UNIT- IV

Information Theory

The fonction of any communication system is to convey, from transmitter to receiver, a sequence of messages which are selected from a finite number of predetermined messages. within some specified time interval one of these messages is transmitted. during next time interval another of these messages is transmitted.

Information theory is related to the concepts of statistical properties of messages/sources, channels, noise interference etc., and is used for mathematical modeling and analysis of communication systems.

- Discrete messages have two sources.

(i) Memory Sources

produces output depends upon previous data.

(ii) Memoryless sources.

produces output depends upon present data, it doesn't dependent on previous data.

Amount of Information:-

Let us consider a communication system in which the allowable messages are $m_1, m_2 \cdots\cdots$, with probabilities of occurrences $P_1, P_2 \cdots\cdots$ $[P_1 + P_2 + \cdots = 1]$.

Let the transmitter select message m_k of probability P_k. Assume that the receiver has correctly identified the message.

The probability of an event and the amount of information are inversely proportional.

[The principle of improbability (which is one of the basic principles of the media world). "if a dog bites a man, it's no news, but if a man bits a dog, it's a news".

The probability of a dog biting a man is quite high, so this is not a news, i.e. very little amount of information is communicated by the message "a dog bits a man".

on the other hand, the probability of a man biting a dog is extremely small, so this becomes a news, i.e. quite an amount of information is communicated by the message "a man bites a dog".

Thus, we see that there should be some sort of inverse relationship between the probability of an event and the amount of information associated with it. The more the probability of an event, the less is the amount of information associated with it and vice versa.] → example for explanation.

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

$$I_K = f(1/P_K)$$

function 'f' can be replaced by logarithmic function.

$$\boxed{I_K = \log_2 (1/P_K) \text{ bits}}$$

$$\boxed{I_K = -\log_2 (P_K) \text{ bits}}$$

- If the base is 2, units are "bits".
- If the base is 10, units are "Decits or Hartely".
- If the base is 'e' units are "nat".

Properties of amount of information:-

(1) If $P_K = 1$, then $I_K = 0$

If $P_K = 1$ then message is already known by Receiver, there is really no need for transmission.

2) $0 < P_K < 1 \implies I_K \geq 0$.

As P_K decreases from 1 to 0, I_K increases monotonically, going from 0 to ∞. Thus, a greater amount of information has been conveyed when the receiver correctly identifies a less likely message.

(3) $I_{K,l} = I_K + I_l$.

When two independent messages m_K and m_l are correctly identified with probabilities P_K and P_l. The amount of information conveyed is the sum of the information associated with each of the messages individually.

$$I_K = \log_2 (1/P_K) \text{ bits}$$

$$I_l = \log_2 (1/P_l) \text{ bits.}$$

Since the messages are independent, the probability of composite message is $P_K P_l$.

The corresponding information content of messages m_K and m_l is

$$I_{K,l} = \log_2 1/P_K P_l = \log_2 (1/P_K) + \log_2 (1/P_l) = I_K + I_l.$$

$$\boxed{I_{K,l} = I_K + I_l}$$

(4) Take two messages m_K, m_l with probabilities P_K and P_l.

Let $P_K = 1/4$, $P_l = 3/4$.

$$I_K = -\log_2 1/4 = 2 \text{ bits.}$$

$$I_l = -\log_2 3/4 = 0.41 \text{ bits.}$$

then $\boxed{I_K > I_l \Rightarrow P_K < P_l}$

(5) If equal probability is there then $I_K = 1$.

Ex:- for coin $P(H) = 1/2$, $P(T) = 1/2$

then $I_H = -\log_2 1/2 = 1$ bits

$I_T = -\log_2 1/2 = 1$. bits.

Average Information (Entropy):- (H)

Suppose we have M different and independent messages m_1, m_2 ---- with probabilities of occurences P_1, P_2 --- during a long period of transmission a sequence of L message has been generated.

If L is very large, we may expect that in the L message sequence we transmitted P_1L messages of m_1, P_2L messages of m_2 etc. The total information in such a sequence is be.

$$I_{total} = P_1L \log_2 1/P_1 + P_2L \log 1/P_2 + ----$$

The average information per message interval, represented by the symbol H, will then be

$$H = \frac{I_{total}}{L} = P_1 \log_2 1/P_1 + P_2 \log_2 1/P_2 + ----$$

$$\boxed{H = \sum_{K=1}^{M} P_K \log 1/P_K \text{ bits/message}}$$

The average information is also referred by the term "Entropy".

Properties of Entropy:-

1) If we transmit 'm' messages out of 'm' messages one message probability has '1' and remaining all probabilities are zero.

