Notes for Recitation 8

1 Bipartite Graphs

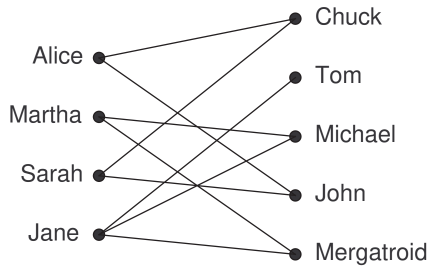
Graphs that are 2-colorable are important enough to merit a special name; they are called ***bipartite graphs***. Suppose that G is bipartite. Then we can color every vertex in G either black or white so that adjacent vertices get different colors. Then we can put all the black vertices in a clump on the left and all the white vertices in a clump on the right. Since every edge joins differently-colored vertices, every edge must run between the two clumps.

Bipartite graphs are both useful and common. For example, every path, every tree, and every even-length cycle is bipartite.

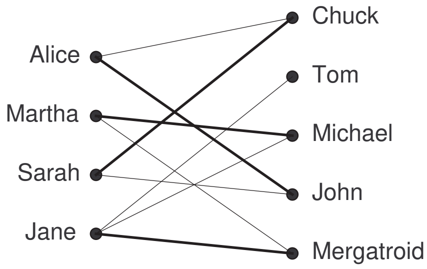
2 Hall's theorem

A class contains some girls and some boys. Each girl likes some boys and does not like others. Under what conditions can each girl be paired up with a boy that she likes?

We can model the situation with a bipartite graph. Create a vertex on the left for each girl and a vertex on the right for each boy. If a girl likes a boy, put an edge between them. For example, we might obtain the following graph:



In graph terms, our goal is to find a ***matching*** for the girls; that is, a subset of edges such that exactly one edge is incident to each girl and at most one edge is incident to each boy. For example, here is one possible matching for the girls:



Hall's Marriage Theorem states necessary and sufficient conditions for the existence of a matching in a bipartite graph. Hall's Theorem is a remarkably useful mathematical tool, a hammer that bashes many problems. Moreover, it is the tip of a conceptual iceberg, a special case of the "max-flow, min-cut theorem", which is in turn a byproduct of "linear programming duality", one of the central ideas of algorithmic theory.

We'll state and prove Hall's Theorem using girl-likes-boy terminology. Define *the set of boys liked by a given set of girls* to consist of all boys liked by at least one of those girls. For example, the set of boys liked by Martha and Jane consists of Tom, Michael, and Mergatroid.

For us to have any chance at all of matching up the girls, the following ***marriage condition*** must hold:

Every subset of girls likes at least as large a set of boys.

For example, we can not find a matching if some 4 girls like only 3 boys. Hall's Theorem says that this necessary condition is actually sufficient; if the marriage condition holds, then a matching exists.

Theorem 1. *A matching for a set of girls G with a set of boys B can be found if and only if the marriage condition holds.*

Proof. First, let's suppose that a matching exists and show that the marriage condition holds. Consider an arbitrary subset of girls. Each girl likes at least the boy she is matched with. Therefore, every subset of girls likes at least as large a set of boys. Thus, the marriage condition holds.

Next, let's suppose that the marriage condition holds and show that a matching exists. We use strong induction on $|G|$, the number of girls. If $|G| = 1$, then the marriage condition implies that the lone girl likes at least one boy, and so a matching exists. Now suppose that $|G| \geq 2$. There are two possibilities:

1. Every proper subset of girls likes a *strictly larger* set of boys. In this case, we have some latitude: we pair an arbitrary girl with a boy she likes and send them both away. The marriage condition still holds for the remaining boys and girls, so we can match the rest of the girls by induction.

2. Some proper subset of girls $X \subset G$ likes an *equal-size* set of boys $Y \subset B$. We match the girls in X with the boys in Y by induction and send them all away. We will show that the marriage condition holds for the remaining boys and girls, and so we can match the rest of the girls by induction as well.

   To that end, consider an arbitrary subset of the remaining girls $X' \subseteq G - X$, and let Y' be the set of remaining boys that they like. We must show that $|X'| \leq |Y'|$. Originally, the combined set of girls $X \cup X'$ liked the set of boys $Y \cup Y'$. So, by the marriage condition, we know:

$$|X \cup X'| \leq |Y \cup Y'|$$

   We sent away $|X|$ girls from the set on the left (leaving X') and sent away an equal number of boys from the set on the right (leaving Y'). Therefore, it must be that $|X'| \leq |Y'|$ as claimed.

