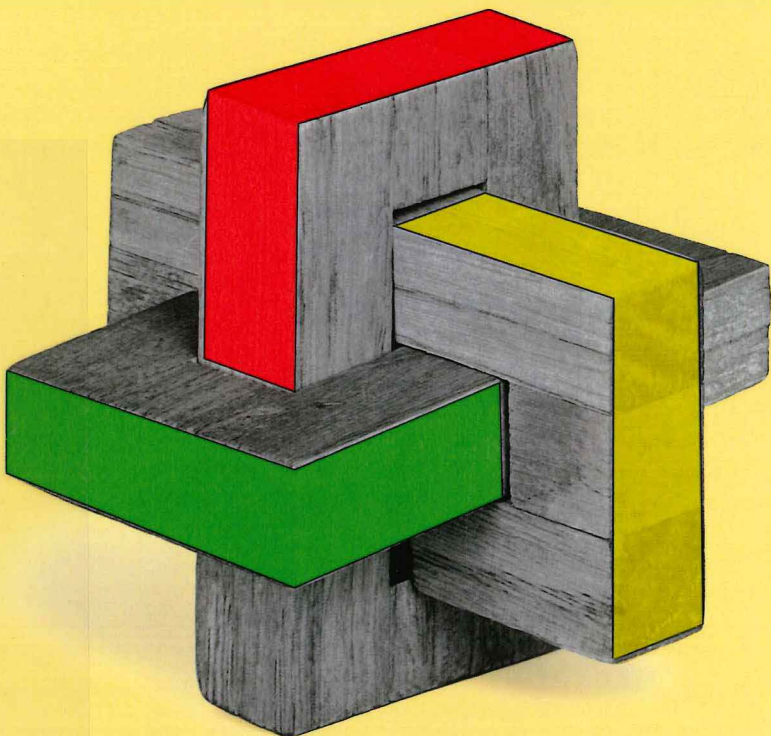


125 Problems in Text Algorithms

Maxime Crochemore
Thierry Lecroq
Wojciech Rytter



สำนักหอสมุด มหาวิทยาลัยเชียงใหม่

p65
2200-

6 16586505
012538504
1 22568232

125 Problems in Text Algorithms

With Solutions

MAXIME CROCHEMORE

Gustave Eiffel University

THIERRY LECROQ

University of Rouen Normandy

WOJCIECH RYTTER

University of Warsaw



สำนักหอสมุดมหาวิทยาลัยเชียงใหม่
CHIANG MAI UNIVERSITY LIBRARY



CAMBRIDGE
UNIVERSITY PRESS

2005 MAR 11

Contents

	<i>Preface</i>	<i>page ix</i>
1	The Very Basics of Stringology	1
2	Combinatorial Puzzles	17
1	Stringologic Proof of Fermat's Little Theorem	18
2	Simple Case of Codicity Testing	19
3	Magic Squares and the Thue–Morse Word	20
4	Oldenburger–Kolakoski Sequence	22
5	Square-Free Game	24
6	Fibonacci Words and Fibonacci Numeration System	26
7	Wythoff's Game and Fibonacci Word	28
8	Distinct Periodic Words	30
9	A Relative of the Thue–Morse Word	33
10	Thue–Morse Words and Sums of Powers	34
11	Conjugates and Rotations of Words	35
12	Conjugate Palindromes	37
13	Many Words with Many Palindromes	39
14	Short Superword of Permutations	41
15	Short Supersequence of Permutations	43
16	Skolem Words	45
17	Langford Words	48
18	From Lyndon Words to de Bruijn Words	50
3	Pattern Matching	53
19	Border Table	54
20	Shortest Covers	56
21	Short Borders	58

