Coding and Information Theory
Chapter 4
Information Channels

Xuejun Liang

This is the first lecture of chapter 4

Quick Review of Last Lecture

- Shannon-Fane Coding examples

$$l_i = \lceil \log_2(1/p_i) \rceil = \min\{n \in \mathbf{Z} \mid 2^n \geq 1/p_i\}$$

- Entropy of Extensions and Products

$$H_r(S^n) = nH_r(S).$$

- Shannon's First Theorem

$$\lim_{n \to \infty} \frac{L_n}{n} = H_r(S)\,.$$

- An Example of Shannon's First Theorem

S has two symbols s_1, s_2 of probabilities $p_i = 2/3,\ 1/3$

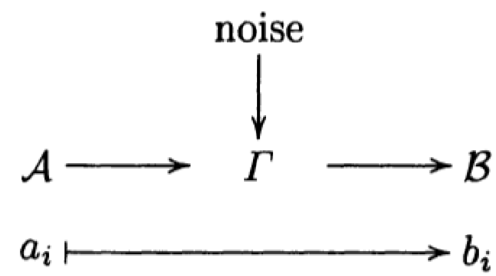
Chapter 4: Information Channels

1. Notation and Definitions
2. The Binary Symmetric Channel
3. System Entropies
4. System Entropies for the Binary Symmetric Channel
5. Extension of Shannon's First Theorem to Information Channels
6. Mutual Information
7. Mutual Information for the Binary Symmetric Channel
8. Channel Capacity

The aim of this chapter

- We Consider
  - a source sending messages through an unreliable (or noisy) channel to a receiver
- Our aim here is
  - to measure how much information is transmitted, and how much is lost in this process, using several different variations of the entropy function, and then
  - to relate this to the average word-length of the code used.

4.1 Notation and Definitions



- Information channel Γ
- Input of Γ: Source A,
  - with finite alphabet A of symbols $a = a_1, \ldots, a_r$, having probabilities

$$p_i = \Pr(a = a_i) \quad \text{where}$$

$$0 \le p_i \le 1 \qquad \text{and} \qquad \sum_{i=1}^{r} p_i = 1$$

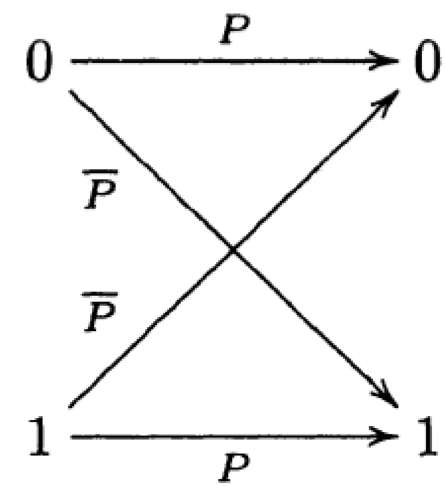
- Output of Γ: Source B,
  - with a finite alphabet B of symbols $b = b_1, \ldots, b_s$, having probabilities

$$q_j = \Pr(b = b_j) \quad \text{where}$$

$$0 \le q_j \le 1 \qquad \text{and} \qquad \sum_{j=1}^{s} q_j = 1$$

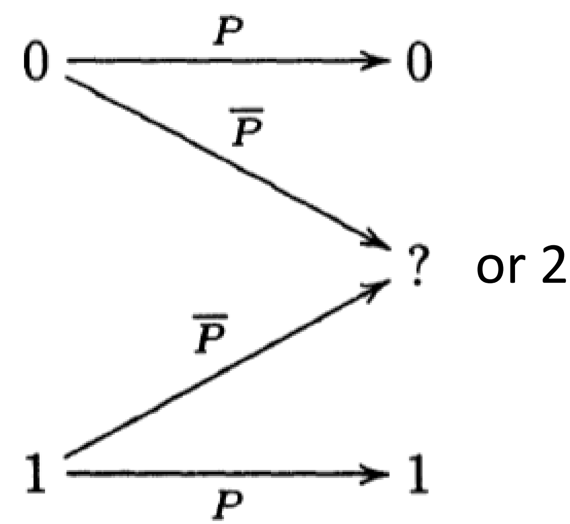
Example 4.1

- Binary symmetric channel (BSC)
  - $\mathrm{A} = \mathrm{B} = Z_2 = \{0, 1\}$.
  - Each input symbol $a = 0$ or 1 is correctly transmitted with probability P, and is incorrectly transmitted (as $\bar{a} = 1 - a$) with probability $\bar{P} = 1 - P$, for some constant P $(0 \leq P \leq 1)$.