If all the probabilities are zero except 1 which must be unity. for this condition the entropy is given by

$$H = 1 \log_2 1/1 + \left\{ \underset{P \to 0}{Lt} \sum_{K=1}^{M} P_K \log (1/P_K) \right\}$$

$$H(x) = 0.$$

(a) If we transmit two messages m_1 and m_2 with probabilities P_1 and P_2 then entropy is given by.

if $P_1 = P$ then $P_2 = 1-P$ $(\because P_1 + P_2 = 1)$

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

$$H = -p\log p - (1-p)\log(1-p).$$

$\frac{dH}{dp} = 0$ for maximum information.

$$\Rightarrow \frac{d}{dp}\left(-p\log_2 P - (1-p)\log_2(1-p)\right) = 0$$

$$\frac{d}{dp}\left(p\log P + (1-p)\log(1-p)\right) = 0.$$

$$p \cdot 1/p + \log_2 P + (1-p)\frac{1}{1-p}(-1) + \log_2(1-p)(-1) = 0$$

$$\cancel{1} + \log_2 P - \cancel{1} - \log_2(1-p) = 0$$

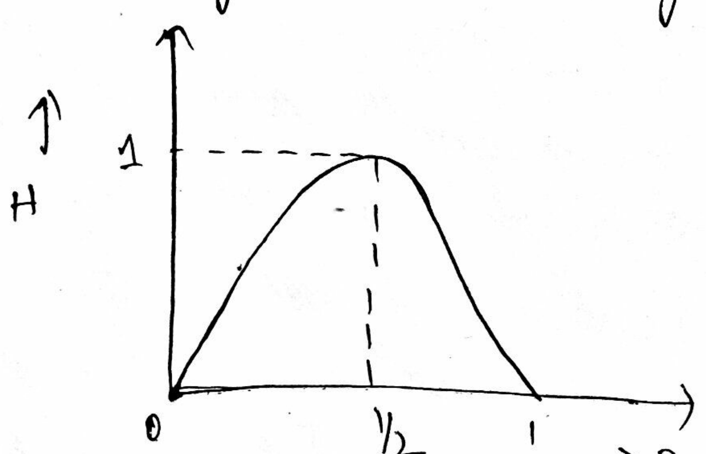
$$\log_2 P = \log_2(1-p)$$

$$P = 1-P$$

$$2p = 1$$

$$\boxed{P = 1/2}$$

when $P = 1/2$ we get maximum Average information.



(3) If messages are equiprobable

i.e $p = 1/M$.

then H is given by

$$H = -p\log_2 p - (1-p)\log(1-p)$$

$$= -\tfrac{1}{2}\log_2 \tfrac{1}{2} - (1-\tfrac{1}{2})\log(1-\tfrac{1}{2})$$

$$H = \tfrac{1}{2} + \tfrac{1}{2} = 1 \text{ bit/message.}$$

$H(X) = \log M$ if all the probabilities are equal so that $P(X) = 1/M$

4) $0 \le H(X) \le \log_2 M$.

- probability for average Information varies from 0 to $1/M$.
- probability for amount of information varies from 0 to 1.

Information rate:-

Information rate is defined as the number of bits transmitted per second. It is denoted by 'R'.

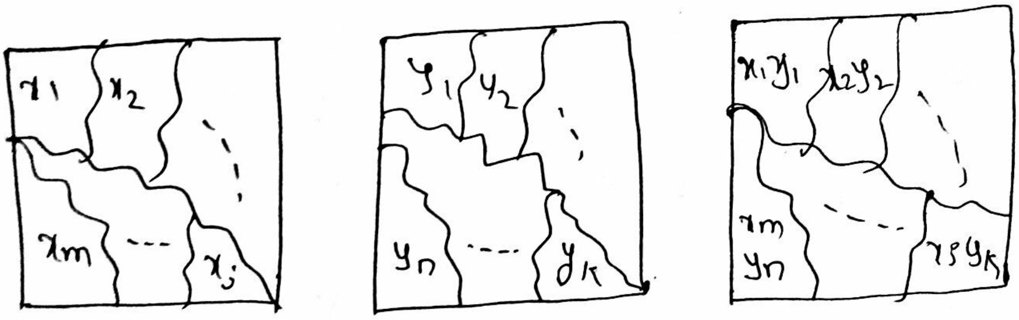
If transmitter transmits r messages/sec and H is bits/message then R is given by.

$$\boxed{R = rH \text{ bits/sec}}$$

Joint And Conditional Entropy:-

To study the behaviour of a communication system we must simultaneously study the behaviour of the transmitter and the receiver. This gives rises to the concept of a two-dimensional probability scheme.

Let there are two finite discrete sample spaces S_1 and S_2 and let their product space be $S = S_1 S_2$.



Sample space S_1 Sample space S_2 Sample space $S = S_1S_2$.

