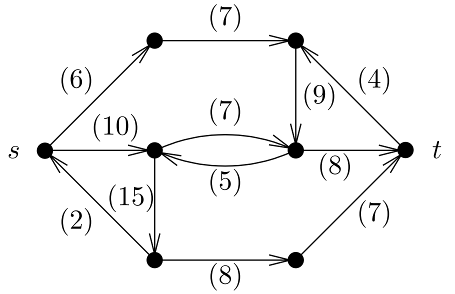
4. Flows and circulations

Let G be a directed graph with a nonnegative **capacity** $c(e)$ assigned to each edge $e \in E(G)$, and let $s, t \in V(G)$ be two distinct vertices.

The quadruple $N = (G, c, s, t)$ is a **flow network** with **source** s and **sink** t.



Flows

A **flow** on a flow network $N = (G, c, s, t)$ is a mapping $f : E(G) \to \mathbb{R}$ that satisfies the two conditions:

(F1) $0 \le f(e) \le c(e)$ for each edge $e \in E(G)$; and

(F2) for each vertex $v \ne s, t$, we have that

$$\sum_{e^+ = v} f(e) = \sum_{e^- = v} f(e),$$

where e^- and e^+ denote the start and end vertex of e.

Condition (F1) requires that the flow is **feasible**, i.e., is nonnegative and does not exceed the capacity of an edge. Condition (F2) requires **flow conservation**, i.e., that the flow entering any vertex $v \ne s, t$ is equal to the flow leaving v.

Value of a flow, maximum flow

The **value** $w(f)$ of a flow f on N is the net flow leaving the source s (equivalently, the net flow entering the sink), i.e.,

$$w(f) := \sum_{e^- = s} f(e) - \sum_{e^+ = s} f(e) = \sum_{e^+ = t} f(e) - \sum_{e^- = t} f(e).$$

A flow f is a **maximum flow** if $w(f) \ge w(f')$ for all flows f' on N.

The **maximum flow problem** is to determine a maximum flow for a given flow network.

In this lecture we study the maximum flow problem and its applications. Our main topics of interest are:

1. Maximum flows
   - ▷ Characterization using cuts and augmenting paths
   - ▷ Max-flow min-cut theorem, integral flow theorem
   - ▷ Edmonds-Karp algorithm (other algorithms not considered)
2. Zero-one flows, applications in graph theory
   - ▷ Menger's theorem, Hall's theorem
3. Generalizations: minimum-cost flow, circulations (brief overview, see [Ahu] and [Jun] for details).

Sources for this lecture

The material for this lecture has been prepared with the help of [Jun, Chapters 6, 7, 9] and [Wes, Section 4.3]. Two excellent references to network flow theory are [Ahu] and the classic [For].

[Ahu] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network flows*, Prentice Hall, Englewood Cliffs NJ, 1993.

[For] L. R. Ford, Jr. and D. R. Fulkerson, *Flows in Networks*, Princeton University Press, Princeton NJ, 1962.

Cut, capacity of a cut

A **cut** of a flow network $N = (G, c, s, t)$ is a partition of the vertex set $V(G)$ into two disjoint sets S, T such that $s \in S$ and $t \in T$.

The **capacity** of a cut (S, T) is

$$c(S,T) := \sum_{\substack{e^- \in S,\\ e^+ \in T}} c(e).$$

Lemma A.8 *For any flow f and any cut (S, T), we have*

$$w(f) = \sum_{\substack{e^- \in S,\\ e^+ \in T}} f(e) - \sum_{\substack{e^- \in T,\\ e^+ \in S}} f(e) \le c(S,T).$$

Proof: By (F2) we have

$$w(f) = \sum_{e^- = s} f(e) - \sum_{e^+ = s} f(e) = \sum_{v \in S} \Bigg(\sum_{e^- = v} f(e) - \sum_{e^+ = v} f(e) \Bigg).$$

