NDMI012: Combinatorics and Graph Theory 2

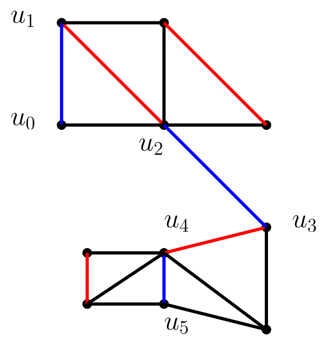
Lecture #2
Edmonds' Blossom algorithm

Irena Penev

Convention: In all our figures, edges of the matching in question are in red.[1]

1 M-augmenting paths

Let M be a matching in a graph G. An *M-alternating path* is a path $u_0, u_1, \dots, u_t$ in G such that every other edge of the path belongs to M (and the remaining edges do not). An *M-augmenting path* is an M-alternating path $u_0, u_1, \dots, u_t$ ($t \neq 0$) such that u_0, u_t are both unsaturated by M. For instance, in the picture below, $u_0, u_1, u_2, u_3, u_4, u_5$ is an M-augmenting path (as usual, the edges of the matching M are in red; the edges of the M-augmenting path that do not belong to M are in blue).



Lemma 1.1. *Let M be a matching in a graph G, and let $u_0, u_1, \dots, u_t$ be an M-augmenting path. Then t is odd and*

$$M' \;:=\; \Big(M \setminus \{u_1u_2, u_3u_4, \dots, u_{t-2}u_{t-1}\}\Big) \cup \{u_0u_1, u_2u_3, \dots, u_{t-1}u_t\}$$

is a matching of G satisfying $|M'| = |M| + 1$.

[1] This is not a standard convention. We simply use it in these lecture notes.

Proof. This follows from the relevant definitions. □

Theorem 1.2. *[Berge, 1957] Let M be a matching in a graph G. Then M is a maximum matching of G if and only if G has no M-augmenting path.*

Proof. We will prove the contrapositive: the matching M is **not** maximum if and only if G has an M-augmenting path.

If G has an M-augmenting path, then Lemma 1.1 guarantees that M is not a maximum matching of G.

Suppose now that M is not a maximum matching, and let M' be matching of G such that $|M'| > |M|$. Let $F := M \Delta M'$,[2] and let H be the graph with vertex set $V(H) = V(G)$ and edge set $E(H) = F$. Clearly, $\Delta(H) \leq 2$.[3] So, H is the disjoint union of paths and cycles.

Now, since $|M'| > |M|$, some component P of H has more edges of M' than of M. If P is a cycle, then we see that some vertex of P is incident two edges of M', contrary to the fact that M' is a matching. So, P is a path, and it is easy to see that it is in fact an M-augmenting path in G.[4] □

2 Blossoms and stems

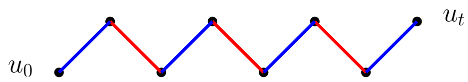
Our goal is to give a polynomial-time algorithm that finds a maximum matching in a graph. The basic idea is to start with an empty matching, and then repeatedly find augmenting paths and use them to find larger matchings (as in Lemma 1.1). We do this until no augmenting path remains, at which point Theorem 1.2 guarantees that our matching is maximum. We now need to show how we can find a maximum matching. In this section, we describe the basic tools that we need, and in the subsequent section, we describe the algorithm.

We begin with a definition. Suppose that M is a matching in a graph G. A *blossom* is a cycle $c_0, c_1, \dots, c_{2k}, c_0$ of length $2k+1$ (with $k \geq 1$) in G

[2]By definition, $M \Delta M' = (M \setminus M') \cup (M' \setminus M)$.

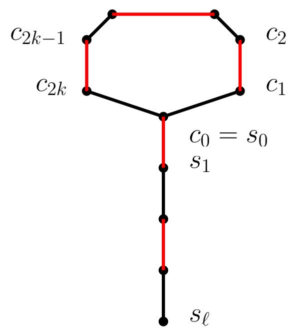
[3]Recall that $\Delta(H)$ is the maximum degree in H, i.e. $\Delta(H) = \max\{d_H(v) \mid v \in V(H)\}$. Let us check that $\Delta(H) \leq 2$. Since M and M' are matchings, we see that every vertex v of G is incident with at most one edge of M and at most one edge of M'. Since $V(H) = V(G)$ and $E(H) \subseteq M \cup M'$, it follows that every vertex of H is incident with at most two edges; thus, $\Delta(H) \leq 2$.