In both cases, there is a matching for the girls. The theorem follows by induction. $\square$

There is an efficient algorithm for finding a matching in a bipartite graph, if one exists. Thus, if a problem can be reduced to finding a matching, the problem is essentially solved from a computational perspective.

The Chemist and The Maltster

At Guiness brewery in the early 1900's, W. S. Gosset (a chemist) and E. S. Beaven (a "maltster") were working to improve barley. Gosset and Beavan planned to grow several varieties of barley in a field and compare the yields. However, local variations in the fertility of the field might skew the results. Their solution was to divide the field into many small plots and grow each crop in several different places.

Similar thinking led to the use of *Latin squares* in experiment design. A Latin square is an $n \times n$ array of numbers such that each row and each column contains every number from 1 to n. For example, here is a 4×4 Latin square:

1	2	3	4
3	4	2	1
2	1	4	3
4	3	1	2

You can imagine that this array is an agricultural field where each square is a small plot, and the number inside indicates the variety of barley planted there.

There are some nice connections between Latin squares and graph theory. For example, see if you can construct a graph G_n such that there is a one-to-one correspondence between $n \times n$ Latin squares and valid n-colorings of G_n.

(Recall that an ***n-coloring*** of a graph is a way of assigning one of n colors to each vertex so that vertices joined by an edge are assigned different colors.)

Solution. Create a vertex in G for each entry in the Latin Square. Then connect each vertex to every other vertex in the same row and to every other vertex in the same column. Now color the graph with n colors, each corresponding to a number between 1 and n.

Notice that every pair of vertices in the same row are connected, so no two vertices in the same row can get the same color. Similarly, since every pair of vertices in the same column are connected, no two vertices in the same column can get same color either. These coloring constraint match the constraints on Latin squares, so there is a one-to-one correspondence between colorings of G and $n \times n$ Latin squares.

Perfect Matching from Hall's Theorem

The next question about Latin squares will require a corollary of Hall's Theorem. A bipartite graph is *regular* if every vertex on the left has degree c and every vertex on the right has degree d. Prove the following:

Corollary. *A regular bipartite graph has a matching for the vertices on the left if $c \geq d > 0$.*

Solution. We use Hall's Theorem. Let L denote the set of vertices on the left of the regular bipartite graph, and R denote the set of vertices on the right.

Now let $L' \subseteq L$ be any subset of vertices on the left, and let R' be $N(L')$, namely, the set of all vertices on the right adjacent to some vertex in L'.

Since every edge incident to a vertex in L' is also incident to a vertex in R', the number of edges incident to vertices in L' is at most as large as the number of vertices incident to R'. But the number of edges incident to L' is $c \cdot |L'|$ and the number of edges incident to R' is $d \cdot |R'|$. Thus, we have:

$$c \cdot |L'| \quad \leq \quad d \cdot |R'|$$

Since $c \geq d > 0$ by assumption, this implies $|L'| \leq |R'|$. Thus, the desired condition holds.

Completing Latin Squares

Now we return to Latin squares.

Suppose your teammate wrote only three lines of a 5×5 Latin square before deciding to chuck it all and become a maltster:

2	4	5	3	1
4	1	3	2	5
3	2	1	5	4

Can you fill in the last two lines to extend this "Latin rectangle" to a complete Latin square?

Solution. Here is one possible solution:

2	4	5	3	1
4	1	3	2	5
3	2	1	5	4
1	5	2	4	3
5	3	4	1	2

Show that filling in the next row of a Latin rectangle is equivalent to finding a matching in some bipartite graph.

Solution. Construct a bipartite graph as follows. The vertices on the left are the columns of the Latin rectangle, and the vertices on the right are the numbers 1 to n. Put an edge between a column and a number if the number has *not yet appeared* in the column. Thus, a matching in this graph would associate each column with a distinct number that has not yet appeared in that column. These numbers would form the next row of the Latin rectangle.

Prove that a matching must exist in this bipartite graph and, consequently, a Latin rectangle can always be extended to a Latin square.

Solution. First, we show that the bipartite graph described above has a matching. Each column-vertex on the left has degree $n - k$ and each number-vertex on the right has degree $n - k$ as well. Therefore, a matching for the columns exists by the corollary in the preceding problem. This implies that we can add rows to the Latin rectangle by the procedure described above as long as $k < n$. At that point, we have a Latin square.