The Book Review Column [1]

by **Nicholas Tran** (ntran@scu.edu)
Department of Mathematics & Computer Science, Santa Clara University

1 Notable New Releases

The Computer Always Wins: A Playful Introduction to Algorithms through Puzzles and Strategy Games (The MIT Press, 2025) by Elliot Lichtman (Yale University, '27) is an accessible introduction to advanced computing concepts for young coders.

Quantum Algorithms: A Survey of Applications and End-to-End Complexities (Cambridge University Press, 2025) by Alexander M. Dalzell et al. (AWS Center for Quantum Computing) summarizes the most common quantum algorithmic primitives and demonstrates their various applications.

Guide to Distributed Algorithms: Design, Analysis and Implementation Using Python (Springer, 2025) by K. Erciyes (Yaşar University, Türkiye) is a textbook on distributed algorithms with working Python implementations.

Algorithm Design: A Methodological Approach - 150 Problems and Detailed Solutions (Routledge, 2023) by Patrick Bosc, Marc Guyomard, Laurent Miclet (all at Université de Rennes, France) is the English edition of the French best-selling textbook that aims to teach algorithm design methodologies and analysis through examples.

2 This Column

Peter Ross finds *Graph Theory in America: The First Hundred Years* (Princeton University Press, 2023) by Robin Wilson, John J. Watkins, and David J. Parks to be a well-written and engrossing history of graph theory in the United States. The book covers extensively the development of graph theory from its early days in the 19th century to the proof of the Four Color Theorem in 1976, and it includes many interesting anecdotes about the mathematicians and computer scientists involved in this history.

Sarvagya Upadhyay rates *Introduction to Probability for Computing* (Cambridge University Press, 2023) by Mor Harchol-Balter (Carnegie Mellon University) among the most engaging textbooks on probability theory. He recommends it to students seeking an applied yet rigorous perspective on the subject. Instructors interested in adopting this textbook can obtain a free version on the author's website and many teaching resources (slides, handouts, syllabi) by contacting the author.

Interested in Ramsey theory but don't know which book to read? Bill Gasarch has got you covered with his joint reviews of *Rudiments of Ramsey Theory*, 2nd ed. (AMS, 2015) by Ron

Graham and Steve Butler, *Fundamentals of Ramsey Theory* (CRC Press/Routledge, 2021) by Aaron Robertson, *Elementary Methods of Graph Ramsey Theory* (Springer, 2022) by Yusheng Li and Qizhong Lin, and *Basics of Ramsey Theory* (CRC Press/Routledge, 2023) by Veselin Jungić. Bill provides a detailed comparison of these four books plus the three books on this subject he reviewed earlier in this column, listing their common features as well as pointing out their unique aspects.

3 How to Contribute

Consider writing a book review for SIGACT News when planning your summer reading. Either choose from the books listed below, or propose your own. In either case, the publisher will send you a free copy of the book. Guidelines and a LaTeX template can be found at `https://algoplexity.com/~ntran`.

BOOKS THAT NEED REVIEWERS FOR THE SIGACT NEWS COLUMN

Algorithms, Complexity, & Computability

1. Vaze, R. (2023). *Online Algorithms.* Cambridge University Press.
2. Ferragina. P. (2023). *Pearls of Algorithm Engineering.* Cambridge University Press.
3. Bosc, P., Guyomard, M., & Miclet, L. (2023). *Algorithm Design: A Methodological Approach - 150 Problems and Detailed Solutions.* Routledge.
4. Downey, R. (2024). *Computability and Complexity: Foundations and Tools for Pursuing Scientific Applications.* Springer.
5. Traub, V., & Vygen, J. (2024). *Approximation Algorithms for Traveling Salesman Problems.* Cambridge University Press.
6. Brody, J. (2025). *The Joy of Quantum Computing: A Concise Introduction.* Princeton University Press.
7. Dalzell, A., & McArdle, S., & Berta, M., & Bienias, P., & Chen, C.-F., & Gilyén, A., & Hann, C., & Kastoryano, M., & Khabiboulline, E., & Kubica, A., & Salton, G., & Wang, S., & Brandão, F. (2025). *Quantum Algorithms: A Survey of Applications and End-to-end Complexities.* Cambridge University Press.
8. Erciyes, K. (2025). *Guide to Distributed Algorithms: Design, Analysis and Implementation Using Python.* Springer.

Miscellaneous Computer Science & Mathematics

1. Grechuk, B. (2019) *Theorems of the 21st Century.* Springer.
2. Nisan, N., & Schocken, S. (2021). *The Elements of Computing Systems: Building a Modern Computer from First Principles, 2nd ed.* The MIT Press.

3. Chayka, K. (2024). *Filterworld: How Algorithms Flattened Culture*. Doubleday.

4. Valiant, L. (2024). *The Importance of Being Educable: A New Theory of Human Uniqueness*. Princeton University Press.

5. Lichtman, E. (2025). *The Computer Always Wins: How Algorithms Beat Us at Our Own Games*. The MIT Press.

6. Rojas, R. (2025). *The Language of Mathematics: The Stories behind the Symbols*. Princeton University Press.

Discrete Mathematics and Computing

1. Ross, S., & Peköz, E. (2023). *A Second Course in Probability*. Cambridge University Press.

Cryptography and Security

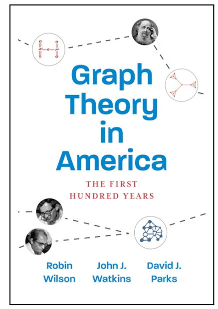
1. Chen, K., & Yang, Q. (2023). *Privacy-Preserving Computing for Big Data Analytics and AI*. Cambridge University Press.

2. Garfinkel, S. (2025). *Differential Privacy*. The MIT Press.

Combinatorics and Graph Theory

1. Landman, B., Luca, F., Nathanson, M., Nešetřil, J., & Robertson, A. (Eds.). (2022). *Number Theory and Combinatorics: A Collection in Honor of the Mathematics of Ronald Graham*. De Gruyter.

Review of [1]



Graph Theory in America:
The First Hundred Years
Robin Wilson, John J. Watkins, and David J. Parks

Princeton University Press, 2023
$35 Hardcover, $35 eBook, 320 pages

Review by
Peter Ross (pross@scu.edu)

1 Introduction

The book focuses on the development of graph theory in the USA and Canada from 1876 to 1976. This is the period from when J. J. Sylvester arrived at Johns Hopkins University from England in 1876, until the proof of the Four Color Theorem by Appel and Haken at the University of Illinois in 1976. But the book covers much more, including a preliminary chapter on early American mathematics, two Interludes that are essentially short chapters on graph theory in Europe, and an Aftermath, an eight-page summary of progress in graph theory in recent years.

Graph Theory in America (GTA) is based on a doctoral dissertation by co-author David Parks under the supervision of co-author Robin Wilson. Its Preface states, “No prior knowledge of graph theory is required in reading this book, which aims to explain the historical development of the subject in simple terms to a general reader interested in mathematics.” It adds that “readers who are interested mainly in the historical narrative, and in the personalities involved, can safely pass over any technical material.” I believe that GTA has done this successfully, and my review will try to focus equally on the history and the mathematics involved.