Let

$$[X] = [x_1, x_2 \; --- \; x_m]$$

$$[Y] = [y_1, y_2 \; ---- \; y_n].$$

each event x_j of S_1 may occur in conjunction with any event y_k of S_2

$$\therefore \; [XY] = \begin{bmatrix} x_1y_1 & x_1y_2 & ---- & x_1y_n \\ x_2y_1 & x_2y_2 & --- & x_2y_n \\ \vdots & & & \\ x_my_1 & x_my_2 & --- & x_my_n \end{bmatrix}$$

Here we have three sets of complete probability

- $P(X) = [P(x_j)]$
- $P(Y) = [P(y_k)]$
- $P(XY) = [P(x_j, y_k)]$.

$$P(x_j) = \sum_{k=1}^{n} P(x_j, y_k)$$

$$P(y_k) = \sum_{j=1}^{m} P(x_j, y_k).$$

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

$$H(Y) = -\sum_{k=1}^{n} P(y_k) \log_2 P(y_k).$$

$$\boxed{H(XY) = -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j, y_k) \log_2 P(x_j, y_k)}$$

$H(X)$ and $H(Y)$ are marginal entropies of X and Y and $H(XY)$ is the joint entropy of X and Y.

Conditional probability is given by

$$\boxed{P(x|y) = \frac{P(x,y)}{P(y)}}$$

Relation between joint entropy and conditional entropy is given by

$$H(XY) = H(X) + H(Y/X)$$
$$= H(Y) + H(X/Y).$$

Proof:-

$$\rightarrow H(XY) = -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j, y_k) \log_2 P(x_j, y_k).$$

$$= -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j, y_k) \log_2 \left[P(x_j/y_k) P(y_k)\right].$$

$$H(X\,Y) = -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\left[\log_2 p(x_j|y_k) + \log_2 p(y_k)\right]$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2 p(x_j|y_k) - \sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2 p(y_k)$$

$$= H(X|Y) - \sum_{k=1}^{n}\left[\sum_{j=1}^{m} p(x_j, y_k)\right]\log_2 p(y_k)$$

$$= H(X|Y) - \sum_{k=1}^{n} p(y_k)\log_2 p(y_k)$$

$$\boxed{H(X\,Y) = H(X|Y) + H(Y)}$$

$$\rightarrow \quad H[X\,Y] = -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2 p(x_j, y_k)$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2\left[p(y_k|x_j)\,p(x_j)\right]$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\left[\log_2 p(y_k|x_j) + \log_2 p(x_j)\right]$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2 p(y_k|x_j) - \sum_{j=1}^{m}\sum_{k=1}^{n} p(x_j, y_k)\log_2 p(x_j$$

$$= H(Y|X) - \sum_{j=1}^{m}\left[\sum_{k=1}^{n} p(x_j, y_k)\right]\log_2 p(x_j).$$

$$= H(Y|X) - \sum_{j=1}^{m} p(x_j)\log_2 p(x_j)$$

$$\boxed{H(X\,Y) = H(Y|X) + H(X)}$$

H(X) = Entropy of the transmitter

H(Y) = entropy of the Receiver

H(X Y) = average information per pair of the transmitted and received characters (or) average uncertainity of the communication system as a whole.

$H(X/Y)$ = A Received character y_k may be the result of transmission of one of the x_i's with a given probability.

$H(Y/X)$ = A transmitted character x_i may result in the reception of one of the y_k's with a given probability.

* $H(X)$ and $H(Y)$ give indications of the probabilistic nature of the transmitter and receiver respectively.

$H(X/Y)$ indicates how well one can recover the transmitted symbols, from the received symbols.

$H(Y/X)$ indicates how well one can recover the received symbols from the transmitted symbols.

Mutual Information:-

We are intrested in the transfer of information from a transmitter through a channel to a receiver.

prior to the reception of a message, the state of knowledge at the receiver about a transmitted symbol x_i is the probability that x_i would be selected for transmission. This is a priori probability $p(x_i)$.

After the reception and selection of the symbol y_k the state of knowledge concerning x_i is the conditional probability $p(x_i/y_k)$ which also known as a posteriori probability.

$\therefore$ before y_k is received, the uncertainity is

$$-\log_2 P(x_i)$$

After y_k is received, the uncertainity becomes

$$-\log_2 P(x_i|y_k).$$

" The information gained about x_i by the reception of y_k is the net reduction in its uncertainity, and is known as mutual information $I(x_i, y_k)$ ".

