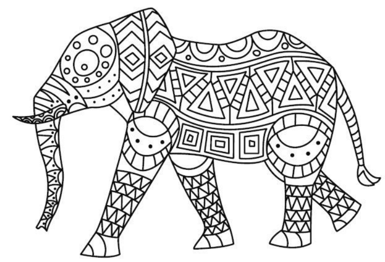
Learning objectives:

- ☐ describe the concept of graph coloring and why it is useful
- ☐ prove that a graph with maximum degree n can be colored with $n+1$ colors

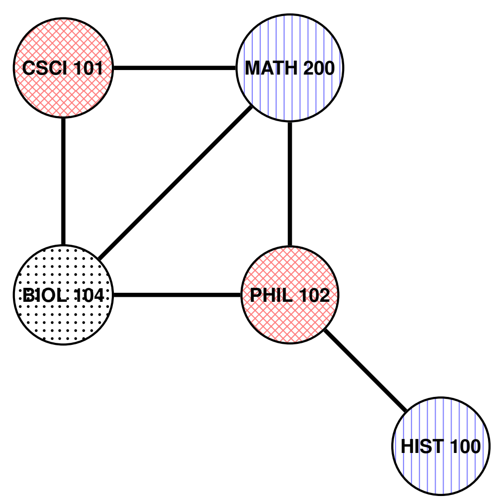
Graph coloring has been around ever since people have tried to color the regions of a map. The main question here is *what is the minimum number of colors needed to color a map, such that no two bordering regions share the same color?* Furthermore, graph coloring has applications to exam scheduling and register allocation. Consider the following example.



Example 1:
Suppose you take up an internship at the Registrar's office and are tasked with scheduling final exams for the end of the semester, which all need to be scheduled during the exam week. There are five courses with final exams. However, some students are enrolled in the same course, so they cannot take exams for different courses in the same time slot! The table below summarizes the cross-enrollment of the various courses (an ✗ means students are in enrolled in both courses for that row and column). How many time slots do you need?

	CSCI 101	MATH 200	PHIL 102	BIOL 104	HIST 100
CSCI 101		✗		✗	
MATH 200	✗		✗	✗	
PHIL 102		✗		✗	✗
BIOL 104	✗	✗	✗		
HIST 100			✗		

Solution:
It helps to write out the graph for the adjacency matrix provided above. All courses that are connected by an edge need to be assigned a different slot. By looking at the coloring of the graph, we only need three colors, so three time slots.



1 Graph coloring problem

Given a graph G and k colors, the graph coloring problems consists of assigning a color to each vertex so that adjacent vertices get different colors. A natural question is: *what is the minimum value for k?*

Definition 1. *The minimum value of k for which such a coloring exists is called the **chromatic number** of the graph. The chromatic number is denoted* *by $\chi(G)$.*

Before trying to solve the problem of coloring a map, we need to make one more definition.

This sounds useful!

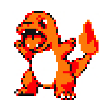
Planar graphs are important because they come up when designing circuits. The problem here is to lay out the circuitry so that no wires cross each other.

Definition 2. *A **planar graph** G is a graph such that the vertices are embedded in the plane and no edges cross each other. In other words, the only intersection points of the edges are at the edge end points. Edges are allowed to be curved.*

Theorem 1 (Four color theorem)**.** *The **four color theorem** states that* any *planar graph can be colored with four colors.*

The problem puzzled mathematicians and cartographers for a long time and was finally proved in 1976 with the help of a computer (so we won't prove it in this course).

Can I use three colors?



Determining whether an arbitrary planar graph is 3-colorable is NP-complete. If you solve it, you win a million dollars.

2 Bounding the chromatic number

You may not always need four colors (see the motivating exam scheduling problem) and three colors might be enough.

This raises the question, *can we place bounds on the chromatic number of a graph?*.

Theorem 2. *A graph with maximum degree at most k can be colored with $(k+1)$ colors.*

Note that we make no assumption about the graph being planar (the theorem applies to any graph).

What's my induction variable?

You might be tempted to use k (the maximum degree of the graph) as your induction variable. A little tip: when doing a proof by induction with a graph, you will most likely want to use the number of vertices or the number of edges as your induction variable.

It's also advisable to start with $n+1$ vertices or edges, then remove one and add it back in as part of your induction step. It's much harder (and will likely be incorrect) to start with an n-vertex (or -edge) graph and add a vertex to create an $(n+1)$-vertex (or -edge) graph.

Proof. We use induction on the number of vertices in a graph $G = (V, E)$, so $n = |V|$. Note that $n \in \mathbb{N}$ since a graph must have at least one vertex (but may have no edges). Let the induction hypothesis be $p(n)$ = *a graph with n vertices and a maximum degree at most k can be colored with $(k+1)$ colors*.

Base case: For $n = 1$, the graph has a single vertex but no edges. So the maximum degree is 0. Since we can use a single color for the vertex, then we can color the graph with $0 + 1 = 1$ colors.

Inductive case: Assume $p(n)$ is true. That is, a graph with n vertices and maximum degree k can be colored with $k+1$ colors. Now, consider a graph $G = (V, E)$ with $n+1$ vertices and assume that the maximum degree of G is k. Remove a vertex v from G to get a n-vertex graph and denote it as $G' = (V', E')$. The maximum degree in G' is at most k since we removed edges (so the degree cannot increase). By our assumption about $p(n)$, G' can be colored with $k+1$ colors.

Now, add v back in. By our assumption, the degree of v will be at most k. Suppose each vertex neighboring v has a different color. Since we have $k+1$ colors to choose from, we can still pick a color for v without it having the same color as one of its neighbors.

Therefore, by induction on n, any graph with maximum degree k is $(k+1)$-colorable. $\square$

Although we can certainly color a graph of maximum degree k with $(k+1)$ colors, it may be possible to use fewer colors.

2.1 Special case: bipartite graphs

In several problems, you can separate the vertices of a graph into **two** distinct sets such that there are no edges connecting vertices of the same set. For example, if you have a set of servers and you need to install software on them, you need to assign variables to registers when writing a compiler, or ensuring that you do not use the same radio frequency over the same territory.

Definition 3. *A graph $G = (V, E)$ is said to be* ***bipartite*** *if the vertices V can be split into V_L and V_R so that all the edges connect a vertex in V_L to a vertex in V_R.*

An important property of bipartite graphs is that the are 2-colorable.

Theorem 3. *A bipartite graph $G = (V, E)$ is 2-colorable.*

Determining whether a graph is 2-colorable reduces to determining if a graph is bipartite, which can be solved in polynomial time (P).