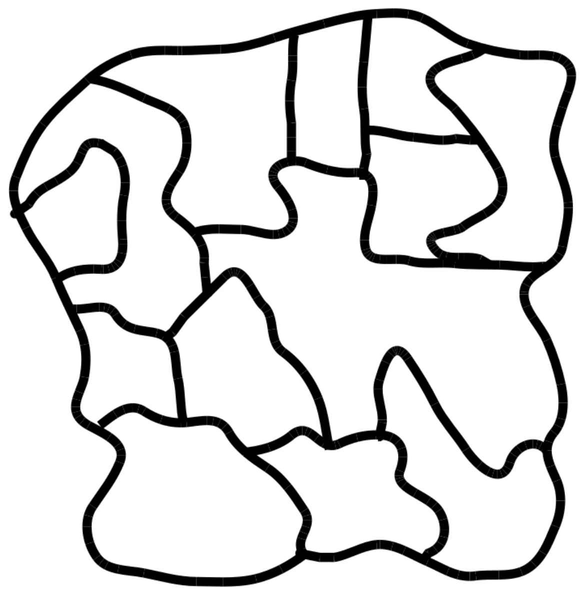
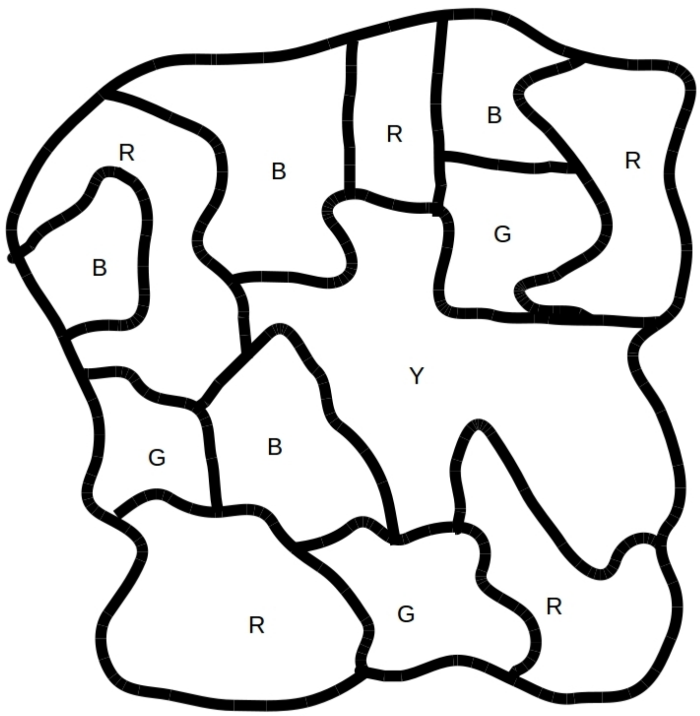
Graph Colouring

Here's a problem that has been intriguing graph theorists for over 250 years. Suppose we have a map of a continent divided up into countries. We will assume that each country consists of a connected piece of land (unlike the USA and Russia, both of which have disconnected bits that are separate from the main territory). The map might look something like this:



Now suppose we are cartographers and we want to make a coloured map of this continent, but we want to make sure that two countries that share a border never have the same colour. The question is, how many colours do we need?

Here’s a colouring of that map that uses 4 colours ... but could we do it with 3? The answer is no ... can you see why?



Now what does this have to do with graph theory? Well we can represent the map with a graph by creating a vertex for each country, and making two vertices adjacent if the countries they represent share a border.

Definition: A **proper colouring** (or **legal** colouring) of a graph is an assignment of colours to vertices so that no adjacent vertices have the same colour.

We have seen that the graph above can be properly coloured using 4 colours, but it can also be properly coloured using 5 or 6 or 7 etc, colours. Since there are 13 countries, we could use 13 colours – giving each vertex an unique colour. We could even say it is possible to properly colour this graph with 50 colours, as long as we don't mind having a bunch of colours left unused.

Definition: The **chromatic number** of a graph G is the smallest integer k such that G can be properly coloured with k colours. We write this as χ(G) = k (that symbol is the Greek letter "chi", pronounced "kye")

So for the graph G shown above, χ(G) = 4.

