Math 3322: Graph Theory[1] Mikhail Lavrov

Lecture 25: Bounds on chromatic number

November 12, 2024 *Kennesaw State University*

1 Lower bounds on chromatic number

Let's summarize what we know about the chromatic number $\chi(G)$ of a graph G so far:

- For any graph G, $\chi(G) \le \Delta(G) + 1$, where $\Delta(G)$ is the maximum degree.
- If G is a planar graph, then $\chi(G) \le 4$. (We only proved that $\chi(G) \le 6$; proving that $\chi(G) \le 4$ is the much harder Four Color Theorem.)
- If G is an interval graph, then $\chi(G) \le \omega(G)$, where $\omega(G)$ is the clique number.
- If you've looked at the practice problems, you also know that if G has m edges, then $\chi(G) \le 1 + \sqrt{2m}$.

These all have one thing in common: they are upper bounds on $\chi(G)$. That's all we're going to get from looking at variants of the greedy algorithm the way that we've been doing. Arguments like that are telling us that there *is* a way to color G using some number of colors, and here's how.

So how do we prove lower bounds on $\chi(G)$? This requires some kind of different argument: we need to show that colorings with a certain number of colors are impossible.

We will begin with two important lower bounds using quantities we've looked at before.

1.1 Lower bound via clique number

What's the most straightforward graph that requires k colors? It is the complete graph K_k: its k vertices are all adjacent, so all of them need different colors.

Moreover, if a large graph G contains a copy of K_k inside it, then we know that $\chi(G) \ge k$ as well. Even the copy of K_k inside G needs k colors: coloring the other vertices can only make things worse. (In general, if H is a subgraph of G, then $\chi(G) \ge \chi(H)$, because to color G, we must first color H.)

We have a word for a copy of K_k inside G: we called it a clique of size k. The size of the largest clique in G is the clique number $\omega(G)$. We conclude:

Theorem 1.1. *For any graph G, we have $\chi(G) \ge \omega(G)$.*

This lower bound is the easiest of all lower bounds to see. In fact, it takes some work to convince yourself that $\chi(G)$ is not always the same as $\omega(G)$.

[1] This document comes from an archive of the Math 3322 course webpage: http://misha.fish/archive/3322-fall-2024

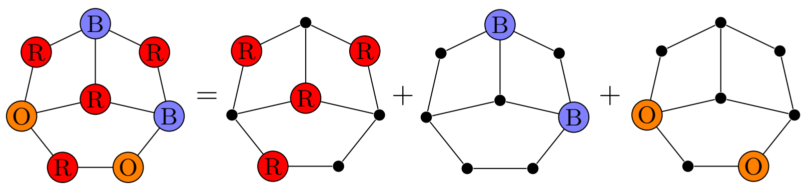
Long odd cycles are the simplest example showing that $\chi(G)$ and $\omega(G)$ can be different. Any odd cycle like C_5 or C_7 or C_{101} has chromatic number at least 3, because it is not bipartite. (In fact, because the maximum degree is 2, the chromatic number can't be more than 3, so it is exactly 3.) However, with the exception of C_3, odd cycles only have clique number 2.

We will see more elaborate examples soon.

1.2 Lower bound via independence number

The independence number $\alpha(G)$ is the exact opposite of the clique number $\omega(G)$: it is the size of the largest set of vertices with *no* edges between them. So it may seem surprising that the independence number can also help us put lower bounds on the chromatic number.

The **color classes** of a coloring are the sets of vertices of each color. For example, if we color the vertices of a graph red, blue, and orange, the set of red vertices is a color class; the set of blue vertices is a color class; the set of orange vertices is a color class. Here is an example illustration:



In a proper coloring, two vertices of the same color can't be adjacent, and therefore the color classes are independent sets. In fact, that's a characterization of proper colorings: they are partitions of the vertices of G into independent sets. (Many important applications of graph coloring come from thinking of colorings in this way.)

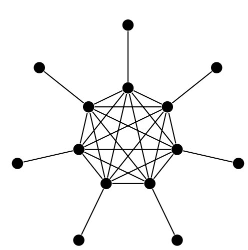
This gives us another lower bound on chromatic number:

Theorem 1.2. *If G is an n-vertex graph with independence number $\alpha(G)$, then $\chi(G) \geq \frac{n}{\alpha(G)}$.*

Proof. The independence number $\alpha(G)$ is the largest number of vertices in any independent set, so in particular every color class in a proper k-coloring has at most $\alpha(G)$ vertices. But the union of all k colors must give us all n vertices, so $k \cdot \alpha(G)$ must be at least n. Rearranging, we get $k \geq \frac{n}{\alpha(G)}$. □

1.3 How good are these bounds?

It is very easy to come up with examples of graphs that "fool" the lower bound $\chi(G) \geq \frac{n}{\alpha(G)}$: graphs that have very large chromatic number, even though $\alpha(G)$ is small. For example, consider the 7-sunlet graph, shown below:



The 7-sunlet graph consists of a 7-vertex clique on the inside, with a "ray" from each vertex of the clique to its own vertex of degree 1. Due to the 7-clique, this graph needs at least 7 vertices to color; in fact, once we color the central clique with 7 colors, it isn't hard to color the rest of the graph without using any other colors. So the chromatic number is 7.

