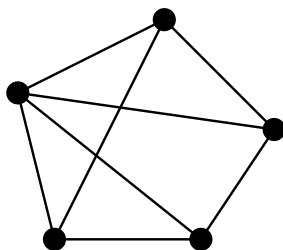


## Graph Theory II

### 1 Coloring Graphs

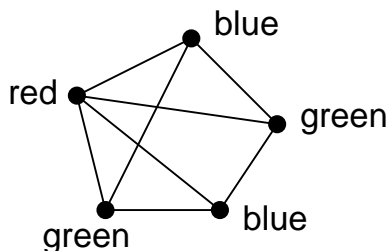
Each term, the MIT Schedules Office must assign a time slot for each final exam. This is not easy, because some students are taking several classes with finals, and a student can take only one test during a particular time slot. The Schedules Office wants to avoid all conflicts, but to make the exam period as short as possible.

We can recast this scheduling problem as a question about coloring the vertices of a graph. Create a vertex for each course with a final exam. Put an edge between two vertices if some student is taking both courses. For example, the scheduling graph might look like this:



Next, identify each time slot with a color. For example, Monday morning is red, Monday afternoon is blue, Tuesday morning is green, etc.

Assigning an exam to a time slot is now equivalent to coloring the corresponding vertex. The main constraint is that adjacent vertices must get different colors; otherwise, some student has two exams at the same time. Furthermore, in order to keep the exam period short, we should try to color all the vertices using as few different colors as possible. For our example graph, three colors suffice:



This coloring corresponds to giving one final on Monday morning (red), two Monday afternoon (blue), and two Tuesday morning (green).

## 1.1 k-Coloring

Many other resource allocation problems boil down to coloring some graph. In general, a graph  $G$  is  $k$ -colorable if each vertex can be assigned one of  $k$  colors so that adjacent vertices get different colors. The smallest sufficient number of colors is called the *chromatic number* of  $G$ . The chromatic number of a graph is generally difficult to compute, but the following theorem provides an upper bound:

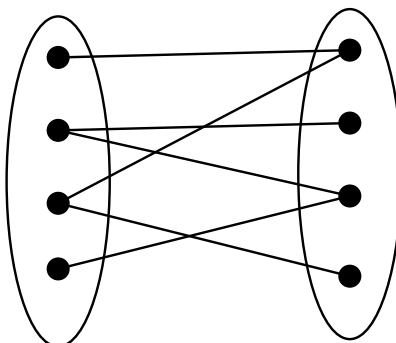
**Theorem 1.** *A graph with maximum degree at most  $k$  is  $(k + 1)$ -colorable.*

*Proof.* We use induction on the number of vertices in the graph, which we denote by  $n$ . Let  $P(n)$  be the proposition that an  $n$ -vertex graph with maximum degree at most  $k$  is  $(k + 1)$ -colorable. A 1-vertex graph has maximum degree 0 and is 1-colorable, so  $P(1)$  is true.

Now assume that  $P(n)$  is true, and let  $G$  be an  $(n + 1)$ -vertex graph with maximum degree at most  $k$ . Remove a vertex  $v$ , leaving an  $n$ -vertex graph  $G'$ . The maximum degree of  $G'$  is at most  $k$ , and so  $G'$  is  $(k + 1)$ -colorable by our assumption  $P(n)$ . Now add back vertex  $v$ . We can assign  $v$  a color different from all adjacent vertices, since  $v$  has degree at most  $k$  and  $k + 1$  colors are available. Therefore,  $G$  is  $(k + 1)$ -colorable. The theorem follows by induction.  $\square$

## 1.2 Bipartite Graphs

The 2-colorable graphs are important enough to merit a special name; they are called *bipartite graphs*. Suppose that  $G$  is bipartite. This means 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. Therefore, every bipartite graph looks something like this:



Bipartite graphs are both useful and common. For example, every path, every tree, and every even-length cycle is bipartite. In turns out, in fact, that every graph not containing an odd cycle is bipartite and vice verse.