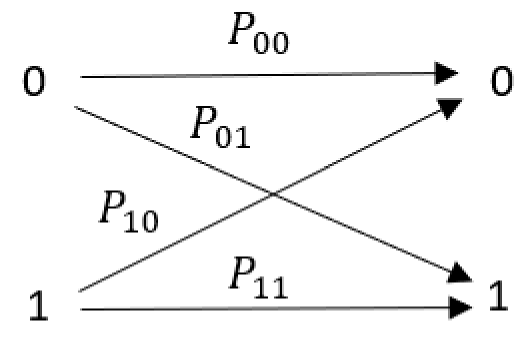
Example 4.2

- Binary erasure channel (BEC)
  - $\mathrm{A} = Z_2 = \{0, 1\}$.
  - $\mathrm{B} = \{0, 1, ?\}$ (or $\{0, 1, 2\}$).
  - Each input symbol $a = 0$ or 1 is correctly transmitted with probability P, and is erased (or made illegible) with probability $\bar{P}$, indicated by an output symbol $b = ?$ (or 2)



Example: Binary Channel (BC)

- Binary channel (BC)
  - $\mathrm{A} = \mathrm{B} = Z_2 = \{0, 1\}$.
  - Input symbol $a = 0$ is correctly transmitted with probability P_{00} and is incorrectly transmitted with probability $P_{01} = 1 - P_{00}$
  - Input symbol $a = 1$ is correctly transmitted with probability P_{11} and is incorrectly transmitted with probability $P_{10} = 1 - P_{11}$



Forward Probabilities

- Forward probabilities of Γ

$$P_{ij} = \Pr(b = b_j \mid a = a_i) = \Pr(b_j \mid a_i)$$

- We have $\sum_{j=1}^{s} P_{ij} = 1$

- The channel matrix $M = (P_{ij}) = \begin{pmatrix} P_{11} & \dots & P_{1s} \\ \vdots & & \vdots \\ P_{r1} & \dots & P_{rs} \end{pmatrix}$

Channel Matrix Example – BC

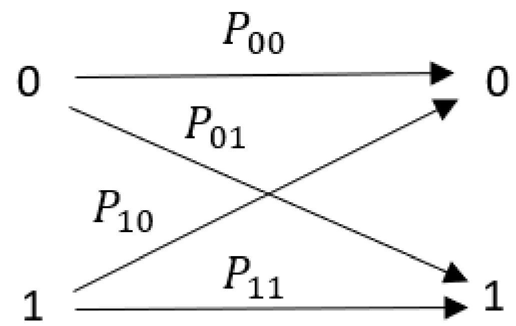
Channel:
r input symbols
s output symbols

$$M = (P_{ij}) = \begin{pmatrix} P_{11} & \dots & P_{1s} \\ \vdots & & \vdots \\ P_{r1} & \dots & P_{rs} \end{pmatrix}$$

$$\sum_{j=1}^{s} P_{ij} = 1$$

$$P_{ij} = \Pr(b = b_j \mid a = a_i) = \Pr(b_j \mid a_i)$$

Binary channel



$$M = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix}$$

$$P_{00} = \Pr(b = 0 \mid a = 0)$$

$$P_{01} = \Pr(b = 1 \mid a = 0)$$

$$P_{10} = \Pr(b = 0 \mid a = 1)$$

$$P_{11} = \Pr(b = 1 \mid a = 1)$$

Channel Matrix Example – BSC

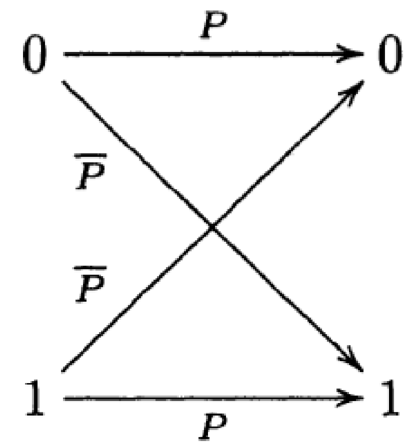
Channel:
r input symbols
s output symbols

