The Book Review Column[1]
by Frederic Green

Department of Mathematics and Computer Science
Clark University
Worcester, MA 01610
email: `fgreen@clarku.edu`

In this column we review the following books:

1. **Factor Man**, by Matt Ginsberg. A fictional account of what might happen if it turns out that P=NP. Review by William Gasarch.

2. **Kolmogorov Complexity and Algorithmic Randomness**, by A. Shen, V. A. Uspensky, and N. Vereshchagin. A textbook on this important topic. Review by Hadi Shafei.

As always, please contact me to write a review; choose from among the books listed on the next pages, or, if you are interested in anything not on the list, just send me a note.

[1] © Frederic Green, 2019.

BOOKS THAT NEED REVIEWERS FOR THE SIGACT NEWS COLUMN

Algorithms

1. *Tractability: Practical approach to Hard Problems*, Edited by Bordeaux, Hamadi, Kohli
2. *Recent progress in the Boolean Domain*, Edited by Bernd Steinbach
3. *Network Flow Algorithms,* by David P. Williamson.

Computability, Complexity, Logic

1. *The Foundations of Computability Theory,* by Borut Robič
2. *Applied Logic for Computer Scientists: Computational Deduction and Formal Proofs*, by Mauricio Ayala-Rincón and Flávio L.C. de Moura.
3. *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*, by Martin Grohe.
4. *Kernelization: Theory of Parameterized Preprocessing,* by Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi.

Miscellaneous Computer Science

1. *Elements of Causal Inference: Foundations and Learning Algorithms*, by Jonas Peters, Dominik Janzing, and Bernhard Schölkopf.
2. *Elements of Parallel Computing*, by Eric Aubanel
3. *CoCo: The colorful history of Tandy's Underdog Computer* by Boisy Pitre and Bill Loguidice
4. *Introduction to Reversible Computing*, by Kalyan S. Perumalla
5. *A Short Course in Computational Geometry and Topology*, by Herbert Edelsbrunner
6. *Partially Observed Markov Decision Processes,* by Vikram Krishnamurthy
7. *Statistical Modeling and Machine Learning for Molecular Biology*, by Alan Moses
8. *The Problem With Software: Why Smart Engineers Write Bad Code,* by Adam Barr.
9. *Language, Cognition, and Computational Models,* Theirry Poibeau and Aline Villavicencio, eds.
10. *Computational Bayesian Statistics, An Introduction,* by M. Antónia Amaral Turkman, Carlos Daniel Paulino, and Peter Müller.
11. *Variational Bayesian Learning Theory,* by Shinichi Nakajima, Kazuho Watanabe, and Masashi Sugiyama.

Cryptography and Security

1. *Cryptography in Constant Parallel Time,* by Benny Appelbaum
2. *A Cryptography Primer: Secrets and Promises*, by Philip N. Klein

Combinatorics and Graph Theory

1. *Finite Geometry and Combinatorial Applications*, by Simeon Ball
2. *Introduction to Random Graphs*, by Alan Frieze and Michał Karoński
3. *Erdős –Ko–Rado Theorems: Algebraic Approaches*, by Christopher Godsil and Karen Meagher
4. *Combinatorics, Words and Symbolic Dynamics,* Edited by Valérie Berthé and Michel Rigo

Miscellaneous Mathematics

1. *Introduction to Probability*, by David F. Anderson, Timo Seppäläinen, and Benedek Valkó.

Review of[2]
Factor Man
by Matt Ginsberg
Published by Zowie Press 2018
281 pages, $45.00 (Hardcover), $14.00 (Softcover, used)

Review by
William Gasarch gasarch@cs.umd.edu
University of Maryland, College Park, MD

1 Introduction

What would you do if you proved P=NP? If you publish it you might be rich and famous… in a world destroyed by cyber terrorists. If you don't publish it someone might beat you to it. Can you make money off of it?

Matt Ginsberg's fictional book, *Factor Man*, grapples intelligently with this question. The premise is that someone (alias Factor Man) has proven P=NP and has a pretty good plan for how to (1) cash in, and (2) not create economic havoc.

2 How Real is the Book?

The book takes place in the real world of today. How real? Some real theorists are characters. My poll on P vs NP is mentioned. Sylvester Stallone wants to know the factors of

$$4731909729799585432707038822764050445110579699728976599571915 53$$

Well, maybe that last item is not likely to happen in our world. But in a world where Factor Man invites people to submit numbers that he can factor, to prove he has P=NP (though fictional theorists points out, as real ones would, that this only helps verify that Factor Man can, uh, factor), I can very much imagine that celebrities would get into the game. More generally, all of the characters act like I think they would act. This book is *not* satire. This book is a realistic speculation of what would happen if someone showed P=NP.

