1 Matchings

A popular question to be asked on graphs, if graphs represent some sort of compatability or association, is how to associate as many vertices as possible into well-matched pairs. It is to this end that we introduce the concept of a *matching*:

Definition 1. A *matching* on a graph G is a subset M of the edge-set $E(G)$ such that every vertex of G is incident on at most one edge of M.

A matching can be thought of as representing a pairing of vertices in which only adjacent vertices are paired, and each vertex is paired with only one of its neighbors. A matching might not include every vertex of the graph, but in investigating the question of how many vertices we can match with good partners, doing so is obviously a desirable effect. We thus introduce two significant concepts with regard to a match's quality:

Definition 2. A matching M is *maximal* in a graph G if there is no matching M' with $|M'| > |M|$.

Definition 3. A mathing M is a *perfect matching*, also called a *1-factor*, if the edges of M are incident to every vertex of G.

The term "1-factor" is a bit mysterious, so, although we will not investigate it further at this time, we present the general concept of which a matching is a special case.

Definition 4. A subgraph H of G is called a k-factor if $V(H) = V(G)$ and H is k-regular.

Since 1-regular graphs are graphs consisting of unions of non-incident edges, a matching corresponds exactly to a 1-regular subgraph, and a perfect matching to a 1-factor.

We may make a few very simple related observations:

- The edges of a matching M are incident on $2|M|$ vertices.
- A matching is perfect in G if and only if $|M| = \frac{1}{2}|G|$.
- If a graph has perfect matchings, the set of perfect matchings is identical to the set of maximal matchings.
- G has no perfect matchings if $|G|$ is odd.

1.1 Matchings in bipartite graphs

The question of finding maximal matchings on arbitrary graphs is sufficiently complicated that at first we should constrain ourselves to a simple but useful special case: when G is bipartite. Here the graph represents a common real-world problem: we have sets A and B of unlike objects, a particular set of pairs elements of A and B which we are allowed to associate with eachother, and a goal of pairing off as many of these elements as possible. Traditionally this problem has often been labeled the "marriage problem" after the traditional context of arranging men and women, not all of whom are necessarily compatible with each other, into appropriate marriages (note: in several U.S. states marriage arrangement is no longer guaranteed to be bipartite, but the appellation stands). Other obvious contexts in which bipartite matching might be those in which one set represents jobs, the other prospective workers, and the edges qualifications to do the job, or those in which one set represents classes at a certain time, the other classrooms, and the edges sufficient capacity to contain the class.

So, in considering matchings in a bipartite graph G, it will be clear that each edge of a matching M will be incident to one vertex in A and one vertex in B, and thus a matching can only be perfect if $|A| = |B|$. A somewhat weaker condition than absolute perfection is to guarantee that at least one set is matched.

Definition 5. A matching M in a bipartite graph with parts A and B is called *A-perfect* if M is incident on every vertex of A.

We can note easily, since each edge in M is incident on a distinct vertex if A, that M is A-perfect if and only if $|M| = |A|$. Thence we can easily determine that an A-perfect matching is maximal, and that a matching is perfect if and only if it is A-perfect and $|A| = |B|$. Thus, investigating A-perfection is sufficient to discover when a matching is perfect.

Clearly there are graphs in which maximal matchings are *not* A-perfect. Suppose $v \in A$ has no neighbors (i.e. degree zero). Then, regardless of what edges are in a matching M, none of them are incident on v, so M cannot be A-perfect. Investigating further, we find a similar if slightly more difficult to illuminate situation if there are v_1 and v_2 in A both of whose only neighbor is some u in B: No matching can include both the edges $\{v_1, u\}$ and $\{v_2, u\}$, since those edges are both incident on the same vertex, so every matching M either lacks an edge incident on v_1 or an edge incident on v_2. Generalizing this argument:

Definition 6. The *neighborhood* $N(S)$ of a vertex-set S is the set of all vertices adjacent to some vertex of S. $N(\{v\})$ may be written as $N(v)$.

