Program : **B.Tech**

Subject Name: **Digital Communication**

Subject Code: **EC-502**

Semester: **5th**

Unit 5

Syllabus:

Information Theory Source Coding: Introduction to information theory, uncertainty and information, average mutual information and entropy, source coding theorem, Huffman coding, Shannon-Fano-Elias coding, Channel Coding: Introduction, channel models, channel capacity, channel coding, information capacity theorem, Shannon limit.

Encoding is the process of converting the data or a given sequence of characters, symbols, alphabets etc., into a specified format, for the secured transmission of data.

Decoding is the reverse process of encoding which is to extract the information from the converted format.

Data Encoding

Encoding is the process of using various patterns of voltage or current levels to represent 1s and 0s of the digital signals on the transmission link.

The common types of line encoding are Unipolar, Polar, Bipolar, and Manchester.

Encoding Techniques

The data encoding technique is divided into the following types, depending upon the type of data conversion.

- **Analog data to Analog signals** − The modulation techniques such as Amplitude Modulation, Frequency Modulation and Phase Modulation of analog signals, fall under this category.
- **Analog data to Digital signals** − This process can be termed as digitization, which is done by Pulse Code Modulation (PCM). Hence, it is nothing but digital modulation. As we have already discussed, sampling and quantization are the important factors in this. Delta Modulation gives a better output than PCM.
- **Digital data to Analog signals** − The modulation techniques such as Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK), Phase Shift Keying (PSK), etc., fall under this category. These will be discussed in subsequent chapters.
- **Digital data to Digital signals** − These are in this section. There are several ways to map digital data to digital signals. Some of them are −

Information is the source of a communication system, whether it is analog or digital. Information theory is a mathematical approach to the study of coding of information along with the quantification, storage, and communication of information.

Conditions of Occurrence of Events

If we consider an event, there are three conditions of occurrence.

- If the event has not occurred, there is a condition of uncertainty.
- If the event has just occurred, there is a condition of surprise.
- If the event has occurred, a time back, there is a condition of having some information.

These three events occur at different times. The differences in these conditions help us gain knowledge on the probabilities of the occurrence of events.

5.01 Information Theory:

Information theory provides a quantitative measure of the information contained in a message signal and allow us to determine the capacity of a communication system to transfer this information from source to destination.

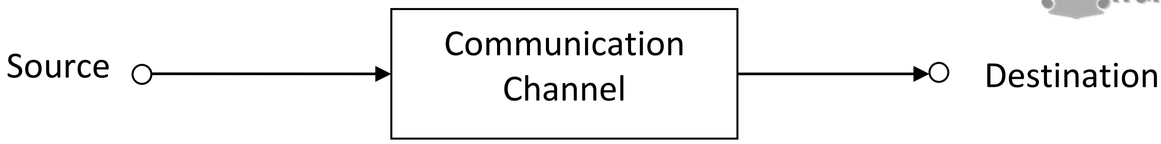


Figure 5.01 A Communication System

1. The information associated with any event depends upon the probability with which it exists.
2. Higher the probability of occurring, lower the information associated with it and vice versa.
3. If the probability of occurring is 1, the information associated with that event is zero since we are certain that a particular bit actually exists at the input of the system.

5.02 Information Source:

An information source is an object that produces an event, the outcome of which is selected at random according to a probability distribution. A discrete information source has only a finite set of symbols as possible outputs.

A source with memory is one for which a current symbol depends on the previous one.

A memory less source is one for which each symbol produced is independent of the previous symbol.

Information: Anything to which some meaning or sense can be attached is called the information.

Example, a written message, a spoken word, a picture etc.

5.03 Unit of Information:

Consider a DMS(discrete memory less source), denoted by X, with alphabet {x_1, x_2, … x_m}. The information content of a symbol x_i, denoted by I(x_i), is defined by

$$I(x_i) = \log_b \frac{1}{P(x_i)} = -\log_b P(x_i) \qquad \text{...5.1.1}$$

Where *P(*x_i*)* is the probability of occurrence of symbol x_i. The *I(*x_i*)* satisfy the following conditions i.e. properties

(1) $I(x_i) = 0 \; for \; P(x_i) = 1$

(2) $I(x_i) \geq 0$

(3) $I(x_i) \geq I(x_j)$ if $P(x_i) \geq P(x_j)$

(4) $I(x_i . x_j) = I(x_i) + I(x_j)$ if x_i and x_j are independent

The unit of *I(*x_i*)* is bit if b=2, Hartley or decit if b=10 and nat (natural unit) if b=e.