[4]Indeed, let P be of the form $u_0, u_1, \dots, u_t$. All edges of P are in $M \Delta M'$, and so since M and M' are both matchings, the edges of $M \setminus M'$ and $M' \setminus M$ alternate on P. Since P has more edges of M' than of M, we have that P has an odd number of edges (so, t is odd), and that $u_0u_1, u_2u_3, \dots, u_{t-1}u_t \in M' \setminus M$ and $u_1u_2, u_3u_4, \dots, u_{t-2}, u_{t-1} \in M \setminus M'$ (see the picture below; edges of $M \setminus M'$ are in red, and edges of $M' \setminus M$ are in blue).



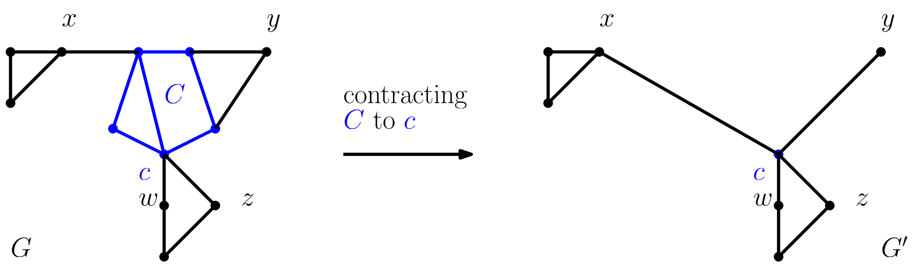
The fact that u_0, u_t are unsaturated by M follows from the construction of H, and from the fact that P is a component of H.

in which edges $c_1c_2, c_3c_4, \dots, c_{2k-1}c_{2k}$ belong to M, and the remaining $k+1$ edges do not belong to M. A *stem* for this blossom is an M-alternating path $s_0, \dots, s_\ell$ of even length[5] such that $s_0 = c_0$ is the unique common vertex of the cycle $c_0, c_1, \dots, c_{2k}, c_0$ and the path $s_0, \dots, s_\ell$, and s_ℓ is unsaturated by M.[6] The union of a blossom and a corresponding stem is called a *flower*.[7] An example is shown below.



Next, let G be a graph, and let $C \subseteq V(G)$ and $c \in C$. We say that G' is the graph obtained form G by *contracting* C to c if

- $V(G') = V(G) \setminus (C \setminus \{c\}) = (V(G) \setminus C) \cup \{c\}$, and
- $E(G') = \Big(\binom{V(G)\setminus C}{2} \cap E(G)\Big) \cup \Big\{xc \mid x \in V(G) \setminus C,\ \exists c' \in C \text{ s.t. } xc' \in E(G)\Big\}$.



Lemma 2.1. *Let M be a matching in a graph G, and let $C = c_0, \dots, c_{2k}, c_0$ be a blossom and $S = s_0, \dots, s_\ell$ a corresponding stem (in particular, $c_0 = s_0$). Let G' be the graph obtained from G by contracting C to c_0,[8] and let $M' = M \setminus E(C)$. Then M' is a matching of G'. Furthermore, M is a maximum matching of G if and only if M' is a maximum matching of G'.*

[5] So, the path has an even number of edges, and therefore, ℓ is even.

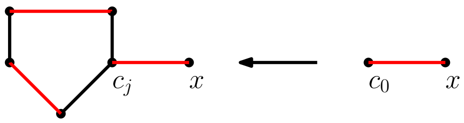
[6] Note that this implies that either $\ell = 0$ and $c_0 = s_0$ is unsaturated by M, or $\ell \geq 2$ and $s_0s_1 \in M$.

[7] Note that there may be more than one stem for a fixed blossom. Nonetheless, all stems attach to the same vertex of the blossom.

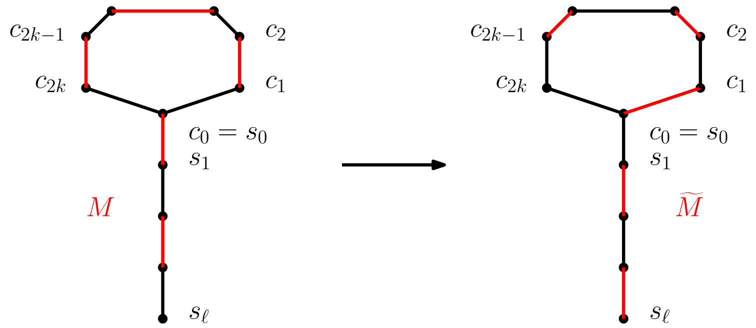
[8] Technically, we mean that G is obtained by contracting $V(C)$ to c.

Proof. The fact that M' is a matching of G' follows from the appropriate definitions.