Proposition 1. *For a bipartite graph $G = (A, B)$, if G has an A-perfect matching, then for every $S \subseteq A$, $|N(S)| \geq |S|$.*

Proof. For an arbitrary S, let us label its vertices $S = \{a_1, a_2, a_3, \ldots, a_n\}$. Let M be an A-perfect matching in G. Then, for each i, M contains $\{a_i, b_i\}$ for some $b_i \in B$. Since each edge must have distinct endpoints, all the b_i are distinct, and since each b_i is adjacent to $a_i \in A$, it follows by definition that $b_i \in N(S)$. Thus $\{b_1, b_2, \ldots, b_n\}$ is an n-element subset of $N(S)$, so $|N(S)| \geq |S|$. □

We have determined a fairly strong necessary condition for a graph to have an A-perfect matching; it may not come as a surprise, then, that this condition is in fact sufficient as well. However, we will need several other tools to prove it. We will start by investigating a relationship between matchings and a somewhat dual concept of a *vertex cover*.

Definition 7. A subset S of the vertex-set of G is called a *vertex cover* if every edge of G is incident on a vertex of S.

Proposition 2. *If M is a matching in G and S is a vertex cover of G, then $|M| \leq |S|$.*

Proof. Denote $M = \{\{u_1, v_1\}, \{u_2, v_2\}, \{u_3, v_3\}, \ldots, \{u_n, v_n\}\}$. By the definition of a matching all of the vertices $u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_n$ are distinct. Since S is a vertex-cover of G, each edge $\{u_i, v_i\}$ must have an endpoint in S, so S contains at least one of u_i, v_i for each i. Since i ranges from 1 to n, this guarantees n distinct elements of S, so $|S| \geq |M|$. □

Since every vertex cover is at least as large as every matching, finding a vertex cover and a matching of equal size would be a strong and useful fact: the matching so found must be maximal, and the vertex-cover minimal. It is always, in fact, possible to do so in a bipartite graph:

Theorem 1 (König-Egerváry, '31)**.** *A bipartite graph G contains a matching M and set-cover S such that $|M| = |S|$.*

Proof. Let $M = \{\{a_1, b_1\}, \dots, \{a_n, b_n\}\}$ be a maximal matching in G. We shall explicitly construct a set-cover S guaranteed to be equal in size to M by selecting one endpoint of each edge.

A path $P = v_1 \sim v_2 \sim v_3 \sim \cdots \sim v_k$ in G is denoted as *alternating* with respect to M if $v_1 \in A$, v_i is not incident on any edge in M, $\{v_i, v_{i+1}\} \notin M$ for all odd i, and $\{v_i, v_{i+1}\} \in M$ for all even i; that is to say, an alternating path is one beginning at a vertex of A not matched by M, and following alternatingly edges not in M and in M.

For every edge $\{a_i, b_i\}$ in M, select b_i as an element of S if there is an alternating path P_i ending in b_i, and select a_i otherwise. By construction, $|S| = |M|$; we shall show furthermore that every edge in G is incident on some vertex of S, so S is a vertex cover.

It is easy to show that every edge in M is incident on a vertex of S, since S was constructed by selecting an endpoint of every edge in M. Let us consider an edge $\{a, b\} \notin M$. Since M is maximal, $M \cup \{\{a, b\}\}$ can not be a matching, so $\{a, b\}$ must share an endpoint with some edge $\{a_i, b_i\}$; i.e., either $a_i = a$ or $b_i = b$ for some i.

If $a_i \neq a$ for all i and there is a $b_j = b$, then a is not incident on any edge in M, so $\{a, b_j\}$ is an alternating path in G with respect to M, so by construction, $b_j \in S$, so $\{a, b\}$ is incident on a vertex of S (specifically, incident on b_j).

On the other hand, if $a_i = a$ for some i, there are two possibilities: either $a_i \in S$, in which case $\{a, b\}$ is clearly covered by S, or $b_i \in S$, which tells us that b_i is the endpoint of some alternating path P_i given by $v_1 \sim v_2 \sim \cdots \sim v_{k-1} \sim v_k = b_i$. Now we shall show that b is also the endpoint of some alternating path: if $b = v_j$ for some j, we could truncate the above path to $v_1 \sim v_2 \sim \cdots v_j = b_i$. If, on the other hand, b_i does not appear on the path, then $v_1 \sim \cdots v_k = b_i \sim a_i = a \sim b$ is an alternating path. Thus, either b is incident on some edge in M, in which case $b \in S$ by the construction of S, yielding a vertex in S incident to $\{a, b\}$, or, b is not incident on any edge in S. We shall show that this last situation contradicts the maximality of S.