**Theorem 2.** *A graph is bipartite if and only if it contains no odd cycle.*

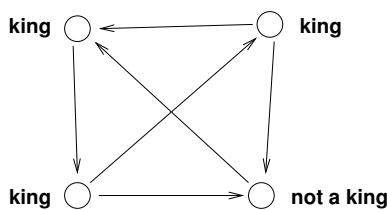
## 2 The King Chicken Theorem

There are  $n$  chickens in a farmyard. For each pair of distinct chickens, either the first pecks the second or the second pecks the first, but not both. We say that chicken  $u$  *virtually pecks* chicken  $v$  if either:

- Chicken  $u$  pecks chicken  $v$ .
- Chicken  $u$  pecks some other chicken  $w$  who in turn pecks chicken  $v$ .

A chicken that virtually pecks every other chicken is called a *king chicken*<sup>1</sup>.

We can model this situation with a tournament digraph. The vertices are chickens, and an edge  $u \rightarrow v$  indicates that chicken  $u$  pecks chicken  $v$ . In the tournament below, three of the four chickens are kings.



Now we're going to prove a theorem about chicken tournaments. The result is not very useful, but the proof involves both induction and digraphs, two of the most common mathematical tools in computer science.

**Theorem 3 (King Chicken Theorem).** *Every  $n$ -chicken tournament has a king, where  $n \geq 1$ .*

*Proof.* The proof is by induction on  $n$ , the number of chickens in the tournament. Let  $P(n)$  be the proposition that in every  $n$ -chicken tournament, there is at least one king.

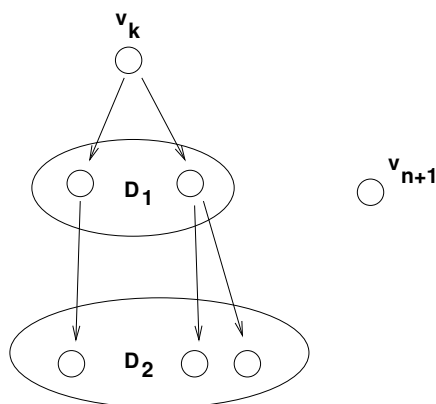
First, we prove  $P(1)$ . In this case, we can safely say that the lone chicken virtually pecks every other chicken, since there are no others. Therefore, the only chicken in the tournament is a king, and so  $P(1)$  is true.

Next, we must show that  $P(n)$  implies  $P(n + 1)$  whenever  $n \geq 1$ . Suppose there is a chicken tournament with chickens  $v_1, \dots, v_{n+1}$ . If we ignore the last chicken for the

<sup>1</sup>But if a chicken is a king, isn't it male? And if it is male, isn't it a rooster? Oh well.

moment, then we are left with a tournament among the first  $n$  chickens. By our induction hypothesis,  $P(n)$ , this tournament has a king chicken,  $v_k$ .

Let  $D_1$  be the set of chickens pecked by the king,  $v_k$ . Let  $D_2$  be the set of chickens virtually pecked by the king, but not pecked directly. Thus, each chicken in  $D_2$  was pecked by some chicken in  $D_1$ . Since  $v_k$  is a king, this accounts for all the chickens; that is,  $\{v_k\}$ ,  $D_1$ , and  $D_2$  form a partition of the set of chickens  $\{v_1, \dots, v_n\}$ . The situation is represented schematically below.



