Contents

Preface vii

1 Introduction 1
1.1 Structure of the chapters 1
1.2 Naming of variables 3
1.3 Probabilities 4
1.4 Asymptotic notation 5
1.5 About the programming languages 5
1.6 On the code for the algorithms 6
1.7 Complexity measures and real timings 7

2 Basic Concepts 9
2.1 Data structure description 9
2.1.1 Grammar for data objects 9
2.1.2 Constraints for data objects 12
2.1.2.1 Sequential order 13
2.1.2.2 Uniqueness 13
2.1.2.3 Hierarchical order 13
2.1.2.4 Hierarchical balance 13
2.1.2.5 Optimality 14
2.2 Algorithm descriptions 14
2.2.1 Basic (or atomic) operations 15
2.2.2 Building procedures 17
2.2.2.1 Composition 17
2.2.2.2 Alternation 21
2.2.2.3 Conformation 22
2.2.2.4 Self-organization 23
2.2.3 Interchangeability 23
3 Searching Algorithms

3.1 Sequential search
  3.1.1 Basic sequential search
  3.1.2 Self-organizing sequential search: move-to-front method
  3.1.3 Self-organizing sequential search: transpose method
  3.1.4 Optimal sequential search
  3.1.5 Jump search

3.2 Sorted array search
  3.2.1 Binary search
  3.2.2 Interpolation search
  3.2.3 Interpolation-sequential search

3.3 Hashing
  3.3.1 Practical hashing functions
  3.3.2 Uniform probing hashing
  3.3.3 Random probing hashing
  3.3.4 Linear probing hashing
  3.3.5 Double hashing
  3.3.6 Quadratic hashing
  3.3.7 Ordered and split-sequence hashing
  3.3.8 Reorganization schemes
    3.3.8.1 Brent’s algorithm
    3.3.8.2 Binary tree hashing
    3.3.8.3 Last-come-first-served hashing
    3.3.8.4 Robin Hood hashing
    3.3.8.5 Self-adjusting hashing
  3.3.9 Optimal hashing
  3.3.10 Direct chaining hashing
  3.3.11 Separate chaining hashing
  3.3.12 Coalesced hashing
  3.3.13 Extendible hashing
  3.3.14 Linear hashing
  3.3.15 External hashing using minimal internal storage
  3.3.16 Perfect hashing
  3.3.17 Summary

3.4 Recursive structures search
  3.4.1 Binary tree search
    3.4.1.1 Randomly generated binary trees
    3.4.1.2 Random binary trees
    3.4.1.3 Height-balanced trees
    3.4.1.4 Weight-balanced trees
    3.4.1.5 Balancing by internal path reduction
    3.4.1.6 Heuristic organization schemes on binary trees
    3.4.1.7 Optimal binary tree search
    3.4.1.8 Rotations in binary trees
    3.4.1.9 Deletions in binary trees
3.4.1.10  m-ary search trees 116
3.4.2  B-trees 117
3.4.2.1  2-3 trees 124
3.4.2.2  Symmetric binary B-trees 126
3.4.2.3  1-2 trees 128
3.4.2.4  2-3-4 trees 129
3.4.2.5  B-tree variations 130
3.4.3  Index and indexed sequential files 130
3.4.3.1  Index sequential access method 132
3.4.4  Digital trees 133
3.4.4.1  Hybrid tries 137
3.4.4.2  Tries for word-dictionaries 138
3.4.4.3  Digital search trees 138
3.4.4.4  Compressed tries 140
3.4.4.5  Patricia trees 140
3.5  Multidimensional search 143
3.5.1  Quad trees 144
3.5.1.1  Quad tries 146
3.5.2  K-dimensional trees 149

4  Sorting Algorithms 153
4.1  Techniques for sorting arrays 153
4.1.1  Bubble sort 154
4.1.2  Linear insertion sort 156
4.1.3  Quicksort 158
4.1.4  Shellsort 161
4.1.5  Heapsort 164
4.1.6  Interpolation sort 166
4.1.7  Linear probing sort 168
4.1.8  Summary 170
4.2  Sorting other data structures 171
4.2.1  Merge sort 173
4.2.2  Quicksort for lists 174
4.2.3  Bucket sort 176
4.2.4  Radix sort 179
4.2.5  Hybrid methods of sorting 180
4.2.5.1  Recursion termination 181
4.2.5.2  Distributive partitioning 181
4.2.5.3  Non-recursive bucket sort 182
4.2.6  Treesort 182
4.3  Merging 183
4.3.1  List merging 184
4.3.2  Array merging 185
4.3.3  Minimal-comparison merging 186
CONTENTS

