Non-bipartite Matching

E.de Klerk
Delft University of Technology

Week 4: Slides about Chapter 5 of the Lecture notes.
Matchings and odd components

Let $G = (V, E)$ be given.

**Definition:** A matching is a subset $M \subseteq E$ such that $e \cap e' = \emptyset$ for all $e, e' \in M, e \neq e'$.

A matching is called *perfect* if it covers all vertices (size $= \frac{1}{2}|V|$).

**Definition:** For any subset $U \subseteq V$ we denote the set of edges with both end nodes in $U$ as $E_U$ (or $EU$). The graph $(U, E_U)$ is denoted as $G_U$. The number of odd components in $G_U$ is denoted as $o(U)$. The complement of $U$ in $V$ is denoted as $\bar{U}$. The graph obtained by removing the nodes in $U$ and all edges connected to it, is denoted as $G - U$. 
Weak duality for matchings

If $M$ is a matching and $U \subseteq V$ then one has

$$2|V| - |U| - o(U) \geq 2|M|$$

**Proof:** Let $U_i \ (1 \leq i \leq k)$ be the connected components of $G_U = (U, E_U)$. Then the number of edges of $M$

in $G_U$, is $\leq \frac{1}{2}|U_i|$ if $|U_i|$ is even

$\leq \frac{1}{2}(|U_i| - 1)$ if $|U_i|$ is odd

not in $G_U$ is $\leq |\bar{U}| = |V| - |U|$

Taking the sum we obtain

$$|M| \leq |V| - |U| + \frac{1}{2}(|U| - o(U)) = |V| - \frac{1}{2}(|U| + o(U)).$$

It follows from the above proof that equality holds if and only the edges in $M$ form perfect matchings for the even components of $G_U$ and 'almost perfect matchings' for the odd components; apart from these edges there are precisely $|\bar{U}|$ edges that connect every node outside $U$ with an odd component of $G_U$. 
Graphical illustration

$|U_i|$ is even

$|U_i|$ is odd

$\bar{U}$
**Tutte's 1-factor theorem**

**Theorem** $G$ has a perfect matching if and only if

$$|U| + o(U) \leq |V|$$

for each $U \subseteq V$.

**Proof:**

'⇒': Assume $G$ has a perfect matching $M$ ($|M| = \frac{1}{2}|V|$). Then

$$2|V| - |U| - o(U) \geq 2|M| = |V|$$

for any $U \subset V$, which implies the desired inequality.

'⇐': Much harder: see course syllabus Thm. 5.1.
Tutte’s theorem and a corollary

**Theorem** Let $\nu(G)$ denote the matching number. We have:

$$\nu(G) = \min_{U \subseteq V} \left( |V| - \frac{|U| + o(U)}{2} \right)$$

$$= |V| - \max_{U \subseteq V} \frac{|U| + o(U)}{2}$$

**Corollary**
Assume $G$ has no isolated vertices. Denote by $\rho(G)$ the edge cover number of $G$. One has:

$$\rho(G) = \max_{U \subseteq V} \frac{|U| + o(U)}{2}.$$

**Proof:** Use $\rho(G) = |V| - \nu(G)$ (Gallai’s theorem: $\rho(G) + \nu(G) = |V|$).
Finding a certificate $U$

For any $G = (V, E)$ let the subsets $D, A$ and $C$ of $V$ be defined by

\[ D = \{ v \in V : v \text{ is missed by a maximal matching} \} \]
\[ A = \{ v \notin D : v \text{ is neighbor of some } u \in D \} \]
\[ C = V \setminus (D \cup A) \]

Then the set $U = V \setminus A = C \cup D$ maximizes $|U| + o(U)$.


