More on Planar Graphs: The Four-Color Theorem

Kevin Angelo 13517086
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13517086@std.stei.itb.ac.id

***Abstract*—Mathematics has long been with men; people coming out with theories, others coming out with problems, and the rest of us struggled with mathematical proofs of those theories mentioned. Although it is now completely rational to utilize computers and technology to prove a certain hypothesis, mathematicians in the past refused doing so. Most of them did not even approve proofs done by computers as they did not give us full understanding of the said theories. This paper will mainly cover The Four-Color Theorem. Being the first ever computer-assisted proof in mathematics, it was controversial at the time.**

***Keywords*—four-color theorem, planar graphs, Kempe's fallacious proof, Appel-Haken proof**

I. Introduction

The Four-Color Theorem stated that, given any separation of a plane into several contiguous regions, producing a map, no more than four colors are required to color the map such that no neighboring regions have the same color. It applies not only on the world maps, but all maps, with tens or even millions of countries/regions of all shapes and sizes.

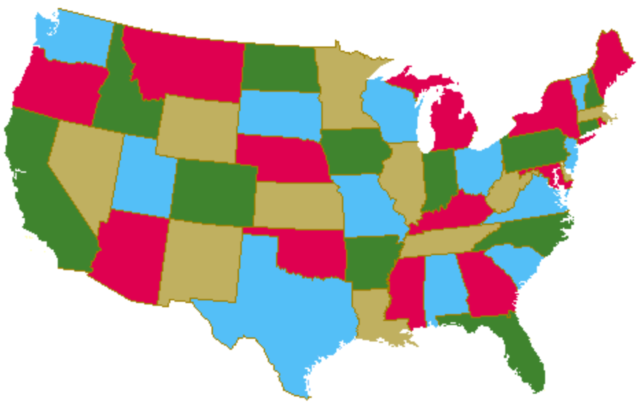


Figure 1. Map of the United States, perfectly colored with only four colors
(http://people.math.gatech.edu/~thomas/FC/fourcolor.html, accessed on December 9, 2018 18:12 GMT+7)

The Four-Color Problem first came out in 1852 when Francis Guthrie, a South African mathematician tried to color a map of the countries on England. He then noticed that it only took him four colors to correctly color the map. By correctly it means no two adjacent regions of the maps are of the same color, otherwise it would be confusing for people reading the map.

He asked his brother Frederick Guthrie if it was true that any map can be colored in four colors such that two regions sharing the same boundary segment have different colors, who then communicated the speculation with his lecturer in University College London, Augustus De Morgan. De Morgan then published the problem. Being a relatively simple problem, it quickly gained public attention.

There were several failed attempts at solving the problems. One was given by Alfred B. Kempe in 1879, followed by Peter G. Tait in 1880. In 1992, an attempt came from George David Birkhoff whose work helped Philip Franklin in proving the four-color conjecture for maps with at most 25 regions in 1922. Until finally, Kennel Appel and Wolfgang Haken at University of Illinois announced that they had proved the theorem on June 21, 1976, making The Four-Color Theorem the first ever proof in mathematics to be done using a computer.

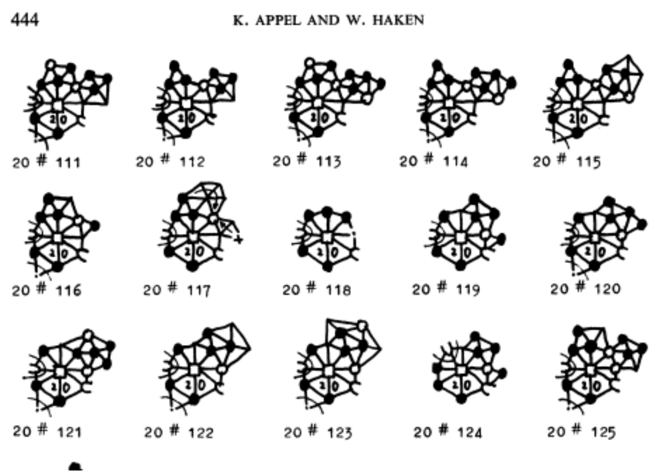


Figure 2. Few among the 1976 network configurations tested by Appel-Haken
(Appel, K. and Haken, W. (1989). Every Planar Map Is Four Colorable: Part I. accessed at https://projecteuclid.org/download/pdf_1/euclid.ijm/1256049011, on December 9, 2018 19:48 GMT+7)

