Let \( G = (V, E) \) be a bipartite graph, \( V = A \cup B \). For any subset \( S \subseteq A \), define \( \Gamma(S) \subseteq B \) as: \( \Gamma(S) = \{ b \in B \mid (a, b) \in E, a \in S \} \).

In the example on the previous slide, \( S = \{ a_1, a_2 \} \), \( \Gamma(S) = \{ b \} \) and \( \#\Gamma(S) = 1 < 2 = \#S \).

**Proposition**

*If there is a perfect matching in a bipartite graph \( G = (V, E) \) with \( V = A \cup B \), then \( \#\Gamma(A) \leq \#A \).*

Does the converse hold?
Hall’s theorem

Theorem

Let $G = (V, E)$ be a bipartite graph, $V = A \cup B$ with $\#A = \#B$. Then,

- either $G$ has a perfect matching,
- or there is a $S \subseteq A$: $\#\Gamma(S) < \#A$.

A perfect matching or a certificate subset $S$ can be found in $O(mn)$ time, where $n = \#V$ and $m = \#E$.

Outline of the proof:

1. The Ford-Fulkerson algorithm gives the maximum flow in $O(mn)$.
2. The value of the maximum flow equals the capacity of the minimum cut.
3. A breadth-first or dept-first search computes the cut in $O(m)$.
4. The minimum cut can be modified to find $S \subseteq A$: $\#\Gamma(S) < \#A$.
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
supplies and demands

Consider a flow network with multiple sources and multiple sinks.
- the sources represent factories who provide supply,
- the sinks represent retail outlets who have demand.

Given a flow network \( G = (V, E) \) with integer capacities \( c_e \) on the edges.
To each vertex \( v \) corresponds a demand \( d_v \):
- if \( d_v > 0 \), \( v \) has demand of \( d_v \) and is a sink,
- if \( d_v < 0 \), \( v \) has supply of \( d_v \) and is a source,
- if \( d_v = 0 \), \( v \) is an internal node.

A *circulation* with demands \( \{d_v\} \) is a function \( f \) that assigns a nonnegative number of each edge, subject to
- *capacity conditions*: for each \( e \in E \): \( 0 \leq f(e) \leq c_e \).
- *conservation conditions*: for each \( v \in V \): \( f^{\text{in}}(v) - f^{\text{out}}(v) = d_v \).
a feasibility problem

A *circulation* with demands \( \{d_v\} \) is a function \( f \) that assigns a nonnegative number of each edge, subject to

- *capacity conditions*: for each \( e \in E \): \( 0 \leq f(e) \leq c_e \).
- *conservation conditions*: for each \( v \in V \): \( f^{\text{in}}(v) - f^{\text{out}}(v) = d_v \).

Instead of an optimization problem, we are interested in *feasibility*: does there exist a circulation \( f \) which satisfies the capacity and conservation conditions?
Lemma

If there is a feasible circulation in a graph \( G = (V, E) \) with demands \( \{d_v\} \), then \( \sum_{v \in V} d_v = 0 \).

Proof. In a feasible solution \( f \), the conservation conditions are satisfied: for each \( v \in V \): \( f^{\text{in}}(v) - f^{\text{out}}(v) = d_v \). Summing over all \( v \):

\[
\sum_{v \in V} d_v = \sum_{v \in V} f^{\text{in}}(v) - \sum_{v \in V} f^{\text{out}}(v).
\]

Consider the edge \((u, v) \in E\). Along \((u, v)\), what goes out a vertex is \( f^{\text{out}}(u) \) and what comes in is \( f^{\text{in}}(v) \). Summing over all vertices,

- the term \( f^{\text{in}}(v) \) is canceled by \( f^{\text{out}}(v) \) in the other sum, and
- the term \( f^{\text{out}}(u) \) is canceled by \( f^{\text{in}}(u) \) in the other sum.

So the total sum equals zero. Q.E.D.
the capacity $D$

Corollary

For a feasible circulation with demands $\{d_v\}$:

$$\sum_{v \in V: d_v > 0} d_v = \sum_{v \in V: d_v < 0} -d_v.$$ 

The common value is denoted by

$$D = \sum_{v \in V: d_v > 0} d_v = \sum_{v \in V: d_v < 0} -d_v.$$
Two Applications of Maximum Flow

1. The Bipartite Matching Problem
   - a bipartite graph as a flow network
   - maximum flow and maximum matching
   - alternating paths
   - perfect matchings

2. Circulation with Demands
   - flows with multiple sources and multiple sinks
   - reduction to a flow problem
reduction to a flow problem

- Let $S$ denote the set of all vertices $v$ with $d_v < 0$.
- Let $T$ denote the set of all vertices $v$ with $d_v > 0$.

Attach a super source $s^*$ to each vertex $u \in S$:
- for each $u \in S$, add the edge $(s^*, u)$ with capacity $-d_u$.

Attach a super sink $t^*$ to each vertex $v \in T$:
- for each $v \in T$, add the edge $(v, t^*)$ with capacity $d_v$.

In this graph $G'$, we search for a maximum $s^*-t^*$ flow.
Exercise 4:
Run the Ford-Fulkerson algorithm on the flow network shown on the previous slide.
Show all steps in the execution of the algorithm.
circulation and maximum flow

**Theorem**

Let $G = (V, E)$ be a flow network with demands $\{d_v\}$. Let $G'$ be the graph with added source $s^*$ and sink $t^*$. There is a feasible circulation if and only if the maximum $s^* - t^*$ flow in $G'$ has value $D$.

**Proof.** We cannot have an $s^* - t^*$ flow in $G'$ of value larger than $D$ because the cut $(A, B)$ with $A = \{s^*\}$ has only capacity $D$.

1. $\Rightarrow$: A feasible circulation $f$ sends a flow value of $-d_v$ on each edge $(s^*, v)$ and a flow value $d_v$ on each edge $(v, t^*)$. Summing over all vertices, we obtain an $s^* - t^*$ flow of value $D$, which is maximal.

2. $\Leftarrow$: Consider a maximum $s^* - t^*$ flow in $G'$ of value $D$. Then, every edge out of $s^*$ and every edge into $t^*$ is completely saturated with flow. If we delete those edges, we obtain a circulation $f$ in $G$ which satisfies $f^{\text{in}}(v) - f^{\text{out}}(v) = d_v$, for every $v \in V$, so the conservation conditions are met. Q.E.D.
Theorem

Let $G = (V, E)$ be a flow network with demands $\{d_v\}$. Let $G' = (V', E')$ be the graph with added source $s^*$ and sink $t^*$. If all capacities and demands are integer numbers, then the feasible circulation has an integer value and can be computed in $O(mD)$ time, where $m = \#E'$.

The theorem follows from the application of the Ford-Fulkerson algorithm and the observation that the constant $C = \sum_{e \text{ out of } s} c_e$ equals $D$ in the graph $G'$. 