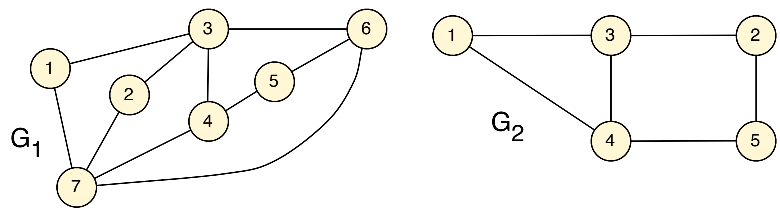
Bipartite Graphs and Matchings

(Revised Thu May 22 10:59:19 PDT 2014)

A graph $G = (V, E)$ is called *bipartite* if its vertex set V can be partitioned into two disjoint subsets L and R such that all edges are between L and R. For example, the graph G_1 below on the left



is bipartite, because we can partition its vertex set into $L = \{1, 2, 4, 6\}$ and $R = \{3, 5, 7\}$, and then each edge will have one endpoint in L and the other endpoint in R. On the other hand, the graph G_2 on the right is not bipartite. Why? This can be justified by considering vertices 1, 3 and 4. No matter how we partition the vertices of this graph into two sets L and R, two of the three vertices 1, 3 and 4 will have to belong to the same set. But these vertices are connected by an edge, so we would have an edge inside L or inside R, which contradicts the definition of bipartite graphs.

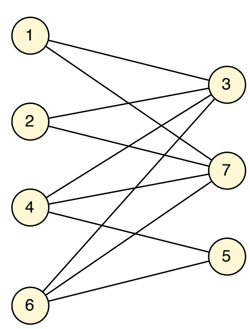
The observation in the above paragraph can be generalized, giving us the following characterization of bipartite graphs.

Theorem 1 *A graph G is bipartite if and only if it does not contain any cycle of odd length.*

Proof: ($\Rightarrow$) Easy: each cycle alternates between left-to-right edges and right-to-left edges, so it must have an even length.

($\Leftarrow$) Pick any vertex v_0. For each other vertex v, let d_v be the length of the shortest path from v_0 to v. Let L be the set of vertices with d_v even and R the set of vertices with d_v odd. Then show that there is no edge inside L or inside R ... (otherwise an odd-length cycle would exist). $\square$

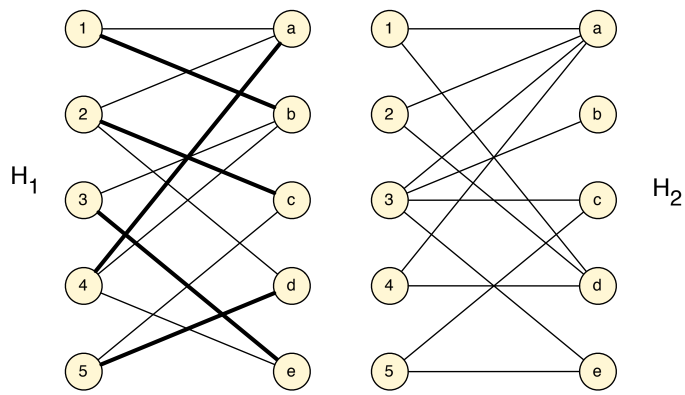
For bipartite graphs it is convenient to use a slightly different graph notation. If $G = (V, E)$ is bipartite and $V = L \cup R$ is the partition of the vertex set such that all edges are between L and R then we will write $G = (L, R, E)$. We will also typically draw these bipartite graphs with L on the left-hand side, R on the right-hand side, and edges going across. For example, graph G_1 above can be redrawn as follows:



Perfect matchings. We will now restrict our attention to bipartite graphs $G = (L, R, E)$ where $|L| = |R|$, that is the number of vertices in both partitions is the same. A *perfect matching* in such a graph is a set M of edges such that no two edges in M share an endpoint and every vertex has an edge that belongs to M. The two conditions imply that M must have the same cardinality as L and R.

For some intuition, think of L as a set of boys and R a set of girls, with each edge (x, y) representing a pair x, y that like each other. If G has a perfect matching, this means that we can marry all boys and girls happily, with everyone getting a partner they actually like.

For example, the graph H_1 below on the left has a perfect matching, shown using thick lines:



What about the graph H_2 on the right-hand side? As it turns out, H_2 does not have a perfect matching. The reason is this: consider vertices (boys) 1, 2 and 4. Each of them likes girls a and d and no other girls. So we have three boys that only like two girls, and therefore it is not possible to make all three of them happy (even in Utah).

For a set of vertices $X \subseteq L$, denote by $N(X)$ the set of neighbors of X. In H_2 above, the reason the graph did not have a perfect matching was that the graph had a "bottleneck" set $X \subseteq L$, that is X satisfying $|N(X)| < |X|$. Because of this inequality it is impossible to match all vertices in X to their neighbors in a 1-1 fashion. Thus the existence of such a bottleneck set implies that the graph cannot have a perfect matching. A natural question arises here: if G does not have such a bottleneck set, does this guarantee that G must have a perfect matching? It turns out that yes, as we show below, although the proof of this is quite subtle.