Appel-Haken made an enormous list of maps and checked each of them using a computer. The Appel-Haken proof consisted of massive test cases, around 1936 configurations of unavoidable set (which was then narrowed down to 1482 configurations), and they proved that all of them can be colored, or recolored, using no more than four colors. Many mathematicians of the time rejected the proof, as massive case checking, despite being valid, did not grant better understanding of the problem.

II. Graph Theory and Terminology

A. Graph

A graph G = (V, E) consists of a non-empty set of vertices V and a set of edges E. Each edge has one or a pair of vertices connected with it, called endpoints.

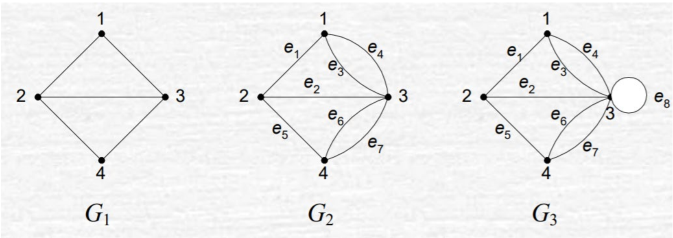


Figure 3. Graphs
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 20:21 GMT+7)

For instance, graph G1 in Figure 3 can be considered as G = (V, E) where V = {1, 2, 3, 4} and E = {{1, 2}, {1, 3}, {2, 3}, {2, 4}, {3, 4}} and G2 can also be considered as G = (V, E) with the same V as G1. However, G2 differs from G1 in terms of its sets of edges E, where E in G2 has 2 sets of parallel or multiple edges, one being e3 & e4, another being e6 & e7. Multiple edges are two different edges connecting the same two vertices; in this case, edges e3 and e4 both connect vertices 1 and 3, while edges e6 and e7 both connect vertices 3 and 4. Edge e8 in G3 is called a loop, as it starts from a vertex (vertex 3) and goes back to the same vertex (back to vertex 3). We can also say that a loop edge connects a vertex to itself.

Based on the presence of parallel edges and loop edges, there are different types of graphs:

a. Simple Graph, which has no parallel edges or loop edges. Graph G1 in Figure 3 is a simple graph.
b. Multi-Graph, which is a graph having at least one loop or parallel edges. Graph G2 and G3 in Figure 3 are examples of multi-graph.

A graph may have 'direction' in its edges, meaning it has sets of edges made of ordered vertex pair. Such graph is called a directed graph. If it has sets of edges made of unordered vertex pair on the other hand, it is an undirected graph. Graphs G1, G2, G3 in Figure 3 are all undirected graphs as they do not have direction associated with their edges.

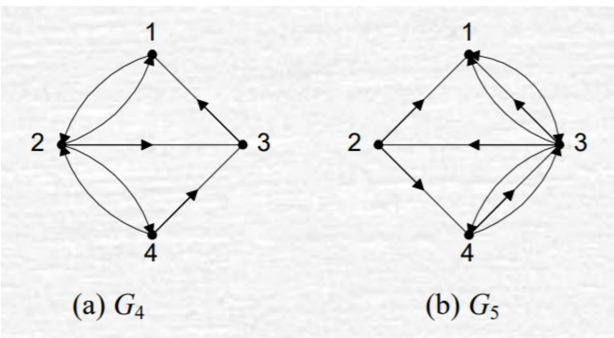


Figure 4. Examples of directed graphs
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 20:44 GMT+7)

Graph G4 and G5 from Figure 4 has sets of edges made of ordered pair of vertex, which can be denoted as G = (V, E) where E = (sets of {a, b}), implying a 'direction' from vertex a to vertex b, and further implying that {a, b} ≠ {b, a}.