2 Summary

The six main chapters cover chronologically the 1800s through the 1960s and 1970s. Each chapter discusses advances in graph theory and the key mathematicians involved, as well as progress on the four color problem (FCP) and eventually the Four Color Theorem. A **preliminary chapter**, “Setting the Scene,” focuses on Benjamin Peirce at Harvard and especially Eliakim Hastings Moore at the University of Chicago, who was there from 1892 until his death in 1932. Moore worked in algebra, the foundations of geometry, and analysis, but never in graph theory. However, two of his 31 doctoral students, Oswald Veblen and George Birkhoff, made significant contributions to graph theory that are discussed in Chapter 2 on the 1900s and 1910s.

The story of the FCP begins in **Chapter 1**, on the 1800s, when Francis Guthrie, a former student of De Morgan, stated it in 1852. The problem intrigued Hamilton and also Cayley, who showed in 1878 that to prove that four colors can color any (planar) map, it suffices to show that they can color any cubic map, one with 3 countries at each meeting point or vertex. In 1879 Cayley's former student Alfred Bray Kempe published a purported solution to the FCP in the new *American Journal of Mathematics*, whose editor in chief was J. J. Sylvester, a friend of Cayley. Kempe's "proof" was generally accepted until 1890, when Percy Heawood exposed its fatal error. Much of Chapter 1 discusses Sylvester's fascinating career, including his back-and-forth moves to England and his innovative uses of graphs to represent molecules in chemistry and "binary quantics" in algebra, and an analogy between them. The book mentions that his 1873 note in *Nature* on this analogy included the first use of the word "graph" in our modern sense.

The **first Interlude** on graph theory in Europe discusses Heawood's ground-breaking paper "Map-colour Theorem". The paper not only exposed the fatal flaw in Kempe's long accepted "proof" but also showed that 5 colors suffice for a map on a plane or sphere by modifying Kempe's arguments. Heawood went on to study coloring of other orientable surfaces like a torus. He proved that the chromatic number of a torus, the smallest number of colors to color every map on it, is 7. He further conjectured the general formula $\lfloor \frac{7+\sqrt{1+48g}}{2} \rfloor$ for the chromatic number of an orientable surface of genus g (g handles or holes) where $g \geq 1$. This so-called Heawood Conjecture was not proven for over seventy years, until Gerhard Ringel and Ted Youngs proved it at the University of California at Santa Cruz in 1968. The *Interlude* chapter discusses other notable European mathematicians like Julius Petersen of Denmark, for whom the Petersen graph is named, Heinrich Tietze of Austria, who studied the chromatic number of non-orientable surfaces, and Hermann Minkowski of Germany. It also repeats an amusing story from Constance Reid's biography of Hilbert, of how Minkowski interrupted his topology lecture at Göttingen to assert that the Four Color Theorem "has not yet been proved, but that is because only mathematicians of the third rank have occupied themselves with it." Minkowski added that he believed he could prove it, and proceeded to work on it in class for several weeks, before giving up and admitting his arrogance.

Chapter 2 on the 1900s and 1910s discusses significant contributions to graph theory by George D. Birkhoff at Harvard and Oswald Veblen at Princeton. Veblen used algebra to discuss maps, like modular equations and incidence matrices for graphs, following Henri Poincaré. Veblen's influential 1922 monograph *Analysis Situs* on combinatorial topology was based on his 1916 *AMS Colloquium Lectures in Cambridge, MA*. Among other things the lectures introduced the concepts of rank and nullity of a map or planar graph. They also developed some 19th century ideas of G.R. Kirchhoff who, in studying charges in electrical networks, introduced the idea of a spanning tree. George D. Birkhoff, "the leading mathematician of his time," had a mild obsession with the FCP according to the GTA book, but in later life Birkhoff rued the time and effort he spent on it. However, Birkhoff once told Hassler Whitney that "every great mathematician (of the time) had studied the problem, and thought at some time that he had proved the theorem." Birkhoff's first paper on the subject introduced the chromatic polynomial of a map $P(\lambda)$, which counts the possible colorings of the map with λ colors. $P(\lambda)$ is of degree n, where n is the number of regions of the map. His main objective was to prove $P(4) > 0$, the Four Color Theorem, but his methods involving determinants were cumbersome. However, twenty years later Whitney discovered a simpler procedure for determining the coefficients of $P(\lambda)$. Birkhoff was regarded as North America's leading mathematician after he solved a restricted form of the three-body problem, like finding the motion of the sun, Earth, and moon, which Poincaré had been unable to solve.

Chapter 3 on the 1920s focuses a lot on Philip Franklin, whose doctoral thesis, supervised by Veblen, was on the FCP. It recounts an amusing story by David Widder, who had a bunk in the same barracks as fellow mathematicians Franklin and Norbert Wiener during World War I, while working at the Aberdeen Proving Ground in Maryland. Franklin and Wiener sometimes talked mathematics so far into the night that, in order to get some sleep, Widder once hid the light bulb! In later life Franklin and Wiener became brothers-in-law at MIT, where both spent the majority of their careers, and where each brought different forms of topology to the institute. Franklin's 1922 paper on the FCP proved that the conjecture was true for every map on the plane (or sphere) with 25 or fewer regions. In the late 1930s Franklin improved this to 31 or fewer regions, followed by C.E. Winn, who showed 35 or fewer regions sufficed, a result that held for thirty years. Also in the 1930s Franklin finished off a Heawood conjecture for non-orientable surfaces of genus g, where g is the number of cross-caps added to a sphere, so g is 1 for the projective plane and 2 for the Klein bottle. Work by H. Tietze in 1910 had led to a Heawood conjecture for non-orientable surfaces of genus $g \geq 1$, that their chromatic number was $\lfloor \frac{7+\sqrt{1+24g}}{2} \rfloor$. Franklin proved that this failed for $g = 2$, and in 1952 Gerhard Ringel showed that Franklin had found the only instance where Tietze's formula fails.

A **second Interlude** on graph theory in Europe discusses the work of Polish topologist Kazimierz Kuratowski on planar graphs. His central theorem of 1922 asserted that a graph is planar if and only if it has no subgraph that is homeomorphic to the complete graph K_5 or the complete bipartite graph $K_{3,3}$. Kuratowski had "a most distinguished career," collaborating with major figures like Stefan Banach, Max Zorn (of Zorn's Lemma), and John von Neumann, and he served as "a world ambassador for Polish mathematics."

