

---

## *Contents*

Preface . . . . .	v
1 Introduction to Information Theory . . . . .	3
2 Probability, Entropy, and Inference . . . . .	22
3 More about Inference . . . . .	48
<b>I Data Compression . . . . .</b>	<b>65</b>
4 The Source Coding Theorem . . . . .	67
5 Symbol Codes . . . . .	91
6 Stream Codes . . . . .	110
7 Codes for Integers . . . . .	132
<b>II Noisy-Channel Coding . . . . .</b>	<b>137</b>
8 Dependent Random Variables . . . . .	138
9 Communication over a Noisy Channel . . . . .	146
10 The Noisy-Channel Coding Theorem . . . . .	162
11 Error-Correcting Codes and Real Channels . . . . .	177
<b>III Further Topics in Information Theory . . . . .</b>	<b>191</b>
12 Hash Codes: Codes for Efficient Information Retrieval . . . . .	193
13 Binary Codes . . . . .	206
14 Very Good Linear Codes Exist . . . . .	229
15 Further Exercises on Information Theory . . . . .	233
16 Message Passing . . . . .	241
17 Communication over Constrained Noiseless Channels . . . . .	248
18 Crosswords and Codebreaking . . . . .	260
19 Why have Sex? Information Acquisition and Evolution . . . . .	269
<b>IV Probabilities and Inference . . . . .</b>	<b>281</b>
20 An Example Inference Task: Clustering . . . . .	284
21 Exact Inference by Complete Enumeration . . . . .	293
22 Maximum Likelihood and Clustering . . . . .	300
23 Useful Probability Distributions . . . . .	311
24 Exact Marginalization . . . . .	319
25 Exact Marginalization in Trellises . . . . .	324
26 Exact Marginalization in Graphs . . . . .	334
27 Laplace's Method . . . . .	341

Cambridge University Press

978-0-521-64298-9 - Information Theory, Inference, and Learning Algorithms

David J.C. MacKay

Table of Contents

[More information](#)

28	Model Comparison and Occam's Razor . . . . .	343
29	Monte Carlo Methods . . . . .	357
30	Efficient Monte Carlo Methods . . . . .	387
31	Ising Models . . . . .	400
32	Exact Monte Carlo Sampling . . . . .	413
33	Variational Methods . . . . .	422
34	Independent Component Analysis and Latent Variable Modelling . . . . .	437
35	Random Inference Topics . . . . .	445
36	Decision Theory . . . . .	451
37	Bayesian Inference and Sampling Theory . . . . .	457
<b>V</b>	<b>Neural networks . . . . .</b>	<b>467</b>
38	Introduction to Neural Networks . . . . .	468
39	The Single Neuron as a Classifier . . . . .	471
40	Capacity of a Single Neuron . . . . .	483
41	Learning as Inference . . . . .	492
42	Hopfield Networks . . . . .	505
43	Boltzmann Machines . . . . .	522
44	Supervised Learning in Multilayer Networks . . . . .	527
45	Gaussian Processes . . . . .	535
46	Deconvolution . . . . .	549
<b>VI</b>	<b>Sparse Graph Codes . . . . .</b>	<b>555</b>
47	Low-Density Parity-Check Codes . . . . .	557
48	Convolutional Codes and Turbo Codes . . . . .	574
49	Repeat–Accumulate Codes . . . . .	582
50	Digital Fountain Codes . . . . .	589
<b>VII</b>	<b>Appendices . . . . .</b>	<b>597</b>
A	Notation . . . . .	598
B	Some Physics . . . . .	601
C	Some Mathematics . . . . .	605
	Bibliography . . . . .	613
	Index . . . . .	620