Information Theory: Entropy, Coding, and Channel Capacity

Computational Science Templates

November 24, 2025

1 Introduction

Information theory quantifies information content, compression limits, and communication channel capacity. Founded by Claude Shannon in 1948, it provides the mathematical framework for data compression (source coding) and error-correcting codes (channel coding). This analysis computes entropy for various probability distributions, demonstrates Huffman and arithmetic coding efficiency, explores mutual information, and examines channel capacity.

2 Mathematical Framework

2.1 Shannon Entropy

The entropy of a discrete random variable X with probability mass function $p(x)$:

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x) \tag{1}$$

2.2 Conditional Entropy and Mutual Information

Conditional entropy and mutual information:

$$H(X|Y) = -\sum_{x,y} p(x,y) \log_2 p(x|y) \tag{2}$$

$$I(X;Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y) \tag{3}$$

2.3 Channel Capacity

For a discrete memoryless channel:

$$C = \max_{p(x)} I(X;Y) \tag{4}$$

For the binary symmetric channel with crossover probability p:

$$C = 1 - H_b(p) = 1 + p \log_2 p + (1-p) \log_2(1-p) \tag{5}$$

3 Environment Setup

4 Entropy and Information Content

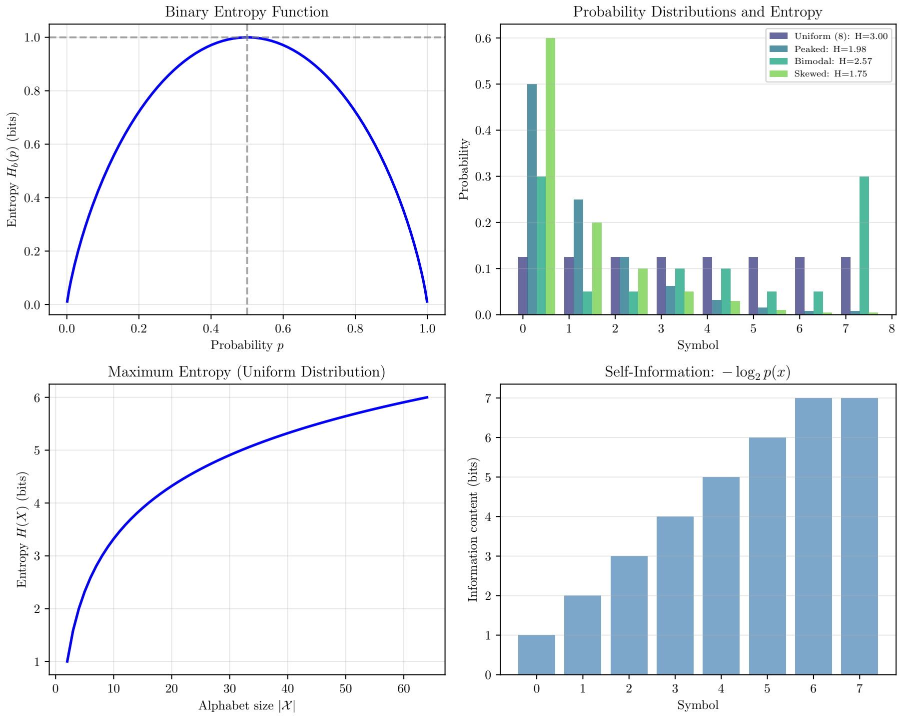

Figure 1: Entropy fundamentals: binary entropy, distributions, and information content.

