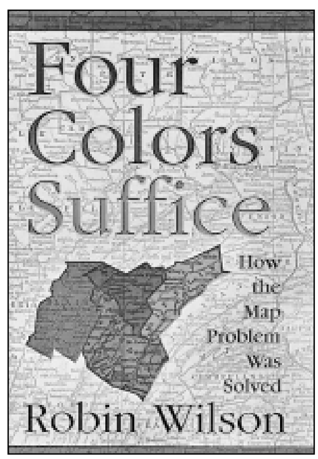
Four Colors Suffice: How the Map Problem Was Solved

Reviewed by Bjarne Toft

Four Colors Suffice: How the Map Problem Was Solved
Robin Wilson
Princeton University Press, 2002
288 pages, Cloth, $24.95
ISBN 0-691-11533-8

This is an attractive book telling the story of the Four Color Problem and Four Color Theorem. It has nice typography, contains numerous illustrations (black and white halftones), and sports a fine binding and beautiful dustcover. The elegant appearance is matched by the general exposition of the book. Robin Wilson explains in simple terms the mathematical ideas involved in the solution of the Four Color Problem, explores history, tells anecdotes, and exhibits controversies. The reader thus gets a picture not only of the mathematics itself but also of the sociology of mathematics—mathematicians' successes and defeats, and how they collaborate and compete. Philosophical discussions on computer-aided proofs in mathematics are also presented. The main text itself is not hampered by footnotes or references; these are given towards the end in an 18-page chapter "Notes and References". The interested reader will find here much additional material to pursue, should he want to. All in all, the book is a pleasure to read or browse through.

The book focuses mainly on one line of investigation, namely the one leading to the solution of the problem. This line was emphasized by Heinrich Heesch (1906–1995) from the early 1930s, based on work by Paul Wernicke (1866–19??) and George David Birkhoff (1884–1944). The idea of Heesch was to find a set of small maps (configurations) with two properties. Firstly, the set should be unavoidable in the sense that any (nontrivial) map contains at least one of the configurations as a submap. Secondly, each configuration in the set should be reducible in the sense that a map containing it can be colored in four colors provided that a certain smaller map (depending on the given map and configuration) can be colored in four colors. If such a set exists, there is no smallest counterexample to the Four Color Theorem; therefore there is no counterexample at all; therefore the Four Color Theorem is true.



Heesch worked painstakingly for the rest of his life to provide such a set. And maybe he succeeded! In 1970 he completed the first part by finding an unavoidable set of 8,904 configurations. Heesch proposed to work through these 8,904 configurations one by one, using computers, proving reducibility. For each configuration the process is a finite one; hence Heesch believed that he had finitized the

Bjarne Toft is a professor in the Department of Mathematics and Computer Science of the University of Southern Denmark. His email address is `btoft@imada.sdu.dk`.

Four Color Problem—but had he really? If some of the configurations should turn out to be nonreducible, they have to be replaced by some other configurations (maintaining unavoidability). If some of these new configurations are nonreducible, they have to be replaced, etc. Will this process lead, in a finite number of steps, to a set of unavoidable and reducible configurations? To this reviewer, it is still an interesting open question to ask whether one can prove, without using computers, that there exists a number N such that the truth of the Four Color Theorem can be established in at most N steps. Heesch had difficulties in the 1970s getting economic support to carry out the necessary reducibility computations. One reason was probably that Yoshio Shimamoto in 1971 had presented a proof of the Four Color Theorem, based on an earlier reducibility computation, which subsequently turned out to be flawed; thus the computer had in at least one case produced a wrong answer. Another reason was Heesch's failure to produce clear evidence that the problem had indeed been reduced to a finite one. These two reasons made funders skeptical about putting money into computer proofs of the Four Color Theorem. Wilson's book does not address the finitization in detail but seems to accept the idea that Heesch did obtain it.

In 1976 Kenneth Appel and Wolfgang Haken were able to complete Heesch's program. By further developing Heesch's method to find unavoidable sets, they succeeded in finding a set U of only 1,936 unavoidable and reducible configurations. The method consists of assigning to each vertex x of a planar triangulating graph a value called the charge of x. The number of edges meeting at x is the degree of x, and the charge of x is simply the integer 6 minus its degree. By Euler's Formula the total sum of all such charges of any triangulation is 12. The charges are redistributed according to specific rules (for this process Haken invented the now generally accepted and used term "discharging"). Since no charge disappears, the sum of all charges is still 12 after discharging; hence there are still vertices with a positive charge. By analyzing the positive vertices, one constructs the set U. Each configuration in U is thereafter tested for reducibility. This process is a clear and straightforward one, but it involves for each configuration so many cases that the use of computers is indispensable. Using simple examples, Wilson's book explains these methods well.

