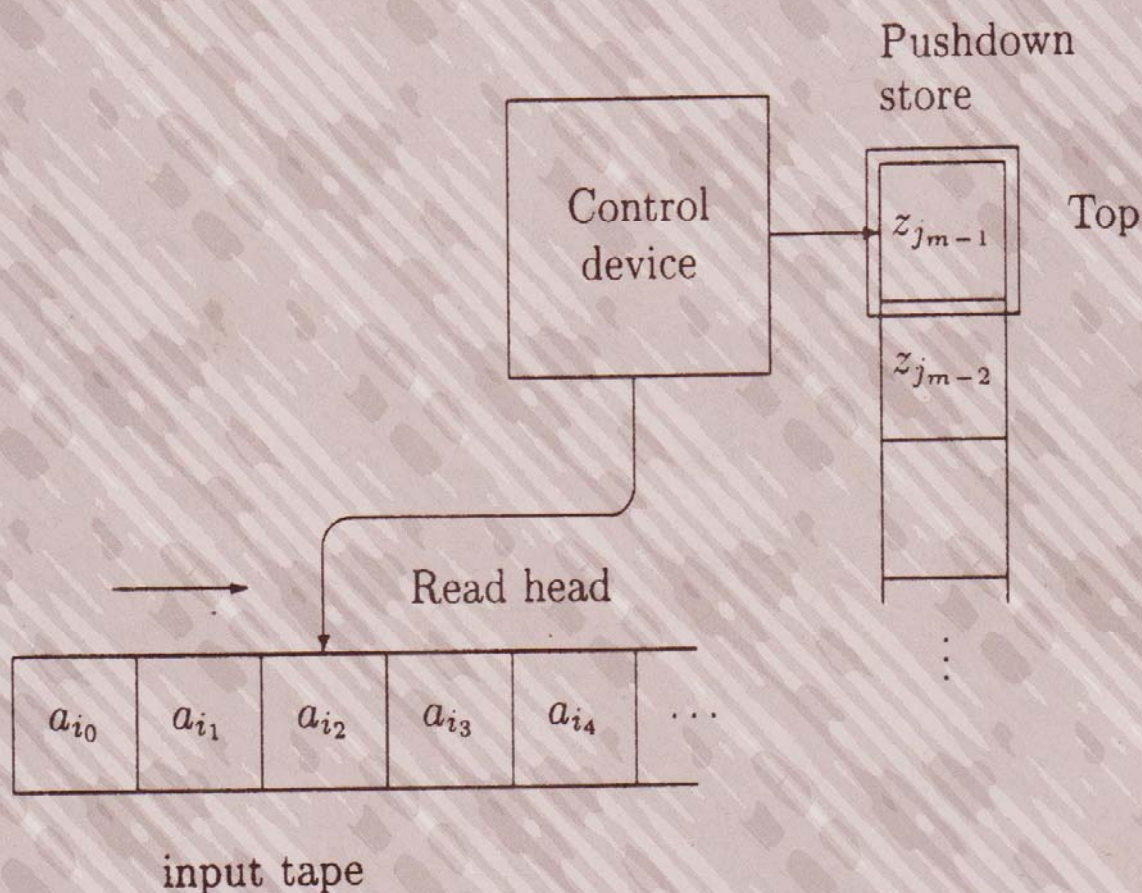


THEORY OF FORMAL LANGUAGES WITH APPLICATIONS

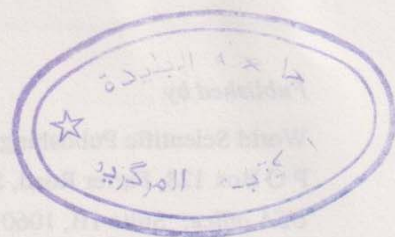


Dan A Simovici
Richard L Tenney

World Scientific

2-005-516-1

2-005-516-1



THEORY OF FORMAL LANGUAGES WITH APPLICATIONS

Dan A Simovici

Richard L Tenney

*Department of Mathematics and Computer Science,
University of Massachusetts at Boston*



World Scientific

Singapore • New Jersey • London • Hong Kong

Contents

Preface	ix
Introduction	xi
I Introductory Notions	1
1 Preliminaries	3
1.1 Introduction	3
1.2 Sets, Relations, and Functions	3
1.2.1 Sets	3
1.2.2 Ordered Pairs and Cartesian Products	4
1.2.3 Relations	6
1.2.4 Equivalence Relations	9
1.2.5 Partial Orders	11
1.2.6 Functions	12
1.3 Operations and Algebras	16
1.3.1 Operations	17
1.3.2 Algebras, Semigroups, and Monoids	19
1.3.3 Morphisms and Subalgebras	21
1.3.4 Congruences	22
1.4 Sequences	24
1.4.1 The Monoid of Sequences	26
1.4.2 Arithmetic Progressions	29
1.5 Graphs	30
1.6 Cardinality	37
1.7 Exercises	45
1.8 Bibliographical Comments	55
2 Words and Languages	57
2.1 Introduction	57
2.2 Words	57
2.3 Languages	60
2.4 Substitutions and Morphisms	65
2.5 Matrices and Languages	67
2.6 Polynomial Functions	71