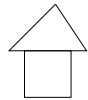
Now we reintroduce the last chicken,  $v_{n+1}$ , and show that the full tournament on  $n + 1$  chickens has a king. There are two cases:

1. Suppose that  $v_k$  pecks  $v_{n+1}$ . Then  $v_k$  is a king of the full tournament.
2. Otherwise,  $v_{n+1}$  pecks  $v_k$ . There are then two subcases:
  - (a) If some chicken in  $D_1$  pecks  $v_{n+1}$ , then  $v_k$  virtually pecks  $v_{n+1}$  and so  $v_k$  is again a king of the full tournament.
  - (b) Otherwise,  $v_{n+1}$  pecks every chicken in  $D_1$ . In this case,  $v_{n+1}$  is a king of the full tournament; he directly pecks  $v_k$  and all the chickens in  $D_1$ , and he virtually pecks all the chickens in  $D_2$ .

In every case, a chicken tournament with  $n + 1$  chickens has a king, and so  $P(n + 1)$  holds. Thus, by the principle of induction, the claim is proved.  $\square$

### 3 Planar Graphs

Here are three dogs and three houses.



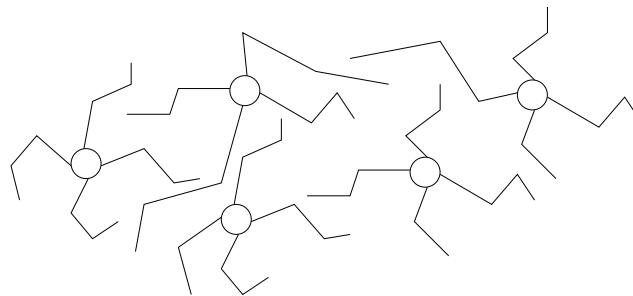
**Dog**

**Dog**

**Dog**

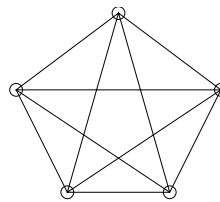
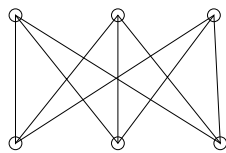
Can you find a path from each dog to each house such that no two paths intersect?

A *quadapus* is a little-known animal similar to an octopus, but with four arms. Here are five quadapi resting on the seafloor:



Can each quadapus simultaneously shake hands with every other in such a way that no arms cross?

Informally, a *planar graph* is a graph that can be drawn in the plane so that no edges cross. Thus, these two puzzles are asking whether the graphs below are planar; that is, whether they can be redrawn so that no edges cross.



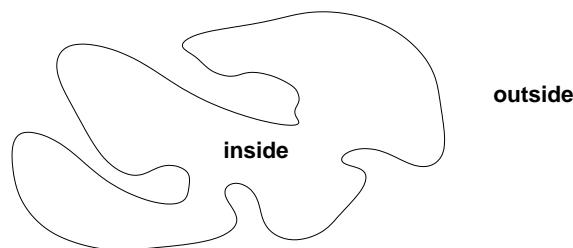
In each case, the answer is, “No— but almost!” In fact, each drawing *would* be possible if any single edge were removed.

More precisely, graph is planar if it has a *planar embedding* (or *drawing*). This is a way of associating each vertex with a distinct point in the plane and each edge with a continuous, non-self-intersecting curve such that:

- The endpoints of the curve associated with an edge  $(u, v)$  are the points associated with vertices  $u$  and  $v$ .

- The curve associated with an edge  $(u, v)$  contains no other vertex point and intersects no other edge curve, except at its endpoints.

This scary definition hints at a theoretical problem associated with planar graphs: while the idea seems intuitively simple, rigorous arguments about planar graphs require some heavy-duty math. The classic example of the difficulties that can arise is the Jordan Curve Theorem. This states that every simple, closed curve separates the plane into two regions, an inside and an outside, like this:



Up until the late 1800's, mathematicians considered this obvious and implicitly treated it as an axiom. However, in 1887 Jordan pointed out that, in principle, this could be a theorem proved from simpler axioms. Actually nailing down such a proof required more than 20 years of effort. (It turns out that there are some *nasty* curves that defy simple arguments.) Planar graphs come up all the time and are worth learning about, but a several-month diversion into topology isn't in the cards. So when we need an "obvious" geometric fact, we'll handle it the old fashioned way: we'll assume it!

