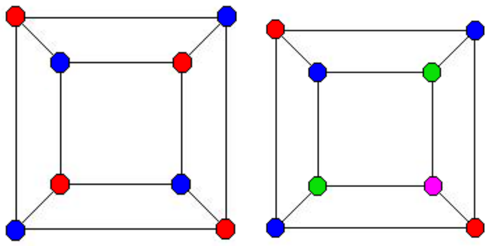
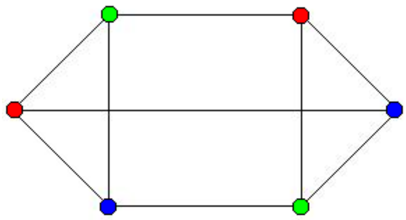
The Four Color Problem

A graph has been colored if a color has been assigned to each vertex in such a way that adjacent vertices have different colors. In other words, a graph has been colored if each edge has two differently colored endpoints.

For example the figure below shows two colorings of the cube.



The cube has chromatic number (It is the smallest number of colors with which it can be colored and will be denoted by "X".) X=2. But the graph below needs at least three colors.



The Four Color Problem dates back to 1852 when Francis Guthrie, while trying to color the map of counties of England noticed that four colors sufficed. He asked his brother Frederick if it was true that any map can be colored using four colors in such a way that adjacent regions (i.e. those sharing a common boundary segment, not just a point) receive different colors. Frederick Guthrie then communicated the conjecture to DeMorgan. The first printed reference is due to Cayley in 1878.

A year later the first `proof' by Kempe appeared; its incorrectness was pointed out by Heawood 11 years later. Another failed proof is due to Tait in 1880; a gap in the argument was pointed out by Petersen in 1891. Both failed proofs did have some value, though. Kempe discovered what became known as Kempe chains, and Tait found an equivalent formulation of the Four Color Theorem in terms of 3-edge-coloring.

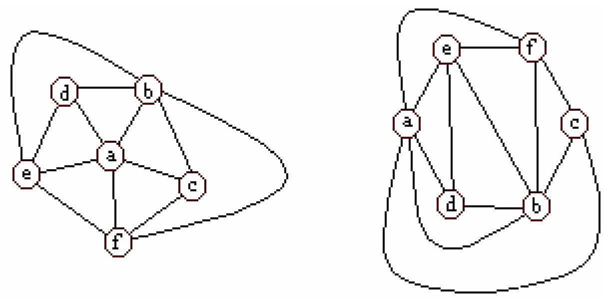
The next major contribution came from Birkhoff whose work allowed Franklin in 1922 to prove that the four color conjecture is true for maps with at most 25 regions. It was also used by other mathematicians to make various forms of progress on the four color problem. We should specifically mention Heesch who developed the two main ingredients needed for the ultimate proof - reducibility and discharging. While the concept of reducibility was studied by other researchers as well, it appears that the idea of discharging, crucial for the unavoidability part of the proof, is due to Heesch, and that it was he who conjectured that a suitable development of this method would solve the Four Color Problem.

This was confirmed by Appel and Haken in 1976, when they published their proof of the Four Color Theorem.

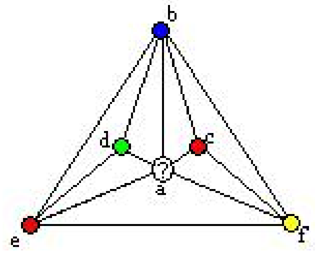
However, the last word on the four – color problem has not been said. The ingenious solution by K. Appel, W. Haken, and J. Koch is a major achievement, but to some mathematicians the solution is unsatisfactory and raises new questions, both mathematical and philosophical.

Also their proof is not approachable enough because it uses computer, and cannot be verified by hand, and even the part that is supposedly hand – checkable is extraordinarily complicated and tedious, and obviously very few people have probably verified it entirely.

