

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

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*.

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

AUTOMATIC SEQUENCES

Theory, Applications, Generalizations

JEAN-PAUL ALLOUCHE

CNRS, LSI, Orsay

JEFFREY SHALLIT

University of Waterloo



CAMBRIDGE
UNIVERSITY PRESS

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

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>

© Jean-Paul Allouche Jeffrey Shallit 2003

This book is in copyright. Subject to statutory exception
 and to the provisions of relevant collective licensing agreements,
 no reproduction of any part may take place without
 the written permission of Cambridge University Press.

First published 2003

Printed in the United States of America

Typeface Times 11/14 pt. *System* L^AT_EX 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

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

*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

<i>Preface</i>	<i>page</i>	xiii
1 Stringology		1
1.1 Words		1
1.2 Topology and Measure		5
1.3 Languages and Regular Expressions		7
1.4 Morphisms		8
1.5 The Theorems of Lyndon and Schützenberger		10
1.6 Repetitions in Words		14
1.7 Overlap-Free Binary Words		16
1.8 Additional Topics on Repetitions		23
1.9 Exercises		24
1.10 Open Problems		30
1.11 Notes on Chapter 1		31
2 Number Theory and Algebra		39
2.1 Divisibility and Valuations		39
2.2 Rational and Irrational Numbers		39
2.3 Algebraic and Transcendental Numbers		41
2.4 Continued Fractions		44
2.5 Basics of Diophantine Approximation		48
2.6 The Three-Distance Theorem		53
2.7 Algebraic Structures		55
2.8 Vector Spaces		56
2.9 Fields		56
2.10 Polynomials, Rational Functions, and Formal Power Series		58
2.11 p -adic Numbers		62
2.12 Asymptotic Notation		63
2.13 Some Useful Estimates		63
2.14 Exercises		64
2.15 Open Problems		67
2.16 Notes on Chapter 2		67

3	Numeration Systems	70
3.1	Numeration Systems	70
3.2	Sums of Digits	74
3.3	Block Counting and Digital Sequences	77
3.4	Representation of Real Numbers	84
3.5	Sums of Sums of Digits	86
3.6	Base- k Representation with Alternate Digit Sets	101
3.7	Representations in Negative Bases	103
3.8	Fibonacci Representation	105
3.9	Ostrowski's α -Numeration System	106
3.10	Representations in Complex Bases	107
3.11	Exercises	112
3.12	Open Problems	118
3.13	Notes on Chapter 3	119
4	Finite Automata and Other Models of Computation	128
4.1	Finite Automata	128
4.2	Proving Languages Nonregular	136
4.3	Finite Automata with Output	137
4.4	Context-Free Grammars and Languages	143
4.5	Context-Sensitive Grammars and Languages	146
4.6	Turing Machines	146
4.7	Exercises	148
4.8	Open Problems	150
4.9	Notes on Chapter 4	150
5	Automatic Sequences	152
5.1	Automatic Sequences	152
5.2	Robustness of the Automatic Sequence Concept	157
5.3	Two-Sided Automatic Sequences	161
5.4	Basic Properties of Automatic Sequences	165
5.5	Nonautomatic Sequences	166
5.6	k -Automatic Sets	168
5.7	1-Automatic Sequences	169
5.8	Exercises	170
5.9	Open Problems	171
5.10	Notes on Chapter 5	171
6	Uniform Morphisms and Automatic Sequences	173
6.1	Fixed Points of Uniform Morphisms	173
6.2	The Thue–Morse Infinite Word	173
6.3	Cobham's Theorem	174
6.4	The Tower of Hanoi and Iterated Morphisms	177
6.5	Paperfolding and Continued Fractions	181
6.6	The k -Kernel	185

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

<i>Contents</i>		ix
6.7	Cobham's Theorem for (k, l) -Numeration Systems	187
6.8	Basic Closure Properties	189
6.9	Uniform Transduction of Automatic Sequences	192
6.10	Sums of Digits, Polynomials, and Automatic Sequences	197
6.11	Exercises	201
6.12	Open Problems	207
6.13	Notes on Chapter 6	208
7	Morphic Sequences	212
7.1	The Infinite Fibonacci Word	212
7.2	Finite Fixed Points	213
7.3	Morphic Sequences and Infinite Fixed Points	215
7.4	Two-Sided Infinite Fixed Points	218
7.5	More on Infinite Fixed Points	226
7.6	Closure Properties	228
7.7	Morphic Images of Morphic Words	231
7.8	Locally Catenative Sequences	237
7.9	Transductions of Morphic Sequences	240
7.10	Exercises	242
7.11	Open Problems	244
7.12	Notes on Chapter 7	245
8	Frequency of Letters	247
8.1	Some Examples	247
8.2	The Incidence Matrix Associated with a Morphism	248
8.3	Some Results on Non-negative Matrices	249
8.4	Frequencies of Letters and Words in a Morphic Sequence	266
8.5	An Application	276
8.6	Gaps	278
8.7	Exercises	280
8.8	Open Problems	282
8.9	Notes	282
9	Characteristic Words	283
9.1	Definitions and Basic Properties	283
9.2	Geometric Interpretation of Characteristic Words	290
9.3	Application: Unusual Continued Fractions	291
9.4	Exercises	293
9.5	Open Problems	295
9.6	Notes on Chapter 9	295
10	Subwords	298
10.1	Introduction	298
10.2	Basic Properties of Subword Complexity	300
10.3	Results for Automatic Sequences	304
10.4	Subword Complexity for Morphic Words	306