3 How is it as a Book to, You Know, Read?

I started this book on a Saturday and finished it on a Sunday. And I am not speed reader. The pacing is good and it's never dull. You begin reading it and you want to see what will happen next. The beginning, middle, and end are all satisfying (you are invited to insert a joke about satisfiability here). As they said at the website Goodreads, it's a Good Read.

I was asked if it's a thriller or a mystery or science fiction or what. I find it hard to classify; however, I would call it both a thriller and science fiction.

[2] ©2019, William Gasarch

4 Is the book Technically Correct?

I've seen some reviews with some (correct) quibbles; however, where it matters, the book is spot on. None of the technical errors have any bearing on the plot.

Getting the math right is especially impressive in light of other attempts. I'll just give a pointer to one. There was an episode of *Elementary* where it seems as though P=NP has been established. How much did the episode get wrong? It might be easier to tell you how much they got right. See my blog post about it:
`https://blog.computationalcomplexity.org/2013/10/p-vs-np-is-elementary-no-p-vs-np-`
`html`

5 Who Will Enjoy this Book?

Anyone who has even a passing acquaintance with P vs NP will enjoy this book. The book also has some explanation of P and NP for those who do not know the problem. I tend to think that if one does not know the problem, best to skip those sections and just know that if P=NP then (1) all known crypto systems can be broken, (2) given enough warning people can change those systems to hopefully not be broken, (3) lots of awesome science and other great technical achievements will be possible, though one would have to take the proof that P=NP and modify it to be practical.

Review of[3]
Kolmogorov Complexity and Algorithmic Randomness
by A. Shen, V. A. Uspensky, and N. Vereshchagin
AMS, 2017
511 pages, $124.00 ($99.20, AMS Member Price; Hardcover)

Review by
Hadi Shafei (hshafei@nmu.edu)
Department of Mathematics and Computer Science
Northern Michigan University

1 Introduction

This is a book about Kolmogorov complexity and algorithmic randomness. The main idea in Kolmogorov complexity, introduced by A. N. Kolmogorov and others in the 1960's, is to use an algorithmic approach to measure the amount of information in finite objects [1]. Kolmogorov wanted to create an information theory that, unlike Shannon information theory, is founded independently from probability theory. This approach results in a notion of randomness that is independent from the notion of probability and is closer to the intuition that "randomness is the absence of regularities" [2, 1].

As authors stated in the introduction, this book is not intended to be comprehensive, especially with respect to the more recent results; However, the authors cover a wide range of topics and results in Kolmogorov complexity theory and whenever a result is not rigorously investigated, sufficient references are provided. The book starts from the basic concepts and builds its way up toward the more advanced concepts and techniques. The progress from the introductory to the advanced topics is seamless which makes the reading an enjoyable experience. Concepts and techniques are explained very well and proofs are rigorous, detailed, and clear. Important and complicated proofs are usually followed by excellent discussions about the main ideas and techniques used in them. The book also contains many interesting and insightful philosophical discussions.

2 Summary of Contents

2.1 Introduction

The main idea of Kolmogorov Complexity theory and its formal definition using optimal description modes are covered in the introduction. A description mode (a decompressor) is simply a computable function $f : \{0,1\}^* \to \{0,1\}^*$. If $f(x) = y$ then we say x is a description of y with respect to f. The Kolmogorov complexity of a string y with respect to f, denoted as $C_f(y)$, is defined as the length of the shortest string x such that $f(x) = y$, i.e., the length of the shortest description of y with respect to f. It turns out that, up to an additive constant, we can find an optimal description mode. We fix an optimal description mode and define $C(x)$ to be the Kolmogorov complexity of x with respect to this fixed description mode. Some basic properties of Kolmogorov complexity are proved in the introduction which gives the reader a flavor of the topic. The connection between Kolmogorov complexity and randomness is also discussed briefly in the introduction. More interesting properties like non-computability of Kolmogorov complexity and applications like a proof of Gödel's incompleteness theorem using Kolmogorov complexity are covered in the introduction as well. The introduction is designed masterfully such that it gives the user a flavor of ideas and techniques that will appear frequently in the subsequent chapters and also shows applications of Kolmogorov complexity in areas like probability theory, combinatorics, and logic that at first may seem unrelated to Kolmogorov complexity.

2.2 Chapters 1 and 2

In chapter 1, plain Kolmogorov complexity is defined and some of its basic properties are proved. Some of these definitions are based on concepts like enumerable sets and semicomputable functions from computability theory. In these cases, the prerequisites from computability theory are provided and explained thoroughly.

