Planar Graphs and Graph Coloring

Margaret M. Fleck

1 December 2010

These notes cover facts about graph colorings and planar graphs (sections 9.7 and 9.8 of Rosen)

1 Introduction

So far, we've been looking at general properties of graphs and very general classes of relations. We'll now concentrate on a limited class of graph: simple finite undirected connected graphs. Recall that a simple graph contains no self-loops or multi-edges. and connected means that there's a path between any two vertices.

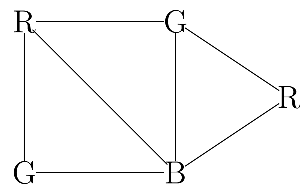
We'll look at two different problems in analyzing these graphs: finding vertex colorings and determining whether a graph can be drawn in the plane without edges crossing.

2 Graph coloring

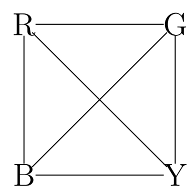
Remember that two vertices are *adjacent* if they are directly connected by an edge.

A *coloring* of a graph G assigns a color to each vertex of G, with the restriction that two adjacent vertices never have the same color. The *chromatic number* of G, written $\chi(G)$, is the smallest number of colors needed to color G.

For example, only three colors are required for this graph:



But K_4 requires 4:



3 Why should I care?

Graph coloring is required for solving a wide range of practical problems. For example, there is a coloring algorithm embedded in most compilers. Because the general problem can't be solved efficiently, the implemented algorithms use limitations or approximations of various sorts so that they can run in a reasonable amount of time.

For example, suppose that we want to allocate broadcast frequencies to local radio stations. In the corresponding graph problem, each station is a vertex and the frequencies are the "colors." Two stations are connected by an edge if they are geographically too close together, so that they would interfere if they used the same frequency. This graph should not be too bad to color in practice, so long as we have a large enough supply of frequencies compared to the numbers of stations clustered near one another.

We can model a sudoku puzzle by setting up one vertex for each square. The colors are the 9 numbers, and some are pre-assigned to certain vertices. Two vertices are connected if their squares are in the same block or row or column. The puzzle is solvable if we can 9-color this graph, respecting the pre-assigned colors.

Probably the oldest sort of coloring problem is coloring the different countries on a map. In this case, the original map is a graph, in which the regions are the countries. To set up the coloring problem, we need to convert this to the *dual graph*, where each region is a vertex and two regions are connected by an edge exactly when they share a border.[1] In this case, the graph is planar (see below), which makes coloring much easier to do.

We can model exam scheduling as a coloring problem. The exams for two courses should not be put at the same time if there is a student who is in both courses. So we can model this as a graph, in which each course is a vertex and courses are connected by edges if they share students. The question is then whether we can color the graph with k colors, where k is the number of exam times in our schedule.

In the exam scheduling problem, we actually expect the answer to be "no," because eliminating conflicts would require an excessive number of exam times. So the real practical problem is: how few students do we have to take out of the picture (i.e. give special conflict exams to) in order to be able to solve the coloring problem with a reasonable value for k. We also have the option of splitting a course (i.e. offering a scheduled conflict exam) to simplify the graph.

A particularly important use of coloring in computer science is register allocation. A large java or C program contains many named variables. But a computer has a smallish number (e.g. 32) of fast registers which can feed basic operations such as addition. So variables must be allocated to specific registers.

The vertices in this coloring problem are variables. The colors are registers. Two variables are connected by an edge if they are in use at the same time and, therefore, cannot share a register. As with the exam scheduling problem, we actually expect the raw coloring problem to fail. The compiler then uses so-called "spill" operations to break up the dependencies and create

[1]Two regions touching at a point are not considered to share a border.

a colorable graph. The goal is to use as few spills as possible.

4 More examples of chromatic numbers

For very small graphs and certain special classes of graphs, we can easily compute the chromatic number. For example, the chromatic number of K_n is n, for any n. Notice that we have to argue two separate things to establish that this is its chromatic number:

- K_n can be colored with n colors.
- K_n cannot be colored with less than n colors

For K_n, both of these facts are fairly obvious. Assigning a different color to each vertex will always result in a well-formed coloring (though it may be a waste of colors). Since each vertex in K_n is adjacent to every other vertex, no two can share a color. So fewer than n colors can't possibly work.

Similarly, the chromatic number for $K_{n,m}$ is 2. We can color one side of the graph with one color and the other side with a second color. And there is clearly no hope of coloring this graph with only one color.

