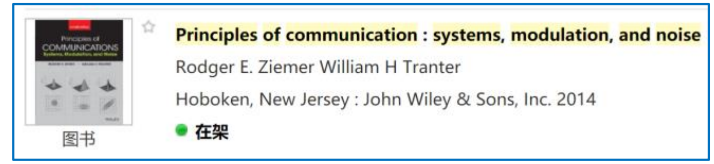
EE140 Introduction to Communication Systems

Lecture 9

Instructor: Prof. Lixiang Lian

ShanghaiTech University, Fall 2025



- Syllabus (second half)

Content	Hours	Week
Introduction to digital communication sys (Chapter 1)	1	9
Information Theory and Source Coding (Chapter 2, Chapter 12)	5	9&10
Sampling and Quantization (Chapter 3, 4)	6	10&11
Vector space and signal space (Chapter 5, Chapter 11)	6	12&13
Modulation and Demodulation (Chapter 6, Chapter 10)	6	13&14
Detection and Channel Coding (Chapter 8, Chapter 9,11,12)	6	15&16
Wireless Communication (Chapter 9)	2	16

What is Digital Communications?

Use a digital sequence as an interface between the source and the channel

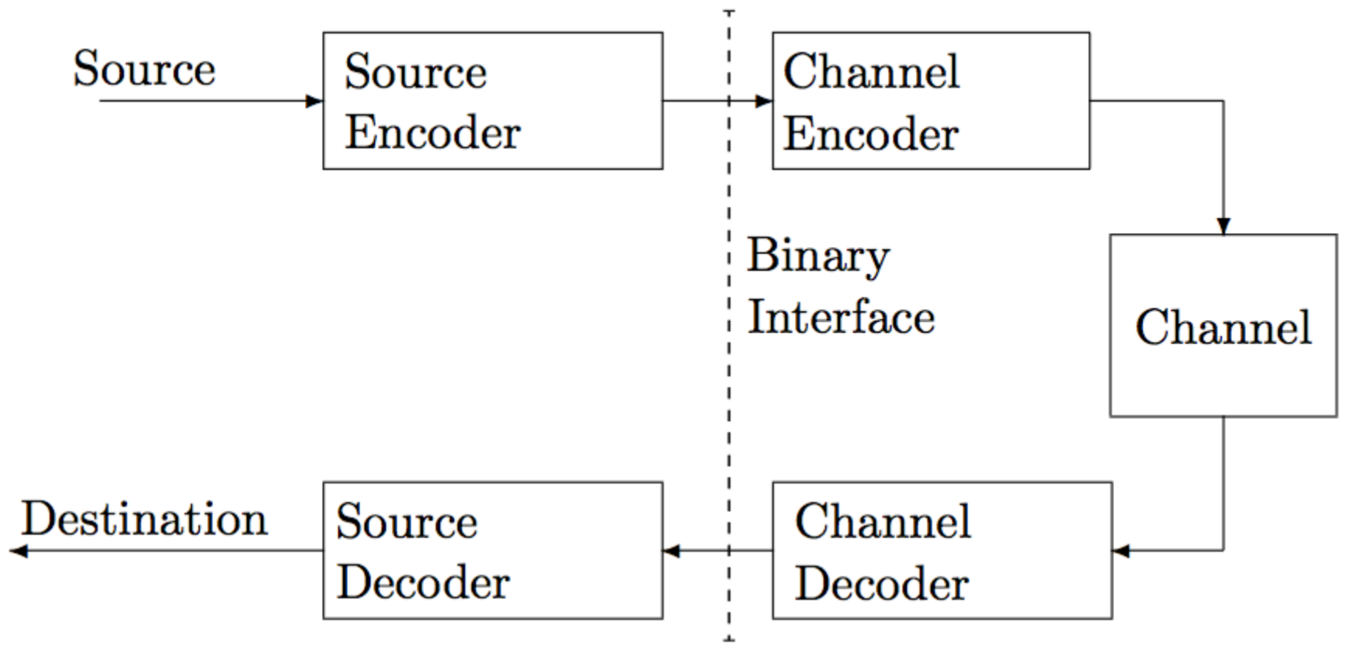
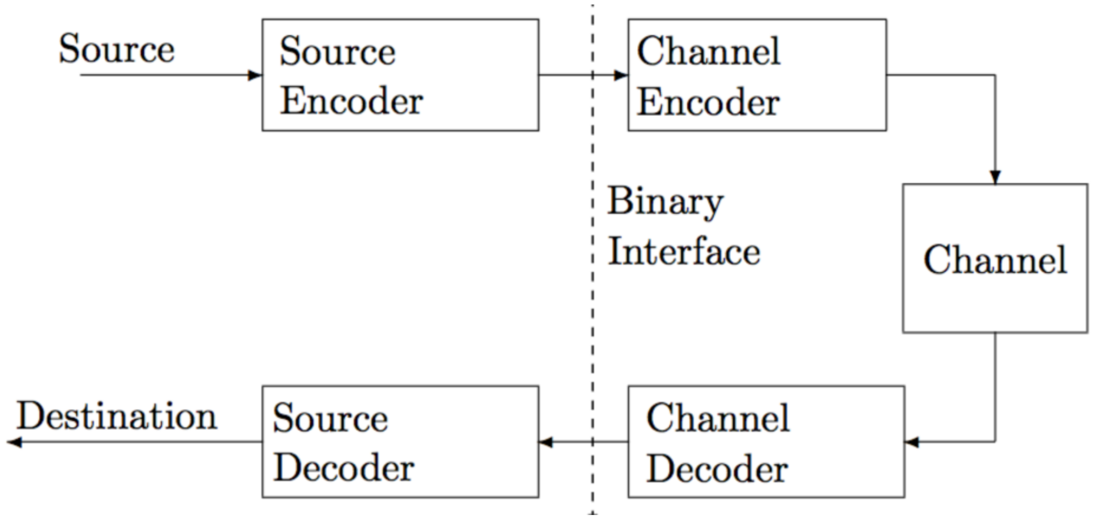


Figure: Separation of source and channel coding [Gallagar'Book]

Why need Digital Communications?

- Digital hardware has become so cheap, reliable and miniaturized.
- Simplify implementation and understanding
- Security
- Doing this won’t decrease the rate performance

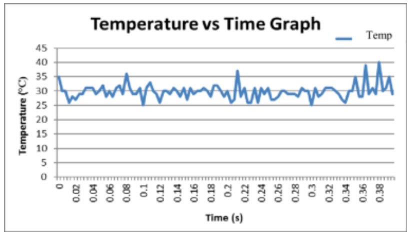
Digital Communication System



- Source
- Source Encoder ←→ Source Decoder
- Channel Encoder ←→ Channel Decoder
- Binary/Digital interface
- Channel

Digital Communication System

- Part 1: Source
- Important Classes of Sources:
  - Analog sources. E.g., voice, music, video and images etc. (We restrict to wave form sources, i.e. voice and music)
  - Discrete sources: A sequence of symbols from a known discrete alphabet. E.g. English letters, Chinese characters, binary digits etc.



