Sixteenth Lecture: 11/5

Definitions 16.1. Let $G = (V, E)$ be a graph and k a positive integer. A *(vertex) k-coloring* of G is a function $f : V \to \{1, 2, \ldots, k\}$ such that $f(v_1) \neq f(v_2)$ whenever $\{v_1, v_2\} \in E$.

We say that $f(v)$ is the *color* assigned to vertex v. So, in words, a k-coloring of G is an assignment of colors to vertices so that at most k colors are used in all and vertices which are *neighbors*, i.e.: joined by an edge, always get different colors.

The *chromatic number* of a graph G, denoted $\chi(G)$, is the smallest k for which there exists a k-coloring of G.

Example 16.2. $\chi(K_n) = n$ since every pair of vertices are neighbors in K_n and so any coloring must use a different color at each vertex. This observation is generalised in Proposition 16.6 below.

Example 16.3. Let C_n denote a cycle of length $n \geq 3$. Then

$$\chi(C_n) = \begin{cases} 2, & \text{if } n \text{ is even,} \\ 3, & \text{if } n \text{ is odd.} \end{cases}$$

To see this, observe that if one tries to 2-color a cycle one vertex at a time then one will have no choice but to alternate back and forth between the two colors. This will work if n is even, but if n is odd, then we'll need to use a third color at the last vertex.

Remark 16.4. **(i)** To design a general algorithm for computing the chromatic number of a graph G, and/or to find an explicit $\chi(G)$-coloring of it, is another classic example of an NP-complete problem. So it's a a hard problem.
(ii) This problem is often popularised as the *Scheduling Problem*. Suppose you want to make a schedule of exams for courses at GU in Lp4. The basic constraint is that you can't schedule two exams in the same time slot if there is at least one student taking both the courses involved. One would like to make a schedule involving as few slots as possible, simply because each slot costs money to run (rooms must be booked, guards employed etc.). This can be modeled as a graph coloring problem. One lets $G = (V, E)$, where V is the set of all courses which will have an exam this period, and an edge is placed between courses v_1 and v_2 if and only if there is at least one student taking both of them. Then to determine $\chi(G)$ is to determine the minimum number of slots needed in an exam schedule and to determine a specific such schedule is to determine an explicit $\chi(G)$-coloring of the graph. As noted in **(i)**, this problem is extremely hard for general graphs. However, for smallish graphs one can do a kind of simple trial-and-error and "hope for the best", since there is a very simple algorithm for finding *some* coloring, which if you're lucky may not use too many colors. This *greedy coloring algorithm* will be described below. In practice, one is usually satisfied with any coloring which uses "sufficiently few" colors, even if one doesn't know if it's optimal.

Lower bounds for the chromatic number. The most obvious lower bound is gotten from the following concept:

Definition 16.5. Let $G = (V, E)$ be a graph. The *clique number* of G, denoted $\omega(G)$, is the maximum number of vertices in a complete subgraph of G.

Proposition 16.6. *For any graph G one has*

$$\chi(G) \geq \omega(G). \tag{16.1}$$

Proof. Obvious, since the vertices of a clique are all neighbors and must therefore all be assigned different colors in any coloring. ☐

It is easy to find examples of graphs in which the inequality in (16.1) is strict. An important example is the *wheel graph* W_n ($n \geq 3$), see Exercise 15.1.1. It is obvious that $\omega(W_n) = 3$. However, $\chi(W_n) = 4$ when n is odd since the outer n-cycle will then require three colors (see Example 16.3) and a fourth color must be used at the central vertex, as this is joined to everything on the cycle.

The gap between $\omega(G)$ and $\chi(G)$ can be arbitrarily large. To see how pathetic a bound (16.1) is in general, let's first make a definition:

Definition 16.7. The *girth* of a graph G, denoted $g(G)$, is the minimum length of a cycle in G. If G has no cycles, one sets $g(G) := +\infty$.

The following result is due to Erdős:

Theorem 16.8. *For each positive integer* t*, there exists a graph G such that*

$$\min\{\chi(G), g(G)\} \geq t. \tag{16.2}$$

The proof of this theorem is beyond the scope of the course though it's not "too hard". What is interesting is that Erdős' proof uses a probabilistic method and hence, in particular, he only proves the existence of graphs satisfying (16.2) for large t, and does not find any explicit examples. As far as I'm aware, it is still an open problem to construct such graphs explicitly, though Erdős' proof is 70+ years old ! If you're interested, you can find his proof in the book

N. Alon and J. Spencer, *The Probabilistic Method*, 3rd edition (2008), Wiley.

For our purposes, the point is that any graph with girth at least 4 has clique number 2. So one can have $\omega(G) = 2$ and $\chi(G)$ arbitrarily large. Note by the way that all the graphs in Erdős' theorem have *finite* girth, since it follows from the next result that a graph with no cycles at all has chromatic number at most 2:

Theorem 16.9. *For a graph G with at least one edge, the following are equivalent:*
(i) G is bipartite
(ii) G has no cycles of odd length
(iii) $\chi(G) = 2$.