B. Graph Terminology

B.1. Adjacency and Incidence

Two vertices a and b are said to be adjacent if a and be are both the endpoints of the same edge. For instance, vertex 1 and 2 in graph G1 from Figure 3 are adjacent. Vertex 1 and 4 from the same graph, however, are not adjacent.

An edge is incident on a vertex if said vertex is an endpoint of said edge. Thus, for each e = {a, b}, e incident on a, or, e is incident on b. e = {1, 3} in graph G1 from Figure 3 is incident on vertex 1 and vertex 3.

B.2. Isolated Vertex and Null Graph

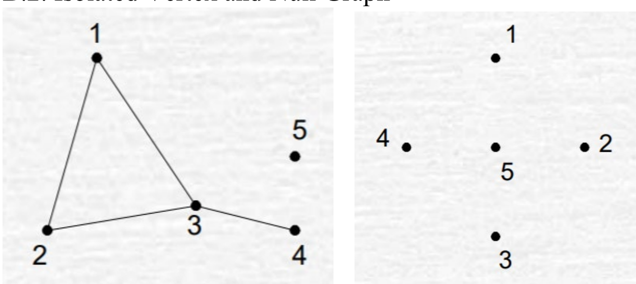


Figure 5. Isolated Vertex (Left) and Null Graph (Right)
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 21:14 GMT+7)

A vertex is isolated when there are no incident edge on it, or when it is not adjacent with another vertex. Vertex 5 from the left graph on Figure 5 is an isolated vertex as it is not connected to any other vertices.

A null graph is a graph with null set of edges, meaning it has no edges. The right graph on Figure 5 is an example of null graph.

B.3. Degree

The degree of a vertex v is the amount of edges incident with v, noted as d(v), d(v) ϵ Z. Looking at the left graph in Figure 5: d(1) = 2, d(2) = 2, d(3) = 3, d(4) = 1, and d(5) = 0.

B.4. Path and Cycle

A path with length n from vertex a to vertex b in a graph G is a sequence of alternating vertices and edges such that each successive vertex is connected by the edge. Frequently, only the vertices are listed. Path length n is the amount of edges the path contains. The path from vertex 2 to vertex 4 in the left graph from Figure 5 can be noted as P = 2, 3, 4, with its path length n = 2.

A cycle or a circuit is a path that initiates and terminates at the same vertex. Looking at the left graph in Figure 5: P = 1, 2, 3, 1 is a cycle, with its cycle length n = 3.

B.5. Connectivity

Vertex a and vertex b in a graph G is said to be connected if there is a valid path from vertex a to vertex b. Vertex 1 in the left graph from Figure 5 is connected to vertex 4, but not to vertex 5.

A graph G is a connected graph if each vertex in G is connected to all other vertices in said graph. Graph G1 in Figure 3 is a connected graph, while the left graph in Figure 5 is not a connected graph, or a disconnected graph.

B.6. Subgraphs and Spanning Subgraphs

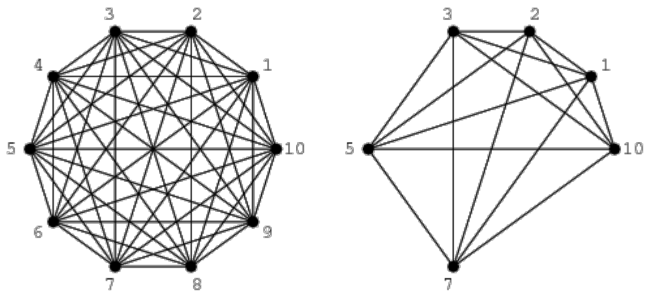


Figure 6. Subgraph (Right) of a Graph (Left)
(http://mathworld.wolfram.com/Vertex-InducedSubgraph.html, accessed on December 9, 2018 21:39 GMT+7)

Say we have a graph G = (V, E). G1 = (V1, E1) is a subgraph of graph G if V1 ⊆ V and E1 ⊆ E. In Figure 6, the graph on the right is a subgraph of the graph on the left.