On the other hand, there are 14 vertices and a 7-vertex independent set: take all the outer vertices. So the bound of Theorem 1.2 only tells us that the chromatic number is at least 2: not very helpful!

We can make this example arbitrarily bad by generalizing from the 7-sunlet graph to the n-sunlet graph, built around a clique of size n.

It is much harder to find examples where $\chi(G)$ and $\omega(G)$ are far apart. We will see two situations where this happens.

2 The Mycielski construction

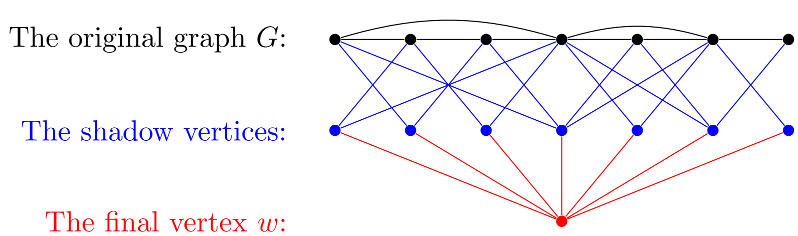
The Mycielski construction is an iterative construction for building graphs where $\chi(G)$ is high, but $\omega(G)$ is low. In fact, the graphs we construct in this section will have $\omega(G) = 2$. A clique of size 3 is also called a **triangle**, because that's what it looks like: 3 vertices with all 3 edges between them. So a graph with $\omega(G) = 2$ is often called **triangle-free**.

To find these graphs, we first define a new operation on graphs. Given a graph G, the **Mycielskian of** G is a graph $M(G)$ constructed as follows.

1. Start with a copy of G, with vertices named $v_1, v_2, \dots, v_n$.
2. For each vertex v_i, add a "shadow vertex" u_i adjacent to all of v_i's neighbors in the copy of G.

   (We never add edges between two different shadow vertices.)
3. Finally, add a vertex w adjacent to all the shadow vertices $u_1, u_2, \dots, u_n$.

Here is an example of this construction in action, with an arbitrary starting graph G. (This is not the graph G we'll ultimately want to apply the construction to.)



Claim 2.1. *If G is triangle-free, then so is $M(G)$.*

Proof. Where could we try to find a triangle in $M(G)$, when there are no triangles in G?

If one vertex of the triangle were the final vertex w, then the other two vertices would both have to be shadow vertices u_i, u_j. This does not work, because u_i and u_j are not adjacent to each other.

Similarly, we cannot have a triangle with more than one shadow vertex in it. So the only hope for creating a triangle is to take one shadow vertex u_i, and two original vertices v_j, v_k. (You can see some triangles like this in the diagram above.)

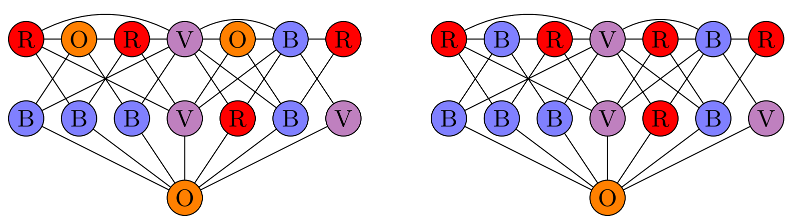
But if all three of these vertices are adjacent, then v_i is also adjacent to v_j and v_k, so v_i, v_j, v_k form a triangle as well! Therefore this cannot happen if G is triangle-free: $M(G)$ must also be triangle-free. □

Claim 2.2. $\chi(M(G)) = \chi(G) + 1$*: the Mycielskian operation increases chromatic number by* 1*.*

Proof. One direction of this claim (which we don't particularly need in the end, but which we'll do anyway because it's easy) is to show that $\chi(M(G)) \le \chi(G) + 1$: if we can color G with $\chi(G)$ colors, we can color $M(G)$ with $\chi(G) + 1$ colors.