Proof. (i) $\Rightarrow$ (ii) is obvious since, in a bipartite graph, any path must cross back and forth between the two sides and hence, after an odd number of steps we'd be at the opposite side from where we started and could not have completed a cycle. It is also pretty obvious that (i) $\Leftrightarrow$ (iii). For in a bipartite graph there are no edges between vertices on the same side and hence all such vertices can be given the same color. Conversely, if G can be 2-colored, then the vertices in each color form the two sides of a bipartition.

To complete the proof of the theorem, it thus suffices to show that (ii) $\Rightarrow$ (iii). So suppose G has no odd cycles. We may assume G is connected, since G can be 2-colored provided each of its connected components can. We will describe an explicit step-by-step procedure for 2-coloring G. Call the two available colors red and blue.

Step 0: Pick a vertex v at random and color it red.
Step 1: Color each neighbor of v blue. The point is that we can do this, since if any two of v's neighbors were themselves neighbors, then together with v they would form a 3-cycle, contradicting the assumption that G has no odd cycles.
...
Step k: At the k:th step we use the color not used at step $k-1$. The vertices colored at this step are those which are neighbors of the vertices colored at step $k-1$ and which have not yet been colored. Equivalently, the vertices colored at step k are those whose distance (see Definition 14.2) from the vertex v is exactly k. The point is that no two of these vertices are neighbors. For if two of them were, say v_1 and v_2, then G would possess a cycle of length $2k+1$ got by taking a shortest path from v_1 back to v, followed by a shortest path back to v_2 and finally the edge $\{v_2, v_1\}$. This completes the proof. □

The "greedy coloring procedure" described in the above proof, and also hinted at in Example 16.3, can be applied more generally to yield some *upper* bound on the chromatic number of an arbitrary graph. Before we turn to upper bounds, one final remark on lower bounds however. One comment one might have on Theorem 16.8 is that, while it shows that there exist graphs for which $\omega(G)$ and $\chi(G)$ differ by an arbitrarily large amount, is doesn't say anything about whether (16.1) is a good or bad bound "usually". However, it turns out that for "most graphs", $\chi(G)$ is, in fact, much larger than $\omega(G)$. To make this more precise, first some definitions:

Definition 16.10. Let $G = (V, E)$ be a graph. A subset W of V is said to be an *independent set* if $\{w_1, w_2\} \notin E$ whenever $w_1, w_2 \in W$. In other words, there are no edges whatsoever between the vertices of W.

The *independence number* of G, denoted $\alpha(G)$, is the maximum number of vertices in an independent set.

Definition 16.11. Let $G = (V, E)$ be a graph. The *(graph) complement* of G, denoted $\overline{G}$, is the graph whose vertices are the same as those of G and such that $E(\overline{G}) = \binom{V}{2} \backslash E(G)$. In other words, the edges of $\overline{G}$ are precisely those which are absent from G.

Proposition 16.12. *Let $G = (V, E)$ be a graph. Then*
(i) $\alpha(G) = \omega(\overline{G})$.
(ii) $\chi(G) \geq \frac{|V|}{\alpha(G)}$.

Proof. (i) is immediate from the definitions above. For (ii), note that in any coloring of G, the vertices assigned a fixed color must form an independent set. Hence, the number of colors used times the maximum size of an independent set cannot be less than the total number of vertices in the graph (since each vertex gets *some* color). In particular, $\chi(G)\alpha(G) \geq |V|$, v.s.v. $\square$

Now let $V = V_n = \{1, 2, \ldots, n\}$. For each of the $\binom{n}{2}$ pairs $\{i, j\}$ toss a fair coin and insert the edge $\{i, j\}$ if and only if it shows heads. This produces what's called an *(Erdős-Renyí) random graph* and is denoted $G(n, 1/2)$. What it is, more precisely, is a uniform distribution over all possible graphs with vertex set V_n. Note that we are distinguishing here between graphs which don't have exactly the same set of edges, even if they are isomorphic.