$I(x_i) = -\log_2 P(x_i)$ bits

$I(x_i) = -\log_{10} P(x_i)$ hartley or decit

$I(x_i) = -\log_e P(x_i)$ nats

It is standard to use b=2.

The unit bit (b) is a measure of information content.

Conversion of Units

$$\log_2 a = \frac{\ln a}{\ln 2} = \frac{\log a}{\log 2}$$

$$\log_2 6 = \frac{\log_{10} 6}{\log_{10} 2} = \frac{1}{\log_{10} 2}\log_{10} 6$$

5.03 Average Information or Entropy

In a practical Communication system, we usually transmit long sequences of symbol from an information source. Hence it is more important to find the average information that a source produces than the information content of a single symbol.

The mean value of *I(*x_i*)* over the alphabet of source X, with in different symbol is given by-

$$H(x) = E[I(x_i)] = \sum_{i=1}^{m} P(x_i).I(x_i)$$

$$H(x) = E[I(x_i)] = -\sum_{i=1}^{m} P(x_i).\log_b P(x_i) \quad \text{b/symbol}$$

The quantity H(x) is called the entropy of source X. It is a measure of the average information content per source symbol. The source entropy H(x) can be considered as the average amount of uncertainty with in the source X.

For a binary source X, that generates independent symbols 0 & 1, with equal probability, the source entropy H(x) is

$$H(x) = -\frac{1}{2}log_2\left(\frac{1}{2}\right) - \frac{1}{2}log_2\left(\frac{1}{2}\right)$$

$$H(x) = \frac{1}{2}log_2 2 + \frac{1}{2}log_2 2 \qquad \text{1 b/symbol}$$

The source entropy H(x) satisfy the following relation

$$0 \le H(x) \le log_2 m$$

Where m is the number of symbols (also called size of alphabet of source (X)

Case 1 If in input only one digit is occurring i.e. m=1; P(i)=1

$$H(x) = 0$$

Case 2 If all m digits are equal probable i.e. P(i) = 1/m

$$H(x) = -\sum_{i=1}^{m} P(i).\log_2 P(i) = -\sum_{i=1}^{m} \frac{1}{m}.\log_2 \frac{1}{m}$$

$$= -\log_2 \frac{1}{m} = \log_2 m$$

This is the maximum value, therefore

$$H(x)|_{max} = \log_2 m$$

$$0 \le H(x) \le log_2 m$$

Entropy

When we observe the possibilities of the occurrence of an event, how surprising or uncertain it would be, it means that we are trying to have an idea on the average content of the information from the source of the event.

Entropy can be defined as a measure of the average information content per source symbol. Claude Shannon, the "father of the Information Theory", provided a formula for it as –

$$H = -\sum_i p_{i \log_b p_i}$$

Where p_i is the probability of the occurrence of character number i from a given stream of characters and b is the base of the algorithm used. Hence, this is also called as Shannon's Entropy.

The amount of uncertainty remaining about the channel input after observing the channel output, is called as Conditional Entropy. It is denoted by $H(X \mid Y)$.

5.04 Mutual Information

Let us consider a channel whose output is Y and input is X

Let the entropy for prior uncertainty be X = H(x)

(This is assumed before the input is applied)

To know about the uncertainty of the output, after the input is applied, let us consider Conditional Entropy, given that Y = y_k

$$H(x \mid y_k) = \sum_{j=0}^{j-1} p(x_j \mid y_k)\, log_2[\frac{1}{p(x_j \mid y_k)}]$$

$$H(X \mid y = y_0) \ldots \ldots \ldots \ldots \ldots \ldots \ldots . H(X \mid y = y_k)$$

This is a random variable for $p(y_0) \ldots \ldots \ldots \ldots \ldots \ldots \ldots . p(y_{k-1})$ with probabilities respectively. The mean value of $H(X / y = y_k)$ for output alphabet y is –

$$H(X \mid Y) = \sum_{k=0}^{k-1} H(X \mid y = y_k)\, p(y_k)$$

$$= \sum_{k=0}^{k-1} \sum_{j=0}^{j-1} p(x_j \mid y_k)\, p(y_k) log_2[\frac{1}{p(x_j \mid y_k)}]$$

$$= \sum_{k=0}^{k-1} \sum_{j=0}^{j-1} p(x_j, y_k)\, log_2[\frac{1}{p(x_j \mid y_k)}]$$

Now, considering both the uncertainty conditions (before and after applying the inputs), we come to know that the difference, i.e.

$$H(x) - H(\mathrm{x} \mid \mathrm{y})$$

must represent the uncertainty about the channel input that is resolved by observing the channel output. This is called as the Mutual Information of the channel.
Denoting the Mutual Information as *I*(*x*;*y*), we can write the whole thing in an equation, as follows

$$I(x; y) = H(x) - H(\mathrm{x} \mid \mathrm{y})$$

Hence, this is the equational representation of Mutual Information.