Thus,

$$I(x_i, y_k) = \text{initial uncertainity} - \text{final uncertainity}.$$

$$= -\log_2 P(x_i) - [-\log_2 P(x_i|y_k)]$$

$$= -\log_2 P(x_i) + \log_2 P(x_i|y_k)$$

$$\boxed{I(x_i, y_k) = \log_2 \left[\frac{P(x_i|y_k)}{P(x_i)}\right]}$$

$$I(x_i, y_k) = \log_2 \frac{P(x_i, y_k)}{P(y_k)P(x_i)}$$

$$= \log_2 \left(\frac{P(y_k|x_i)}{P(y_k)}\right)$$

$$\boxed{I(x_i, y_k) = I(y_k, x_i)}$$

from the above relation, we see that mutual information is symmetrical in x_i and y_k.

Self Information:-

Self information may be treated as a special case of mutual information when $y_K = x_j$.

$$\therefore I(x_j, y_k) = \log_2 \left[\frac{P(x_j/x_j)}{P(x_j)} \right]$$

$$= \log_2 \left[1/P(x_j) \right]$$

$$\boxed{I(x_j, y_K) = I(x_j)}$$

Entropy of mutual Information:-

The average of mutual information i.e the entropy corresponding to mutual information, is given by.

$$I[X;Y] = \overline{I(x_j; y_K)}$$

$$= \sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K)\, I(x_j; y_K).$$

$$= \sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K) \log \frac{P(x_j|y_K)}{P(x_j)}.$$

$$= \sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K) \left[\log P(x_j|y_K) - \log P(x_j) \right].$$

$$= -\sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K) \log P(x_j) + \sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K) \log P(x_j|y_K)$$

$$I[X;Y] = -\sum_{j=1}^{m} \left[\sum_{K=1}^{n} P(x_j, y_K) \right] \log P(x_j) - \left[-\sum_{j=1}^{m} \sum_{K=1}^{n} P(x_j, y_K) \log P(x_j|y_K) \right]$$

$$= -\sum_{j=1}^{m} P(x_j) \log P(x_j) - H(X/Y).$$

$$\boxed{I[X;Y] = H(X) - H(X|Y)}$$

$$\boxed{\begin{aligned} I[X;Y] &= H(Y) - H(Y/X) \\ &= H(X) + H(Y) - H(XY) \end{aligned}}$$

$I[X;Y]$ does not depend on the individual symbols x_i or y_K; it is a property of the whole communication system. On the other hand $I(x_i;y_K)$ depends on the individual symbols x_i or y_K. In practice, we are intrested in the whole communication system and not in individual symbols.

For this reason, although $I[X;Y]$ is an average mutual information, it is usually referred as mutual information.

Since $I[X;Y]$ indicates a measure of information transferred through the channel, it is also known as "transferred" information or "transinformation" of the channel.

=o=

Noise-free channel:-

Let us consider the communication channel shown in figure. It is known as a Noise free channel.

In such channels there is one to one correspondence between input & output i.e each input symbol is received

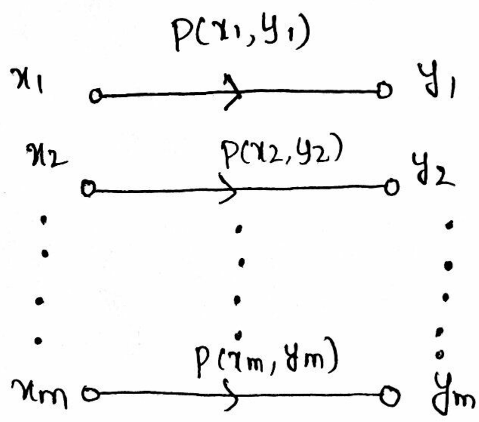


fig:- Noise free channel.

as one and the only one output symbol. Also, $n = m$. A discrete noise-free channel is shown in figure. The joint probability matrix $p(x,y)$ is of the diagonal form.

$$[P(X,Y)] = \begin{bmatrix} p(x_1,y_1) & 0 \dots & 0 & 0 \\ 0 & p(x_2,y_2) & 0 & 0 \\ --- & ------ & --- & --- \\ 0 & 0 \dots & 0 & p(x_m,y_m) \end{bmatrix}$$

and the channel probability matrices $[P(Y/X)]$ and $[P(X/Y)]$ are unity-diagonal matrices.

$$[P(Y/X)] = [P(X/Y)] = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

$$H[X,Y] = -\sum_{j=1}^{m}\sum_{k=1}^{m} p(x_j,y_k)\log p(x_j,y_k)$$

$$= -\sum_{j=1}^{m} p(x_j,y_j)\log p(x_j,y_j).$$

$$(\because\ p(x_j,y_k) = 0 \text{ for } j \neq k).$$

also from the matrix $[P(X,Y)]$

$$p(x_j,y_j) = p(x_j) = p(y_j).$$

Hence

$$H(X,Y) = H(X) = H(Y).$$

$$\& \quad H(Y/X) = H(X/Y) = -m\,(1\log 1) = 0.$$

Because there are m unity terms in the matrix and remaining terms are zero. Thus for Noise free channel,

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

problems

(1) An event has six possible outcomes with the probabilities $P_1 = 1/2$, $P_2 = 1/4$, $P_3 = 1/8$, $P_4 = 1/16$, $P_5 = 1/32$ and $P_6 = 1/32$. Find the entropy of the system. Also find the rate of information if there are 16 outcomes per second.