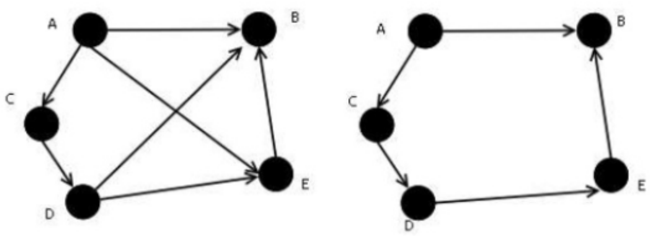


Figure 7. Spanning Subgraph (Right) of a Graph (Left)
(https://www.slideshare.net/Tech_MX/graph-theory-1, accessed on December 9, 2018 21:46 GMT+7)

A subgraph G1 = (V1, E1) is a spanning subgraph of graph G = (V, E) if G1 contains all the vertices of G, or V1 = V. In Figure 7, the subgraph on the right is a spanning subgraph of the graph on the left.

B.7. Weighted Graphs

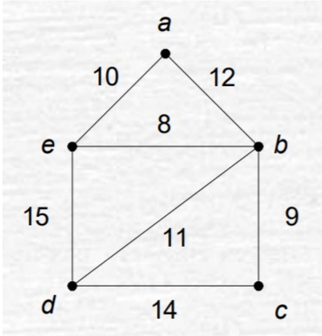


Figure 8. Weighted Graph
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 22:21 GMT+7)

Weighted graph is a graph in which each edge is given a numerical weight. A common operation on weighted graphs is the shortest-path computation; deciding paths from vertex a to vertex b such that the sum of the weight of the path is minimal. Graph in Figure 8 is a weighted graph, with each number beside each edge representing the respective edges' values.

C. Planar Graph

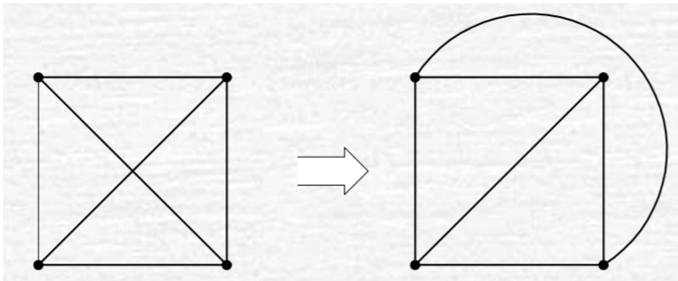


Figure 9. A Planar Graph
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 22:41 GMT+7)

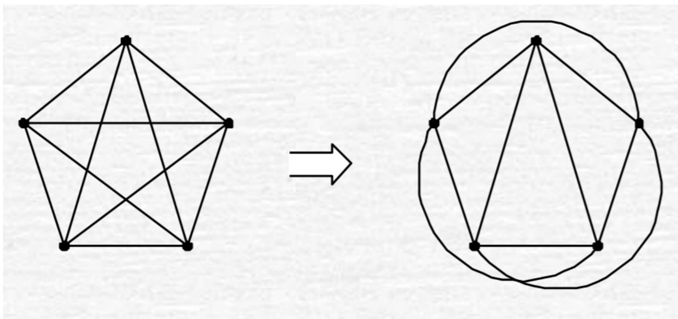


Figure 10. A Non-Planar Graph
(http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf, accessed on December 9, 2018 22:41 GMT+7)

Graph G is a planar graph if it can be drawn in a plane such that no edges are crossed with each other. Else, it is called a non-planar graph. The graph in Figure 9 is a planar graph as it can be redrawn and have its edges restructured such that there are no crossing edges. The graph in Figure 10, however, is a non-planar graph as there is no such configuration that allows non-crossing edges in said graph when drawn in a plane.

The concept of planar graphs is widely used in other discipline, mainly engineering. In the design of an Integrated Circuit (IC) for instance, no crossing edges may be present within the board as such configuration might lead to circuit malfunction.

All separation of planes into several regions, including, but not restricted to, maps, can be converted into their corresponding planar graphs, with each vertex in the graph representing regions, and each edge in the graph representing respective regions' borders.

Suppose we have this map of Austria:

Figure 11. Map of Austria
(https://commons.wikimedia.org/wiki/File:Austria_states_english.png, accessed on December 9, 2018 23:01 GMT+7)

We can convert the map in Figure 11 into a planar graph by putting a dot, representing the vertices of the graph, and drawing a line connecting two respective vertices whenever two regions border each other, representing the edges of the graph.

