PRIMES Circle: Graph Theory

Tal Berdichevsky and Corinne Mulvey

May 17, 2019

Contents

1	Introduction to Graphs	2
	1.1 Basics	
	Planar Graphs	6
	Graph Colorings	8
	3.1 Introduction to Vertex Coloring	
	3.2 Four and Five Color Theorems	10

1 Introduction to Graphs

How can we model Facebook friends? Here's a diagram of some friends. We can let each point represent an individual person and each line drawn between two points represent a friendship between two people.

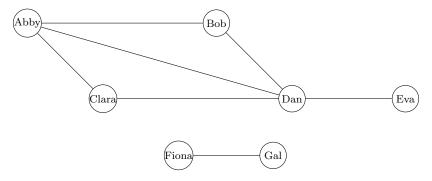


Figure 1: Graph of Facebook Friends

1.1 Basics

Let's formalize what we just said. The representation we created is an example of a graph. Each friend, represented by the points, can be thought of as a vertex of the graph and the friendships, represented by the lines drawn between points, are edges of the graph. More specifically,

Definition 1.1. A graph G is a set of vertices V(G) and edges E(G). Each edge is a "connection" between two vertices, and can be represented as a pair of vertices (u, v). Each edge is said to be incident to the two vertices it connects.

Example 1.1. In Figure 1, the vertices are people and edges represent friendships. Abby, Bob, Clara, Dan, Eva, Fiona, and Gal are all vertices and make up the vertex set of the graph.

Definition 1.2. Two edges u and v are adjacent if there is an edge between them; in other words, $(u, v) \in E(G)$.

Example 1.2. In Figure 1, any two adjacent vertices, or people, are friends.

Definition 1.3. The *order* and *size* of a graph G is the cardinality of its vertex set |V(G)| and the cardinality of its edge set |E(G)|, respectively.

Example 1.3. The size of the graph in Figure 1 representing Facebook friends is the number of friendships in the network or edges in the graph. The order of graph G is the number of friends in the group. The order and the size of the graph are therefore both equal to 7.

Definition 1.4. For a graph G and vertex $v \in V(G)$, the degree $\deg_G(v)$ of v is the number of edges incident with v. We denote the minimum degree of all vertices in G as $\delta(G)$, and the maximum degree we denote as $\Delta(G)$.

Example 1.4. In Figure 1, a person's degree is the number of friends (s)he has. Clara has two friends and therefore $\deg_G(Clara)=2$. The numbers $\delta(G)$ and $\Delta(G)$ of the social network in Figure 1 represent the people with the least and most friends, respectively. Dan has five friends which is the most friends of anyone in the social network, therefore $\Delta(G)=5$. Fiona, Gal, and Eva, all each have only one friend, which is the least number of friends anyone in the network has. Therefore $\delta(G)=1$.

Theorem 1.1. For any graph G of size m,

$$\sum_{v \in V(G)} \deg_G(v) = 2m.$$

Proof. When we count the degrees of each vertex in the set V(G), we account for each edge twice, once for each of its two incident vertices.

1.2 Common Types of Graphs

What happens if Bob and Clara are enemies? Fakebook is an additional social network much like Facebook except that it connects its users if they are enemies instead of friends. As shown in Figure 1.2, we can represent this Fakebook with a graph similarly to how we represented Facebook. We make each person a vertex of the graph and draw an edge between them if they are enemies.

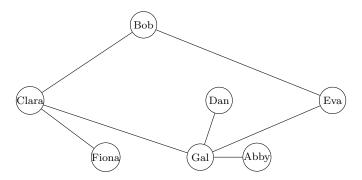


Figure 2: Graph of Fakebook Enemies

Say that you wanted to have a party and want to only invite people who do not hate one another. You would need to divide the people into groups so that no "hatred connections" exist in each group. In other words, you would need sets of people who are not adjacent in the Figure 2. Is this possible?