Sol. Given probabilities are

$$P_1 = 1/2,\ P_2 = 1/4,\ P_3 = 1/8,\ P_4 = 1/16,\ P_5 = 1/32,\ P_6 = 1/32.$$

Entropy H is given by the formula

$$H = -\sum_{K=1}^{n} P_K \log_2 P_K$$

$$= -\sum_{K=1}^{6} P_K \log_2 P_K$$

$$= -\left[P_1 \log P_1 + P_2 \log P_2 + P_3 \log P_3 + P_4 \log P_4 + P_5 \log P_5 + P_6 \log P_6\right]$$

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

$$= \tfrac{1}{2}\log_2 2 + \tfrac{1}{4}\log_2 2^2 + \tfrac{1}{8}\log_2 2^3 + \tfrac{1}{16}\log_2 2^4 + 1/32 \log_2 2^5 + \tfrac{1}{32}\log_2 2^5$$

$$= \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{4}{16} + \frac{5}{32} + \frac{5}{32}.$$

$$= \frac{31}{16} \text{ bits/message.}$$

Given $r = 16$ outcomes/sec.

Hence the rate of information R is

$$R = rH$$

$$= 16 \times \frac{31}{16}$$

$$R = 31 \text{ bits/message}$$

A continuous signal is bandlimited to 5KHz. The signal is quantized in 8 levels of a PCM system with the probabilities 0.25, 0.2, 0.2, 0.1, 0.1, 0.05, 0.05 and 0.05 calculate the entropy and the rate of information.

g. Given frequency $f_m = 5KHz$.

Sampling fq $f_s = 2f_m = 10KHz$.

$\therefore$ message rate $r = 10,000$ messages/sec.

Given probabilities $P_1 = 0.25$, $P_2 = 0.2$, $P_3 = 0.2$, $P_4 = 0.1$

$P_5 = 0.1$, $P_6 = 0.05$, $P_7 = 0.05$, $P_8 = 0.05$.

entropy 'H' is given by

$$H = -\sum_{K=1}^{n} P_K \log_2 P_K$$

$$= -\sum_{K=1}^{8} P_K \log_2 P_K$$

$$= -[P_1 \log P_1 + P_2 \log P_2 + P_3 \log P_3 + P_4 \log P_4 + P_5 \log P_5$$
$$+ P_6 \log P_6 + P_7 \log P_7 + P_8 \log_2 P_8]$$

$$H = -\left[0.25\log_2 0.25 + 0.2\log_2 0.2 + 0.2\log_2 0.2 + 0.1\log_2 0.1 + 0.1\log_2 0.1 + 0.05\log_2 0.05 + 0.05\log_2 0.05 + 0.05\log_2 0.05\right]$$

$H = 2.74$ bits/message.

Rate of information 'R' is given by

$'R' = rH = 10,000\,(2.74)$ bits/sec

$R = 27,400$ bits/sec

(3) A discrete memoryless source emits one of the 5 symbols once every millisec with probabilities 1/2, 1/4, 1/8, 1/16 and 1/16 respectively. find the source entropy & information rate

[Hint: $H = -\sum_{K=1}^{5} P_K \log_2 P_K$; $r = 1/\text{millisec} = 1/10^{-3} = 10^3$ mes/se

<u>Ans</u>:- $H = 1.875$ bits/message

$R = 1875$ bits/sec.

4) An analog signal is Bandlimited to B Hz sampled at Nyquist rate and samples are quantized into 4 levels. The quantization levels are Q_1, Q_2, Q_3 and Q_4, assume independent and occur with probabilities $P_1 = P_4 = 1/8$ and $P_2 = P_3 = 3/8$. find rate of information.

[Hint: $H = -\sum_{K=1}^{4} P_K \log P_K$. $= 1.811$ bits/message

$r = 2B$ messages/sec.

$R = 3.622B$ bits/sec].

⑤ A source emits 4 symbols with probabilities 1/2, 1/4, 1/8 & 1/8 respectively. (10)

(i) calculate the entropy

(ii) calculate the entropy if the symbols are equiprobable.
and compare 1 & 2.

Sol:- Given probabilities $P_1 = 1/2$

$P_2 = 1/4$

$P_3 = 1/8$

$P_4 = 1/8$.

