UNIT 3

FUNDAMENTAL LIMITS ON PERFORMANCE AND CHANNELS FOR COMMUNICATION

3.1 COMMUNICATION CHANNELS - INTRODUCTION

A *"channel"*, in the broad sense, is defined as the medium through which the coded signals generated by an information source are transmitted. The complete block diagram of a communication system consisting of various terminological channels is shown in figure 3.1.

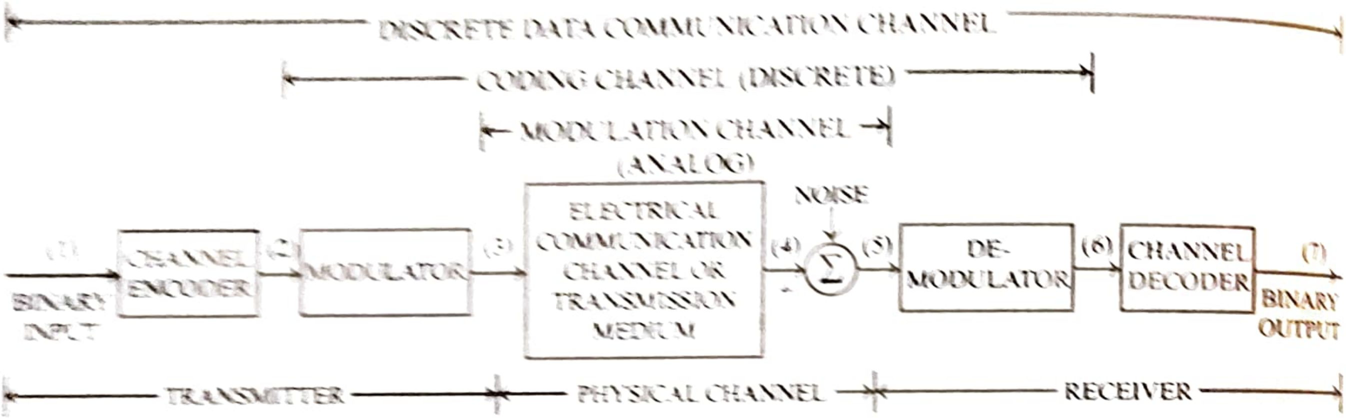

Fig. 3.1 : Block diagram of communication system

A practical communication system consists of a transmitter, physical channel and a receiver. The transmitter consists of an encoder and modulator, while the receiver consists of a demodulator and a decoder.

Between points (2) and (6) in figure 3.1, we have the discrete channel referred to as *"coding channel"*, which accepts a sequence of symbols at its input and produces a sequence of symbols at its output.

The communication channel between points (3) and (5) provides the electrical connection between the transmitter and the receiver. The input and output are analog electrical waveforms. This portion of the channel is often called a ***"continuous or modulation channel"***. Examples are voiceband and wideband telephone systems, high frequency radio systems and troposcatter systems.

Between points (1) and (7) in the diagram, we have the binary data being presented at the input and binary data back at the output and hence the channel is called ***"Data Communication Channel (discrete)"***.

In the channel, there are various kinds of disturbances. These disturbances may be due to amplitude and frequency response variations of the channel within its passband, variations in channel characteristics and also due to non-linearities in the channel. In addition, channel can also corrupt the signal due to various types of additive and multiplicative noise. All these disturbances introduce errors in data transmission and limit the maximum rate at which data can be transferred over the channel.

An important characteristic of a data communication system is the ***"channel capacity"*** which represents the maximum rate at which data is transferred across the channel, with an arbitrarily small probability of error.

In the following sections, we shall discuss discrete communication channels and understand the concept of channel capacity through Shannon-Hartley law.

3.2 DISCRETE COMMUNICATION CHANNELS

The communication channel between points 2 and 6 in figure 3.1 is discrete in nature. In the general case, the input to the channel is a symbol belonging to an alphabet "A" with "r" symbols. The output of the channel is a symbol belonging to some other alphabet "B" with "s" symbols. Due to errors in the channel, the output symbol may be different from the input symbol during any symbol interval. Errors are mainly due to noise in the analog portion of the channel. Let us see how a discrete communication channel is completely modelled in the following section.

3.3 REPRESENTATION OF A CHANNEL

As indicated in the previous section, a communication channel may be represented by a set of input alphabet $A = \{a_1, a_2, \ldots\ldots a_r\}$ consisting of 'r' symbols, a set of output alphabet $B = \{b_1, b_2, \ldots\ldots b_s\}$ consisting of 's' symbols and a set of conditional probabilities $P(b_j/a_i)$ with $i = 1, 2, \ldots, r$ and $j = 1, 2, \ldots, s$, as shown in figure 3.2.

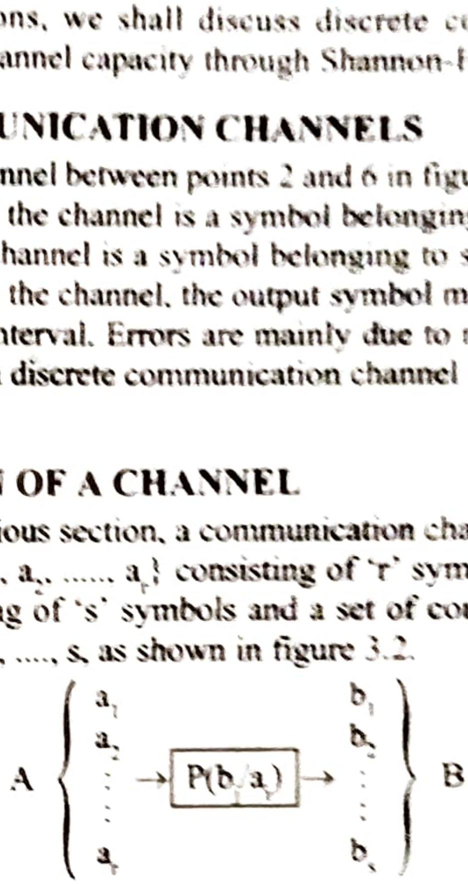

Fig. 3.2 : Representation of a channel.

These conditional probabilities come into existence due to the presence of noise in the channel. Because of this noise, there will be some amount of uncertainty about the reception of any symbol. For this reason, we have a different number of symbols 's' at the receiver from 'r' symbols at the transmitter.

Totally, we have, r × s number of conditional probabilities which are represented in a ***"matrix"*** form with all the input symbols represented row-wise and output symbols column-wise. Such a matrix is called ***"CHANNEL MATRIX"*** or ***"NOISE MATRIX"*** as given below:

$$P(b_j/a_i) \text{ or } P(B/A) = \begin{array}{c} \\ a_1 \\ a_2 \\ \vdots \\ a_r \end{array} \begin{array}{c} \begin{array}{cccc} b_1 & b_2 & \cdots\cdots & b_s \end{array} \\ \begin{bmatrix} P(b_1/a_1) & P(b_2/a_1) & \cdots\cdots & P(b_s/a_1) \\ P(b_1/a_2) & P(b_2/a_2) & \cdots\cdots & P(b_s/a_2) \\ \vdots & \vdots & \cdots\cdots & \vdots \\ P(b_1/a_r) & P(b_2/a_r) & \cdots\cdots & P(b_s/a_r) \end{bmatrix} \end{array} \quad \ldots\ldots (3.1)$$

To understand clearly the effect of noise and intersymbol conversion, let us consider the following example.