22	Prefix Table	60
23	Border Table to the Maximal Suffix	62
24	Periodicity Test	65
25	Strict Borders	67
26	Delay of Sequential String Matching	70
27	Sparse Matching Automaton	72
28	Comparison-Effective String Matching	74
29	Strict Border Table of the Fibonacci Word	76
30	Words with Singleton Variables	78
31	Order-Preserving Patterns	81
32	Parameterised Matching	83
33	Good-Suffix Table	85
34	Worst Case of the Boyer-Moore Algorithm	88
35	Turbo-BM Algorithm	90
36	String Matching with Don't Cares	92
37	Cyclic Equivalence	93
38	Simple Maximal Suffix Computation	96
39	Self-Maximal Words	98
40	Maximal Suffix and Its Period	100
41	Critical Position of a Word	103
42	Periods of Lyndon Word Prefixes	105
43	Searching Zimin Words	107
44	Searching Irregular 2D Patterns	110
4	Efficient Data Structures	111
45	List Algorithm for Shortest Cover	112
46	Computing Longest Common Prefixes	113
47	Suffix Array to Suffix Tree	115
48	Linear Suffix Trie	119
49	Ternary Search Trie	122
50	Longest Common Factor of Two Words	124
51	Subsequence Automaton	126
52	Codicity Test	128
53	LPF Table	130
54	Sorting Suffixes of Thue-Morse Words	134
55	Bare Suffix Tree	137
56	Comparing Suffixes of a Fibonacci Word	139
57	Avoidability of Binary Words	141
58	Avoiding a Set of Words	144

59	Minimal Unique Factors	146
60	Minimal Absent Words	148
61	Greedy Superstring	152
62	Shortest Common Superstring of Short Words	155
63	Counting Factors by Length	157
64	Counting Factors Covering a Position	160
65	Longest Common-Parity Factors	161
66	Word Square-Freeness with DBF	162
67	Generic Words of Factor Equations	164
68	Searching an Infinite Word	166
69	Perfect Words	169
70	Dense Binary Words	173
71	Factor Oracle	175
5	Regularities in Words	180
72	Three Square Prefixes	181
73	Tight Bounds on Occurrences of Powers	183
74	Computing Runs on General Alphabets	185
75	Testing Overlaps in a Binary Word	188
76	Overlap-Free Game	190
77	Anchored Squares	192
78	Almost Square-Free Words	195
79	Binary Words with Few Squares	197
80	Building Long Square-Free Words	199
81	Testing Morphism Square-Freeness	201
82	Number of Square Factors in Labelled Trees	203
83	Counting Squares in Combs in Linear Time	206
84	Cubic Runs	208
85	Short Square and Local Period	210
86	The Number of Runs	212
87	Computing Runs on Sorted Alphabet	214
88	Periodicity and Factor Complexity	219
89	Periodicity of Morphic Words	220
90	Simple Anti-powers	222
91	Palindromic Concatenation of Palindromes	224
92	Palindrome Trees	225
93	Unavoidable Patterns	227

6	Text Compression	230
94	BW Transform of Thue–Morse Words	231
95	BW Transform of Balanced Words	233
96	In-place BW Transform	237
97	Lempel–Ziv Factorisation	239
98	Lempel–Ziv–Welch Decoding	242
99	Cost of a Huffman Code	244
100	Length-Limited Huffman Coding	248
101	Online Huffman Coding	253
102	Run-Length Encoding	256
103	A Compact Factor Automaton	261
104	Compressed Matching in a Fibonacci Word	264
105	Prediction by Partial Matching	266
106	Compressing Suffix Arrays	269
107	Compression Ratio of Greedy Superstrings	271
7	Miscellaneous	275
108	Binary Pascal Words	276
109	Self-Reproducing Words	278
110	Weights of Factors	280
111	Letter-Occurrence Differences	282
112	Factoring with Border-Free Prefixes	283
113	Primitivity Test for Unary Extensions	286
114	Partially Commutative Alphabets	288
115	Greatest Fixed-Density Necklace	290
116	Period-Equivalent Binary Words	292
117	Online Generation of de Bruijn Words	295
118	Recursive Generation of de Bruijn Words	298
119	Word Equations with Given Lengths of Variables	300
120	Diverse Factors over a Three-Letter Alphabet	302
121	Longest Increasing Subsequence	304
122	Unavoidable Sets via Lyndon Words	306
123	Synchronising Words	309
124	Safe-Opening Words	311
125	Superwords of Shortened Permutations	314
	<i>Bibliography</i>	318
	<i>Index</i>	332