
Contents

1	The Science of Information	1
----------	-----------------------------------------	---

Part I Components of Information Theory

2	Information Measures	7
2.1	Independence and Markov Chains	7
2.2	Shannon's Information Measures	12
2.3	Continuity of Shannon's Information Measures for Fixed Finite Alphabets	18
2.4	Chain Rules	21
2.5	Informational Divergence	23
2.6	The Basic Inequalities	26
2.7	Some Useful Information Inequalities	28
2.8	Fano's Inequality	32
2.9	Maximum Entropy Distributions	36
2.10	Entropy Rate of a Stationary Source	38
	Appendix 2.A: Approximation of Random Variables with Countably Infinite Alphabets by Truncation	41
	Chapter Summary	43
	Problems	45
	Historical Notes	50
3	The I-Measure	51
3.1	Preliminaries	52
3.2	The I -Measure for Two Random Variables	53
3.3	Construction of the I -Measure μ^*	55
3.4	μ^* Can Be Negative	59
3.5	Information Diagrams	61
3.6	Examples of Applications	67
	Appendix 3.A: A Variation of the Inclusion–Exclusion Formula	74
	Chapter Summary	76

Problems	78
Historical Notes	80
4 Zero-Error Data Compression	81
4.1 The Entropy Bound	81
4.2 Prefix Codes	86
4.2.1 Definition and Existence	86
4.2.2 Huffman Codes	88
4.3 Redundancy of Prefix Codes	93
Chapter Summary	97
Problems	98
Historical Notes	99
5 Weak Typicality	101
5.1 The Weak AEP	101
5.2 The Source Coding Theorem	104
5.3 Efficient Source Coding	106
5.4 The Shannon–McMillan–Breiman Theorem	107
Chapter Summary	109
Problems	110
Historical Notes	112
6 Strong Typicality	113
6.1 Strong AEP	113
6.2 Strong Typicality Versus Weak Typicality	121
6.3 Joint Typicality	122
6.4 An Interpretation of the Basic Inequalities	131
Chapter Summary	131
Problems	132
Historical Notes	135
7 Discrete Memoryless Channels	137
7.1 Definition and Capacity	141
7.2 The Channel Coding Theorem	149
7.3 The Converse	152
7.4 Achievability	157
7.5 A Discussion	164
7.6 Feedback Capacity	167
7.7 Separation of Source and Channel Coding	172
Chapter Summary	175
Problems	176
Historical Notes	181
8 Rate-Distortion Theory	183
8.1 Single-Letter Distortion Measures	184
8.2 The Rate-Distortion Function $R(D)$	187

8.3	The Rate-Distortion Theorem	192
8.4	The Converse	200
8.5	Achievability of $R_I(D)$	202
	Chapter Summary	207
	Problems	208
	Historical Notes	209
9	The Blahut–Arimoto Algorithms	211
9.1	Alternating Optimization	212
9.2	The Algorithms	214
9.2.1	Channel Capacity	214
9.2.2	The Rate-Distortion Function	219
9.3	Convergence	222
9.3.1	A Sufficient Condition	222
9.3.2	Convergence to the Channel Capacity	225
	Chapter Summary	226
	Problems	227
	Historical Notes	228
10	Differential Entropy	229
10.1	Preliminaries	231
10.2	Definition	235
10.3	Joint Differential Entropy, Conditional (Differential) Entropy, and Mutual Information	238
10.4	The AEP for Continuous Random Variables	245
10.5	Informational Divergence	248
10.6	Maximum Differential Entropy Distributions	249
	Chapter Summary	252
	Problems	255
	Historical Notes	256
11	Continuous-Valued Channels	257
11.1	Discrete-Time Channels	257
11.2	The Channel Coding Theorem	260
11.3	Proof of the Channel Coding Theorem	262
11.3.1	The Converse	262
11.3.2	Achievability	265
11.4	Memoryless Gaussian Channels	270
11.5	Parallel Gaussian Channels	272
11.6	Correlated Gaussian Channels	278
11.7	The Bandlimited White Gaussian Channel	280
11.8	The Bandlimited Colored Gaussian Channel	287
11.9	Zero-Mean Gaussian Noise Is the Worst Additive Noise	290
	Chapter Summary	294
	Problems	296
	Historical Notes	297