Chapter 4 on the 1930s sees the prestige of graph theory so low that one mathematician then called it "the slums of topology." Two important American mathematicians, Hassler Whitney and Saunders Mac Lane, changed this. In the early 1930s Whitney wrote a dozen papers on major areas of graph theory like coloring, planarity, duality, and matroids. Duality included the combinatorial dual of a graph, which is sometimes called the Whitney dual. Matroids arose from Whitney noticing similarities between the concepts of rank and independence in graph theory and those of dimension and linear independence for vector spaces. GTA briefly surveys Whitney's major contributions to other topics than graph theory, such as his strong embedding theorem for n-dimensional differentiable manifolds, the cup product in cohomology rings, and work on singular spaces and the singularities of smooth maps, which eventually led to catastrophe theory and chaos theory. Saunders Mac Lane met Whitney while both were young faculty members at Harvard in the mid-30s. Mac Lane extended Whitney's work on combinatorial graphs in three papers, before moving on to other fields like algebraic topology and category theory, where he had a significant impact. His texts also were a major influence on mathematics education in universities. Along with Ron Graham, Mac Lane was President at times of both the AMS and the MAA.

Chapter 4 ends with 8 pages on academic life in the 1930s, discussing the effects of the Great Depression and the early years of the war in Europe. This included the birth of the AMS's *Mathematical Reviews*, which was proposed by Veblen after G.H. Hardy, Harald Bohr and others resigned from the reviewing journal *Zentralblatt für Mathematik* due to its actions that were influenced by the Nazi Party. Another controversy, which 8 pages deal with, is that of antisemitism in some mathematics departments against Jews, especially refugees from the Nazis, who were viewed as taking jobs away from budding American mathematicians. A prominent leader with this opinion was George Birkhoff at Harvard, who in 1934 even initially opposed Solomon Lefschetz's becoming the first Jewish president of the AMS.

The meaty fifty-page **Chapter 5** on the 1940s and 1950s focuses on further contributions to graph theory by established mathematicians like Birkhoff, and new contributors like Canadian Bill Tutte and Americans Claude Shannon and Frank Harary. Bill Tutte left Cambridge University in January 1941 to join Alan Turing at the highly secret Bletchley Park near London working as a codebreaker. The book says that his work in unraveling the internal workings of the German cipher machine that replaced the Enigma machines was "an astonishing feat of cryptanalysis that is sometimes believed to have shortened the war by two years or more." After the war Tutte returned to Cambridge, earning his doctorate with a 417-page dissertation *An Algebraic Theory of Graphs*, which he later said attempted to reduce graph theory to linear algebra. GTA points out that the thesis included the first major advances on matroids since their introduction by Whitney in 1935. After getting his doctorate, Tutte joined H.S.M. (Donald) Coxeter at the University of Toronto for 14 years, before moving to the University of Waterloo that's also in Ontario, for the last 36 years of his long career. GTA devotes 11 pages to some of Tutte's work, including showing his example of a cubic polyhedron (so of degree 3) with no Hamiltonian cycle, thus disproving Tait's 1884 claim that every cubic polyhedron has one. The book also discusses Tutte's discovery of what became known as the Tutte polynomial, a polynomial in two variables, and his later work at Waterloo on chromatic polynomials.

Chapter 5 includes a concise 11-page summary of progress on algorithms, which arose from practical problems in World War II and later from the budding computer industry. These algorithms include ones for matching and assignment, transportation, linear programming, flows in networks, finding minimum spanning trees, search algorithms, and path problems such as finding the shortest or longest paths in graphs. The final pages of the chapter are devoted to Frank Harary, who came to be known as "the father of modern graph theory." Harary wrote eight books and more than seven hundred papers in his long career at Michigan and later New Mexico State University in Las Cruces. GTA discusses just three of the many areas that Harary worked in: signed graphs, graph enumeration that was based on earlier work by Cayley and Polya, and Ramsey graph theory. Ramsey theory became popular, according to the GTA, after a coloring problem for the complete graph K_6 appeared on the 1953 Putnam exam that was proposed by Harary. Harary may have gotten the problem from this similar one from the Hungarian Mathematical Olympiad of 1947: *If there are six people at a party, prove that there must be at least three mutual friends or three mutual non-friends.* GTA presents an elegant two-sentence solution to the graph theory version of the problem.

Chapter 6, the 1960s and the 1970, opens by observing that graph theory was increasingly becoming part of mainstream mathematics, and the two decades saw the long-awaited proofs of both the Heawood conjecture and the Four Color Theorem. Oystein Ore, a Norwegian whose long career was at Yale, wrote three books on graph theory. The first, his 1962 *AMS Colloquium Lecture series on the Theory of Graphs*, was one of the first two texts in English on the subject. Ore along with two doctoral students, well-known combinatorialist Marshall Hall, Jr., and Joel G. Stemple, showed that a planar map not colorable in four colors must have at least 40 countries. Others extended this result to 48 and then 96 countries, but the GTA authors mischievously note, "there was still a long way to go." As noted earlier, the Heawood formula for the chromatic number of any orientable surface of genus g was finally proven by Ringel and Youngs in 1968. The book examines some of the cases and the ideas in their proof, which it calls a *tour de force*, and notes that the theorem is now known as the Ringel-Youngs theorem. Surprisingly, the Heawood conjecture for non-orientable surfaces had been completely settled by Ringel earlier in 1954, and GTA briefly discusses why those surfaces were easier to deal with. Gerhard Ringel had a very unusual background for a mathematician. Raised in Czechoslovakia and graduating from Charles University in Prague, he was drafted into the Wehrmacht in World War II and later spent four years as a POW in a Soviet jail. After being released he obtained a doctoral degree from the University of Bonn in 1951. He taught for many years in Germany, then moved to the University of California in Santa Cruz, where he succeeded his friend Ted Youngs who had retired.

Another "colorful character" who's described in Chapter 6 is Ron Graham, whom G.-C. Rota called "the leading problem-solver of his generation" in his 1991 nomination of Graham for AMS President. Graham never graduated from high school, due to his father changing jobs around the country. He funded his doctoral studies at Berkeley in the early 1960s with a trampolining troupe and a juggling act that he formed and performed in circuses and elsewhere. Persi Diaconis once said, "Ron, as much as anybody, is responsible for bringing high-powered math to bear on computer science." GTA discusses in detail one of Graham's approximately 400 papers, that was written with his Bell Labs colleague Henry O. Pollak, and that was motivated by telephone switching theory. Graham and Pollak introduced the distance matrix of a graph, whose (i, j) entry was the length of the shortest path between vertices i and j. Pollak later recalled that people had previously only studied the adjacency matrices of graphs. Graham and Pollak's main theorem had wide-ranging consequences, including a simple formula for the determinant of the distance matrix of a tree, which turned out to be independent of the structure of the tree. GTA points out that a clever proof of the formula is based on a determinant rule discovered by the 19th century English mathematician C.L. Dodgson, better known as writer Lewis Carroll. GTA briefly discusses Graham's extensive writings on Ramsey theory, many with his wife Fan Chung, and his very unusual friendship and collaboration with, and stewardship of, Paul Erdős. Graham even defined the amusing Erdős number of an author, as the length of the shortest chain of collaborators from Erdős to the author.