(i)

$$H = -\sum_{K=1}^{n} P_K \log_2 P_K = -\sum_{K=1}^{4} P_K \log_2 P_K$$

$$= -[P_1 \log_2 P_1 + P_2 \log_2 P_2 + P_3 \log_2 P_3 + P_4 \log_2 P_4]$$

$$= -[1/2 \log_2 1/2 + 1/4 \log_2 1/4 + 1/8 \log_2 1/8 + 1/8 \log_2 1/8]$$

$H = 1.75$ bits/message

(ii)

$$P_1 = P_2 = P_3 = P_4 = 1/4$$

$$\therefore H = -\sum_{K=1}^{4} P_K \log P_K = -[4 \cdot 1/4 \cdot \log_2 1/4]$$

$H = 2$ bits/message

If all the messages have equal probability then we get perfect message.

⑥. A code is composed of dots and dashes. assume that the dash is 3 times as long as the dot and has 1/3 the probability of occurence

(i) calculate the amount of information in a dot and in a dash

(ii) calculate the average information in a dot dash word.

(iii) Assume that the dot lasts for 10 msec and that the same time interval is allowed between symbols. calculate the rate of information for the message "•—•" (dot dash dot).

sol:-

Given probability of dash is

$$p(-) = 1/3.$$

$$\therefore p(\cdot) = 2/3.$$

(i)

$$I_{(-)} = \log_2 1/p(-)$$

$$= \log_2 1/1/3 = \log_2 3 = 1.584 \text{ bits.}$$

$$I_{(\cdot)} = \log_2 1/p(\cdot) = \log_2 1/2/3 = \log_2 3/2 \text{ bits.}$$

(ii)

$$H = -\sum_{K=1}^{n} P_K \log_2 P_K = -\sum_{K=1}^{2} P_K \log_2 P_K$$

$$= -\left[1/3 \log_2 1/3 + \frac{2}{3} \log_2 2/3\right] = 0.917 \text{ bits/message}$$

(iii)

dot lasts for 10msec

$\therefore$ dash lasts for 30msec. & 10msec is allowed b/w the symbols

10 10

• — •

10 30 10

$$\therefore T = 70 \text{ msec.}$$

$\therefore$ 1 message $\rightarrow$ 70msec.

? $\leftarrow$ 10000msec (1sec)

$$\therefore \; r = \frac{10000}{70} \times 1$$

$$r = 14.2 \text{ messages/sec.}$$

$$H = -\sum_{K=1}^{2} P_K \log_2 P_K$$

$$= -\left[\underbrace{\tfrac{1}{3}\log_2 \tfrac{1}{3}}_{-0.528} + \tfrac{2}{3}\log_2 \tfrac{2}{3}\right]$$

$$= \underline{0.917} \text{ bits/message}$$

$\therefore$ $R = rH$ bits/sec

$= 14.2 \times H$ bits/sec. $= 13.03$ bits/sec.

=0=

(7) A transmitter has an alphabet of 4 letters x_1, x_2, x_3, x_4 and the receiver has an alphabet of 3 letters y_1, y_2, y_3. The joint probability matrix is

$$p(x,y) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 0.3 & 0.05 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0.15 & 0.05 \\ 0 & 0.05 & 0.15 \end{bmatrix} \end{array}$$

calculate all the entropies $H(x)$, $H(y)$, $H(xy)$, $H(x/y)$, $H(y/x)$

Sol:- Given joint probability matrix is

$$p(x,y) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 0.3 & 0.05 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0.15 & 0.05 \\ 0 & 0.05 & 0.15 \end{bmatrix} \end{array}$$

$$p(x_i) = \sum_{k=1}^{n} p(x_i, y_k)$$

$$\rightarrow \quad p(x_1) = \sum_{k=1}^{3} p(x_1, y_k)$$

$$= p(x_1, y_1) + p(x_1, y_2) + p(x_1, y_3)$$

$$= 0.3 + 0.05 + 0$$

$$\boxed{p(x_1) = 0.35}$$

$$\rightarrow \quad p(x_2) = \sum_{k=1}^{3} p(x_2, y_k)$$

$$= p(x_2, y_1) + p(x_2, y_2) + p(x_2, y_3)$$

$$= 0 + 0.25 + 0$$

$$\boxed{p(x_2) = 0.25}$$

$$\rightarrow \quad p(x_3) = \sum_{k=1}^{3} p(x_3, y_k)$$

$$= p(x_3, y_1) + p(x_3, y_2) + p(x_3, y_3)$$

$$= 0 + 0.15 + 0.05$$

$$\boxed{p(x_3) = 0.2}$$

$$\rightarrow \quad p(x_4) = \sum_{k=1}^{3} p(x_4, y_k)$$

$$= p(x_4, y_1) + p(x_4, y_2) + p(x_4, y_3)$$

$$= 0 + 0.05 + 0.15$$

$$\boxed{p(x_4) = 0.2}$$

$$\therefore \quad p(x_i) = \begin{bmatrix} 0.35 & 0.25 & 0.2 & 0.2 \end{bmatrix}$$