As it turns out, this is possible. Dividing the Fakebook enemies into two sets where Fiona, Gal, and Bob are in one set and Clara, Eva, Dan, and Abby are in the other yields a partition where no one hates another person in their respective set. Asking whether these sets exists is the equivalent of asking whether the Fakebook graph is bipartite.

Definition 1.5. A bipartite graph is a graph G whose vertex set V(G) can be partitioned into two subsets U and W where no edges of G exist between any two vertices in U or two vertices in W.

Definition 1.6. The partite sets of a graph G are the subsets of V(G) in which no vertex is adjacent to another vertex in its respective subset.

Example 1.5. The partite sets in the Fakebook example can be described as the sets of people who don't hate one another. As mentioned before, Fiona, Gal, and Bob would make up one partite set and Clara, Eva, Dan, and Abby would make up the other. Therefore, we can say the graph of these Fakebook enemies is a bipartite graph. The graph can be redrawn so these two sets can be seen more clearly. The two partite sets of people who do not hate one another are circled. Notice that in Figure 3 every edge of the graph connects a vertex of one partite set to the vertex of the other partite set.

The idea of a bipartite graph can be extended to include graphs which have multiple partite sets or multiple sets of people who don't hate one another.

Definition 1.7. A k-partite graph (or a k-colorable graph) is a graph with k partite sets.

If every person in each partite set hated every person from the other partite set, then we would have another special type of partite graph.

Definition 1.8. A complete bipartite graph is a graph G where every vertex of G is adjacent to every vertex of the other partite set. These graphs are denoted by $K_{U,W}$ with U and W being the cardinality of each partite set.

We could also imagine a situation where each person hates everyone else. A graph of this situation on Fakebook would be a graph in which every two vertices (people) are adjacent. In other words, each person hates everyone else in the group.

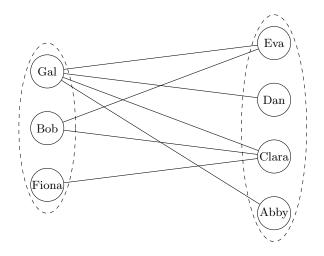


Figure 3: The Bipartite Graph of Fakebook Enemies

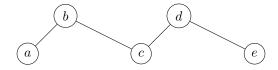
Definition 1.9. A complete graph is a graph G where every vertex $v \in V(G)$ is adjacent to every other vertex $u \in V(G)$. A complete graph is denoted by K_n where n is the order of graph.

Example 1.6. Here is the complete graph of order 4, K_4 :



Definition 1.10. A path P_n is a sequence of n vertices where every two consecutive vertices have an edge between them.

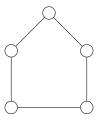
Example 1.7. The path $P = \{a, b, c, d, e\}$ with $E(P) = \{ab, bc, cd, de\}$ as shown.



Every vertex in path P has a degree of two except for the first and last vertices a and e which only have a degree of one. If an edge ae was added to P making every vertex have a degree of two, then the resulting graph would no longer be a path.

Definition 1.11. A cycle is a graph G similar to a path except the first and last vertices of the path form an edge in G. A cycle of order n is denoted by C_n and can be referred to as a n-gon.

Example 1.8. The cycle of order five C_5 or the 5-gon, which is simply a pentagon.



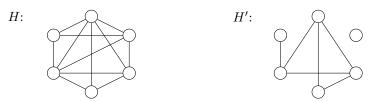
We might also want to have a new graph H that was a "piece" of another graph G, consisting of some of the vertices and edges of G.

Definition 1.12. A graph H = (V', E') is a *subgraph* of the graph G = (V, E) if $V' \subseteq V$, $E' \subseteq E$, and every edge $e \in E'$ has its endpoints in V'.

If a subgraph G' of a graph G is just the graph G with one edge $e \in E(G)$ or vertex $v \in V(G)$ missing, then G' can be denoted as G - e or G - v, respectively.

