Matchings in bipartite graphs, continued

A matching is a graph \( M \) all of whose vertices have degree 1. By a matching in a graph \( G \) we mean a subgraph \( M \) that is a matching. A matching \( M \) is a maximum matching in \( G \) when \( |E(M)| \) is the largest possible number of edges of matchings in \( G \). A perfect matching in \( G \) is a spanning matching in \( G \). A perfect matching is clearly a maximum matching.

A vertex-cover in a graph \( G \) is a set \( C \subseteq V(G) \) so that every edge of \( G \) has at least one end in \( C \). (It is the edges that are “covered”, but the vertices do the “covering”. ) A vertex-cover is minimum when it has the smallest number of vertices of any vertex cover.

If \( M \) is any matching in \( G \) and \( C \) is any vertex-cover in \( G \), then

\[
|E(M)| \leq |C|
\]

because \( C \) must contain at least one end of each of the edges of \( M \). Theorem 2 says that if \( G \) is bipartite, there exist a matching \( M_0 \) and a cover \( C_0 \) with \( |E(M_0)| = |C_0| \). Then we know \( M_0 \) is a maximum matching and \( C_0 \) is a minimum cover.

**Theorem 2.** (König’s theorem) In a bipartite graph \( G \), the size of a maximum matching is equal to the size of a minimum cover.

If your boss (someday) asks you to find a maximum matching in a bipartite graph \( G \) with 10000 vertices and you come back with a matching \( M \) with 4987 edges, your boss can easily check that you found a matching; but how does he/she know it is maximum? At the same time you turn in your matching \( M \), you should consider also turning in a vertex-cover \( C \) with 4987 vertices. It is relatively easy to check that \( C \) is a vertex-cover, and its existence PROVES no matching with more than 4987 vertices exists.

An alternating path with respect to a matching \( M \) is a path (not a closed path) whose edge terms are alternately not in \( E(M) \) and in \( E(M) \). It doesn’t matter in our definition whether the first edge of the path is in or not in \( E(M) \), but we will be primarily interested in alternating paths \( p \) with initial and terminal vertices not in \( V(M) \); these we call special alternating paths w.r.t. \( M \). (Biggs in Section 17.5 uses ‘alternating’ paths for what we call ‘special alternating paths’.) Clearly, a special alternating path is of odd length and has its first and last edge terms not in \( E(M) \).

It is important to understand that if we can find a special alternating path w.r.t. \( M \) in a graph \( G \), bipartite or not, then we can find a matching with more edges than \( M \). If \( p = (x_0, x_1, \ldots, x_r) \) is a special alternating path w.r.t. \( M \) in \( G \), then let \( M' \) be the subgraph whose vertices are \( V(M) \cup \{x_0, x_r\} \) and whose edges are

\[
(E(M) \setminus \{\{x_1, x_2\}, \{x_3, x_4\}, \ldots, \{x_{r-2}, x_{r-1}\}\}) \cup \{\{x_0, x_1\}, \{x_2, x_3\}, \ldots, \{x_{r-1}, x_r\}\}.
\]

It is easy to see that \( M' \) is a matching and that it has one more edge than \( M \) has.
Proposition 3 of Notes #9 claims that if $M$ is a matching in $G$, bipartite or not, and if $M$ is not maximum, then there exists a special alternating path w.r.t. $M$ in $G$. A proof is provided by Problem 4 of Set 4.

We can create an algorithm for finding a maximum matching in a bipartite graph out of these ideas by describing a systematic way to find special alternating paths w.r.t. a matching $M$, applying the edge-switching idea of two paragraphs back to get a matching $M'$ with more edges, and repeating w.r.t. $M'$ to find bigger matchings $M''$ etc., until we are sure no further apecial alternating paths exist, in which case we have a maximum matching.

A systematic way to search is given by a variation of the labeling procedure mentioned in Notes #8.

Let $M$ be a matching in a bipartite graph $G$. Let $X$ be the red vertices and $Y$ the blue vertices in a proper 2-coloring of $V(G)$. Label all vertices of $X \setminus V(M)$ with 0. Label all vertices in $Y$ that are adjacent to a vertex labeled 0 but with 1. After some vertices in $Y$ have been labeled with an odd number $j$, give vertices in $X$ the label $j+1$ if they are adjacent to a vertex labeled $j$ by an edge of $M$. After some vertices in $X$ have been labeled with an even number $\ell$, give vertices in $Y$ the label $\ell+1$ if they are not already labeled and are adjacent to a vertex $x$ in $X$ labeled $\ell$ (by an edge NOT IN $M$, but we don’t need to say this because the other end of the matching edge incident with $x$ is already labeled with $\ell - 1$). Continue until no new labels can be given.

This is more simple than it sounds when I try to write it down. See Biggs, page 220, for another description. Note that even labels are only given to (some) vertices in $X$ and odd labels to (some) vertices in $Y$.

**Proposition 3.** If a vertex $z$ is labeled $r$, then the length of a shortest alternating path from a vertex of $X \setminus V(M)$ to $z$ in $G$ is $r$.

**Proof:** The complete proof is similar to that of Proposition 4 in Notes #8 and is omitted. \(\square\)

Remark: In order to efficiently find an alternating path from a vertex labeled 0 to a vertex $z$, it is recommended that whenever a label $r \geq 1$ is given to a vertex $x$, one should also ‘record’ or ‘remember’ a particular vertex $x'$ labeled $r-1$ that is adjacent to $x$.

The question of whether there is a special augmenting path w.r.t. a matching $M$ is the same as asking whether any vertex $y$ of $Y \setminus V(M)$ gets a label, since an alternating path from a vertex labeled 0 to that vertex $y$ is special.

**Proof of König’s Theorem:** Let $M_0$ be a maximum matching in $G$. With the notation as above, let $A$ be the set of vertices in $x \in X$ for which there are NO alternating paths w.r.t. $M_0$ from a vertex labeled 0 to $x$, and let $B$ be the set of vertices in $y \in Y$ for which there ARE alternating paths w.r.t. $M_0$ from a vertex labeled 0 to $y$. We claim that $C_0 = A \cup B$ is a vertex-cover in $G$ and that $|C_0| = |E(M_0)|$.

It is possible to continue the proof without reference to the algorithm above, but perhaps it will be good to use the labeling terminology since we have introduced it. Of course, $A$ is the set of unlabeled vertices in $X$ and $B$ is the set of labeled vertices in $Y$. 
To see that $C_0$ is a cover, we must check that there are no edges $\{x, y\} \in E(G)$ with $x \in X \setminus A$ and $y \in Y \setminus B$, that is, with $x \in X$ labeled but $y \in Y$ unlabeled. But this is immediate, because once $x$ is labeled, the algorithm would label $y$, if it is not already labeled.

Trivial, $A \subseteq V(M_0)$, and because $M_0$ is maximum, there are no special alternating paths w.r.t. $M_0$, and so $B \subseteq V(M_0)$. In summary, $C_0 \subseteq V(M_0)$.

Let $\{x, y\}$ be an edge of $M_0$. If $y \in B$ is labeled $\ell$, the vertex $x$ receives a label $\ell + 1$ if it is not already labeled. In any case, $x \notin A$. If $y \notin B$, then $x$ cannot be labeled because it is not labeled 0 and the only other way it can get labeled is for $y$ to have a label; so $x \in A$. In summary, each edge of $M_0$ contains exactly one vertex of $C_0$; hence $|C_0| = |E(M_0)|$. □