The last sum can be partitioned into sums over edges with (i) both endpoints in S and (ii) exactly one endpoint in S:

$$\begin{aligned} w(f) &= \overbrace{\sum_{\substack{e^- \in S,\\ e^+ \in S}} f(e) - \sum_{\substack{e^+ \in S,\\ e^- \in S}} f(e)}^{\text{(i)}} + \overbrace{\sum_{\substack{e^- \in S,\\ e^+ \in T}} f(e) - \sum_{\substack{e^- \in T,\\ e^+ \in S}} f(e)}^{\text{(ii)}} \\ &= \sum_{\substack{e^- \in S,\\ e^+ \in T}} f(e) - \sum_{\substack{e^- \in T,\\ e^+ \in S}} f(e) \\ &\le c(S,T), \end{aligned}$$

where the last inequality follows from (F1). □

Augmenting paths

Let G be a digraph and let $u, v \in V(G)$. An **undirected path** P from u to v is an acyclic subgraph of G such that the underlying graph of P is a path with endpoints u, v.

An edge in P is either a **forward edge** (directed from u to v) or a **backward edge** (directed from v to u).

Let $N = (G, c, s, t)$ be a flow network with flow f. An undirected path P from u to v is (f-) **incrementing** if

(I1) $f(e) < c(e)$ for every forward edge $e \in E(P)$; and

(I2) $f(e) > 0$ for every backward edge $e \in E(P)$.

An f-incrementing path from s to t is an (f-)**augmenting path**.

Let P be an f-augmenting path. The residual capacity of a forward edge $e \in E(P)$ is $c(e) - f(e)$; the residual capacity of a backward edge $e \in E(P)$ is $f(e)$. The **residual capacity** of P is the minimum of the residual capacities of the edges in P.

Let d be the residual capacity of P. Clearly, $d > 0$. Define

$$f'(e) := \begin{cases} f(e) + d & \text{if } e \text{ is a forward edge in } P; \\ f(e) - d & \text{if } e \text{ is a backward edge in } P; \\ f(e) & \text{otherwise, i.e., } e \notin E(P). \end{cases}$$

It is straightforward to check that f' satisfies (F1) and (F2), i.e., is a flow. Moreover, $w(f') = w(f) + d$.

Thus, whenever an f-augmenting path exists, there exists a flow f' on N with larger value.

© Petteri Kaski 2006

Three theorems of Ford and Fulkerson

The following theorems are all due to Ford and Fulkerson (1956).

Theorem A.35 (Augmenting Path Theorem) *A flow f on a flow network $N = (G, c, s, t)$ is a maximum flow if and only if there exists no f-augmenting path.*

Proof: ($\Rightarrow$) If there exists an f-augmenting path with residual capacity d, then f is not a maximum flow because the flow f' defined earlier satisfies $w(f') = w(f) + d > w(f)$.

($\Leftarrow$) Let S be the set of all vertices reachable from s by an f-incrementing path. Since there exists no f-augmenting path, $t \notin S$. Put $T := V(G) - S$. Clearly, (S, T) is a cut of N. We show that $w(f) = c(S, T)$ to establish that f is a maximum flow.

© Petteri Kaski 2006

Proof: (continued) Lemma A.8 applied to the cut (S, T) gives:

$$w(f) = \sum_{\substack{e^- \in S, \\ e^+ \in T}} f(e) - \sum_{\substack{e^- \in T, \\ e^+ \in S}} f(e) \le \sum_{\substack{e^- \in S, \\ e^+ \in T}} c(e) = c(S, T).$$

Suppose that $w(f) < c(S, T)$. Then, there exists an edge e that satisfies either (i) $e^- \in S$, $e^+ \in T$, and $f(e) < c(e)$; or (ii) $e^- \in T$, $e^+ \in S$, and $f(e) > 0$.

Suppose case (i) occurs. By definition of S, e^- is reachable from s by an f-incrementing path P. But then e^+ is reachable from s by the f-incrementing path $P + e$, which contradicts the definition of S. For case (ii) we obtain a similar contradiction, so $w(f) = c(S, T)$.