If b is not incident on any edge of S, we may select different edges appearing in the above alternating path from v_1 to b while preserving M as a matching. Recall that by the definition of alternation, v_1 is some vertex in A not incident on any element of M, every edge $\{v_i, v_{i+1}\}$ for even i is in M, and every edge $\{v_i, v_{i+1}\}$ for odd i is not in M; furthermore, since $v_j = b \in B$, it is necessary that j is even. We shall construct a new matching by replacing which edges of the alternating path are in M:

$$M' = M - \{\{v_2, v_3\}, \{v_4, v_5\}, \dots, \{v_{j-2}, v_{j-1}\} \cup \{\{v_1, v_2\}, \{v_3, v_4\}, \dots, \{v_{j-1}, v_j\}$$

Note that M' is a matching, since it has the same incidence upon every vertex of G as M did *except* for the vertices v_1 and v_j, which it has exactly one edge incident upon instead of zero. Furthermore, $|M'| = |M| + 1$, so $|M|$ is not maximal.

We have thus shown that, for every edge in G, some vertex of our constructed set S is incident to that edge, so S is a vertex cover. □

This is a pretty messy and long proof, but it is actually the standard proof, since it contains within it the seeds of a matching algorithm based on selecting an arbitrary matching, selecting an associated set S and seeing which of the above failures-of-maximality occur. There is, however, for fans of brevity, and for those who like clever conenctions among different branches of the same discipling, a very clever proof.

The seed of this idea is that we have in fact proved a structurally similar result before! We have two structures, somewhat dual to eachother, one of which is consistently at least as large as the other. Our unusual result is that these structures are sometimes equivalent in size. If we replace "bipartite graph" with "poset", "vertex cover" with "chain partition", and "matching" with "antichain", we have Dilworth's Theorem!

Alternative proof. For bipartite $G = (A, B)$ with $|G| = n$, let us define the poset $(V(G), \preceq)$ by the relation $a \preceq b$ iff $a \in A$, $B \in B$, and $a \sim b$. Visually speaking, this is essentially rotating our graph 90 degrees from a conventional bipartite representation and calling the result a Hasse diagram.

By Dilworth's Theorem, this poset has an antichain A and partition into chains $\mathcal{C}$ such that $|A| = |\mathcal{C}| = k$ for some k. This poset obviously has height of no more than 2, since every element of A is minimal and every element of B is maximal. Thus, the elements of $\mathcal{C}$ are of two types: single vertices in G (chains of size 1) and edges in G (chains of size 2). Let us denote $\mathcal{C} = \{C_1, \ldots, C_\ell, C_{\ell+1}, \cdots, C_k\}$, where $C_1, \ldots, C_\ell$ have size 2, and $C_{\ell+1}, \ldots, C_k$ have size 1. Since these chains partition the poset, we know that $\sum_{i=1}^{k} |C_i| = |G|$, which can be simplified to $2\ell + (k-\ell) = n$ or $\ell = n - k$. Since $\mathcal{C}$ is a partition of $V(G)$, the chains are all vertex-disjoint, so in particular $\{C_1, \ldots, C_\ell\}$ is a vertex-disjoint collection of edges; that is to say, a matching of size $n - k$.

Now we shall derive from A a vertex cover of size $n - k$ to prove the König-Egerváry Theorem. Since A is an antichain, no two elements of A are comparable, which means by the construction of our poset that in G, there are no edges between elements of A (i.e., in graph parlance, A is an independent set). Thus, every single edge in G is incident on at most one member of A, and thus is incident on at least one member of $V(G) - A$. Thus $V(G) - A$ is a vertex cover of G, and $|V(G) - A| = n - k$. □

Incidentally, it is also possible to prove Dilworth's Theorem *from* König-Egerváry: for a poset $(S, \preceq)$, associate each $s \in S$ with two vertices a_s and b_s of a bipartite graph G, and let $a_s \sim b_t$ iff $s \preceq t$. Then we reverse the process in the proof above: take the complement of a vertex cover to get an antichain, and kludge edges from a matching together to build a system of disjoint chains.