Example 3.1 : Let us consider a transmitter emitting discrete symbols from an input alphabet A. Let the symbols be $\{a_1, a_2, a_3\}$ which are encoded using binary coding as 00, 01 and 10 respectively. Due to noise present in the channel, we may receive 4 symbols at the receiver with output alphabet B given by $\{b_1, b_2, b_3, b_4\}$ with codewords 00, 01, 10 and 11. Note that we are not transmitting any symbol corresponding code-word "11" at the transmitter but we are receiving it due to noise present in the channel. Complete diagram showing all the transmitter symbols between them is picturised as shown in figure 3.3. Such a diagram is called ***"CHANNEL DIAGRAM"*** or ***"NOISE DIAGRAM"***.

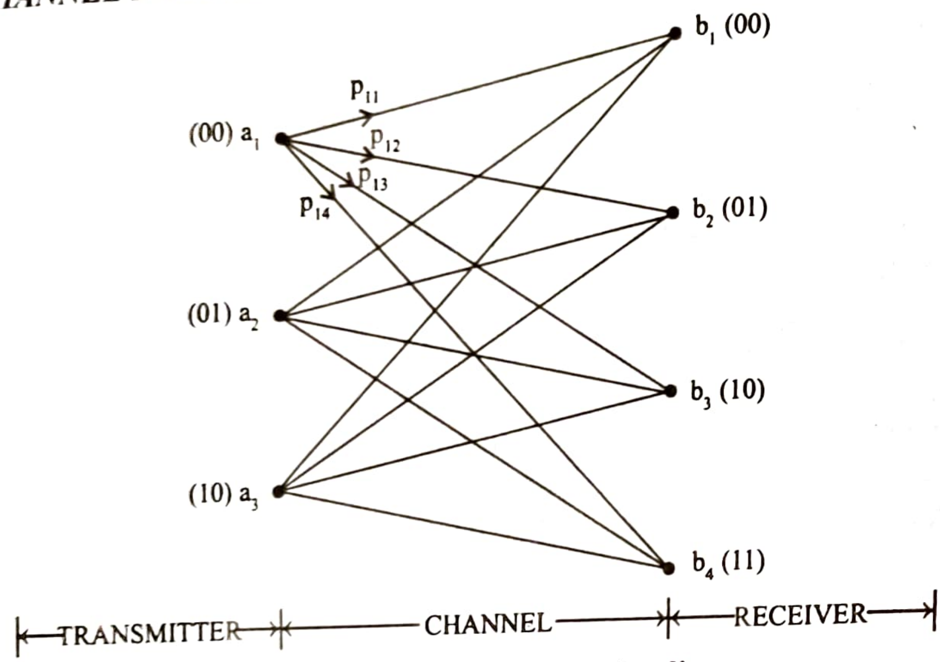

Fig. 3.3 : Illustrating channel or noise diagram

In the channel diagram of figure 3.3, let us consider a particular case of the input symbol "a_1" being transmitted having a code-word "00". When these two binary symbols are transmitted over the channel, there is some chance, that noise might not affect both the symbols and it is received as it is at the receiver. This ***"chance"*** is the conditional probability that the symbol b_1 is received given that a_1 is transmitted at the transmitter denoted by $P(b_1/a_1)$. This is shown as p_{11} on the line connecting a_1 and b_1 in the channel diagram.

$$\therefore \quad p_{11} = P(b_1/a_1) \qquad \text{...... (3.2)}$$

When '00' is transmitted, there may be a chance that the second '0' is converted into '1' by the noise and "01" is received at the output. This chance is again another conditional probability of receiving b_2 given that a_1 is transmitted denoted by $P(b_2/a_1)$ shown as p_{12} in the channel diagram.

$$\therefore \quad p_{12} = P(b_2/a_1) \qquad \text{...... (3.3)}$$

$$\text{Similarly} \quad p_{13} = P(b_3/a_1) \qquad \text{...... (3.4)}$$

= Conditional probability of receiving $b_3(10)$ given that a_1 is transmitted with noise affecting the first '0'

and $p_{14} = P(b_4/a_1)$ ...... (3.5)

= Conditional probability of receiving $b_4(11)$ given that a_1 is transmitted with noise affecting both the symbols.

It is evident that when "00" (a_1) is transmitted, we have to receive only one of the four output symbols $b_1(00)$ or $b_2(01)$ or $b_3(10)$ or $b_4(11)$. [Note that noise can only convert a symbol]. Because of this we must have

$$p_{11} + p_{12} + p_{13} + p_{14} = 1 \quad \text{...... (3.6)}$$

Using equation (3.2) to (3.5) in (3.6), we get

$$P(b_1/a_1) + P(b_2/a_1) + P(b_3/a_1) + P(b_4/a_1) = 1 \quad \text{...... (3.7)}$$

Generalizing equation (3.7) for "s" output symbols and "r" input symbols corresponding to figure 3.2, we get

$$P(b_1/a_1) + P(b_2/a_1) + P(b_3/a_1) + \text{.........} + P(b_s/a_1) = 1 \quad \text{...... (3.8)}$$

Writing with summation sign,

$$\sum_{j=1}^{s} P(b_j / a_1) = 1 \quad \text{...... (3.9)}$$

But equation (3.9) must hold good for all "r" input symbols

$$\sum_{j=1}^{s} P(b_j / a_i) = 1 \quad \text{...... (3.10)}$$

From equation (3.10), it is clear that ***"the sum of all the elements in any row of the channel matrix is equal to unity"***.

Let us suppose that the probabilities of the input symbols namely $P(a_1)$, $P(a_2)$, ..., $P(a_r)$ are also known such that

$$P(a_1) + P(a_2) + \ldots + P(a_r) = 1$$

$$\therefore \quad \sum_{i=1}^{r} P(a_i) = 1 \quad \text{...... (3.11)}$$

Knowing the channel matrix elements $P(b_j/a_i)$ for all i and j and the input probabilities $P(a_i)$ for all i, the probabilities of the output symbols $P(b_j)$ for all j, can be found using the ***"theorem of total probability"*** [refer section 1.5] as

$$\left.\begin{array}{lcl} P(b_1) & = & P(b_1/a_1)\,P(a_1) + P(b_1/a_2)\,P(a_2) + \ldots. + P(b_1/a_r)\,P(a_r) \\ P(b_2) & = & P(b_2/a_1)\,P(a_1) + P(b_2/a_2)\,P(a_2) + \ldots. + P(b_2/a_r)\,P(a_r) \\ \vdots & = & \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \vdots & = & \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ P(b_s) & = & P(b_s/a_1)\,P(a_1) + P(b_s/a_2)\,P(a_2) + \ldots. + P(b_s/a_r)\,P(a_r) \end{array}\right\} \text{... (3.12)}$$

Knowing the input probabilities $P(a_i)$, channel matrix elements $P(b_j/a_i)$ and the output probabilities $P(b_j)$, the *"input conditional probabilities $P(a_i/b_j)$"* can be found by using Baye's Rule.

$$P(a_i/b_j) = \frac{P(b_j/a_i)P(a_i)}{P(b_j)} \quad \text{...... (3.13)}$$

JOINT PROBABILITY :

The joint probability between any input symbol "a_i" and any output symbol "b_j" is given by

$$P(a_i, b_j) = P(a_i \cap b_j) = P(b_j/a_i)\, P(a_i) = P(a_i/b_j)\, P(b_j) \quad \text{...... (3.14)}$$

