6.1 INTRODUCTION

Statistical study of communication engineering will be dealt with in the concept of Information Theory.

The most general basic communication system can be expressed as

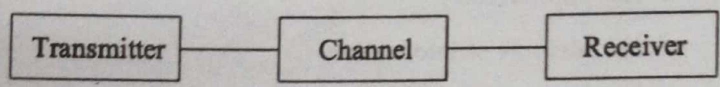

FIGURE 6.1:

Practically, some part of the information (useful) transmitted does not reach the destination or it may undergo some distortion and reach the receiver. So, it is concluded that the communication channel is generally lossy.

The unwanted sources or signals which disturb the useful signal or introduces distortion to it are referred to as noise sources or simply noise or additive noise. A communication system which is less effected by noise is a good communication system i.e. the important consideration in the design of a communication system is to minimise the effect of noise.

This is achieved by the process called "Encoding" at the transmitter and "Decoding" at the receiver.

An encoded message is less sensitive to channel noise and the decoder is used to transform the encoded messages into their pre-encoded version i.e. into a form that is acceptable to the receiver.

Considering all these into consideration, the general communication system can be modelled as

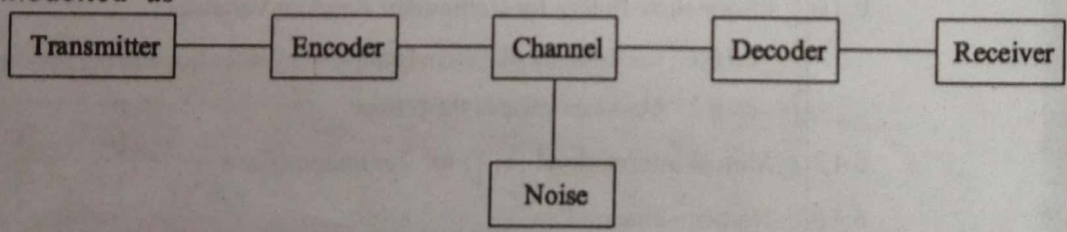

FIGURE 6.2:

6.2 MEASURE OF INFORMATION

Unpredictability is the most significant feature of a practical communication system.

The basic principle involved in determining the information content of a message is that "the information content of a message increases with its uncertainity."

i.e. more the uncertainity of an event (occurrence of the message), more is the information content associated with it.

For justification, consider the message as "Sun sets in the West". Since, the above is a regular, day to day phenomenon, the above event is certain to occur and hence the probability of occurrence of the above event is 1. Since there is nothing new in the above event, we will not show any interest in it i.e. since the uncertainity of above event is less (zero), information content is also less (zero).

If some news is there saying that Sun is going to set in the East on a particular day which is most uncertain, we all will eagerly wait for that day and to see that uncertain event i.e. we can say that the above event is carrying a lot of amount of information.

i.e. More uncertain the event is, more is the information content associated with it.

So, the information content (I) of a message varies inversely with its certainity i.e. its probability of occurrence.

$$\therefore \qquad I \propto \frac{1}{P}$$

where I = Amount of information received from the knowledge of occurrence of the event.

P = Probability of occurence of the event.

Consider a communication system in which the allowable messages are $m_1, m_2 \ldots$ with prob. of occurrence $P_1, P_2 \ldots$ such that $P_1 + P_2 + \ldots = 1$

Let the transmitter selects a message m_k, with probability P_k and let the receiver has correctly identified the message.

Then, the system has conveyed an amount of information

$$I_k \propto \frac{1}{P_K}$$

With two independent messages, m_1 and m_2 correctly identified, the amount of information conveyed is the sum of the information asociated with each of the messages

i.e. $$I(m_1, m_2) = I(m_1) + I(m_2) \qquad \ldots\ldots\ldots\ldots (1)$$

where $I(m_1, m_2)$ = Information associated with joint occurrence of the events m_1 and m_2.

we have $$I(m_1, m_2) \propto \frac{1}{P(m_1, m_2)}$$

where $P(m_1, m_2)$ = Joint probability of occurrence of the events m_1, m_2.

$$\therefore \qquad I(m_1, m_2) \propto \frac{1}{P(m_1) \cdot P(m_2)}$$

because m_1 and m_2 are two independent events.

According to equation (1), we require a mathematical operation which converts the product into addition and it is the logarithmic function.