Chapter 6 includes an excellent six-page section on Complexity, which features important contributions by Jack Edmonds, who may have originated the fundamental question *Is P = NP?* in 1967, and Stephen Cook's definition of NP-complete problems, along with some of his startling results. GTA mentions that despite a Clay Mathematics Institute prize of one million dollars for deciding if P = NP, "since the 1970s, little progress has been made in settling this general problem." I was surprised in reading the concise section on Complexity to find out that John von Neumann caused interest to increase in the efficiency of algorithms, when in 1953 he apparently was the first to distinguish between polynomial-time and exponential-time algorithms in an article linking game theory to the optimal assignment problem.

The final sentence of Chapter 6, at the end of 18 pages of progress on the FCP by Wolfgang Haken, Heinrich Heesch, Kenneth Appel and others, simply states "The Four Color Theorem was proved." Haken made a name for himself in the early 1950s by solving a long-standing unsolved knot problem, determining if a given tangle of string contains a knot. This led to his giving an invited lecture at the 1954 International Congress of Mathematicians in Amsterdam. Afterwards he moved from the University of Kiel in Germany to the University of Illinois, where in 1967 he got four color problem expert Heesch to visit. Heesch had used a CDC 1604A computer for the FCP in Germany, and together Appel and Heesch got time on the Cray Control Data 660 computer, the most powerful machine of the day, at the Atomic Energy Commission's Brookhaven National Laboratory on Long Island. Coincidentally the computer center's director Yoshio Shimamoto was a FCP enthusiast who himself discovered an apparently key configuration of map regions in 1971, which became known as the Shimamoto horseshoe. With help from the "most distinguished graph theorists of the day" Hassler Whitney and Bill Tutte, Shimamoto's approach was found to lead to a dead end. By the early 1970s Haken considered giving up on the FTP, and in a lecture in Illinois he described computer experts' negative views on the massive calculations that his research plan needed. Fortunately Kenneth Appel, a mathematician in the audience, had extensive programming experience and differed with the experts' opinions and volunteered to help. Appel and Haken started working together in 1972, but it took four more years, with a new approach and a powerful new computer at Illinois that only Appel seemed to be able to get to run properly, before in late June 1976 Appel placed a notice on the department's blackboard saying, "Modulo careful checking, it appears that four colors suffice." The phrase *four colors suffice* became the department's postal meter slogan. At the same time, others using similar methods at the University of Waterloo, the University of Rhodesia, and Harvard expected success within months. Appel and Haken, using their children to help check the final configurations, went public on July 22, 1976. A day later *The Times* of London reported their proof, adding that it contained 10,000 diagrams and that the computer printout stood four feet high on the floor.

An eight-page **Aftermath** chapter in GTA briefly summarizes the huge increase in graph theory and combinatorics in America since the proof. It lists fourteen research topics, about half of which are new to the book. The chapter ends by asserting "The development of graph theory in America over the century from 1876 to 1976 was truly remarkable ...," listing ten mathematicians who were undoubtedly among the most significant. Each of these has been mentioned in this review.

3 Opinion

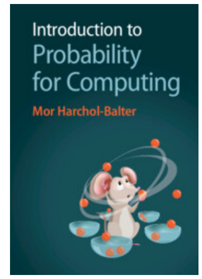
It's probably clear by now that I think very highly of the book. It will likely interest readers who don't know graph theory but do know some history of mathematics, and those who don't know much history of mathematics but do know some graph theory. The authors are good story tellers who make graph theory and history come alive. I noticed no technical errors, and in fact spotted not a single misspelling or typo.

The book has many nice features:

- short summaries of some influential papers that influenced the subject's development;
- a paragraph or two at the beginning or end of each chapter, giving a preview of what's coming and serving as an advanced organizer;
- a ten-page glossary of technical terms and a five-page index;
- twenty pages of detailed notes, references, and further reading;
- a nine-page chronology of events, beginning with the founding of Harvard in 1636 and ending with the awarding of the Abel Prize to László Lovász and Avi Wigderson in 2021;
- photos, many just head shots, of almost all of the major mathematicians discussed in the book.

Finally, the hardcover book is well-made, surviving my extensive annotations, and is very reasonably priced, available from Amazon for less than $25.

Review of [1]



Introduction to Probability for Computing
Mor Harchol-Balter
Cambridge University Press, 2023
Hardcover, 555 pages, $69.99

Review by
Sarvagya Upadhyay (supadhyay@fujitsu.com)
Fujitsu Research of America
4655 Great America Way, Suite 410
Santa Clara CA 95054, USA

1 Overview

Probability theory provides a rigorous mathematical foundation for analyzing and modeling uncertainty. It is a fundamental tool with widespread applications across numerous scientific disciplines. Its applications extend far beyond the multitude of science disciplines, playing a central role in two of the most prominent industrial sectors of today – finance and technology. In finance, sophisticated stochastic models are utilized for evaluating complex financial instruments and understanding market dynamics. In machine learning and statistics, probability is deeply embedded in models where decisions are not made deterministically, but instead are guided by the likelihoods of various outcomes – enabling systems to learn from data and adapt under uncertainty.

This book is a gentle introduction to probability theory from a computing perspective. It covers basic concepts along with several computing applications of the topic.

2 Summary of Contents

This book is structured into eight distinct parts. The initial three parts provide a comprehensive foundation in probability theory, encompassing topics such as discrete and continuous random variables. The subsequent five parts are dedicated to diverse applications of probability theory. As outlined in the preface, these sections may be utilized to develop curricula for four separate courses with an emphasis on applied probability.

Part I: Fundamentals and Probability of Events This part lays the mathematical groundwork essential for the study of probability theory. It is divided into two chapters. The first chapter revisits several fundamental mathematical concepts such as series, limits, integrals, and asymptotic notations. These tools are crucial for both understanding and solving problems in subsequent chapters. The second chapter delves into the basic principles of probability theory, introducing key topics such as the definition of sample spaces and events, the assignment of probabilities to events, conditional probability, independence, and Bayes' law. Throughout this part, carefully chosen examples are used to highlight important nuances. The structure and pedagogical approach adopted here perfectly sets up the presentation style of remainder of the book.

Part II: Discrete Random Variables This part of the book consists of four chapters, each dedicated to the study of discrete random variables. In Chapter 3, the formal definition of discrete random variables is introduced. This chapter also presents several commonly encountered discrete probability distributions such as the Bernoulli, Binomial, Geometric, and Poisson distributions. The next two chapters extend this foundation by examining moments of random variables, including expectations and linearity of expectations, variances, and higher-order moments. These concepts are essential for quantifying the central tendency and variability of random phenomena. Building upon this, the fifth chapter introduces the sum of random variables, particularly useful in contexts like algorithm analysis and queuing systems. It also introduces the fundamental tail inequalities such as Markov's and Chebyshev's inequalities, which provide probabilistic bounds on the likelihood of extreme outcomes. In addition, the concepts of stochastic dominance and Jensen's inequality are discussed, offering powerful tools for comparing random variables and analyzing convex functions of them. The final chapter of this part introduces the z-transform, an analytical tool widely used in the analysis of discrete-time systems. The author demonstrates how z-transforms can be effectively employed to compute moments and solve recurrence relations.