This theorem tells us a lot about maximal matchings, but not when those matchings are perfect. However, it is a simple derivation from König-Egerváry to a perfection criterion:

Theorem 2 (Hall '35). *For a bipartite graph $G = (A, B)$, G has an A-perfect matching, if and only if for every $S \subseteq A$, $|N(S)| \geq |S|$.*

Proof. The necessity of this condition was established in Proposition 1; here we shall show that the given condition is also sufficient.

Let G be such that for all $S \subseteq A$, $|N(S)| \geq |S|$. Clearly A is a vertex cover of G; we shall show that it is the smallest such. Suppose $|S| < |A|$ and S is a vertex cover of A. Let $S_A = S \cap A$ and $S_B = S \cap B$. Then, $|S_A| + |S_B| = |S| < |A|$. Let $T = A - S_A$; note that $|T| = |A| - |S_A| = |A| - (|S| - |S_B|) > |S_B|$. By the condition given in the theorem, $|N(T)| > |T|$, so $|N(T)| > |S_B|$. Thus, there is necessarily at least one vertex in T adjacent to a vertex not in S_B. The edge between these vertices is clearly not covered: the vertex on the A end of the edge is in T, and by definition thus not in S_A, and on the B end of the edge the vertex is shown not to be in S_B. Thus, there is at least one edge in G not covered by this set S.

Thus, there is no vertex cover of size less than $|A|$, and there is clearly a vertex cover of size $|A|$, so $|A|$ is the minimum size of a vertex cover in G. By König-Egerváry, it follows that the maximum size of a matching in G is $|A|$; such a matching is necessarily A-perfect. □

This proof of Hall's Theorem is not actually traditional. A more illuminating if less elegant proof results from algorithmically developing a maximal matching.

Constructive proof. Suppose M is a matching which is *not* A-perfect. We will show that, from the Hall criterion, it is possible to build a larger matching.

Let a_0 be a vertex in A unmatched by M; the Hall Criterion guarantees that a_0 has at least one neighbor, which we shall call b_1. If b_1 is unmatched, then the edge $\{a_0, b_1\}$ is a clear improvement on M. Otherwise, let a_1 be the vertex matched with b_1 by M. Now, since $|N(\{a_0, a_1\})| \geq 2$, there must be some b_2 which is a neighbor of either a_0 or a_1. If b_2 is matched, we can include a_3 and therefrom find a b_4, and so forth. We can thus build a sequence of distinct vertices $a_0, b_1, a_1, b_2, a_2, \ldots, b_{k-1}, a_{k-1}, b_k$ such that:

- a_0 is unmatched.
- Each $b_i \in N(\{a_0, a_1, \ldots, a_{i-1}\})$.
- Each $\{a_i, b_i\} \in M$.
- b_k is unmatched.

The last of these conditions arises from the only situation in which a b_i cannot give us some a_i to continue the sequence.

What we shall now do is construct an alternating path from b_k to a_0 using these properties, and since b_k and a_0 are unmatched, toggling membership in M over this path will make a larger matching.

Starting from b_k, we know, since $b_k \in N(\{a_0, \ldots, a_{k-1}\})$, that some a_{i_1} is b_k's neighbor for $i_1 < k$. Then by the construction $\{a_{i_1}, b_{i_1}\} \in M$, and we repeat the above for b_{i_1}: $b_{i_1} \in N(\{a_0, \ldots, a_{i_1 - 1}\})$, so b_{i_1} is adjacent to some a_{i_2} with $i_2 < i_1$. We continue onwards until it is impossible to progress further, which will happen only when either some b_i has no neighbor with smaller index (which is impossible by our construction) or there is some a_i with no matched partner, which only occurs at a_0. Thus there is an alternating path

$$b_k \sim a_{i_1} \sim b_{i_1} \sim a_{i_2} \sim b_{i_2} \sim \cdots \sim a_{i_\ell} \sim b_{i_\ell} \sim a_0$$

And we can thus improve the matching M by replacing the ℓ edges $\{a_{i_j}, b_{i_j}\}$ with the $\ell + 1$ edges $\{b_k, a_{i_1}\}, \{b_{i_1}, a_{i_2}\}, \{b_{i_2}, a_{i_3}\}, \cdots, \{b_{i_\ell}, a_0\}$.

