Shannon Entropy

Systems Thinking · Module 7 — Information Flow

Node 7.1

Position in Knowledge Graph

This node formalizes Shannon entropy as the foundational quantitative measure of information.

It builds on: - Information-Theoretic Complexity (6.3)

It prepares for: - Channel Capacity (7.2) - Signal Propagation (7.3)

Shannon entropy provides the fundamental unit for analyzing information flow in systems.

Formal Statement / Definition / Construction

Let X be a discrete random variable with finite alphabet $\mathcal{X}$.

Let probability mass function:

$$p(x) = \mathbb{P}(X = x), \quad x \in \mathcal{X}.$$

Definition (Shannon Entropy)

The Shannon entropy of X is:

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x).$$

Logarithm base determines units: - base 2 $\to$ bits, - base e $\to$ nats.

Joint Entropy

For random variables X, Y:

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x, y).$$

Conditional Entropy

$$H(X|Y) = -\sum_{x,y} p(x,y) \log p(x|y).$$

Equivalent identity:

$$H(X|Y) = H(X,Y) - H(Y).$$

Chain Rule

$$H(X,Y) = H(X) + H(Y|X).$$

Assumptions and Conditions

1. Alphabet $\mathcal{X}$ finite.
2. $0 \log 0$ defined as 0 by continuity.
3. All sums finite.
4. Logarithm base fixed.

Continuous variables require differential entropy (not treated here).

Proof or Mathematical Development

Maximum Entropy

Suppose $|\mathcal{X}| = m$.

Maximize entropy subject to:

$$\sum_x p(x) = 1.$$

Using Lagrange multipliers:

$$\mathcal{L} = -\sum_x p(x) \log p(x) + \lambda \left(\sum_x p(x) - 1 \right).$$

Take derivative:

$$\frac{\partial \mathcal{L}}{\partial p(x)} = -(1 + \log p(x)) + \lambda = 0.$$

Thus:

$$p(x) = e^{\lambda - 1}.$$

Since sum equals 1:

$$p(x) = \frac{1}{m}.$$

Therefore:

$$H(X) = \log m.$$

Uniform distribution maximizes entropy.

Non-Negativity

Since $0 \leq p(x) \leq 1$, we have $-\log p(x) \geq 0$, thus:

$$H(X) \geq 0.$$

Equality occurs if and only if X is deterministic.

Structural Role

Shannon entropy:

- Quantifies uncertainty.
- Measures dispersion of probability mass.
- Provides basis for information transmission limits.
- Connects randomness to complexity.

All information flow analysis builds upon entropy identities.

Structural Compression

Invariant

Entropy:

$$H(X) = -\sum_x p(x) \log p(x).$$

Mechanism

Logarithmic weighting of probability produces additive information measure.

Cross-System Echo

- Physics: thermodynamic entropy.
- Economics: uncertainty in markets.
- Biology: genetic diversity.
- Networks: communication capacity limits.

Systems Thinking · Module 7 — Information Flow · Node 7.1
Concepts: shannon-entropy, information-theory, random-variables
Prerequisites: 6.3
Enables: 7.2, 7.3

hbar.university — own.your.cognition