5 A general result

We can also prove a useful general fact about colorability:

Claim 1 *If all vertices in a graph G have degree $\leq D$, then G can be colored with $D + 1$ colors.*

Notice that this is only an upper bound. $D+1$ might not be the chromatic number of G because it might be possible to color G with fewer colors. For example, this theorem says that W_8 (recall: 8 vertices in a circle, one in the center) can be colored with 9 colors. But this graph actually requires only three colors. To do this coloring, alternate two colors around rim of the wheel, then assign a third color to the central vertex.

Proof: by induction on the number of vertices in G.

Base: The graph with just one vertex has maximum degree 0 and can be colored with one color.

Induction: Suppose that any graph with $\leq k$ vertices and maximum vertex degree $\leq D$ can be colored with $D+1$ colors.

Let G be a graph with $k+1$ vertices and maximum vertex degree D. Remove some vertex v (and its edges) from G to create a smaller graph G'.

The maximum vertex degree of G' is no larger than D, because removing a vertex can't increase the degree. So, by the inductive hypothesis, G' can be colored with $D+1$ colors.

Because v has at most D neighbors, its neighbors are only using D of the available colors, leaving a spare color that we can assign to v. The coloring of G' can be extended to a coloring of G with $D+1$ colors.

6 Computing colorings

Unfortunately, in general, graph coloring requires exponential time. There's a couple specific versions of the theoretical problem. I could give you a graph and ask you for its chromatic number. Or I could give you a graph and an integer k and ask whether k colors is enough to be able to color G. These problems are all "NP-complete" or "NP-hard." That is, they apparently require exponential time to compute, even though in some cases it's easy to check that their output is correct. (They are exponential if the "P" and "NP" classes of algorithms are actually different, which is one of the big outstanding problems in computer science.) So, except for small examples and special cases, coloring problems can't be solved exactly.

However, for most practical applications, we can get a reasonable good coloring using a "greedy" method. In this method, we take our graph G and remove vertices one-by-one, creating a series of smaller and smaller graphs. The goal is to ensure that each vertex has a low degree when removed. So we remove low-degree vertices first in hopes that this will simplify the graph structure around vertices with high degree.

We start by coloring the smallest graph and add the vertices back one-by-one, each time extending the coloring to the new vertex. If each vertex has degree $\leq d$ at the point when it's added to the graph, then we can complete the whole coloring with $d + 1$ colors.

This algorithm is quite efficient. Its output might use more colors than the optimal coloring, but it apparently works quite well for problems such as register allocation.

7 Planar graphs

Let's switch our attention to a second question about simple undirected graphs: are they "planar"? A planar graph is a graph which can be drawn in the plane without any edges crossing? For example, K_4 is planar, cube (Q_3) is planar, but $K_{3,3}$ isn't.

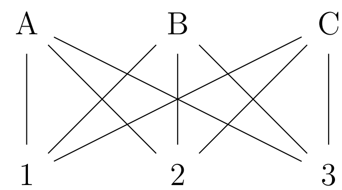
Notice that some pictures of a planar graph may have crossing edges. What makes it planar is that you can draw at least one picture of the graph with no crossings.

Why should we care? Planar graphs have some interesting mathematical properties, e.g. can be colored with only 4 colors. Also, we can use facts about planar graphs to show that there are only 5 Platonic solids.

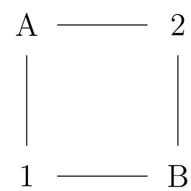
There are also many practical applications with a graph structure in which crossing edges are a nuisance, including design problems for circuits, subways, utility lines. Two crossing connections normally means that the edges must be run at different heights. This isn't a big issue for electrical wires, but it creates extra expense for some types of lines e.g. burying one subway tunnel under another (and therefore deeper than you would ordinarily need). Circuits, in particular, are easier to manufacture if their connections live in fewer layers.

8 $K_{3,3}$ isn't planar

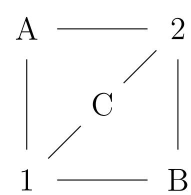
I just claimed that $K3, 3$ isn't planar. Let's see why this is really true. First, let's label the vertices:



The four vertices A, B, 1, and 2 form a cycle.



So C must live inside the cycle or outside the cycle. Let's suppose it lives inside. (The argument is similar if it lives outside.) Our partial graph then looks like:



The final vertex 3 must go into one of the three regions in this diagram. And it's supposed to connect to A, B, and C. But none of the three regions has all three of these vertices on its boundary. So we can't add C and its connections without a crossing.