Digital Communication System

- Part 2: Source Encoder

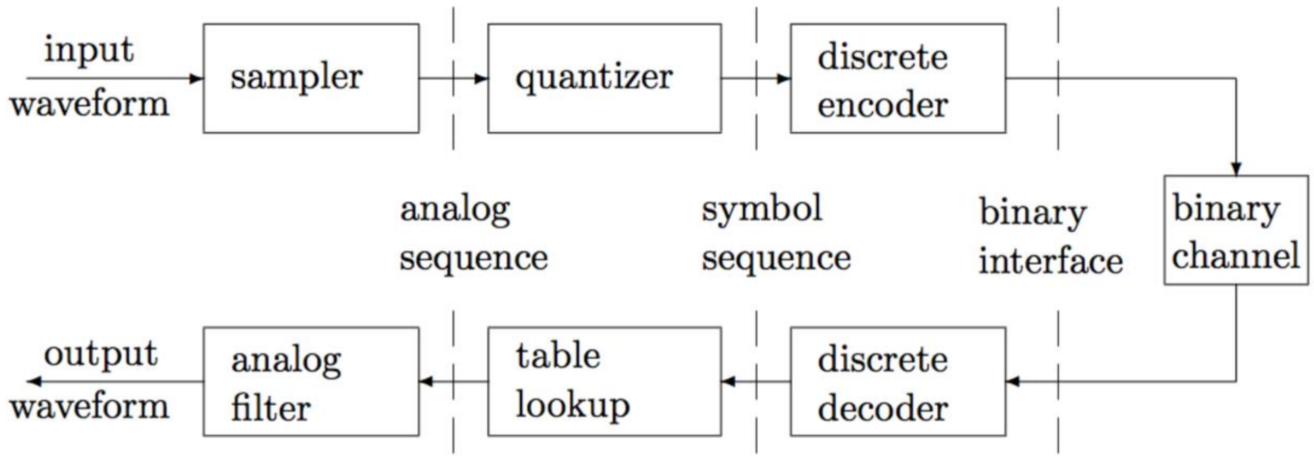


Figure: Layering of Source coding [Gallagar'Book]

- Converting the input to a sequence of bits
  - Discrete source: fixed length codes/variable-length codes
  - Analog source:
    - Sampling: Analog signal to sequence (Chapter 4)
    - Quantizer: Analog sequence into symbols (Chapter 3)
    - Encoder: Symbols to bits (Chapter 2)

Digital Communication System

- Part 3: Channel Encoder
  - Mapping the binary sequence into a channel waveform

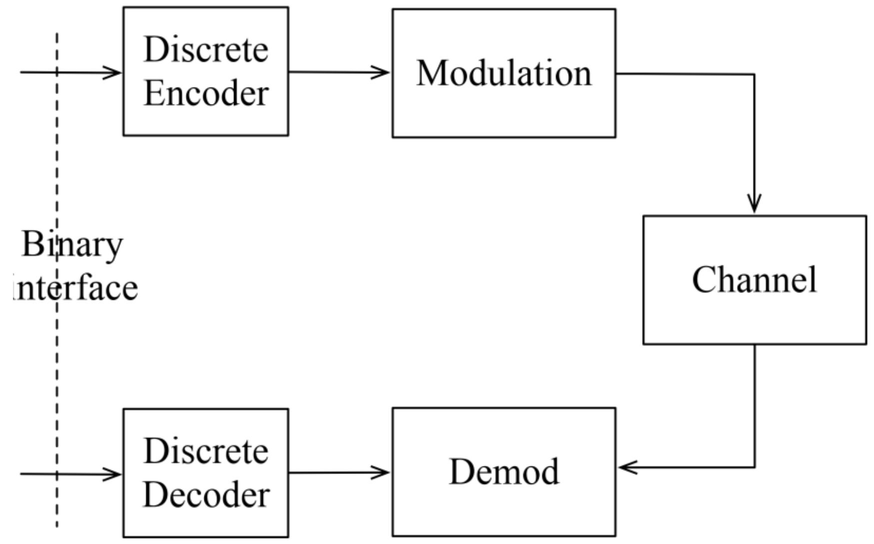
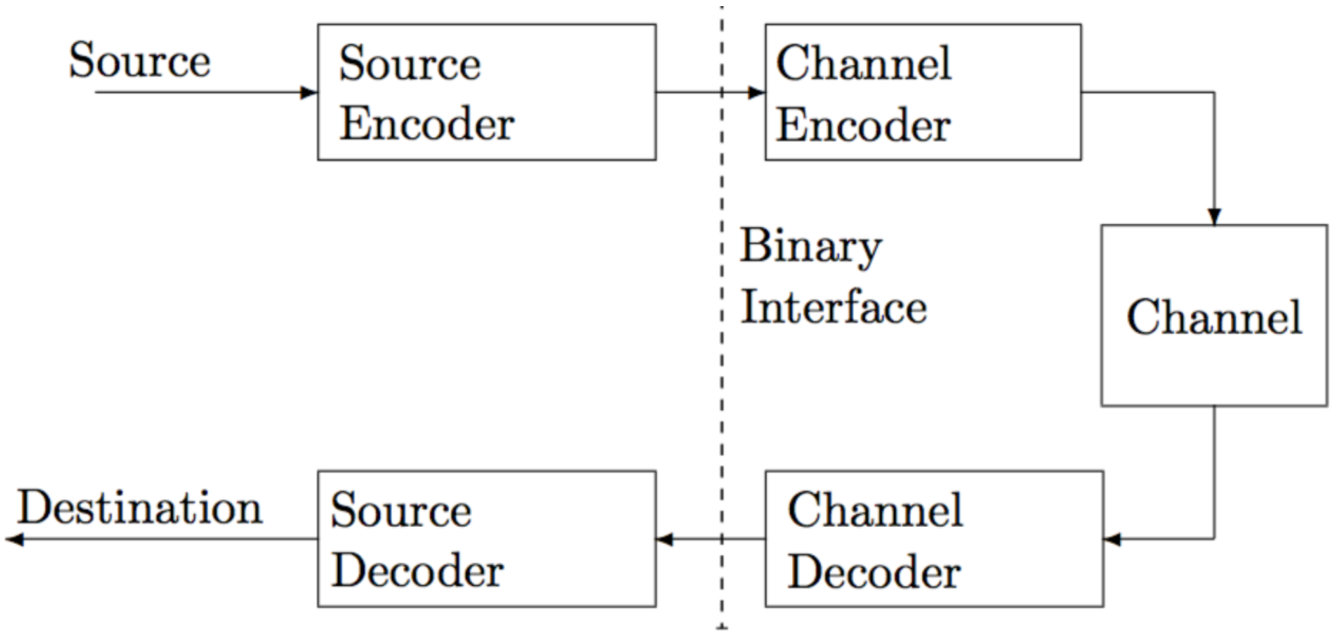