Properties of Mutual information

These are the properties of Mutual information.

1. Mutual information of a channel is symmetric.

$$I(x; y) = I(y; x)$$

2. Mutual information is non-negative.

$$I(x; y) \geq 0$$

3. Mutual information can be expressed in terms of entropy of the channel output.

$$I(x; y) = H(y) - H(y \mid x)$$

   Where $H(\, y \mid x\,)$is a conditional entropy.

4. Mutual information of a channel is related to the joint entropy of the channel input and the channel output.

$$I(x; y) = H(x) + H(y) - H(x, y)$$

   Where the joint entropy $H(x, y)$is defined by

$$H(x, y) = \sum_{\mathrm{j}=0}^{\mathrm{j}-1} \sum_{\mathrm{k}=0}^{\mathrm{k}-1} \mathrm{p}(\mathrm{x_j}, \mathrm{y_k}) \log_2[\frac{1}{\mathrm{p}(\mathrm{x_j}, \mathrm{y_k})}]$$

5.05 Conditional Entropy

Using the input probabilities *P(x$_i$)*, output probabilities *P(y$_i$)*, transition probabilities *P(x$_i$/ y$_i$)* and joint probabilities *P(x$_i$, y$_i$)*, we can define the following various entropy functions for a channel with m input and n outputs,

Source Entropy
$$H(X) = -\sum_{i=1}^{m} P(x_i)\log_2 P(x_i)$$

Destination Entropy
$$H(Y) = -\sum_{j=1}^{n} P(y_j)\log_2 P(y_j)$$

The conditional entropy H(X/Y) is a measure of the average uncertainty about the channel input after the channel output has been observed.

$$H(X/Y) = -\sum_{j=1}^{n}\sum_{i=1}^{m} P(x_i, y_j)\log_2 P(x_i/y_j)$$

The conditional entropy H(Y/X) is the average uncertainty of the channel output given that X was transmitted.

$$H(Y/X) = -\sum_{j=1}^{n}\sum_{i=1}^{m} P(x_i, y_j)\log_2 P(y_j/x_i)$$

The joint entropy H(X,Y) is the average uncertainty of the communication channel as a whole

$$H(X,Y) = -\sum_{j=1}^{n}\sum_{i=1}^{m} P(x_i, y_j)\log_2 P(x_i, y_j)$$

5.06 Efficiency

The transmission efficiency or the channel efficiency is defined as

$$\eta = \frac{actual\ transinformation}{maximum\ transinformation}$$

$$\eta = \frac{I(X;Y)}{\max I(X;Y)} = \frac{I(X;Y)}{C_s}$$

Case I If I(X;Y) = max I(X;Y)
Then η = 100%
The channel is fully utilized.

Case II If I(X;Y) < max I(X;Y)
Then η < 100%

Case III If I(X;Y) > max I(X;Y)
Then η > 100%
The situation is avoided.
In the second case, to increase η, we code the input data using

1. Shannon Fano Coding
2. Huffman Coding Procedure

5.07 Discrete Memory less Source

A source from which the data is being emitted at successive intervals, which is independent of previous values, can be termed as discrete memory less source.

This source is discrete as it is not considered for a continuous time interval, but at discrete time intervals. This source is memory less as it is fresh at each instant of time, without considering the previous values.
The Code produced by a discrete memory less source, has to be efficiently represented, which is an important problem in communications. For this to happen, there are code words, which represent these source codes.

5.08 Source Coding

A conversion of the output of a DMS into a sequence of binary symbols is called source coding. The device that performs this conversion is called the source encoder.

Let us take a look at the block diagram.

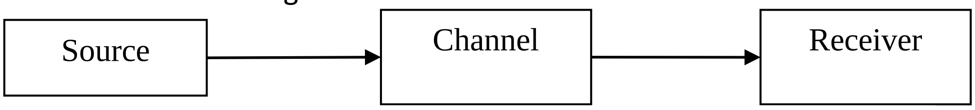


Figure 5.08.01 (a) Without Source Encoder

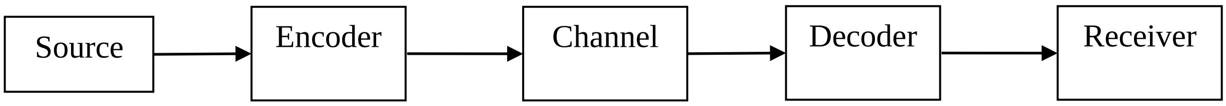


Figure 5.08.01 (b) With Encoder- Decoder

An objective of the source coding is to minimize the average bit rate required for representation of the source by reducing the redundancy (i.e. increasing the efficiency) of the information source.