How large can χ(G) be? In the complete graph K_n, every vertex is adjacent to all other vertices so every vertex must have its own colour. Thus $\chi(K_n)$ = n

How small can χ(G) be? In the complement of K_n, there are no edges at all so every vertex can have the same colour. Thus $\chi(\overline{K_n})$ = 1

Now we can formalize the question we started with. Let G be a graph corresponding to a map as we defined it above. What is the maximum possible chromatic number of G?

People started to ask this question in the 1800's. It was easy to make maps like the one above that have chromatic number = 4, but nobody ever found one with chromatic number = 5. Thus a conjecture was born: No map needs more than 4 colours. (In our notation, if G is a graph corresponding to a map, then χ(G) ≤ 4.) This became known as **the 4-colour conjecture**.

In 1852, a man named **Francis Guthrie** sent the question to the man who had taught him mathematics: **Augustus De Morgan** (the De Morgan who discovered De Morgan's Laws). He sent the question to **Hamilton**, who was a famous graph theorist, and Hamilton sent it to **Cayley**, an extremely important mathematician. None of them could resolve the conjecture.

Then in 1879, a mathematician named **Kempe** published a proof of the conjecture, which then became **the 4-colour theorem** ... except it wasn't. Kempe's proof had an error. The error was discovered in 1890 by **Heawood** – at which point the theorem became a conjecture again.

And there it stayed until 1976, when two mathematicians named **Appel** and **Haken** finally and completely proved that the conjecture is true: there is no map that requires more than 4 colours. (Except that their proof was not quite complete ... but they patched the gaps within a few months.) So the conjecture is a theorem again, and will remain so.

On a personal note, I remember when this happened (the Appel and Haken proof, not the Kempe one!) I was an undergraduate student at the time, but already fascinated by graph theory. It was very exciting to know that such a long-standing mystery had finally been solved. The post-office in the city where Appel and Haken did their work changed their stamp to read "4 Colors Suffice" (since they were in the USA, the word "colours" was spelled incorrectly).

The Appel and Haken proof was controversial because it was the first "computer proof": their proof involved finding proper 4-colourings for a very large set of fairly large graphs – too many for a human to ever check by hand. Some critics rejected the proof on the grounds that it could not be verified by humans. However, soon other people wrote independent programs to repeat and verify the result, and now I don't think there is anyone who disputes it. Nonetheless I suspect that a lot of mathematicians still long for a short and elegant proof of the theorem.

You can read a more detailed history of the conjecture and how it became a theorem here: http://www-history.mcs.st-and.ac.uk/HistTopics/The_four_colour_theorem.html

This is as far as we got in class on November 27, so this is the cut-off point for material that will be on Thursday's test. I'm including notes on the rest of the material I had planned to cover, in case you are interested (I think this graph colouring material is fascinating, but maybe that's just me ...

Anyway, back to our introduction to colouring. We have seen that χ(G) can be any number from 1 to n (where n is the number of vertices in G).

We can put a better upper bound on χ(G)

Theorem: Let Δ be the maximum vertex degree in G. Then $\chi(G) \leq \Delta + 1$

Proof: We will present an argument that it is possible to construct a proper $\Delta + 1$ colouring of G.

Pick any vertex and give it a colour. Now pick another vertex and give it any colour that is not already assigned to one of its neighbours. Continue to pick vertices and colour them. On each iteration of this process, the neighbours of the vertex can use at most Δ different colours because the vertex has at most Δ neighbours. Thus there must be a colour available for the vertex we are colouring.

So for example, if G has 10^6 vertices but $\Delta = 8$, we know $\chi(G) \leq 9$ without having to do any work at all.

It is natural to ask "Given G, how can we compute χ(G) exactly?" ... and this is a great question.

We can easily answer the specific question "Given G, does χ(G) = 1?" ... because χ(G) = 1 if and only if G has vertices but no edges (since any edge forces its ends to have different colours).

We can also answer the question "Given G, does χ(G) = 2?" ... but it's a bit harder.