Figure: Layering of channel coding

- Discrete Enc (Chapter 8):
  - Add redundancy to improve reliability of communication
- Modulation (Chapter 6):
  - Maps the binary sequence to a baseband waveform
  - Maps the baseband to bandpass waveform

Digital Communication System

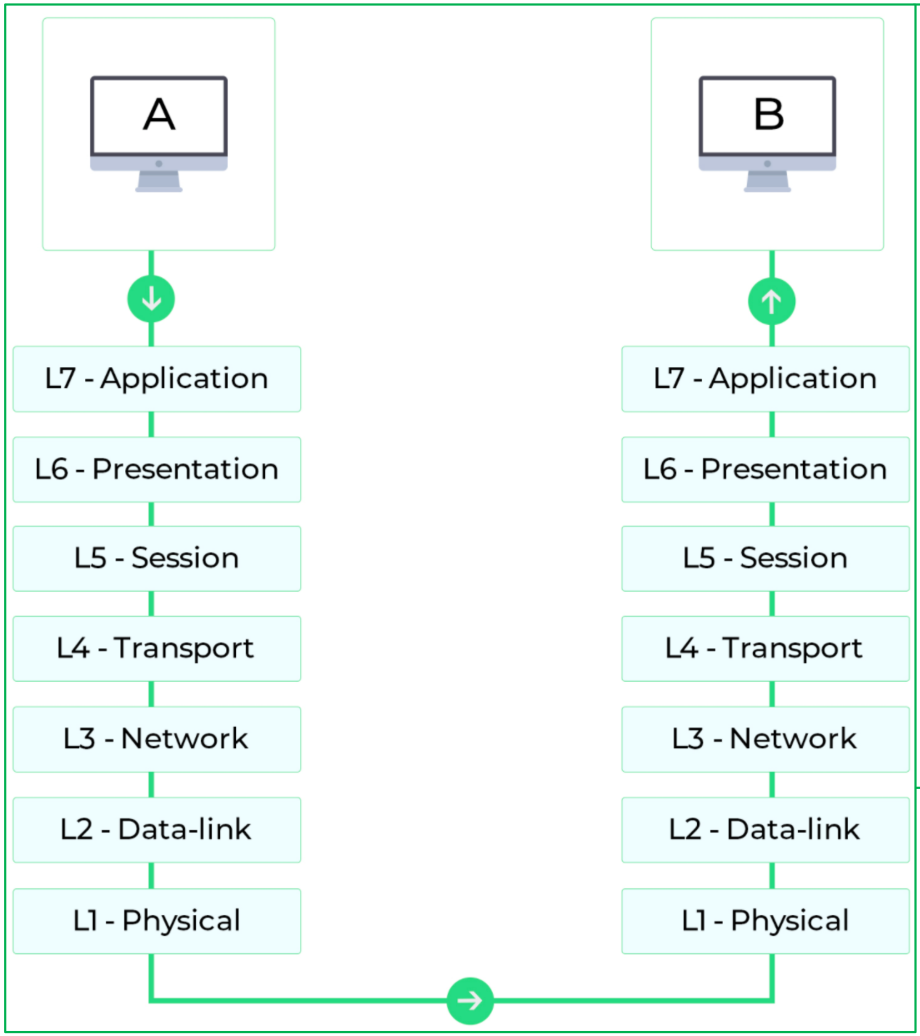
- Part 4 : Digital/Binary Interface



  - Complicating factors:
    - Unequal rates: the rate from source encoder doesn't match channel encoder (Solution: Buffer, queuing)
    - Errors: channel decoder makes errors which causes errors in source decoder (Solution: Good channel codes)
    - Networks: encoded source outputs are for various networks (Solution: Network protocol design)

Digital Communication System

- Part 4 : Digital/Binary Interface



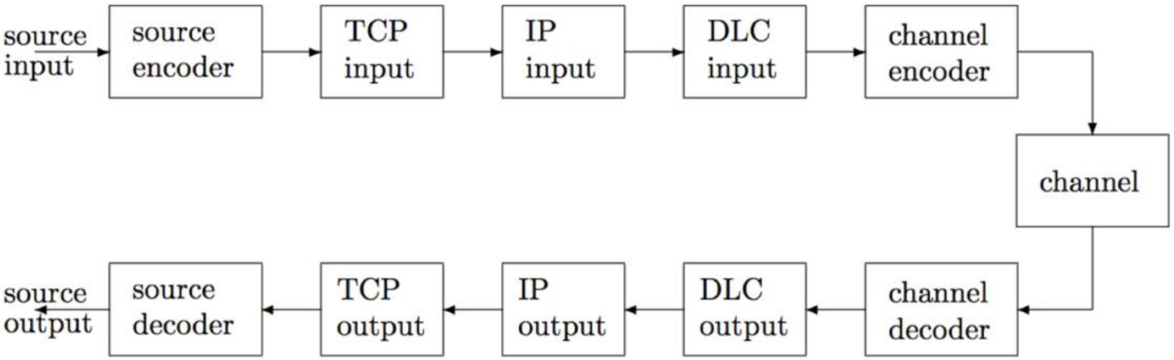


Figure: Binary interface for internet

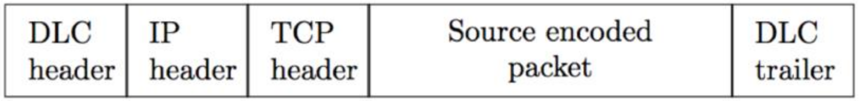


Figure: Structure of data frame

Digital Communication System

- Part 5 : Channel

- Properties on channel:
  - Channel is the part between the transmitter and receiver
  - Channel is given (not under control of designer)
  - Given the inputs, and outputs, the channel is a description of how the input affect the output. The description is usually probabilistic.