To do this, just take a proper k-coloring of G and apply it to the copy of G inside $M(G)$. Then, give every shadow vertex u_i the same color as the color of v_i: this works fine, because it has all the same neighbors as v_i. Finally, give w a new color we did not use in G. This also cannot create any conflicts, so we get a proper $(k+1)$-coloring of $M(G)$.

The other direction is harder: given a proper k-coloring of $M(G)$, we must construct a proper $(k-1)$-coloring of G. The algorithm to do so is this: given a proper k-coloring of $M(G)$, if any vertex v_i has the same color as w, change it to have the same color as u_i (which must be different from w's color, because u_i is adjacent to w). Here is an example:



You'll notice that when we do this, the new coloring is no longer proper: the new color of v_i might conflict with some of its shadow vertex neighbors. However:

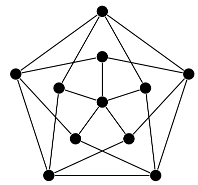
If v_i is adjacent to v_j and we recolor v_i, then v_i and v_j are different colors.

That's because u_i was also adjacent to v_j in the proper coloring we started with, so the color of u_i (which is the new color of v_i) is different from the color of v_j. Also, none of the vertices we recolored were adjacent (because they were all the same color as w), so we don't get any conflicts between two vertices that change colors.

Another way to say the bolded statement is that if we take the new coloring of $M(G)$ and only look at the copy of G inside it, we get a proper coloring of G. That proper coloring uses one fewer color, because the color of w no longer appears on any of $v_1, v_2, \ldots, v_n$. This gives us the $(k-1)$-coloring of G we wanted. □

As a consequence, if we start with an arbitrary triangle-free graph, and apply the Mycielski construction over and over and over, we get a sequence of triangle-free graphs with growing chromatic number. This demonstrates that $\chi(G)$ might be much larger than $\omega(G)$.

Traditionally, we start with K_2: two vertices with an edge. Then $M(K_2) = C_5$: the 5-cycle. $M(C_5)$ is the **Grötzsch graph**, shown below:



The Grötzsch graph is the smallest triangle-free graph with chromatic number 4.

Collectively, the sequence of graphs $K_2, M(K_2), M(M(K_2)), M(M(M(K_2))), \dots$ are sometimes called the **Mycielski graphs**.

3 Coloring random graphs

When we were looking at Ramsey numbers a few lectures ago, we saw that a randomly chosen graph is a good candidate for a graph G where both $\alpha(G)$ and $\omega(G)$ are low.

Let's return to this, but make it a bit more concrete. And to get a large-scale view of the problem, let's consider graphs with $n = 1000000$ (a million) vertices.

There are many such graphs. To pick one of them at random, we flip a coin for each pair of vertices to decide if there is an edge between them. There is a mindbogglingly large number of graphs like this; every time you go to the trouble of flipping all $\binom{1000000}{2}$ coins to get a random 1000000-vertex graph, it is overwhelmingly likely that you are looking at a graph no human has ever seen before. But in some ways, these graphs are very predictable. They all have about $\frac{1}{2}\binom{1000000}{2}$ edges, give or take a few million. And all of them have fairly small clique number and independence number.

Claim 3.1. *If G is a random graph chosen in this way, it is very unlikely that $\omega(G) \geq 40$.*

Proof. There are $\binom{1000000}{40} \approx 1.22 \times 10^{192}$ ways to choose 40 of the vertices of G. Each one of those 40-vertex sets *could* be a 40-vertex clique, if we're lucky.

It *could* be a 40-vertex clique, but that's very unlikely. There are $\binom{40}{2} = 780$ edges between those 40 vertices, so we flip 780 coins to decide which edges between the vertices are present. In order for us to get a clique, all the coinflips have to go one way, which has a probability of $2^{-780} \approx 1.57 \times 10^{-235}$.

As an upper bound on the probability of getting a 40-vertex clique, it's enough to just multiply these two numbers together. If there are about 1.22×10^{192} potential 40-vertex cliques, and each one of them has about a 1.57×10^{-235} chance of actually being a clique, then *even if those are all disjoint events*, the total probability of a 40-vertex clique is only $(1.22 \times 10^{192}) \cdot (1.57 \times 10^{-235})$ or about 1.93×10^{-43}.

That's a tiny probability: we're looking at an event that almost never happens. Our randomly chosen graph on 1000000 vertices almost never has any 40-vertex cliques. $\square$

Similarly:

Claim 3.2. *If G is a random graph chosen in this way, it is very unlikely that $\alpha(G) \geq 40$.*

Proof. The probability here is exactly the same as for cliques. For any given 40-vertex set, if we are flipping coins for each of the 780 edges, there is (again) a 2^{-780} probability that *none* of the edges are present.

