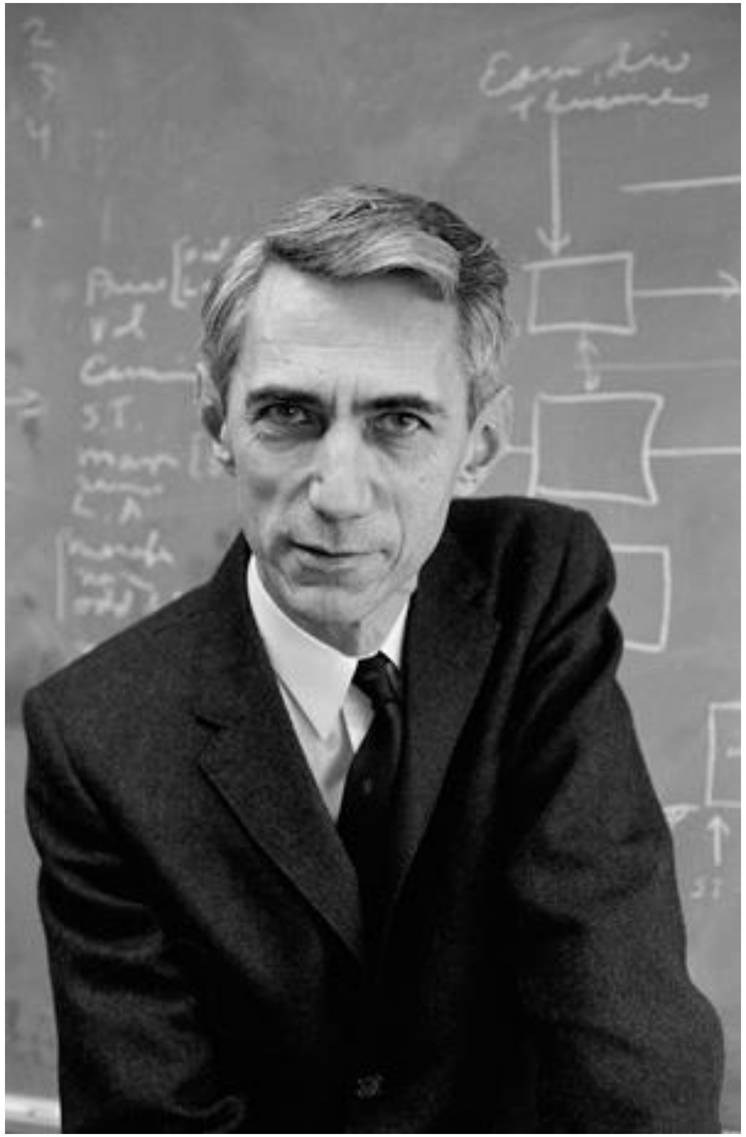
Week 13-15: Engineering II – Shannon’s Information Theory