By Lemma A.8, any flow f' on N satisfies $w(f') \le c(S, T)$. Thus, f is a maximum flow. $\square$

© Petteri Kaski 2006

Theorem A.36 (Max-Flow Min-Cut Theorem) *In any flow network, the maximum value of a flow is equal to the minimum capacity of a cut.*

Proof: Lemma A.8 shows that the capacity of any cut (S, T) is an upper bound on the value of a flow. The proof of Theorem A.35 shows that the maximum flow on a network reaches this bound for a particular cut (S, T). $\square$

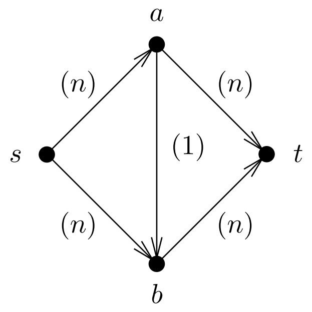
© Petteri Kaski 2006

Theorem A.37 (Integral Flow Theorem) *Let $N = (G, c, s, t)$ be a flow network in which all capacities $c(e)$ are integers. Then, there exists a maximum flow on N such that all values $f(e)$ are integers.*

Proof: Starting from the all zero flow f_0, repeatedly augment the flow using an augmenting path until no augmenting path exists (in which case the flow is a maximum flow by Theorem A.35). Clearly, the residual capacity in each augmentation is an integer. Thus, there are at most $\sum_e c(e)$ augmentations and the resulting maximum flow is integral. $\square$

- The proof of the integral flow theorem suggests a straightforward algorithm for computing maximum flows: starting with the all zero flow, repeatedly augment the flow until no augmenting path exists; output the flow and halt.
- This is essentially the labeling algorithm of Ford and Fulkerson (1957); see [Jun, p. 159–160] for pseudocode.
- There are two problems with the above approach:
  1. for general real-valued edge capacities, the algorithm may not halt at all (and converge to a flow that is not a maximum flow; see e.g. [Ahu, p. 205–206] or [For, p. 21]).
  2. even when all capacities are integral, the algorithm may require time proportional to $\sum_e c(e)$ with a poor choice of augmenting paths, which is very inefficient.

Example. In the flow network below, $2n$ augmenting steps are required to reach the maximum flow of value $2n$ if we use alternately the augmenting paths sa, ab, bt and sb, ba, at.



Edmonds-Karp algorithm

Edmonds and Karp (1972) gave a simple modification to the repeated augmentation algorithm (which ensures termination also in the general real-valued case):

> In each augmenting step, use a shortest possible augmenting path (i.e. an augmenting path with the least number of edges).

Theorem A.38 *The Edmonds-Karp algorithm performs at most $O(n(G)e(G))$ augmenting steps.*

Proof: See [Jun, Theorem 6.2.1]. $\square$

Remarks

- A shortest augmenting path can be located in time $O(e(G))$ using BFS.
- Thus, the Edmonds-Karp algorithm computes a maximum flow in time $O(n(G)e(G)^2)$ (assuming that the arithmetic is constant-time).
- More efficient maximum flow algorithms exist, see [Ahu] and [Jun] for details.

© Petteri Kaski 2006

Zero-one flows

A **zero-one flow** in a flow network N is a flow f with $f(e) \in \{0, 1\}$ for all $e \in E(G)$.

Zero-one flows occur in many combinatorial applications of flow theory.

We will use zero-one flows and the max-flow min-cut theorem to give simple proofs of Hall's theorem and Menger's theorem.

© Petteri Kaski 2006

Hall's theorem

Let H be a bipartite graph and let X, Y be a bipartition of its vertex set. We prove that there exists a matching in H that saturates every vertex in X if and only if $|N(A)| \geq |A|$ for all $A \subseteq X$.