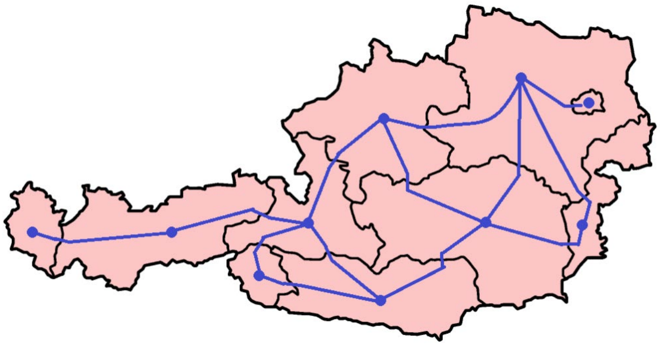


Figure 12. Planar graph converted from the map in Figure 11
(https://math.stackexchange.com/questions/2206138/proof-that-every-map-produces-a-planar-graph-four-colour-theorem, accessed on December 9, 2018 23:01 GMT+7)

It also works the other way around; for every planar graph, there is at least one map-like representation of said graph on a plane.

III. Approaches on The Four-Color Theorem: Proofs and Boundaries

A. Problem Boundaries

The statement of The Four-Color Theorem "that given any separation of a plane into contiguous regions, called a map, the regions can be colored using at most four colors so that no two adjacent regions have the same color" has several boundaries in its interpretation.

First, points that belong to three or more countries must be ignored. Second, no two disconnected regions representing a single country (which will then require same coloring) can exist. Every country has to be a connected region, or contiguous, meaning there is no obligation that two separated regions have the same colors.

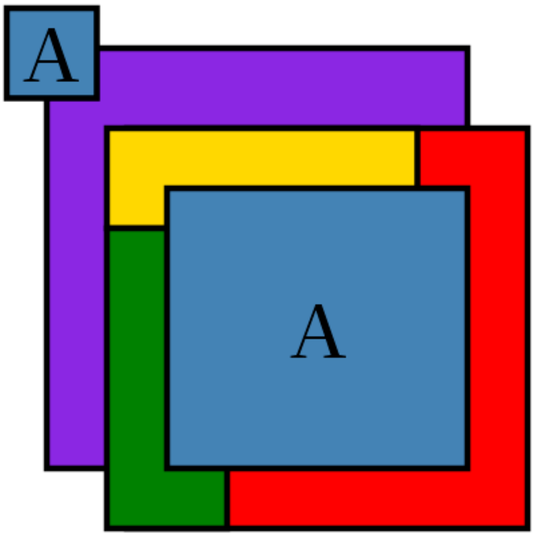


Figure 13. Example of an invalid test case for The Four-Color Theorem
(https://en.wikipedia.org/wiki/Four_color_theorem#CITEREFHudson2003, accessed on December 10, 2018 05:03 GMT+&7)

In the map in Figure 13, there are two regions representing a single segment, both labeled A, and both must be of the same color. Such map requires five colors, thus cannot exist in the discussion of The Four-Color Theorem.

B. Kempe Chain Method

Albert B. Kempe wrote his proof of The Four-Color Theorem in 1879, which was later shown incorrect by Percy Heawood in 1890. Despite being inaccurate, Kempe Chain was later used as an extremely important tool in proving The Five-Color Theorem.

The formal definition of Kempe's Chain theory goes as following:

Definition 1 (C1C2-Kempe chain). Let G be a planar graph whose vertices have been properly four-colored and suppose v ∈ V(G) is colored C1. The C1C2-Kempe chain containing v is the maximal connected component of G that contains v and contains only vertices colored C1 or C2. [Gethner and Springer 2003].

Definition 2 (C1C2-Kempe chain switch). Let K be a C1C2-Kempe chain. A C1C2-Kempe chain switch interchanges all values of C1 and C2 in K. [Gethner and Springer 2003].