Example 1.9. For example, the path P from before could have a subgraph P' with vertex set $V(P') = \{a, b, c\}$ and edge set $E(P') = \{ab, bc\}$. The edge set of P' is a subset of P's edge set and the vertex set of P' is a subset of P's vertex set.

Example 1.10. Subgraphs can be formed from some or all of the vertices of a graph. The subgraph H' of H has the same vertex set as H but only some of the edges of H.



In most of the graphs we have described, every vertex is "connected" to every other vertex of the graph through some sequence of vertices of edges. Sometimes this connection is just an edge, as when two vertices are adjacent, and sometimes it is a series of adjacent vertices. In a connected graph, there is a path between every two vertices. However, there could be a graph that by simply removing one edge would disconnect one vertex or collection of vertices from the rest of the graph. This edge would "bridge" the two sections of the graph.

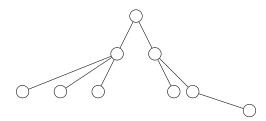
Definition 1.13. A graph G is connected if for any pair of vertices in V(G) there is a path between them. A graph G is disconnected if there exists a pair of vertices in G which is not connected by a path.

Definition 1.14. A bridge is an edge e of a graph G such the graph G - e is disconnected.

It is also possible to have a graph in which every edge is a bridge.

Definition 1.15. A tree is a connected graph where every edge is a bridge.

Example 1.11. The path P is an example of a tree since the removal of any edge would yield a disconnected graph. The graph shown below is also tree.



Similar to how the size and order were related in the fundamental theorem of graph theory, the size and order of a tree can be related.

Theorem 1.2. For every tree with order n and size m, we have m = n - 1.

Proof. The following proof uses induction on the order of a tree. The theorem holds for a base case of a tree of order 1, which has size 0. We then assume for a positive integer k that every tree of order k does, indeed, have size k-1. Let the tree T have order k+1. There a minimum of two vertices in T which have degree 1 and are the "end vertices" of the tree. ¹ By removing the one of these end vertices from T, we yield a tree T' of order k. We previously hypothesized that every tree of order k would have size k-1, so we know T' has size m=k-1. However, since the end vertex we removed from T had degree 1, we know that T has size m+1. Since T has size m+1 and order k+1, we have (m+1)=(k+1)-1, which simplifies to m=k-1.

¹We do not prove this fact, but the argument is simple.

2 Planar Graphs

Imagine there is a King who wants to divide his kingdom into five regions for his five sons. This King wants each son to build a palace in their region. Additionally, he wants each son's palace to be connected by road to every other palace and he wants no two connecting roads to cross. How could the sons build their palaces to fit the King's conditions?

This question naturally lends itself to being represented as a graph with the palaces as the vertices and connecting roads as edges. Since every palace must be connected to every other palace, the graph will be the complete graph, K_5 of order five.

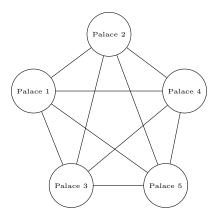


Figure 4: The Problem of the Five Palaces

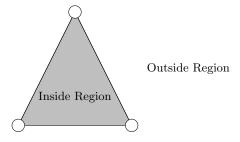
While the graph in Figure 4 shows roads connecting every two palaces, the roads cross multiple times, which is not allowed. The answer to whether there is any construction which fits the king's conditions lies in whether we can draw the graph of K_5 so that no two edges cross. This is the equivalent of asking if the graph K_5 is planar.

Definition 2.1. A planar graph is a graph which can be drawn in a plane so that no edges intersect.

By looking at some plane graphs, it is easy to see one property of planar graphs. The edges divide the graph into different sections or regions.

Definition 2.2. A region in a plane graph G is a section of the plane enclosed by edges of G.

Example 2.1. The plane graph of K_3 has two regions, the outside region and the region enclosed by three edges.



