5 Graph Theory

Graph theory – the mathematical study of how collections of points can be connected – is used today to study problems in economics, physics, chemistry, sociology, linguistics, epidemiology, communication, and countless other fields. As complex networks play fundamental roles in financial markets, national security, the spread of disease, and other national and global issues, there is tremendous work being done in this very beautiful and evolving subject.

The first several sections covered number systems, sets, cardinality, rational and irrational numbers, prime and composites, and several other topics. We also started to learn about the role that definitions play in mathematics, and we have begun to see how mathematicians prove statements called theorems – we've even proven some ourselves. At this point we turn our attention to a beautiful topic in modern mathematics called graph theory. Although this area was first introduced in the 18th century, it did not mature into a field of its own until the last fifty or sixty years. Over that time, it has blossomed into one of the most exciting, and practical, areas of mathematical research.

Many readers will first associate the word 'graph' with the graph of a function, such as that drawn in Figure 4. Although the word graph is commonly

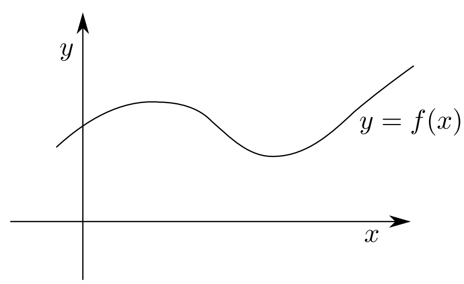


Figure 4: The graph of a function $y = f(x)$.

used in mathematics in this sense, it is also has a second, unrelated, meaning. Before providing a rigorous definition that we will use later, we begin with a very rough description and some examples.

The elementary pieces of graphs are quite simple: points and connections between them. Figure 5 shows several points and connections between some pairs of points. Each point can represent a person, a city, a webpage, or any other object. A connection between two points indicates some relationship between those vertices. For example, connections between two points might represent an existing flight route between two cities, a Facebook friendship between two people, or a pair of webpages that are linked to one another. Graph theory studies networks which can be described in this simplistic manner, as a set of points and connections between them.

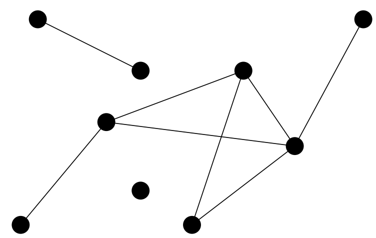


Figure 5: A graph, consisting of points and connections between them.

Example 1. We begin by considering several concrete examples to motivate our interest in studying graphs. Figure 6 shows a 1972 map of flight routes

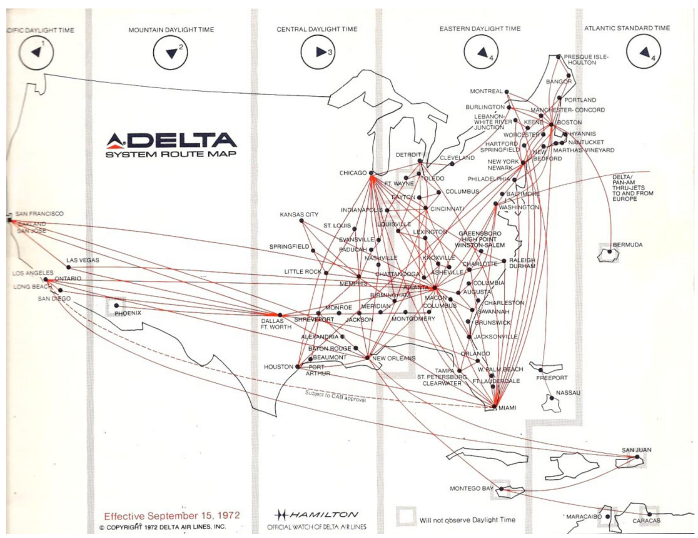


Figure 6: US cities connected by direct Delta Air Lines flights.

serviced by Delta Air Lines. Each black dot indicates a city or airport, and red curves indicates flight routes between pairs of cities. In studying such a map we might consider several very practical questions: Can a traveler leaving Lexington, Kentucky reach his or her destination in Los Angeles flying only Delta flights? If yes, what is the shortest route? What is the maximum number of flights one would need to take between any pair of cities in this map? The reader might notice that some cities are directly connected by flight routes to very few other cities, while other (such as Chicago, Atlanta, and Miami) are connected to many; cities directly connected to many other cities are oftentimes called hubs. One might wonder how adding extra flights between hubs will minimize the average number of connections necessary to travel between different cities. These, any many more questions, are routinely asked when studying maps of this kind.

Example 2. A second exceptionally interesting class of graphs are those found in social networks. In such a graph, each vertex might represent a person, and edges can represent pairs of people who are friends on Facebook. Such graphs raise many interesting questions studied by applied mathematicians and social scientists. For example, is there a good way to estimate the potential influence of various people in a social network? Naively, we might consider the number of friends each person has (or, equivalently, the number of edges incident with each vertex) in determining potential influence. However, the number of Facebook friends might be an inaccurate means of determining influence for the following reason. Consider Jack and Jill, two random Facebook users. Jack has 500 friends, but each of those friends has only 10 friends each. Jill has only 200 friends, but each of her friends has roughly 500 friends themselves. While in some sense Jack is more popular than Jill, he has many fewer friends of friends than does Jill. Of the two, Jill might be the more influential.

