19th Lecture: 8/3

Definitions 19.1. Let $G = (V, E)$ be a graph. A *matching* in G is a subset $M \subseteq E$ such that no two edges of M share a vertex. In other words, M is a subgraph of G in which every vertex has degree zero or one. A vertex of degree one is said to be *included* in the matching, or simply to be *matched*.

The *size* of a matching M is the number of edges in it, and is denoted $|M|$. If $|M| \geq |M'|$ for any other matching M', then M is said to be a *maximum matching*.

A matching is said to be *perfect* or *complete* if $|M| = |V|/2$, in other words, if every vertex in G is matched.

Observation 19.2. Note the simple but useful observation that a graph cannot possess a perfect matching if it has an odd number of vertices.

Remark 19.3. Do not confuse a *maximum* matching with a *maximal* matching. A matching M in a graph $G = (V, E)$ is said to be *maximal* if there doesn't exist any $e \in E \backslash M$ such that $M \cup \{e\}$ is still a matching.

Obviously, every maximum matching is also maximal. However, the reverse need not apply. For example, let G be a chain of length 3, i.e.: a chain with 4 vertices and 3 edges. If M consists of the middle edge, then it is maximal, because we can't add either of the other two edges to M and still have a matching. But it is not maximum, since M' consisting of the first and third edges has size 2, whereas $|M| = 1$. This simple example is generalised in Definition 19.4 below.

The *Matching Problem* asks for a procedure to determine a maximum matching in a graph and thus, in particular, to decide if the graph has a perfect matching. In studying this problem, the key concept turns out to be the following:

Definition 19.4. Let M be a matching in a graph. A path $v_1 \to v_2 \to \cdots \to v_k$ in G is called an *M-augmenting path* if

(i) k is even, i.e.: the length of the path is odd

(ii) every second edge in the path lies in M and every second edge lies outside M

(iii) the first edge lies outside M, i.e.: $\{v_1, v_2\} \notin M$.

(iv) v_1 and v_k are unmatched vertices in M.

One immediately observes the following: Let M be a matching and $\mathcal{P}$ an M-augmenting path, Let $M' \subseteq E$ be the set of edges obtained from M by replacing the edges along $\mathcal{P}$ which lie in M by those which don't. Then M' is also a matching and $|M'| = |M| + 1$.

Hence, if there exists an M-augmenting path, M is not a maximum matching. Crucially, the converse also holds:

Proposition 19.5. *Let $G = (V, E)$ be any graph and M a matching in G. If M is not a maximum matching, then there exists an M-augmenting path.*

PROOF: Let M^* be a maximum matching and let $H = M \Delta M^*$, i.e.: H is the subgraph of G consisting of those edges which occur in precisely one of M and M^*, but not both. Since M and M^* are both matchings, the degree of any vertex in H is at most 2. Hence, H has no branching points and each of its connected components must be either a cycle or a chain. Moreover, in any such component, edges must alternate between M and M^*. This implies that any cycle must have even length and include an equal number of edges from M and M^*. But $|M^*| > |M|$, so there must be at least one component of H which is a chain of odd length, in which edges alternate between M and M^* and the first edge is in the latter. Moreover, the first and last vertices along the chain are unmatched in M, as otherwise this chain could be extended to a larger connected subgraph of H. Hence the chain is an M-augmenting path, v.s.v.

The proposition immediately leads to a general procedure for finding a maximum matching in an arbitrary graph. See Demo6 for a worked example.

Augmenting path algorithm. Start with the empty matching $M = \phi$. Perform a breadth-first search for an M-augmenting path. If no such path is found, conclude that M is a maximum matching and stop. If such a path is found, replace M by the augmented matching M', got by exchanging edges along the M-augmenting path. Repeat.

The matching problem is most natural in the setting of bipartite graphs $G = (X, Y, E)$. Among the myriad interpretations of the problem in this setting, here are three of the most common ones:

A. X is a set of men and Y a set of women. An edge represents a pair such that each regards the other as an "acceptable" spouse. Hence, a maximum matching represents an optimal solution to the *Marriage Problem* of marrying off as many couples as possible.