Part III: Continuous Random Variables As the author notes, much of the material presented in this part parallels the concepts introduced earlier in the context of discrete random variables, but is now extended to the continuous domain. This section comprises five chapters, each focusing on different aspects of continuous random variables and their applications. The first two chapters lay the theoretical foundation. The opening chapter explores continuous random variables drawn from a single distribution, introducing the probability density function (PDF), cumulative distribution function (CDF), and methods for computing expectations and variances. The following chapter expands on this by examining jointly distributed continuous random variables, discussing marginal and conditional distributions, and properties of moments. The proof of these properties are deferred to the exercises at the end of the chapter. Chapters 9 and 10 are devoted to two of the most widely used continuous distributions in both theory and practice: the Gaussian (Normal) and Pareto distributions. Chapter 9 also introduces one of the most important results in probability theory, the Central Limit Theorem (CLT), which underpins much of statistical inference and analysis. Chapter 10 begins with a particularly engaging introduction, a personal anecdote from the author. The story, centered around the author attending an operating systems course, creates a memorable entry point into the material. I must add that the author's sentiment resonated with me! The final chapter of this part introduces the Laplace transform, a powerful analytical tool for continuous random variables. The author demonstrates how Laplace transforms can be used to compute moments much like the z-transform is used for discrete random variables.

Part IV: Computer Systems Modeling and Simulations This part consists of three chapters, each focused on foundational concepts in stochastic modeling and random process simulation. It begins in Chapter 12 with a review of the exponential distribution, a key continuous distribution widely used in modeling the time between independent events. Building on this, the chapter introduces the Poisson process, where the author presents a few of important properties of such processes. The subsequent chapter delves into two primary techniques for generating instances of random variables from arbitrary distributions. The inverse transform method is introduced first, demonstrating how to transform uniformly distributed random numbers into samples from a desired distribution. This is followed by the accept-reject method, which is particularly useful when the inverse of the cumulative distribution function is difficult or impossible to compute directly. The final chapter of this part shifts the focus to event-driven simulation, a technique used to model systems where state changes occur at discrete points in time, often triggered by stochastic events. Within this framework, the author introduces key definitions from queuing theory. While this chapter provides an introductory overview, the author provides a more in-depth treatment of the topic in Chapter 27.

Part V: Statistical Inference This part of the book is conceptually independent from the preceding parts and shifts the focus from probability theory to the domain of statistical inference. The central aim here is to estimate unknown parameters based on observed data—a fundamental task in data analysis, machine learning, and many areas of applied computing. This part comprises three chapters, each systematically building the foundation for understanding and applying different estimation techniques. The first chapter introduces the most widely used point estimators, particularly the sample mean and variance. These estimators serve as intuitive and mathematically tractable tools for summarizing data and making inferences about the underlying distribution. Their properties, such as unbiasedness and consistency, are discussed to highlight their effectiveness in various practical settings. The second chapter delves into Maximum Likelihood Estimation (MLE), one of the most powerful and widely applied methods for parameter estimation. The author explains the importance of MLE through various detailed examples. The final chapter in this part, Chapter 17, introduces estimation from a Bayesian perspective, where parameters are treated as random variables with prior distributions. The chapter presents two key Bayesian estimators: the Maximum A Posteriori (MAP) estimator, which identifies the most probable parameter value given the observed data and prior beliefs, and the Minimum Mean Squared Error (MMSE) estimator, which minimizes the expected squared error between the estimate and the true parameter.

Part VI: Tail Bounds and Applications This part of the book, composed of three chapters, provides a deeper and more rigorous exploration of the tail behavior of random variables, a topic that was introduced in a preliminary form back in Chapter 5 through general tail inequalities such as Markov's and Chebyshev's bounds. Here, the author significantly extends this discussion by introducing stronger and more refined probabilistic bounds that are essential in both theoretical analysis and practical applications. The first chapter in this part, Chapter 18, focuses on two of the most powerful concentration inequalities in probability theory: the Chernoff bound and the Hoeffding inequality. These bounds provide exponentially decreasing probabilities for large deviations of sums of independent random variables from their expected values. Building on this, the next chapter demonstrates how tail bounds can be used to derive confidence intervals for statistical estimates. These results are particularly valuable in situations where one needs to quantify uncertainty in empirical observations. To illustrate the implications of these techniques, the author presents a series of classic balls-and-bins problems (including exercises). The final chapter in this part applies the previously developed tools to the domain of hashing algorithms, a fundamental component in computer science. By leveraging both tail bounds and balls-and-bins analysis, the author shows how to estimate the probability of hash collisions and estimate the size of hash buckets.

Part VII: Randomized Algorithms This part of the book is devoted to the study of randomized algorithms, a powerful class of computational techniques that incorporate randomness as a core component in algorithm design. The part comprises three chapters, each focusing on a distinct type of randomized algorithm and its applications. The discussion begins in Chapter 21, which introduces Las Vegas algorithms – randomized algorithms that always produce the correct result, but whose runtime is a random variable that depends on the outcomes of internal coin tosses. These algorithms are particularly useful when correctness is of paramount importance. The chapter presents two classic examples of this algorithmic variety: randomized quicksort and randomized median-finding. The subsequent chapter turns to Monte Carlo algorithms, which differ from Las Vegas algorithms in that their runtime is typically fixed or bounded, but the correct output can only be obtained with high probability. The author provides a variety of examples throughout the main text and exercises, including randomized matrix multiplication checking, the MAX-CUT and MIN-CUT problems. The final chapter in this part addresses a classic problem in theoretical computer science: primality testing. The chapter culminates with a detailed presentation of the Miller-Rabin primality test, a probabilistic algorithm that determines whether a number is prime with high probability.

Part VIII: Discrete-Time Markov Chains The final part of the book is dedicated to one of the most versatile and mathematically rich topics in probability theory: Markov chains. These models, which describe systems that transition between states in a memoryless fashion, have profound applications across a wide array of disciplines—including computer science, operations research, economics, and biology. While the book primarily emphasizes application-oriented learning, this section provides a meaningful foray into the theory of Markov chains to support and deepen that practical understanding. Of all the sections in the book, this one is arguably the most mathematically intensive. In Chapter 24, finite-state discrete-time Markov chains are introduced. Key concepts such as transition probability matrix, stationary distributions, and limiting distributions are developed. The author also shows that, under certain conditions, the stationary and limiting distributions coincide in the finite-state case. In the next chapter, we delve more deeply into the theoretical properties of finite-state Markov chains. It introduces crucial concepts such as irreducibility and aperiodicity, and explains how these properties link to the existence of limiting distributions. The chapter also explores topics such as mean return times to a given state and long-run time averages. A highlight of this chapter is the application of Markov chains to the PageRank algorithm used by Google. The next chapter shifts focus to infinite-state discrete-time Markov chains. Since the theoretical framework is more complex, the author focuses on prioritizing intuition and conceptual clarity over technical rigor. The concluding chapter of both this part and the book, Chapter 27, is devoted to queuing theory, an application area for Markov chains. This chapter synthesizes concepts developed throughout the part to show how stochastic processes can model and analyze waiting lines and service systems. Topics such as Little's Law and performance metrics for queues are introduced. Admittedly, the material here is dense and can be challenging, especially for readers like me who are unfamiliar with the subject. However, the author's pedagogical style helps in going through this chapter.