When we multiply all the elements of the 1st row of channel matrix of equation (3.1) by $P(a_1)$, 2nd row by $P(a_2)$ ......, r^{th} row by $P(a_r)$, we get

$$P(b_j/a_i)\,P(a_i) = \begin{array}{c c} & \begin{array}{cccc} b_1 & b_2 & \cdots\cdots & b_s \end{array} \\ \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ \\ a_r \end{array} & \begin{bmatrix} P(b_1/a_1)\,P(a_1) & P(b_2/a_1)\,P(a_1) & \cdots\cdots & P(b_s/a_1)\,P(a_1) \\ P(b_1/a_2)\,P(a_2) & P(b_2/a_2)\,P(a_2) & \cdots\cdots & P(b_s/a_2)\,P(a_2) \\ \vdots & \vdots & \cdots\cdots & \vdots \\ & & \cdots\cdots & \\ P(b_1/a_r)\,P(a_r) & P(b_2/a_r)\,P(a_r) & \cdots\cdots & P(b_s/a_r)\,P(a_r) \end{bmatrix} \end{array} \quad \text{... (3.15)}$$

Using equation 3.14 in 3.15, we get

$$P(a_i, b_j) \text{ or } P(A, B) = \begin{array}{c c} & \begin{array}{cccc} b_1 & b_2 & \cdots\cdots & b_s \end{array} \\ \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ \\ a_r \end{array} & \begin{bmatrix} P(a_1, b_1) & P(a_1, b_2) & \cdots\cdots & P(a_1, b_s) \\ P(a_2, b_1) & P(a_2, b_2) & \cdots\cdots & P(a_2, b_s) \\ \vdots & \vdots & \cdots\cdots & \vdots \\ & & \cdots\cdots & \\ P(a_r, b_1) & P(a_r, b_2) & \cdots\cdots & P(a_r, b_s) \end{bmatrix} \end{array} \quad \text{... (3.16)}$$

The matrix of equation (3.16), whose elements are the various joint probabilities between input and output symbols, is called ***"JOINT PROBABILITY MATRIX"*** popularly written as JPM denoted by $P(a_i, b_j)$ or $P(A, B)$. The JPM has very interesting properties which are quite useful. They are listed below:

Property - 1 : From the first equation of equation (3.12) we have

$$P(b_1) = P(b_1/a_1)\,P(a_1) + P(b_1/a_2)\,P(a_2) + ... + P(b_1/a_r)\,P(a_r) \quad \text{...... (3.17)}$$

Making use of equation (3.14) in (3.17), we get

$$P(b_1) = P(a_1, b_1) + P(a_2, b_1) + ... + P(a_r, b_1) \quad \text{...... (3.18)}$$

These joint probabilities of R.H.S. of equation (3.18) are observed to be present in the ***"1st column of JPM"*** of equation (3.16). Hence, we can conclude that by adding all the elements of the 1st column, we get the probability of 1st output symbol b_1.

Similarly, $$P(b_2) = P(a_1, b_2) + P(a_2, b_2) + ... + P(a_r, b_2) \qquad ...... (3.19)$$

$$\vdots \quad = \quad \vdots \qquad \vdots \qquad \vdots$$

$$P(b_s) = P(a_1, b_s) + P(a_2, b_s) + ... + P(a_r, b_s) \qquad ...... (3.20)$$

From equations (3.18) to (3.20) we can derive the first property of JPM as below:

"By adding the elements of JPM columnwise, we can obtain the probability of output symbols".

This property can be expressed as

$$\sum_{i=1}^{r} P(a_i, b_j) = P(b_j) \qquad ...... (3.21)$$

Property - 2 : Using the theorem of total probability for the input symbols, we have for the first symbol a_1,

$$P(a_1) = P(a_1/b_1) P(b_1) + P(a_1/b_2) P(b_2) + ... + P(a_1/b_s) P(b_s)$$

$$\therefore \quad P(a_1) = P(a_1, b_1) + P(a_1, b_2) + ... + P(a_1, b_s) \qquad ...... (3.22)$$

Similarly $$P(a_2) = P(a_2, b_1) + P(a_2, b_2) + ... + P(a_2, b_s) \qquad ...... (3.23)$$

$$\vdots \quad \vdots \qquad \vdots \qquad \vdots$$

$$P(a_r) = P(a_r, b_1) + P(a_r, b_2) + ... + P(a_r, b_s) \qquad ...... (3.24)$$

These joint probabilities of R.H.S. of equations (3.22) to (3.24) are observed to be present in successive ***"rows"*** of the matrix of equation (3.16). Hence, we have the second property of JPM as given below:

"By adding the elements of JPM row wise we can obtain the probability of input symbols".

This property can also be expressed as

$$\sum_{j=1}^{s} P(a_i, b_j) = P(a_i) \qquad ...... (3.25)$$

Property - 3 : By adding equations (3.22) to (3.24), we get

$$\begin{aligned} P(a_1) + P(a_2) + ... + P(a_r) = & [P(a_1, b_1) + P(a_1, b_2) + ... + P(a_1, b_s)] \\ & + [P(a_2, b_1) + P(a_2, b_2) + ... + P(a_2, b_s)] \\ & + \dots\dots\dots\dots\dots \\ & + [P(a_r, b_1) + P(a_r, b_2) + ... + P(a_r, b_s)] \qquad ...... (3.26) \end{aligned}$$

The L.H.S. of equation (3.26) = 1 from equation (3.11) and the R.H.S. of equation (3.26) is the sum of all the elements of JPM. Thus, we have the 3rd property of JPM as:

"The sum of all the elements of JPM is equal to unity".

This is again expressed as

$$\sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) = 1 \qquad \text{...... (3.27)}$$

Example 3.2 : In a communication system, a transmitter has 3 input symbols $A = \{a_1, a_2, a_3\}$ and receiver also has 3 output symbols $B = \{b_1, b_2, b_3\}$. The matrix given below shows JPM with some marginal probabilities:

$a_i \backslash b_j$	b_1	b_2	b_3
a_1	$\frac{1}{12}$	*	$\frac{5}{36}$
a_2	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{5}{36}$
a_3	*	$\frac{1}{6}$	*
$P(b_j)$	$\frac{1}{3}$	$\frac{14}{36}$	*

(i) Find the missing probabilities (*) in the table.

(ii) Find $P(b_3/a_1)$ and $P(a_1/b_3)$

(iii) Are the events a_1 and b_1 statistically independent? Why?

Solution

Addition of R.H.S. of equations (3.18) to (3.20) also yields all the elements of JPM which add up to unity.

$$P(b_1) + P(b_2) + ... + P(b_s) = 1$$

$$\therefore \quad \sum_{j=1}^{s} P(b_j) = 1 \qquad \text{...... (3.28)}$$

(i) In the problem given, s = 3

$$\therefore \quad \sum_{j=1}^{3} P(b_j) = P(b_1) + P(b_2) + P(b_3) = 1$$

$$\therefore \quad \frac{1}{3} + \frac{14}{36} + P(b_3) = 1$$

$$\therefore \quad P(b_3) = 1 - \frac{1}{3} - \frac{14}{36}$$

$$\therefore \quad \mathbf{P(b_3) = \frac{5}{18}}$$

From property - 1 of JPM.