Kyeong Soo (Joseph) Kim

XJTLU

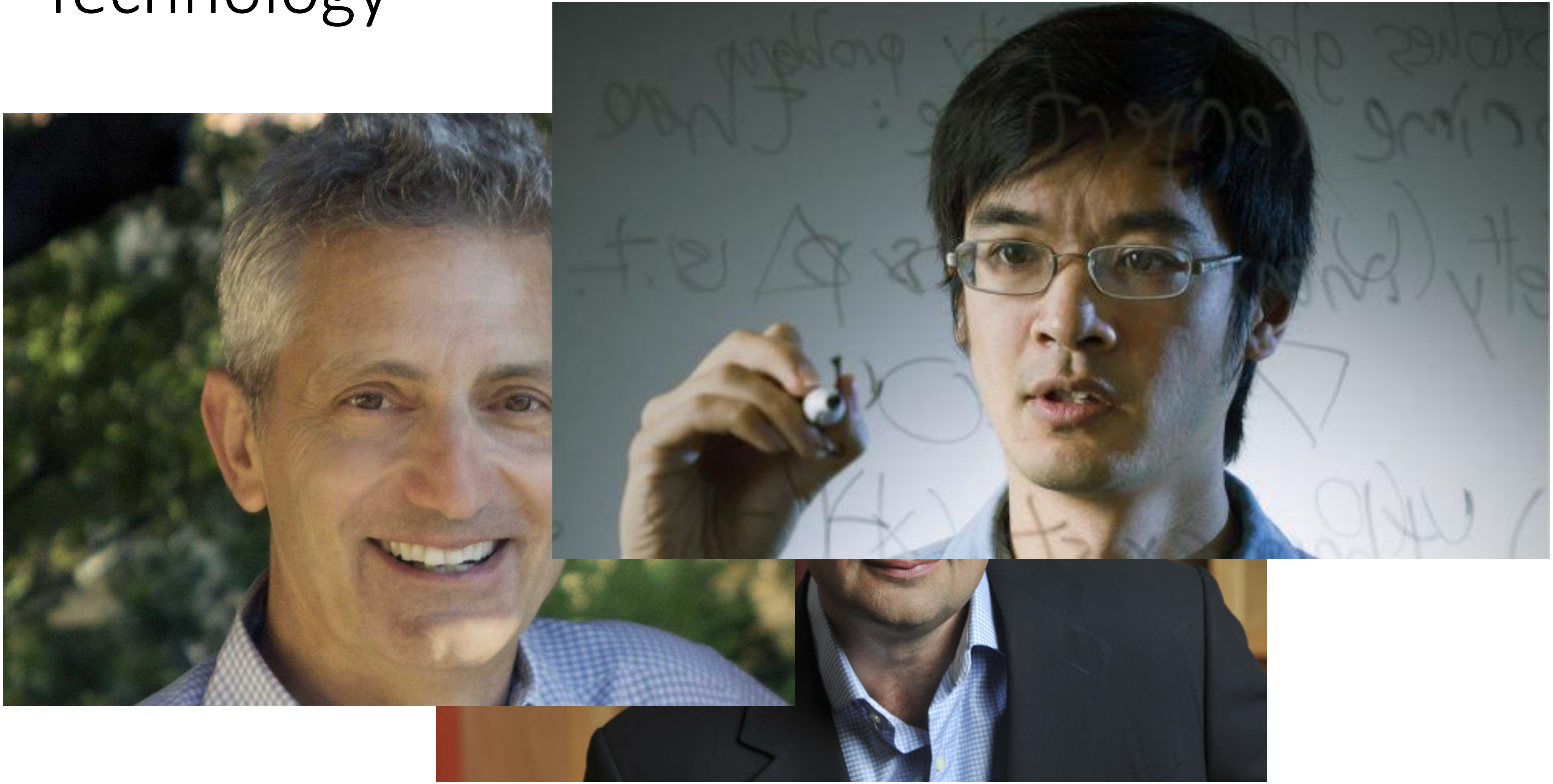
Claude E. Shannon (1916-2001)[*]: Father of Information Theory