Planar graphs are worthy of study for several reasons. One is rooted in human psychology: many kinds of information can be presented as a graph (family relations, chemical structures, computer data structures, contact data for study of disease spread, flow of cash in money laundering trials, etc.). Big graphs are typically incomprehensible messes, but planar graphs are relatively easy for humans to grasp since there are no crisscrossing edges. Sometimes the advantages of planarity are more concrete; for example, when wires are arranged on a surface, like a circuit board or microchip, crossings require troublesome three-dimensional structures. When Steve Wozniak designed the disk drive for the early Apple II computer, he struggled mightily to achieve a nearly planar design:

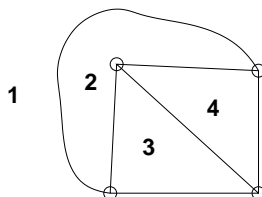
For two weeks, he worked late each night to make a satisfactory design. When he was finished, he found that if he moved a connector he could cut down on feedthroughs, making the board more reliable. To make that move, however, he had to start over in his design. This time it only took twenty hours. He then saw another feedthrough that could be eliminated, and again started over on his design. "The final design was generally recognized by computer engineers as brilliant and was by engineering aesthetics beautiful. Woz later said, 'It's something you can only do if you're the engineer and the PC board layout person yourself. That was an artistic layout. The board has virtually no feedthroughs.'"<sup>2</sup>

<sup>2</sup>From apple2history.org which in turn quotes *Fire in the Valley* by Freiburger and Swaine.

Finally, as we'll see shortly, planar graphs reveal a fundamental truth about the structure of our three-dimensional world.

### 3.1 Euler's Formula

A drawing of a planar graph divides the plane into *faces*, regions bounded by edges of the graph. For example, the drawing below has four faces:



Face 1, which extends off to infinity in all directions, is called the *outside face*. It turns out that the number of vertices and edges in a connected planar graph determine the number of faces in every drawing:

**Theorem 4 (Euler's Formula).** *For every drawing of a connected planar graph*

$$v - e + f = 2$$

where  $v$  is the number of vertices,  $e$  is the number of edges, and  $f$  is the number of faces.

For example, in the drawing above,  $|V| = 4$ ,  $|E| = 6$ , and  $f = 4$ . Sure enough,  $4 - 6 + 4 = 2$ , as Euler's Formula claims.

*Proof.* We use induction on the number of edges in the graph. Let  $P(e)$  be the proposition that  $v - e + f = 2$  for every drawing of a graph  $G$  with  $e$  edges.

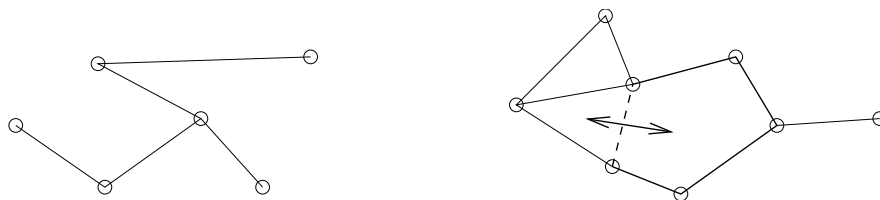
*Base case:* A connected graph with  $e = 0$  edges has  $v = 1$  vertices, and every drawing of the graph has  $f = 1$  faces (the outside face). Thus,  $v - e + f = 1 - 0 + 1 = 2$ , and so  $P(0)$  is true.