Code Length

Let X be a DMS with finite entropy H(X) and an alphabet $\{x_1, x_2, \ldots, x_m\}$ with the corresponding probabilities of occurrence $P(x_i)$ $\{i=1,2,\ldots,m\}$.
Let the binary code word assigned to symbol x_i have n_i length, measured in bits. Then the average code word length L per source symbol is given by:

$$\boxed{L = \sum_{i=1}^{m} P(x_i)\, n_i}$$

L represent the average number of bits per source symbol.
The code efficiency η can be defined as

$$\boxed{\eta = \frac{L_{min}}{L}}$$

Where L_{min} is the minimum possible value of L.

Code redundancy γ

$$\boxed{\gamma = 1 - \eta}$$

5.09 Source Coding Theorem

It states that for a DMS-X, with entropy H(X), the average code word length L per symbol is bounded as

$$L \geq H(X)$$

And further L can be made as close to H(X) as desired for some suitable chosen code.
Thus with

$$L_{min} = H(X)$$

$$\boxed{\eta = \frac{H(X)}{L}}$$

Example 5.01 A source delivers six digits with the following probabilities:

i	A	B	C	D	E	F
P(i)	1/2	1/4	1/8	1/16	1/32	1/32

Find (1) H(X) (2) H'(X) = r.H(X) (3) H'(X)$_{max}$=C (4) Efficiency

Solution:

(1)

$$H(X) = -\sum_{i=1}^{m} P(x_i).\log_2 P(x_i) \qquad \text{Bits/ Symbol}$$

$$H(X) = -\sum_{i=1}^{6} P(i).\log_2 P(i)$$

$$= -\left[\frac{1}{2}log_2\frac{1}{2} + \frac{1}{4}log_2\frac{1}{4} + \frac{1}{8}log_2\frac{1}{8} + \frac{1}{16}log_2\frac{1}{16} + \frac{1}{32}log_2\frac{1}{32} + \frac{1}{32}log_2\frac{1}{32}\right]$$

$$= -\left[\frac{1}{2}\times(-1) + \frac{1}{4}\times(-2) + \frac{1}{8}\times(-3) + \frac{1}{16}\times(-4) + \frac{2}{32}\times(-5)\right]$$

$$= 1.938\ bits/symbol$$

(2)

$$H'(X) = r.H(X)$$

$$H(x) = -\sum_{i=1}^{6} P(i).\log_2 P(i)$$

$$H'(X) = 1.938\ bits/sec$$

(3)

$$H(X)|max = \log_2 m = \log_2 6 = \frac{\log_{10} 6}{\log_{10} 2} = 2.58\ bits/symbol$$

$$H'(X)|max = 2.58\ bits/symbol \qquad \text{r=1 symbol/sec}$$

(4)

$$\eta = \frac{H'(X)}{H'(X)|max}\times 100\% = \frac{1.938}{2.58}\times 100\%$$

$$\boxed{\eta = 75.11\%}$$

5.10 Shannon-Fano Coding

An efficient code can be obtained by the following simple procedure, known as Shannon-Fano coding:

1. First write the source symbols in order of decreasing probability,
2. Partition the set into two most equi-probable sub sets and assign a '0' to the upper set and '1' to the lower one,
3. Continue this procedure, each time partitioning the sets with as nearly as equal probabilities as possible until further partitioning is not possible.

Example 5.02 Apply the Shannon Fano coding procedure for the following message ensemble:

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
P	1/4	1/8	1/16	1/16	1/16	1/4	1/16	1/8

Take M=2. Find the Code Efficiency.