Since every matching with an unmatched vertex in A can be made larger, it follows that the largest matching has no unmatched vertices in A, and is thus A-perfect. □

This proof is in some ways superior to the König-Egerváry argument, because it contains the seeds of an algorthm for *building* an A-perfect (and thus actually perfect, where possible) matching. Our algorithm is as follows:

1. Pick a matching M (you can start with the empty matching, or a greedily-constructed matching, both of which are pretty easy.
2. If every vertex of A is matched by M, we have an A-perfect matching and are done. Otherwise, let a_0 be the unmatched vertex, and let i be 1.
3. If there is a vertex adjacent to $\{a_0, \ldots, a_{i-1}\}$ not already selected as some b_k, call it b_i. Otherwise, our graph fails the Hall criterion and has no A-perfect matching.
4. If b_i is in our matching, let a_i be the vertex matched with it, and increment i and return to step 3. Otherwise, continue to step 5.
5. Let $u_1 = b_i$, and let $j = 1$.
6. Let v_j be the element of $\{a_0, \ldots, a_{i-1}\}$ of lowest index which is adjacent to u_j.
7. If $v_j = a_0$, then continue to the next step. Otherwise, let u_{j+1} be the vertex matched with v_j in the matching M, increment j, and return to step 6.
8. Remove all edges $\{v_1, u_2\}, \{v_2, u_3\}, \ldots, \{v_{j-1}, u_j\}$ from M, and put the edges $\{u_1, v_1\}, \{u_2, v_2\}, \ldots, \{u_j, v_j\}$ in M.
9. Forget what all these variables mean, and return to step 2.

OK, this gets us to a good place as regards matchings in bipartite graphs: we have an algorithm for constructing perfect matchings when they exist (and a moderate tweak of this algorithm can even find maximal matchings in graphs that don't meet the Hall criterion). Now we move on to a harder problem.

1.2 Matchings in non-bipartite graphs

This is a pairing problem for not-necessarily-unlike items: one traditional name for it is "the roommate problem", since unlike for job assignments or heterosexual marriage, roommate suitability is compatability among a single group with no intrinsic bipartition.

Up front, we can work out a few simple facts:

- If G has an odd number of vertices, it has no perfect matching.
- If G has any isolated vertices, it has no perfect matching.
- G has a perfect matching if and only if each of its components has a perfect matching.

The first and second of these observations can actually be united by way of the third to the observation that if G has any component with an odd number of vertices, then G has no perfect matching.

In fact, using the same arguments we used to show necessity of the Hall condition on bipartite graphs, we can find a fairly strong necessary condition for a graph to have a perfect matching. Note that the below condition includes our above observation when $S = \emptyset$.

Proposition 3. *If there is a set S such that $G - S$ has more than $|S|$ components with an odd number of vertices, then G has no perfect matching.*

Proof. Let M be a matching on G and let $S = \{v_1, v_2, \ldots, v_k\}$. Now consider the restriction M' of M to $G - S$, so any edges incident on the v_i no longer appear in M'. Given that $G - S$ has odd components $C_1, C_2, \ldots, C_{k+1}$ (and possibly more), we know by our above observation that there is at least vertex u_i in each C_i which is unmatched by M'. Now, we know that M consists only of the edges in M' together with at most one edge incident on each v_i, so M contains at most k edges not in M'; however, there are at least $2k + 1$ vertices of G which were unmatched by M' (all the u_i and v_i), so M does not match every vertex in G. □

In fact, much as Hall's Marriage Theorem showed that the analogous condition was sufficient in bipartite graphs, so will we see that this condition is sufficient for non-bipartite graphs.