- Types of channel:
  - Memoryless (main focus) v.s. Memory
  - Discrete v.s. Continuous

Digital Communication System

- Part 5 : Channel

- Discrete memoryless channel (DMC)

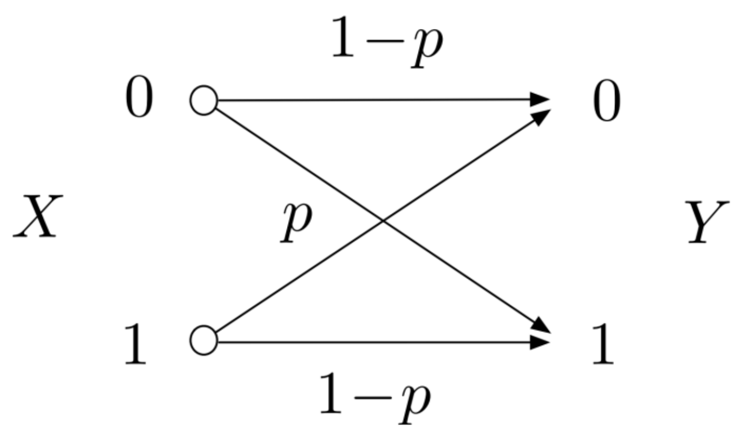


Figure: Binary symmetry channel

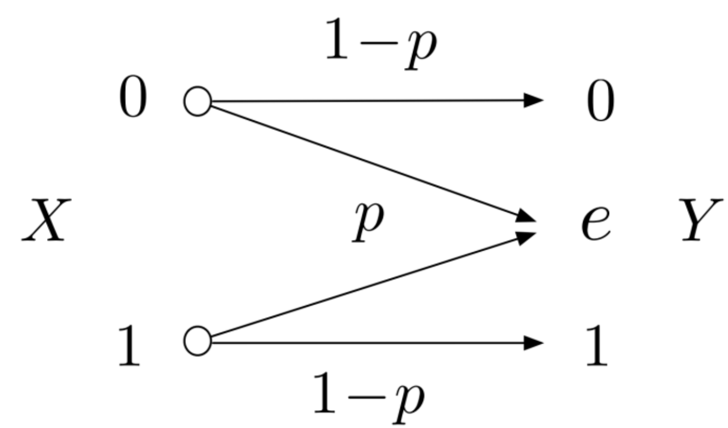


Figure: Binary erasure channel

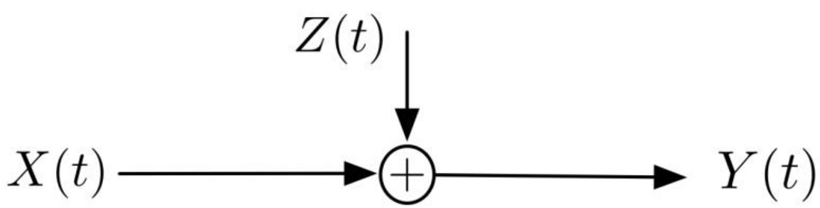
Digital Communication System

- Part 5 : Channel

- Continuous Channel

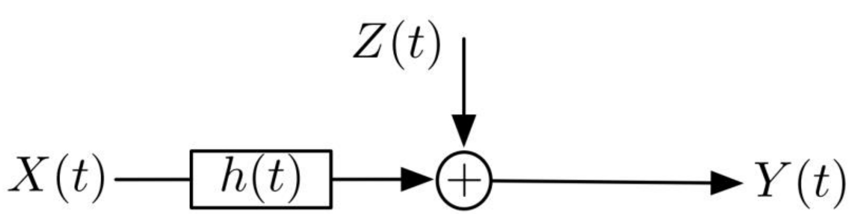
Given Gaussian noise $Z(t)$:

- Additive white Gaussian noise (AWGN) channel:



$$Y(t) = X(t) + Z(t)$$

- Linear Gaussian channel (with linear filter $h(t)$):

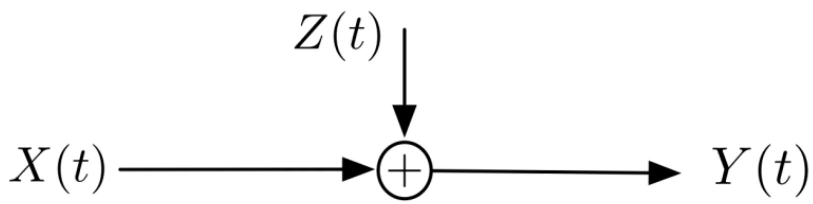


$$Y(t) = X(t) * h(t) + Z(t)$$

Digital Communication System

- Part 5 : Channel

- AWGN Channel



$$Y(t) = X(t) + Z(t)$$

- For the AWGN channel with bandwidth W, the capacity (in bps) is

$$C = W \log_2 \left(1 + \frac{P}{N_0 W}\right)$$

- This is the ultimate, but it is essential achievable in practice
- Wireless channels have added complications (Chapter 9)
  - Multiple physical paths from input to output
  - Random fluctuation in the strength of multipath.

Outline

- Information Theory
- Coding for Discrete Sources
- Sampling
- Quantization
- Vector spaces and signal space
- Channel, Modulation and Demodulation
- Detection, coding and decoding

Information Theory

- Reference books
- "A Mathematical Theory of Communication" by C. E. Shannon
- "Elements of Information Theory" by T. Cover (Chapt. 2&8)
- "Principle of Communications" by R. Ziemer
- "Information Theory and Network Coding" by R. Yeung

Q1: How to measure the quantity of information?