*Inductive step:* Now we assume that  $P(e)$  is true in order to prove  $P(e + 1)$  where  $e \geq 0$ . Consider a connected graph  $G$  with  $e + 1$  edges. There are two cases:

1. If  $G$  is acyclic, then the graph is a tree. Thus, there are  $e + 2$  vertices and every drawing has only the outside face. Since  $(e + 2) - (e + 1) + 1 = 2 - 1 + 1 = 2$ ,  $P(n + 1)$  is true.
2. Otherwise,  $G$  has at least one cycle. Select a spanning tree and an edge  $(u, v)$  in the cycle, but not in the tree. (The spanning tree can not contain all edges in the cycle, since trees are acyclic.) Removing  $(u, v)$  merges the two faces on either side of the edge and leaves a graph  $G'$  with only  $e$  edges and some number of vertices  $v$  and faces  $f$ . Graph  $G'$  is connected, because there is a path between every pair of vertices within the spanning tree. So  $v - e + f = 2$  by the induction assumption  $P(e)$ . Thus, the original graph  $G$  had  $v$  vertices,  $e + 1$  edges, and  $f + 1$  faces. Since  $v - (e + 1) + (f + 1) = v - e + f = 2$ ,  $P(n + 1)$  is again true.

The theorem follows by the principle of induction. □

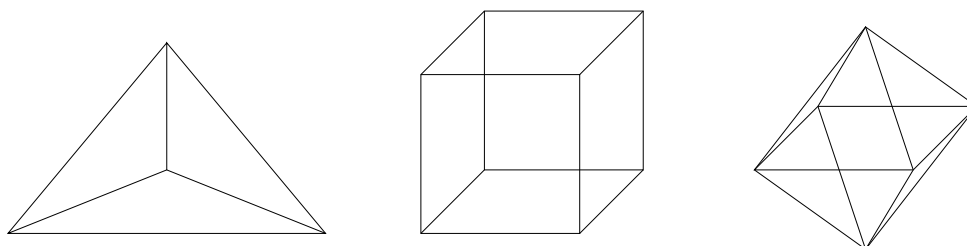
In this argument, we implicitly assumed two geometric facts: a drawing of a tree can not have multiple faces and removing an edge on a cycle merges two faces into one.



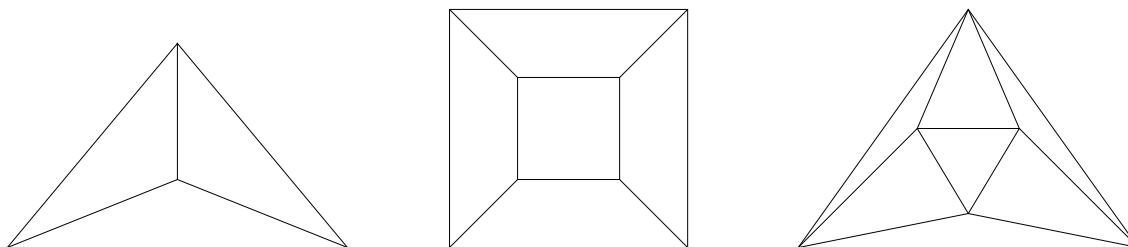
### 3.2 Classifying Polyhedra

The Pythagoreans had two great mathematical secrets, the irrationality of  $\sqrt{2}$  and a geometric construct that we're about to rediscover!

A *polyhedron* is a convex, three-dimensional region bounded by a finite number of polygonal faces. If the faces are identical regular polygons and an equal number of polygons meet at each corner, then the polyhedron is *regular*. Three examples of regular polyhedra are shown below: the tetrahedron, the cube, and the octahedron.



How many more polyhedra are there? Imagine putting your eye very close to one face of a translucent polyhedron. The edges of that face would ring the periphery of your vision and all other edges would be visible within. For example, the three polyhedra above would look something like this:



Thus, we can regard the corners and edges of these polyhedra as the vertices and edges of a planar graph. (This is another logical leap based on geometric intuition.) This means Euler's formula for planar graphs can help guide our search for regular polyhedra.



