AUTOMATIC SEQUENCES

Uniting dozens of disparate results from different fields, this book combines concepts from mathematics and computer science to present the first integrated treatment of sequences generated by the simple model of computation called the finite automaton.

The authors develop the theory of automatic sequences and their generalizations, such as Sturmian words and k-regular sequences. Further, they discuss applications to number theory (particularly formal power series and transcendence in finite characteristic), physics, computer graphics, and music.

Results are presented from first principles wherever feasible, and the book is supplemented by a collection of 460 exercises, 85 open problems, and over 1600 citations to the literature. Thus this book is suitable for graduate students or advanced undergraduates, as well as for mature researchers wishing to know more about this fascinating subject.

Jean-Paul Allouche is Directeur de Recherche at CNRS, LRI, Orsay. He has written some 90 papers in number theory and combinatorics on words. He is on the editorial board of *Advances in Applied Mathematics* and on the scientific committee of the *Journal de Théorie des Nombres de Bordeaux*.

Jeffrey Shallit is Professor of Computer Science at the University of Waterloo. He has written 80 articles on number theory, algorithms, formal languages, combinatorics on words, computer graphics, history of mathematics, algebra and automata theory. He is the editor-in-chief of the *Journal of Integer Sequences* and coauthor of *Algorithmic Number Theory*.

AUTOMATIC SEQUENCES

Theory, Applications, Generalizations

JEAN-PAUL ALLOUCHE
CNRS, LSI, Orsay

JEFFREY SHALLIT
University of Waterloo

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United States of America

Typeface Times 11/14 pt. *System* LaTeX 2_{ε} [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data
Allouche, Jean-Paul, 1953-
Automatic sequences : theory, applications, generalizations / Jean-Paul Allouche, Jeffrey Shallit.
p. cm.
Includes bibliographical references and index.
ISBN 0-521-82332-3
1. Sequential machine theory. 2. Sequences (Mathematics) I. Shallit, Jeffrey Outlaw. II. Title.
QA267.5.S4 A55 2003
515′.24 – dc21 2002041262

ISBN 0 521 82332 3 hardback

Nous dédions ce livre à Michel Mendès France
en signe de gratitude et d'amitié

Cambridge University Press
0521823323 - Automatic Sequences: Theory, Applications, Generalizations
Jean-Paul Allouche and Jeffrey Shallit
Frontmatter
More information

Contents

Preface

Goals of This Book

Sequences, both finite and infinite, are ubiquitous in mathematics and theoretical computer science. Sloane and Plouffe's book, *The Encyclopedia of Integer Sequences*, lists over 5,000 interesting sequences from the mathematical literature. Sloane's web site,

`http://www.research.att.com/~njas/sequences/index.html`

gives access to more than 69,000 sequences. There is a web-based scholarly journal, the *Journal of Integer Sequences*, devoted to sequence-related topics, and even a periodic international conference, SETA (Sequences and Their Applications), devoted to the study of sequences.

Sequences come in all flavors. Some, such as periodic sequences, are highly ordered and very easy to describe, while others, such as random sequences, are unordered and have no simple description.

The subject of this book is *automatic sequences* and their generalizations. Automatic sequences form a class of sequences somewhere between simple order and chaotic disorder. This class contains such celebrated sequences as the Thue–Morse sequence (see Chapters 1 and 6) and the Rudin–Shapiro sequence (see Chapter 3), which play important roles in many different areas of mathematics.

Automatic sequences are generated by finite automata, one of the most basic models of computation. Finite automata and other computational models are introduced in Chapter 5. Automatic sequences are also generated by iterating a simple kind of map, called a uniform morphism; see Chapter 6. By generalizing this to arbitrary morphisms, we obtain another interesting class of sequences called morphic sequences, which are discussed in Chapter 7.

Other generalizations of automatic sequences discussed in this book include multidimensional sequences (Chapter 14), sequences over infinite alphabets (Chapter 16), and sequences that are generated by slowly growing automata (Chapter 15).

One of the main reasons to study automatic sequences and their generalizations is the large number of interesting connections with number theory. To cite the most prominent example, methods of automata theory have recently been applied to prove new results in transcendence theory in positive characteristic; see Chapter 12.

While hundreds of papers discussing the relationship between automata theory and number theory have appeared in the literature, up to now there has been little attempt to bring these results together in any sort of consistent framework, using a unified notation. Books on automata theory rarely discuss results with a number-theoretic flavor, and when they do, these applications are often relegated to footnotes. On the other hand, the techniques of theoretical computer science are rarely incorporated in books on number theory, since they require unfamiliar language and notation.