B. X is a set of job seekers and Y a set of vacancies. An edge represents a job for which the corresponding person is qualified. Hence, a maximum matching represents an optimal solution to the *Job Assignment Problem* of getting jobs for as many people as possible.

C. X is a set of high-school leavers and Y a set of university degree programs. An edge represents a program which the corresponding student is both qualified for and interested in. Hence, a maximum matching represents a way of getting the maximum number of students into university.

Note that in all of these settings one can reasonably argue that a "matchmaker" doesn't just want to match as many pairs as possible, but also needs to take account of the fact that people have preferences. This viewpoint will be explored in Lecture 21.

Terminology 19.6. In a bipartite graph $G = (X, Y, E)$, the size of any matching cannot exceed $\min\{|X|, |Y|\}$. In particular, there cannot exist a perfect matching if $|X| \neq |Y|$. If $|X| \leq |Y|$ (resp. if $|Y| \leq |X|$) and there exists a matching of size $|X|$ (resp. of size $|Y|$), then this matching is said to be *perfect for X, or X-perfect (resp.*

perfect for Y, or Y-perfect).

Definition 19.7. Let $G = (V, E)$ be any graph and $v \in V$. The *neighborhood* of v, denoted $N(v)$, is the set of all vertices joined to v by an edge, i.e.:

$$N(v) = \{w \in V : \{v, w\} \in E\}.$$

More generally, if $A \subseteq V$, the neighborhood of A, denoted $N(A)$, is the union of the neighborhoods of its elements, i.e.: $N(A) = \cup_{v \in A} N(v)$.

The classical result on matchings is the following:

Theorem 19.8. (Hall's Marriage Theorem) *Let $G = (X, Y, E)$ be a bipartite graph. Then there exists a perfect matching for X if and only if*

$$|N(A)| \geq |A| \quad \forall\, A \subseteq X. \tag{19.1}$$

Eq. (19.1) is called *Hall's condition*. Note how the following proof makes crucial use of the idea of an M-augmenting path.

Proof of Hall's theorem: It is obvious that Hall's condition is necessary since, if $A \subseteq X$ and $x \in A$, then x can a priori only be matched with a vertex in $N(A)$. In an X-perfect matching, every vertex of A must be matched, so there must be at least as many vertices in $N(A)$ as there are in A.

So suppose Hall's condition is satisfied. Let M be any matching such that $|M| < |X|$. It suffices to prove the existence of an M-augmenting path. Since M is not X-perfect, there is at least one unmatched node in X. Pick one, call it x_0. Set $A := \{x_0\}$. Hall's condition says $|N(A)| \geq |A| = 1$, so x_0 has at least one neighbor. Pick a neighbor, call it y_0. If y_0 is unmatched, then $x_0 y_0$ is an M-augmenting path.

So we may suppose y_0 is matched, say to x_1. Set $A := \{x_0, x_1\}$. Hall's condition says $|N(A)| \geq |A| = 2$, so there is at least one more vertex, other than y_0, which is a neighbor of either x_0 or x_1. Pick such a vertex, call it y_1. If y_1 is unmatched and a neighbor of x_0, then $x_0 y_1$ is an M-augmenting path. If y_1 is unmatched and a neighbor of x_1, then $x_0 y_0 x_1 y_1$ is an M-augmenting path.

So we may assume that y_1 is matched, say to x_2. Keep iterating the above procedure to produce a sequence of distinct vertices $x_0, y_0, x_1, y_1, x_2, y_2, \ldots, x_k, y_k$, until you hit an unmatched vertex y_k. Note that this must eventually happen since applying Hall's condition with $A = X$ implies that $|Y| \geq |X|$ and hence, if there is an unmatched vertex in X there must also be one in Y. Once we hit an unmatched y_k, there will be a shortest path from y_k back to x_0 which only passes through vertices among those in the above sequence, and such that every second edge along this path is of the form $\{x_k, y_{k-1}\}$ and included in M. Hence every other edge is not in M, since M is a matching. The path starts in Y and ends in X, so it has odd length. The first and last vertices, y_k and x_0, are unmatched. Hence this is an M-augmenting path, v.s.v.