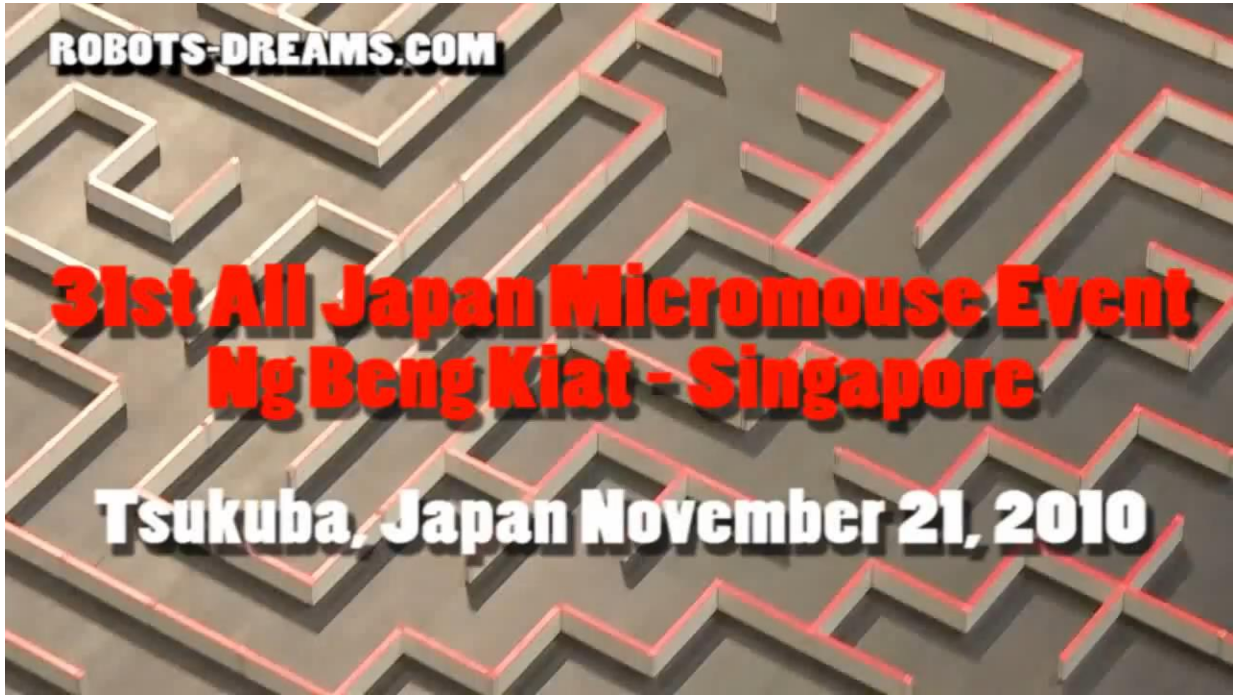
- An American mathematician, electrical engineer, and cryptographer.
- Founded information theory with a landmark paper "A Mathematical Theory of Communication" published in 1948.
- Provided a framework where we can study how to handle information, i.e.,
  - Quantification
  - Storage
  - Communication

**** Based on Wikipedia.***

Mathematicians Active in Science & Technology



Shannon’s Mouse



Fundamental Questions

- What is the ultimate data compressions?
  - Answer: Entropy ***H***.

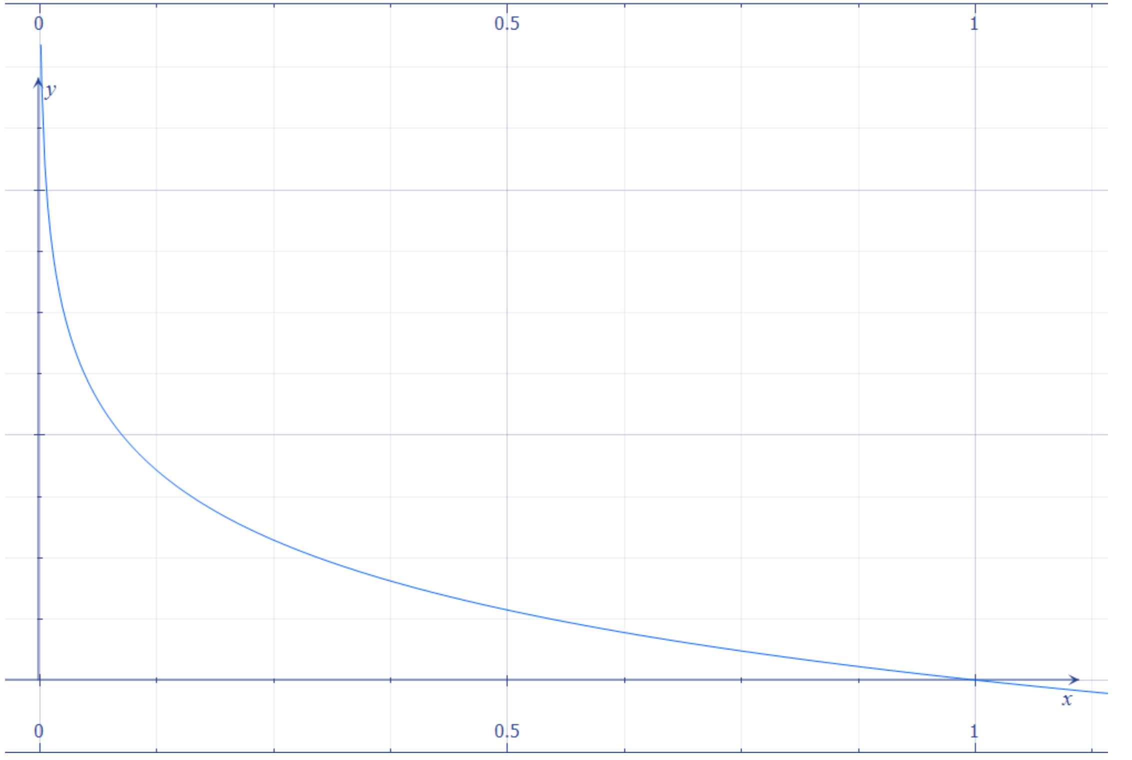
- What is the ultimate transmission rate of communication?
  - Answer: Channel capacity ***C***.