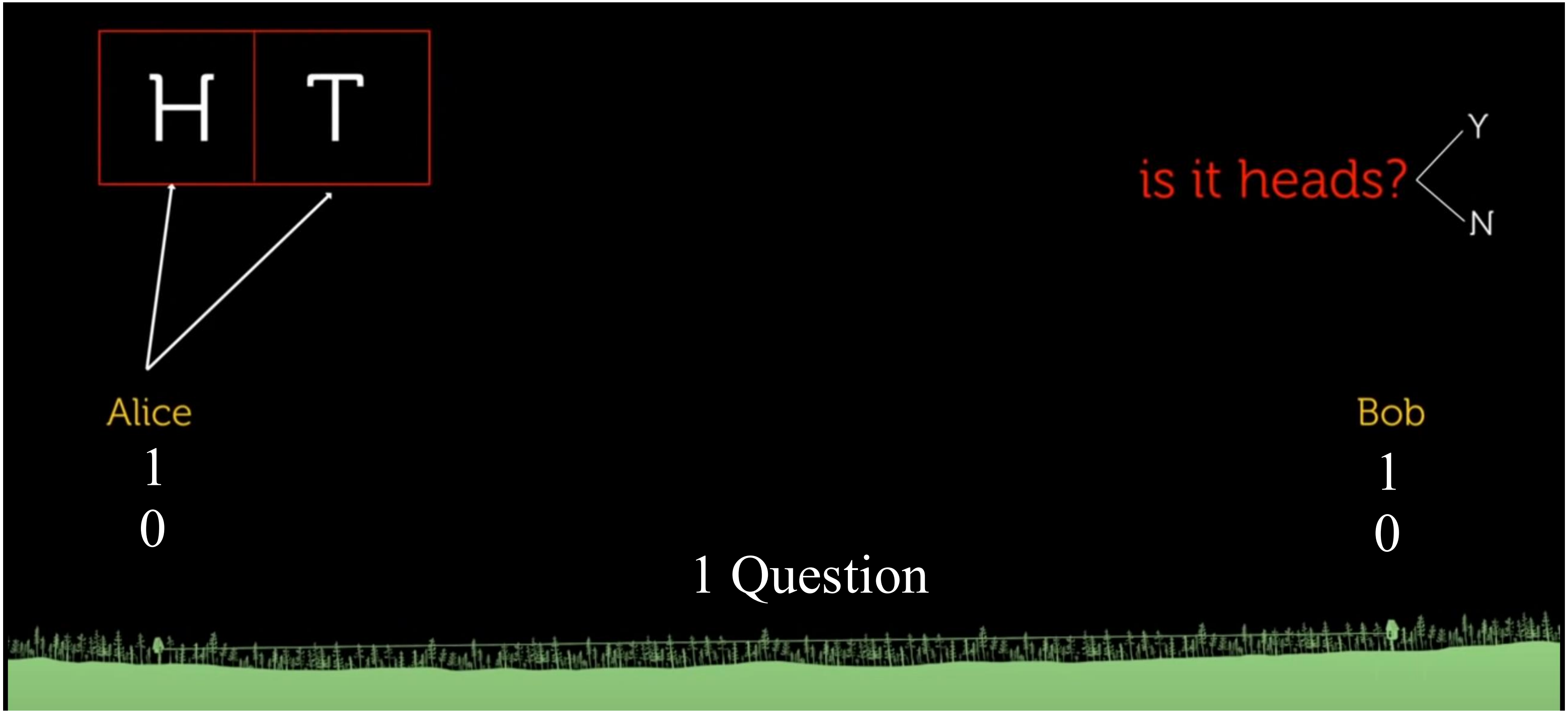
Example (*Football Games*):

China sucks at playing soccer, while France and Brazil both are very good at it. Which game result below contains more uncertainty?

- China V.S. Brazil
- France V.S. Brazil

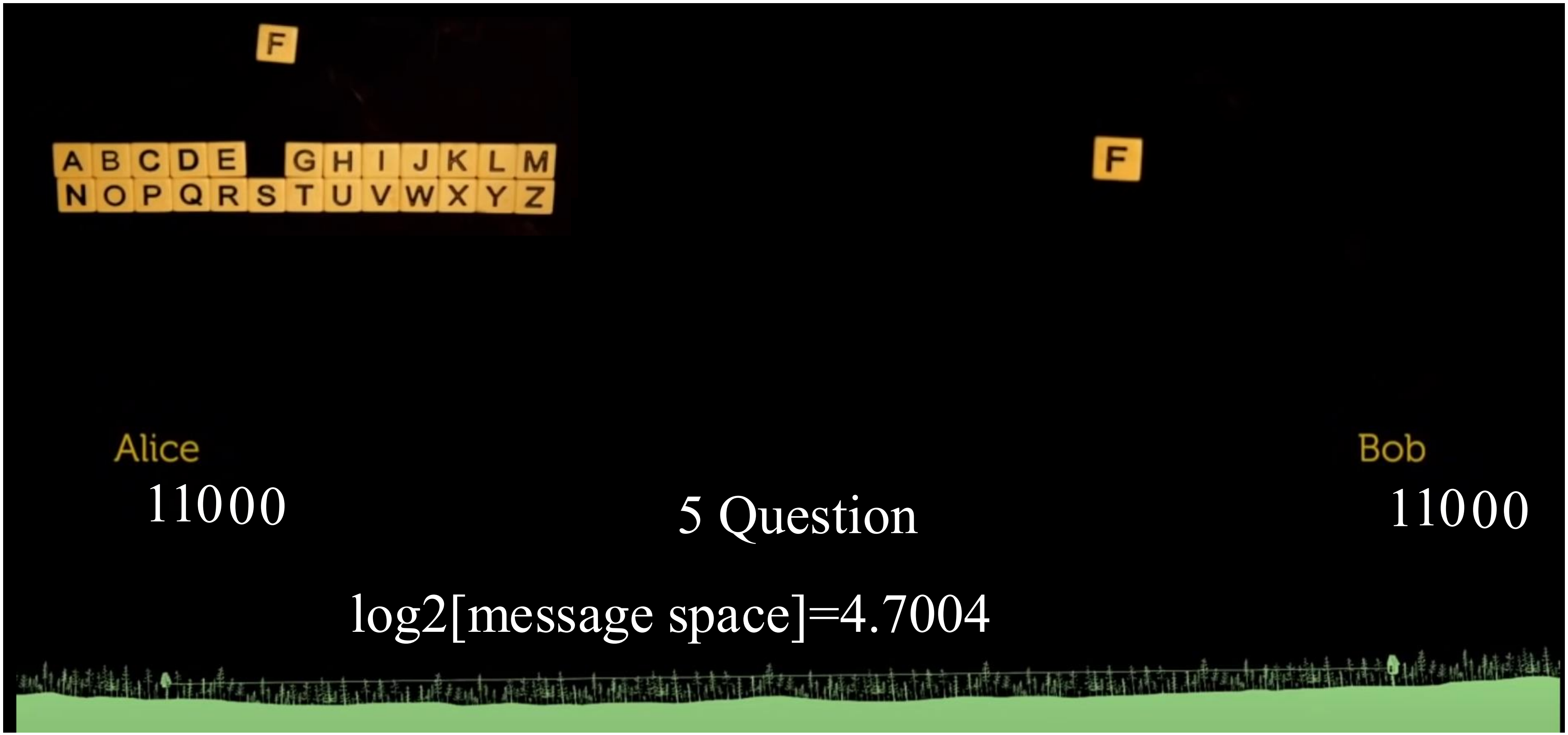
The more uncertain an event is, the more information it contains.

Q1: How to measure the quantity of information?



