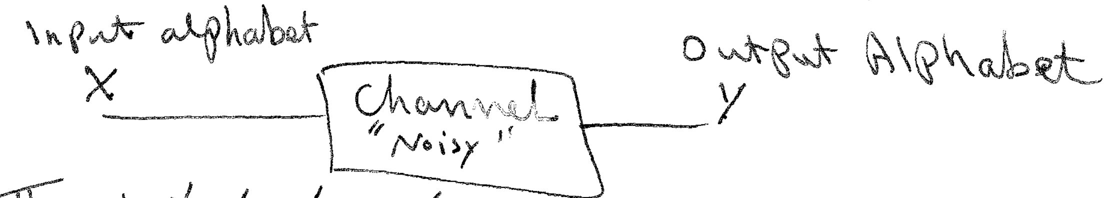
Chapter 2

Channels and channel Capacity.



The simplest channel Model used in communication is the AWGN channel.

$$y = x + \eta \text{ ; where } \eta \text{ is a zero-mean Gaussian Random Variable.}$$

Channels in general have memory. That means that the output sequence can will be correlated. A special case that is used a lot in analyzing communication systems is the Memoryless channels.

2.1 Discrete Memoryless Channel Model

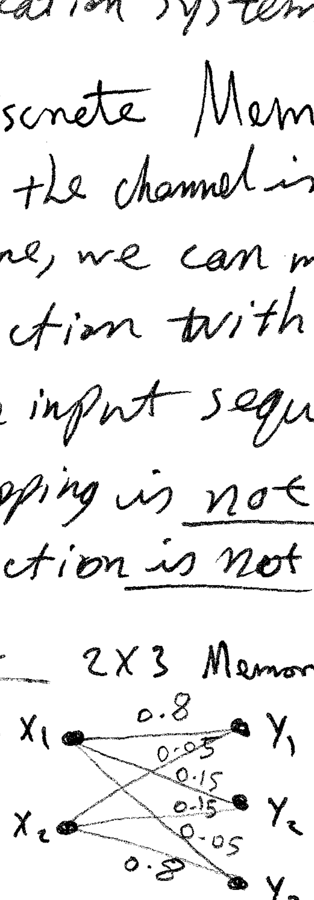
Since the channel is noisy, errors in transmission will occure, we can model a Memoryless channel as a function with transition probabilities that maps the input sequence X to output sequence Y. This mapping is <u>not</u> one-to-one and reversing the function <u>is not</u> possible.

<u>Example</u> 2X3 Memoryless channel Model

Let P_x be the Prob. that symbol x is transmitted.

$$\Rightarrow P_y = \sum_{x \in X} P_{y/x} P_x$$

Prob. that the received symbol is y.



the forward transition Prob. can be represented by a matrix

$$P_{Y/X} = \begin{bmatrix} 0.8 & 0.05 \\ 0.15 & 0.15 \\ 0.05 & 0.8 \end{bmatrix} = \begin{bmatrix} P_{Y_1/X_1} & P_{Y_1/X_2} \\ P_{Y_2/X_1} & P_{Y_2/X_2} \\ P_{Y_3/X_1} & P_{Y_3/X_2} \end{bmatrix}$$

In Matrix Form:

the bar denotes a Matrix or a vector.

$$\bar{P}_Y = \bar{P}_{Y/X}\,\bar{P}_X$$

$$\begin{bmatrix} P_{y_1} \\ P_{y_2} \\ P_{y_3} \end{bmatrix} = \begin{bmatrix} P_{y_1/x_1} & P_{y_1/x_2} \\ P_{y_2/x_1} & P_{y_2/x_2} \\ P_{y_3/x_1} & P_{y_3/x_2} \end{bmatrix} \begin{bmatrix} P_{x_1} \\ P_{x_2} \end{bmatrix}$$

Notice that

$$\sum_{y \in Y} P_{Y/x_1} = \sum_{y \in Y} P_{Y/x_2} = 1$$

$\Rightarrow$ columns of $P_{Y/X}$ sum to unity.

For the previous example

$$\bar{P}_Y = \begin{bmatrix} 0.425 \\ 0.15 \\ 0.425 \end{bmatrix} \text{ if } P_X = [0.5 \;\; 0.5] \rightarrow \text{equally probable.}$$

— Notice that the calculations of the transition probabilities are part of the communication systems performance evaluation covered in Digital Communication Courses.

— Notice that ~~the~~ $|Y| \neq |X|$ in the above example.

if $|Y| = |X| \rightarrow$ Hard-decision Decoding

if $|Y| > |X| \rightarrow$ Soft-decision Decoding.

2.1.2 Output Entropy and Mutual information

What is the Mutual information between the input ~~to~~ the channel and the output?

First, examine the previous example,

Entropy $\rightarrow$ $H(Y) = \sum_{Y \in Y} P_Y \log_2\left(\frac{1}{P_Y}\right) = 1.4598$ bits

while $H(X) = 1$ bits since P_X are equally prob.