$$M = (P_{ij}) = \begin{pmatrix} P_{11} & \dots & P_{1s} \\ \vdots & & \vdots \\ P_{r1} & \dots & P_{rs} \end{pmatrix}$$

$$\sum_{j=1}^{s} P_{ij} = 1$$

$$P_{ij} = \Pr(b = b_j \mid a = a_i) = \Pr(b_j \mid a_i)$$

BSC



$$M = \begin{pmatrix} P & \overline{P} \\ \overline{P} & P \end{pmatrix}$$

Channel Matrix Example – BEC

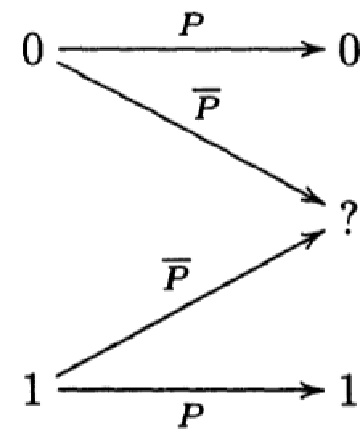
Channel:
r input symbols
s output symbols

$$M = (P_{ij}) = \begin{pmatrix} P_{11} & \dots & P_{1s} \\ \vdots & & \vdots \\ P_{r1} & \dots & P_{rs} \end{pmatrix}$$

$$\sum_{j=1}^{s} P_{ij} = 1$$

$$P_{ij} = \Pr(b = b_j \mid a = a_i) = \Pr(b_j \mid a_i)$$

BEC



or 2

$$M = \begin{pmatrix} P & 0 & \overline{P} \\ 0 & P & \overline{P} \end{pmatrix}$$

Combining two channels

- **Sum** $\Gamma + \Gamma'$
  - If Γ and Γ' have disjoint input alphabets A and A', and disjoint output alphabets B and B', then the **sum** $\Gamma + \Gamma'$ has input and output alphabets $A \cup A'$ and $B \cup B'$.
  - Each input symbol is transmitted through Γ or Γ', so the channel matrix is a block matrix

$$\begin{pmatrix} M & O \\ O & M' \end{pmatrix}$$

where M and M' are the channel matrices for Γ and Γ'.

Combining two channels

- **Product** $\Gamma \times \Gamma'$
  - The input and output alphabets are A x A' and B x B’
  - The sender transmits a pair $(a, a') \in$ A x A' by simultaneously sending a through Γ and a' through Γ'
  - A pair $(b, b') \in$ B X B' is received
  - Thus the forward probabilities are

$$\Pr((b,b') \mid (a,a')) = \Pr(b \mid a) \,.\, \Pr(b' \mid a')$$

  - So the channel matrix is the **Kronecker product** $M \otimes M'$ of the matrices M and M' for Γ and Γ'.
    - if $M = (P_{ij})$ and $M' = (P'_{kl})$ are $r \times s$ and $r' \times s'$ matrices, then $M \otimes M'$ is an $rr' \times ss'$ matrix, with entries $P_{ij}P'_{kl}$

Example

- If Γ and Γ' are binary symmetric channels, with channel matrices

$$M = \begin{pmatrix} P & \overline{P} \\ \overline{P} & P \end{pmatrix} \quad \text{and} \quad M' = \begin{pmatrix} P' & \overline{P'} \\ \overline{P'} & P' \end{pmatrix}$$