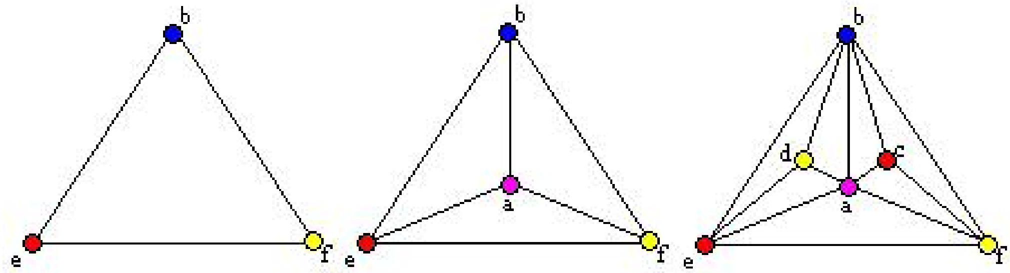
If we construct a graph by adding vertices, one at a time, and make the maximum number of connections at each stage, we will always find the graph plane divided into "triangular" regions, each of which has access to only three vertices. Thus it follows that four colors suffice for any graph that can be constructed in this way. For example, consider the two graphs shown below.



These two graphs each have V = 6 vertices and E = 12 edges, and in fact the two graphs are topologically identical. The plane regions represented by the vertices "a" and "b" each have five adjoining neighbours, whereas the vertices "e" and "f" each have four, and the vertices "c" and "d" each have three. These are complete graphs and, as noted above, a complete graph divides the entire graph plane into "triangular" regions, i.e., regions bounded by three edges connecting three vertices. Nevertheless, depending on how we add the points, it is possible that when a vertex is added to a graph it has more than three neighbours, so we cannot say automatically that four colors would suffice. For example, if vertex "a" is the last to be added, it would have five pre-existing neighbours, and if four colors have already been used in those five vertices, the vertex "a" would require a fifth color.

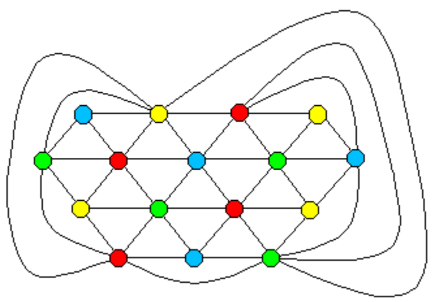


However, we need not add vertex "a" last. Another topologically equivalent way of drawing the above graph is shown below



This shows that we could first assign three distinct colors to the vertices e,b,f, and then place the vertex "a" in this triangle, connect it to each of the three surrounding vertices, and give it a fourth color. Then we can place vertex d inside the triangle abe and give it the same color as f. Then we can place vertex c inside the triangle abf and give it the color of e. Hence the graph is 4-colorable. Moreover, any graph, or portion of a graph bounded by a triangle such as ebf , and having this hierarchical pattern of nested triangles, is 4-colorable. This is the case when the graph contains some vertices with only three edges, and when those vertices and edges are removed, the remaining graph has some vertices with only three edges, which can be removed, and so on, until finally all that remains is a single triangle.

Unfortunately (for the prospect of a simple proof), not every graph is of this hierarchical form. For example, consider the complete graph shown below.



If we denote the number of vertices, edges, and faces (i.e., the bounded regions) of a planar graph by V, E, and F respectively, then Euler's formula for a plane (or a sphere) is V - E + F = 2. Furthermore, each face of a complete graph is bounded by three edges, and each edge is on the boundary of two faces, so we have F = 2E/3, and Euler's formula for a complete planar graph is simply E = 3V - 6. Now, each edge is connected to two vertices, so the total number of attachments (in a complete graph) is 2E = 6V - 12, and hence the average number of attachments per vertex is 6 - 12/V. For any incomplete graph, the total number of attachments is less. Consequently the average number of attachments per vertex for any graph (with a finite number of vertices) is less than 6, which implies that at least one vertex has only five or fewer attachments.