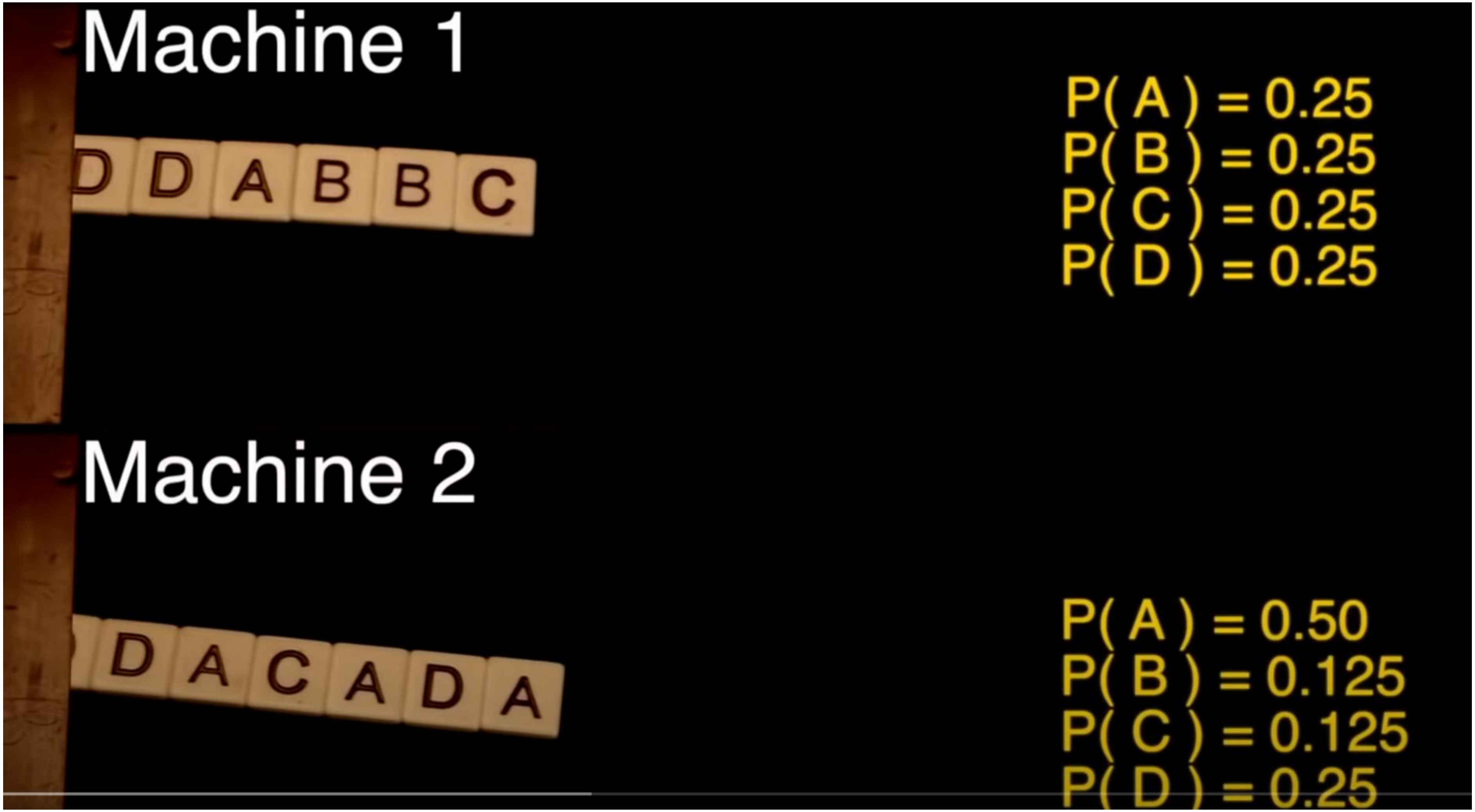
For 10 flips, what's the minimum number of questions? → 10 questions (10 binary digit to send the message)

Q1: How to measure the quantity of information?



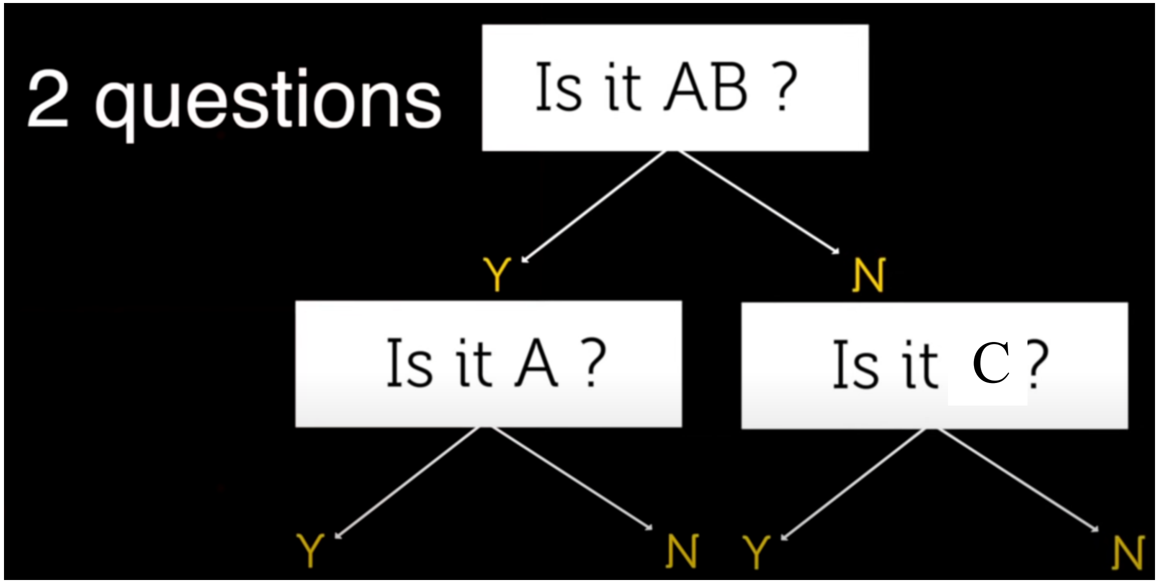
For 6 letters, what's the minimum number of questions? →
6*4.7=28.2 questions (28.2 binary digit to send the message)

Q1: How to measure the quantity of information?



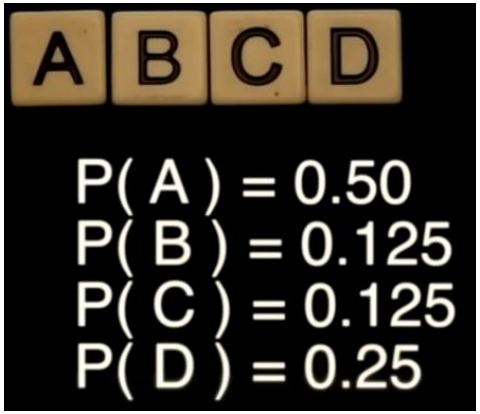
Q1: How to measure the quantity of information?

