

Information Theory

A Tutorial Introduction

James V Stone

Contents

1. What Is Information?	1
1.1. Introduction	1
1.2. Information, Eyes and Evolution	2
1.3. Finding a Route, Bit by Bit	3
1.4. A Million Answers to Twenty Questions	8
1.5. Information, Bits and Binary Digits	10
1.6. Example 1: Telegraphy	11
1.7. Example 2: Binary Images	13
1.8. Example 3: Grey-Level Images	15
1.9. Summary	20
2. Entropy of Discrete Variables	21
2.1. Introduction	21
2.2. Ground Rules and Terminology	21
2.3. Shannon's Desiderata	31
2.4. Information, Surprise and Entropy	31
2.5. Evaluating Entropy	38
2.6. Properties of Entropy	41
2.7. Independent and Identically Distributed Values	43
2.8. Bits, Shannons, and Bans	43
2.9. Summary	44
3. The Source Coding Theorem	45
3.1. Introduction	45
3.2. Capacity of a Discrete Noiseless Channel	46
3.3. Shannon's Source Coding Theorem	49
3.4. Calculating Information Rates	50
3.5. Data Compression	54
3.6. Huffman Coding	57
3.7. The Entropy of English Letters	61
3.8. Why the Theorem is True	71
3.9. Kolmogorov Complexity	77
3.10. Summary	78

4. The Noisy Channel Coding Theorem	79
4.1. Introduction	79
4.2. Joint Distributions	80
4.3. Mutual Information	88
4.4. Conditional Entropy	92
4.5. Noise and Cross-Talk	95
4.6. Noisy Pictures and Coding Efficiency	98
4.7. Error Correcting Codes	101
4.8. Capacity of a Noisy Channel	104
4.9. Shannon's Noisy Channel Coding Theorem	104
4.10. Why the Theorem is True	109
4.11. Summary	110
5. Entropy of Continuous Variables	111
5.1. Introduction	111
5.2. The Trouble With Entropy	112
5.3. Differential Entropy	115
5.4. Under-Estimating Entropy	118
5.5. Properties of Differential Entropy	119
5.6. Maximum Entropy Distributions	121
5.7. Making Sense of Differential Entropy	127
5.8. What is Half a Bit of Information?	128
5.9. Summary	132
6. Mutual Information: Continuous	133
6.1. Introduction	133
6.2. Joint Distributions	135
6.3. Conditional Distributions and Entropy	139
6.4. Mutual Information and Conditional Entropy	143
6.5. Mutual Information is Invariant	147
6.6. Kullback–Leibler Divergence and Bayes	148
6.7. Summary	150
7. Channel Capacity: Continuous	151
7.1. Introduction	151
7.2. Channel Capacity	151
7.3. The Gaussian Channel	152
7.4. Error Rates of Noisy Channels	158
7.5. Using a Gaussian Channel	160
7.6. Mutual Information and Correlation	162
7.7. The Fixed Range Channel	164
7.8. Summary	170

8. Thermodynamic Entropy and Information	171
8.1. Introduction	171
8.2. Physics, Entropy and Disorder	171
8.3. Information and Thermodynamic Entropy	174
8.4. Ensembles, Macrostates and Microstates	176
8.5. Pricing Information: The Landauer Limit	177
8.6. The Second Law of Thermodynamics	179
8.7. Maxwell's Demon	180
8.8. Quantum Computation	183
8.9. Summary	184
9. Information As Nature's Currency	185
9.1. Introduction	185
9.2. Satellite TVs, MP3 and All That	185
9.3. Does Sex Accelerate Evolution?	188
9.4. The Human Genome: How Much Information?	193
9.5. Enough DNA to Wire Up a Brain?	194
9.6. Are Brains Good at Processing Information?	195
9.7. A Very Short History of Information Theory	206
9.8. Summary	206
Further Reading	207
A. Glossary	209
B. Mathematical Symbols	217
C. Logarithms	221
D. Probability Density Functions	223
E. Averages From Distributions	227
F. The Rules of Probability	229
G. The Gaussian Distribution	233
H. Key Equations	235
Bibliography	237
Index	241