Information and Entropy

- The entropy of $H(X)$ of a discrete random variable X with alphabet $\mathcal{X}$ and probability mass function $p(x) = \Pr\{X = x\}, x \in \mathcal{X}$, is defined as

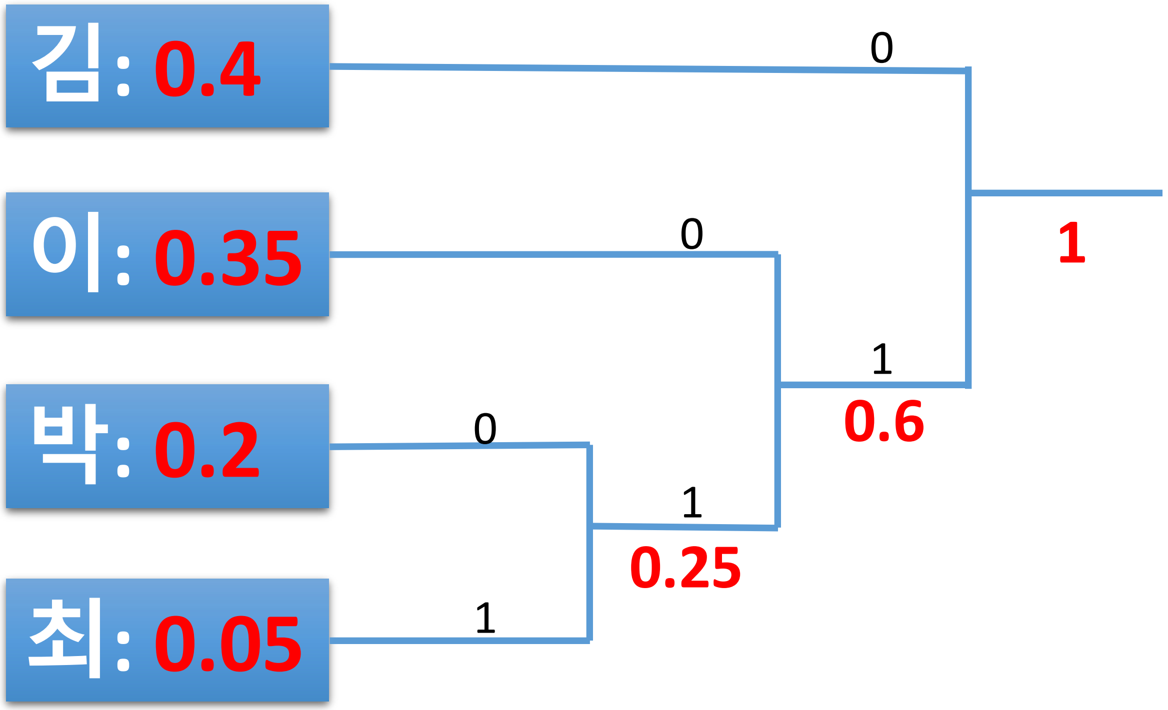
$$H(X) \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} \{-\log p(x) \times p(x)\} = E[-\log p(x)]$$

- The entropy is the ***measure of uncertainty*** of a random variable and interpreted as the expected value of the information carried by each alphabet, i.e., $-\log p(x)$ (shown in the figure).
  - Entropy is expressed in ***bits*** when the log is to the base 2, i.e., $-\log_2 p(x)$.