Solution:

As per the procedure explained in section 5.10, The code can be obtained as under,

Message	Prob.	Step 1	Step 2	Step 3	Step 4	Code	Code Length
x_1	0.25	0	0			00	2
x_6	0.25	0	1			01	2
x_2	0.125	1	0	0		100	3
x_8	0.125	1	0	1		101	3
x_3	0.0625	1	1	0	0	1100	4
x_4	0.0625	1	1	0	1	1101	4
x_5	0.0625	1	1	1	0	1110	4
x_7	0.0625	1	1	1	1	1111	4

$$H(X) = -\sum_{i=1}^{m} P(i).\log_2 P(i)$$

$$= -\left[\frac{1}{4}log_2\frac{1}{4} + \frac{1}{8}log_2\frac{1}{8} + \frac{1}{16}log_2\frac{1}{16} + \frac{1}{16}log_2\frac{1}{16} + \frac{1}{16}log_2\frac{1}{16} + \frac{1}{4}log_2\frac{1}{4} + \frac{1}{16}log_2\frac{1}{16} + \frac{1}{8}log_2\frac{1}{8}\right]$$

$$= -\left[\frac{1}{4}\times(-2) + \frac{1}{8}\times(-3) + \frac{1}{16}\times(-4)\times 3 + \frac{1}{4}\times(-2) + \frac{1}{16}\times(-4) + \frac{1}{8}\times(-3)\right]$$

$$= \frac{1}{2} + \frac{3}{8} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{2} + \frac{1}{4} + \frac{3}{8}$$

$$= 2.75\ bits/symbol$$

Now the average code word length L per source symbol is

$$L = \sum_{i=1}^{m} P(x_i).n_i$$

$$= \frac{1}{4}\times 2 + \frac{1}{4}\times 2 + \frac{1}{8}\times 3 + \frac{1}{8}\times 3 + \frac{1}{16}\times 4\times 4$$

$$= 2.75\ bits/symbol$$

Then Efficiency

$$\eta = \frac{H(X)}{L} = \frac{2.75}{2.75}100\% = 100\%\ (\boldsymbol{Answer})$$

Example 5.03 A DMS has seven messages with probabilities

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7
P	**0.4**	**0.2**	**0.12**	**0.08**	**0.08**	**0.08**	**0.04**

Apply the Shannon Fano coding procedure and calculate the efficiency of the code. Take M=2

Solution:

Message	Prob.	Step 1	Step 2	Step 3	Step 4	Code	Code Length
X_1	0.4	0	0			0	1
X_2	0.2	1	0	0		100	3
X_3	0.12	1	0	1		101	3
X_4	0.08	1	1	0	0	1100	4
X_5	0.08	1	1	0	1	1101	4
X_6	0.08	1	1	1	0	1110	4
X_7	0.04	1	1	1	1	1111	4

Then Entropy is

$$H(X) = -\sum_{i=1}^{m} P(x_i).\log_2 P(x_i)$$

$$= -[(0.4)log_2(0.4) + (0.2)log_2(0.2) + (0.12)log_2(0.12) + (0.08)log_2(0.08) + (0.08)log_2(0.08) + (0.08)log_2(0.08) + (0.04)log_2(0.04)]$$

$$= \quad 2.42\ letters/symbol$$

Now the average code word length L per source symbol is

$$L = \sum_{i=1}^{m} P(x_i).n_i$$

$$= 0.4 \times 1 + 0.2 \times 3 + 0.12 \times 3 + 0.08 \times 4 + 0.08 \times 4 + 0.08 \times 4 + 0.04 \times 4$$

$$= \quad 2.48\ bits/message$$

Then Efficiency

$$\eta = \frac{H(X)}{Lavg} \times 100\% = \frac{2.42}{2.48} 100\% = 97.58\%\ (\boldsymbol{Answer})$$

5.10 Huffman Encoding

Huffman Encoding results in a code that has the highest efficiency. The Huffman Coding procedure is as under:

1. List the source symbol in decreasing probability,
2. Combine the probabilities of the two symbols having the least probabilities and reorder the resultant probabilities. This step is called reduction. The same procedure is repeated until there are two ordered probabilities remaining,
3. Start reduction with the last reduction, which consists of exactly two ordered probabilities. Assign '0' to the first probabilities and a '1' to the second probability.
4. Now assign '0' and '1' for the probabilities that were combined in the previous reduction step, until the first reduction step.