$$P(b_1) = P(a_1, b_1) + P(a_2, b_1) + P(a_3, b_1) \rightarrow 1^{st} \text{ column}$$

$$\therefore \quad \frac{1}{3} = \frac{1}{12} + \frac{5}{36} + P(a_3, b_1)$$

$$\therefore \quad P(a_3, b_1) = \frac{1}{3} - \frac{1}{12} - \frac{5}{36}$$

$$\therefore \quad \mathbf{P(a_3, b_1) = \frac{1}{9}}$$

Again from property - 1 of JPM,

$$P(b_2) = P(a_1, b_2) + P(a_2, b_2) + P(a_3, b_2) \rightarrow 2^{nd} \text{ column}$$

$$\frac{14}{36} = P(a_1, b_2) + \frac{1}{9} + \frac{1}{6}$$

$$\therefore \quad P(a_1, b_2) = \frac{14}{36} - \frac{1}{9} - \frac{1}{6}$$

$$\therefore \quad \mathbf{P(a_1, b_2) = \frac{1}{9}}$$

$$\text{and} \quad P(b_3) = P(a_1, b_3) + P(a_2, b_3) + P(a_3, b_3) \rightarrow 3^{rd} \text{ column}$$

$$\frac{5}{18} = \frac{5}{36} + \frac{5}{36} + P(a_3, b_3)$$

$$\therefore \quad P(a_3, b_3) = \frac{5}{18} - \frac{5}{36} - \frac{5}{36}$$

$$\therefore \quad \mathbf{P(a_3, b_3) = 0}$$

(ii) From definition of conditional probability (Refer equation (0.16)) and joint probability (equation (3.14))

$$P(b_3/a_1) = \frac{P(a_1 \cap b_3)}{P(a_1)} = \frac{P(a_1, b_3)}{P(a_1)} \qquad \text{...... (3.29)}$$

From property - 2 of JPM,

$$P(a_1) = P(a_1, b_1) + P(a_1, b_2) + P(a_1, b_3) \rightarrow 1^{st} \text{ row}$$

$$= \frac{1}{12} + \frac{1}{9} + \frac{5}{36} = \frac{1}{3}$$

$$\text{and} \quad P(a_1, b_3) = \frac{5}{36}$$

Substituting in equation (3.29), we get

$$P(b_3/a_1) = \frac{(5/36)}{(1/3)}$$

$$\text{or} \quad \mathbf{P(b_3/a_1) = \frac{5}{12}}$$

Again from equation (3.14),

$$P(a_1/b_3) = \frac{P(a_1,b_3)}{P(b_3)} = \frac{(5/36)}{(5/18)}$$

$$\therefore \quad \mathbf{P(a_1/b_3) = \frac{1}{2}}$$

(iii) For a_1 and b_1 to be statistically independent, we must have (refer equation (0.25))

$$P(a_1 \cap b_1) = P(a_1, b_1) = P(a_1)\,P(b_1)$$

$$\text{Consider } P(a_1)\,P(b_1) = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{9}$$

$$\text{and } \quad P(a_1, b_1) = \frac{1}{12}$$

$$\therefore \quad P(a_1, b_1) \neq P(a_1)\,P(b_1)$$

$\therefore$ a_1 and b_1 are not statistically independent. It is because, we are going to receive b_1 by transmitting a_1 with some probability and hence b_1 is dependent on a_1.

3.4 ENTROPY FUNCTIONS AND EQUIVOCATION

1. PRIORI ENTROPY :

The entropy of the input symbols a_1, a_2, ..., a_r before their transmission is defined as ***"priori entropy"*** denoted by H(A) given by

$$H(A) = \sum_{i=1}^{r} P(a_i) \log \frac{1}{P(a_i)} \text{ bits/message-symbol} \qquad \text{...... (3.30)}$$

2. POSTERIORI (CONDITIONAL) ENTROPY :

The entropy of the input symbols a_1, a_2, ..., a_r, after transmission and reception of a particular output symbol "b_j" is defined as ***"posteriori or conditional entropy"*** denoted by $H(A/b_j)$ given by

$$H(A/b_j) = \sum_{i=1}^{r} P(a_i/b_j) \log \frac{1}{P(a_i/b_j)} \text{ bits/message-symbol} \qquad \text{...... (3.31)}$$

Note : Since j varies from 1 to s, we have 's' number of conditional entropies.

3. EQUIVOCATION :

Equivocation is defined as the average value of all the conditional entropies specified in equation (3.31) when "j" is varied from 1 to s. It is denoted by H(A/B) given by

$$H(A/B) = E\,[H(A/b_j)]$$

$$= \sum_{j=1}^{s} P(b_j)\, H(A/b_j)$$

$$= \sum_{j=1}^{s} P(b_j) \sum_{i=1}^{r} P(a_i/b_j) \log \frac{1}{P(a_i/b_j)} \text{ using equation (3.31)}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(b_j)\, P(a_i/b_j) \log \frac{1}{P(a_i/b_j)}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j) \log \frac{1}{P(a_i/b_j)} \text{ by using equation (3.14)}$$

$$\therefore \quad H(A/B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j) \log \frac{1}{P(a_i/b_j)} \text{ bits/message-symbol} \quad \text{...... (3.32)}$$

Equivocation specifies the average amount of information needed to specify an input symbol provided we are allowed to make an observation of the output produced by that input symbol. Therefore, equivocation is a measure of ***"uncertainty"*** when symbols are transmitted over the channel and hence represents the amount of information lost due to noise (due to intersymbol conversions) with respect to any of the output symbols.

Interchanging A and B in equation (3.32), we get

$$H(B/A) = \sum_{j=1}^{s}\sum_{i=1}^{r} P(b_j,a_i) \log \frac{1}{P(b_j/a_i)}$$

But $P(b_j, a_i) = P(a_i, b_j)$

$$\therefore \quad H(B/A) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j) \log \frac{1}{P(b_j/a_i)} \text{ bits/message-symbol} \quad \text{...... (3.33)}$$

H(B/A) is also called equivocation which is a measure of information about the receiver.

3.5 MUTUAL INFORMATION

From equation (3.30), we have

$$H(A) = \sum_{i=1}^{r} P(a_i) \log \frac{1}{P(a_i)} \text{ bits/message-symbol}$$

and from equation (3.32), we have

$$H(A/B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j) \log \frac{1}{P(a_i / b_j)} \text{ bits/message-symbol}$$

From equation (3.30), we observe that on an average, we would require H(A) bits of information to specify one input symbol at the transmitter. If we are allowed to observe the output symbol produced by that input symbol, we would then require only H(A/B) bits of information to specify that input symbol. Therefore, on an average, observation of a particular output symbol provides us with [H(A) – H(A/B)] bits of information. This difference is defined as ***"Mutual Information"*** or ***"Transinformation"*** of the channel denoted by I(A, B).

$$\therefore \quad I(A, B) = H(A) - H(A/B) \qquad \text{...... (3.34)}$$

Mutual Information can also be interpreted as follows:

When an average information of H(A) is transmitted over the channel, an average amount of information equal to equivocation H(A/B) is lost in the channel due to intersymbol conversion which is due to noise. The balance of information received at the receiver with respect to an observed output symbol is the mutual information given by equation (3.34).