Suppose first that M' is not a maximum matching of G'; we must show that M is not a maximum matching of G. Let M'' be a matching of G' of size greater than $|M'|$. If c_0 is unsaturated by M'', then $M'' \cup \Big(M \cap E(C)\Big)$ is a matching of G of size greater than $|M|$. Suppose now that c_0 is saturated by M''. Then there exists some vertex $x \in V(G) \setminus V(C)$ and an index $j \in \{0, \dots, 2k\}$ such that $xc_j \in E(G)$. But now the matching $(M'' \setminus \{xc_0\}) \cup \{xc_j\} \cup \{c_{j+1}c_{j+2}, c_{j+3}c_{j+4}, \dots, c_{j+2k-1}c_{j+2k}\}$ is a matching of G of size greater than $|M|$ (see the picture below).



Suppose now that M is not a maximum matching of G; we must show that M' is not a maximum matching of G'. First, let $\widetilde{M} := \Big(M \setminus (E(C) \cup E(S)\Big) \cup \{c_0c_1, c_2c_3, \dots, c_{2k-2}c_{2k-1}\} \cup \{s_1s_2, s_3s_4, \dots, s_{\ell-1}s_\ell\}$ and $\widetilde{M}' = \Big(M' \setminus E(S)\Big) \cup \{s_1s_2, s_3s_4, \dots, s_{\ell-1}s_\ell\}$.



Clearly, $\widetilde{M}$ is a matching of G of the same size as M, and $\widetilde{M}'$ is a matching of G' of the same size as M'. Since the matching M of G is not maximum, neither is $\widetilde{M}$; so, by Theorem 1.2, there exists an $\widetilde{M}$-augmenting path in G, say $P = p_0, \dots, p_t$. It now suffices to exhibit an $\widetilde{M}'$-augmenting path in G', for Theorem 1.2 will then imply that the matching $\widetilde{M}'$ is not maximum in G', and consequently, M' is not maximum in G', either.

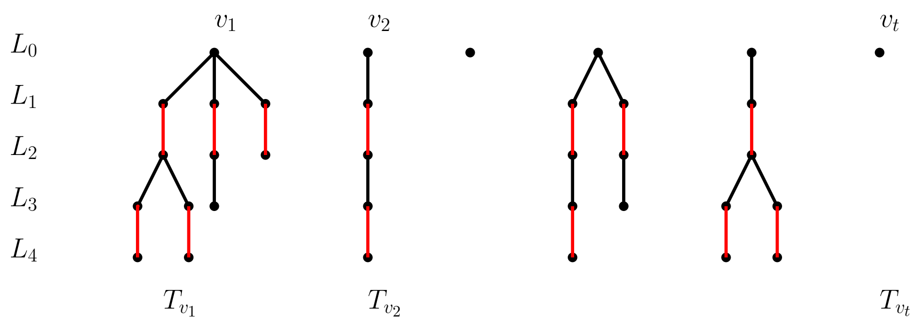
If $V(P) \cap V(C) = \emptyset$, then P is an $\widetilde{M}'$-augmenting path in G', and we are done. So, we may assume that $V(P) \cap V(C) \neq \emptyset$. First of all, c_{2k} is the only vertex in $V(C)$ that is unsaturated by $\widetilde{M}$; since both p_0, p_t are unsaturated by $\widetilde{M}$, we see that at most one of p_0, p_t belongs to $V(C)$. By symmetry, we may assume that $p_0 \notin V(C)$. Now, set $t_1 := \min\{i \in \{1, \dots, t\} \mid p_i \in V(C)\}$. But then $p_0, \dots, p_{t_1-1}, c_0$ is an $\widetilde{M}'$-augmenting path in G',[9] and we are done. $\square$

3 Edmonds' Blossom algorithm

In what follows, we will use the following notation: for a tree T and vertices $x, y \in V(T)$, we denote by $x - T - y$ the unique path between x and y in T.

Let G be an input graph. Initially, we start with the empty matching, and we iteratively increase the size of the matching until this is no longer possible, at which point, our matching is maximum. All we need to do is show how, given a matching M in G, we either produce a larger matching, or determine that no larger matching exists. We proceed as follows.

Step 1. First, we form an auxiliary forest F (which is a subgraph of G) as follows. $V(F)$ is partitioned into levels, $L_0, L_1, L_2, \dots$. Level L_0 consists of all vertices of G that are unsaturated by M. If $L_0 = \emptyset$, then M is a perfect (and therefore maximum) matching of G, and we are done. So, we may assume that $L_0 \neq \emptyset$. Then, using breadth-first-search, we form a (maximal) forest F in such a way that, for each integer $k \geq 0$, L_k is the set of all vertices of F that are at distance k from L_0 in F,[10] and moreover, for all even $k \geq 0$, edges between L_k and L_{k+1} in F do not belong to M, and edges between L_{k+1} and L_{k+2} in F do belong to M. For each $v \in L_0$, the unique component of F that contains v is the tree T_v rooted at v.