While it is easy to show some graphs, like K_3 , are planar since they are often drawn with no intersecting edges, proving a graph is not planar is requires showing that all possible drawings of the graph have intersecting edges. However, there are some properties of planar graphs which places conditions on what graphs are planar.

Theorem 2.1. (The Euler Identity) If G is a connected plane graph of order n, size m, and r regions, then n - m + r = 2.

Proof. First, we treat the cases of trees. For any given tree, m = n - 1 according to Theorem 1.2. As a tree does not close off any regions, r = 1 (representing the exterior), which results in n - m + r = n - (m - 1) + 1 = 2, as desired.

Let's assume for the sake of contradiction that there exists a graph G of the smallest size that does not satisfy the Euler Identity. Therefore in G, having size m, order n, and r regions, $n-m+r\neq 2$. Since the identity held for trees, we know G is not a tree and therefore must have at least one edge e where G-e is connected. The graph G-e is a planar graph and has size m-1. Since e is not a bridge, we also know that G-e has r-1 regions. We assumed that G was the graph of smallest size which the identity did not hold, so the Euler Identity must hold for G-e whose size is less than G's. Why? Because G-e is smaller, and we assumed G was the smallest counter example. Applying the identity yields n-(m-1)+(r-1)=2 which simplifies to n-m+r=2. This is a contradiction since we claimed the identity did not hold for a graph of order n, size m, and having r regions.

The Euler Identity could be used to derive many useful ideas in Graph Theory. How, for example, could we prove that a graph is *not* planar? One way we could tell if a graph is not planar is if it contains too many edges.

Theorem 2.2. For any planar graph G of size m and order $n \geq 3$, we have

$$m < 3n - 6$$
.

Proof. First, assuming that graph G is connected, we know that the inequality holds for $G = P_3$ with m = 3 and n = 3. So, without the loss of generality, we can assume that G has a size of $m \ge 3$. We proceed to draw graph G and denote each of its r regions by $R_1, R_2, R_3, ..., R_r$. Naturally, each region r is constructed from at least 3 edges. We denote the number of edges forming the boundary of each region R_i by m_i , where $m_i \ge 3$. We introduce a new variable M, where

$$M = \sum_{i=1}^{r} m_i \ge 3r.$$

Variable M counts each edge according to the number of times it was used as a boundary of a region. More specifically, M counts an edge once if it is a bridge and twice if said edge is not a bridge.

Since no edge is counted more than twice, $M \leq 2m$. By combining the two inequalities, it could then be said that $3r \leq M \leq 2m$. The expression simplifies into $3r \leq 2m$, which could be applied in the Euler Identity as follows:

$$6 = 3n - 3m + (3r) < 3n - 3m + (2m) = 3n - m.$$

Next, we assume that graph G is disconnected. Edges could then be added into graph G in order to make it connected, producing a new graph G' of order n and size m', where, necessarily, $m' \ge m$. Since G' is connected, we could assume that $m' \le 3n - 6$ from what we derived earlier. We can then say that m < 3n - 6.

Corrolary 2.1. If G is a graph of order $n \geq 3$ and size m where m > 3n - 6, then G is nonplanar.

П

Proof. This is the contrapositive of Theorem 2.2.

Theorem 2.2 is additionally useful as it could be used to derive additional conditions and identities of planar graphs, as it does in Corollary 2.1.

Corrolary 2.2. Every planar graph contains a vertex of degree 5 or less.

Proof. We will prove this corollary using contradiction. Then there exists a graph G of order n and size m in which every vertex has degree 6 or more. Since each vertex is adjacent to at least 6 other vertices by definition, $n \geq 7$. Integrating our conditions into the inequality presented in Theorem 1.1, we have

$$2m = \sum_{v \in G(V)} \deg_G(v) \ge 6n.$$

Thus $m \ge 3n > 3n - 6$, and graph G is nonplanar by Theorem 2.2.