3 Evaluation and Opinion

In contrast to many traditional probability texts, this book deliberately steers away from overly abstract or purely theoretical treatments. Instead, it emphasizes practical insight and real-world intuition, aligning closely with the author's goal of delivering an application-oriented approach to probability. Drawing on more than two decades of teaching experience, the author brings a distinct pedagogical style to the material—one that is rich in illustrative examples, deeply intuitive, and highly accessible. The book is written in a conversational yet rigorous manner, making it suitable for a broad audience without compromising on the mathematical depth required for serious study.

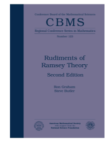
For readers used to traditional mathematical books, this approach of mixing mathematical concepts with real-world applications and examples might seem a bit odd. However, as one progresses through the chapters, these practical illustrations become invaluable. There are plenty of exercises and examples, which are thoughtfully designed to reinforce the material covered in the chapters. They are worth doing!

One of the standout features of the book is its thoughtfully modular structure. For instance, rather than covering discrete and continuous random variables into two big chapters, the author opts to divide them into nine concise chapters. This approach makes the material more digestible, especially for beginners and practitioners who may be revisiting probability concepts. It seems that the aim of such conciseness is to allow the readers to focus on specific concepts without feeling overwhelmed. The content of most chapters and their exercises can be covered easily in a day. This structure not only aids in building a solid foundational understanding but also encourages readers to progress through the material with confidence.

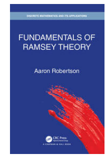
This book is a great resource for undergraduate students seeking an applied perspective on probability. Among the probability books that I've read, this one has been the most engaging one. I thoroughly enjoyed reading the book. This book is borne out of a *Probability and Computing* class that Harchol-Balter has offered for over two decades. Given the presentation style of the book, I can only wish that I had the opportunity to attend one of her classes.

Joint Review of [1]

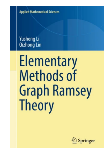
Rudiments of Ramsey Theory (2nd ed.)
by **Ron Graham** and **Steve Butler**
AMS, 2015, Softcover, 83 pages, $33



Fundamentals of Ramsey Theory
by **Aaron Robertson**
CRC Press, 2021, Paperback, 255 pages, $74.99



Elementary Methods of Graph Ramsey Theory
by **Yusheng Li** and **Qizhong Lin**
Springer, 2022, Hardcover, 349 pages, $53.99



Basics of Ramsey Theory
by **Veselin Jungić**
CRC Press, 2023, Hardcover, 238 pages, $91

Review by
William Gasarch (gasarch@cs.umd.edu)
Department of Computer Science
University of Maryland, College Park, MD

1 Introduction

If I want to start an argument on my blog, I might post

> *When I teach automata theory, I do not cover context-free grammars!*

There is a large group of people who teach automata theory and a large set of textbooks on the topic. Hence there are many different ideas of what should be in the course. Debate is possible and desirable.

If I were to post

> *When I teach Ramsey theory, I do not cover the Hales-Jewett theorem!*

I doubt I would get an argument going. There are very few people who teach a course in Ramsey theory. There are a few textbooks on it. In this column I review four of them. This could be the starting point for an argument.

I reviewed three other books on Ramsey theory in the past:

- *Ramsey Theory for Discrete Structures* by Prömel. The review is in the column:

  `https://mathcs.clarku.edu/~fgreen/SIGACTReviews/bookrev/48-4.pdf`

- *An Introduction to Ramsey Theory: Fast Functions, Infinity, and Metamathematics* by Katz and Reimann. The review is in the column:

  `https://mathcs.clarku.edu/~fgreen/SIGACTReviews/bookrev/50-2.pdf`

- *Ramsey Theory over the Integers*, 2nd ed. by Landman and Robertson. The review is in the column:

  `https://mathcs.clarku.edu/~fgreen/SIGACTReviews/bookrev/47-2.pdf`

Of the four books we review, three of them are textbooks, and the fourth one (by Graham & Butler) is notes for a short course aimed at mathematicians.

2 Definitions and Basic Theorems

Def 2.1

1. N is the naturals, $\{1, 2, 3, \ldots\}$.
2. If $n \in \mathsf{N}$ then $[n] = \{1, \ldots, n\}$.
3. If A is a set and $a \in \mathsf{N}$ then $\binom{A}{a}$ is the set of a-sized subsets of A. Note that $\binom{[n]}{a}$ is the edges of the complete a-hypergraph.
4. Let $A \subseteq \mathsf{N}$ (it will either be $[n]$ or N).
5. In Ramsey theory we are often given, as a premise of a theorem, a *coloring*, which is just a map to colors, often c of them. For notation we usually use $[n]$ rather than (say) $\{\mathrm{RED}, \mathrm{BLUE}\}$. We usually denote the coloring function COL. This is all a prelude to the next definition:

   Let $a, c \in \mathsf{N}$ and let $\mathrm{COL}\colon \binom{A}{a} \to [c]$. Let $H \subseteq A$. A is *homogeneous* if COL restricted to $\binom{H}{a}$ is constant.

We describe several theorems from Ramsey theory.

1. *Ramsey's Theorem [14]:* For all a, k, c there exists $n = R_a(k, c)$ such that for all $\mathrm{COL}\colon \binom{[n]}{a} \to [c]$ there exist a homogeneous set of size k.

   There has been much work on upper and lower bounds on $R_a(k, c)$, especially $R_2(k, 2)$, which we denote $R(k)$. Using elementary methods (hence suitable for a course) one can show

   $$(1 + o(1))\frac{k}{\sqrt{2}e}2^{k/2} \le R(k) \le (1 + o(1))\frac{4^{s-1}}{\sqrt{\pi k}}$$

   The upper bounds is due to Erdös & Szekeres [3]. The lower bound was obtained by Erdös [4] using the probabilistic method. That last sentence is true but odd. Erdös invented the probabilistic method in order to get this lower bound on $R(k)$.

   For both the upper and lower bounds, better results are known.

2. *Van der Waerden's theorem (henceforth VDW's theorem) [17]:* For all k, c there exists $W = W(k, c)$ such that for all $\mathrm{COL}\colon [W] \to [c]$ there is a monochromatic arithmetic sequence of length k (henceforth a k-AP).

   The *Gallai-Witt theorem* is a generalization to more dimensions. (There is no publication by Gallai that contains it; however, Rado [12],[13] proved it and credited Gallai. Witt [18] proved it independently.)