**Example:**

Matching $M$ (red) and sets $D$ (green) and $C$ (yellow)

\[ D = \{2, 4, 5, 8, 10, 12, 13\}, \quad A = \{3, 11\}, \quad C = \{1, 6, 7, 9\} \]
\[ |U| = 11, \ o(U) = 3 \quad \Rightarrow \quad |M| \leq 13 - \frac{1}{2}(11 + 3) = 6. \]
Weak duality for edge covers

Assume $G = (V, E)$ has no isolated vertices. If $F$ is an edge cover and $U \subseteq V$ then one has

$$|U| + o(U) \leq 2|F|$$

**Proof:** Let $U_i$ ($1 \leq i \leq k$) be the connected components of $G_U = (U, E_U)$. Then the number of edges in $F$ needed to cover the nodes of the $i$-th component is

$$\geq \frac{1}{2}|U_i| \quad \text{if } |U_i| \text{ is even}$$

$$\geq \frac{1}{2}(|U_i| + 1) \quad \text{if } |U_i| \text{ is odd}.$$

Taking the sum over all $i$ we obtain

$$|F| \geq \frac{1}{2} \left( \sum_{i=1}^{k} |U_i| + \# \text{ of odd components} \right).$$

It follows from the above proof that equality holds if and only the edges in $F$ fall apart in perfect matchings for the even components of $G_U$ and 'almost perfect matchings' for the odd components; each odd component has exactly one node that is covered by an edge of $F$ whose other end node is in the complement of $U$. 
Example

\[ G = (V, E) \]

Edge cover \( F \) (red arcs) and set \( U \) (green nodes)

\[ |F| = 5, \quad |U| = 6, \quad o(U) = 4. \]
Example

$G = (V, E)$

Edge cover $F$ (red arcs) and set $U$ (green nodes)

$|F| = 13, \quad |U| = 22, \quad o(U) = 4.$
Proof of Tutte's theorem for edge covers

To include the case where $G = (V, E)$ has isolated vertices, we let $\rho(G)$ denote the minimal number of nodes and edges needed to cover all nodes. Then we have

$$\rho(G) = \max_{U \subseteq V} \frac{|U| + o(U)}{2}$$

**Proof:** First we show the $\geq$ part. For any $U \subseteq V$ we have

$$\rho(G) \geq \rho(U) \geq \frac{1}{2} (|U| + o(U)).$$

For the inverse inequality we use induction on $|V|$. The case $|V| = 0$ is trivial. We proceed by considering a graph $G = (V, E)$ with $|V| > 0$, while assuming that the statement holds for all graphs with fewer than $|V|$ nodes. Without loss of generality we assume that $G$ is connected (verify!). This implies $o(V) = 0$ if $|V|$ is even and $o(V) = 1$ if $|V|$ is odd. We also assume that if a vertex can be covered by itself (without increasing the size of the node-edge cover) than we do so. So we minimize the number of edges in any node-edge cover.

First assume that $G$ has a vertex $v$ that is covered by an edge in each minimal node-edge cover. Then $\rho(G - v) = \rho(G)$, and by induction there exists a subset $U'$ of $V \setminus \{v\}$ such that

$$\rho(G - v) = \frac{|U'| + o(U')}{2}.$$

Taking $U = U'$ we obtain

$$\rho(G) = \rho(G - v) = \frac{1}{2} (|U'| + o(U')) = \frac{1}{2} (|U| + o(U)).$$
Proof of Tutte's theorem for edge covers (cont.)

\[ \rho(G) = \max_{U \subseteq V} \frac{|U| + o(U)}{2} \]

Now assume that for each vertex \( v \) there exists a minimal node-edge cover that covers \( v \) by a node. We show that a minimal node-edge cover then contains precisely one node.

Suppose on the contrary that a minimal node-edge cover \( F \) contains two nodes, \( u \) and \( v \) say. Choose \( F, u \) and \( v \) such that \( d(u, v) \) is minimal.

If \( d(u, v) = 1 \), then we can add the edge \( uv \) to \( F \), decreasing the size of \( F \) with 1, and hence contradicting that \( F \) is minimal. So \( d(u, v) \geq 2 \).

Let \( t \in V \) be an intermediate node on a shortest \( u-v \) path. By assumption, there exists a minimal node-edge cover \( H \) containing \( \{t\} \). Choose \( H \) such that the number of edges in \( F \cap H \) is maximal.