Let  $m$  be the number of faces that meet at each corner of a polyhedron, and let  $n$  be the number of sides on each face. In the corresponding planar graph, there are  $m$  edges incident to each of the  $v$  vertices. Since each edge is incident to two vertices, we know:

$$mv = 2e$$

Also, each face is bounded by  $n$  edges. Since each edge is on the boundary of two faces, we have:

$$nf = 2e$$

Solving for  $v$  and  $f$  in these equations and then substituting into Euler's formula gives:

$$\begin{aligned} \frac{2e}{m} - e + \frac{2e}{n} &= 2 \\ \frac{1}{m} + \frac{1}{n} &= \frac{1}{e} + \frac{1}{2} \end{aligned}$$

The last equation places strong restrictions on the structure of a polyhedron. Every non-degenerate polygon has at least 3 sides, so  $n \geq 3$ . And at least 3 polygons must meet to form a corner, so  $m \geq 3$ . On the other hand, if either  $n$  or  $m$  were 6 or more, then the left side of the equation could be at most  $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$ , which is less than the right side. Checking the finitely-many cases that remain turns up five solutions. For each valid combination of  $n$  and  $m$ , we can compute the associated number of vertices  $v$ , edges  $e$ , and faces  $f$ . And polyhedra with these properties do actually exist:

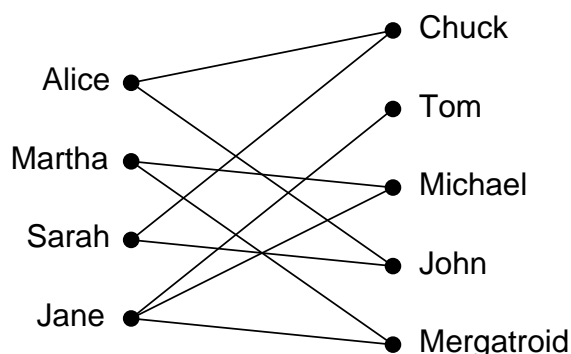
$n$	$m$	$v$	$e$	$f$	polyhedron
3	3	4	6	4	tetrahedron
4	3	8	12	6	cube
3	4	6	12	8	octahedron
3	5	12	30	20	icosahedron
5	3	20	30	12	dodecahedron

The last polyhedron in this list, the dodecahedron, was the other great mathematical secret of the Pythagorean sect!

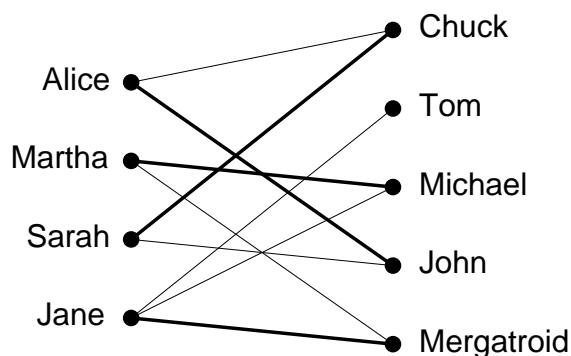
## 4 Hall's Marriage Theorem

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 5.** *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.  $\square$

The proof of this theorem gives an algorithm for finding a matching in a bipartite graph, albeit not a very efficient one. However, efficient algorithms for finding a matching in a bipartite graph do exist. Thus, if a problem can be reduced to finding a matching, the problem is essentially solved from a computational perspective.

## 4.1 A Formal Statement

Let's restate Hall's Theorem in abstract terms so that you'll not always be condemned to saying, "Now this group of little girls likes at least as many little boys..." Suppose  $S$  is a set of vertices in a graph. Define  $N(S)$  to be the set of all neighbors of  $S$ ; that is, all vertices that are adjacent to a vertex in  $S$ , but not actually in  $S$ .

**Theorem 6 (Hall's Theorem).** *Let  $G = (L \cup R, E)$  be a bipartite graph such that every edge has one endpoint in  $L$  and one endpoint in  $R$ . There is a matching for the  $L$  vertices if and only if  $|N(S)| \geq |S|$  for every set  $S \subseteq L$ .*