2.7	Exercises	82
2.8	Bibliographical Comments	92
II Regular and Context-Free Languages		95
3	Regular Languages	97
3.1	Introduction	97
3.2	Finite Automata	98
3.2.1	Deterministic Automata	98
3.2.2	Nondeterministic Automata	107
3.2.3	Configurations	114
3.3	Transition Systems	116
3.4	Closure Properties	122
3.5	The Pumping Lemma	128
3.6	Minimal Automata	132
3.7	Syntactic Monoids	136
3.7.1	Automata and Monoids	137
3.7.2	The Syntactic Monoid of a Language	139
3.8	Fixed Points and Regular Languages	141
3.9	Regular Expressions	147
3.9.1	The Unique Readability of Regular Expressions	147
3.9.2	Regular Expressions as Notations for Regular Languages	150
3.9.3	Closure Properties and Regular Expressions	152
3.9.4	A Formal System for Regular Expressions	158
3.10	Transducers	165
3.11	Automata and String Patterns	171
3.12	Applications of Regular Expressions	184
3.12.1	Regular Expressions and UNIX	184
3.12.2	The <code>grep</code> Utility and Its Relatives	186
3.12.3	The <code>aux</code> Text Processing Program	187
3.12.4	The <code>lex</code> Lexical Analyzer Generator	189
3.13	Exercises	191
3.14	Bibliographical Comments	221
4	Rewriting Systems and Grammars	223
4.1	Introduction	223
4.2	Semi-Thue and Thue Systems	223
4.3	Grammars and Chomsky Hierarchy	228
4.3.1	Equivalent Grammars	233
4.4	Regular Operations	237
4.5	Properties of Type-2 Grammars	242
4.6	Regular Languages and Type-3 Grammars	254
4.7	Exercises	258
4.8	Bibliographical Comments	267

5	Context-Free Languages	269
5.1	Introduction	269
5.2	Derivations and Derivation Trees	270
5.3	Fixed-Points and Context-Free Languages	281
5.4	Normal Forms	286
5.4.1	Chomsky Normal Form	286
5.4.2	Greibach Normal Form	289
5.5	The Pumping Lemmas	298
5.6	Closure Properties	302
5.7	Regular and Context-Free Languages	306
5.8	Ambiguity	308
5.9	Parikh Theorem	314
5.10	The Chomsky-Schützenberger Theorem	319
5.11	Exercises	322
5.12	Bibliographical Comments	335
6	Pushdown Automata	337
6.1	Introduction	337
6.2	Nondeterministic Pushdown Automata	337
6.3	Deterministic Context-Free Languages	352
6.4	Exercises	370
6.5	Bibliographical Comments	376
III	Algorithmic Aspects	377
7	Partial Recursive Functions	379
7.1	Computable Functions	379
7.2	Primitive Recursive Functions	380
7.3	Primitive Recursive Predicates	385
7.4	Bounded Minimalization	391
7.5	Extensions	393
7.6	Numerical Primitive Recursive Functions	395
7.7	Transformations between Alphabets	401
7.8	Primitive Recursive Languages	411
7.9	Partial Recursive Functions	414
7.10	Exercises	422
7.11	Bibliographical Comments	430
8	Recursively Enumerable Languages	431
8.1	Introduction	431
8.2	Labeled Markov Algorithms	432
8.3	Turing Machines	439
8.4	Systems of Deterministic Turing Machines	444
8.5	Church's Thesis	448
8.5.1	Functions Computable by Turing Machines	451
8.5.2	Closing the Circle	458
8.5.3	Recursive Languages	463

8.5.4	Universality	464
8.6	Recursive Enumerable Languages	471
8.7	Rice's Theorem	489
8.8	Post Correspondence Problem	491
8.9	Multitape Turing Machines	497
8.10	Nondeterministic Turing Machines	503
8.11	Exercises	506
8.12	Bibliographical Comments	521
9	Context-Sensitive Languages	523
9.1	Introduction	523
9.2	Linear Bounded Automata	523
9.3	Closure Properties	531
9.4	Normal Forms for Context-Sensitive Grammars	546
9.5	Exercises	548
9.6	Bibliographical Comments	549
IV	Applications	551
10	Codes	553
10.1	Introduction	553
10.2	Unique Decipherability	554
10.3	The Kraft-McMillan Inequality	561
10.4	Huffman Codes and Data Compression	567
10.5	Exercises	571
10.6	Bibliographical Comments	573
11	Biological Applications	575
11.1	Introduction	575
11.2	L -Systems	575
11.3	Nucleic Acids	589
11.4	Exercises	601
11.5	Bibliographical Comments	606
	Bibliography	607
	Notation Index	613
	Topic Index	619