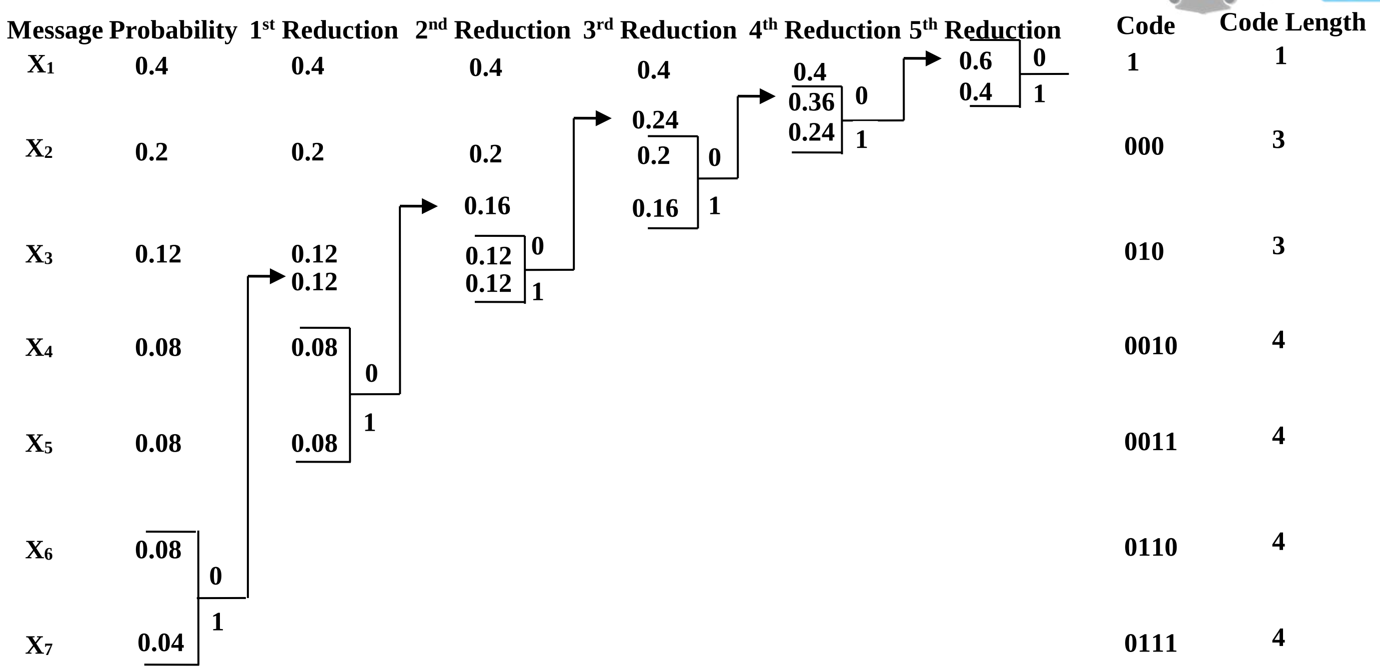
Example 5.04 A DMS has seven messages with probabilities

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7
P	0.4	0.2	0.12	0.08	0.08	0.08	0.04

Apply Huffman coding procedure and calculate the efficiency of the code.

Solution:

Applying the Huffman's Coding procedure and determining the codes as under



Now, the Entropy

$$H(X) = -\sum_{i=1}^{m} P(x_i).\log_2 P(x_i)$$
$$= -[(0.4)log_2(0.4) + (0.2)log_2(0.2) + (0.12)log_2(0.12) + (0.08)log_2(0.08) + (0.08)log_2(0.08) + (0.08)log_2(0.08) + (0.04)log_2(0.04)]$$
$$= \quad 2.42\ letters/symbol$$

Now the average code word length L per source symbol is

$$L = \sum_{i=1}^{m} P(x_i).n_i$$
$$= 0.4 \times 1 + 0.2 \times 3 + 0.12 \times 3 + 0.08 \times 4 + 0.08 \times 4 + 0.08 \times 4 + 0.04 \times 4$$
$$= \quad 2.48\ bits/message$$

Then Efficiency

$$\eta = \frac{H(X)}{Lavg} \times 100\% = \frac{2.42}{2.48} 100\% = 97.58\%\ (\boldsymbol{Answer})$$

5.11 Shannon's Theorem

Given a source of equally likely messages, with M>>1, which is generating information at a rate R. Given a channel with channel capacity C. Then if R≤C, there exists a coding technique such that the output of the source may be transmitted over the channel with the probability of error of receiving the message signal very small.

Thus according to the theorem if R≤C, then the noise free transmission is possible in the presence of noise. The negative theorem states that if the information rate R exceeds C (R>C), error probability approaches to unity as M increases.

5.12 Channel Capacity

We have so far discussed mutual information. The maximum average mutual information, in an instant of a signaling interval, when transmitted by a discrete memory less channel, the probabilities of the rate of maximum reliable transmission of data, can be understood as the channel capacity.

It is denoted by C and is measured in bits per channel use.