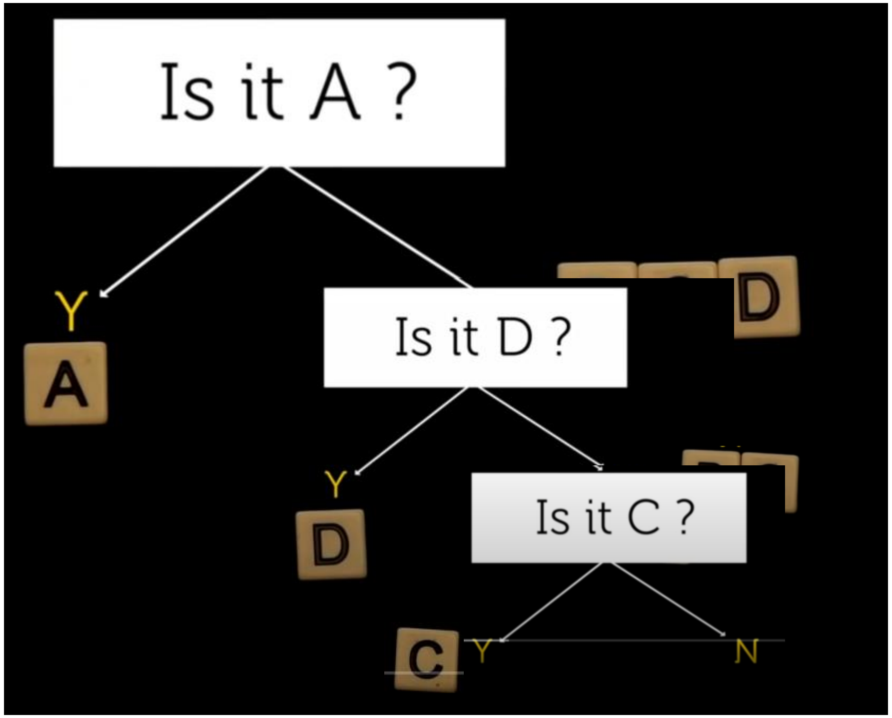
Machine 1:



Q1: How to measure the quantity of information?

Machine 2:



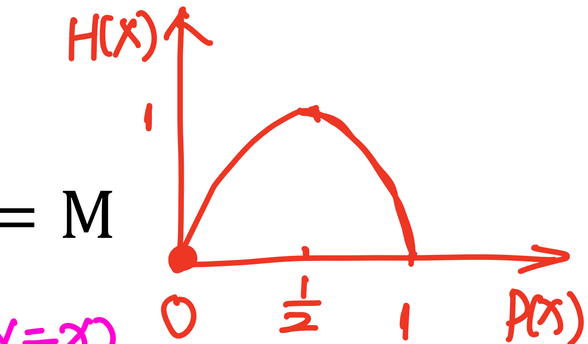


On average, how many questions to determine the symbol of Machine 2? ->0.5*1+0.25*2+0.125*3+0.125*3=1.75

Entropy

Pmf: P(x)

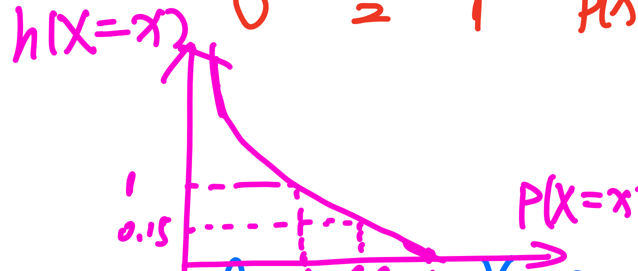
- Assume a discrete r.v. $X \in \mathcal{X}$, and $|\mathcal{X}| = \mathrm{M}$



$$H(X) = \sum_{x\in\mathcal{X}} p(x) \boxed{\log_2\left(\frac{1}{p(x)}\right)}$$

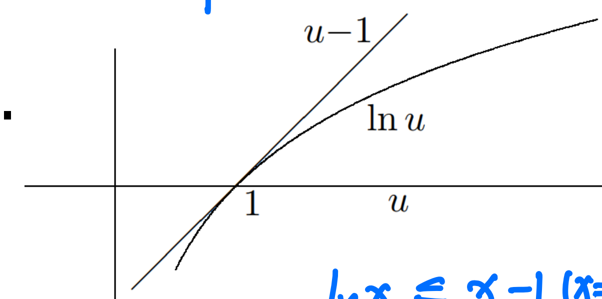
$E[-\log_2(P(x))]$

$h(X=x)$



- $H(X) = E_{p(x)}\left[\log_2\left(\frac{1}{p(x)}\right)\right]$

Shannon information content of outcome $X=x$



- $H(X) \geq 0$. Equality holds if X is deterministic.
- $\log_2$: bits; $\log_e$: nats.
- $H(X) \leq \log_2 M$. Equality holds if X is equiprobable.

$\ln x \leq x-1$ $(x=1)$

$\frac{\log_2 x}{\log_2 e} \leq x-1$

$(x=1)$ $\log_2 x \leq \log_2 e\,(x-1)$

  - Proof: $H(X) - \log M = \sum_{x\in\mathcal{X}} p(x) \log_2\left(\frac{1}{p(x)M}\right)$

$$\leq \log e \sum_{x\in\mathcal{X}} p(x)\left(\frac{1}{p(x)M} - 1\right) = 0$$

$\log x \leq \log e\,(x-1)$

$\sum P(x)\log_2\frac{1}{P(x)} - \sum P(x)\log_2 M$

$= \sum P(x) \boxed{\log_2 \frac{1}{P(x)M}}$

$\leq \log_2 e \sum P(x)\left(\frac{1}{P(x)M} - 1\right) = 0$

$\frac{1}{P(x)M} = 1 \Rightarrow P(x) = \frac{1}{M}, \forall x \in \mathcal{X}$

Entropy

- Example

$$X = \begin{cases} a & \text{with probability } \frac{1}{2}, \\ b & \text{with probability } \frac{1}{4}, \\ c & \text{with probability } \frac{1}{8}, \\ d & \text{with probability } \frac{1}{8}. \end{cases}$$

The entropy of X is

$$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{8}\log\frac{1}{8} = \frac{7}{4} \text{ bits.}$$

$-E[\log_2 P(x)]$

Joint Entropy and Conditional Entropy

- **Joint Entropy**: Assume $(X,Y) \sim p(x,y)$, the joint entropy H(X,Y) is defined as

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

$H(X,Y) = -E[\log p(x) + \log p(y)]$

- If X and Y are independent, we have H(X,Y)=H(X)+H(Y).
- Question: How to measure the quantity of information on X, when we already knew Y?

