Sept 13: Introduction and Eulerian/Hamiltonian Cycles

MAT332 – Geoffrey Scott

We will start with examples of real-world phenomena that can be expressed mathematically using graph theory.

Social Networks: Suppose you want to study a network of five friends: Amadou, Bob, Carol, Deepak, and Esolde. We can represent each person by a dot, and draw an edge between two dots if the corresponding people are friends:

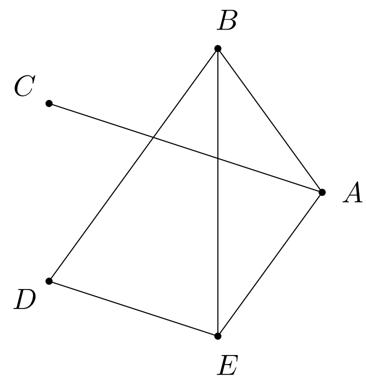

Figure 1: One way to visualize a social network

Is there an efficient algorithm for recommending friends to people? Are there efficient algorithms for grouping someone's friends into different social groups? If there are known "terrorists" in the social network, can you predict who else might be a terrorist? Each of these questions can be answered by abstracting the social network into a dot-and-edge diagram (called a *graph*) as shown above, and applying techniques of graph theory.

Scheduling You are the operations manager for a business with 10 conference rooms, each with different capacities and equipped with different audiovisual technology (projectors, microphones, etc.). One day, ten groups all want to have a meeting at the same time, and each group tells you which rooms would work for them. You can represent each group with a dot, and each room with a dot, and draw an edge between a group's dot and a room's dot if the room is acceptable to the group.

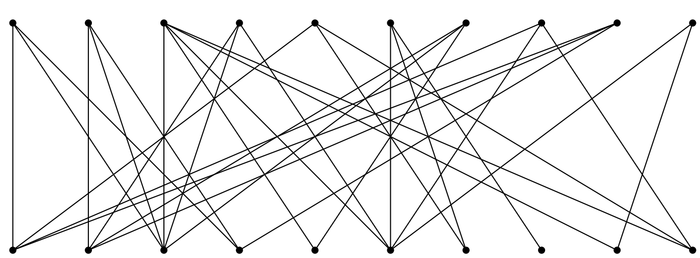

Figure 2: Groups and Rooms

Can you find a collection of edges with the property that match all groups to a different room? Again, this question can be solved using techniques from graph theory.

Puzzles A classic puzzle asks you to suppose that you're on one side of a river with a fox, a goose, and a bag of beans. Your goal is to transport all three to the other side using a boat that can only fit one item at a time. However, the fox cannot be left unsupervised with the goose, and the goose cannot be left unsupervised with the beans. One solution is to travel with the goose, then return empty. Then travel with the fox, and return with the goose. Then travel with the beans, and return empty. Finally, travel with the duck.

To generalize this puzzle, suppose now we have n items, and some pairs of these items cannot be left together unsupervised. You can represent each item with a dot, and put an edge between items that cannot be left unsupervised.

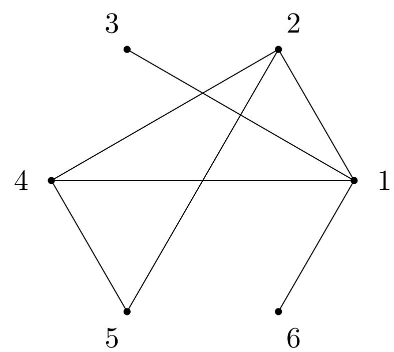

Figure 3: A “supervision” diagram for six items

Is it possible to successfully transport all items across the river? What if your boat can fit 2 items instead of 1? These questions can be answered just by looking at the dot-and-edge diagrams corresponding to the items, again using the techniques of graph theory.

Route Planning Suppose you want to construct a rail network connecting a bunch of cities. If you represent each city by a dot, and each potential route with an edge labeled with the cost of constructing the route, you might end up with a diagram like this.