This proof is ok, but it requires some care to make it convincing. Moreover, it's not going to generalize easily to more complex examples. So we're going to work out some algebraic properties of planar graphs. This will let us prove that certain graphs aren't planar.

9 Faces

A planar graph divides the plane into a set of regions, also called *faces*. Each region is bounded by a simple cycle of the graph: the path bounding each region starts and ends at the same vertex and uses each edge only once. The numer of edges in this boundary is the *degree* of the face. By convention, we also count the unbounded area outside the whole graph as one region.

Examples: a cycle (2 regions), a figure 8 graph (3 regions), two nodes connected by a single edge (1 region).

This neat division of the plane into a set of regions seems intuitively obvious, but actually depends on a result from topology called the "Jordan curve theorem" which states that any simple closed curve (i.e. doesn't cross itself, starts and ends at the same place) divides the plane into exactly two regions. Proving this theorem requires worrying about the possibility that the curve has infinitely complex patterns of maze-like wiggles, but we won't go there.

Since planar graphs are more tightly constrained than general simple graphs, we have two basic formulas beyond our normal handshaking theorem. Specifically, if e is the number of edges, v is the number of vertices, and f is the number of faces/regions, then

- Euler's formula says that $v - e + f = 2$.
- Handshaking theorem: sum of vertex degrees is $2e$
- Handshaking theorem for faces: sum of the face degrees is also $2e$.

To see why the handshaking theorem for faces holds, notice that each edge normally forms part of the boundary of two faces, one to each side of it. The few exceptions involve cases where the edge appears twice as we walk around the boundary of a single face. We'll prove Euler's formula below.

10 Trees

Before we try to prove Euler's formula, let's look at one special type of planar graph: trees. In graph theory, a tree is any connected graph with no cycles. When we normally think of a tree, it has a designated root (top) vertex. In graph theory, these are called *rooted trees*. For what we're doing this class, we don't need to care about which vertex is the root.

A tree doesn't divide the plane into multiple regions, because it doesn't contain any cycles. In graph theory jargon, a tree has only one face: the entire plane surrounding it. So Euler's theorem reduces to $v - e = 1$, i.e. $e = v - 1$. Let's prove that this is true, by induction.

> Proof by induction on the number of vertices in the graph.
>
> Base: If the graph contains no edges and only a single vertex, the formula is clearly true.
>
> Induction: Suppose the formula works for all trees with up to n vertices. Let T be a tree with $n + 1$ vertices. We need to show that T has n edges.
>
> Now, we find a vertex with degree 1 (only one edge going into it). To do this start at any vertex r and follow a path in any direction, without repeating edges. Because T has no cycles, this path can't return to any vertex it has already visited. So it must eventually hit a dead end: the vertex at the end must have degree 1. Call it p.
>
> Remove p and the edge coming into it, making a new tree T' with n vertices. By the inductive hypothesis, T' has $n - 1$ edges. Since T has one more edge than T', T has n edges. Therefore our formula holds for T.

11 Proof of Euler's formula

We can now prove Euler's formula ($v - e + f = 2$) works in general, for any connected planar graph.

> Proof: by induction on the number of edges in the graph.
>
> Base: If $e = 0$, the graph consists of a single vertex with a single region surrounding it. So we have $1 - 0 + 1 = 2$ which is clearly right.
>
> Induction: Suppose the formula works for all graphs with no more than n edges. Let G be a graph with $n + 1$ edges.
>
> Case 1: G doesn't contain a cycle. So G is a tree and we already know the formula works for trees.
>
> Case 2: G contains at least one cycle. Pick an edge p that's on a cycle. Remove p to create a new graph G'.
>
> Since the cycle separates the plane into two regions, the regions to either side of p must be distinct. When we remove the edge p, we merge these two regions. So G' has one fewer regions than G.
>
> Since G' has n edges, the formula works for G' by the induction hypothesis. That is $v' - e' + f' = 2$. But $v' = v$, $e' = e - 1$, and $f' = f - 1$. Substituting, we find that
>
> $$v - (e - 1) + (f - 1) = 2$$
>
> So
>
> $$v - e + f = 2$$

12 A corollary of Euler's formula

Suppose G is a connected simple planar graph, with v vertices, e edges, and f faces, where $v \geq 3$. Then $e \leq 3v - 6$.

Proof: The sum of the degrees of the regions is equal to twice the number of edges. But each region must have degree ≥ 3. So we have $2e \geq 3f$. Then $\frac{2}{3}e \geq f$.