So, if $$I(m_1, m_2) = \log \frac{1}{P(m_1) \cdot P(m_2)}$$

then, $$I(m_1, m_2) = \log \frac{1}{P(m_1)} + \log \frac{1}{P(m_2)}$$

$\therefore$ So, the information associated with an event is defined as

$$I = \log \frac{1}{(Probability\ of\ occurrence\ of\ event\ )}$$

i.e.,

$$I = \log \frac{1}{P} \quad \text{or} \quad I = -\log P$$

Consider the case of selection between two events E_1 and E_2.

The amount of information associated with the selection of one out of two equiprobable events is $\log \frac{1}{1/2} = \log 2$.

6.3 UNITS OF INFORMATION

If the base of above logarithm is 2, the units for I are bits.

If the base is e, units are Nats and for the base of 10, units are Hartleys.

6.4 ENTROPY OF A SOURCE

Consider that a discrete source transmitting N no. of messages $m_1, m_2, \ldots m_N$, whose probabilities of occurrence are given by $P_1, P_2, \ldots P_n$.

Assume that a long sequence of L messages is needed from the source.

In the above message sequence of length L, we transmit $P_1 \cdot L$ times m_1, $P_2\, L$ times m_2 and so on.

So, the total information content the above sequence of messages is the sum of the information contents associated with the individual messages.

$$\therefore \quad I_{total} = P_1\, L \cdot \log \frac{1}{P_1} + P_2\, L \cdot \log \frac{1}{P_2} + \ldots$$

Average information / message is

$$\frac{I_{total}}{L} = P_1 \log \frac{1}{P_1} + P_2 \log \frac{1}{P_2} + \ldots$$

$$\therefore \quad \sum_{i=1}^{N} P_i \log \frac{1}{P_1} = -\sum_{i=1}^{N} P_i \log P_i$$

and is referred to as entropy of the source denoted by H.

Consider a source transmitting messages m_1 and m_2 with probabilities P_1 and P_2 respectively.

$$P_1 = 0.01\,,\ P_2 = 0.99 \quad \Rightarrow \quad H = 0.08\ bits/message$$

Here, it is easy to guess whether m_1 or m_2 will occur. Most of the time, m_2 will occur.

$\therefore$ uncertainity is less and hence entropy is also less.

Let $\quad P_1 = 0.4\ ;\ P_2 = 0.6; \quad \Rightarrow \quad H = 0.97$

It is some what difficult to guess whether m_1 will occur or m_2, because P_1 and P_2 are nearly equal. So, uncertainity is more, and hence uncertainity is more.

If $\quad P_1 = 0.5$ and $P_2 = 0.5 \quad \Rightarrow \quad H = 1$

It is extremely difficult to guess whether m_1 or m_2 will occur, as their probabilities are equal. Thus, uncertainity is maximum, and hence the entropy.

Thus Entropy is a measure of uncertainity.

6.5 MAXIMUM VALUE OF ENTROPY

Consider the case of two messages with probability P and $(1-P)$. The average information/message is

$$H = P \log \frac{1}{P} + (1-P) \log \left(\frac{1}{1-P}\right)$$

To find H_{max},

$$\frac{dH}{dP} = \frac{d}{dP}[-P \log P - (1-P) \log (1-P)] = 0$$

$$\Rightarrow \quad \left[-P \times \frac{1}{P} - \log P\right] - [-(1-P)/(1-P) - \log (1-P)] = 0$$

$$\Rightarrow \quad \log \frac{1-P}{P} = 0$$

$$\Rightarrow \quad \frac{1-P}{P} = 1$$

$$\Rightarrow \quad P = \frac{1}{2}$$

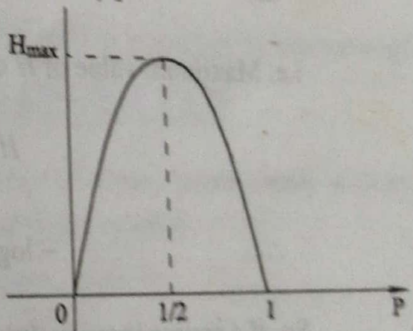

FIGURE 6.3:

So, maximum value of H occurs when the two messages are equally likely and

$$H_{max} = \frac{1}{2} \log 2 + \frac{1}{2} \log 2$$

$$= 1 \text{ bit/message}$$

Similarly, consider the case of three messages with probabilities P_1, P_2 and P_3.