$H(Y|X) \le H(Y)$

$H(X|Y) \le H(X)$

- **Conditional Entropy:**

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

Joint Entropy and Conditional Entropy

- Chain rule:

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

$$\boxed{H(X,Y) = H(Y) + H(X|Y)}$$

- Proof:

- $H(X,Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log p(x,y)$
- $= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\left(p(x|y)p(y)\right)$
- $= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\left(p(x|y)\right) - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log p(y)$
- $= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\left(p(x|y)\right) - \sum_{y\in\mathcal{Y}} p(y)\log p(y)$
- $= H(X|Y) + H(Y)$

$$I(X;Y) = E\left[\log_2 \frac{P(x,y)}{P(x)P(y)}\right]$$

$$= E\left[-\log_2 \frac{P(x)P(y)}{P(x,y)}\right]$$

$$\geq -\log_2 \boxed{E\left[\frac{P(x)P(y)}{P(x,y)}\right]}$$

$$\sum_{x,y} P(x,y) \frac{P(x)P(y)}{P(x,y)}$$

$$= 0$$

Jensen's Inequality

if $f(x)$ is convex

$$f(E(x)) \leq E[f(x)]$$

$|a-b|$

$\|\vec{x}-\vec{y}\|_2$

$\|\vec{X}-\vec{Y}\|_F$

$f_1(x) \quad f_2(x)$

KL Divergence

$$I(X;Y) \Leftrightarrow D_{KL}(P(x,y) \,\|\, P(x)P(y))$$

$$\begin{cases} \geq 0 \\ = 0 \quad P(x,y) = P(x)P(y),\ X, Y \text{ Independent} \end{cases}$$

Mutual Information

- How to measure the dependence between X and Y?
- **Mutual Information:** Assume $(X, Y) \sim p(x, y)$, and $X \sim p(x), Y \sim p(y)$. The mutual information I(X;Y) is defined as

$$\mathrm{I}(X;Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \begin{cases} \geq 0 \\ 0, \text{ independent} \end{cases}$$

- $I(X;Y) = I(Y;X)$
- $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
- $I(X;Y) = H(X) + H(Y) - H(X,Y)$
- $I(X;X) = H(X)$

$I(X;Y) = 0$, Independent

$$I(X;Y) = E\left[\log_2 \frac{p(x,y)}{p(x)p(y)}\right]$$

$$= E\left[\log_2 \frac{p(y|x)}{p(y)}\right]$$

$$= E \log_2(p(y|x)) - E \log_2(p(y))$$

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

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

Mutual Information

- **Mutual Information and entropy**

Joint Entropy $H(X,Y)$

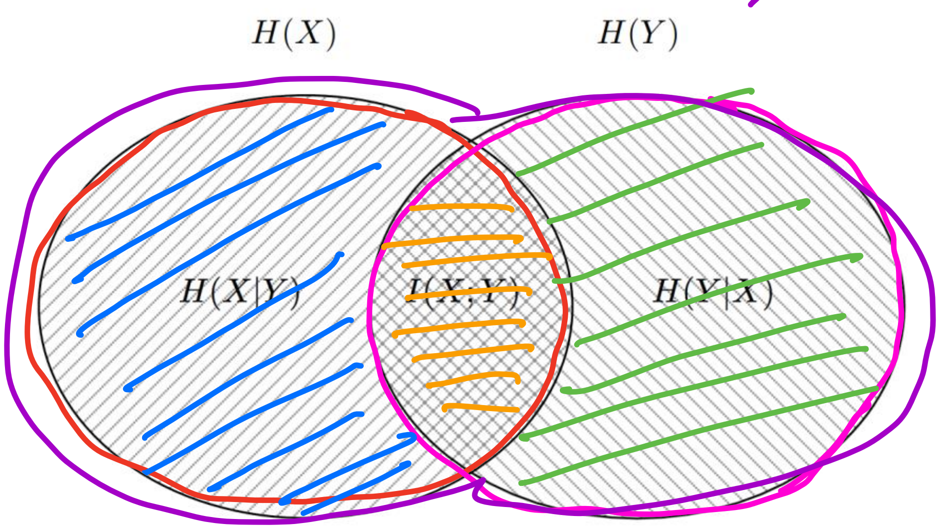


Figure: Entropy and mutual information

- If X and Y are independent $\Longrightarrow$ $I(X;Y) = 0$

Differential Entropy and Mutual Information

- Assume a continuous r.v. X with pdf f(x). The **differential entropy** h(X) is defined as

$$h(X) = -\int f(x) \log f(x) \, dx$$

- $h(X) = E\left[-\log\left(f(X)\right)\right]$
- $h(X)$ could be negative or infinite;
- **Mutual Information** I(X;Y) with f(x,y) is defined as

$$\mathrm{I}(X;Y) = \int f(x,y) \log \frac{f(x,y)}{f(x)f(y)} \, dxdy$$

- I(X;Y)=h(X)-h(X|Y)

Differential Entropy and Mutual Information

- Example
- Uniform Distribution

Given a RV X, with $a \le X \le b$. Its PDF follows

$$\boxed{f_X(x) = \frac{1}{b-a}}$$

And,

$$\boxed{E(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(a-b)^2}{12}}$$

Check: $h(X)$

Differential Entropy and Mutual Information

- Example
- Uniform Distribution
- Check:

- $h(X) = \int_a^b -\frac{1}{b-a} \log \frac{1}{b-a} dx = \log(b-a)$
- When b-a<1, we have h(X)<0.

Differential Entropy and Mutual Information

- Example
- Gaussian Distribution

Given a RV $X \sim \mathcal{N}(u, \sigma^2)$, its PDF follows

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

And,

$$\boxed{E(X) = \mu, \quad \text{Var}(X) = \sigma^2}$$

- Check: h(X)

Differential Entropy and Mutual Information

- Example
- Gaussian Distribution
- Check: h(X)

$\mu = 0$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x)^2}{2\sigma^2}}$$

Normal distribution:

$$\begin{aligned} h(x) &= -\int f(x) \ln f(x) dx \\ &= -\int f(x) \left[-\frac{x^2}{2\sigma^2} - \ln \sqrt{2\pi\sigma^2} \right] dx \\ &= \frac{EX^2}{2\sigma^2} + \frac{1}{2}\ln 2\pi\sigma^2 \\ &= \frac{1}{2}\ln 2\pi e\sigma^2 \quad \text{nats} \\ &= \frac{1}{2}\log 2\pi e\sigma^2 \quad \text{bits} \end{aligned}$$

nats

$\frac{1}{2} + \frac{1}{2}\ln 2\pi\sigma^2$

$h(x) = -\int f(x) \log_2 f(x) dx$

$= \log_2 e \left(-\int f(x) \ln f(x) dx \right)$

$= \log_2 e \left(\frac{1}{2} \ln 2\pi e \sigma^2 \right)$

$= \frac{1}{2} \log_2 (2\pi e \sigma^2)$

bits

Compare: $H_b(X) = \log_b a H_a(X)$

Differential Entropy and Mutual Information

- $\max_{E(\mathbf{X}\mathbf{X}^{\mathrm{T}})=\mathbf{K}} h(X) = \frac{1}{2}\log(2\pi e)^n|\mathbf{K}|$, with equality iff $\mathbf{X} \sim N(0, \mathbf{K})$.
- P254 of T. Cover
- Gaussian Distribution maximizes the entropy over all distributions with the same variance.

Thanks for your kind attention!

Questions?