When considering the Facebook graph, we might also look for groups of people, each of whom is friends with everyone else in that group. In graph theory, such a subset of vertices is known as a clique. Much cutting-edge research in graph theory studies structural features of graphs, such as cliques.

One important difference between the two examples is the way in which they are drawn. In the first example, cities can be placed on a map in a location corresponding to its actual geographic location. There is a strong geometrical component to the information conveyed by this graph. In the second example, on the other hand, there is no clear way in which to draw the corresponding graph. People do not have unambiguous positions. Therefore, if we are to study a graph of a social network, we will only care about the "graph structure", that is the way in which points are connected, but ignore data regarding positions of the particular points.

Example 3. A third graph that is even more ubiquitous than social networks is that associated with the world-wide web itself. Imagine that we abstract each webpage on the internet as a vertex. An edge between vertex A and B might indicate the existence of some link on A that directs a web-surfer to B. Of course this graph is extremely large (some estimates place this number at over one trillion), and its structure is very complex. However, we might again ask about certain features of this graph. In the 1990's, two graduate students at Stanford developed a way of estimating the importance of a webpage by considering structural features of the graph of all webpages. These graduate students, Larry Page and Sergey Brin, ultimately converted their understanding of graphs into a very practical algorithm called PageRank, which has transformed Google into the largest and most successful search engine of the 21st century. This example introduced a new level of structure beyond what we have seen in the previous examples. In particular, in prior examples, relationships were symmetric in the sense that if A was connected to B, then B, of course, was connected to A. This is certainly the case in the Facebook network, as Jack cannot be friends with Jill if Jill is not also friends with Jack. Likewise, in the airline industry, if a traveler can fly from city A to city B on a given airline, it is generally the case that they can also fly from city B to city A. The world-wide web, however,

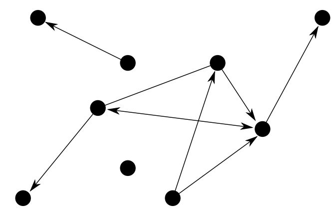


Figure 7: A directed graph, consisting of points and directed connections between them.

provides a more sophisticated kind of graph. It is certainly possible that webpage A links to B without webpage B linking to webpage A. The reader can certainly think for themselves of examples of networks that are non-symmetric. Graphs of such networks are called **directed graphs**. The graphs that we will consider in this class will all be **undirected graphs**.

5.1 Basics

We begin by describing some of the basics of graphs. Roughly speaking, a graph is a set of points and connections between those points; the points are called **vertices** and the connections are called **edges**. A more formal approach to defining a graph is given by the following:

Definition 13. *A* **graph** *G is an ordered pair of sets (V, E). Each element of E is a two-element subset of V. Each element of V is called a* **vertex** *and each element of E is called an* **edge**.

The definition uses sets to define a new mathematical object called a graph. Keeping in mind our previous example, V can be a set of cities, people, or websites. Edges are two-element subsets of V indicating some relationship between those two elements.

Although sets can be thought of abstractly – as a pair of sets with certain properties – it is often convenient to think about graphs visually. For relatively small graphs, we can draw a point for each vertex, and can draw edges between vertices to indicate edges. Consider for example a graph given by $V = \{a, b, c, d\}$ and $E = \{\{a, b\}, \{a, c\}, \{a, d\}\}$. A representation of that graph is shown in Figure 8.

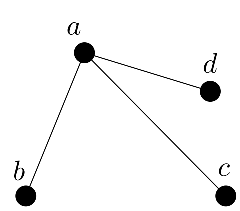


Figure 8: A graph consisting of four vertices and three edges.

Degree

One important property, perhaps the most important property, of a vertex v is the number of other vertices with which it is connected.

Definition 14. *The* **degree** *of a vertex v is the number of edges incident with v; equivalently, it is the number of elements of E of which v is itself an element. The degree of a vertex v is oftentimes written* $\deg(v)$*.*

In the example illustrated in Figure 8, the degree of a is 3, and the degree of b, c, and d is 1. Note that graphs can have vertices that are not connected to any other vertex; the degree of such vertices is 0. If a graph has no edges, then all of its vertices have degree 0. Note also that a graph with n vertices ($|V| = n$) can have vertices with degree at most $n - 1$, since any vertex can be connect to at most the other $n - 1$ vertices.

Relationship of Degrees to Edges

The degrees of the vertices give us one way of counting the number of edges in a graph. More specifically, the following is true for all graphs.