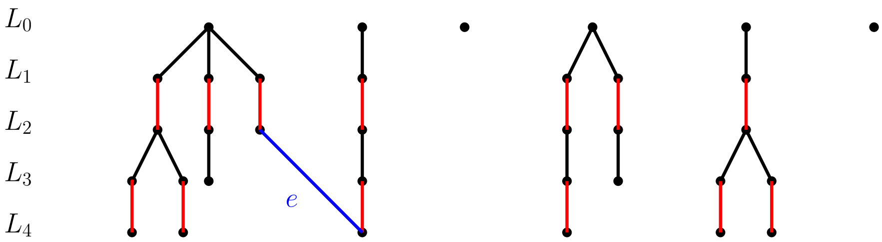


Step 2. If there exists an edge $e \in E(G)$ between even levels of two distinct trees, we immediately obtain an M-augmenting path,[11] and then we obtain a matching of size $|M| + 1$, as in Lemma 1.1.

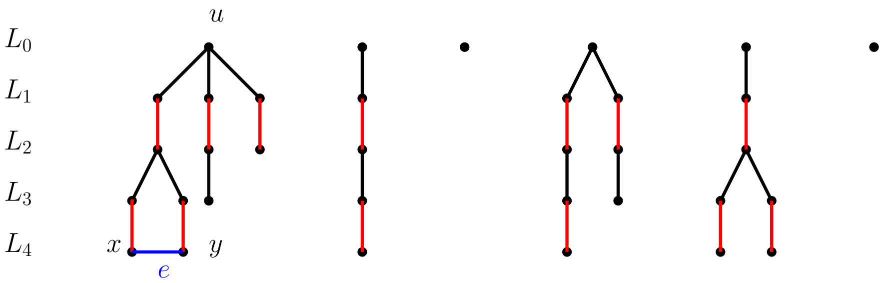
[9] We are using the fact that, by construction, c_0 is unsaturated by $\widetilde{M}'$ in G'.

[10] So: distance is counted in the forest F, and not in the whole graph G.

[11] Indeed, suppose that for distinct $u, v \in L_0$, and some even p, q, we have an edge e between a vertex $u' \in V(T_u) \cap L_p$ and a vertex $v' \in V(T_v) \cap L_q$. Then $u - T_u - u' - v' - T_v - v$ is an M-augmenting path in G.



If there exists an edge $e \in E(G)$ between two vertices, say x and y, belonging to even levels of the same tree T_u, then we can find a flower (i.e. a blossom with a corresponding stem), as follows.



We consider the (unique) path in T_u between x and u in T_u, and the (unique) path in T_u between y and u. The union of these two paths, together with the edge e, is a flower in G, say, with blossom $C = c_0, \dots, c_{2k}, c_0$ and stem $S = s_0, \dots, s_\ell$, where $c_0 = s_0$ and $s_\ell \in L_0$. Let G' be the graph obtained from G by contracting C to a vertex c_0, and let $M' = M \setminus E(C)$ (as in Lemma 2.1). We now call the algorithm with input G' and M'. Then there are two cases.

- If we obtain the answer that M' is a maximum matching in G', then (by Lemma 2.1) M is a maximum matching in G, and we are done.
- Suppose we obtained a matching M'' in G' that is of size greater than $|M'|$. If c_0 is unsaturated by M'', then $(E(C) \cap M) \cup M''$ is a matching in G of size greater than $|M|$, and we are done. Suppose now that c_0 is saturated by M'', and let $x \in V(G) \setminus V(C)$ be such that $xc_0 \in M''$. Let v be some vertex of C such that $xv \in E(G)$, and let M_C be the (unique) matching of size $\frac{|V(C)|-1}{2}$ in C, chosen so that v is M_C-unsaturated. Then $(M'' \setminus \{xc_0\}) \cup \{xv\} \cup M_C$ is a matching in G of size greater than $|M|$.

Next, suppose that some edge $e \in M \setminus E(F)$ has at least one endpoint in $V(F)$. Set $e = xy$. Then e in fact has both its endpoints in $V(F)$, for otherwise, it would have been added to F via our breadth-first-search construction. Moreover, both endpoints of e must belong to odd levels. If both endpoints of e belong to the same tree T_u (for some $u \in L_0$), then similarly to the previous case, we obtain a flower containing e, and we then proceed as in the previous case. So, we may assume that e does not have both its endpoints in the same tree. Then there exist distinct $u, v \in L_0$ such that $x \in V(T_u)$ and $y \in V(T_v)$, and so $u - T_u - x - y - T_v - v$ is an M-augmenting path in G. We can now obtain a matching of size $|M| + 1$, as in Lemma 1.1.