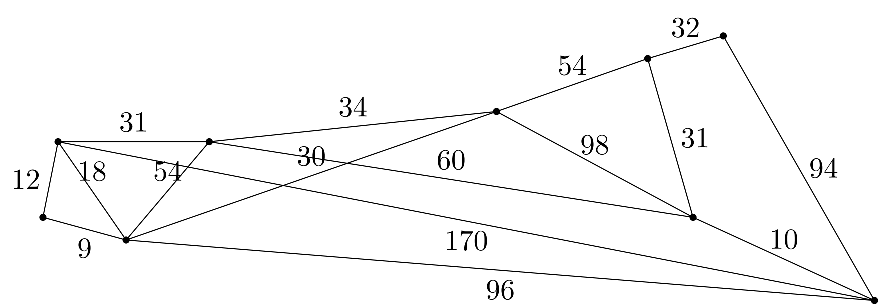

Figure 4: Costs of routes

What is the cheapest rail network for which it’s possible to travel from any city to any other city (perhaps stopping in intermediate cities along the way)? Alternatively, suppose you have two favourite cities: what’s the cheapest rail network you could build that would connect them? Or maybe this diagram represents train routes that have already been built, and you love trains so much that your goal is to take a vacation where you travel on every single train route without repeating any routes. Is it possible?

The last question we posed (about how to find a vacation that will take you on every train route) is how the study of graph theory started, with the so-called *Bridges of Konigsburg* puzzle. That’s how we’ll start the course too, but first we need some notation.

Notation and Handshake Theorem

We want a mathematically rigorous definition of what we mean by "a diagram of dots and of edges between the dots".

> **Definition:** A **graph** G consists of a set $V(G)$ called the **vertex set**, a set $E(G)$ called the **edge set**, and a function that takes each edge $e \in E(G)$ to two vertices (or the same vertex twice). These vertices are called the **endpoints** of the edge e. An edge e is **incident** to a vertex v if v is an endpoint of e.

You've seen the term *graph* before in sentences like "The graph of $y = x^2$ is a parabola." This other definition of the word *graph* is unrelated to the definition above.

As we've seen in the examples, sometimes graphs have some extra structure to them that is relevant to the application they describe. There are is also a lot of language we need to learn to start talking about graphs.

> **Notation**
>
> **Loop:** An edge is called a **loop** if its endpoints are the same.
>
> **Degree:** The **degree** of a vertex $v \in V(G)$ is the number of edges having v as an endpoint; if a *loop* has v as an endpoint, it contributes two to the degree of v.
>
> **Degree sequence:** The **degree sequence** of a graph G is the list of the degrees of the vertices of G, written in decreasing order.
>
> **Regular graph:** A graph G is n**-regular** if every vertex has degree n.

Theorem: Let G be a graph. There are an even number of vertices of G whose degree is odd.

Proof: Each edge in G contributes 2 to $\sum_{v \in V(G)} d(v)$, so

$$\sum_{v \in V(G)} d(v) = 2|E(G)|$$

This shows that $\sum_{v \in V(G)} d(v)$ is even, so the number of v for which $d(v)$ is odd must be even.

Corollary: If some people in a room shake each others hands (and perhaps some pairs of people do not shake hands), then the number of people who have shaken hands with an odd number of people is even.

Proof: Consider the graph G whose vertices are the people in the room, and two vertices are connected by an edge if those two people have shaken hands. The theorem applied to G shows that the number of people that have shaken hands with an odd number of people is even.

More Notation

Simple graph: A graph G is **simple** if no edge of G is a loop and no two edges of G have the same set of endpoints.

Bipartite graph: A graph G is **bipartite** if it is possible to find two disjoint subsets A, B of $V(G)$, such that $V(G) = A \cup B$ and every $e \in E(G)$ has one endpoint in A and one endpoint in B (this last condition is equivalent to: "there are no edges that connect two vertices in A, and there are no edges that connect two vertices in B")

Weighted graph: A **weighted graph** is a graph G and a function $w : E(G) \to \mathbb{R}$ called the **weight function**.

Subgraph: A **subgraph** of a graph G is a subset $V' \subseteq V(G)$ and a subset $E' \subseteq E(G)$ such that all endpoints of edges in E' are in V'. In other words, V' and E' form the vertices and edges of a new graph G'.

Walk/trail/circuit: A **walk** in a graph G is a sequence of the form $v_0 e_1 v_1 e_2 v_2 \ldots e_n v_n$, where the v_i are vertices and the e_i are edges, such that the endpoints of e_i are v_{i-1} and v_i. If no edges are repeated in a walk, it is called a **trail**. A walk (or trail) is **closed** if $v_0 = v_n$. A closed trail is also called a **circuit**.

Connected/Connected component: Two vertices $v, w \in V(G)$ are **connected** if there is a walk in G from v to w. A graph G is called **connected** if every two vertices in G are connected. The **connected components** of G are the maximal connected subgraphs of G – i.e. the subgraphs that are connected and are not contained in a strictly larger connected subgraph.