Euler's formula says that $v-e+f=2$, so $f=e-v+2$. Combining this with $\frac{2}{3}e \geq f$, we get

$$e - v + 2 \leq \frac{2}{3}e$$

So $\frac{e}{3} - v + 2 \leq 0$. So $\frac{e}{3} \leq v - 2$. Therefore $e \leq 3v - 6$.

We can also use this formula to show that the graph K_5 isn't planar. K_5 has five vertices and 10 edges. This isn't consistent with the formula $e \leq 3v - 6$. Unfortunately, this trick doesn't work for $K_{3,3}$, which isn't planar but satisfies the equation (with 6 vertices and 9 edges).

13 Another corollary

In a similar way, we can show that if G is a connected planar simple graph with e edges and v vertices, with $v \geq 3$, and if all cycle in G have length ≥ 4, then $e \leq 2v - 4$.

Proof: The sum of the degrees of the regions is equal to twice the number of edges. But each region must have degree ≥ 4 because all cycles have length ≥ 4. So we have $2e \geq 4f$. Then $\frac{1}{2}e \geq f$.

Euler's formula says that $v-e+f=2$, so $e-v+2=f$. Combining this with $\frac{1}{2}e \geq f$, we get

$$e - v + 2 \leq \frac{1}{2}e$$

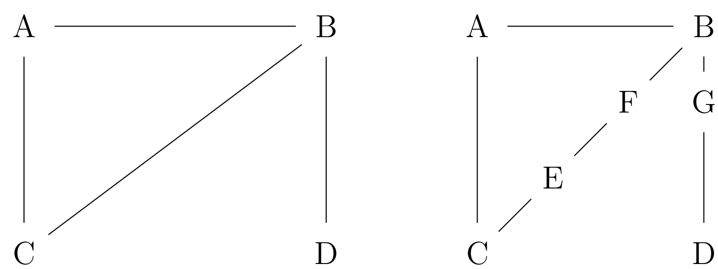
So $\frac{e}{2} - v + 2 \leq 0$. So $\frac{e}{2} \leq v - 2$. Therefore $e \leq 2v - 4$.

$K_{3,3}$ has 9 edges and 6 vertices, which isn't consistent with this formula. So $K_{3,3}$ can't be planar.

14 Kuratowski's Theorem

The two example non-planar graphs $K_{3,3}$ and K_5 weren't picked randomly. It turns out that any non-planar graph must contain a copy of one of these two graphs. Or, sort-of. The copy of $K_{3,3}$ and K_5 doesn't actually have exactly the literal vertex and edge structure of one of those graphs (i.e. be isomorphic). We need to define a looser notion of graph equivalence, called *homeomorphism*.

A graph G is a *subdivision* of another graph F if G is just like F except that you've divided up some of F's edges by adding vertices in the middle of them. For example, in the following picture, the righthand graph is a subdivision of the lefthand graph.



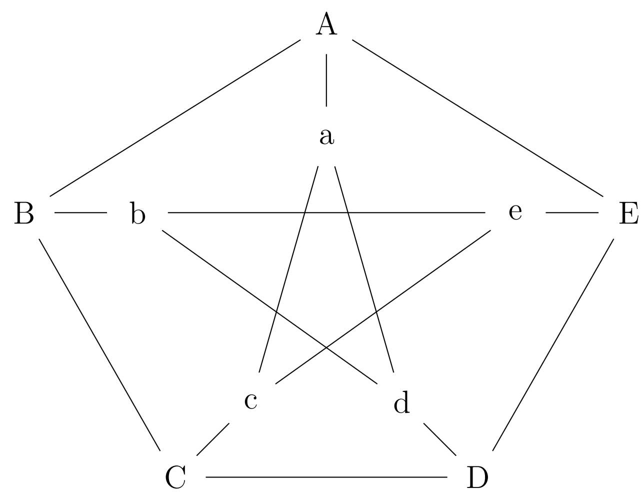
Two graphs are *homeomorphic* if one is a subdivision of another, or they are both subdivisions of some third graph. Graph homeomorphism is a special case of a very general concept from topology: two objects are homeomorphic if you can set up a bijection between their points which is continuous in both directions. For surfaces (e.g. a rubber ball), it means that you can stretch or deform parts of the surface, but not cut holes in it or paste bits of it together.

We can now state our theorem precisely.