By the choice of \( F, u \) and \( v \), \( d(u, v) \) is minimal. Since \( d(u, t) < d(u, v) \) and \( d(t, v) < d(u, v) \), \( H \) covers both \( u \) and \( v \) by an edge. Since \( F \) and \( H \) contain the same number of nodes (and edges), there exists a node \( x \neq t \) that \( F \) covers by some edge \( e = xy \) and \( H \) by \( \{x\} \). Then \( y \) is covered by some edge \( f = yz \in H \), since otherwise \( H \) would contain \( \{x\} \) and \( \{y\} \), and we could replace these by the edge \( xy \), yielding a node-edge cover smaller than \( H \). Replacing \( H \) by \( (H \setminus (\{x\} \cup \{f\})) \cup (\{e\} \cup \{z\}) \) increases the number of edges in \( F \cap H \), yielding a contradiction with the choice of \( H \).

Hence we have shown that a minimal node-edge cover \( F \) contains at most 1 vertex, whence \( |V| = 2|F| \) or \( |V| = 2|F| - 1 \). In both case \( U := V \) yields the desired equality.
Finding a certificate $U$

For any $G = (V, E)$ let the subsets $D$ and $C$ of $V$ be defined by

$$
D = \{ v \in V : \{v\} \text{ occurs in some minimal node-edge cover} \} \\
C = \{ v \notin D : v \text{ is has no neighbor } u \in D \}
$$

Then the set $U = C \cup D$ maximizes $|U| + o(U)$.

**Proof:** We use induction to $|V|$. If $|V| \leq 2$ the statement is obvious. We proceed by considering a graph $G = (V, E)$ with $|V| > 0$, while assuming that the statement holds for all graphs with fewer than $|V|$ nodes. Without loss of generality we assume that $G$ is connected (verify!). This implies $o(V) = 0$ if $|V|$ is even and $o(V) = 1$ if $|V|$ is odd. We also assume that if a vertex can be covered by itself (without increasing the size of the node-edge cover) than we do so. In other words, we assume that in any node-edge cover the number of edges is minimized.

It suffices to show that there is a minimal node-edge cover $F$ of $G$ such that for the given set $U$ one has

$$
|U| + o(U) = 2|F|.
$$

We consider two cases: $A \neq \emptyset$ and $A = \emptyset$. 

Finding a certificate $U$: $A \neq \emptyset$

For any $G = (V, E)$ let the subsets $D, A$ and $C$ of $V$ be defined by

$$
D = \{ v \in V : \{v\} \text{ occurs in some minimal node-edge cover} \}
$$

$$
C = \{ v \notin D : v \text{ is has no neighbor } u \in D \}
$$

Then the set $U = C \cup D$ maximizes $|U| + o(U)$.

First consider $A \neq \emptyset$. Let $x \in A$. Then $xz \in E$ for some $z \in D$. Since $z \in D$, there is a minimal node-edge cover $F$ such that $\{z\} \in F$. Since $x \notin D$, $xy \in F$ for some $y$. This implies $y \neq z$. On the other hand, replacing in $F$ edge $xy$ by $xz$ and node $z$ by $y$ we get another minimal node-edge cover that contains $y$ as a node. Therefore, $y \in D$. 

Finding a certificate $U$: $A = \emptyset$

For any $G = (V, E)$ let the subsets $D, A$ and $C$ of $V$ be defined by

$$D = \{ v \in V : \{v\} \text{ occurs in some minimal node-edge cover} \}$$

$$C = \{ v \notin D : v \text{ is has no neighbor } u \in D \}$$

Then the set $U = C \cup D$ maximizes $|U| + o(U)$.
Finding maximal matchings

We saw in week 2 how to find maximal matchings in bipartite graphs.

Given a matching $M$, we used an $M$-augmenting path $P$ to obtain a larger matching

$$M' = M \Delta EP.$$ 

We want to use the same technique for general $G$, but must take care of odd cycles. (Bipartite graphs do not contain odd cycles (why?))