$$p(y_k) = \sum_{j=1}^{m} p(x_j, y_k)$$

$\rightarrow$ $$p(y_1) = \sum_{j=1}^{4} p(x_j, y_1)$$

$$= p(x_1, y_1) + p(x_2, y_1) + p(x_3, y_1) + p(x_4, y_1)$$

$$= 0.3 + 0 + 0 + 0$$

$$\boxed{p(y_1) = 0.3}$$

$\rightarrow$ $$p(y_2) = \sum_{j=1}^{4} p(x_j, y_2) = p(x_1, y_2) + p(x_2, y_2) + p(x_3, y_2) + p(x_4, y_2)$$

$$= 0.05 + 0.25 + 0.15 + 0.05$$

$$\boxed{p(y_2) = 0.5}$$

$\rightarrow$ $$p(y_3) = \sum_{j=1}^{4} p(x_j, y_3)$$

$$= p(x_1, y_3) + p(x_2, y_3) + p(x_3, y_3) + p(x_4, y_3)$$

$$= 0 + 0 + 0.05 + 0.15$$

$$\boxed{p(y_3) = 0.2}$$

$$\therefore \; p(y_k) = [0.3 \quad 0.6 \quad 0.2].$$

$\rightarrow$ $$H(X) = -\sum_{j=1}^{4} p(x_j) \log_2 p(x_j).$$

$$= -\Big[p(x_1) \log_2 p(x_1) + p(x_2) \log_2 p(x_2) + p(x_3) \log_2 p(x_3) + p(x_4) \log_2 p(x_4) \Big]$$

$$= -\Big[0.35 \log_2 0.35 + 0.25 \log_2 0.25 + 0.2 \log_2 0.2 + 0.2 \log_2 0.2 \Big]$$

$$\boxed{H(X) = 1.95 \text{ bits/message}}$$

$\rightarrow$ $H(Y) = -\sum_{K=1}^{3} p(y_K)\log_2 p(y_K)$

$= -[p(y_1)\log_2 p(y_1) + p(y_2)\log_2 p(y_2) + p(y_3)\log_2 p(y_3)]$

$= -[0.3\log_2 0.3 + 0.5\log_2 0.5 + 0.2\log_2 0.2]$

$\boxed{H(Y) = 1.485 \text{ bits/message}}$

$\rightarrow$ $H(XY) = -\sum_{j=1}^{4}\sum_{K=1}^{3} p(x_j, y_K)\log_2 p(x_j, y_K)$

$= -[p(x_1,y_1)\log_2 p(x_1,y_1) + p(x_1,y_2)\log_2 p(x_1,y_2) + p(x_1,y_3)\log_2 p($

$+ p(x_2,y_1)\log_2 p(x_2,y_1) + p(x_2,y_2)\log_2 p(x_2,y_2) + p(x_2,y_3)\log_2 p(x_2,y_3)$

$+ p(x_3,y_1)\log_2 p(x_3,y_1) + p(x_3,y_2)\log_2 p(x_3,y_2) + p(x_3,y_3)\log_2 p(x_3,y_3)$

$+ p(x_4,y_1)\log_2 p(x_4,y_1) + p(x_4,y_2)\log_2 p(x_4,y_2) + p(x_4,y_3)\log_2 p(x_4,y_3)$

$= -[0.3\log_2 0.3 + 0.05\log_2 0.05 + 0.25\log_2 0.25 + 0.15\log_2 0.15 +$

$0.05\log_2 0.05 + 0.05\log_2 0.05 + 0.15\log_2 0.15]$

$H[XY] = 2.48$ bits/message

$\rightarrow$ $H(X/Y) = H(X:Y) - H(Y) = 2.48 - 1.485$

$\boxed{H(X/Y) = 1 \text{ bit/message}}$

$\rightarrow$ $H(Y/X) = H(XY) - H(X) = 2.48 - 1.95$

$\boxed{H(Y/X) = 0.53 \text{ bits/message}}$

13.

8) Complete the following probability matrix in all possible ways.

$$\begin{bmatrix} 0.1 & a & 0.2 & 0.4 \\ 0.3 & 0.1 & b & 0.5 \\ c & 0.4 & 0.1 & 0.3 \\ 0.2 & 0.2 & 0.1 & d \end{bmatrix}$$

Sol.