Notice that $H(Y) > H(X) \Rightarrow$ The channel adds information to the source. More randomness

this is not always the case, some times we get $H(Y) < H(X)$ $\Rightarrow$ information loss.

Assume we have a 2X2 ~~Discrete~~ DMC {Discrete Memoryless Channel}

$$P_{Y|X} = \begin{bmatrix} 0.98 & 0.05 \\ 0.02 & 0.95 \end{bmatrix} \text{ and } P_X = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

$$\Rightarrow \quad P_Y = \begin{bmatrix} 0.515 \\ 0.485 \end{bmatrix} \Rightarrow H(Y) = 0.99935 < H(X)$$

$\Rightarrow$ we have information Loss.

The information in the above example was lost during the transmission process. ~~That~~

Information Lossy channels $\Rightarrow H(Y) < H(X)$

Mutual Information

The receiver observes Y, How much can it tell us about the transmitted information sequence?

This is the Mutual information of X observing Y.

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P_{xy} \log_2 \left(\frac{P_{xy}}{P_y P_x} \right)$$

$$I(X;Y) = H(X) - H(X/Y)$$

— If Y and X are independent $\Rightarrow I(X;Y) = 0$

$\Rightarrow$ ~~As~~ Y tells us nothing at all about X.

— Upper bound on Mutual information is

$$I(X;Y) \leq H(X)$$

with equality if and only if $H(X/Y) = 0$.

that means that there is no information Loss, and Y contains sufficient information to tell us what the transmitted sequence was.

Example 2.1.4

For the 2X3 DMC in Example 2-1-1

$I(X;Y) = 0.57566$ bits

$\Longrightarrow$ since I is much less than $H(X)=1$, the channel has a high level of information loss.

Example 2-1-5

In this case, 2X2 DMC $I(X;Y) = 0.78543$

Also, this 2X2 DMC causes information loss.

2.2 Channel Capacity and the Binary Symmetric Channel.

2.2.1 Channel Capacity

Channel capacity is the maximum average amount of information that can be sent per channel use. Each time the transmitter sends a symbol, it is said to use the channel.

Channel capacity is

$$C_x = \max_{P_x} I(X;Y)$$

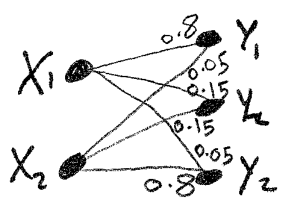
$\Rightarrow$ is the maximum Mutual information achieved for a given channel.

Maximization is done over the input probabilities.

In otherwords, the channel capacity is the maximum information rate that can be supported by the channel.

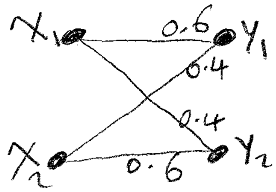
2.2.2 Symmetric Channels.

Example



$$P_{Y|C} = \begin{bmatrix} 0.8 & 0.05 \\ 0.15 & 0.15 \\ 0.05 & 0.8 \end{bmatrix}$$

Example



$\Rightarrow$ Binary symmetric channel

$$P_{Y|C} = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix}$$

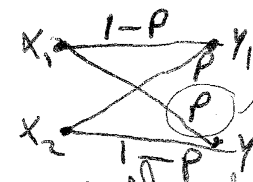
* The capacity of Binary symmetric channel is
(BSC)

$$C_x = 1 - H(p)$$
$$= 1 + (1-p)\log_2(1-p) + p\log_2(p)$$

where p is the crossover probability (Prob. of error)

The BSC ~~Prob.~~ transition Prob. is

$$P_{Y|X} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$



— The above capacity is the maximum Mutual information
And that happens when the the input distribution (P_x) is uniform [equally likely].

— $0 \leq C_x \leq 1$ for BSC

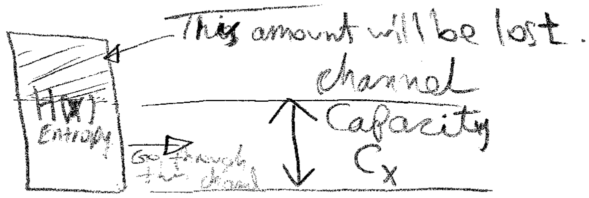
$1 \to H(x)$

— the upper bound is achieved only if $p = 0$ or $p = 1$. At these values, $C_x = 1$