Kempe attempted to show that a graph with n number of vertices can be four-colored and assumed that all graphs with less than n vertices can too, be four-colored. With Euler's Theorem for planar graphs, he showed that a graph in which all its vertices have degree 6 or higher is never a planar graph. This leads him to prove that any maximal planar graph – a planar graph in which no new edges can be added without disturbing its planarity – must have at least one node with degree 5 or less.

"If we denote the number of nodes, edges, and faces (i.e., the bounded regions) of a planar graph by V, E, and F respectively, then Euler's theorem for a plane (or a sphere) is $V - E + F = 2$ (the outside face counts too). Each face of a maximal planar graph is bounded by three edges, and each edge is on the boundary of two faces, so we have $F = 2E/3$. Euler's formula for a complete planar graph then becomes simply $E = 3V - 6$. The degree of a node is the number of edges incident to it. Now each edge is connected to two nodes, so the sum of the degrees of all the nodes $2E = 6V - 12$, and hence the average degree per node is $6 - 12/V < 6$ (assuming the graph is finite). This implies that at least one vertex has degree five or less." Retrieved from http://web.stonehill.edu/compsci/lc/four-color/four-color.htm.

Suppose we have a map in which each region is colored with one of our chosen four colors; in this case, say we have red, blue, green, and yellow. Each region of said map is colored with one of the colors, except for one single region, say region R. If this region R is not surrounded by four other regions each with different colors, then there is a color left for R. Thus, suppose region R is surrounded by regions of all four colors, namely A, B, C, and D, each with their colors in order: red, blue, green, and yellow.

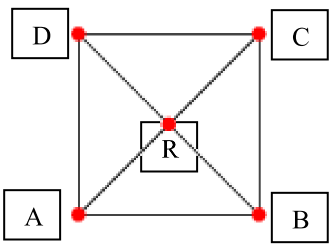


Figure 14. Region R surrounded by four regions, each with different colors
(http://mathworld.wolfram.com/WheelGraph.html, accessed on December 10, 2018 04:02 GMT+7)

There are two possible scenarios to consider:

i. Regions A and C are not adjacent, meaning there is no chain or edge connecting A to C.
If this is the case, then change A to green. Then interchange the color of the red/green regions in the chain connecting A. There is now no red region adjacent with R. R can now be colored red.

ii. Regions A and C are adjacent, meaning there is a chain or edge connecting A to C.
If this is the case, then there must be no chain connecting region B and D, else the graph is no longer a planar graph, and non-planar graphs cannot be represented on a plane. The property in point (i) becomes true for B and D, and we change the colors accordingly.

This method applies to all regions in given map, with vertices having degree 4 or less.

For graph with vertices having degree 5, we use the same approach as the degree 4 situation. Our uncolored region R is now surrounded by five regions.

Figure 15. Case i. (Left) and Case ii. (Right)
(http://web.stonehill.edu/compsci/lc/four-color/four-color.htm, accessed on December 10, 2018 06:03 GMT+7)

There are another two scenarios to consider:

1. Two nodes with the same color are next to each other as in the left part of Figure 15, where red colored regions are next to each other.
If this is the case, the argument is similar with the degree 4 situation, since the two red regions can be treated as a single region, producing a semi four-degree region R.
2. Two nodes with the same color are not next to each other as in the right part of Figure 15.

The main idea in proving case (ii) is to create two Kempe chains: (a) starting from the region colored blue and following all edges connecting vertices colored blue or yellow. As before, if this chain does not contain the vertex adjacent to R that is colored yellow, we then toggle the colors of the chain, resulting in the top region (previously blue) being colored yellow. There is now no blue region adjacent to R. R can now be colored blue. However, if the chain contains the yellow-colored vertex, we move on to the other chain; (b) starting from the region colored blue and following all edges connecting vertices colored blue or green. If this chain does not contain the vertex adjacent to R that is colored green, we can then toggle the colors of this chain, once again allowing us to color R blue. If this second chain contains the green-colored vertex, we must reject this chain as well. We then can no longer be using Kempe Chain starting from the blue region.