Construct the following flow network $N = (G, c, s, t)$. Include to G all vertices in H; add two new vertices s, t. For each edge $uv \in E(H)$ with $u \in X$ and $v \in Y$, add the directed edge uv to G. Furthermore, add all edges of the form su, vt to G, where $u \in X$ and $v \in Y$. Each edge has capacity one.

© Petteri Kaski 2006

First we prove that N has a flow of value $|X|$ if and only if there exists a matching in H that saturates every vertex in X:

($\Leftarrow$) Let M be a matching that saturates every vertex in X. Put $f(e) = 1$ for every edge $e \in M$, and put $f(su) = 1$ (respectively, $f(vt) = 1$) for each saturated vertex $u \in X$ (respectively, $v \in Y$). For all other edges e in G, put $f(e) = 0$. Clearly, f is a flow on N with $w(f) = |X|$.

($\Rightarrow$) A flow f on N with $w(f) = |X|$ is clearly a maximum flow on N. By the integral flow theorem we can assume f to be a zero-one flow. Thus, we obtain a matching M that saturates all vertices in X by simply including all edges $e = uv$ with $u \in X$, $v \in Y$ and $f(e) = 1$ into M.

© Petteri Kaski 2006

Suppose that a maximum flow on N has value less than $|X|$ (i.e. H does not have matching that saturates all vertices in X). By the max-flow min-cut theorem, there exists a cut (S, T) of N with $c(S, T) < |X|$.

We show that Hall's condition fails for some $A \subseteq X$. Moreover, by the construction of N,

$$c(S, T) \geq |T \cap X| + |T \cap N(S \cap X)| + |S \cap Y|.$$

Also,

$$|N(S \cap X)| \leq |T \cap N(S \cap X)| + |S \cap Y|.$$

Combining the inequalities, we obtain

$$|N(S \cap X)| < |X| - |T \cap X| = |S \cap X|.$$

Thus, Hall's condition fails for $A := S \cap X$.

 © Petteri Kaski 2006

Menger's theorem

Let G be a (directed) graph and let s, t be two distinct vertices of G.

A set of paths from s to t is **edge-disjoint** (**vertex-disjoint**) if no two paths have an edge (a vertex $v \neq s, t$) in common.

A set X of edges (vertices) **separates** s from t if every path from s to t in G contains an edge (a vertex $v \neq s, t$) from X.

The following theorem is essentially due to Menger (1927).

Theorem A.39 *Let G be a directed graph and let s, t be two distinct vertices of G. Then, the maximum number of edge-disjoint paths from s to t is equal to the minimum number of edges separating s from t.*

 © Petteri Kaski 2006

Proof: Let $N = (G, c, s, t)$ be the flow network with $c(e) = 1$ for all $e \in E(G)$.

Clearly, any k edge-disjoint paths from s to t give a flow of value k on N. Conversely, let f be a maximum flow on N. We can assume that f is a zero-one flow by the integral flow theorem. A zero-one flow of value k can be used to construct k edge-disjoint paths from s to t: walk along unmarked edges with $f(e) = 1$ from s until t is reached, mark each edge as it is traversed; when t is reached, decrement the flow on each marked edge by one and construct a path from the marked edges by removing cycles caused by repeated vertices. By (F2) the walk can terminate only at t. The value of the flow decreases by one as each path is constructed.

Thus, the maximum number of edge-disjoint paths from s to t in G is equal to the maximum value of a flow on N, which is by the max-flow min-cut theorem equal to the minimum capacity of a cut of N.

 © Petteri Kaski 2006

Proof: (continued) Clearly, any cut (S, T) of N defines a set of $c(S, T)$ edges $X = \{e \in E(G) \ : \ e^- \in S, \ e^+ \in T\}$ that separates s from t.

Conversely, suppose $X \subseteq E(G)$ separates s from t and that X is minimal (i.e. no edge can be removed from X without violating the separation property). Let S be the set of vertices reachable from s along a path that does not contain an edge from X. By definition of X, $t \notin S$. Put $T := V(G) - S$. Clearly, (S, T) is a cut of N. We have $e \in X$ for all edges e with $e^- \in S$ and $e^+ \in T$ because otherwise $e^+ \in S$, a contradiction. Because X is minimal, $|X| = c(S, T)$.