Chapter 2 contains more properties of plain Kolmogorov complexity including properties related to complexity of pairs and conditional complexity. There are a lot of properties related to the complexity of tuples and conditional complexity and the authors decided to express many of them as exercises. These exercises are organized such that there are problems at different levels of difficulty and whenever needed, hints are given. For some of the more important theorems, alternative perspectives are given. For example, a combinatorial interpretation of the following statement, known as Kolmogorov-Levin Theorem is discussed:

Theorem 2.1 *For all strings x, y of length at most n,* $C(x, y) = C(x) + C(y|x) + O(\log n)$

Along the same lines, a game theoretic argument for the following statement is provided:

Proposition 2.2 *For every n there exists an n-bit string x such that* $C(C(x)|x) = \log n - O(1)$.

Needless to say that seeing the same problem from different angles gives the reader a much better understanding of the topic.

2.3 Chapter 3

Effectively null sets and their properties are the subject of chapter 3. The concept of Martin-Löf randomness and its connection to effectively null sets is discussed. Some properties of Martin-Löf random sequences are explored in this chapter, including the fact that after changing a finite number of bits (or flipping zeros and ones) in a sequence that is Martin-Löf random, the sequence remains Martin-Löf random.

2.4 Chapter 4

A different notion of complexity, namely prefix complexity, is discussed in this chapter. Prefix complexity can be defined in two different ways:

- By using optimal prefix-stable decompressors, i.e. functions like f where if $f(x)$ is defined for some string x, then $f(y)$ is defined for every string y where x is a prefix of y and $f(x) = f(y)$.
- By using prefix-free decompressors, i.e. functions like f where if $f(x)$ is defined for some string x, then f is not defined on any extension of x.

The main result in this chapter is that these two definitions of prefix complexity and the negative logarithm of the a priori probability coincide up to an additive constant [2]. As mentioned in the introduction, the discussions before or after long and/or complicated proofs in this book are priceless. In this case, after finishing a detailed proof of the aforementioned result, the authors provide a short summary of the main ideas in the proof which helps the reader see the structure of the proof that might be hidden behind all the detail. After proving the equivalence of two definitions of prefix complexity, properties of prefix complexity are explored. When proving these properties, the authors go back and forth between these two definitions, sometimes giving two proofs for a property, which helps the reader gain a solid understanding of the concept and the proof techniques involved.

2.5 Chapter 5

Two more notions of complexity are defined in this chapter: a priori complexity and monotone complexity. For readers who are just learning Kolmogorov complexity, by now there have been too many different notions of complexity. To make matters worse, sometimes the difference between these notions is very subtle. One can find the definitions and theorems in any textbook, but what makes this book stand out is the discussions on comparisons between these different notions and consequences of these differences and similarities.

2.6 Chapter 6

In chapter 6, plain complexity, prefix complexity, and monotone complexity along with decision complexity, which is defined in this chapter, are compared in an interesting way. Using a topological point of view, the authors define a parameterized version of description modes such that any of the complexity notions above can be defined by changing the parameters. This classification helps to see the big picture and puts the definitions of these four complexities in context.

2.7 Chapter 7

Shannon entropy and its properties are introduced in this chapter. Related subjects like entropy of pairs, conditional entropy, and mutual information are investigated. The connection between Kolmogorov complexity and Shannon entropy is discussed. In particular, the following bound is proved:

Theorem 2.3 *Let x be a string of length N and let $p_1, ..., p_k$ be the frequencies of the letters in x. Then:*

$$\frac{C(x)}{N} \leq h(p_1, ..., p_k) + \frac{O(\log N)}{N}$$

where $h(p_1, ..., p_k)$ is the entropy of the distribution $p_1, ..., p_k$.

2.8 Chapter 8

In this chapter, we see how Kolmogorov complexity can be used to prove results in several areas of mathematics including number theory, real analysis, automata theory, analysis of Turing machines, combinatorics, and probability theory. Here are some of these results:

- Euclid's theorem that there are infinitely many primes.
- Duplicating an n-bit string on the tape of a one-tape Turing machine requires ϵn^2 steps in the worst case.
- The Loomis-Whitney inequality.
- No Lipschitz mapping can be Besicovitch-transitive.

Given the fact that Kolmogorov complexity started as an algorithmic approach for measuring the amount of information in strings, it is very interesting to see how diverse its applications are.

2.9 Chapter 9

Richard von Mises defined probability theory based on a notion of random sequences that he called *Kollektivs*. The main property of these random sequences is *frequency stability* which means the limit frequency of ones in the sequence exists and remains the same for subsequences selected by some rule [2]. The problem with this approach is that von Mises did not explicitly say what kind of selections are admissible. Depending on the definition of admissible selection rules, we get different notions of Kollektivs. This chapter explores different possibilities for selection rules and the resulting notions of randomness. Examples include: Mises-Church randomness, martingales, semicomputable martingales, computable martingales, partial selection rules, and non-monotonic selection rules. All these definitions and their properties are investigated in different sections. A very insightful discussion in the last section ties all these concepts together.