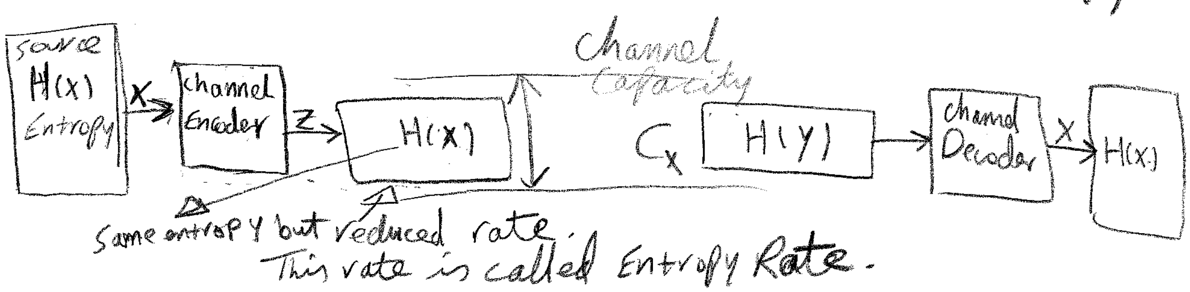
— the Lower bound, $C_x = 0$, is achieved when $p = 0.5 \Rightarrow$ 50% chance to get the output correct. Thus, Error and correct transmition are equally likely. And the information loss is total.

- If we send data with information $H(x) > C_x$ then for sure we will loose some information over the channel. However, we can reduce the entropy of the data by adding redundant symbols.

Thus, this the key idea of coding. We add redundant bits using a coding algorithm so that we reduce the information of the source and make it able to pass the channel with very low prob. of lost information.



But, Add more redundant bits ⟹ Lower Entropy



2.3 Block Coding and Shannon's Second Theorem:

2.3 Block Coding and Shannon's Second Theorem.

The channel capacity is

$$C_x = \max_{P_x} I(X;Y)$$
$$= \max_{P_x} \left(H(X) - H(X/Y) \right)$$

The maximum capacity is $C_x = H(X)$,

thus, the term $H(X/Y)$ is the information loss from the Maximum. This conditional entropy is called Equivocation. Recall that

$$H(X/Y) \leq H(X).$$

2.3.2 Entropy Rate and Channel Coding Theorem.

- The entropy of a block of n symbols is

$$H(X_0, X_1, X_2, \ldots, X_{n-1}) \leq n H(X)$$

if they are drawn from the same source and the prob. doesn't change with time.

- So, this is the information content of the block.
- The average information per channel use is the Rate

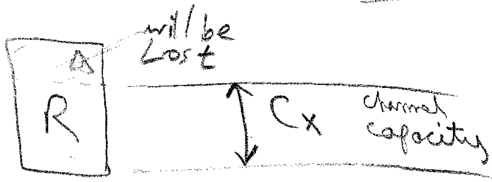
$$R = \frac{H(X_0, X_1, \ldots, X_{n-1})}{n} \leq H(X)$$

- Taking the limit $n \to \infty$, we get Entropy Rate

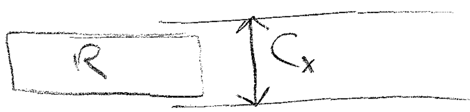
Entropy Rate $\longrightarrow$ $$R = \lim_{n \to \infty} \frac{H(X_0, X_1, \ldots, X_{n-1})}{n} \leq H(X)$$

What is the relation between R and C_x?
Can We control R inorder to have almost zero information loss?

Theoritically, the answer is Yes. We can reduce R by adding redundant symbols such that $R \leq C_x$.
This Process is called channel coding.



However, if we let $R \leq C_x$, we can pass the information with Low Prob. of error.



Shannon's Theorem

suppose $R < C_x$, where C_x is the capacity of a memoryless channel. Then for any $\epsilon > 0$, there exists a block length n and a code of block length n and rate R whose probability of block decoding error P_e satisfies $P_e \leq \epsilon$ when the code is used on this channel.

Cutoff rate, R_0

Another bound for practical error-correcting codes is called the cutoff rate (R_0). For a binary symmetric channel, the cutoff rate (R_0) is

$$R_0 = -\log_2\left(0.5 + \sqrt{P(1-P)}\right)$$

This bound applies for most of the codes, However, new codes, such as, Turbo codes or LDPC codes can operate very close to channel capacity.

Example 2.3.2

Find C_X and R_0 for the BSC with

a) $p=0.1$ b) $p=0.01$ c) $p=0.001$ d) $p=0.0001$

a) $C_X=0.531$ $R_0=0.322$ b) $C_X=0.919$ $R_0=0.738$

c) $C_X=0.988$ $R_0=0.911$ d) $C_X=0.998$ $R_0=0.971$

Notice that the cutoff rate approaches the capacity as the crossover Prob. p grows smaller.

— See Attached paper "Shannon's Noisy Channel Coding Theorem".

— The capacity of continuous time, bandlimited, AWGN channels is

$$C_x = W \log_2(1 + SNR) \text{ bps}.$$

where W is the channel Bandwidth in Hertz.

$$SNR = \frac{P_s}{P_N} = \frac{\text{Signal Power}}{\text{Noise Power}}$$

— See the examples in the paper.