Huffman Encoding

- Entropy of the source: 1.74 bits/symbol
- Average length of binary coding: 2 bits/symbol
- Average length of Huffman coding: 1.85 bits/symbol



Symbol	Probability	Binary Coding	Huffman Coding
김	0.4	00	0
이	0.35	01	10
박	0.2	01	110
최	0.05	11	111

Channel Capacity

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

- C: Maximum channel capacity [bits/second]
- B: Bandwidth of the channel [Hz]
- S: Signal power [Watts]
- N: Noise power [Watts]

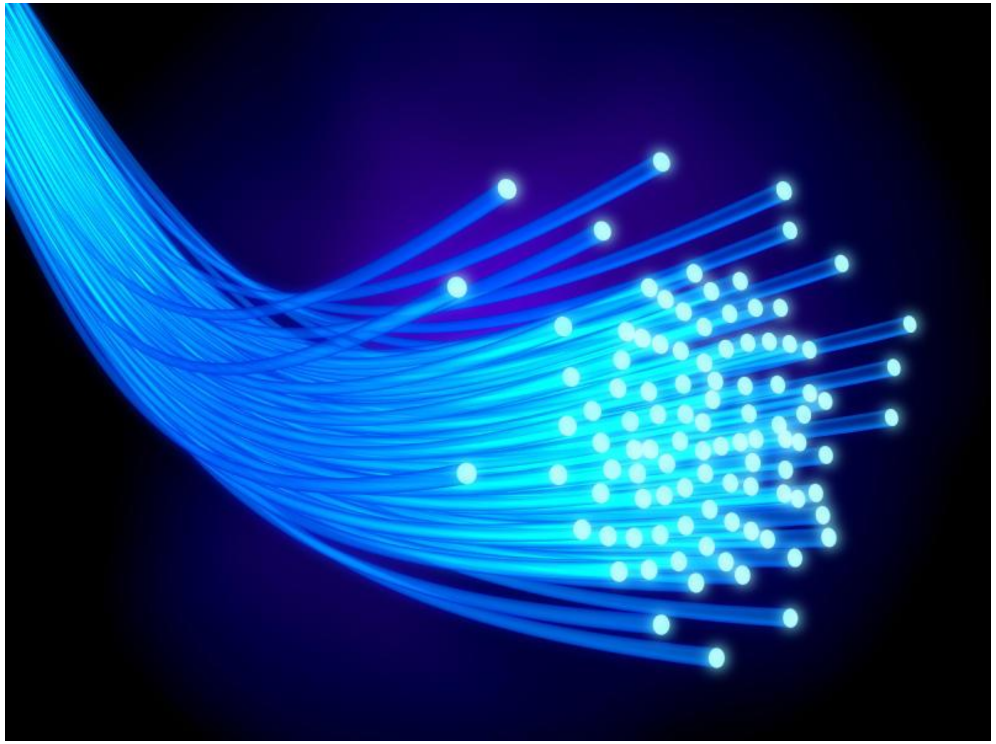
Capacity of Optical Fiber

- It is known that the maximum capacity of one strand of optical fiber is greater than 100 Terabits/second*, i.e.,

$$C > 100 \times 10^{12} bits/second$$

- The size of 2-hour-long HD-quality movie is around 4 Gigabytes, i.e., 32×10^9bits.
  - The download time is as follows:

$$\frac{32 \times 10^9}{100 \times 10^{12}} = 32 \times 10^{-5} s = 320\ \mu s = 0.32\ ms$$



*** Mitra & Stark, Nature, vol 411, June 28, 2001.***