Substituting equations (3.30) and (3.32) into (3.34), we can get an expression for I(A, B) in terms of probabilities.

$$\therefore \quad I(A, B) = \sum_{i=1}^{r} P(a_i)\log\frac{1}{P(a_i)} - \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\log\frac{1}{P(a_i/b_j)} \qquad \text{...... (3.35)}$$

From the equation (3.10), we have

$$\sum_{j=1}^{s} P(b_j/a_i) = 1$$

Multiplying 1st term of equation (3.35) by 1 and substituting the above expression in place of 1, we get

$$I(A, B) = \sum_{i=1}^{r} P(a_i)\left[\sum_{j=1}^{s} P(b_j/a_i)\right]\log\frac{1}{P(a_i)} - \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\log\frac{1}{P(a_i/b_j)}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i)\, P(b_j/a_i)\log\frac{1}{P(a_i)} - \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\log\frac{1}{P(a_i/b_j)}$$

From equation (3.14), joint probability $P(a_i, b_j) = P(a_i)\, P(b_j/a_i)$

$$\therefore \quad I(A, B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\log\frac{1}{P(a_i)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\log P(a_i/b_j)$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\left[\log\frac{1}{P(a_i)} + \log P(a_i/b_j)\right]$$

Again from equation (3.14), $P(a_i/b_j) = \dfrac{P(a_i, b_j)}{P(b_j)}$

$$\therefore \quad I(A, B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\left[\log\frac{1}{P(a_i)} + \log\frac{P(a_i, b_j)}{P(b_j)}\right]$$

$$\therefore \quad I(A, B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) \log \frac{P(a_i, b_j)}{P(a_i)\, P(b_j)} \text{ bits/message symbol} \quad \text{...... (3.36)}$$

Interchanging A and B in equation (3.36), we get

$$I(B, A) = \sum_{j=1}^{s}\sum_{i=1}^{r} P(b_j, a_i) \log \frac{P(b_j, a_i)}{P(b_j) P(a_i)}$$

$$\text{or} \quad I(B, A) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) \log \frac{P(a_i, b_j)}{P(a_i)\, P(b_j)} \text{ bits/message-symbol} \quad \text{..... (3.37)}$$

Equations (3.36) and (3.37) are identical

$$\therefore \quad I(B, A) = I(A, B) \quad \text{...... (3.38)}$$

Thus, we find that the mutual information is symmetrical with respect to its arguments and hence the name mutual.

From equation (3.34), we have

$$I(A, B) = H(A) - H(A/B)$$

Interchanging A and B, we get

$$I(B, A) = H(B) - H(B/A) \quad \text{...... (3.39)}$$

where H(B) is called the ***"Entropy of the output symbols"*** given by

$$H(B) = \sum_{j=1}^{s} P(b_j) \log \frac{1}{P(b_j)} \text{ bits/message-symbol} \quad \text{...... (3.40)}$$

Substituting equations (3.34) and (3.39) in (3.38), we get

$$H(B) - H(B/A) = I(B, A) = I(A, B) = H(A) - H(A/B)$$

$$\therefore \quad H(B) - H(B/A) = H(A) - H(A/B)$$

$$\therefore \quad H(B) + H(A/B) = H(A) + H(B/A) \quad \text{...... (3.41)}$$

The L.H.S. of equation (3.41) is

$$H(B) + H(A/B) = H(B, A) \quad \text{...... (3.42)}$$

and the R.H.S. of equation (3.41) is

$$H(A) + H(B/A) = H(A, B) \quad \text{...... (3.43)}$$

$\therefore \quad H(A, B) = H(B, A)$ is defined as the ***"Joint Entropy"***.

The expression for Joint entropy can be obtained by using equations (3.30) and (3.33) in equation (3.43).

$$\therefore \quad H(A, B) = \sum_{i=1}^{r} P(a_i) \log \frac{1}{P(a_i)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) \log \frac{1}{P(b_j / a_i)}$$

$$\therefore \quad H(A,B) = \sum_{i=1}^{r} P(a_i)\left[\sum_{j=1}^{s} P(b_j/a_i)\right]\log\frac{1}{P(a_i)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(b_j/a_i)}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i)P(b_j/a_i)\log\frac{1}{P(a_i)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(b_j/a_i)}$$

From equation (3.14), $P(a_i)\,P(b_j/a_i) = P(a_i, b_j)$ = Joint probability

$$\therefore \quad H(A,B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(a_i)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(b_j/a_i)}$$

$$\therefore \quad H(A,B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\left[\log\frac{1}{P(a_i)} + \log\frac{1}{P(b_j/a_i)}\right]$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(a_i)P(b_j/a_i)}$$

$$\therefore \quad H(A,B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i,b_j)\log\frac{1}{P(a_i,b_j)} \text{ bits/message-symbol} \quad \ldots\ldots (3.44)$$

From equation (3.39),

$$H(B/A) = H(B) - I(B, A)$$

$$\therefore \quad H(B/A) = H(B) - I(A, B) \quad \ldots\ldots (3.45)$$

Using (3.45) in equation (3.43), we get

$$H(A) + H(B) - I(A, B) = H(A, B)$$

$$\text{or} \quad I(A, B) = H(A) + H(B) - H(A, B) \quad \ldots\ldots (3.46)$$

All the entropy relations given above can be represented graphically shown in figure 3.4

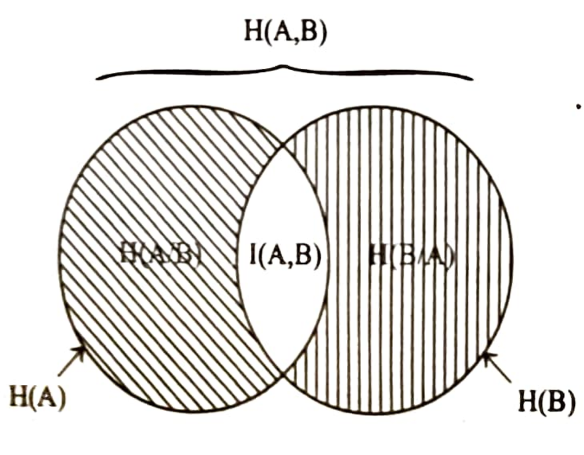

Fig. 3.4 : Illustrating entropy relations

In figure 3.4, the entropy H(A) is represented by a circle on the left side and entropy H(B) by a circle on the right side. The overlap between the two circles is the mutual information I(A, B). The hatched portion on the left circle is the equivocation H(A/B) and that on the right circle is H(B/A). These hatched portions satisfy the equations (3.34) and (3.39).

The joint entropy H(A, B) is the sum of H(A) and H(B) except for the fact that the overlap is added twice so that H(A, B) = H(A) + H(B) – I(A, B) which is same as equation (3.46).

3.6 PROPERTIES OF MUTUAL INFORMATION

In the previous section (3.5), we have obtained interrelationships between various entropies and also mutual information. Based on these interrelationships, we can now list the properties of mutual information as below:

Property - 1 :

The mutual information of a channel is symmetric.

$$\text{i.e.,} \quad I(A, B) = I(B, A)$$

Property - 2 :

The mutual information is always non-negative.

$$\text{i.e.,} \quad I(A, B) \geq 0 \qquad \text{......} (3.47)$$

Proof:

From equation (3.36), we have

$$I(A, B) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) \log_2 \frac{P(a_i, b_j)}{P(a_i)P(b_j)}$$

$$\therefore \quad I(A, B) = \log e \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\, ln \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \qquad \text{......} (3.48)$$

But from equation (1.14) we have

$$ln \frac{1}{x} \geq 1 - x$$

$$\text{with} \quad x = \frac{P(a_i)P(b_j)}{P(a_i, b_j)}, \text{ we get}$$

$$ln \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \geq \left[1 - \frac{P(a_i)P(b_j)}{P(a_i, b_j)}\right] \qquad \text{......} (3.49)$$

Multiplying both sides of equation (3.49) by $P(a_i, b_j)$ and taking double summation, we get

$$\sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\, ln \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \geq \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)\left[1 - \frac{P(a_i)P(b_j)}{P(a_i, b_j)}\right]$$

Multiplying both sides by log e, we get

$$\log e \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) ln \frac{P(a_i, b_j)}{P(a_i)P(b_j)} \geq \log e \left[\sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) - \sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i)P(b_j)\right]$$

The L.H.S. of the above equation is I(A, B) from equation (3.48)

$$\therefore \quad I(A, B) \geq \log e \left[\sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j) - \sum_{i=1}^{r} P(a_i) \sum_{j=1}^{s} P(b_j)\right]$$

We have $\sum_{i=1}^{r}\sum_{j=1}^{s} P(a_i, b_j)$ = Sum of all the elements of JPM

= 1 from property-3 of JPM

$\sum_{i=1}^{r} P(a_i)$ = 1 from equation (3.11)

and $\sum_{j=1}^{s} P(b_j)$ = 1 from equation (3.28)

$$\therefore \quad I(A, B) \geq \log e\,[1 - (1)(1)]$$

$$\therefore \quad I(A, B) \geq 0 \rightarrow \text{proved.}$$

Property - 3 :

The mutual information of a channel may be expressed in terms of the entropy of the channel output as,

$$I(A, B) = H(B) - H(B/A)$$

Property - 4 :

The mutual information is related to the joint entropy of the channel by

$$I(A, B) = H(A) + H(B) - H(A, B)$$

Example 3.3 : Show that H (X, Y) = H (X/Y) + H (Y).

