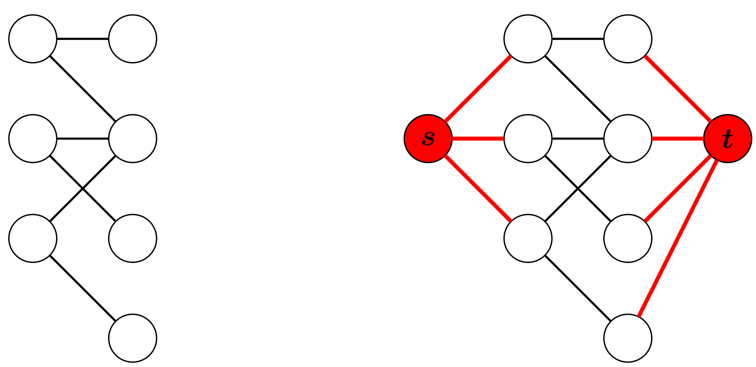
Graph Theory

Matchings and the max-flow min-cut theorem

Instructor: *Nicolò Cesa-Bianchi* version of April 11, 2020

A set of edges in a graph $G = (V, E)$ is independent if no two edges have an incident vertex in common. Independent sets of edges are called **matchings**. M is a matching of $U \subseteq V$ if every vertex in U is incident with some edge in M. The vertices in U are then called matched (by M); vertices not incident with any edge of M are unmatched. A **perfect matching** in $G = (V, E)$ is a matching of all vertices in V. We want to find conditions ensuring the existence of large or perfect matchings in arbitrary graphs.

Matchings in bipartite graphs. An important special case of matching considers bipartite graphs $G = (V, E)$ where $V = X \cup Y$ and $X \cap Y = \emptyset$. We study this special case through the lens of the max-flow min-cut theorem. First, we transform G in a flow network $G' = (V', E')$ where $V' \equiv V \cup \{s, t\}$ and $E' = E \cup \{(s, x) : (x \in X\} \cup \{(y, t) : y \in Y\}$, see figure below here.



The **max flow** problem in G' is to find an admissible flow of maximum value between s and t under a capacity constraint $c : E \to \mathbb{R}$, where $c(e) \geq 0$ is a nonnegative capacity assigned to each edge.

A **flow** is a function $f : E \to \mathbb{R}$ assigning $f(e) \geq 0$ to each e of G'. A flow f is **admissible** when the two following set of constraints are satisfied:

1. **capacity** constraints, $f(e) \leq c(e)$ for all $e \in E$.
2. **flow conservation** constraints, for all $v \in V$,

$$f(s, x) = \sum_{y \in Y \,:\, (x,y) \in E} f(x, y) \qquad x \in X$$

$$f(y, t) = \sum_{x \in X \,:\, (x,y) \in E} f(x, y) \qquad y \in Y$$

The **value** of an admissible flow f is

$$V_f = \sum_{x \in X} f(s, x) = \sum_{y \in Y} f(y, t)$$

A **cut** of G' is a partition S, T of V' such that $s \in S$ and $t \in T$. The **cost** of a cut is $c(\Gamma) = \sum_{e \in \Gamma} c(e)$, where $\Gamma = \Gamma(S,T) = \{(u,v) \in E \,:\, u \in S,\, v \in T\}$. The **min-cut** problem is to find a cut of minimum cost. The next result, which we do not prove here, is a fundamental consequence of linear programming duality.

Theorem 1 (Max-flow min-cut) *In any flow network, the maximum value of an admissible flow equals the minimum cost of a cut.*

We also use (without proof) this important fact.

Theorem 2 (Integral flow) *If each edge in a flow network has integral capacity, then there exists an integral admissible flow of maximum value.*

Fact 3 *Let G be a bipartite graph and let G' be the flow network derived by G such that $c(e) = 1$ for all $e \in E'$. Then the value of the maximum flow equals the size of a maximum matching in E.*

Proof. Due to the integral flow theorem, and recalling that $c(e) = 1$ for all $e \in E'$, there exists a maximum flow f^* such that $f^*(e) \in \{0,1\}$ for all $e \in E'$. Due to the flow conservation constraints, if $f^*(x,y) = 1$ for some $(x,y) \in E$, then it must be $f^*(x,y') = 0$ for all $y' \in Y \setminus \{y\}$ and $f^*(x',y) = 0$ for all $x' \in X \setminus \{x\}$. Hence f^* defines a set of V_f edge-disjoint paths from s to t of the form $P = \{(s,x),(x,y),(y,t)\}$ with $f(e) = 1$ if and only if $e \in P$. This implies that

$$M = \{(x,y) \in E \,:\, f^*(x,y) = 1\}$$

is a matching in G of size $|M| = V_{f^*}$. Hence the maximum matching M^* satisfies $|M^*| \geq V_{f^*}$.