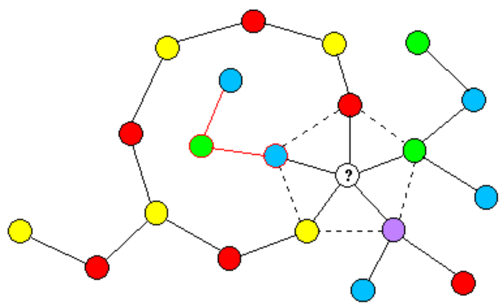
If we have six available colors, a vertex with only five neighbours obviously imposes no constraint on the coloring of the other vertices, because, regardless of the colors of its five (or fewer) neighbours, we can assign it a color without exceeding the six available colors. Therefore if we delete this vertex and all its connections from the graph, creating a graph with one fewer vertices, it's clear that if the resulting graph is 6-colorable, then so was the original graph. Moreover, Euler's formula assures us that this reduced graph also contains at least one vertex with five or fewer neighbours, so we can apply this procedure repeatedly, reducing the graph eventually to one with just 6 vertices, which is obviously 6-colorable. Hence the original graph is 6-colorable.

So, we've seen that Euler's formula immediately implies that no graph can require more than six colors. Furthermore, with just a little more work, we can also show that no graph can require more than five colors. (Ultimately we will see that no graph can require more than four colors, but it's worthwhile to begin with the proof of the 5-colorability of every planar graph.) Obviously every graph with five or fewer vertices is 5-colorable, so if there exists a finite graph that requires more than five colors, it must have more than five vertices. Let the positive integer V6 denote the smallest number of vertices on which there exists a graph that requires six (or more) colors. Conversely, every graph with fewer than V6 vertices is 5-colorable.

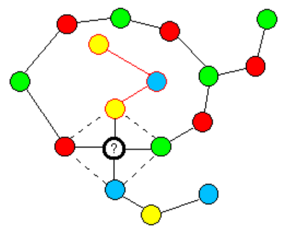
Now, assuming the existence of a graph that requires more than five colors, we can consider one that has exactly V6 vertices. By Euler's formula, this graph must contain at least one vertex with five or fewer connections. However, it cannot contain any vertex with just four (or fewer) connections, because if it did, we could delete such a vertex and leave a graph with just V6 - 1 vertices, which is 5-colorable by definition. Re-inserting the deleted vertex would clearly have no effect on the 5-colorability of the graph, because the vertex has only four (or fewer) neighbours, so the original graph must be 5-colorable, contradicting our assumption. Therefore, a graph with V6 vertices that requires more than five colors cannot contain any vertex with just four or fewer neighbours.

Since Euler's formula implies that the graph contains at least one vertex with five or fewer connections, the only remaining possibility is that the graph contains a vertex with exactly five connections. However, this too leads to a contradiction. To show this, it's helpful to introduce the notion of a k-cluster, which is specified by a set of k distinct colors and one particular vertex that has one of those colors. The original vertex is included in the cluster, and, in addition, every vertex with one of the k specified colors that neighbours a vertex in the cluster is also in the cluster. By definition the only vertices outside a cluster that are directly connected to vertices inside the cluster have colors that are not in the specified set of k colors. Therefore, we can apply any permutation of the k colors to the vertices in a cluster without invalidating the coloration. In particular, a 2-cluster is a contiguous set of vertices, each with one of two specified colors, and we can transpose these two colors without upsetting the coloration of a graph.

Now consider a graph containing a vertex with exactly five immediate neighbours, of five distinct colors, as illustrated below.



Since the uncolored vertex in the center has neighbours of five distinct colors, it might seem that a sixth color is required. However, notice that we can transpose the blue and green colors in the blue/green 2-cluster attached to the upper left vertex of the central pentagon, so that the upper left vertex is green instead of blue. Once we have done this, the uncolored vertex in the center has neighbours of only four distinct colors. If we delete the central vertex, the overall graph has V6 - 1 vertices, so it is 5-colorable, but re-inserting this vertex requires no sixth color (once we have transposed the blue and green in the small 2-cluster as described), so the original graph is 5-colorable.



Respectively, it can be shown that if a planar graph exists, requiring five distinct colors, then it actually needs only four colors.

Neil Robertson, Daniel P. Sanders, Paul Seymour and Robin Thomas announced another proof simpler than Appel and Haken's.

The proof is based on this idea: If a minimal counterexample means a plane graph G that is not 4-colorable, then they show that there is no minimal counterexample.

