

Jan Daciuk

Optimization of Automata

Gdańsk 2014

PRZEWODNICZĄCY KOMITETU REDAKCYJNEGO
WYDAWNICTWA POLITECHNIKI GDAŃSKIEJ

Janusz T. Cieśliński

REDAKTOR PUBLIKACJI NAUKOWYCH

Michał Szydłowski

RECENZENCI

Marek Kubale

Rogério Reis

PROJEKT OKŁADKI

Katarzyna Olszonowicz

Wydano za zgodą
Rektora Politechniki Gdańskiej

Oferta wydawnicza Politechniki Gdańskiej jest dostępna pod adresem
<http://pg.edu.pl/wydawnictwo/katalog>
zamówienia prosimy kierować na adres wydaw@pg.gda.pl

Utwór nie może być powielany i rozpowszechniany, w jakiegokolwiek formie
i w jakiegokolwiek sposób, bez pisemnej zgody wydawcy

© Copyright by Wydawnictwo Politechniki Gdańskiej
Gdańsk 2014

ISBN 978-83-7348-564-8

Contents

Acronyms	5
Glossary	7
1 Introduction	9
1.1 Motivation	9
1.2 Incrementality	9
1.3 Hashing	10
1.4 Compression	10
1.5 Organization of This Book	10
1.6 Acknowledgements	11
2 Definitions	13
2.1 Finite-State Automata	13
2.2 Tree Automata	16
3 Sorted Incremental Construction	19
3.1 Traditional, Non-Incremental Construction	20
3.2 Incremental Construction of Minimal, Acyclic DFAs from sorted data	28
3.3 Extension to Pseudo-Minimal Automata	40
3.4 Extension to Cyclic Automata	45
4 Incremental Construction from Unsorted Data	63
4.1 Incremental MADFA Construction from Unsorted Data	63
4.2 Extension to Pseudo-Minimal Automata	79
4.3 Extension to Cyclic Automata	89
4.4 Extension to Minimal Bottom-Up Tree Automata	98
4.5 Extension to Pseudo-Minimal Bottom-Up Tree Automata	116
5 Semi-Incremental Construction Algorithms	131
5.1 Watson's Algorithm	131
5.1.1 Extension to Pseudo-Minimal Automata	137
5.1.2 Extension to Cyclic Automata	144
5.1.3 Extension to Minimal Bottom-Up Tree Automata	152
5.1.4 Extension to Pseudo-Minimal Bottom-Up Tree Automata	164

5.2	Revuz's Algorithm	166
5.3	Summary	167
6	Hashing	171
6.1	Perfect Hashing with Minimal DFAs --- Numbers in States	172
6.2	Perfect Hashing with Minimal DFAs --- Numbers in Transitions	176
6.3	Arbitrary Hashing with Pseudo-Minimal DFAs	178
6.4	Variable Output Transducers	180
6.5	Dynamic Perfect Hashing	182
6.6	Perfect Hashing with Minimal DTAs	185
6.7	Arbitrary Hashing with Pseudo-Minimal DTAs	194
6.8	Summary	194
7	Incremental Minimization Algorithms	197
7.1	Original Algorithm by Bruce Watson	197
7.2	Faster Minimization	200
7.2.1	Presorting of States	200
7.2.2	Saving Fewer State Pairs in S	201
7.2.3	Full Memoization	201
7.2.4	Improved Version	203
7.3	Simplified Version	208
7.4	New Version	209
8	Memory-Efficient Representations	211
8.1	Sharing Space	212
8.2	Compressing Fields	216
8.3	Using Frequency	216
8.4	Summary	217
9	Summary	219
	Bibliography	221
	Index	225