The *Hales-Jewett theorem (HJ theorem)* [8] is a generalization about coloring sequences. The initial proof of VDW's theorem yielded enormous bounds on $W(k,c)$ that were Ackermann-like. Shelah [15] obtained primitive recursive (though still quite large) bounds on the HJ numbers and hence on $W(k,c)$. His proof is elementary (though difficult). Gowers [7] later obtained a non-elementary proof of bounds you can actually write down.

3. *Szemerédi's theorem [16]* is a density version of VDW's theorem. This proof is elementary in that it only uses combinatorial methods, but the definition of *elementary* is stretched to the breaking point here, since it is quite difficult. There is also a density version of the HJ theorem [10] with an elementary proof.

   There are also non-elementary proofs of Szemerédi's theorem. They may be easier to understand than the elementary proof.

4. *Rado's theorem*: Let $a_1, \ldots, a_n \in \mathsf{Z}$. The following are equivalent:

   - For all finite colorings of N there exists $x_1, \ldots, x_n \in \mathsf{N}$ that are the same color such that $a_1x_1 + \cdots + a_nx_n = 0$.
   - Some subset of $\{a_1, \ldots, a_n\}$ sums to 0.

   There is a generalization of this theorem to systems of linear equations. Some nonlinear equations have also been studied.

3 How the Books are Similar

All of the theorems stated in the last section are covered very well in all of the books; the only exceptions are Szemerédi's theorem and Gowers's theorem, which are stated but not proved. All of the textbooks have good exercises and good explanations. The non-textbook (by Graham & Butler) was not intended as a textbook—it was intended to introduce Ramsey theory to mathematicians. It does a fine job at that.

That takes care of the similarities. In the following, for each book we discuss something unique in it.

4 Rudiments of Ramsey Theory, 2nd ed., Graham and Butler, 1979, 2015

In 1979 Ronald Graham gave a series of lectures at a regional math conference (at St. Olaf College) on Ramsey theory. A set of notes came out of that, which formed the first edition of this book. Over the next 35 years there was much progress in Ramsey theory, perhaps inspired by these notes and the book *Ramsey Theory* by Graham, Rothchild, and Spencer. The book under review is a second edition of these notes. A lot of updating was needed and was done.

These notes are from talks given to other mathematicians to introduce them to the subject. Many of the proofs are informal. This is both good and bad.

- Because of the informality, much ground can be covered in a mere 80 pages. Indeed, there are some topics in this book that are not in the other, longer books, reviewed in this column.

- Some of the proofs are incomplete. Some of the references are to articles that were never published.

I mention a line of research that I learned from this book and was so inspired, I made up slides on it, which I will present to my Ramsey Theory class. Note that the book only had proofs of the first few results, but it pointed me in the right direction for more material.

The line of research is about the following question:

> *Is there an n such that, for all 2-colorings of R^n, there is a monochromatic unit square (side of length 1, all four corners the same color)?*

1. The following result is due to Burr; however, he did not publish it. The result appeared in Erdös et al. [5].

   For all COL: $\mathsf{R}^6 \rightarrow [2]$ there is a monochromatic unit square. The complete proof is in the book. I blogged about the statement and proof and learned what else was known from commenters.

2. For all COL: $\mathsf{R}^5 \rightarrow [2]$ there is a monochromatic unit square. This can be done with a small trick on top of the R^6 result.

3. Kent Cantwell [1] showed the following: For all COL: $\mathsf{R}^4 \rightarrow [2]$ there is a monochromatic unit square. This is a completely different proof.

4. There exists COL: $\mathsf{R}^2 \rightarrow [2]$ with no monochromatic unit square. This is easy.

5. The case of R^3 is open. That's too bad since we live in R^3.

5 Fundamentals of Ramsey Theory by Aaron Robertson, 2021

Most of the results in the early history of Ramsey theory were proved using combinatorial techniques. This was great for having a low barrier to entry. I have seen high school students, and one 10-year old, learn Ramsey theory. By contrast, very few high school students do projects on Functional Analysis.

While it was great to have a modern field of math that was accessible to high school students, eventually non-combinatorial techniques were needed to solve some problems.

So should these techniques be put into a textbook for undergraduates? Can they be?

Robertson's book takes up that challenge. Section 2.2 is titled *Density Theorems.* Chapter 5 is titled *Other Approaches to Ramsey Theorem.* Both contain proofs that use non-combinatorial techniques.

1. Section 2.2: *Roth's Theorem.* If $A \subseteq \mathsf{N}$ has positive upper density, then A has a 3-AP. (Szemerédi later proved that, for all k, A has a k-AP.) This is proved with techniques from analysis, namely Fourier transforms over finite fields. A complete proof is given. This is the way Roth proved it. This Roth's theorem via Fourier transforms is accessible and serves as the starting point for later results in Ramsey theory.

   A combinatorial proof for the $k = 3$ case is also known. It is not in this book. The only account I've ever seen of that is in Graham-Rothchild-Spencer's book *Ramsey Theory.*

2. Section 5.1: A topological proof of VDW's theorem. Note that VDW's theorem has a combinatorial proof (which is in all four books); however, it's good to see a topological proof to get the reader's feet wet with those techniques.

3. Hindman's theorem is as follows: *For all* r*, for all* COL: $\mathsf{Z} \rightarrow [r]$*, there exists an infinite* $A \subseteq \mathsf{Z}$ *and a color* c *such that, for all finite non-empty* $B \subseteq A$*,* $\mathrm{COL}(\sum_{x \in B} x) = c$*.* Earlier in the book (Section 2.4.2) there is a complicated combinatorial proof of Hindman's theorem. In Section 5.1.3 there is a complicated topological proof of Hindman's theorem. In Section 5.3.1 there is a (debatably) simpler topological proof that uses Stone-Čech compactification.

4. In Section 5.4.1. The circle method, a technique from analysis, is used to show that the primes contain an infinite number of 3-term arithmetic progressions. The complete proof is given.

This book is an excellent place to see well-written (aimed at undergraduates) expositions of proofs involving non-combinatorial methods in Ramsey theory.

6 Elementary Graph Ramsey Theory by Li and Lin, 2022

This book begins with the standard topics but then has several chapters on non-standard topics that are not in any of the other books. Note that this is the longest of the books reviewed. Here are just some of the non-standard topics:

1. *Constructive Lower Bounds.* Recall that the lower bound $R(k) \geq \Omega(k2^k)$ was proven by the probabilistic method. Hence the proof does not give a way to construct the needed colorings of $\binom{[n]}{2}$. This chapter investigates constructive lower bounds including the results of Nagy's [9] $R(k) \geq \Omega(k^3)$ (the original paper is in Hungarian) and the Frankl-Wilson [6] constructive bound: $R(k) \geq 2^{\Omega(\log^2 k/\log\log k)}$. As a bonus this chapter contains a disproof of Borsuk's conjecture (for all d every $X \subseteq \mathsf{R}^d$ can be partitioned into $d+1$ sets of smaller dimension) which uses ideas from constructive Ramsey theory.