The proof by Appel and Haken has been surrounded by controversies. Firstly, it is forbiddingly long, even before reducibility. For example, the discharging is based on a set of 487 different rules. Secondly, the use of computers in mathematical proofs has been questioned. Does the use of computers reduce mathematics to an empirical science, where mathematical truths are established by running machines in laboratories? Should such use of computers therefore be avoided? Or are computers acceptable physical tools, like pen and paper, for creating eternal mathematical truths? Wilson presents the pros and cons of the debate. In the reviewer's opinion the use of computers is acceptable in mathematical proofs; when discussing this with my students, I present them with a page containing the first 2,000 digits of the decimal expansion of π. I tell them that these digits have been found using computers—different methods, different programs, and different computers produce this same list of digits. Should we doubt that these are the true digits of π because computers are involved? Do we need mathematicians to use several years of their lives to check the whole computation by hand (pocket calculators are not allowed!) to be really sure? Should we not trust a computer more than a human being testing a very large number of logically clear cases?

The Appel and Haken proof as presented in the book by Wilson seems to imply a polynomial algorithm for 4-coloring a planar graph. But this is deceptive. A main problem is the concept of configuration containment. A configuration C may "wrap around", so that vertices and/or edges on the boundary of C, which are different in the drawing of C, are in fact the same in a graph G containing C. Or, vertices on the boundary of C not joined by an edge in the drawing of C may in fact be joined by an edge in a graph G containing C. These possibilities complicate the concept of and use of configurations, and thus also the proof of Appel and Haken and its translation into a polynomial algorithm. Nevertheless, in 1989 Appel and Haken succeeded in presenting a polynomial (quartic) planar graph 4-coloring algorithm.

A strength of the Appel and Haken proof is that it is not only one proof, but in Appel and Haken's own words "there are thousands of different proofs of the theorem.…Thus an understanding of the principles involved in the proof makes the reader somewhat less concerned about the horrendous bookkeeping necessary to give all details in a particular proof.…" Therefore it does not matter so much if one particular proof contains an error. There is an error-correcting routine that will change the proof into a new one. But these many proofs may be considered a weakness also; what we need is not thousands of proofs with problematic details, but one proof without them. It would have been interesting if Wilson had taken this up in a more critical light. The many proofs of Appel and Haken are reflected in many different possible sets U. In their original 1976 announcement, Appel and Haken used a U that had 1,936 configurations; in the published proof from 1977 the size of U was 1,834 or 1,482, and later it became 1,478 or 1,476 (and Wilson mentions that 1,405 is possible).

So where do we stand today? Fortunately, in 1997 a new proof of the Four Color Theorem was published by Neil Robertson, Paul Seymour, Daniel Sanders, and Robin Thomas (See "An Update on the Four-Color Theorem," by Robin Thomas, *Notices*, August 1998, pages 848–859). It is also based on Heesch's ideas and runs along the same lines as the Appel and Haken proof. But its set U is of size only 633; it has only 32 discharging rules; and, last but not least, its concept of configuration containment is precise and without the problems encountered by Appel and Haken. A quadratic polynomial algorithm follows. Moreover, each configuration in U is "small" (it is within the second neighborhood of a vertex, a possibility suggested by Heesch). This new proof is a major achievement and a main argument for four colors sufficing. I would have liked to have seen the new proof emphasized and described in more detail in Wilson's book.

At the end of the book Wilson cites William T. Tutte's words: "The Four Color Theorem is the tip of the iceberg, the thin end of the wedge and the first cuckoo of spring." This is not substantiated, except in general terms. It would certainly have been fitting in a book like this to mention a few of the many simple and intriguing open problems, such as Gerhard Ringel's Earth and Moon Problem: Consider maps on two spheres such that each country has a connected part on each sphere. How few colors are needed to color all such maps? (As usual, neighboring countries get different colors, and the two parts of each country get the same color). A proof similar to the six color theorem for planar graphs gives 12 as an upper bound on the number of colors needed. Strangely, this has never been improved. Otherwise, all we know is that the number must be at least 9.

In conclusion, this book is an attractive and well-written account of the solution of the Four Color Problem, justly emphasizing the major achievement of Appel and Haken, the important role they played, and the immense work they carried out. It tells in simple terms an exciting story. It is not the intention of the book to be critical, but rather to give the reader a view into the world of mathematicians, their ideas and methods, discussions, competitions, and ways of collaboration. As such it is warmly recommended.

About the Cover

Heawood's demonstration of Kempe's error

This month's cover is redrawn from Robin Wilson's book on the four-color problem, reviewed in this issue. Two of the more pleasant features of the book are that Wilson has taken great care with his figures and that he has also taken care to follow the history of the problem in some detail. One of the examples of this is his discussion of Percy Heawood's demolition in 1890 of Alfred Kempe's attempted proof of the four-color conjecture, which had stood for eleven years. Counterexamples simpler than Heawood's were discovered a bit later, but his is perhaps the more interesting for its complexity. The map on the cover has been taken directly from the engraving that accompanies Heawood's original article of 1890.

As the cover illustrates, the problem is to extend the coloring of the outer rings to the central region, which requires partial recoloring of the outer regions. Kempe asserted that this could be done by two color swaps, as indicated in the two middle rows of the cover illustration, allowable because certain features of the diagram now called 'Kempe chains' separate the figure nicely. In Heawood's map, either of the two color swaps is perfectly admissible on its own, but performing both together is not, since the adjacent top (green) and lower right (yellow) regions end up colored the same (red).

—Bill Casselman

(notices-covers@ams.org)