Theorem 7. *For all graphs, the sum of degrees over all vertices is equal to twice the number of edges. In symbols, $\sum_i \deg(v_i) = 2|E|$, where v_i are the vertices of the graph.*

Proof. If an edge is added between vertices u and v of a graph, then the degrees of u and v each increase by 1. Therefore, each edge increases the sum of all degrees by two. □

Example 1

Some graphs have the property that all vertices are connected in some sense:

Definition 15. *A graph is* **connected** *if one can "travel" from any vertex to any other vertex along a series of edges. A graph that is not connected is called* **disconnected***.*

If we think of a graph of the New York City subway system, in which vertices are subway stops and edges indicate direct trains from one subway stop to the next, then this graph is connected. It is always possible to reach any station from any other station in the system. On the other hand, if we consider the Facebook graph, in which vertices are people and edges indicate a pair of people that are friends, then such a graph is disconnected, as there are certainly Facebook users that have 0 friends. Note also that the graph pictured in Figure 5 is disconnected, while that pictured in Figure 8 is connected.

Example 2

Some graphs have the property that every vertex has the same degree. Such graphs are called regular:

Definition 16. *A graph is called* k**-regular** *if the degree of every vertex is* k*.*

Notice that a graph on n vertices can only be k-regular for certain values of k. First, of course k must be less than n, since the degree of any vertex is at most $n-1$. Furthermore, consider a graph with an odd number of vertices. If such a graph were k-regular for an odd value of k, then $\sum_i \deg(v_i)$ would be odd, which is not possible since it is equal to an even number $2|E|$. Therefore, it is not possible to have a k-regular graph on n vertices if both k and v are odd. Note that a regular graph need not be connected.

Example 3

A special type of regular graph is one in which any two vertices are connected by an edge.

Definition 17. *A* **complete graph** *is one such that for any two vertices* $u, v \in V$*, there exists an edge* $\{u, v\} \in E$*.*

A complete graph on n vertices is commonly denoted by K_n. Note that K_n is a $(n-1)$-regular graph. Also note that a complete graph on n vertices has $n(n-1)/2$ edges. We can see this by considering that the degree of each of the n vertices is $n-1$. Therefore, the sum of degrees of all vertices is $n(n-1)$, which according to Theorem 7 is equal to $2|E|$. Therefore, the number of edges is $|E| = n(n-1)/2$. Note also that all complete graphs are connected.

Example 4

Another important kind of graph is one in which the set of vertices and edges can be divided in a very particular way:

Definition 18. *A* **bipartite** *graph is one whose vertices* V *can be divided into two sets* A *and* B *such that there are no edges between any two vertices in* A *or between any two vertices in* B*.*

Notice that in Figure 9, all edges connect one vertex from group A to one vertex from group B; there are no edges between pairs of vertices in A or between

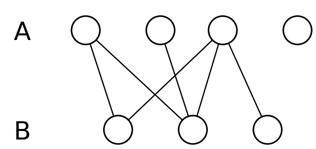
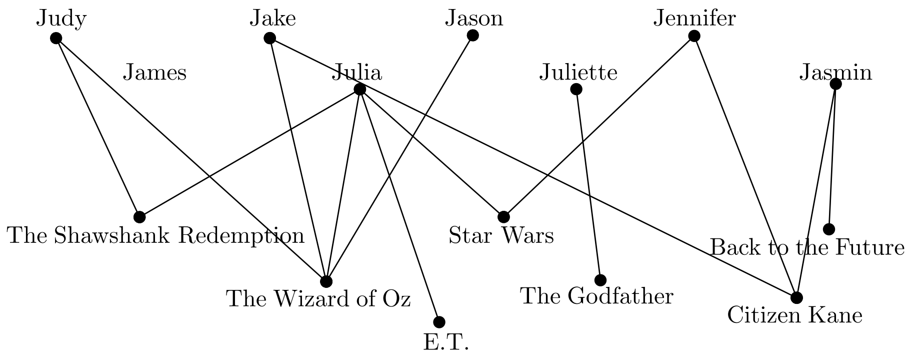


Figure 9: A bipartite graph with six edges on seven vertices.

pairs of vertices in B. One real-life example of a bipartite graph can be seen in the "Netflix Prize" problem. In 2009 Netflix offered a \$1 million dollar prize to the team that could best predict how much a viewer would enjoy a particular movie given their movie preferences. One could view this problem as studying a graph with two sets of vertices: viewers and movies. Edges are drawn between individuals who have viewed a particular movie in the past.



In such a graph it does not make sense to draw an edge between two people or between two movies. Much can be learned about this bipartite graph. We note that it should be clear that a bipartite graph need not be connected. This was clear in Figure 9 and is clear in this figure as well.

We might wonder about the largest number of edges in a bipartite graph. We recall that for a general graph on n vertices, we could have $n(n-1)/2$ edges. In the case of a bipartite graph, however, we should expect the number to be smaller, since there are no edges between vertices of each subset. If we want to add as many edges as possible, we should draw an edge from every vertex in A to every vertex in B. We will then have a total of $|A||B|$ edges. One way to see this is by noticing that the degree of every vertex in A is $|B|$ and the degree of every vertex in B is $|A|$. Using the formula $\sum \deg(v_i) = 2e = 2|E|$, we have $|A||B| + |B||A| = 2|E|$, which of course gives us $|E| = |A||B|$. If you don't find this convincing, experiment a bit using different sets A and B, and think about the maximal number of edges you can draw in such a graph.