[VI Sem EC/TE, August 2001, Q.2 (c)], [VI Sem EC/TE, February 2002, Q.3 (a)]

Proof

Replacing A by X and B by Y in equation (3.32), we get

$$H(X/Y) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j) \log \frac{1}{P(x_i/y_j)} \text{ bits/message symbol} \quad .... (3.50)$$

From equation (3.40) with B replaced by Y, we have

$$H(Y) = \sum_{j=1}^{s} P(y_j) \log \frac{1}{P(y_j)} \text{ bits/message-symbol} \quad ...... (3.51)$$

Again from equation (3.44) with A replaced by X and B by Y, we have

$$H(X, Y) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j) \log \frac{1}{P(x_i, y_j)}$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j) \log \frac{1}{P(x_i/y_j)\, P(y_j)}$$

[Since $P(x_i, y_j) = P(x_i/y_j)\, P(y_j)$ from equation (3.14)]

$$= \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j) \log \frac{1}{P(x_i/y_j)} + \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j) \log \frac{1}{P(y_j)}$$

$$= H(X/Y) + \sum_{j=1}^{s}\left[\sum_{i=1}^{r} P(x_i, y_j)\right] \log \frac{1}{P(y_j)} \quad \text{by using equation (3.50)}$$

$$= H(X/Y) + \sum_{j=1}^{s} P(y_j) \log \frac{1}{P(y_j)}$$

[by using equation (3.21) with a_i replaced by x_i and b_j replaced by y_j]

By using equation (3.51), we get

$\therefore \quad H(X, Y) = H(X/Y) + H(Y) \rightarrow$ proved

Example 3.4 : A transmitter has an alphabet consisting of 5 letters $\{a_1, a_2, a_3, a_4, a_5\}$ and the receiver has an alphabet of four letters $\{b_1, b_2, b_3, b_4\}$. The joint probabilities of the system are given below.

$$P(A, B) = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{array} \begin{array}{c} \begin{array}{cccc} b_1 & b_2 & b_3 & b_4 \end{array} \\ \begin{bmatrix} 0.25 & 0 & 0 & 0 \\ 0.10 & 0.30 & 0 & 0 \\ 0 & 0.05 & 0.10 & 0 \\ 0 & 0 & 0.05 & 0.1 \\ 0 & 0 & 0.05 & 0 \end{bmatrix} \end{array}$$

Compute different entropies of this channel.

Solution

From property-1 of JPM, addition of elements of JPM columnwise results in probability of output symbols.

$\therefore P(b_1) = 0.35, P(b_2) = 0.35, P(b_3) = 0.2$, and $P(b_4) = 0.1$

From property-2 of JPM, addition of elements of JPM rowwise results in probability of input symbols.

$\therefore P(a_1) = 0.25, P(a_2) = 0.4, P(a_3) = 0.15, P(a_4) = 0.15, P(a_5) = 0.05$

From equation (3.30), the entropy of input symbols is given by

$$H(A) = \sum_{i=1}^{5} P(a_i) \log \frac{1}{P(a_i)}$$

$$= 0.25 \log \frac{1}{0.25} + 0.4 \log \frac{1}{0.4} + 0.15 \log \frac{1}{0.15} + 0.15 \log \frac{1}{0.15}$$

$$+ 0.05 \log \frac{1}{0.05}$$

$\therefore$ **H(A) = 2.066 bits/message-symbol**

From equation (3.40), the entropy of the output symbols is given by,

$$H(B) = \sum_{j=1}^{4} P(b_j) \log \frac{1}{P(b_j)}$$

$$= 0.35 \log \frac{1}{0.35} + 0.35 \log \frac{1}{0.35} + 0.2 \log \frac{1}{0.2} + 0.1 \log \frac{1}{0.1}$$

$\therefore$ **H(B) = 1.857 bits/message-symbol**

From equation (3.44), the joint entropy is given by

$$H(A, B) = \sum_{i=1}^{5} \sum_{j=1}^{4} P(a_i, b_j) \log \frac{1}{P(a_i, b_j)}$$

In the above expansion, there are total 20 terms. But the JPM contains many zero elements. When we consider the remaining elements, we get

$$H(A, B) = 0.25 \log \frac{1}{0.25} + 0.1 \log \frac{1}{0.1} + 0.3 \log \frac{1}{0.3} + 0.05 \log \frac{1}{0.05}$$

$$+ 0.1 \log \frac{1}{0.1} + 0.05 \log \frac{1}{0.05} + 0.05 \log \frac{1}{0.05} + 0.1 \log \frac{1}{0.1}$$

$\therefore$ **H(A, B) = 2.666 bits/message-symbol**

The equivocation H(B/A) is found using equation (3.43) as

$$H(B/A) = H(A, B) - H(A)$$

$$= 2.666 - 2.066$$

$\therefore$ **H(B/A) = 0.6 bits/message-symbol**

The other equivocation H(A/B) is found using equation (3.42) as

$$H(A/B) = H(A, B) - H(B)$$

$$= 2.666 - 1.857$$

$\therefore$ **H(A/B) = 0.809 bits/message-symbol**

The mutual information I(A, B) is found from equation (3.34) as

$$I(A, B) = H(A) - H(A/B)$$
$$= 2.066 - 0.809$$
$$\therefore \quad \mathbf{I(A, B) = 1.257 \text{ bits/message-symbol}}$$

Equation (3.39) can also be used to find I(A, B) as

$$I(A, B) = H(B) - H(B/A)$$
$$= 1.857 - 0.6$$
$$\therefore \quad \mathbf{I(A, B) = 1.257 \text{ bits/message-symbol}}$$

Example 3.5 : A transmitter transmits five symbols with probabilities 0.2, 0.3, 0.2, 0.1 and 0.2. Given the channel matrix P(B/A), calculate (i) H(B) (ii) H(A, B).

$$P(B/A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Solution

From equation (3.14), we have

$$P(a_i, b_j) = P(a_i)\, P(b_j/a_i)$$

The JPM may now be constructed by multiplying 1st row elements by $P(a_1) = 0.2 = \frac{1}{5}$, 2nd row by $P(a_2) = 0.3 = \frac{3}{10}$, 3rd row by $P(a_3) = 0.2 = \frac{1}{5}$, 4th row by $P(a_4) = 0.1 = \frac{1}{10}$ and 5th row by $P(a_5) = 0.2 = \frac{1}{5}$.