As a minimal counterexample they consider an internally 6-connected triangulation. They define a near-triangulation as a non-null connected loopless plane graph G such that every finite region is a triangle. A configuration K consists of a near-triangulation G(K) and a map $\gamma_K : V(G(K)) \rightarrow \mathbb{Z}$ with the following properties:

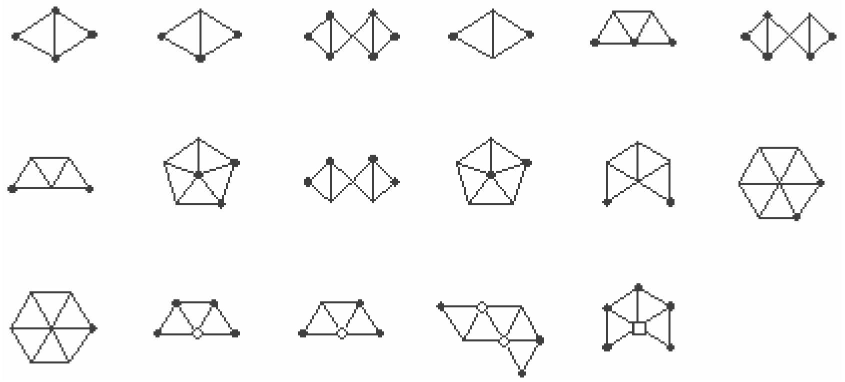
(i) For every vertex v, G(K)\v has at most two components, and if there are two, then $\gamma_K(v) = d(v) + 2$,

(ii) For every vertex v, if v is not incident with the infinite region, then $\gamma_K(v) = d(v)$, and otherwise $\gamma_K(v) > d(v)$; and in either case $\gamma_K(v) \geq 5$,

(iii) K has ring-size ≥ 2

Two configurations are isomorphic if there is a homeomorphism of the plane mapping G(K) to G(L) and γ_K to γ_L.

They exhibit a set of 633 configurations that are essential to their proof. Some of them are these below:



Any configuration isomorphic to one of the 633 configurations is called a good configuration. Let T be a triangulation. A configuration K appears in T if G(K) is an induced subgraph of T, every finite region of G(K) is a region of T, and $\gamma_K(v)=d_T(v)$ for every vertex V in G(K).

They prove the two theorems below:

1. If T is a minimal counterexample, then no good configuration appears in T.
2. For every internally 6-connected triangulation T, some good configuration appears in T.

From these two theorems it follows that no minimal counterexample exists, and so the four color theorem is true.

The way they prove the first theorem is the following:

By a circuit they mean a non-null connected graph in which every vertex has degree two. Let K be a configuration. A near-triangulation S is a free completion of K with ring R if:

1. R is an induced circuit of S, and bounds the infinite region of S,
2. G(K) is an induced subgraph of S, G(K)=S\V(R), every finite region of G(K) is a finite region of S, and the infinite region of G(K) includes R and the infinite region of S,
3. Every vertex v of S not in V(R) has degree $\gamma_K(v)$ in S.

It is easy to check that every configuration has a free completion. Also because there is a homeomorphism between two free completions of K, we may speak of "the" free completion without serious ambiguity.

So, let S be the free completion of a configuration K with ring R. Let C* be the set of all 4-colourings of R, and let C be the set of all restrictions to V(R) of 4-colourings of S. Let C′ be the maximal consistent subset of C*-C. The configuration K is said to be *D-reducible* if C′ is null, and is said to be *C-reducible* if there exists a near-triangulation S′, called reducer, obtained from S by replacing G(K) by a smaller graph (and possibly identifying some vertices of R) such that no member of C′ is the restriction to V(R) of a 4-coloring of S′.

The 633 configurations that they use are either D-reducible of C-reducible.

Then they proved the second theorem:

A configuration K appears in a configuration L if G(K) is an induced subgraph of G(L), every finite region of K is a finite region of L and $\gamma_K(v) = \gamma_L(v)$ for every v in V(G(K)).