So once again, there is at most a 1.93×10^{-43} probability of getting a 40-vertex independent set. □

Let G be a 1000000-vertex graph with $\alpha(G) \leq 40$ and $\omega(G) \leq 40$. What do we know about its chromatic number?

- Theorem 1.1, our bound via clique number, says that $\chi(G) \geq 40$.
- However, Theorem 1.2, our bound via independence number, says that $\chi(G) \geq \frac{1000000}{40} = 25000$. That's much bigger!

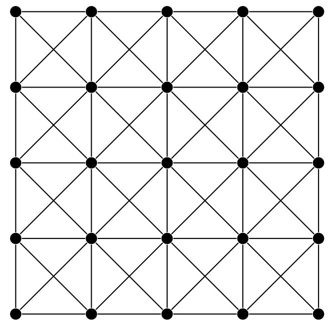
In some sense, this result tells us that for **almost all** large graphs, Theorem 1.2 gives a much more useful bound than Theorem 1.1. Even though clique numbers often seem useful in small examples, those small examples are misleading: large graphs are very different.

This is not to say that Theorem 1.1 is useless. Not every graph we encounter is a randomly-generated graph. And we've already seen that for interval graphs, for example, the clique number always tells us the exact truth about the chromatic number.

In summary, both bounds are useful—and neither should be assumed to be correct without doing more investigation.

4 Practice problems

1. Let G be the 25-vertex graph shown below:



   (a) What is the clique number $\omega(G)$? What lower bound on $\chi(G)$ does it give?

   (b) What is the independence number $\alpha(G)$? What lower bound on $\chi(G)$ does it give?

   (c) What is the maximum degree $\Delta(G)$? What upper bound on $\chi(G)$ does it give?

   (d) What is the actual chromatic number of G?

2. A **wheel graph** W_k is the graph obtained from the cycle graph C_n by adding a vertex adjacent to all vertice of the cycle.

   (a) Draw a diagram of W_5.

   (b) Show that $\omega(W_5) = 3$ but $\chi(W_5) = 4$.

   (c) Generalize this construction to find a graph G in which $\omega(G) = k$ but $\chi(G) = k + 1$, for every $k \geq 3$.

3. (a) Draw the graph $M(K_3)$.

   (b) What are all the triangles in this graph?

   Does $M(K_3)$ have any cliques with 4 or more vertices?

4. Prove the four-color theorem for planar graphs G with $\omega(G) = 2$: triangle-free planar graphs.

   (Note: Grötzsch's theorem—this is the same Grötzsch as the namesake of the Grötzsch graph—proves that in fact all triangle-free planar graphs are 3-colorable.)

5. The Mycielski graphs have another interesting property. They are **edge-critical**: if you delete any edge, the chromatic number decreases.

   Equivalently, a graph G with $\chi(G) = k$ is edge-critical if, for every edge e, there is a $(k-1)$-coloring of G in which both endpoints of e have the same color, but which is otherwise proper.

   (a) Show that C_5 is edge-critical.

   (b) Show that if G is edge-critical, then $M(G)$ is edge-critical. By induction, conclude that all Mycielski graphs are edge-critical.

6. It is interesting to think about coloring the **co-bipartite** graphs: graphs whose complement is bipartite. A graph G is a co-bipartite graph if and only if we can partition $V(G)$ into two sets A and B such that the subgraphs induced by A and by B are both complete graphs (and, additionally, there may be some edges between A and B).

   (a) Show that in any proper coloring of a co-bipartite graph, each color class contains at most 2 vertices.

   (b) Suppose that a co-bipartite graph G has a proper coloring where k color classes have 2 vertices.

   What are those k color classes doing in the complement graph $\overline{G}$ (a bipartite graph)? Relate this to a problem we've already studied this semester.

   (c) What is the connection between a clique in a co-bipartite graph G and a *vertex cover* in the bipartite graph $\overline{G}$?

   (d) Put these things together to prove that if G is a co-bipartite graph, then $\chi(G) = \omega(G)$.

7. The following inequality is true whenever $n_1, n_2, \ldots, n_k$ are positive integers with $n_1 + n_2 + \cdots + n_k = n$:

$$\sum_{i=1}^{k} \binom{n_i}{2} \geq k\binom{n/k}{2} = \frac{n(n-k)}{2k}.$$

   Use this to find the maximum number of edges in an n-vertex graph with chromatic number k.

   (Note: it turns out that the lower bound on chromatic number we get in this way can never be more powerful than the lower bound on chromatic number via clique number. But proving this is harder.)