We can determine if $\chi(G) = 2$ by attempting to build a proper 2-colouring of the graph, as follows:

> Choose a vertex and colour it Red (the choice of vertex and colour doesn't matter)
>
> Now all neighbours of this vertex must be coloured Blue. If any of these vertices are adjacent to each other, they ***can't*** both be Blue, but neither of them can be Red because they are adjacent to the vertex we started with. Thus if any of these vertices are adjacent to each other, G cannot be properly 2-coloured.
>
> If this test is passed, we look at all the neighbours of the vertices we just coloured Blue. All of these vertices ***must*** be coloured Red ... and if any of them are adjacent to each other, the colouring fails and G cannot be properly 2-coloured.
>
> We continue to colour sets of vertices in this manner. We either end with a proper 2-colouring of G, or we find two adjacent vertices to which we must assign the same colour, which tells us that G does not have a proper 2-colouring.

The reason this gives us a definite answer about whether G can be properly 2-coloured is because we never have any options or alternatives. Every time we reach a new vertex, we **must** assign it the colour that is different from the vertices we coloured on the previous iteration.

This is why this process doesn't work if we want to answer "Given G, does $\chi(G) = 3$?" ... with 3 colours we have options.

We can start the same way: pick a vertex and colour it Red. But now when we look at the neighbours of this vertex, we don't know whether we should colour them all Blue, all Green, or some combination of the two. If we guess, we may be wrong ... we may have to try out all possible combinations of Blue and Green on these vertices, and then for their neighbours, we might have to try all the combinations of Red, Blue and Green, and so on for each new group of vertices ... and trying all those combinations will take a long time. Even for relatively small graphs, this is infeasible.

The truth is that this question: "Given G, does $\chi(G) = 3$?" is recognized as ***another*** one of the hardest problems in all of computing science. It is in the same category of difficulty of determining if a graph has Hamilton Cycle and finding the largest k-clique in a graph. (It probably seems as if all the interesting problems are too hard to solve – this is not quite true.)

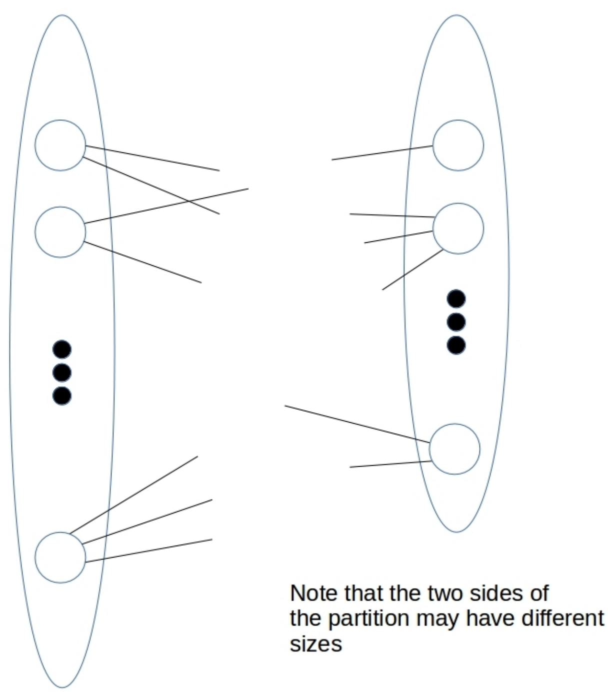
Now this may seem a little bit familiar. Recall that when we looked at the dimension of posets, we found that it is easy to determine if a poset's dimension is either 1 or 2 (although we did not discuss and algorithm to do so), but we have no fast method to determine if a poset's dimension is 3.

These two problems have the strange property that they are easy to solve for 1 and 2, but effectively impossible to solve for any value ≥ 3. There are many other problems that have exactly the same property. Trying to figure out why this is – what causes this sudden and overwhelming increase in difficulty as we go from 2 to 3 – is an active field of research in computing science.

Coming back to Earth, we have one more bit of notation:

Definition: If a graph G is 2-colourable, we call G **bipartite**.

We often draw bipartite graphs like this, with all the Red vertices on one side and all the Blue vertices on the other. Drawn in this fashion, all the edges of the graph go between the two sets of vertices.



Suppose we are told a graph is bipartite – can it be 2-coloured? Since a bipartite graph has no edges between vertices that are in the same part, it is clear that the answer is yet: we can colour all vertices in the first part red, and all the vertices in the second part can be coloured blue, giving a legal 2-colouring.

Thus a graph can be 2-coloured if and only if it is bipartite.

And one final question: All trees on ≥ 2 vertices have the same chromatic number. What is it?