Let T be an internally 6-connected triangulation. For a vertex v of T they define the vicinity of v in T to be the configuration K with G(K) the subgraph of T induced by all vertices u of T such that there is a path P in T with ends u and v with at most three vertices, and such that the interior vertex of P (if there is one) has degree at most eight, and $\gamma_K(v)=d_T(v)$ for v in V(G(K)). To prove the second theorem they show that a good configuration appears in the vicinity of some vertex.

Their proof gives a quadratic algorithm to four-color planar graphs:

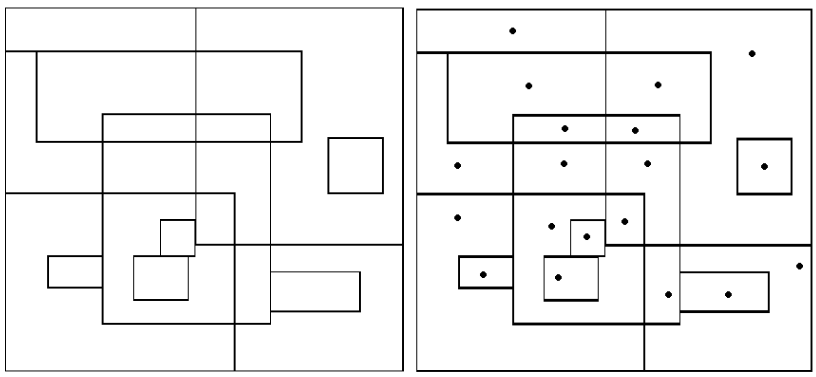
Given an input plane graph G on n vertices (which may as well be assumed to be a triangulation) they find a good configuration appearing in G, replace G by a smaller graph G′, and apply recursion. A reducible configuration can be found in linear time, and a 4-colouring of G can be constructed from a 4-colouring of G′ in linear time. Since there are at most n recursive steps, the overall running time would be quadratic.

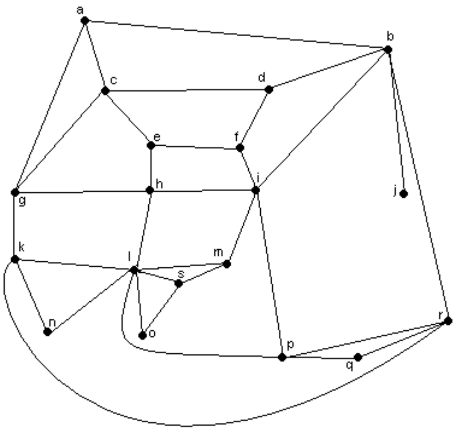
The problem is that this works on an internally 6-connected triangulation. But for each graph they find a set X of vertices of G violating the definition of internal 6-connection, in which case they apply recursion to two carefully selected subgraphs of G, and then obtain a 4-colouring of G by piecing together 4-colourings of the two subgraphs.

There is another algorithm to 4-color a planar map:

Input: A map
Output: 4 – coloring of the map

1. Every region of the map is replaced by a vertex of the graph G, and two vertices are connected by an edge if and only if the two regions share a border segment (not just a corner). (complexity O(n) if n is the number of vertices)





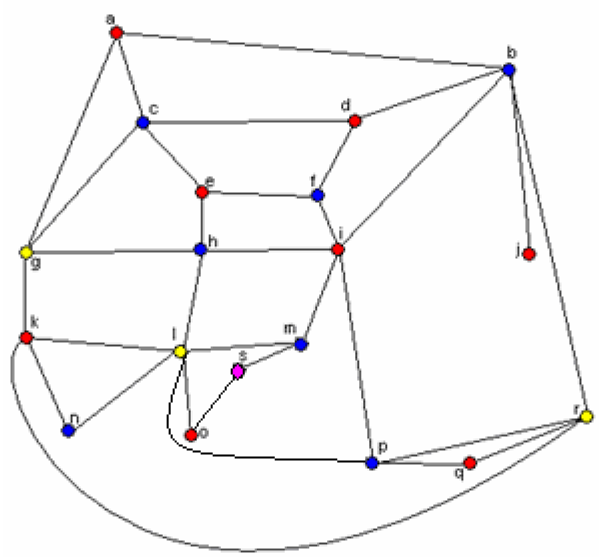
2. We construct a n×n matrix V (actually we use only the part of the matrix under the diagonal), where