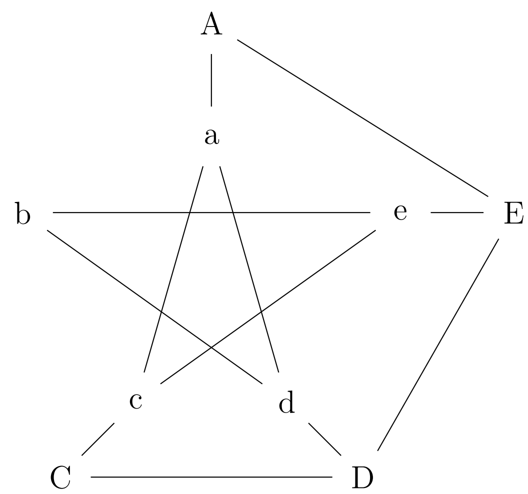
Claim 2 *Kuratowski's Theorem: A graph is nonplanar if and only if it contains a subgraph homeomorphic to* $K_{3,3}$ *or* K_5.

This was proved in 1930 by Kazimierz Kuratowski, and the proof is apparently somewhat difficult. So we'll just see how to apply it.

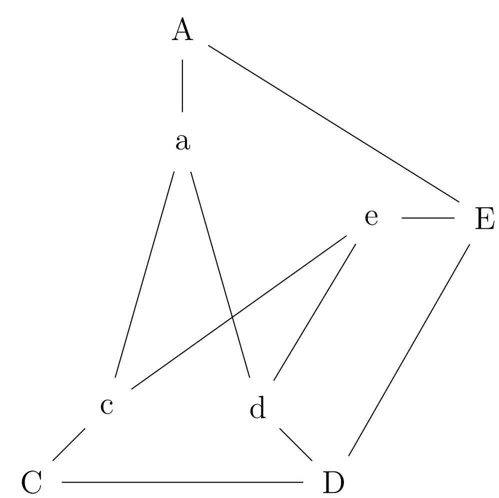
For example, here's a graph known as the Petersen graph (after a Danish mathematician named Julius Petersen).



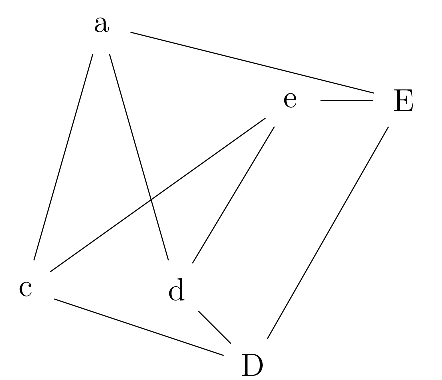
This isn't planar. The offending subgraph is the whole graph, except for the node B (and the edges that connect to B):



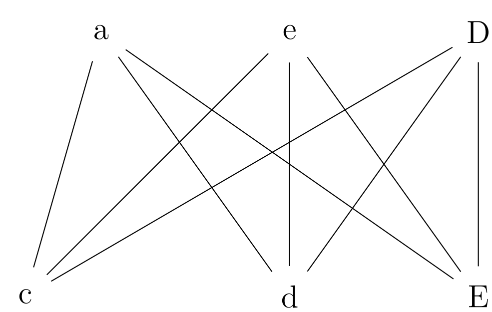
This subgraph is homeomorphic to $K_{3,3}$. To see why, first notice that the node b is just subdividing the edge from d to e, so we can delete it. Or, formally, the previous graph is a subdivision of this graph:



In the same way, we can remove the nodes A and C, to eliminate unnecessary subdivisions:



Now deform the picture a bit and we see that we have $K_{3,3}$.

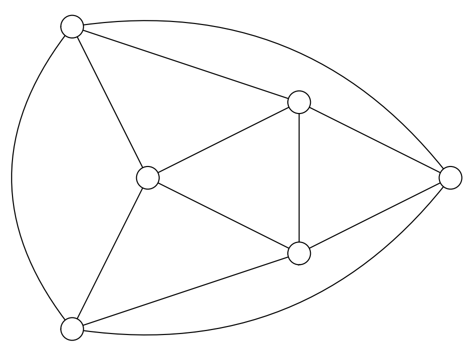


15 Application: Platonic solids

A fact dating back to the Greeks is that there are only five *Platonic solids*. These are convex polyhedra whose faces all have the same number of sides (k) and whose vertices all have the same number of edges going into them (d).

Show a picture of the five Platonic solids from the web: cube, dodecahedron, tetrahedron, icosahedron, octahedron, e.g. wikipedia "Platonic solids".