For the other direction, let M^* be a maximum matching in G. Then there exists an edge-disjoint path $P = \{(s,x),(x,y),(y,t)\}$ in G' for each $(x,y) \in M^*$. Let f be the flow such that $f(e) = 1$ if and only if e belongs to one of these paths. This f is admissible and has value $V_f = |M^*|$. This implies that the maximum flow f^* satisfies $V_{f^*} \geq |M^*|$, and the proof is concluded. $\square$

A **vertex cover** of a graph $G = (V,E)$ is any subset $U \subseteq V$ such that for any $(x,y) \in E$ it holds that $U \cap \{x,y\} \not\equiv \emptyset$. We use the max-flow min-cut theorem to prove the next result.

Theorem 4 (König, 1931) *The maximum cardinality of a matching in a bipartite graph G is equal to the minimum cardinality of a vertex cover of its edges.*

Proof. Let M be a maximum matching in G. We first show that there exists a vertex cover of size equal to $|M|$. Let G'_∞ be the flow network derived from G with capacity c such that $c(s,x) = c(y,t) = 1$ for all $x \in X$ and $y \in Y$, and $c(x,y) = \infty$ for all $(x,y) \in E$. Now, the minimum cut S, T must be such that

$$\Gamma(S,T) \equiv \{(y,t) \,:\, y \in S\} \cup \{(x,s) \,:\, x \in T\}$$

because all edges between X and Y have infinite capacity. Therefore, $c(S,T) = |A|$ where $A = (X \cap T) \cup (Y \cap S)$. We know show that A is a vertex cover. Indeed, if there exists $(x,y) \in E$ not incident in A, then $y \notin S$ and $x \notin T$, which is equivalent to $x \in S$ and $y \in T$. But then

$c(S,T) = \infty$ contradicting the minimality of S, T. So A is a vertex cover of size $c(S,T)$. By the max-flow min-cut theorem, this is also the value of the maximum flow, which because of Fact 3 is equal to the size of the maximum matching.

Next, we show that no vertex cover can be smaller than $|M|$. Indeed, if A is a vertex cover, then it must cover the edges in M. As each $v \in A$ can cover at most one edge of M (because M is a matching), we conclude that $|A| \geq |M|$. $\square$

The next theorem gives a characterization of bipartite graphs that contain a perfect matching. If $G = (V, E)$ with $V = X \cup Y$ is bipartite, then it can contain a perfect matching only if $|X| = |Y|$. For all $W \subseteq V$, let $N(W) \equiv \bigcup_{w \in W} N(w)$.

Theorem 5 (Hall, 1935) *A bipartite graph G such that $|X| = |Y|$ contains a perfect matching of if and only if $|N(W)| \geq |W|$ for all $W \subseteq X$.*

Hall's theorem is also known as the marriage theorem, where the vertices are viewed as individuals in two disjoint groups and edges represent a potential relationship between two individuals.

PROOF. Assume G has a perfect matching and fix any $W \subseteq X$. Then each $w \in W$ is uniquely matched to a $y \in N(w)$. But this is impossible unless $|N(W)| \geq |W|$.

Vice versa, assume $|N(W)| \geq |W|$ for all $W \subseteq X$ holds and build the flow network G'_∞ from G as we did in the proof of König's theorem. We consider two cases.

Case 1. There exists a flow f with value $|X|$. Then Fact 3 implies that there exists a matching of size $|X|$ which must then be perfect.

Case 2. Any flow f has value $V_f < |X|$. Then the max-flow min-cut theorem implies that the minimum cut S, T has cost $k = c(S,T) < |X|$. From the proof of König's theorem, we also know that

$$\left|X \cap T\right| + \left|Y \cap S\right| = k < |X| = \left|X \cap T\right| + \left|X \cap S\right|$$

Therefore, $\left|Y \cap S\right| < \left|X \cap S\right|$. Now set $W \equiv X \cap S$ and note that $N(W) \subseteq Y \cap S$ must hold. Otherwise, there exists $y \in T$ such that $(x, y) \in E$ for some $x \in S$. But this implies $c(S,T) = \infty$ and we have a contradiction. Hence, $|N(W)| \leq |Y \cap S| < |W|$, which contradicts our initial assumption. $\square$

Matching in arbitrary graphs. We move on to state and prove a generalization of Hall's theorem, characterizing the existence of a perfect matching in an arbitrary graph.