5.13 Shannon Limit

Shannon-Hartley equation relates the maximum capacity (transmission bit rate) that can be achieved over a given channel with certain noise characteristics and bandwidth. For an AWGN the maximum capacity is given by

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

Here C is the maximum capacity of the channel in bits/second otherwise called Shannon's capacity limit for the given channel, B is the bandwidth of the channel in Hertz, S is the signal power in Watts and N is the noise power, also in Watts. The ratio S/N is called Signal to Noise Ratio (SNR)**.** It can be ascertained that the maximum rate at which we can transmit the information without any error, is limited by the bandwidth, the signal level, and the noise level. It tells how many bits can be transmitted per second without errors over a channel of bandwidth B Hz, when the signal power is limited to S Watt and is exposed to Gaussian White Noise of additive nature.

Example 5.05 Calculate the channel capacity of a channel with BW 3kHz & S/N ratio given as 10^3, assuming that the noise is white Gaussian noise.

Solution:

We know that

$$C = B\log_2\left(1+\frac{S}{N}\right)$$
$$= 3000\log_2(1+10^3)$$
$$= 30000\ bits/sec$$

5.14 Binary Symmetric Channel (BSC)

A symmetric channel is defined as the one for which

(i) $H(Y/x_j)$ is independent of j: the entropy corresponding to each row of P(Y/X) is the same. &

(ii) $\sum_{j=1}^{m} P\left(y_k/x_j\right)$ is independent of k i.e. the sum of all columns of P(Y/X) is the same.

For a symmetric channel

$$I(X{:}Y) = H(Y) - H(Y/X)$$
$$= H(Y) - A\sum_{j=1}^{m} H\left(\frac{Y}{x_j}\right)P\left(x_j\right)$$
$$= H(Y) - A\sum_{j=1}^{m} P\left(x_j\right)$$

Where $A=H(Y/x_j)$ is independent of j and hence taken out of the summation sign. Also

$$\sum_{j=1}^{m} P\left(x_j\right) = 1$$

therefore $$I(X{:}Y) = H(Y) - A$$

Hence the channel capacity

$$C = \max I(x{:}y)$$
$$= \max[H(Y) - A]$$

$$= \max[H(Y)] - A$$
$$C = \log n - A$$

Where n is the total number of receiver symbols, and max[H(Y)] = logn.

Binary Symmetric Channel;

The most important case of a symmetric channel is BSC. In this case m=n=2, and the channel matrix is

$$D = \left[P\left(\frac{Y}{X}\right)\right] = \begin{bmatrix} P & 1-P \\ 1-P & P \end{bmatrix} = \begin{bmatrix} P & q \\ q & P \end{bmatrix}$$

BSC is shown graphically as shown

The channel is symmetric because the probability of receiving a 1, if 0 was transmitted is same as the probability of receiving a 0 if a 1 was transmitted. This common transition probability is given by P.

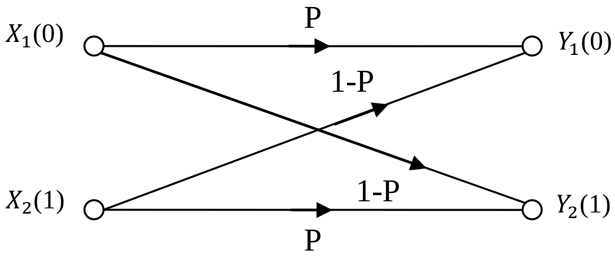
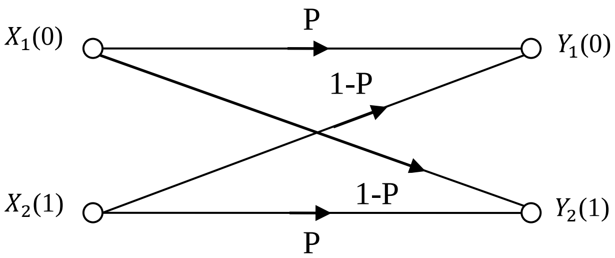


Figure 5.14.1 BSC

Example 5.06 For the Binary Symmetric Channel Calculate the channel capacity for (i) P=0.9 and (ii) P= 0.6



Solution:

We know that

$$C = \log n - A$$
$$C = \log 2 - H\left(\frac{Y}{x_j}\right)$$
$$C = \log_2 2 - \left[-\sum_{j=1}^{2} P\left(\frac{y_k}{x_j}\right)\log_2 P\left(\frac{y_k}{x_j}\right)\right]$$
$$C = \log 2 + P\log P + \ (1-P)\log(1-P)$$
$$C = \log 2 + (P\log P + q\log q)$$
$$C = 1 + (P\log P + q\log q)$$
$$C = 1 + H(P) = 1 - H(q)$$