- then $\Gamma + \Gamma'$ and $\Gamma \times \Gamma'$ have channel matrices

$$\begin{array}{c} \\ 0 \\ 1 \\ 0' \\ 1' \end{array} \begin{array}{c} \begin{array}{cccc} 0, & 1, & 0', & 1' \end{array} \\ \begin{pmatrix} P & \overline{P} & 0 & 0 \\ \overline{P} & P & 0 & 0 \\ 0 & 0 & P' & \overline{P'} \\ 0 & 0 & \overline{P'} & P' \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{c} \begin{array}{cccc} (0,0'), & (1,0'), & (0,1'), & (1,1') \end{array} \\ \begin{pmatrix} PP' & \overline{P}P' & P\overline{P'} & \overline{P}\,\overline{P'} \\ \overline{P}P' & PP' & \overline{P}\,\overline{P'} & P\overline{P'} \\ P\overline{P'} & \overline{P}\,\overline{P'} & PP' & \overline{P}P' \\ \overline{P}\,\overline{P'} & P\overline{P'} & \overline{P}P' & PP' \end{pmatrix} \end{array} \begin{array}{c} \\ (0,0') \\ (1,0') \\ (0,1') \\ (1,1') \end{array}$$

The channel relationships

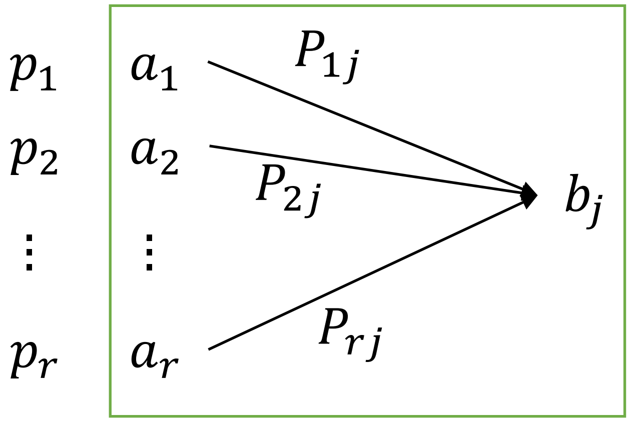
- The channel relationships

$$\sum_{i=1}^{r} p_i P_{ij} = q_j \qquad (4.2)$$

Where $p_i = \Pr(a = a_i) = \Pr(a_i)$

$q_j = \Pr(b = b_j) = \Pr(b_j)$ and

$P_{ij} = \Pr\left(b = b_j \middle| a = a_i\right) = \Pr(b_j | a_i)$



$$q_j = p_1 P_{1j} + p_2 P_{2j} + \cdots + p_r P_{rj}$$

The channel relationships: Cont.

- The channel relationships

$$\sum_{i=1}^{r} p_i P_{ij} = q_j \qquad (4.2)$$

(4.2) can be written as

$$\mathbf{p}M = \mathbf{q}\,. \qquad (4.2')$$

Where, $p = (p_1, p_2, \dots, p_r)$, $q = (q_1, q_2, \dots, q_s)$, and $M = \left(P_{ij}\right)_{r\times s}$

- The backward probabilities

$$Q_{ij} = \Pr\left(a = a_i \mid b = b_j\right) = \Pr\left(a_i \mid b_j\right)$$

- The joint probabilities

$$R_{ij} = \Pr\left(a = a_i \text{ and } b = b_j\right) = \Pr\left(a_i, b_j\right)$$

Bayes' Formula

- **Bayes' Formula**

$$Q_{ij} = \frac{p_i}{q_j} P_{ij} \qquad (4.3)$$

provided $q_j \neq 0$.