Theorem 6 (Tutte, 1947) *A graph $G = (V, E)$ has a perfect matching if and only if for every subset $U \subseteq V$, the subgraph induced by $V \setminus U$ has at most $|U|$ connected components with an odd number of vertices.*

Note that a graph of odd order cannot clearly have a perfect matching, and this is captured by the choice $U \equiv \emptyset$ since in this case G must have at least one odd component.

PROOF. In the following, for any graph $G = (V, E)$ and for all $S \subseteq V$ we write $G - S$ to denote the subgraph induced by $V \setminus S$. Also $o(G)$ denotes the number of components of odd order in G.

Consider a graph G, with a perfect matching. Pick any $U \subseteq V$ and let C be an arbitrary odd component in $G - U$. Since G had a perfect matching, at least one vertex in C must be matched to a vertex in U. Hence, each odd component has at least one vertex matched with a vertex in U. Since each vertex in U can be in this relation with at most one connected component (because of it being matched at most once in a perfect matching), $o(G - U) \leq |U|$.

Now let $G = (V, E)$ be a graph without a perfect matching. A **bad set** is a set $S \subseteq V$ violating Tutte's condition $o(G - S) \leq |S|$. Our task is to find a bad set.

We may assume that G is of even order (otherwise we know the empty set is a bad set) and edge-maximal (adding any edge would create a perfect matching). Indeed, if G' is obtained from G by adding edges and $S \subseteq V$ is bad for G', then S is also bad for G: any odd component of $G' - S$ is the union of components of $G - S$ and at least one of these must be odd, so $o(G' - S) \leq o(G - S)$.

Claim 7 *Assume $G = (V, E)$ has even order, is edge-maximal, and has no perfect matching. Then $S \subseteq V$ is bad if and only if*

1. *$G - S$ is a union of disjoint cliques,*
2. *every $s \in S$ has degree $|V| - 1$.*

Proof of Claim. Assume S is a bad set and $G - S$ has a component with a missing edge. Note that adding this edge cannot turn S into a good set (odd components remain odd components and $|S|$ does not increase). Since G is edge-maximal, adding the edge creates a perfect matching. This contradicts what we already proved, namely that bad sets prevent perfect matchings. Now assume S has a vertex of degree smaller than $|V| - 1$. Just like before, adding this missing edge does not turn S into a good set. However, since G is edge-maximal, adding the edge creates a perfect matching and we have again a contradiction.

Thanks to Alberto Boggio for simplifying this part of the proof.

Conversely, if a set $S \subseteq V$ satisfies both conditions in the claim, then S must be bad. For the purpose of contradiction, assume S is good. If $S \equiv \emptyset$ satisfies the conditions, then the $o(G) = 0$ and so G has zero odd components. So G is a set of disjoint cliques all of even order, contradicting the hypothesis that G has no perfect matching. If $S \not\equiv \emptyset$ satisfies the conditions, then $o(G - S) \leq |S|$. Hence we can fix each odd components of $G - S$ using a vertex from S (a different vertex for each odd component), and pair up all the remaining vertices in S because $|V|$ is even. Again this contradicts the hypothesis that G has no perfect matching. $\square$

So it suffices to prove that any even-order and edge-maximal G without a perfect matching has a set S of vertices satisfying the two conditions of Claim 7. For the purpose of contradiction, let assume that G has no perfect matching and there is no S satisfying the two conditions. Let S be the (possibly empty) set of vertices that are adjacent to every other vertex. If $G - S$ is not a union of disjoint cliques, then some component of $G - S$ has non-adjacent vertices x, y. Let x, a, b be the first three vertices on a shortest x–y path in this component; then $(x, a), (a, b) \in E$ but $(x, b) \notin E$. Since $a \notin S$, there is a vertex $c \in V$ such that $(a, c) \notin E$. By the maximality of G, there is a matching M_1 of V in $G + (x, b)$, and a matching M_2 of V in $G + (a, c)$. Observe that surely $(x, b) \in M_1$ and $(a, c) \in M_2$.

Let P be the edges of a maximal path in G that starts from c with an edge from M_1 and whose edges alternate between M_1 and M_2. How can P end? Unless we are arrive at x, a or b, which are adjacent to edges not in G, we can always continue (recall that M_1 and M_2 are perfect matchings).

Let v denote the last vertex in P. If the last edge of P is in M_1, then v has to be a, since otherwise we could continue with an edge from M_2. In this case we let C be the cycle $P(a, c)$. If the last edge of P is in M_2, then surely $v \in \{x, b\}$ for analogous reasons, and we let C be the cycle $P(v, a)(a, c)$.