Now that we have established some properties of planar graphs, we can return to our original question of whether the 5 princes can build roads between their palaces so that no roads cross. Whether this was possible depended on if the graph K_5 is planar. Using the corollary to Euler Identity we proved makes this proof simple.

Corrolary 2.3. The complete graph K_5 is nonplanar.

Proof. The complete graph K_5 has order n=5 and size m=10. Since 10>9=3(5)-6, by Corollary 2.2, the graph is non planar.

Having proven that K_5 is nonplanar, we have also proven that there is no way the princes can build connecting roads to all their palaces which do not cross. There are other problems which can be solved using the properties of planar graphs like the three utilities problem. Say there are three utility stations, gas, water, and electricity, and three houses, each of which need access to all three utilities. The utility company needs to connect each house each utility by a pipeline, however, they must build the pipelines in such a way so no two pipelines cross. Like the previous problem involving the princes, we can model this problem with a graph, as shown in Figure 5.

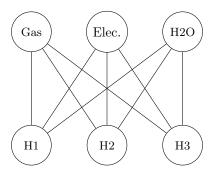


Figure 5: The three utilities problem

While this graph does not fit the condition of the utility company, we can make some useful observations. First, we recognize this graph as the complete bipartite graph $K_{3,3}$ with the utility stations as one partite set and the houses as the other. Much like the previous example, whether or not the pipelines to each house can be constructed in such a way so no two cross depends on whether the graph $K_{3,3}$ is planar. Since $K_{3,3}$ has order n=6 and size m=6, the inequality we used to prove K_5 will not be useful. Instead, we will begin by assuming that $K_{3,3}$ is planar and then arrive at a contradiction.

Theorem 2.3. The graph $K_{3,3}$ is nonplanar.

Proof. For the sake of contradiction, assume the graph $K_{3,3}$ is planar, therefore we can apply the Euler Identity to find the number of regions in the graph. Using n=6 and m=9, we find that the number of regions r=5. Since every region is enclosed by some edges, we can now find a bound for the minimum number of edges $K_{3,3}$ must have if it is planar. Since the graph is bipartite and there are no edges within each partite set, we know that there can be no regions enclosed by only 3 edges. Therefore each of the 5 regions is enclosed by at least 4 edges. We also know that since $K_{3,3}$ is a complete bipartite graph there are no bridges which means every edge encloses 2 regions. We can then conclude that $2m \ge 4r$ and therefore if $K_{3,3}$ is planar, then $m \ge 10$. This is a contradiction since we know that the size of $K_{3,3}$ is 9. We can then conclude that $K_{3,3}$ cannot be planar.

3 Graph Colorings

3.1 Introduction to Vertex Coloring

Here is a blank map of the New England states. Say we wanted to color it so that each state is a different color than the states it borders. How could we go about doing this? How many colors would we need?



Figure 6: Blank Map of the New England States

To help think about these questions, we can represent Figure 6 as a graph with the vertices representing states and edges representing the shared borders of states. This graphical representation, as shown in Figure 7, is known as the dual of the map.

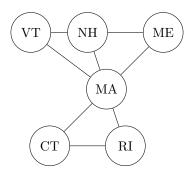


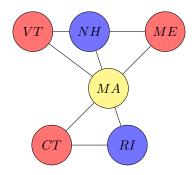
Figure 7: Dual of the Map

Now, instead of thinking about coloring regions on a map, we can think about coloring the vertices of a graph. Since each edge in the graph represents a border between states, coloring the map so no two bordering states are the same color is the same as coloring each vertex in the graph a color different from the colors of its adjacent vertices. More simply, in a *coloring* of our graph, no two adjacent vertexes can be the same color. It is easy to imagine many ways which we could go about coloring the vertices so we satisfy this condition. We could use six colors and assign each vertex a color, we could use five colors assign VT and RI the same color, and so on. However, we want to consider the coloring of the graph that uses the fewest number of different colors. What is the minimum number of colors we can use to color the vertices of the graph so no adjacent vertices are the same color? We can expand on this question by defining some essential characteristics for vertex coloring.