x	<i>Contents</i>	
	10.5 Sturmian Words	312
	10.6 Sturmian Words and k th-Power-Freeness	320
	10.7 Subword Complexity of Finite Words	323
	10.8 Recurrence	324
	10.9 Uniform Recurrence	328
	10.10 Appearance	333
	10.11 Exercises	334
	10.12 Open Problems	340
	10.13 Notes on Chapter 10	340
11	Cobham's Theorem	345
	11.1 Syndetic and Right Dense Sets	345
	11.2 Proof of Cobham's Theorem	347
	11.3 Exercises	350
	11.4 Notes on Chapter 11	350
12	Formal Power Series	351
	12.1 Examples	352
	12.2 Christol's Theorem	354
	12.3 First Application to Transcendence Results	359
	12.4 Formal Laurent Power Series and Carlitz Functions	359
	12.5 Transcendence of Values of the Carlitz–Goss Gamma Function	362
	12.6 Application to Transcendence Proofs over $\mathbb{Q}(X)$	365
	12.7 Furstenberg's Theorem	367
	12.8 Exercises	371
	12.9 Open Problems	375
	12.10 Notes on Chapter 12	376
13	Automatic Real Numbers	379
	13.1 Basic Properties of the Automatic Reals	379
	13.2 Non-closure Properties of $L(k, b)$	382
	13.3 Transcendence: An Ad Hoc Approach	385
	13.4 Transcendence of the Thue–Morse Number	387
	13.5 Transcendence of Morphic Real Numbers	391
	13.6 Transcendence of Characteristic Real Numbers	393
	13.7 The Thue–Morse Continued Fraction	394
	13.8 Exercises	400
	13.9 Open Problems	402
	13.10 Notes on Chapter 13	403
14	Multidimensional Automatic Sequences	405
	14.1 The Sierpiński Carpet	405
	14.2 Formal Definitions and Basic Results	408
	14.3 Subword Complexity	412
	14.4 Formal Power Series	413
	14.5 Automatic Sequences in Base $-1 + i$	414

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

<i>Contents</i>		xi
14.6	The Pascal Triangle Modulo d	420
14.7	Exercises	424
14.8	Open Problems	425
14.9	Notes on Chapter 14	426
15	Automaticity	428
15.1	Basic Notions	428
15.2	Nondeterministic Automaticity	431
15.3	Unary Automaticity	433
15.4	Automaticity of Sequences	434
15.5	Exercises	436
15.6	Open Problems	436
15.7	Notes on Chapter 15	437
16	k -Regular Sequences	438
16.1	Basics	438
16.2	Robustness of the k -Regular Concept	441
16.3	Further Results	444
16.4	k -Regular Power Series	445
16.5	Additional Examples	447
16.6	Exercises	449
16.7	Open Problems	453
16.8	Notes on Chapter 16	454
17	Physics	455
17.1	The One-Dimensional Ising Model	457
17.2	The Rudin–Shapiro Sequence and the One-Dimensional Ising Model	459
17.3	Distribution Results for the Rudin–Shapiro Sequence	462
17.4	The One-Dimensional Schrödinger Operator	464
17.5	Exercises	466
17.6	Notes on Chapter 17	467
Appendix		
Hints, References, and Solutions for Selected Exercises		471
A.1	Chapter 1	471
A.2	Chapter 2	472
A.3	Chapter 3	472
A.4	Chapter 4	474
A.5	Chapter 5	474
A.6	Chapter 6	474
A.7	Chapter 7	475
A.8	Chapter 8	476
A.9	Chapter 9	476
A.10	Chapter 10	476
A.11	Chapter 11	477

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

xii	<i>Contents</i>	
A.12	Chapter 12	477
A.13	Chapter 13	477
A.14	Chapter 14	478
A.15	Chapter 15	478
A.16	Chapter 16	478
A.17	Chapter 17	479
	<i>Bibliography</i>	481
	<i>Index</i>	555

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.

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)*Preface*

xv

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 $\text{T}_{\text{E}}\text{X}$ advice. Andy Poon verified many of the references, and found many typographical errors. Eric Bach, Céline Barbery, 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

Cambridge University Press

0521823323 - Automatic Sequences: Theory, Applications, Generalizations

Jean-Paul Allouche and Jeffrey Shallit

Frontmatter

[More information](#)

xvi

Preface

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