Thus, the minimum capacity of a cut of N is equal to the minimum number of edges separating s from t in G, which completes the proof. $\square$

 © Petteri Kaski 2006

Vertex-disjoint paths

Theorem A.40 *Let G be a directed graph and let s, t be two distinct nonadjacent vertices of G. Then, the maximum number of vertex-disjoint paths from s to t is equal to the minimum number of vertices separating s from t.*

Proof: Construct from G a directed graph G' as follows: the vertices of G' are s, t and two vertices v', v'' for each vertex $v \neq s, t$ in G. For any edge sv or vt in G, G' contains the edge sv' or $v''t$, respectively. Furthermore, G' contains the edge $u''v'$ for each edge uv in G, and the edge $v'v''$ for each vertex $v \neq s, t$.

Clearly, a set of k vertex-disjoint paths from s to t in G can be transformed into a set of k edge-disjoint paths from s to t in G'.

Proof: (continued) Conversely, a set of k edge-disjoint paths in G' can be transformed to vertex-disjoint paths in G. This is because any path from s to t in G' that contains v' or v'' must also contain the edge $v'v''$ by the structure of G'.

Any set $X' \subseteq E(G')$ that separates s from t in G' can be transformed to a set $X \subseteq V(G)$ with $|X| = |X'|$ that separates s from t in G as follows. First, replace each edge $u''v' \in X'$ with either $u'u''$ or $v'v''$. Similarly, replace each edge $su' \in X'$ with $u'u''$, and each $v''t \in X'$ with $v'v''$. The resulting set still separates s from t in G'. Now, for each edge $v'v'' \in X'$, put $v \in X$ to obtain a set X of vertices that separates s from t in G. The converse construction from $X \subseteq V(G)$ to $X' \subseteq E(G')$ with $|X| = |X'|$ is obvious: for each $v \in X$ put $v'v'' \in X'$ to obtain a set of edges that separates s from t in G'.

The claim now follows since Theorem A.39 holds for G'. $\square$

The undirected case

Theorems A.39 and A.40 hold also in the case when G is undirected. We first consider Theorem A.40.

Construct from G a directed graph G' with $V(G) = V(G')$ that contains for each undirected edge uv in G the two directed edges uv and vu.

Clearly, a set $X \subseteq V(G')$ separates s from t in G' if and only if it does so in G. Moreover, any set of k vertex-disjoint directed paths in G' can be transformed into a set of k vertex-disjoint paths in G and vice versa. Thus, Theorem A.40 holds for undirected graphs.

Theorem A.39 requires a bit more work in the undirected case.

Consider a set of k edge-disjoint paths in G'. Suppose two paths P, Q use the same undirected edge uv in the opposite directions, say P in the direction uv and Q in the direction vu. Replace these paths with paths

$$P' = P_{su} + Q_{ut}, \quad Q' = Q_{sv} + P_{vt}$$

and iterate until no such paths P, Q exist. The resulting k paths use each pair of edges uv, vu in at most one direction and thus can be transformed into k edge-disjoint paths in G.

Thus, any k edge-disjoint paths in G' can be transformed into k edge-disjoint paths in G and vice versa.

Let k be the maximum number of edge-disjoint paths from s to t in G. By the previous construction, k is also the maximum number of edge-disjoint paths from s to t in G'. By Theorem A.39, G' contains a set $X' \subseteq E(G')$ that separates s from t and $|X'| = k$. Let X be the corresponding set of edges in G. Clearly, X separates s from t in G (otherwise X' would not separate s from t in G'). Furthermore, $|X| \geq k$, because X must contain at least one edge from each path in a set of the k edge-disjoint paths in G. By construction of X, $|X| \leq |X'| = k$. Thus, $|X| = k$. This shows that Theorem A.39 holds also in the undirected case.

© Petteri Kaski 2006