The average information/message is

$$H = -\left[P_1 \log P_1 + P_2 \log P_2 + P_3 \log P_3\right]$$

$$= -\left[P_1 \log P_1 + P_2 \log P_2 + (1 - P_1 - P_2) \log (1 - P_1 - P_2)\right]$$

$$\left[\because \ P_1 + P_2 + P_3 = 1\right]$$

Consider $\frac{dH}{dP_1} = 0$

i.e. $\frac{d}{dP_1}\left[-P_1 \log P_1 - P_2 \log P_2 - (1 - P_1 - P_2) \log (1 - P_2 - P_2)\right] = 0$

$\Rightarrow \quad 2P_1 + P_2 = 1$

Consider $\frac{dH}{dP_2} = 0$

$\Rightarrow \quad 2P_2 + P_1 = 1$

$\therefore \quad 2P_1 + P_2 = 2P_2 + P_1$

$\therefore \quad P_1 = P_2$

So, $2P_1 + P_1 = 1$

$\Rightarrow \quad P_1 = \frac{1}{3}$

and $P_2 = \frac{1}{3}$

$\Rightarrow \quad P_3 = \frac{1}{3}$

i.e. Maximum value of H occurs when the three messages are equally likely and

$$H_{\max} = -\frac{1}{3}\log\frac{1}{3} - \frac{1}{3}\log\frac{1}{3} - \frac{1}{3}\log\frac{1}{3}$$

$$\therefore \quad -\log\frac{1}{3} = \log_2 3 \text{ bits/messages}$$

So, if a source is transmitting 'm' symbols $m_1, m_2, \ldots$ etc with probabilities $P_1, P_2 \ldots$ *etc*, then for

$$H_{\max}, \quad P_1 = P_2 = \ldots = \frac{1}{m}$$

$$\therefore \quad H_{\max} = -\left[\frac{1}{m}\log\frac{1}{m} + \frac{1}{m}\log\frac{1}{m} + \ldots\right]$$

$$= -\log\frac{1}{m} = \log_2 m \text{ bits/message}$$

6.6 PROPERTIES OF ENTROPY

Consider a source 'X' transmitting 'm' no. of messages, with probabilities $P_1, P_2 \ldots$ *etc.*

1. $H(X) = \log_2 m$ if all the 'm' messages are equally likely
2. $\log_2 m \geq H(X) \geq 0$

3. If $m = 1$, then $H(X) = 0$ since, in the case of a single possible message, the reception of the message conveys no information since the uncertainity is zero because we can make decision in favour of the available single message with more certainity (i.e. $P = 1$).

4. The measure of entropy of the source X is invariant with respect to the order of the messages of the source.

i.e. $$H(P_1, P_2 \ldots\ldots\ldots P_m) = H(P_2, P_1 \ldots\ldots\ldots P_m)$$

6.7 INFORMATION RATE

If the source of the messages generates at the rate 'r' messages/sec then the information rate is defined to be $R = r \cdot H$

= Average no. of bits of information/sec

Consider two sources of equal entropy H, generating r_1 and r_2 messages/sec, then $R_1 = r_1 \cdot H$; $R_2 = r_2 \cdot H$

If $r_1 > r_2$ then $$R_1 > R_2$$

The source having high r will require the channel for most of the time than the source, with small 'r' i.e. the source with more 'r' will have greater demand on the communication channel.

So, two sources can be differentiated by not only based on their entropies but also depending on the basis of their information rate.

R is sometimes referred to as entropy in bits/sec and H is referred to as entropy in bits/message.

6.8 MEASURE OF INFORMATION FOR TWO DIMENSIONAL CASE

... length codes can be generated by usig the following techniques.

16.22 SHANON - FANO ENCODING

The different steps involved in the code generation are

1. Arrange the messages in the order of decreasing probability. If there are equal probabilities choose any of the various orderings.
2. Divide the messages into subgroups with as nearly equal probability as possible. This division determines the first binary digit of the code.
3. Each group is then redivided into sub groups with as nearly equal probability as possible and this division determines the 2^{nd} binary digit of the code group.
4. This process is continued till each subgroup contains only one message.

PROBLEM 21

A source is transmitting messages A, B, C, D, E and F with probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$ *and* $\frac{1}{32}$ *respectively. Find the source code words using SHANON-FANO method and also the coding efficiency.*

Solution:

Following the above steps

1.

$$\left.\begin{matrix} A & 1/2 \\ B & 1/4 \\ C & 1/8 \\ D & 1/16 \\ E & 1/32 \\ F & 1/32 \end{matrix}\right\} \text{Messages are arranged in descending order of their probability}$$

2.

$$\underline{A \quad \tfrac{1}{2} \quad 0} \rightarrow \textit{Group 1}$$

$$\left.\begin{matrix} B & \tfrac{1}{4} & 1 \\ C & \tfrac{1}{8} & 1 \\ D & \tfrac{1}{16} & 1 \\ E & \tfrac{1}{32} & 1 \\ F & \tfrac{1}{32} & 1 \end{matrix}\right\} \rightarrow \textit{Group 2}$$

Here the messages are divided into two groups of equal probability and '0' is the first digit of the code word for the messages in group 1 and '1' is the first digit of the code word for the messages in group 2.