$$p_i P_{ij} = \Pr(a_i) \Pr(b_j \mid a_i) = \Pr(a_i, b_j) = R_{ij}$$

$$q_j Q_{ij} = \Pr(b_j) \Pr(a_i \mid b_j) = \Pr(a_i, b_j) = R_{ij}$$

- Combining this with (4.2) we get

$$Q_{ij} = \frac{p_i P_{ij}}{\sum_{k=1}^{r} p_k P_{kj}} \qquad (4.4)$$

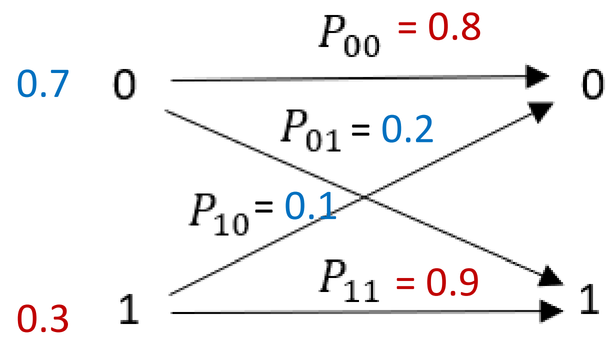
$$\sum_{i=1}^{r} p_i P_{ij} = q_j \qquad (4.2)$$

Example: In a binary communication system below. Given

$p_0 = P(a = 0) = 0.7,$

$P_{01} = P(b = 1 \mid a = 0) = 0.2$ and

$P_{10} = P(b = 0 \mid a = 1) = 0.1,$



(a) Find the channel matrix M

(b) Find $q_0 = P(b = 0)$ and $q_1 = P(b = 1)$.

(c) Find $R_{ij} = P(a = i \mid b = j)$, for $i, j = 0, 1$

(d) Find $Q_{ij} = P(a = i \mid b = j)$, for $i, j = 0, 1$

To solve (a) just using definition

(a) $$M = \left(P_{ij}\right) = \left(P(b = j \mid a = i)\right) = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix}$$

To solve (b) using (4.2) $\sum_{i=1}^{r} p_i P_{ij} = q_j$ or (4.2') $\mathbf{p}M = \mathbf{q}$.

(b) $$(q_0 \quad q_1) = (p_0 \quad p_1)\left(P_{ij}\right) = (0.7 \quad 0.3)\begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix} = (0.59 \quad 0.41)$$

Example: In a binary communication system below. Given

$p_0 = P(a = 0) = 0.7,$

$P_{01} = P(b = 1 \mid a = 0) = 0.2$ and

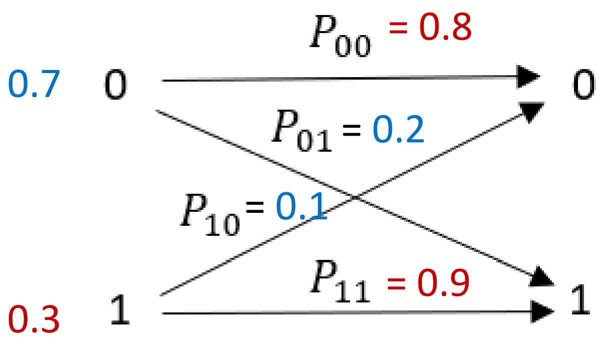
$P_{10} = P(b = 0 \mid a = 1) = 0.1,$

(a) Find the channel matrix M

(b) Find $q_0 = P(b = 0)$ and $q_1 = P(b = 1)$.

(c) Find $R_{ij} = P(a = i \mid b = j)$, for $i, j = 0, 1$

(d) Find $Q_{ij} = P(a = i \mid b = j)$, for $i, j = 0, 1$



To solve (c) and (d) using $R_{ij} = p_i P_{ij} = q_j Q_{ij}$

(c) $$(R_{ij}) = \begin{pmatrix} p_0 & 0 \\ 0 & p_1 \end{pmatrix} (P_{ij}) = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.3 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{pmatrix} = \begin{pmatrix} 0.56 & 0.14 \\ 0.03 & 0.27 \end{pmatrix}$$

(d) $$(Q_{ij}) = (R_{ij}) \begin{pmatrix} 1/q_0 & 0 \\ 0 & 1/q_1 \end{pmatrix} = \begin{pmatrix} 0.56 & 0.14 \\ 0.03 & 0.27 \end{pmatrix} \begin{pmatrix} 1/0.59 & 0 \\ 0 & 1/0.41 \end{pmatrix}$$
$$= \begin{pmatrix} 0.9492 & 0.3415 \\ 0.0508 & 0.6585 \end{pmatrix}$$