12 Markov Structures	299
12.1 Conditional Mutual Independence	300
12.2 Full Conditional Mutual Independence	309
12.3 Markov Random Field	314
12.4 Markov Chain	317
Chapter Summary	319
Problems	320
Historical Notes	321
13 Information Inequalities	323
13.1 The Region I_n^*	325
13.2 Information Expressions in Canonical Form	326
13.3 A Geometrical Framework	329
13.3.1 Unconstrained Inequalities	329
13.3.2 Constrained Inequalities	330
13.3.3 Constrained Identities	332
13.4 Equivalence of Constrained Inequalities	333
13.5 The Implication Problem of Conditional Independence	336
Chapter Summary	337
Problems	338
Historical Notes	338
14 Shannon-Type Inequalities	339
14.1 The Elemental Inequalities	339
14.2 A Linear Programming Approach	341
14.2.1 Unconstrained Inequalities	343
14.2.2 Constrained Inequalities and Identities	344
14.3 A Duality	345
14.4 Machine Proving – ITIP	347
14.5 Tackling the Implication Problem	351
14.6 Minimality of the Elemental Inequalities	353
Appendix 14.A: The Basic Inequalities and the Polymatroidal Axioms	356
Chapter Summary	357
Problems	358
Historical Notes	360
15 Beyond Shannon-Type Inequalities	361
15.1 Characterizations of I_2^* , I_3^* , and \bar{T}_n^*	361
15.2 A Non-Shannon-Type Unconstrained Inequality	369
15.3 A Non-Shannon-Type Constrained Inequality	374
15.4 Applications	380
Chapter Summary	383
Problems	383
Historical Notes	385

16 Entropy and Groups 387

16.1 Group Preliminaries 388

16.2 Group-Characterizable Entropy Functions 393

16.3 A Group Characterization of \overline{T}_n^* 398

16.4 Information Inequalities and Group Inequalities 401

Chapter Summary 405

Problems 406

Historical Notes 408

Part II Fundamentals of Network Coding

17 Introduction 411

17.1 The Butterfly Network 412

17.2 Wireless and Satellite Communications 415

17.3 Source Separation 417

Chapter Summary 418

Problems 418

Historical Notes 419

18 The Max-Flow Bound 421

18.1 Point-to-Point Communication Networks 421

18.2 Examples Achieving the Max-Flow Bound 424

18.3 A Class of Network Codes 427

18.4 Proof of the Max-Flow Bound 429

Chapter Summary 431

Problems 431

Historical Notes 434

19 Single-Source Linear Network Coding:

Acyclic Networks 435

19.1 Acyclic Networks 436

19.2 Linear Network Codes 437

19.3 Desirable Properties of a Linear Network Code 443

19.3.1 Transformation of a Linear Network Code 447

19.3.2 Implementation of a Linear Network Code 448

19.4 Existence and Construction 449

19.5 Generic Network Codes 460

19.6 Static Network Codes 468

19.7 Random Network Coding: A Case Study 473

19.7.1 How the System Works 474

19.7.2 Model and Analysis 475

Chapter Summary 478

Problems 479

Historical Notes 482

20	Single-Source Linear Network Coding: Cyclic Networks	485
	20.1 Delay-Free Cyclic Networks	485
	20.2 Convolutional Network Codes	488
	20.3 Decoding of Convolutional Network Codes	498
	Chapter Summary	503
	Problems	503
	Historical Notes	504
21	Multi-source Network Coding	505
	21.1 The Max-Flow Bounds	505
	21.2 Examples of Application	508
	21.2.1 Multilevel Diversity Coding	508
	21.2.2 Satellite Communication Network	510
	21.3 A Network Code for Acyclic Networks	511
	21.4 The Achievable Information Rate Region	512
	21.5 Explicit Inner and Outer Bounds	515
	21.6 The Converse	516
	21.7 Achievability	521
	21.7.1 Random Code Construction	524
	21.7.2 Performance Analysis	527
	Chapter Summary	536
	Problems	537
	Historical Notes	539
	Bibliography	541
	Index	561