$$P(A, B) = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{array} \begin{array}{c} \begin{array}{cccc} b_1 & b_2 & b_3 & b_4 \end{array} \\ \begin{bmatrix} 1/5 & 0 & 0 & 0 \\ 3/40 & 9/40 & 0 & 0 \\ 0 & 1/15 & 2/15 & 0 \\ 0 & 0 & 1/30 & 1/15 \\ 0 & 0 & 1/5 & 0 \end{bmatrix} \end{array}$$

Adding the elements of each column, we get

$$P(b_1) = \frac{1}{5} + \frac{3}{40} = \frac{11}{40}$$

$$P(b_2) = \frac{9}{40} + \frac{1}{15} = \frac{7}{24}$$

$$P(b_3) = \frac{2}{15} + \frac{1}{30} + \frac{1}{5} = \frac{11}{30}$$

$$P(b_4) = \frac{1}{15}$$

(i) From equation (3.40), H(B) is found from

$$H(B) = \sum_{j=1}^{4} P(b_j)\log\frac{1}{P(b_j)}$$

$$= \frac{11}{40}\log\frac{40}{11} + \frac{7}{24}\log\frac{24}{7} + \frac{11}{30}\log\frac{30}{11} + \frac{1}{15}\log 15$$

$$\therefore \quad \mathbf{H(B) = 1.822\ bits/message\text{-}symbol}$$

(ii) From equation (3.11), H(A, B) is found from

$$H(A, B) = \sum_{i=1}^{5}\sum_{j=1}^{4} P(a_i, b_j)\log\frac{1}{P(a_i, b_j)}$$

$$= \frac{1}{5}\log 5 + \frac{3}{40}\log\frac{40}{3} + \frac{9}{40}\log\frac{40}{9} + \frac{1}{15}\log 15 + \frac{2}{15}\log\frac{15}{2}$$

$$+ \frac{1}{30}\log 30 + \frac{1}{5}\log 5 + \frac{1}{15}\log 15$$

$$\mathbf{H(A, B) = 2.7653\ bits/message\text{-}symbol}$$

Example 3.6 : For the JPM given below, compute individually H(X), H(Y), H(X, Y), H(X/Y), H(Y/X) and I(X,Y). Verify the relationship among these entropies.

$$P(X, Y) = \begin{bmatrix} 0.05 & 0 & 0.20 & 0.05 \\ 0 & 0.10 & 0.10 & 0 \\ 0 & 0 & 0.20 & 0.10 \\ 0.05 & 0.05 & 0 & 0.10 \end{bmatrix}$$

[VI sem EC/TE, July/August 2003, Q.2 (b)]

Solution

The given JPM can be rewritten with input and output symbols as below:

$$P(X, Y) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array}\begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{bmatrix} 0.05 & 0 & 0.20 & 0.05 \\ 0 & 0.10 & 0.10 & 0 \\ 0 & 0 & 0.20 & 0.10 \\ 0.05 & 0.05 & 0 & 0.10 \end{bmatrix} \end{array}$$

From property-1 of JPM, addition of elements of JPM columnwise results in probability of output symbols.

$$\therefore \quad P(y_1) = 0.05 + 0.05 = 0.10$$

$$P(y_2) = 0.10 + 0.05 = 0.15$$

$$P(y_3) = 0.2 + 0.1 + 0.2 = 0.50$$

$$P(y_4) = 0.05 + 0.1 + 0.1 = 0.25$$

From property-2 of JPM, addition of elements of JPM rowwise results in probability of input symbols.

$$\therefore \quad P(x_1) = 0.05 + 0.2 + 0.05 = 0.30$$
$$P(x_2) = 0.10 + 0.10 = 0.20$$
$$P(x_3) = 0.20 + 0.10 = 0.30$$
$$P(x_4) = 0.05 + 0.05 + 0.1 = 0.20$$

From equation (3.30), the entropy of input symbols is given by

$$H(X) = \sum_{i=1}^{4} P(x_i) \log \frac{1}{P(x_i)}$$

$$= \left(0.3 \log \frac{1}{0.3}\right)(2) + \left(0.2 \log \frac{1}{0.2}\right)(2)$$

$$\therefore \quad \mathbf{H(X) = 1.971 \text{ bits/message-symbol}}$$

From equation (3.40), the entropy of the output symbols is given by

$$H(Y) = \sum_{j=1}^{4} P(y_j) \log \frac{1}{P(y_j)}$$

$$= 0.1 \log \frac{1}{0.1} + 0.15 \log \frac{1}{0.15} + 0.5 \log \frac{1}{0.5} + 0.25 \log \frac{1}{0.25}$$

$$\therefore \quad \mathbf{H(Y) = 1.743 \text{ bits/message-symbol}}$$

From equation (3.44), the joint entropy is given by

$$H(X, Y) = \sum_{i=1}^{4}\sum_{j=1}^{4} P(x_i, y_j) \log \frac{1}{P(x_i, y_j)}$$

$$= \left(0.05 \log \frac{1}{0.05}\right)(4) + \left(0.10 \log \frac{1}{0.10}\right)(4) + \left(0.2 \log \frac{1}{0.2}\right)(2)$$

$$\therefore \quad \mathbf{H(X, Y) = 3.122 \text{ bits/message-symbol}}$$

The equivocation H(X/Y) is given by equation (3.32) [with A replaced by X, B by Y, a_i by x_i and b_j by y_j] as,

$$H(X/Y) = \sum_{i=1}^{4}\sum_{j=1}^{4} P(x_i, y_j) \log \frac{1}{P(x_i/y_j)} \text{ bits/message symbol.}$$

Using the relationship $P(x_i/y_j) = \dfrac{P(x_i, y_j)}{P(y_j)}$, the matrix P(X/Y) is constructed as below:

$$P(X/Y) = \begin{bmatrix} \frac{0.05}{0.10} & 0 & \frac{0.2}{0.5} & \frac{0.05}{0.25} \\ 0 & \frac{0.10}{0.15} & \frac{0.1}{0.5} & 0 \\ 0 & 0 & \frac{0.2}{0.5} & \frac{0.10}{0.25} \\ \frac{0.05}{0.10} & \frac{0.05}{0.15} & 0 & \frac{0.10}{0.25} \end{bmatrix}$$

$$\therefore \quad P(X/Y) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{bmatrix} 1/2 & 0 & 2/5 & 1/5 \\ 0 & 2/3 & 1/5 & 0 \\ 0 & 0 & 2/5 & 2/5 \\ 1/2 & 1/3 & 0 & 2/5 \end{bmatrix} \end{array}$$

$$\therefore \quad H(X/Y) = 0.05 \log 2 + 0.05 \log 2 + 0.1 \log 3/2 + 0.05 \log 3 + 0.2 \log 5/2$$
$$+ 0.1 \log 5 + 0.2 \log 5/2 + 0.05 \log 5 + 0.1 \log 5/2 + 0.1 \log 5/2$$

$\therefore$ **$H(X/Y)$ = 1.379 bits/message symbol**

The equivocation $H(Y/X)$ is given by equation (3.33) [with A replaced by X, B by Y, a_i by x_i and b_j by y_j] as,

$$H(Y/X) = \sum_{i=1}^{4}\sum_{j=1}^{4} P(x_i, y_j) \log \frac{1}{P(y_j/x_i)} \text{ bits/message-symbol.}$$