2.10 Chapter 10

The main approaches to quantifying information are combinatorial, probabilistic, and algorithmic. The relationship between these approaches is explored in more depth in this chapter. The main focus is on linear inequalities about entropy and complexity. For each inequality, all approaches are examined. For some inequalities, multiple interpretations are given with respect to each approach.

2.11 Chapter 11

In this chapter the authors explore mutual information in more depth. More specifically, they look at questions like:

- Given a string x, is there an incompressible string x' that has the same information as x?
- Given a string x, can we divide the information in x into two equal parts in strings x_1 and x_2?
- Given strings x and y, is there an incompressible string z that has the same information as the pair $\langle x, y \rangle$?

2.12 Chapter 12

Transmission of information in a network has been studied extensively from the classical Shannon approach where source nodes are considered as random variables and the goal is to find conditions that make a given transition request possible. This chapter addresses the same problem from the algorithmic information theory perspective. From this point of view, input nodes receive outside information and every node processes the information it receives and transmits it into some other nodes.

2.13 Chapter 13

The complexity of a set of strings, also known as a problem, is defined as the minimal complexity of its elements. Conjunction and disjunction of two problems can be defined as the product and disjoint union of the sets, respectively. It turns out that there is an interesting connection between the complexity of problems and provability of formulas in Intuitionistic Propositional Calculus (IPC). On one hand, if a formula $\Phi(p, q, ...)$ is provable in IPC, then the complexity of the problem $\Phi(X, Y, ...)$ is bounded by a constant (depending only on Φ and not on X, Y,...). On the other hand, the reverse statement is also true: if $\Phi(p, q, ...)$ is a propositional formula with no negation that is not provable in IPC, then the complexity of $\Phi(X, Y, ...)$ is not bounded.

2.14 Chapter 14

This chapter is about algorithmic statistics. The problem here is that given a string x, we want to find a finite set A containing x where A is a reasonable explanation for x. This means A has small Kolmogorov complexity and x is a typical element of A.

3 Opinion

The intended goal of this book, stated by the authors in the introduction, is to provide clear explanations of the most important topics and results in the field, and they definitely achieved this goal. Every definition or concept is followed by a discussion about the main idea behind it. Proofs are very well-explained and detailed. Longer proofs are usually followed by short summaries of the main ideas so that the trees do not prevent the readers from seeing the forest. Kolmogorov complexity theory is closely related to probability theory, computability theory, and combinatorics. Kolmogorov, in his first paper about the notion of complexity called "Three approaches to the quantitative definition of complexity," mentions three main approaches: combinatorial, algorithmic, and probabilistic. Following this idea, in many cases we see a theorem in this book that has multiple proofs where each proof is based on ideas and techniques from algorithmic complexity, probability theory, combinatorics, or geometry. This provides the reader the complete picture of Kolmogorov complexity as envisioned by Kolmogorov. The authors did a great job with the philosophical discussion in the book providing a more in-depth insight to the concepts. Having a rigorous proof in full detail (sometimes multiple proofs) followed by an intuitive discussion about the ideas and techniques used in the proof, and a philosophical discussion, results in a thorough understanding of the topic.

The authors cover a wide range of topics and results in this book, including almost all standard results in the field. The only exceptions, as stated by the authors in the introduction of the book, are historical remarks and more recent results. The clarity of proofs and discussions in this book is remarkable. No prior knowledge of Kolmogorov complexity or even computability theory is assumed by the authors as they define and explain every concept they use; however, to be able to follow the arguments, some mathematical sophistication is required. In particular, prior knowledge on topics like enumerable sets, computable functions, and Turing machines, which are typically covered in an undergraduate course on computability and complexity theory, can be useful. In every chapter unproved properties are stated. These are problems of different levels of difficulty and hints are provided for many of them. In my opinion, reading this book will be an enjoyable and instructive experience for anyone who is interested in learning Kolmogorov complexity. Moreover, this book is a great candidate for a graduate textbook on Kolmogorov complexity.

In Memorium: In the course of preparing this review, we were saddened to learn of the passing of one of the book's authors, Vladimir Uspensky (himself a student of Kolmogorov), on June 27th, 2018.

References

[1] Andrei Nikolaevich Kolmogorov. Three approaches to the quantitative definition of information. *International journal of computer mathematics*, 2(1-4):157–168, 1968.

[2] Alexander Shen, Vladimir A. Uspensky, and and Nikolay Vereshchagin. Kolmogorov complexity and algorithmic randomness. *Kolmogorov complexity and algorithmic randomness*, volume 220. American Mathematical Soc., 2017.