

An Introduction to Kolmogorov Complexity and Its Applications

Ming Li
Paul Vitányi

With 38 Illustrations

Technische Hochschule Darmstadt
FACHBEREICH INFORMATIK
B I B L I O T H E K
Inventar-Nr.: 21074
Sachgebiete: F.V
Standort: 1994



Springer-Verlag
New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

Contents

Preface	v
How to Use This Book	viii
Acknowledgements	x
Outlines of One-Semester Courses	xii
List of Figures	xix
1 Preliminaries	1
1.1 A Brief Introduction	1
1.2 Mathematical Preliminaries	6
1.2.1 Prerequisites and Notation	6
1.2.2 Numbers and Combinatorics	7
1.2.3 Binary Strings	11
1.2.4 Asymptotics Notation	14
1.3 Basics of Probability Theory	16
1.3.1 Kolmogorov Axioms	17
1.3.2 Conditional Probability	18
1.3.3 Continuous Sample Spaces	19
1.4 Basics of Computability Theory	22
1.4.1 Effective Enumerations and Universal Machines	26
1.4.2 Undecidability of the Halting Problem	32
1.4.3 Enumerable Functions	34
1.4.4 Feasible Computations	35
1.5 The Roots of Kolmogorov Complexity	45
1.5.1 Randomness	46

1.5.2	Prediction and Probability	55
1.5.3	Information Theory and Coding	61
1.5.4	State \times Symbol Complexity	79
1.6	History and References	80
2	Algorithmic Complexity	87
2.1	The Invariance Theorem	90
2.2	Incompressibility	95
2.3	C as an Integer Function	101
2.4	Random Finite Sequences	105
2.5	*Random Infinite Sequences	112
2.6	Statistical Properties of Finite Sequences	126
2.6.1	Statistics of 0's and 1's	127
2.6.2	Statistics of Blocks	130
2.6.3	Length of Runs	132
2.7	Algorithmic Properties of C	134
2.8	Algorithmic Information Theory	140
2.9	History and References	165
3	Algorithmic Prefix Complexity	169
3.1	The Invariance Theorem	171
3.2	Incompressibility	175
3.3	K as an Integer Function	177
3.4	Random Finite Sequences	177
3.5	*Random Infinite Sequences	180
3.6	Algorithmic Properties of K	188
3.7	*The Complexity of the Complexity Function	190
3.8	*Symmetry of Algorithmic Information	194
3.9	History and References	209
4	Algorithmic Probability	211
4.1	Enumerable Functions Revisited	212
4.2	A Nonclassical Approach to Measures	214
4.3	Discrete Sample Space	216
4.3.1	Universal Enumerable Semimeasure	217
4.3.2	<i>A Priori</i> Probability	221
4.3.3	Algorithmic Probability	223
4.3.4	The Coding Theorem	223
4.3.5	Randomness by Sum Tests	228
4.3.6	Randomness by Payoff Functions	232
4.4	Continuous Sample Space	234
4.4.1	Universal Enumerable Semimeasure	234
4.4.2	<i>A Priori</i> Probability	238
4.4.3	*Solomonoff Normalization	242
4.4.4	*Monotone Complexity and a Coding Theorem	243

4.4.5	*Relation Between Complexities	246
4.4.6	*Randomness by Integral Tests	247
4.4.7	*Randomness by Martingale Tests	254
4.4.8	*Randomness by Martingales	256
4.4.9	*Relations Between Tests	258
4.5	History and References	268
5	Inductive Reasoning	275
5.1	Introduction	275
5.2	Bayesian Reasoning	279
5.3	Solomonoff's Induction Theory	282
5.3.1	Formal Analysis	284
5.3.2	Application to Induction	290
5.4	Recursion Theory Induction	291
5.4.1	Inference of Hypotheses	291
5.4.2	Prediction	292
5.4.3	Mistake Bounds	293
5.4.4	Certification	294
5.5	Pac-Learning	295
5.5.1	Definitions	296
5.5.2	Occam's Razor Formalized	296
5.6	Simple Pac-Learning	300
5.6.1	Discrete Sample Space	301
5.6.2	Continuous Sample Space	305
5.7	Minimum Description Length	308
5.8	History and References	318
6	The Incompressibility Method	323
6.1	Two Examples	324
6.2	Combinatorics	328
6.3	Average Case Complexity of Algorithms	334
6.3.1	Heapsort	334
6.3.2	Longest Common Subsequence	338
6.3.3	m-Average Case Complexity	340
6.4	Languages	344
6.4.1	Formal Language Theory	344
6.4.2	On-Line CFL Recognition	349
6.4.3	Multihead Automata	351
6.5	Machines	356
6.5.1	*Turing Machine Time Complexity	356
6.5.2	Parallel Computation	362
6.6	History and References	370

7	Resource-Bounded Complexity	377
7.1	Mathematical Theory	378
7.1.1	Computable Majorants	381
7.1.2	Resource-Bounded Hierarchies	386
7.2	Language Compression	392
7.2.1	With an Oracle	393
7.2.2	Without an Oracle	396
7.2.3	Ranking	399
7.3	Computational Complexity	401
7.3.1	Constructing Oracles	402
7.3.2	P-Printability	405
7.3.3	Instance Complexity	406
7.4	<i>Kt</i> Complexity	410
7.4.1	Universal Optimal Search	411
7.4.2	Potential	413
7.5	Logical Depth	421
7.6	History and References	428
8	Physics and Computation	433
8.1	Reversible Computation	434
8.1.1	Energy Dissipation	434
8.1.2	Reversible Logic Circuits	435
8.1.3	A Ballistic Computer	436
8.1.4	Reversible Turing Machines	439
8.2	Information Distance	441
8.2.1	Max Distance	442
8.2.2	Picture Distance	446
8.2.3	Reversible Distance	448
8.2.4	Sum Distance	450
8.2.5	Metrics Relations and Dimensional Properties	452
8.2.6	Thermodynamics of Computing	455
8.3	Thermodynamics	458
8.3.1	Classical Entropy	458
8.3.2	Statistical Mechanics and Boltzmann Entropy	461
8.3.3	Gibbs Entropy	467
8.4	Entropy Revisited	468
8.4.1	Algorithmic Entropy	469
8.4.2	Algorithmic Entropy and Randomness Tests	473
8.4.3	Entropy Stability and Nondecrease	478
8.5	Chaos, Biology, and All That	486
8.6	History and References	490
	Bibliography	493
	Index	527