From now on, we assume that there are no edges (of G) between vertices in even levels, and moreover, that every edge of M that has an endpoint in $V(F)$ is in fact an edge of F. We now claim that G contains no M-augmenting path, and that M is therefore (by Theorem 1.2) a maximum matching in G. Since L_0 is the set of all vertices that are unsaturated by M, it suffices to show that no non-trivial M-alternating path has more than one endpoint in L_0.[12] So, fix an M-alternating path $P = p_0, \dots, p_t$, with $t \geq 1$. We must show that at most one of p_0, p_t belongs to L_0. If neither p_0 nor p_t belongs to L_0, then we are done. So, by symmetry, we may assume that $p_0 \in L_0$, and we must show that $p_t \notin L_0$.

> **Claim.** For all $i \in \{0, \dots, t-1\}$, one of the following holds:
>
> (1) $p_i p_{i+1} \in E(F)$, and there exists an integer k such that $p_i \in L_k$ and $p_{i+1} \in L_{k+1}$;
>
> (2) $p_i p_{i+1} \notin E(F)$, p_i belongs to an even level, and p_{i+1} belongs to an odd level.[13]

Proof of the Claim. We proceed by induction on i. First of all, $p_0 \in L_0$. So, if $p_0 p_1 \in E(F)$, then $p_1 \in L_1$, and (1) holds for $i = 0$. So, we may assume that $p_0 p_1 \notin E(F)$. Since vetices of L_0 are unsaturated by M, we know that $p_0 p_1 \notin M$. Now $p_1 \in V(F)$, for otherwise, our breadth-first-search construction of F would have added $p_0 p_1$ to F. Since there are no edges between even levels, we see that p_1 belongs to an odd level, and so (2) holds for $i = 0$.

Now, fix $i \in \{0, \dots, t-2\}$, and assume that the claim holds for i. We must show it holds for $i + 1$.

Suppose first that (1) holds for i, i.e. that $p_i p_{i+1} \in E(F)$, and there exists an integer k such that $p_i \in L_k$ and $p_{i+1} \in L_{k+1}$. If $p_{i+1} p_{i+2} \in E(F)$, then $p_{i+2} \in L_{k+2}$, and (1) holds for $i+1$. So, assume that $p_{i+1} p_{i+2} \notin E(F)$. Then $p_{i+1} p_{i+2} \notin M$,[14] and so since P is M-alternating, we see that $p_i p_{i+1} \in M$. But then k is odd and $k + 1$ is even. Note that $p_{i+2} \in V(F)$, for otherwise, our breadth-first-search construction of F would have added $p_{i+1} p_{i+2}$ to F. Since there are no edges between even levels of F, and since p_{i+1} belongs to an even level, it follows that p_{i+2} belongs to an odd level. So, $i + 1$ satisfies (2).

[12] A path is *non-trivial* if it has at least one edge.

[13] Note that both (1) and (2) imply that $p_i, p_{i+1} \in V(F)$.

[14] This is because $p_{i+1} \in V(F)$, but $p_{i+1} p_{i+2} \notin E(F)$.

Suppose now that (2) holds for i, i.e. that $p_ip_{i+1} \notin E(F)$, p_i belongs to an even level, and p_{i+1} belongs to an odd level. Since p_ip_{i+1} has an endpoint in $V(F)$, but does not belong to $E(F)$, we see that $p_ip_{i+1} \notin M$. Therefore, $p_{i+1}p_{i+2} \in M$, since P is M-alternating. So, $p_{i+1}p_{i+2} \in E(F)$.[15] Since p_{i+1} belongs to an odd level, say L_k, we see that p_{i+2} belongs to the even level L_{k+1}. So, (1) holds for $i+1$. This proves the Claim. ■

In view of the Claim, p_0 is the only vertex of P that belongs to L_0.[16] So, $p_t \notin L_0$, and we are done.

Remark: The running time of Edmonds' Blossom algorithm is $O(n^4)$, if the algorithm is implemented in the obvious way. We omit the details.

[15]This is because $p_{i+1}p_{i+2} \in M$ and $p_{i+1} \in V(F)$.

[16]Indeed, fix $i \in \{1, \dots, t\}$. In view of the Claim, p_i either belongs to an odd level, or it belongs to a level that is one higher than the level that p_{i-1} belongs to. In either case, $p_i \notin L_0$.