Using the relationship $P(y_j/x_i) = \dfrac{P(x_i, y_j)}{P(x_i)}$, the channel matrix $P(Y/X)$ is constructed as below:

$$P(Y/X) = \begin{bmatrix} \frac{0.05}{0.30} & 0 & \frac{0.20}{0.30} & \frac{0.05}{0.30} \\ 0 & \frac{0.10}{0.20} & \frac{0.10}{0.20} & 0 \\ 0 & 0 & \frac{0.20}{0.30} & \frac{0.10}{0.30} \\ \frac{0.05}{0.20} & \frac{0.05}{0.20} & 0 & \frac{0.10}{0.20} \end{bmatrix}$$

$$\therefore \quad P(Y/X) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{bmatrix} 1/6 & 0 & 2/3 & 1/6 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 2/3 & 1/3 \\ 1/4 & 1/4 & 0 & 1/2 \end{bmatrix} \end{array}$$

$$\therefore \quad H(Y/X) = 0.05 \log 6 + 0.2 \log 3/2 + 0.05 \log 6 + 0.1 \log 2 + 0.1 \log 2$$
$$+ 0.2 \log 3/2 + 0.1 \log 3 + 0.05 \log 4 + 0.05 \log 4 + 0.1 \log 2$$

$\therefore$ **$H(Y/X)$ = 1.151 bits/message-symbol**

Verification : From equation (3.43), we have,

$$H(Y/X) = H(X, Y) - H(X)$$
$$= 3.122 - 1.971$$

$\therefore$ $$\mathbf{H(Y/X) = 1.151 \text{ bits/message-symbol as before}}$$

From equation (3.42), we have

$$H(X/Y) = H(X, Y) - H(Y)$$
$$= 3.122 - 1.743$$

$\therefore$ $$\mathbf{H(Y/X) = 1.379 \text{ bits/message-symbol as before}}$$

From equation (3.34), the mutual information I(X, Y) is given by

$$I(X, Y) = H(X) - H(X/Y)$$
$$= 1.971 - 1.379$$

$\therefore$ $$\mathbf{I(X, Y) = 0.592 \text{ bits/message-symbol}}$$

Verification : The mutual information I(X, Y) can also be verified using equation (3.39) as given below.

$$I(X, Y) = H(Y) - H(Y/X)$$
$$= 1.743 - 1.151$$

$\therefore$ $$\mathbf{I(X, Y) = 0.592 \text{ bits/message-symbol}} \quad \text{– verified.}$$

3.7 RATE OF INFORMATION TRANSMISSION OVER A DISCRETE CHANNEL

From equation (3.30), we have the entropy of the input symbols given by

$$H(A) = \sum_{i=1}^{r} P(a_i) \log \frac{1}{P(a_i)} \text{ bits/message-symbol}$$

When we consider a discrete memoryless channel accepting symbols at the rate of ***"r_s message-symbols/sec"***, then the average rate at which information is going into the channel is given by

$$R_{in} = H(A)\, r_s \text{ bits/sec} \qquad \text{...... (3.52)}$$

At the receiver, we have seen in section (3.5) that, in general, it is not possible to reconstruct the input symbol sequence with ***"certainty"*** by operating on the receiving sequence. This is due to errors introduced when the signals pass through the channel. These errors are in fact, introduced due to the noise present in the channel. Thus some amount of information is lost in the channel due to noise. This information which is lost in the channel has been called "equivocation H(A/B)" in section (3.5). Hence the net amount of information which is the mutual information I(A, B) is given by equation (3.34) as

$$I(A, B) = H(A) - H(A/B) \text{ bits/message-symbol}$$

Then, the average rate of information transmission R_t is then given by

$$R_t \triangleq I(A, B)\, r_s \text{ bits/sec}$$

or $$R_t = [H(A) - H(A/B)]\, r_s \text{ bits/sec} \quad \text{...... (3.53)}$$

Since, $I(A, B) = I(B, A) = H(B) - H(B/A)$, the average rate of information transmission R_t can also be expressed as ...... (3.54)

$$R_t = [H(B) - H(B/A)]\, r_s \text{ bits/sec} \quad \text{...... (3.55)}$$

This definition of equation (3.54) takes into account the case when the channel is so noisy that the output may become statistically independent of the input. When 'B' and 'A' become independent, then $H(B/A) = H(B)$ and hence all the information going into the channel is lost and no information is transmitted over the channel (Refer problem number 12 - unit 3).

3.8 CAPACITY OF A DISCRETE MEMORYLESS CHANNEL

The capacity of a discrete memoryless noisy channel is defined as the maximum possible rate of information transmission over the channel. The maximum rate of transmission occurs when the source is ***"matched"*** to the channel.

Therefore the ***"channel capacity C"*** is defined as

$$C = \text{Max } \{R_t\}$$

$$\therefore \quad C = \text{Max } [H(A) - H(A/B)]\, r_s \quad \text{...... (3.56)}$$

3.9 SHANNON'S THEOREM ON CHANNEL CAPACITY [SHANNON'S SECOND THEOREM]

From equation (3.54), we have rate of information transmission given as

$$R_t = [H(A) - H(A/B)]\, r_s \text{ bits/sec}$$

and the channel capacity from equation (3.56) as

$$C = \text{Max } [\{H(A) - H(A/B)\}\, r_s] \text{ bits/sec}$$

Shannon's theorem on channel capacity is stated in two ways.

Positive Statement :

Shannon's theorem on channel capacity states that "when the rate of information transmission $R_t \le C$, then there exists a coding technique which enables transmission over a channel with as small a probability of error as possible, even in the presence of noise in the channel".

This theorem indicates that for $R_t \le C$, transmission of information is achieved without errors, even in the presence of noise. This situation is analogous to amplitude modulation with "m" as the modulation index defined as the ratio of peak value of modulating voltage to the peak value of carrier voltage. As long as $m \le 1$, transmission of modulated signal is possible without errors. But when $m > 1$, transmission is possible but there will be errors introduced due to over-modulation and there will be loss of information.

Negative Statement :

If $R_t > C$, then reliable transmission of information is not possible without errors. Thus when $R_t > C$, then the errors cannot be controlled by any coding technique and the probability of error of receiving the correct message becomes close to unity.

CHANNEL EFFICIENCY AND REDUNDANCY

The ***"channel efficiency"*** denoted by η_{ch} is given by

$$\eta_{ch} = \frac{R_t}{C} \times 100\% \qquad ......(3.57)$$

Substituting for R_t and C from respective equations (3.54) and (3.56), we get

$$\eta_{ch} = \frac{[H(A)-H(A/B)]\, r_s}{\text{Max}\,[H(A)-H(A/B)]\, r_s} \times 100\%$$

$$\text{or} \quad \eta_{ch} = \frac{I(A,B)}{\text{Max}\,[I(A,B)]} \times 100\% \qquad ......(3.58)$$

The ***"Channel redundancy"*** denoted by $R_{\eta_{ch}}$ is given by

$$R_{\eta_{ch}} = 1 - \eta_{ch} \text{ where } \eta_{ch} \text{ is a fraction} \qquad ......(3.59)$$

3.10 SPECIAL CHANNELS

There are several special channels which are of great interest in the field of communication systems. Some of them are listed below. Let us find out the channel capacity of these channels.

(i) Symmetric/uniform channels

(ii) Binary Symmetric Channels (BSC)

(iii) Binary Erasure Channels (BEC)

(iv) Noiseless Channel

(v) Deterministic Channel

(vi) Cascaded Channel.

(i) Symmetric/Uniform Channel :

A channel is said to be symmetric or uniform channel, if the second and subsequent rows of the channel matrix contain the same elements as that of first row, but in a different order.

For example : Figures 3.5 (a) and (b) represent the channel diagrams and their corresponding channel matrices of two different symmetric or uniform channels.

$$P(Y/X) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/2 \\ 1/6 & 1/2 & 1/3 \end{bmatrix} \end{array}$$