“Information is the resolution of uncertainty.”

Shannon

Erfan Mirzaei

Information Theory Mini-Course

Sep 2022

Session1

Information Entropy

Information

- What is information?
  - Philosophical question
  - All animals, humans, even plants communicate information. [Language]
  - What is conveyed or represented by a particular arrangement or sequence of things.
- Can we measure information? [What is our intuition?]
- You are lost in an island and can not hear the voice of each other:
-  OR  [You have enough of them.]
- Persian Language: 32 letters

  Sol: Send All the letters respectively, with how many stones?

Information

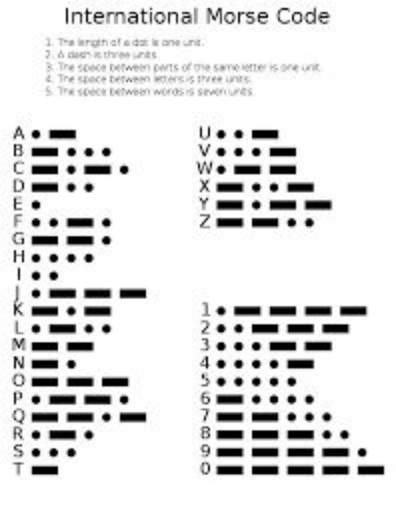
- Can you do better?
  - Yes, you can :)
  - Log2 26 = 4.8 vs. 4.1257
- Underlying Assumption:

  # bits = Log N —> They're equiprobable.

  # bits = Log 1/(1/N) = Log 1/p

  *Note: Base of logarithm depends on your language.

  Probability of a random event is reciprocal proportional with number of bits needed to represent it.



Information vs. Probability vs. Surprise

- Imagine yourself in the describe situation, and now think you didn't receive stones. Now you should guess that what had been the letter?

- It is easier to guess in Persian or English?

- If an oracle say the right answer to you, in which case you gain more information?

- Less probable events → more surprising events → events which contain more information if happen/reveal → # bits need to describe events

Information measure: Entropy

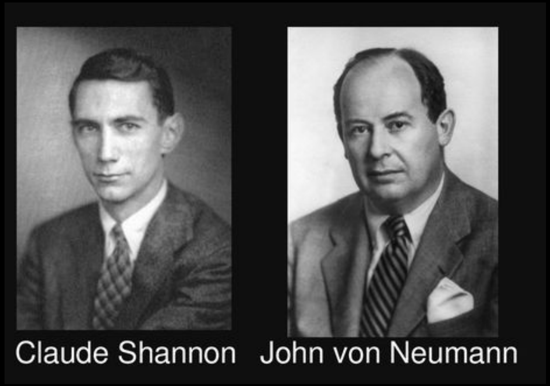
- Properties of Information
  - Deterministic Outcome → zero information
  - Probability is Decreasing → information increasing
  - Information content of independent random variables is additive.
- 1948: “Mathematical Theory of communication”,Shannon

$$H(X = x) = \text{Log2 } 1/P(X{=}x) = -\text{ Log2 } P(X = x)$$

I think our example can explain why we should use log

for information measure and its relation to bits.

Do you agree with me?



Information measure: Entropy

- Def of Entropy of a random variable X is: [Expected Value of Log2 -P(X)]

$$H(x) = -\sum_x P(x)\cdot \log P(x)$$

$$= \sum_x \underbrace{P(x)}_{\text{The prob. of event X}} \cdot \underbrace{\log\left(\frac{1}{P(x)}\right)}_{\text{WHAT IS THIS ?}}$$

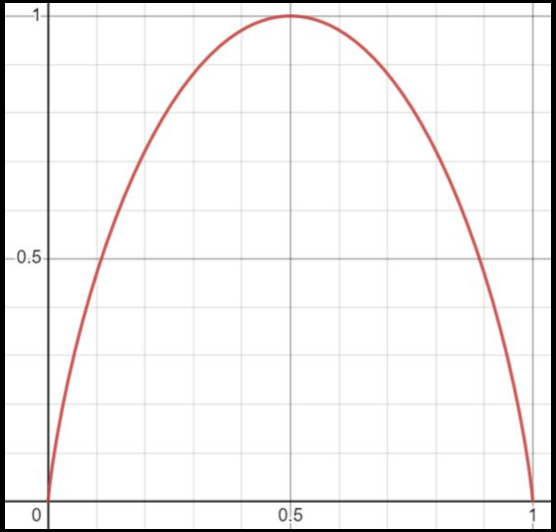
$X = \{1,2,3,4,5,6\}$

$H(X) = 6 * \log(1/6) = 2.48$ bits

Entropy properties

- H(X) >= 0
- H(X) = H(2X) —> H(X), g is one-2-one func. → H(g(X)) = H(X)
- X = {0 with p,1 with 1-p} →

$$H(X) = -p * \log2\ p - (1-p) * \log2\ (1-p)$$



Its argmax is 0.5. When events are equiprobable.

i.e when probability distribution is uniform.

p = 0, p = 1 → H(X) = 0

Thanks for your attention

Erfan Mirzaei
erfunmirzaei@gmail.com