Now one can show, though it is beyond our scope to do so, that $\omega(G(n, 1/2)) \sim 2\log_2 n$, that is, a uniformly chosen labelled graph on n vertices will, with high probability, have a clique number close to $2\log_2 n$. But now observe that $G(n, 1/2)$ has the same probability distribution as its complement $\overline{G(n, 1/2)}$, since to choose edges independently with probability $1/2$ is the same as to *not* choose them independently with probability $1/2$. It follows, by part (i) of Proposition 16.12, that also $\alpha(G(n, 1/2)) \sim 2\log_2 n$. But then, by part (ii), the chromatic number of a uniformly random n-vertex labelled graph is, with high probability, on the order of $\frac{n}{2\log_2 n}$ or larger. When n is large, this is way bigger than $2\log_2 n$ so, with high probability, the chromatic number is way bigger than the clique number[1]. Good to know !

Upper bounds for the chromatic number. For a graph $G = (V, E)$ denote

$$\Delta(G) := \max_{v \in V} \deg(v).$$

Theorem 16.13. *(i) For any graph G one has* $\chi(G) \leq \Delta(G) + 1$.
(ii) Moreover, if G is connected and not regular of degree $\Delta(G)$*, then* $\chi(G) \leq \Delta(G)$.

Proof. The theorem follows from an analysis of the so-called *greedy coloring algorithm*, which we have already hinted at in several places.

(i): Order the vertices arbitrarily, say $v_1, v_2, \ldots, v_n$. Denote the available colors as $c_1, c_2, \ldots, c_{\Delta(G)+1}$. We will show how to color G using the available colors. The algorithm is:

Color the vertices in order, at each step using the smallest color not already used on one of the previously colored neighbors of the current vertex.

To see that this works, consider a general step k. At this step we color v_k. Some of its neighbors may have already been colored, but by definition v_k has no more than $\Delta(G)$ neighbors, hence no more than $\Delta(G)$ colors can have already been used on its previously colored neighbors. So we'll be able to color v_k with one of the $\Delta(G) + 1$ available colors.

(ii): The same greedy algorithm will work if we are careful to order the vertices in

[1] I don't see any reason why the same conclusion would not hold if one only distinguished between non-isomorphic graphs, though I haven't (yet) consulted the literature to see if there is a rigorous proof of this.

such a way that

For every vertex, at most $\Delta(G) - 1$ of its neighbors appear before it in the ordering.

So it remains to prove that such an ordering exists. We create the list "backwards". For v_n choose any vertex of degree strictly less than $\Delta(G)$ - our assumption in part (ii) is that at least one such vertex exists. If v_n has r neighbors, say, then put these as $v_{n-1}, \dots, v_{n-r}$, in any order. Proceed in this way: at each step prepend to the list, in any internal order, the neighbors of the vertices prepended at the previous step which are not yet listed.

First of all, since G is connected, every vertex will eventually be listed, thus giving a complete ordering of all n vertices. The procedure ensures that every vertex, apart from v_n, has at least one neighbor appearing after it on the list, hence at most $\Delta(G) - 1$ neighbors appearing before it. But this is also true of v_n, since it has degree at most $\Delta(G) - 1$ in the first place. □

Applications of the greedy algorithm to explicitly coloring graphs can be found in the exercises on the homepage. We finish by noting a very famous result which gives a striking upper bound on the chromatic number of planar graphs:

Theorem 16.14. (Four-Color Theorem) *If G is a planar graph then $\chi(G) \leq 4$.*

This was first proven in the mid-1970s, though the bound $\chi(G) \leq 5$ had been known for a long time and it is easy[2] to prove the bound $\chi(G) \leq 6$. One reason for the fame of this theorem is that it is generally considered to be the first proof of a major result which used electronic computing power in a crucial way. The idea of the proof was a (complicated) kind of induction on the number of vertices in the graph, whereby the authors eventually managed to reduce the proof to checking several thousand specific graphs. There were far too many to check by hand, but by the 1970s computers could manage it. Over the years, various simpler proof of the Four Color Theorem have appeared, though as far as I know every one still reduces eventually to checking a few thousand or so specific graphs. By 2016 this can be done in a matter of milliseconds !

[2] I will leave this as an exercise. HINT: Use Theorem 13.7 to prove that a planar graph must satisfy $e \leq 3v - 6$ and deduce that it must contain a vertex of degree at most 5. Then apply induction on the number of vertices.