Because our subject incorporates results from both mathematics and computer science, papers are scattered widely in the literature and often use inconsistent notation. Sometimes important results have appeared in obscure journals or remained unpublished because they did not find a home in more mainstream journals devoted to pure mathematics or theoretical computer science. Furthermore, since many researchers in the area are French, some important results have appeared only in the French language, making them less accessible to non-Francophones. Many of these results appear in this book in English for the first time.

We have attempted to present the material in as self-contained a way as feasible. Unfortunately, some results, such as Roth's theorem and Ridout's theorem, require rather detailed and complicated proofs, and we have chosen to omit the proofs.

Since this book is intended as an introduction, we do not always present results in the most general possible formulation. For example, in Chapter 9 we focus on characteristic words, and do not prove many theorems on the more general case of Sturmian words. Sometimes results are presented largely for their illustrative and pedagogical value. In particular, material in this book intersects with symbolic dynamics and ergodic theory, as well as other fields, but this book is not intended to be an introduction to those fields.

Each chapter ends with sections entitled Exercises and Notes. Some exercises are very easy, while the solution of others is a significant accomplishment. (Indeed, researchers should not be insulted if they find their own favorite results listed as exercises.) Exercises are arranged more or less randomly, with order having no implication for difficulty. Hints and solutions to selected exercises, as well as references, can be found in the Appendix. The Notes sections provide the reader with a detailed set of over 1600 references to pursue further work. Finally, dozens of unsolved research problems are listed in a section of each chapter entitled Open Problems.

Prerequisites

We hope the material in this book will be useful to readers at many levels, from advanced undergraduates to experts in the area. Experts may want to turn immediately to the new results in Chapters 12 and 13, for example, while novices may first need the background material in Chapters 1–5.

The main prerequisite is a degree of mathematical sophistication. Familiarity with the basic concepts of formal languages and number theory will be useful, but not absolutely essential. We have attempted to make the material self-contained wherever feasible.

A typical graduate or advanced undergraduate course on formal languages and number theory might cover the material in Chapters 1–6, 9, 12, and 13. Readers familiar with number theory and algebra might plan to skip Chapter 2, while those familiar with theoretical computer science would skip Chapter 4.

Algorithm Descriptions

Algorithms in this book are described in a pseudocode notation similar to Pascal or C, which should be familiar to most readers. We do not provide a formal definition of this notation.

Acknowledgments

We learned much of this material from Michel Mendès France, our good friend and colleague. We also benefited greatly from conversations with Jean Berstel, James Currie, David Damanik, Will Gilbert, and Luca Zamboni.

Drew Vandeth typed a very early version of the manuscript, suggested many improvements and changes, drew some of the diagrams, and wrote solutions to some of the exercises. Jennifer Keir provided expert TEX advice. Andy Poon verified many of the references, and found many typographical errors. Eric Bach, Céline Barberye, Jean Berstel, Valérie Berthé, Richard Crandall, Larry Cummings, Michael Domaratzki, Anna Frid, Marton Kósa, Bryan Krawetz, Scott Lindhurst, Andrew Martinez, Michel Mendès France, Simon Plouffe, Joe Rideout, Patrice Séébold, Gentcho Skordev, Troy Vasiga, Ming-wei Wang, David Wilson, and David Yeung were kind enough to read drafts of this book; they pointed out many errors and suggested many improvements. We are grateful to all of them. We also owe a huge debt to Jia-Yan Yao, who performed a Herculean task by reading every chapter with great care and sending us 67 pages of corrections. Of course, any errors that remain are our responsibility. Doug Bowman generously shared with us some of his unpublished results for use as exercises. We also thank Christophe Genolini for his assistance in creating Figures 17.1 and 17.2. Finally, we thank Mikael Kristiansen and Thor Bak for sharing their knowledge about the work of composer Per Nørgård with us.

The authors would appreciate hearing about any errors, no matter how trivial. Current errata for the book can be found in Shallit's home page;

`http://www.math.uwaterloo.ca/~shallit/`

is the URL.

This work was supported in part by grants from the Natural Sciences and Engineering Research Council (Canada). Part of this book was prepared during several visits of the first author to the University of Waterloo; he thanks that institution for its assistance. Part of this book was prepared while the second author enjoyed the kind hospitality of the Departments of Mathematics and Computer Science at the University of Arizona; he thanks John Brillhart and Peter Downey for their assistance.

Jean-Paul Allouche
Paris, France; September 2002
Jeffrey Shallit
Waterloo, Ontario; September 2002