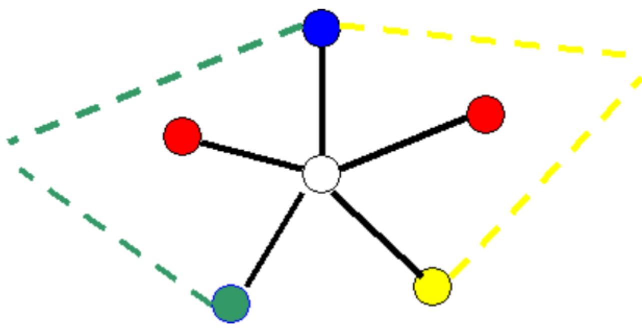


Figure 16. Kempe Chain
(http://web.stonehill.edu/compsci/lc/four-color/four-color.htm, accessed on December 10, 2018 06:27 GMT+7)

Kempe's solution in this case is to start at each of the two vertices colored red and create two Kempe chains, one with colors red and green, and the other with colors red and yellow. From the vertex colored red that is surrounded by the blue-yellow Kempe chain (the red on the right), he creates a

Kempe chain with colors red and green. From the vertex colored red that is surrounded by the blue-green Kempe chain (the red on the left), he creates a Kempe chain with colors red and yellow. The new Kempe chain with colors red and green cannot reach the vertex adjacent to R colored green, so the colors can be toggled, and red becomes green. The new Kempe chain with colors red and yellow cannot reach the vertex adjacent to R colored yellow, so red becomes yellow. The red vertex on the right is colored green and the red vertex on the left is colored yellow. This leaves color red free for R, which is now adjacent to colors yellow-green-blue-yellow-green (in counter-clockwise order around R).

This implies that Kempe uses strong induction in his proof, where he assumed that the statement holds true for all values preceding n.

C. The Flaw in Kempe's Proof

Percy J. Heawood exposed a flaw in Kempe's proof of the Four-Color Theorem in 1890. The flaw lies under the following statements from Kempe's proof: "*The new Kempe chain with colors red and green cannot reach the vertex adjacent to R colored green, so the colors can be toggled, and red becomes green. The new Kempe chain with colors red and yellow cannot reach the vertex adjacent to R colored yellow, so red becomes yellow.*"

The problem is, it is possible to have edges between yellow vertices in the red-yellow Kempe Chain and the green vertices in the red-green Kempe Chain. If we change the colors based on Kempe Chain Method, both yellow and blue vertices, which are adjacent, will be changed to red, resulting in one or more adjacent pair of vertices having the same color. It turns out that a green vertex surrounded by the blue-yellow Kempe Chain can be adjacent to a yellow vertex surrounded by the blue-green Kempe Chain, as shown in Figure 17.

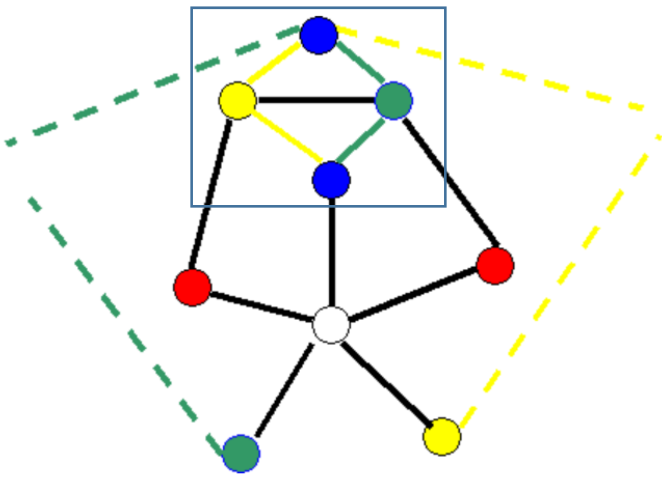


Figure 17. The flaw lies on the squared region
(http://web.stonehill.edu/compsci/lc/four-color/four-color.htm, accessed on December 10, 2018 06:54 GMT+7)

Kempe Chain Method does not go wasted, however. Utilizing Kempe's Chain Theory, Heawood proved a similar but less powerful Five-Color Theorem.