4.4 External sorting
   4.4.1 Selection phase techniques
     4.4.1.1 Replacement selection
     4.4.1.2 Natural selection
     4.4.1.3 Alternating selection
     4.4.1.4 Merging phase
   4.4.2 Balanced merge sort
   4.4.3 Cascade merge sort
   4.4.4 Polyphase merge sort
   4.4.5 Oscillating merge sort
   4.4.6 External Quicksort

5 Selection Algorithms
   5.1 Priority queues
     5.1.1 Sorted/unsorted lists
     5.1.2 P-trees
     5.1.3 Heaps
     5.1.4 Van Emde-Boas priority queues
     5.1.5 Pagodas
     5.1.6 Binary trees used as priority queues
       5.1.6.1 Leftist trees
       5.1.6.2 Binary priority queues
       5.1.6.3 Binary search trees as priority queues
     5.1.7 Binomial queues
     5.1.8 Summary
   5.2 Selection of kth element
     5.2.1 Selection by sorting
     5.2.2 Selection by tail recursion
     5.2.3 Selection of the mode

6 Arithmetic Algorithms
   6.1 Basic operations, multiplication/division
   6.2 Other arithmetic functions
     6.2.1 Binary powering
     6.2.2 Arithmetic-geometric mean
     6.2.3 Transcendental functions
   6.3 Matrix multiplication
     6.3.1 Strassen's matrix multiplication
     6.3.2 Further asymptotic improvements
   6.4 Polynomial evaluation
## Contents

### 7 Text Algorithms 251

7.1 Text searching without preprocessing 251

- 7.1.1 Brute force text searching 253
- 7.1.2 Knuth–Morris–Pratt text searching 254
- 7.1.3 Boyer–Moore text searching 256
- 7.1.4 Searching sets of strings 259
- 7.1.5 Karp–Rabin text searching 260
- 7.1.6 Searching text with automata 262
- 7.1.7 Shift-or text searching 266
- 7.1.8 String similarity searching 267
- 7.1.9 Summary of direct text searching 270

7.2 Searching preprocessed text 270

- 7.2.1 Inverted files 271
- 7.2.2 Trees used for text searching 273
- 7.2.3 Searching text with automata 275
- 7.2.4 Suffix arrays and PAT arrays 277
- 7.2.5 DAWG 279
- 7.2.6 Hashing methods for text searching 280
- 7.2.7 P-strings 281

7.3 Other text searching problems 283

- 7.3.1 Searching longest common subsequences 283
- 7.3.2 Two-dimensional searching 284

### I Distributions Derived from Empirical Observation 289

1.1 Zipf's law 289

- 1.1.1 First generalization of a Zipfian distribution 290
- 1.1.2 Second generalization of a Zipfian distribution 290

1.2 Bradford's law 291

1.3 Lotka's law 293

1.4 80%-20% rule 293

### II Asymptotic Expansions 297

11.1 Asymptotic expansions of sums 298

11.2 Gamma-type expansions 300

11.3 Exponential-type expansions 301

11.4 Asymptotic expansions of sums and definite integrals containing $e^{-z^2}$ 302

11.5 Doubly exponential forms 303

11.6 Roots of polynomials 304

11.7 Sums containing descending factorials 305

11.8 Summation formulas 307

### III References 309

111.1 Textbooks 309

111.2 Papers 311
<table>
<thead>
<tr>
<th>IV Algorithms coded in Pascal and C</th>
<th>375</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV.1 Searching algorithms</td>
<td>375</td>
</tr>
<tr>
<td>IV.2 Sorting algorithms</td>
<td>387</td>
</tr>
<tr>
<td>IV.3 Selection algorithms</td>
<td>399</td>
</tr>
<tr>
<td>IV.4 Text algorithms</td>
<td>408</td>
</tr>
</tbody>
</table>

Index 415