

Contents

Forword	xiii
Preface	xv
Acknowledgments	xxi

PART I PROBLEMS ON WORDS

1 Data Structures and Algorithms on Words	3
1.1 Digital Trees	4
1.2 Data Compression: Lempel-Ziv Algorithms	9
1.2.1 Lempel-Ziv'77 Algorithm	10
1.2.2 Lempel-Ziv'78 Algorithm	11
1.2.3 Extensions of Lempel-Ziv Schemes	12
1.3 Pattern Matching	15
1.4 Shortest Common Superstring	17
1.5 String Editing Problem	20
1.6 Optimization Problems	22
1.7 Exercises	23
2 Probabilistic and Analytical Models	25
2.1 Probabilistic Models of Strings	26
2.2 Review of Probability	30
2.2.1 Some Useful Inequalities	30
2.2.2 Types of Stochastic Convergence	32
2.3 Review of Complex Analysis	35
2.4 Special Functions	39
2.4.1 Euler's Gamma Function	40
2.4.2 Riemann's Zeta Function	44
2.5 Extensions and Exercises	46

PART II PROBABILISTIC AND COMBINATORIAL TECHNIQUES

3	Inclusion-Exclusion Principle	51
3.1	Probabilistic Inclusion-Exclusion Principle	52
3.2	Combinatorial Inclusion-Exclusion Principle	56
3.3	Applications	59
3.3.1	Depth in a Trie	60
3.3.2	Order Statistics	61
3.3.3	Longest Aligned Word	64
3.4	Extensions and Exercises	69
4	The First and Second Moment Methods	73
4.1	The Methods	74
4.2	Applications	80
4.2.1	Markov Approximation of a Stationary Distribution	80
4.2.2	Excursion into Number Theory	82
4.2.3	Height in Tries	84
4.2.4	Height in PATRICIA Tries	94
4.2.5	Height in Digital Search Trees	97
4.2.6	Height in a Suffix Tree	99
4.3	Extensions and Exercises	103
5	Subadditive Ergodic Theorem and Large Deviations	106
5.1	Subadditive Sequence	107
5.2	Subadditive Ergodic Theorem	113
5.3	Martingales and Azuma's Inequality	116
5.4	Large Deviations	125
5.5	Applications	134
5.5.1	Edit Distance	134
5.5.2	Knuth-Morris-Pratt Pattern Matching Algorithms	136
5.5.3	Large Deviations of a Random Sequence	141
5.6	Extensions and Exercises	144
6	Elements of Information Theory	149
6.1	Entropy, Relative Entropy, and Mutual Information	151
6.2	Entropy Rate and Rényi's Entropy Rates	154
6.2.1	The Shannon-McMillan-Breiman Theorem	154
6.2.2	Rényi's Entropy Rates	157
6.3	Asymptotic Equipartition Property	159
6.3.1	Typical Sequences	159

6.3.2	Jointly Typical Sequences	161
6.3.3	AEP for Biased Distributions	163
6.3.4	Lossy Generalizations of AEP	165
6.4	Three Theorems of Shannon	172
6.4.1	Source Coding Theorem	172
6.4.2	Channel Coding Theorem	176
6.4.3	Rate Distortion Theorem	181
6.5	Applications	187
6.5.1	Phrase Length in the Lempel-Ziv Scheme and Depth in a Suffix Tree	187
6.5.2	Shortest Common Superstring Problem	195
6.5.3	Fixed-Database Lossy Lempel-Ziv Algorithm	200
6.6	Extensions and Exercises	204

PART III ANALYTIC TECHNIQUES

7	Generating Functions	213
7.1	Ordinary Generating Functions	215
7.1.1	Formal Power Series	215
7.1.2	Combinatorial Calculus	218
7.1.3	Elements of Analytic Theory	222
7.1.4	Generating Functions Over an Arithmetic Progression	227
7.2	Exponential Generating Functions	229
7.2.1	Elementary Properties	229
7.2.2	Labeled Combinatorial Structures	231
7.3	Cauchy, Lagrange and Borel Formulas	235
7.3.1	Cauchy Coefficient Formula	235
7.3.2	Lagrange Inversion Formula	236
7.3.3	Borel Transform	242
7.4	Probability Generating Functions	243
7.5	Dirichlet Series	246
7.5.1	Basic Properties	247
7.5.2	Euler Products	249
7.5.3	Perron-Mellin Formula	251
7.6	Applications	256
7.6.1	Recurrences Arising in the Analysis of Digital Trees	256
7.6.2	Pattern Occurrences in a Random Text	262
7.6.3	Delange's Formula for a Digital Sum	272
7.7	Extensions and Exercises	275

8	Complex Asymptotic Methods	278
8.1	Introduction to Asymptotic Expansions	281
8.1.1	Definitions and Notations	281
8.1.2	From Taylor Expansion to Asymptotic Series	284
8.2	Basic Methods	290
8.2.1	Euler-Maclaurin Summation Formula	290
8.2.2	Matched Asymptotics and the WKB Method	296
8.2.3	Uniform Distribution of Sequences	304
8.3	Small Singularities of Analytic Functions	312
8.3.1	Polar Singularities	314
8.3.2	Algebraic Singularities: Singularity Analysis	318
8.4	Large Singularities: Saddle Point Method	331
8.4.1	Laplace Method	332
8.4.2	Method of Steepest Descent	337
8.4.3	Local Central Limit Theorem and Large Deviations Revisited	350
8.5	Finite Sums as Complex Integrals	353
8.6	Limiting Distributions	362
8.6.1	Discrete Limit Laws Through PGF	362
8.6.2	Continuous Limit Laws by Integral Transforms	365
8.7	Applications	372
8.7.1	Variations on Approximate Self-Overlapping Words	372
8.7.2	Redundancy Rate for Memoryless Sources	376
8.7.3	Limiting Distribution of the Depth in Digital Search Trees	381
8.8	Extensions and Exercises	388
9	Mellin Transform and Its Applications	398
9.1	Basic Properties	400
9.2	Asymptotic Properties of the Mellin Transform	407
9.3	Extension to the Complex Plane	417
9.4	Applications	421
9.4.1	Variance of the Depth in a Generalized Digital Search Tree	421
9.4.2	Redundancy of Renewal Processes	425
9.5	Extensions and Exercises	435
10	Analytic Poissonization and Depoissonization	442
10.1	Analytic Poissonization and the Poisson Transform	446
10.1.1	Poisson Transform	446
10.1.2	Asymptotic Poissonization	449

10.2	Basic Depoissonization	453
10.2.1	Algebraic Depoissonization	454
10.2.2	Asymptotic Depoissonization	455
10.3	Generalizations of Depoissonization	466
10.3.1	A Full Asymptotic Depoissonization	466
10.3.2	Exponential Depoissonization	472
10.3.3	General Depoissonization Tool	475
10.4	Moments and Limiting Distributions	480
10.4.1	Moments	481
10.4.2	Limiting Distributions	482
10.5	Applications	485
10.5.1	Leader Election Algorithm	486
10.5.2	Depth in a b -Digital Search Tree for the Biased Memoryless Model	491
10.5.3	Depth in a Digital Search Tree with a Markovian Source	501
10.6	Extensions and Exercises	512

Bibliography	520
---------------------	------------

Index	543
--------------	------------