Definition 3.1. For any graph G, a certain coloring of the graph using k colors is defined as a k-coloring.

Definition 3.2. The minimum number of colors needed to properly color G is denoted by $\chi(G)$. A graph G is said to be k-chromatic or k-colorable if $\chi(G) = k$.

Example 3.1. The graph G of map example provided above is 3-chromatic ($\chi(G)=3$). Although the maximum k-coloring of the graph contains 6 colors (for the size of V(G) is 6), the minimum number of colors which could be used to represent all of the vertices of graph G without having any two adjacent vertices be of the same color is 3. We use blue, red, and yellow to present one 3-coloring of G below.



In the example above, you might notice that certain groups of vertices necessarily require to have all different colors since they are all adjacent. For example, in order to color VT, NH, and MA, it requires three colors, even if these vertices were their own isolated graph. This observation relates to the idea of a complete graph, where every two vertices are adjacent. If we wanted to color a complete graph K_n of order n, then it would require n colors for the n vertices which are all adjacent to one another. However, since the dual graph of maps are always planar, any dual graph of order n > 4 will not be complete. In order to use this observation on the coloring of complete graphs when thinking about planar graphs, we can modify it slightly.

Definition 3.3. A clique of a graph G is a complete subgraph of G. The clique number, or the clique number $\omega(G)$ of G is the order of the largest clique in G.

If we consider a clique as a set of vertices which all require different colors, then we can also define a set of vertices which are all colored the same. In this set, no two vertices could be adjacent.

Definition 3.4. An independent set is a subset of the vertices of a graph G of which no two vertices in the set are adjacent in G. The independence number $\alpha(G)$ of a graph G is the cardinality of the maximum independent set.

Theorem 3.1. For every graph G of order n, we have

$$\chi(G) \ge \omega(G)$$
 and $\chi(G) \ge \frac{n}{\alpha(G)}$

Proof. Let the subgraph H of order $\omega(G)$ be a clique of G. H is complete which means $\chi(H) = \omega(G)$, since every vertex must be a different color. Since H is a subgraph of G, the coloring G requires at least as many colors as H, or $\chi(G) \geq \chi(H) = \omega(G)$. Therefore $\chi(G) = \omega(G)$.

Furthermore, if a coloring of G requires $\chi(G)$ colors, then V(G) can be divided into $\chi(G)$ subsets where in every subset all the vertices are the same color and therefore not adjacent. The sum of the cardinalities of each subset is the order of G, n. Since the maximum cardinality of these subset is $\alpha(G)$, then the product of the number of subsets, $\chi(G)$ and the maximum cardinality of them will be at least the order of G. Hence

$$\alpha(G)\chi(G) \ge n$$

and therefore

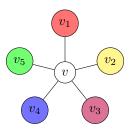
$$\chi(G) \ge \frac{n}{\alpha(G)}.$$

3.2 Four and Five Color Theorems

So far, we have have only placed lower bounds on the chromatic number of a graph. These bounds have been based on other properties of the graph such as its independence number and omega number. We now might want to ask if there is a minimum number of colors which *any* planar graph can be colored with regardless of its omega number or independence number. As it turns out, every planar graph is 4-colorable. However, to prove this is extremely difficult and requires a supercomputer. Fortunately, there is a proof that every planar graph is 5-colorable which can be proven much easier.

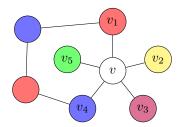
Theorem 3.2. Every planar graph is 5-colorable.

Proof. We prove this theorem by contradiction. For the sake of contradiction, first assume there exists a graph G that is not 5-colorable. We pick G to be the not-5-colorable graph of smallest order. The graph then necessarily has an order of 6 or more. According to Corollary 2.2, the minimum degree $\delta(G)$ in a planar graph G is always less than or equal to 5. Pick v to have that smallest degree. Since G was stated to be of minimum order, the removal of v, which produces the new graph G - v, is necessarily 5-colorable.