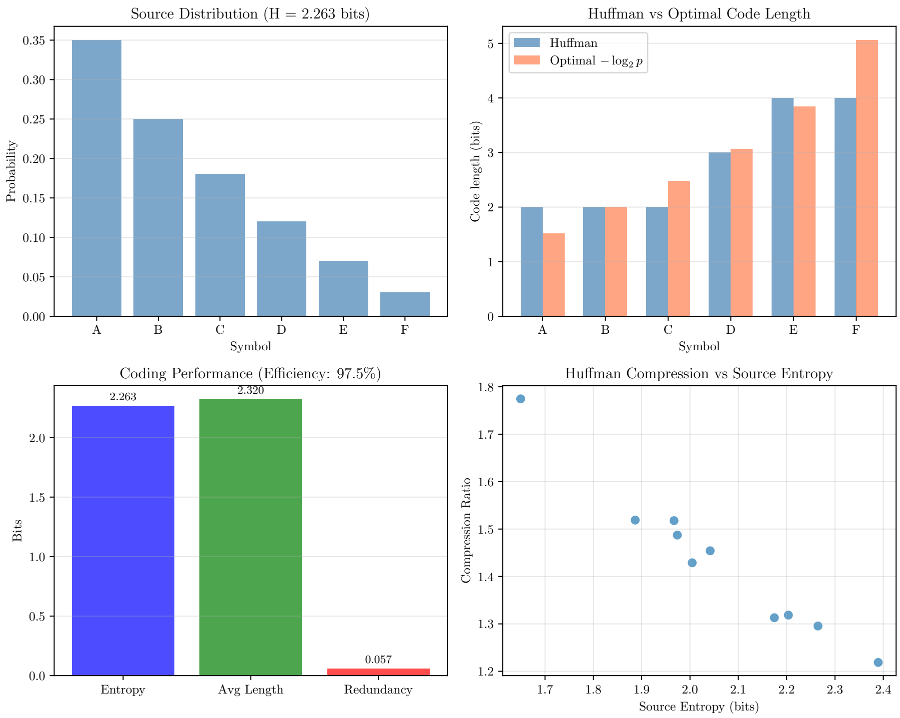

Figure 2: Huffman coding: source distribution, code lengths, and compression performance.

Table 1: Summary of Information Theory Results

Measure	Value	Description
	Huffman Coding	
Source entropy	2.263 bits	$H(X)$
Average code length	2.320 bits	$\bar{L}$
Coding efficiency	97.5%	$H(X)/\bar{L}$
Redundancy	0.057 bits	$\bar{L} - H(X)$
	Mutual Information	
$H(X)$	1.000 bits	Marginal entropy
$H(Y)$	1.000 bits	Marginal entropy
$H(X,Y)$	1.722 bits	Joint entropy
$I(X;Y)$	0.278 bits	Mutual information
	Channel Capacity	
BSC ($p = 0.1$)	0.531 bits/use	$1 - H_b(p)$
BEC ($\epsilon = 0.1$)	0.900 bits/use	$1 - \epsilon$
AWGN (10 dB)	1.730 bits/use	$\frac{1}{2}\log_2(1 + SNR)$

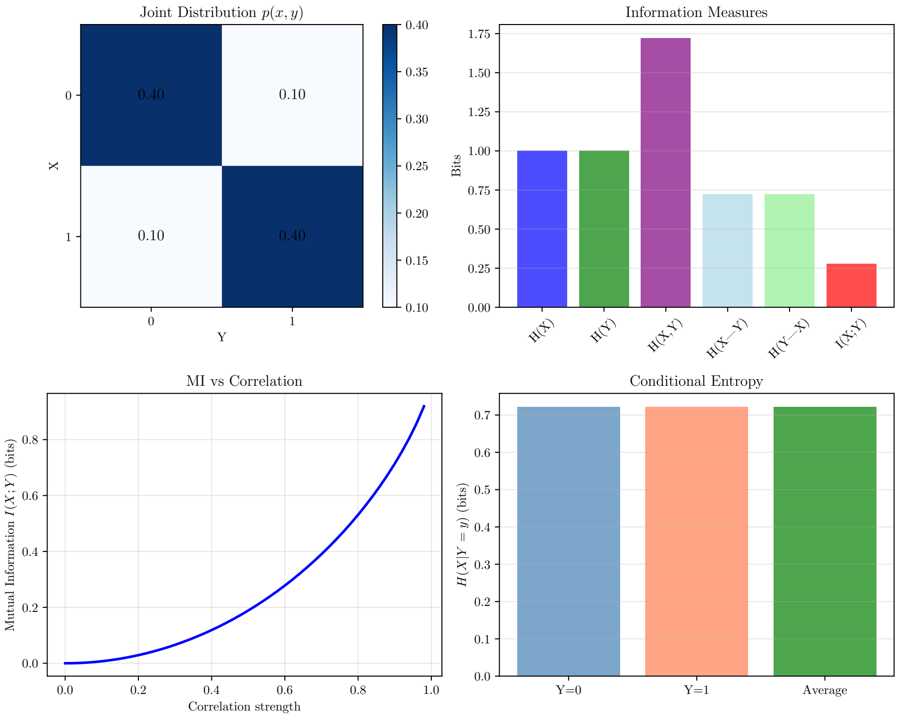

Figure 3: Mutual information: joint distribution, information measures, and correlation.