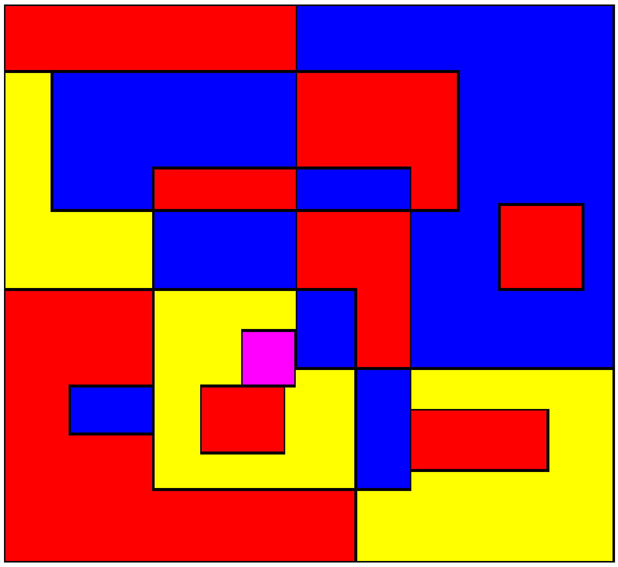
$$V(u,v)=\begin{cases} 1 & \text{, if } u=v \\ 0 & \text{, if } u\neq v \end{cases}$$

The complexity is O(n) because the matrix is zero and the number of the non zero points is actually the number of the edges which is ≤ 3*n−6 according to Euler's formula.

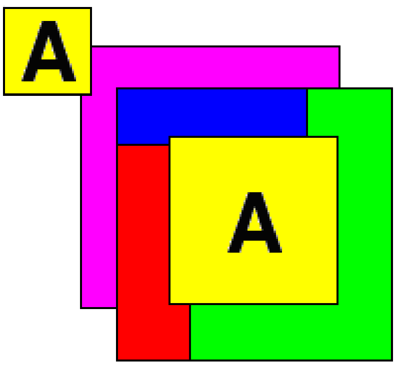
3. Keep the colors RED, BLUE, YELLOW, PURPLE, in this order. O(1)
4. Color every vertex with the first color of the order above, which is not forbidden, and make this color forbidden for its neighbors. O(n)

	a	b	c	d	e	f	g	h	I	j	k	l	m	N	o	p	q	r	s
a		**1**	**1**	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0
	b		0	**1**	0	0	0	0	**1**	**1**	0	0	0	0	0	0	0	**1**	0
		c		**1**	**1**	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0
			d		0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0
				e		**1**	0	**1**	0	0	0	0	0	0	0	0	0	0	0
					f		0	0	**1**	0	0	0	0	0	0	0	0	0	0
						g		**1**	0	0	**1**	0	0	0	0	0	0	0	0
							h		**1**	0	0	**1**	0	0	0	0	0	0	0
								i		0	0	0	**1**	0	0	**1**	0	0	0
									J		0	0	0	0	0	0	0	0	0
										k		**1**	0	**1**	0	0	0	**1**	0
											l		**1**	**1**	**1**	**1**	0	0	**1**
												m		0	0	0	0	0	**1**
													n		0	0	0	0	0
														O		0	0	0	**1**
															p		**1**	**1**	0
																q		**1**	0
																	r		0
																		s	





In the real world, not all countries are contiguous (e.g. Alaska as part of the United States, Nakhichevan as part of Azerbaijan, and Kaliningrad as part of Russia). If the chosen coloring scheme requires that the territory of a particular country must be the same color, four colors may not be sufficient. Conceptually, a constraint such as this enables the map to become non-planar, and thus the four color theorem no longer applies. For instance, consider a simplified map:



In this map, the two regions labeled A belong to the same country, and must be the same color. This map then requires five colors, since the two A regions together are contiguous with four other regions, each of which is contiguous with all the others. If A consisted of three regions, six or more colors might be required; one can construct maps that require an arbitrarily high number of colors.