There are now two cases which we can observe. If all the neighbors of v in G-v can be colored in 4 or less colors, then the theorem holds as there is at least one color left for vertex v, which proves that G is 5-colorable. The second possible case is more challenging to prove, however.

The basic graph that we observed has a vertex v of degree 5 or more. If the vertices in G-v are colored in 5 or more colors, we must prove that some of the colors can be changed by altering the colors assigned to the existing vertices to produce a graph G that is 5-colorable. We look at two vertices, v_1 and v_4 , for example. To produce a 5-colorable graph G, as said earlier, we would need to color one of them the color of the other. If they are connected by a path S of alternating colors, say red and blue, then this would seem impossible.



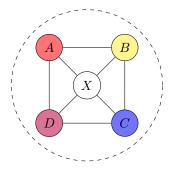
However, in such a case, a similar alternating path S' connecting two other vertices, like v_2 and v_5 , would not be able to exist as the graph is planar by definition. We could then color one of the two vertices the color of the other and simply switch the color of any chain of vertices that may follow it. Whether v_2 , v_5 , or any other vertex of G-v that is adjacent to v in G is colored to match another vertex in the set, the number of colors is reduced by 1. This allows v to be colored in a way that produces a 5-colorable graph G.

Theorem 3.3. Every planar graph is 4-colorable.

Francis Guthrie began the discussion around the topic of the theorem in 1852. Later, in 1879, Alfred Bay Kempe published a proof for the Four Color Theorem, involving what had later been named *Kempe Chains*. His proof goes as follows:

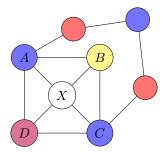
Proof. Assume there exists a region X which is surrounded by regions colored using $c \le 4$ colors. Quite obviously, if c > 4, then there is necessarily a color left to color region X and the Four Color Theorem holds. Therefore, the latter case is of our concern in this proof.

This situation could be represented with the following graph G in which regions that share a border are adjacent vertices:



Although the number of regions neighboring X is not specified as four, we present the base case in which the regions directly bordering X are all colored differently. Their surroundings are unknown and are represented by the dotted line surrounding the graph.

There are now two possibilities. In the first, we assume that there is no alternatively colored red and blue chain that connects A and C. In that case, we switch the color of A to blue and continue altering the color of all of the vertices in the chain that connects A and C.



Now both A and C are colored in blue, allowing for X to be colored in red. In the second possible case in this problem, there does exist an alternatively colored red and blue chain that connects A and C.

In this case, similarly to the one presented in Theorem 3.2, there could not exist a yellow and purple alternating chain connecting B and D as the graph is planar by definition. Therefore, either B or D could be colored to match the other vertex, and any chain that may follow said vertex would be colored in the opposite color to adapt to the change. There would then be an additional remaining color for X, making G 4-colorable.

Kempe's proof is flawed, however, as it overlooks one essential case of the situation that it presents, in which X is surrounded by exactly 5 regions. The regions could then use the four colors essential for the theorem.

When mathematicians attempted to solve said case over the years, they faced severe challenges. The proof simply continued to be divided into many different parts, complicating its solution further.

In 1976, Kenneth Appel and Wolfgang Haken found and tested a set of 1936 reducible configurations.

Definition 3.5. A reducible configuration is an arrangement of regions that could not occur in a minimum counterexample.

These configurations were then tested in three computers throughout a 1200 hour period of computer time in total. Mathematicians were, however, skeptical of the proof and its validity as a mathematical proof.

Finally, the Four Color Theorem had been solved through a simpler, yet computer-based, solution of 633 reducible configurations was published in 1933 by Neil Robertson, Daniel P. Sanders, Paul Seymour, and Robin Thomas.

12