In each case, C is an even cycle in $G + (a, c)$ with every other edge in M_2, and whose only edge not in E is $(a, c) \in M_2$. Note that all vertices of C (including a and c) are matched by M_1, whereas the remaining vertices of V are matched by M_2 by construction. Adding to $M_2 \setminus C$ the edges in $C \cap M_1$, we obtain a matching of V contained in E, a contradiction. □

The **Tutte–Berge formula** says that the size of a maximum matching of a graph $G = (V, E)$ equals

$$\frac{1}{2} \min_{U \subseteq V} \Big(|U| - o(G - U) + |V| \Big)$$

Tutte's condition is equivalent to say that the expression inside the minimum is at least $|V|$. So,

- Tutte's condition holds
- the graph has a matching of size at least $|V|/2$
- the graph has a perfect matching

are equivalent statements.

The first algorithm for finding maximum matchings in arbitrary graphs is the Blossom algorithm by Jack Edmonds (1965) which runs in time $\mathcal{O}\big(|E||V|^2\big)$. For bipartite graphs, John Hopcroft and Richard Karp (1973) designed an improved algorithm with runnning time $\mathcal{O}\big(|E|\sqrt{|V|}\big)$. The algorithm by Silvio Micali and Vijay Vazirani (1980) achieves the same running time $\mathcal{O}\big(|E|\sqrt{|V|}\big)$ and works on arbitrary graphs.

A consequence of König's Theorem for partially ordered sets. A finite partially ordered set (poset) is a pair $(S, \leq)$ where S is a finite set and $\leq$ is a reflexive, antisymmetric, and transitive binary relation:

- $s \leq s$ for all $s \in S$ (reflexivity)
- $s \leq t$ and $t \leq s$ if and only if $s = t$ for all $s, t \in S$ (antisymmetry)
- $r \leq s$ and $s \leq t$ implies $r \leq t$ (transitivity)

We write $s < t$ if $s \leq t$ and $s \neq t$. We can represent a finite poset with a **directed acyclic graph** (DAG) $G = (V, E)$ where $V \equiv S$ and $(s, t) \in E$ if and only if $s < t$. The graph is acyclic because a directed cycle would imply $s < s$ for any s on the cycle due to transitivity.

An **antichain** $A \subseteq S$ is such that for all $s, t \in A$ with $s \neq t$, $s \not< t$ and $t \not< s$. Note that an antichain corresponds to an independent set in the DAG associated with the poset. A **chain** $C \subseteq S$ is such that for all $s, t \in C$ with $s \neq t$, $s < t$ or $t < s$. Note that a chain, which can be a singleton, is a total order over a subset of S. A set of chains covers $(S, \leq)$ if their union is S.

Theorem 8 (Dilworth, 1950) *In a finite poset $(S, \leq)$, the cardinality of the largest antichain A equal to the smallest number N of chains that cover the poset.*

PROOF. If A is an antichain, then clearly S cannot be covered by fewer than $|A|$ chains because every chain intersects at most one element of the antichain, hence $N \geq |A|$. Now define a bipartite graph G where $A \equiv \{a_s : s \in S\}$, $B \equiv \{b_s : s \in S\}$, and where $(a_s, b_t) \in E$ if and only if $s < t$ in S. Fix a vertex cover K and let $A \equiv \{s \in S : a_s, b_s \notin K\}$. Then A is an antichain (no two $s, t \in A$

are such that $s < t$, otherwise $(a_s, b_t) \in E$ is not covered by K). Moreover, if s is such that a_s or b_s are in K (possibly both), then $s \notin A$. Therefore $|A| \geq |S| - |K|$.

Now choose a matching M of G and consider the elements of M as directed edges (a_s, b_s). Every $s \in S$ can have at most one incoming edge (if b_s is matched) and one outgoing edge (if a_s is matched). This means that M corresponds to a set of chains where the end of each chain corresponds to an unmatched vertex in M. Then $N \leq |S| - |M|$. By König's theorem, there exists a matching M, and a vertex cover K such that $|M| = |K|$. So $N \leq |S| - |K| \leq |A|$ and the proof is concluded. $\square$

Other consequences of the max-flow min-cut Theorem. Any undirected graph $G = (V, E)$ can be viewed as a flow network between two distinct $s, t \in V$. Let $\mathcal{P}$ be the set of all paths in G between s and t, and let $F : \mathcal{P} \to \mathbb{R}$ a function assigning a flow $F(P) \geq 0$ to each path $P \in \mathcal{P}$. Then F is admissible when

$$\sum_{P \,:\, e \in \mathcal{P}} F(e) \leq c(e) \qquad e \in E$$

and has value

$$V_F = \sum_{P \in \mathcal{P}} F(P)$$

Theorem 9 (Max-flow min-cut, alternative version) *The maximum value of an admissible flow F equals the minimum cost of a cut.*

When all edge capacities are 1, a consequence of this fact together with the flow integrality theorem is that any admissible flow F corresponds to V_F edge-disjoint paths. This immediately implies the following result.