3. Leave the message A and continue the above procedure

B	1/4	1	0
C	1/8	1	1
D	1/16	1	1
E	1/32	1	1
F	1/32	1	1

4.

C	1/8	1	1	0
D	1/8	1	1	1
E	1/16	1	1	1
F	1/32	1	1	1

5.

D	1/16	1	1	1	0
E	1/32	1	1	1	1
F	1/32	1	1	1	1

6.

E	1/32	1	1	1	1	0
F	1/32	1	1	1	1	1

This is the last stage of division. So, the code words are

A	0	1
B	10	2
C	110	3
D	1110	4
E	11110	5
F	11111	5

→ Length of each code word

Average code word length is

$$\bar{L} = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + 5 \times \frac{1}{32} + 5 \times \frac{1}{32}$$

$$= 1.9375 \textit{ bits/code word}$$

Entropy of the source is

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

$$= 1.9375 \textit{ bits/message}$$

$$\therefore \quad \eta_c = \frac{H}{\bar{L}} \times 100\,\% = 100\,\%$$

PROBLEM 22

The different messages from a source and their probabilities are given as

P (A) = 1/2; P (B) = 1/6; P (C) = 1/12; P (D) = 1/6; P (E) = 1/12

Find the Shanon-Fano Code words and coding efficiency.

Solution:

1.

A	1/2	0
B	1/6	1
D	1/6	1
C	1/12	1
E	1/12	1

2.

B	1/6	1	0
D	1/6	1	1
C	1/12	1	1
E	1/12	1	1

3.

D	1/6	1	1	0
C	1/12	1	1	1
E	1/12	1	1	1

4.

C	1/12	1	1	1	0
E	1/12	1	1	1	1

The code words are

Message	Code word	L
A	0	1
B	10	2
C	1110	4
D	110	3
E	1111	4

Average Code word length $\overline{L}$

$$1 \times \frac{1}{2} + 2 \times \frac{1}{6} + 4 \times \frac{1}{12} + 3 \times \frac{1}{6} + 4 \times \frac{1}{12}$$

$$= 2 \textit{ bits/word}$$

Entropy of the source is

$$H = -\left[\frac{1}{2}\log\frac{1}{2} + \frac{1}{6}\log\frac{1}{6} + \frac{1}{12}\log\frac{1}{12} + \frac{1}{6}\log\frac{1}{6} + \frac{1}{12}\log\frac{1}{12}\right]$$

$$= 1.959 \textit{ bits/message}$$

$$\therefore \quad \eta_c = 97.95\ \%$$

PROBLEM 23

P (A) = 0.4; P (B) = 0.2; P (C) = 0.2; P (D) = 0.1; P (E) = 0.05 P (F) = 0.03;
P (G) = 0.01; P (H) = 0.01

Find the source code words using Shanon-Fano method and also code efficiency.

Solution:

1.

A	0.4	0
B	0.2	1
C	0.2	1
D	0.1	1
E	0.05	1
F	0.03	1
G	0.01	1
H	0.01	1

2.

B	0.2	1	0
C	0.2	1	1
D	0.1	1	1
E	0.05	1	1
F	0.03	1	1
G	0.01	1	1
H	0.01	1	1

3.

C	0.2	1	1	0
D	0.1	1	1	1
E	0.05	1	1	1
F	0.03	1	1	1
G	0.01	1	1	1
H	0.01	1	1	1

4.

D	0.1	1	1	1	0
E	0.05	1	1	1	1
F	0.03	1	1	1	1
G	0.01	1	1	1	1
H	0.01	1	1	1	1

5.

E	0.05	1	1	1	1	0
F	0.03	1	1	1	1	1
G	0.01	1	1	1	[illegible]	1
H	0.01	1	1	1	1	1

6.

F	0.03	1	1	1	1	1	0
G	0.01	1	1	1	1	1	1
H	0.01	1	1	1	1	1	1

7.

G	0.01	1	1	1	1	1	1	0
H	0.01	1	1	1	1	1	1	1

So, the code words are

Message	Code Word	L
a	0	1
B	10	2
C	110	3
D	1110	4
E	11110	5
F	111110	6
G	1111110	7
H	1111111	7

$\therefore$ $I = 2.37$ *bits/word* and

$H = 2.29$ *bits/message*

and $\eta_c = 96.65\ \%$