Entropy Exercise

Multimedia Technology

Tutorial 1, section 2

Entropy Calculation Exercise

Assume an information source that produces eight symbols with probabilities 1/4, 1/4, 1/8, 1/8, 1/8, 1/16, 1/32, and 1/32.

1. Calculate the entropy of this source,
2. Calculate the entropy of a source that produces eight equally probable symbols, and
3. Compare the two entropies.

Entropy Theory - Reminder

- Quantifies the average amount of information of a source,
- Average number of bits of information required to represent the symbols the source produces

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) = -\sum_{i=1}^{n} P(x_i) \log P\,(x_i)$$

  - X represents the symbols of an alphabet.
- "A mathematical theory of communication,"
  - *C. E. Shannon, Bell Systems Technical Journal, 1948*
- Refer to the lecture slides for more information. (04 - Information Theory)

Entropy Calculation Exercise – Question 1

Symbol	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Probability	1/4	1/4	1/8	1/8	1/8	1/16	1/32	1/32

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) = -\sum_{i=1}^{n} P(x_i) \log P(x_i)$$

For x_0 and x_1, we have: $-1/4 \log_2 1/4 = -1/4 \log_2 4^{-1} =$
$= 1/4 \log_2 2^2 = 2/4$, since $\log_2 2 = 1$

For x_2, x_3 and x_4, we have: $-1/8 \log_2 1/8 = -1/8 \log_2 8^{-1} =$
$= 1/8 \log_2 2^3 = 3/8$, since $\log_2 2 = 1$

For x_5 we have: $-1/16 \log_2 1/16 = -1/16 \log_2 16^{-1} =$
$= 1/16 \log_2 2^4 = 4/16$, since $\log_2 2 = 1$

For x_0 and x_1, we have: $-1/32 \log_2 1/32 = -1/32 \log_2 32^{-1} =$
$= 1/32 \log_2 2^5 = 5/32$, since $\log_2 2 = 1$

Thus, we have:
H(X) = 2/4 + 2/4 + 3/8 + 3/8 + 3/8 + 4/16 + 5/32 + 5/32
= (16 + 16 + 12 + 12 + 12 + 8 + 5 + 5) / 32 =
= 86/32 = 2.6875

Entropy Calculation Exercise – Question 2

We now have 8 symbols with equal probability, thus the probability for each one of them is 1/8.

Symbol	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Probability	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) = -\sum_{i=1}^{n} P(x_i) \log P(x_i) \Longrightarrow H(X) = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = \log n$$

In this question we have a source with random behavior, thus we can conclude that it requires $\log_2 8 = \log_2 2^3 = 3$, since $\log_2 2 = 1$

Entropy Calculation Exercise – Question 3

- Obviously, 3 > 2.6875, but *what conclusion can we reach about random sources?*
- Q1: some symbols have higher probabilities, meaning they contribute less to the overall uncertainty of the source.
  - The entropy is lower, indicating less unpredictability in the information produced by this source.
- Q2: there is <u>maximum uncertainty</u> about which symbol will be produced, as no symbol is favored over another.

Source	Q1	Q2
H(X)	2.6875	3

- The more evenly distributed the probabilities across possible outcomes, the higher the entropy.
- Random sources maximize entropy.
- Higher entropy values indicate greater uncertainty or unpredictability in the source's output.
- Lower entropy suggests a more predictable source.