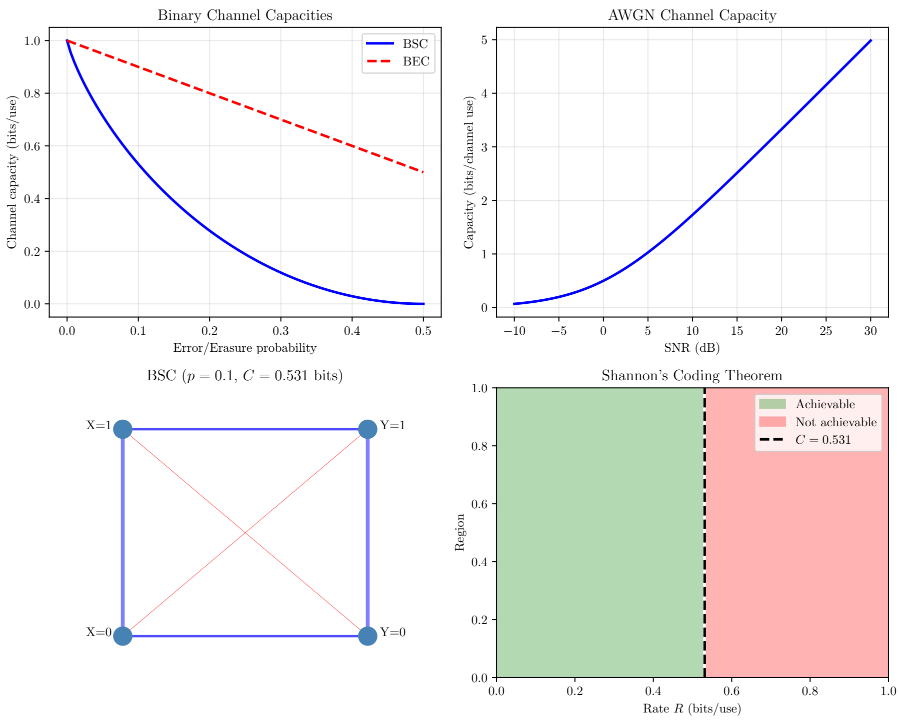

Figure 4: Channel capacity: BSC, BEC, AWGN, and achievable rates.

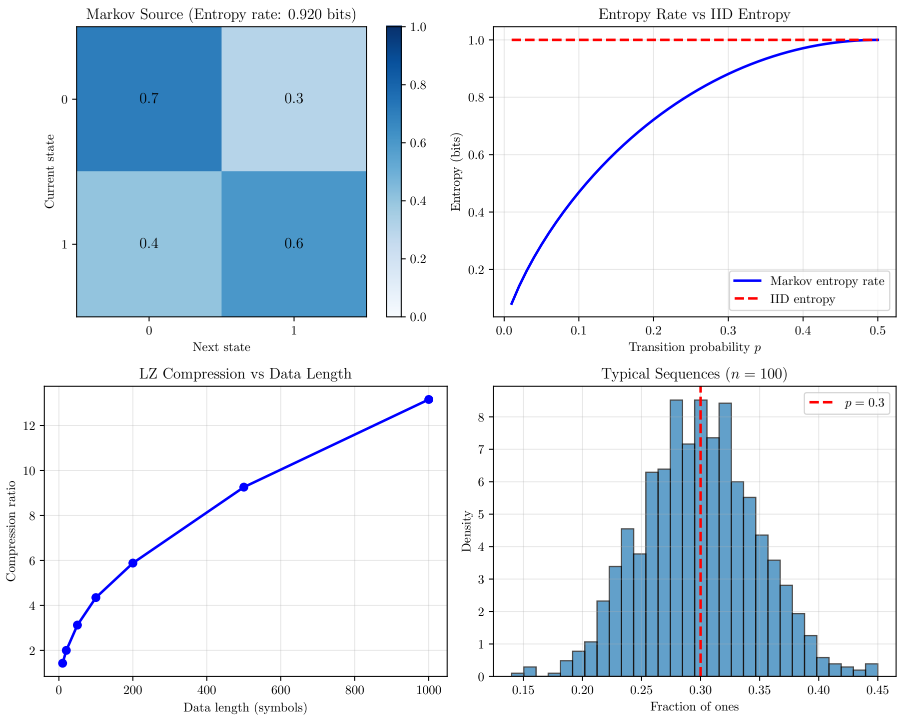

Figure 5: Source coding: Markov entropy, compression ratio, and typical sequences.

5 Huffman Coding

6 Mutual Information and Joint Entropy

7 Channel Capacity

8 Source Coding and Compression

9 Results Summary

10 Statistical Summary

- **Source entropy**: 2.263 bits
- **Maximum entropy** ($|\mathcal{X}| = 6$): 2.585 bits
- **Average Huffman code length**: 2.320 bits
- **Coding efficiency**: 97.5%
- **Joint entropy** $H(X, Y)$: 1.722 bits
- **Mutual information** $I(X; Y)$: 0.278 bits
- **Markov entropy rate**: 0.920 bits
- **BSC capacity** ($p = 0.1$): 0.531 bits/use

11 Conclusion

Information theory provides fundamental limits for data compression and communication. Shannon entropy measures the average information content and sets the minimum average code length. Huffman coding achieves near-optimal compression for known probability distributions. Mutual information quantifies the shared information between random variables, critical for channel coding. The channel coding theorem guarantees error-free communication at rates below capacity, achieved by modern codes like turbo codes and LDPC codes. These principles underpin data compression (ZIP, JPEG), error-correcting codes (WiFi, 5G), and cryptography.