Corollary 10 (Menger, 1927 — edge version) *Let $G = (V, E)$ be a graph and $s, t \in V$. Then the minimum number of edges to cut for separating s from t in G is equal to the maximum number of edge-disjoint paths between s and t.*

We now state and prove a version of Menger's Theorem using vertices instead of edges. Although this version can still be proved using the max-flow min-cut theorem, we provide a direct proof from first principles.

Given $G = (V, E)$ and $A, B \subseteq V$, a AB-**separator** in G is a subset $X \subseteq V$ such that every path from a vertex in A to a vertex in B contains a vertex from X. Note that A and B are both AB-separators. Also, if $A \subseteq B$, then A is an AB-separator. If there are no paths between A and B, then the empty set is the only AB-separator. A path is an AB-path if its first vertex is in A, its last vertex is in B, and none of its internal vertices is in A or B. Two paths are disjoint if the set of their vertices is disjoint. An AB-**connector** is a set of paths between A and B that are pairwise disjoint.

Theorem 11 (Menger, 1927 — vertex version) *Let $G = (V, E)$ be a graph a and $A, B \subseteq V$. Then the size of any minimum AB-separator in G is equal to the size of any maximum number AB-connector in G.*

Proof. Throughout the proof, let $k \geq 0$ be the size of a minimum AB-separator in G. Clearly, G cannot contain a AB-connector of size larger than k. Indeed, if $X \subseteq V$ is a AB-separator of size k and there are more than k disjoint AB-paths, then two of them must share a vertex in X.

By induction on the number of edges in G, we now show how to construct an AB-connector of size k. If G has no edges, then the minimum AB-separator is $A \cap B$, which is also an AB-connector consisting of single-vertex paths. So the theorem holds in this case.

Assume now G has at least one edge $e = (x, y)$. Let $G/e = (V', E')$ be the graph obtained by replacing x, v with a vertex v_e adjacent to $N(x) \setminus \{y\} \cup N(y) \setminus \{x\}$.

Case 1. G/e has a minimal AB-separator of size k. Then, by the induction hypothesis, there is also an AB-connector of size k in G/e. Since any path that goes through v_e in G/e corresponds to a path that goes through e in G, there is an AB-connector of size k also in G.

Case 2. G/e has a minimal AB-separator S with $|S| < k$. Then $v_e \in S$ otherwise $S \subseteq V$ and so S is a AB-separator in G of size less than k. Then $X = (S \setminus \{v_e\}) \cup \{x, y\}$ is a AB-separator in G of size k (it can not be less than k).

We now consider the graph $G - e$. Every AX-separator in $G - e$ is also an AB-separator in G, because any AB-path in G goes through X and $e = (x, y)$ does not create AX-paths in G as $x, y \in X$. Since X is a minimal AB-separator in G, this other separator must contain at least k vertices. So by induction there is a AX-connector of size k in $G - e$, and similarly there is a XB-connector of size k in $G - e$. As X separates A from B, these two path systems do not meet outside X, and can thus be combined to k disjoint AB-paths. $\square$

As a curiosity, we state without proof one of the consequences of the max-flow min-cut theorem that are not expressed in graph-theoretic terms.

A square matrix D with real non-negative entries is doubly stochastic if the sum of the entries in any row and any column equals 1. A **permutation matrix** is a **doubly stochastic** matrix with entries 0 and 1, that is, a matrix with exactly one 1 in each row and in each column. A matrix A is a convex combination of matrices $A_1, \ldots, A_n$ if there exist reals $\lambda_1, \ldots, \lambda_n \geq 0$ such that

$$A = \sum_{i=1}^{n} \lambda_i A_i \qquad \sum_{i=1}^{n} \lambda_i = 1 \ .$$

The set of all convex combinations of permutation matrices is the **Birkhoff polytope**.

Theorem 12 (Birkhoff-Von Neumann) *The Birkhoff polytope is exactly the set of all doubly stochastic matrices.*