Idea: We will shrink some odd cycles, so-called blossoms, to a single vertex.
**Shrinking a graph**

Let $X, Y$ be sets. Define:

$$X/Y := \begin{cases} X & \text{if } X \cap Y = \emptyset \\ (X \setminus Y) \cup \{Y\} & \text{if } X \cap Y \neq \emptyset. \end{cases}$$

For $G = (V, E)$ and $C \subset V$ then $V/C$ replaces $C$ by a single vertex called $C$ (shrinks $C$ to a vertex).

An edge in $C$ becomes a loop, and an edge $(u, v)$ from $V \setminus C$ to $C$ becomes an edge $(u, C)$.

Loops can be removed — they play no role in matchings.
Augmenting paths and matchings

Recall the results of week 2:

Let $M$ be a matching in $G$.

**Definition:** A *path* $P = (v_0, v_1, \ldots, v_t)$ is called $M$-augmenting if:

1. $t$ is odd and $v_0, v_1, \ldots, v_t$ are distinct;
2. $v_1v_2, v_3v_4, \ldots, v_{t-2}v_{t-1} \in M$
3. $v_0, v_t \notin M$.

**Theorem:** Either $M$ has maximal cardinality or there is an $M$-augmenting path.
**Alternating paths**

**Definition:** A *path* 

\[ P = (v_0, v_1, \ldots, v_t) \]

is called *M*-alternating if exactly one of \( v_{i-1}v_i \) or \( v_i v_{i+1} \) belongs to \( M \), for each \( i \).

Every *M*-augmenting path is also *M*-alternating.

Let \( W \) be the vertices missed by \( M \). One can find the shortest *M*-alternating \( W - W \) path: Consider \( D = (V, A) \) where:

\[ A := \left\{ (w, w') \mid \exists x \in V : (w, x) \in E, (x, w') \in M \right\}. \]

*M*-alternating \( W - W \) paths are directed paths from a vertex in \( W \) to a vertex with at least one neighbour in \( W \).
**Alternating paths are not always augmenting**

There is an alternating path from the free node 2 to free node 3:

This path is not augmenting, which is due to the blossom:
Shrinking a blossom

Shrinking the blossom yields the following:

In the new graph there is no augmenting path. Hence the matching is maximal.
**M-blossoms**

**Definition:** An $M$-alternating path is called an $M$-blossom if $v_0, \ldots, v_{t-1}$ are distinct, $v_0 \in M$, and $v_t = v_0$.

**Theorem:** Let $C$ be an $M$-blossom in $G$. Then $M$ is maximal iff $M/C$ is maximal in $G/C$.

**Theorem** Let $P = (v_0, v_1, \ldots, v_t)$ be a shortest even-length $M$-alternating $W - v$ path. Then $P$ is simple or there exist $i < j$ such that $v_i = v_j$, $i$ is even, $j$ is odd, and $v_0, \ldots, v_{j-1}$ are distinct.
Algorithm for maximal matching

Input: A matching $M$.
Output: A matching $N$ with $|N| > |M|$ or a proof that $M$ is maximal.

1. Let $W \subseteq V$ be the vertices missed by $M$. Is there an $M$-alternating $W - W$ path? If no, STOP (no $M$-augmenting path exists). If yes, go to step 2.

2. Let $P = (v_0, v_1, \ldots, v_t)$ be the shortest such path.

3. If $P$ is $M$-augmenting, output $M \Delta EP$.

4. Choose $i < j$ such that $v_i = v_j$ and $j$ as small as possible.
   Reset $M := M \Delta \{v_0v_1, v_1v_2, \ldots, v_{i-1}v_i\}$. Now, $C := (v_i, v_{i+1}, \ldots, v_j)$ is an $M$-blossom. Shrink the blossom and go to step 1.
Weighted matching

General problem: Given a graph $G = (V, E)$ and a weight function:

$$w : E \mapsto \mathbb{Q}.$$ 

find a perfect matching $M$ minimizing

$$\sum_{e \in M} w(e).$$

We can assume $G$ has a perfect matching and that $w(e) \geq 0$ (why?)

Applications

- The Chinese postman problem;
- An approximation algorithm for the traveling salesman problem.