(i) for p=0.9

$$C = 1 + (0.9\log 0.9 + 0.1\log 0.1)$$
$$C = 0.531 \text{ bits / message}$$

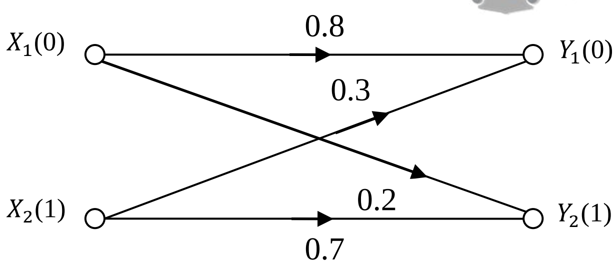
(ii) for p=0.6

$$C = 1 + (0.6\log 0.6 + 0.4\log 0.4)$$
$$C = 0.029 \text{ bits / message}$$

Example 5.07 Find the entropy of the source, information rate, average length and efficiency if the rate of message generation is 300message per seconds and if the symbols and the probabilities are as under

X	**x_1**	**x_2**	**x_3**	**x_4**	**x_5**	**x_6**	**x_7**	**x_8**	**x_9**	**x_{10}**
P	**0.1**	**0.13**	**0.01**	**0.04**	**0.08**	**0.29**	**0.06**	**0.22**	**0.05**	**0.02**

Example 5.08 For the Channel shown in figure, Calculate the channel capacity.



Solution:

The channel matrix is given by

$$[P(Y/X)] = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$$

Since this is an un symmetric channel therefore the channel capacity of a binary channel is

$$C = \log(2^{Q_1} + 2^{Q_2})$$

Where Q_1 and Q_2 are defined by [P].[Q]=[H], therefore

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} P_{11}\log P_{11} + P_{12}\log P_{12} \\ P_{21}\log P_{21} + P_{22}\log P_{22} \end{bmatrix}$$

$$= \begin{bmatrix} 0.8\log 0.8 + 0.2\log 0.2 \\ 0.3\log 0.3 + 0.7\log 0.7 \end{bmatrix}$$

$$= \begin{bmatrix} -0.2576 - 0.4644 \\ -0.5211 - 0.3602 \end{bmatrix} = \begin{bmatrix} -0.722 \\ -0.8813 \end{bmatrix}$$

Therefore $$0.8Q_1 + 0.2Q_2 = -0.722$$

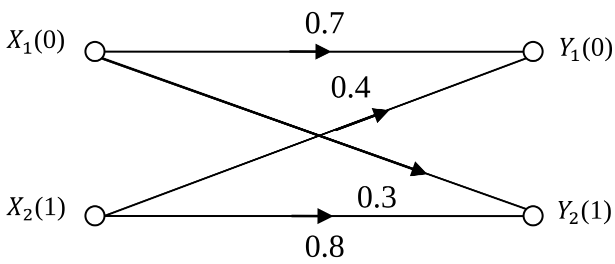
and $$0.3Q_1 + 0.7Q_2 = -0.8813$$

After solving $$Q_1 = -0.6568 \text{ and } Q_2 = -0.9764$$

Then $$C = \log(2^{Q_1} + 2^{Q_2}) = \log(2^{-0.6558} + 2^{-0.9764})$$

$$C = \log(0.6343 + 0.5082) = \log_2(1.14255) = 0.192\ bit/message\ \boldsymbol{Answer}$$

Example 5.09 For the Channel shown in figure, Calculate the channel capacity.



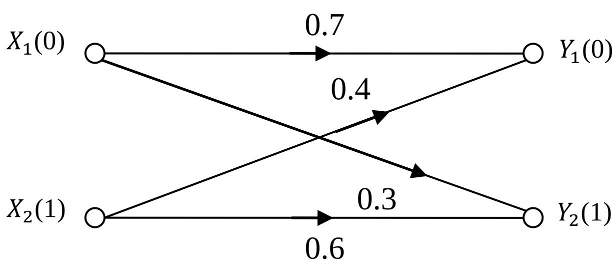
Solution

After solving $$Q_1 = -1.061 \text{ and } Q_2 = -0.456$$

Then $$C = \log(2^{-1.061} + 2^{-0.456})$$

$$C = \log(0.4793 + 0.729) = 0.273\ bit/symbol\ \boldsymbol{Answer}$$

Example 5.10 For the Channel shown in figure, Calculate the channel capacity.



Solution

After solving $$Q_1 = -0.79 \text{ and } Q_2 = -1.09$$

$$C = 0.067\ bit/symbol\ \boldsymbol{Answer}$$