Theorem 2 *(Hall's Theorem.) A bipartite graph $G = (L, R, E)$ with $|L| = |R|$ has a perfect matching if and only if each set $X \subseteq L$ satisfies $|N(X)| \geq |X|$.*

Below, we will refer to the condition "$|N(X)| \geq |X|$ for each $X \subseteq L$" in this theorem as the *no-bottleneck* condition.

Proof: (Sketch) ($\Rightarrow$) This implication is easy, and we have essentially proved it already. Suppose that G has a perfect matching M and pick any set $X \subseteq L$. Denote by $M(X)$ the set of vertices matched to those in X via the edges in M. Recall that we match only neighbors, so $|N(X)| \geq |M(X)| = |X|$, and $|N(X) \geq |X|$ follows.

($\Leftarrow$) This implication is harder. We need to prove the following statement:

($*$) Suppose that $G = (L, R, E)$ is a bipartite graph with $|L| = |R|$ such that each set $X \subseteq L$ satisfies $|N(X)| \geq |X|$. Then G has a perfect matching.

We will prove ($*$) by induction on $n = |L| = |R|$. We start with the base case, when $n = 1$. Then L and R

have each just one vertex and the no-bottleneck condition states that $N(L) \geq |L| = 1$, so these vertices must be connected by an edge. Thus this graph has a perfect matching, consisting of just this one edge.

In the inductive step, fix now some $n > 1$ and suppose that $(*)$ holds for all bipartite graphs with fewer than n vertices in each partition. Consider any graph $G = (L, R, E)$ with $|L| = |R| = n$ that satisfies the no-bottleneck condition, that is $|N(X)| \geq |X|$ for each $X \subseteq L$. Using the inductive assumption, we want to show that there is a perfect matching in G. We have two cases to consider.

Case 1: $|N(X)| \geq |X| + 1$ for all $X \subset L$. Intuitively, this represents the "slack" case, with each set of vertices (except for L itself) having more candidates to match than necessary. In this case we proceed as follows. Chose an arbitrary edge $(x, y) \in E$. Let $G' = (L', R', E')$ be obtained from G by removing x and y. Then, for any $X \subseteq L'$, using appropriate subscripts to distinguish between neighborhoods in G and G', we have

$$|N_{G'}(X)| \geq |N_G(X)| - 1 \geq (\,|X| + 1\,) - 1 = |X|,$$

where the first inequality follows from the fact that $N_G(X)$ can only lose one element when we remove x from G, and the second inequality follows from the case assumption.

What the last inequality shows is that G' satisfies the no-bottleneck condition in the lemma, and since L' has fewer than n vertices, the inductive assumption implies that G' has a perfect matching. Let's call this matching M'. By adding (x, y) to M' we obtain a perfect matching in M, completing the proof for this case.

Case 2: $|N(X)| = |X|$ for some $X \subset L$. This case is a bit more subtle. Introduce the following notations:

$$\begin{aligned} Y &= N_G(X) \\ H &= (X, Y, F) \\ X' &= L - X \\ Y' &= R - Y \\ H' &= (X', Y', F') \end{aligned}$$

where F is the set of edges between X and Y and F' is the set of edges between X' and Y'. We now examine graphs H and H'. Note that both graphs have equal size partitions, namely $|X| = |Y|$ and $|X'| = |Y'|$. We will argue that each of them must have a perfect matching and then we will combine these matchings into one.

Let's start with H. The key property of H is that it inherits all edges of G that have an endpoint in X, because of the way Y is defined. For any $Z \subseteq X$, by the definition of H we have $N_H(Z) = N_G(Z)$, so $|N_H(Z)| = |N_G(Z)| \geq |Z|$. Thus H satisfies the no-bottleneck condition and it has fewer vertices than G, so from the inductive assumption we obtain that H has a perfect matching. Let's call this matching Q.

What about H'? Let $Z \subseteq X'$. It is harder to reason about $N_{H'}(Z)$, because this neighborhood could be different than $N_G(Z)$, if Z has some neighbors in Y. So we use the following trick. We consider the set $Z \cup X$ instead. The no-bottleneck condition for G implies that $|N_G(Z \cup X)| \geq |Z \cup X|$. Further, $N_G(Z \cup X) = Y \cup N_{H'}(Z)$, and these two sets are actually disjoint. Putting this together, we get

$$\begin{aligned} |Y| + |N_{H'}(Z)| &= |N_G(Z \cup X)| \\ &\geq |Z \cup X| \\ &= |Z| + |X|. \end{aligned}$$

Since we have $|Y| = |X|$, this gives us $|N_{H'}(Z)| \geq |Z|$. We thus just showed that the no-bottleneck condition $|N_{H'}(Z)| \geq |Z|$ holds for all $Z \subseteq X'$, which, using the inductive assumption implies that H' has a perfect matching, say Q'.

Joining these two matchings together, that is letting $M = Q \cup Q'$, we obtain that M is a perfect matching in G, proving the inductive claim. $\square$