To turn a Platonic solid into a graph, imagine that it's made of a stretchy material. Make a small hole in one face. Put your fingers into that face and pull sideways, stretching that face really big and making the whole thing flat. For example, an octahedron (8 triangular sides) turns into the following graph. Notice that it still has eight regions, one for each face of the original solid, each with three sides.



Graphs of polyhedra are slightly special planar graphs. Polyhedra aren't allowed to have extra vertices partway along edges, so each vertex in the graph must have degree at least three. Also, since the faces must be flat and the edges straight, each face needs to be bounded by at least three edges.

So, if G is the graph of a Platonic solid, all the vertices of G must have the same degree $d \geq 3$ and all faces must have the same degree $k \geq 3$. I claim that the graphs of the five Platonic solids are the only planar graphs which satisfy these conditions.

Proof: By the handshaking theorem, the sum of the vertex degrees is twice the number of edges. So, since the degrees are equal to d, we have

$$dv = 2e$$

By the handshaking theorem for faces, the sum of the region degrees is also twice the number of edges. That is

$$kf = 2e$$

So this means that $v = \frac{2e}{d}$ and $f = \frac{2e}{k}$

Euler's formula says that $v - e + f = 2$. Substituting into this, we get:

$$\frac{2e}{d} - e + \frac{2e}{k} = 2$$

So

$$\frac{2e}{d} + \frac{2e}{k} = 2 + e$$

Dividing both sides by $2e$:

$$\frac{1}{d} + \frac{1}{k} = \frac{1}{e} + \frac{1}{2}$$

If we analyze this equation, we discover that d and k can't both be larger than 3. If they are both 4 or above, the left side of the equation is at most $\frac{1}{2}$. But since e is positive, the righthand side of the equation must be larger than $\frac{1}{2}$. So one of d and k is actually equal to three and the other is some integer that is at least 3.

Suppose we set d to be 3. Then the equation becomes

$$\frac{1}{3} + \frac{1}{k} = \frac{1}{e} + \frac{1}{2}$$

So

$$\frac{1}{k} = \frac{1}{e} + \frac{1}{6}$$

Since $\frac{1}{e}$ is positive, this means that k can't be any larger than 5.

Similarly, if k is 3, then d can't be any larger than 5.

This leaves us only five possibilities for the degrees d and k: $(3, 3)$, $(3, 4)$, $(3, 5)$, $(4, 3)$, and $(5, 3)$.

Once we've pinned down the degrees of all the vertices in the graph, we've pinned down the basic structure of the graph and of the corresponding solid figure. So there are only five possible graph structures and thus five possible Platonic solids.

At several points in this proof, it's probably not obvious why you would make that step e.g. in the algebra. This is the kind of proof that would have been constructed by trying several ideas and fiddling around with the algebra and the real-world geometrical problem. It's the kind of thing mathematicians do when stuck in the back of a boring committee meeting.

16 Coloring planar graphs

Planar graphs are an important special case for graph coloring, because they are much easier to color than some other graphs. Way back in 1852, Francis Guthrie hypothesized that any planar graph could be colored with only four colors. This is, in fact, correct, but it took a very long time to prove it. Alfred Kempe thought he had proved it in 1879 and it took 11 years for another mathematician to find an error in his proof.

It was finally proved by Kenneth Appel and Wolfgang Haken at UIUC in 1976. They reduced the problem mathematically, but were left with 1936 specific graphs that needed to be checked exhaustively, using a computer program.

5-colorability of planar graphs can be proved without a computer search, but the proof is somewhat messy. However, it's easy to show that all simple planar graphs can be 6-colored.

17 Planar graphs are 6-colorable

Before launching into the proof that any planar graph can be 6-colored, recall from Section 12 that if G is a connected simple planar graph with v vertices and e edges, $v \geq 3$, then $e \leq 3v - 6$. From this, we can deduce that G has a vertex of degree no more than five.

> Proof: This is clearly true if G has one or two vertices.
>
> If G has at least three vertices, we know that $e \leq 3v - 6$. So $2e \leq 6v - 12$.
>
> By the handshaking theorem, $2e$ is the sum of the degrees of the vertices. Suppose that the degree of each vertex was at least 6. Then we would have $2e \geq 6v$. But this contradicts the fact that $2e \leq 6v - 12$.

If our graph G isn't connected, the result still holds, because we can apply our proof to each connected component individually.

This result implies that if we run the greedy coloring algorithm (Section 6) on a planar graph, we'll always be adding back a vertex of degree no more than 5 at each step. This means that the greedy algorithm will always color the graph with (at most) 6 colors.