Mathematical Induction

Suppose you want to prove some claim of the form

For all integers $n \geq 1$, *SOME MATHEMATICAL ASSERTION* is true

for example:

For all integers $n \geq 1$, the sum of the first n positive odd integers equals n^2.

One way to prove statements of this form is a technique called *mathematical induction*, which has two steps:

Steps of Mathematical Induction

Base case: Prove that the statement is true when $n = 1$.

Inductive step: Prove that if you assume the statement is true for an arbitrary $n \geq 1$, then it must also be true for $n + 1$.

Let's use this procedure to prove the claim above

Claim: For all integers $n \geq 1$, the sum of the first n positive odd integers equals n^2.

Proof: The statement is true when $n = 1$ because $1 = 1^2$. Now assume that the statement is true for n. In other words, $n^2 = 1 + 3 + ... + (2n - 1)$. We want to prove that the statement is true for $n + 1$, i.e. we must show $(n+1)^2 = 1 + 3 + ... + (2n - 1) + (2n + 1)$. We expand the left hand side and apply our inductive assumption:

$$(n+1)^2 = n^2 + 2n + 1 = 1 + 3 + ... + (2n - 1) + (2n + 1)$$

Remark: I've written above a recipe for mathematical induction in the box above, but sometimes you need to tweak this recipe. For example, if you want to use induction to prove that some assertion is true for all $n \geq 2$ (or $n \geq 3$) instead of $n \geq 1$, then your base case will be to show that the statement is true for $n = 2$ (or $n = 3$) instead of $n = 1$. Also, sometimes in your inductive step you need to assume the assertion is true for n *and* for $n + 1$ to prove it is true for $n + 2$; this is okay, as long as you prove *two* base cases, one for each of the two smallest values of n. So don't memorize the "induction recipe" in the box above – instead, understand that the general philosophy of induction is "prove that the assertion is true for a sufficient number of simple cases, and then prove that the truth of the assertion for simple cases implies the truth of the assertion for more complicated cases."

Eulerian Circuits

We want to mathematically study the problem of finding a way to take a vacation that travels on all the trains in a railway system. This problem can be stated graph theoretically as trying to find a *trail* in a graph that includes every edge. If you want to end your journey in the same city you started, a *circuit* that includes all edges. Such trails and circuits are called *Eulerian*.

Definition: A trail or circuit is **Eulerian** if it uses all edges of G.

Theorem: A connected graph G has an Eulerian circuit if and only if all vertices of G have even degree.

Proof: Throughout this proof, we will use the metaphor of a circuit $v_0 e_1 \dots v_n$ as a traveler starting at vertex v_0 and "walking" along edges.

- **⇒:** Assume G has an Eulerian circuit, and let v be a vertex which isn't the starting/ending vertex $v_0 = v_n$. Every time the traveler visits v, he leaves v using a different edge than he entered from. Because the traveler eventually uses all edges incident to v, $d(v)$ equals two times the number of times v is visited. Therefore, $d(v)$ is even. If $v = v_0$, the traveler's initial departure from v_0 (in the direction of e_1) contributes 1 to $d(v)$ and the traveler's final arrival back at $v_n = v_0$ (from edge e_n) also contributes 1 to $d(v)$. So the total degree of v equals two times the number of times the vertex is visited in the *middle* of the circuit, plus an extra two to account for the initial departure and final arrival.
- **⇐:** We apply induction on the number of edges in G. If a connected graph has zero edges, it is just a single vertex, so it has an Eulerian circuit. Now assume the claim is true for every connected graph having n edges, and consider a connected graph G with $n + 1$ edges. Let a traveler start at any vertex and walk (without repeating edges) until the traveler cannot walk any more. Because every vertex has even degree, the only way the traveler could be unable to walk farther is if they have returned to

their starting vertex. The traveler's route forms a circuit (not necessarily Eulerian). Delete the edges of this circuit from G to get a new graph G'. This graph may be disconnected, apply the inductive hypothesis to each of its components to get an *Eulerian* circuit in each connected component. Then consider the original circuit in G, but whenever the traveler reaches a vertex from a connected component of G', they travel around the corresponding Eulerian circuit in that connected component.

Corollary: A connected graph G has an Eulerian trail from vertex v to a different vertex w if and only if v and w have odd degree, and all other vertices have even degree.

Proof: Add a new edge e' to G that with endpoints v and w. This new graph has an Eulerian circuit. We may change the starting location (and possibly the direction) so that this Eulerian circuit ends with $we'v$. Delete $e'v$ from the end of this circuit to obtain the desired Eulerian trail.

You can now solve all problems on homework 1 except 1e, 1f, 1g, 2c, 3