D. Appel-Haken Proof

The Four-Color Theorem was finally proven by Kennel Appel and Wolfgang Haken at University of Illinois on June 21, 1976, basing their methods on reducibility using Kempe Chain. They stated that if The Four-Color Theorem were false, then there would be at least one map with the smallest number of regions which require five colors. They proceeded to prove this through: (a) the making of a very long list of networks called the *unavoidable set*, a set of configurations such that every map that satisfies necessary conditions for being a non-four-colorable triangulation (all faces bounded by three edges) must have at least one configuration from said set; (b) the concept of *reducible configuration*, an arrangement of maps that cannot occur in a minimal counter-example. If a map contains a reducible configuration, then it can be reduced into smaller maps. If this smaller map is four-colorable, then the original map is, too. This implies that if the original map is not four-colorable, then the smaller map is not, either, and said original map is not minimal.

Appel and Haken later found an unavoidable set of reducible configurations, proving that a minimal counter-example for the four-color conjecture can never exist. Their proof reduced the number of possible maps down to 1936 configurations (which was then reduced to 1482), and such massive number of cases was checked one by one using a computer. It took over a thousand hours to complete the checking. The reducibility part of the proof was then double checked with different programs, again with different computers. The unavoidability part, however, had to checked manually.

However, there was a flaw in this proof, specifically in the discharging procedure, proposed by Ulrich Schmidt. This was later corrected in 1989, the same year in which they published their greatest work, *Every Planar Map is Four-Colorable*, a book containing a complete and detailed proof of the theorem, including the explanation of Schmidt's discovery and several errors found by others.

IV. Conclusion

The Four-Color Theorem may not be one of the best proofs in mathematics, but it is now widely accepted by public and people are getting advantages from it. One direct application of The Four-Color Theorem is in mobile phone towers. Towers cover certain areas with some overlaps, thus they cannot all transmit on the same frequency. The trick here is to give them all different frequencies. In deciding the minimum possible number of frequencies, one can utilize The Four-Color Theorem. The Four-Color Theorem is not just about graphs and maps, but it can also be used in modeling real-life problems, like expressing some binary relation among objects and test scheduling.

V. Acknowledgment

The author would firstly like to thank the great lecturers of the course IF2120 Matematika Diskrit, especially to Dra. Harlili M.Sc for the great lectures and her strong dedication. A great gratitude towards previous authors whose

work was cited in this paper, towards the community of mathematics, and towards people's dedication to the development of mathematics and science in general.

REFERENCES

[1] K. Appel and W. Haken (1989), *Every Planar Map is Four-Colorable*, Contemporary Mathematics, 98, With the collaboration of J. Koch., Providence, RI: American Mathematical Society.

[2] R. Wilson (2014) [2002], *Four Colors Suffice*, Princeton Science Library, Princeton, NJ: Princeton University Press.

[3] K. Appel and W. Haken (1977), *Every Planar Map is Four-Colorable, Part I: Discharging*, Illinois J. Math. 21

[4] R. Munir. (2015). *Graf*. Retrieved from http://informatika.stei.itb.ac.id/~rinaldi.munir/Matdis/2015-2016/Graf%20(2015).pdf

[5] [Numberphile]. (2017, March 20). *The Four Color Map Theorem – Numberphile*. Retrieved from https://www.youtube.com/watch?v=NgbK43jB4rQ

[6] Prof. R. Bravaco and S. Simonson. Stonehill College. Retrieved from http://web.stonehill.edu/compsci/lc/four-color/four-color.htm

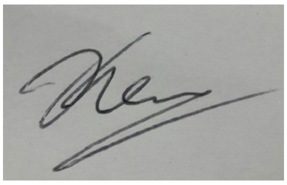
[7] E. Gethner, B. Kallichanda, *et al.* (2009). *How false is Kempe's proof of the Four Color Theorem? Part II*. Vol. 2, No. 3. Mathematical Science Publishers.

[8] J. J. O'Connor and E. F. Robertson. (1996). *The four colour theorem*. Retrieved from http://www-groups.dcs.st-and.ac.uk/history/HistTopics/The_four_colour_theorem.html

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 10 Desember 2018



Kevin Angelo 13517086