2. *Communication Channels.* We do an example. Alice and Bob want to communicate with each other. The set of messages they want to send is $\{0,1\}^3$. They are communicating over a noisy channel where 1 bit might get flipped. Hence, if Bob receives 000, the real message is one of $\{000, 001, 010, 100\}$. The *confusion graph* has as vertices the set of messages that can be sent; its edges connect two messages that might be confused for each other. The *Shannon capacity* of such a graph is, roughly speaking, the amount of information that can be transmitted. This 14-page chapter introduces the topic, gives some result, and has 2 pages on how Ramsey theory can be used in the study of Shannon capacity.

3. *Quasi-Random Graphs.* Informally, *quasi-random graphs* (defined by Chung-Graham-Wilson [2]) satisfy many properties that random graphs do. Indeed, the book presents a theorem that has 7 properties of random graphs that are also satisfied by quasi-random graphs. They differ from random graphs in that they can be generated deterministically, and some of their properties are easier to verify. In this chapter they are defined, many theorems are proved about them, and they are used to get better lower bounds on the multicolor Ramsey numbers. For example, $R(C_4, C_4, K_n)$ is the least N such that for all 3-colorings of the edges of K_N with Red, Blue, and Green there is either a Red C_4 or a Blue C_4, or a Green K_n.

4. *Regularity Lemma and Van der Waerden Numbers.* VDW's theorem and the HJ theorem are standard topics, and they are in this chapter. Szemerédi's regularity lemma is important but difficult, so it is not in most other textbook. This lemma is a key ingredient in the proof of Szemerédi's theorem (every $A \subseteq \mathsf{N}$ of positive upper density has arbitrarily long arithmetic sequences). Surprisingly, this book does not mention Szemerédi's theorem.

This book has much material of interest that is not in any other book.

7 Basics of Ramsey Theory by Jungić, 2023

This book covers the standard material very well. The pace is leisurely, and there are many good exercises. We point out three things that this book has that are not standard.

1. There are short biographies of Ramsey, Erdös, van der Waerden, Schur, and Rado. The biographies are interesting and have material the reader probably does not know.

2. The book has a proof of the *canonical Van der Waerden theorem* from *Szemerédi's theorem.* We state both theorems:

   Canonical VDW For all k there exist $W = W(k)$ such that, for all finite coloring of $[W]$ there is *either* a monochromatic k-AP or a rainbow k-AP (all colors different).

   Szemerédi's theorem If A is a set of positive upper density then A has arbitrarily long arithmetic progressions.

   Surprisingly, the book does not mention that there is an elementary proof of the canonical VDW theorem by Prömel & Rodl [11].

3. There is an extensive treatment of the *Happy Ending Theorem*, which we state:

   Happy Ending Theorem For all $n \geq 3$ there exists $f(n)$ such that for all arrangement of $f(n)$ points in the plane, no three colinear, there is a subset of n of them that form a convex n-gon.

   The first proof of this theorem was by Ramsey theory. That proof is in the other books. There were proofs that yielded lower values of $f(n)$ that are also included (the cups-and-caps method). This is unusual.

8 Opinion

We will discuss the four books reviewed here and the three books we reviewed in prior columns. All of the books are good and serve their purpose. We will classify the seven books into two categories.

Books that Mostly Use Combinatorial Techniques

1. *An Introduction to Ramsey Theory: Fast Functions, Infinity, and Metamathematics* by Katz and Reimann. Bonus: has a complete proof that the Paris-Harrington Theorem is independent of Peano Arithmetic.

2. *Ramsey Theory over the Integers (2nd ed.)* by Landman and Robertson. Caveat and bonus: This book does not contain anything about Ramsey theory on graphs but has many variants of VDW's theorem that are not in any other book.

3. *Basics of Ramsey Theory* by Jungić. Bonus: This book has biographies of some Ramsey theorists and an extensive treatment of the Happy Ending Theorem.

4. *Rudiments of Ramsey Theory (2nd ed.)* by Graham & Butler. Bonus and caveat: lots of material here that is not elsewhere, but it is somewhat terse.

Books that Use Non-Combinatorial Techniques

1. *Fundamentals of Ramsey Theory* by Robertson. The point of the book is to present non-combinatorial techniques in a way that can be understood. Bonus: The book succeeds.

2. *Elementary Methods of Graph Ramsey Theory* by Li and Lin. Covers more material than all of the books here except Prömel's. Rough going but it's worth it.

3. *Ramsey Theory for Discrete Structures* by Prömel. This covers a lot of very abstract material and also the elementary proof of the Density HJ theorem. A good place to read rather advanced material.

References

[1] K. Cantwell. Finite Euclidean Ramsey theory. *Journal of Combinatorial Theory A*, 73:273–285, 1996.
`https://www.cs.umd.edu/~gasarch/COURSES/752/S25/slides/R4square.pdf`.

[2] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[3] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math*, 2(4):463–470, 1935.

[4] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1946.

[5] P. Erdős, R. Graham, P. Montgomery, B. L. Rothchild, J. Spencer, and E. G. Straus. Euclidean Ramsey theory I. *Journal of Combinatorial Theory B*, 14:341–363, 1973.

[6] P. Frankl and R. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1:357–368, 1981. `http://www.springer.com/new+%26+forthcoming+titles+%28default%29/journal/493`.

[7] W. Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11:465–588, 2001.

[8] A. Hales and R. Jewett. Regularity and positional games. *Transactions of the American Math Society*, 106, 1963.

[9] Z. Nagy. A constructive estimate of the Ramsey numbers. *Mat. Lapok*, pages 301–302, 1975.

[10] D. Polymath. A new proof of the density Hales-Jewett Theorem. *Annals of Mathematics*, 175:1283–1327, 2012.

[11] H. J. Prömel and V. Rödl. An elementary proof of the canonizing version of Gallai-Witt's theorem. *Journal of Combinatorial Theory, Series A*, 42:144–149, 1986. `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/vdw.html`.

[12] R. Rado. Studien zur Kombinatorik. *Mathematische Zeitschrift*, 36:424–480, 1933. `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/vdw.html`. Includes Gallai's theorem and credits him.

[13] R. Rado. Notes on combinatorial analysis. *Proceedings of the London Math Society*, 48:122–160, 1943. `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/vdw.html`. Includes Gallai's theorem and credits him.

[14] F. Ramsey. On a problem of formal logic. *Proceedings of the London Math Society*, 30(1):264–286, 1930.

[15] S. Shelah. Primitive recursive bounds for van der Waerden numbers. *Journal of the American Math Society*, 1:683–697, 1988.

[16] E. Szeméredi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:299–345, 1975. `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/szdensity.pdf`.

[17] B. van der Waerden. Beweis einer Baudetschen Vermutung (in Dutch). *Nieuw Arch. Wisk.*, 15:212–216, 1927.

[18] E. Witt. Ein kombinatorischer satz de elementargeometrie. *Mathematische Nachrichten*, 6:261–262, 1951. `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/vdw.html`. Contains Gallai-Witt Theorem, though Gallai had it first so it is now called Gallai's theorem.