Given matrix is

$$\begin{bmatrix} 0.1 & a & 0.2 & 0.4 \\ 0.3 & 0.1 & b & 0.5 \\ c & 0.4 & 0.1 & 0.3 \\ 0.2 & 0.2 & 0.1 & d \end{bmatrix}$$

* The given matrix is a two-dimensional matrix as it is neither a row nor a coloumn matrix.

* The sum of all the given entries is 2.9 which is greater than '1', it cannot be a joint probability matrix.

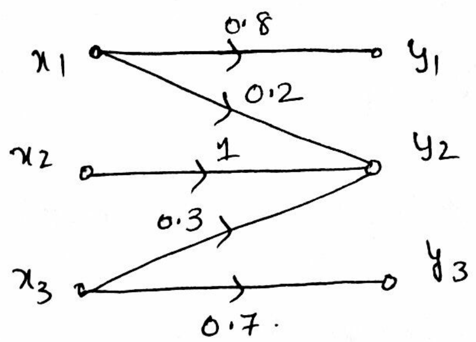
* The sum of all the given entries in the fourth coloumn of the matrix is 1.2 (which is greater than 1), it cannot be a conditional probability matrix $p(X/Y)$.

** As the sum of all the given entries in each row of the matrix is less than 1, it can be a conditional probability matrix $P(Y/X)$ with the value of a, b, c and d so chosen that the sum of each row becomes 1.

Thus when $a = 0.3$, $b = 0.1$, $c = 0.2$ and $d = 0.5$, the given matrix is the conditional matrix $P(Y/X)$. Thus, there is only one possible way to complete the given probability matrix and it is:

$$\begin{bmatrix} 0.1 & 0.3 & 0.2 & 0.4 \\ 0.3 & 0.1 & 0.1 & 0.5 \\ 0.2 & 0.4 & 0.1 & 0.3 \\ 0.2 & 0.2 & 0.1 & 0.5 \end{bmatrix} = P(Y/X).$$

Q. A discrete source transmits messages x_1, x_2 and x_3 with the probabilities 0.3, 0.4 and 0.3. The source is connected to the channel given in figure. Calculate all the entropies.



Sol. The figure gives the conditional probability $p(Y/X)$ as.

$$p(Y/X) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 1 & 0 \\ 0 & 0.3 & 0.7 \end{bmatrix} \end{array}$$

also given

$$p(X) = [0.3 \quad 0.4 \quad 0.3].$$

The joint probability matrix $p(X, Y)$ can be obtained by multiplying the rows of $p(Y/X)$ by $p(x_1)$, $p(x_2)$ and $p(x_3)$ i.e by 0.3, 0.4 and 0.3, respectively, giving

$$p(X, Y) = \begin{bmatrix} 0.8 \times 0.3 & 0.2 \times 0.3 & 0 \\ 0 & 1 \times 0.4 & 0 \\ 0. & 0.3 \times 0.3 & 0.7 \times 0.3 \end{bmatrix}$$

(sum of all the entries in $p(X, Y)$ is 1).

The probabilities $p(y_1)$, $p(y_2)$ and $p(y_3)$ can be obtained by adding the coloumns of $p(X, Y)$ giving

$$p(y_1) = 0.24, \quad p(y_2) = 0.06 + 0.4 + 0.09 = 0.55, \quad p(y_3) = 0.21.$$

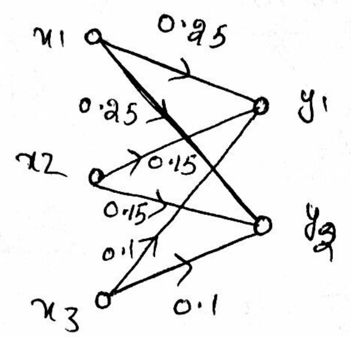
$H(X) = 1.571$ bits/message, $H(Y) = 1.441$ bits/message.

$H(X, Y) = 2.053$ bits/message.

$H(X/Y) = 0.612$ bits/message

$H(Y/X) = 0.482$ bits/message

(10) Find the mutual information for the channel shown in figure. (14)



sol:- The joint probability matrix for the channel shown in figure

$$[P(X,Y)] = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{bmatrix} 0.25 & 0.25 \\ 0.15 & 0.15 \\ 0.1 & 0.1 \end{bmatrix} \end{array}$$

$P(x_1) = 0.25 + 0.25 = 0.5$

$P(x_2) = 0.15 + 0.15 = 0.3$

$P(x_3) = 0.1 + 0.1 = 0.2$

$P(y_1) = P(y_2) = 0.25 + 0.15 + 0.1 = 0.5$.

$H(X) = -\sum_{j=1}^{3} P(x_j)\log_2 P(x_j) = -[0.5\log_2 0.5 + 0.3\log_2 0.3 + 0.2\log_2 0.2]$

$= 1.485$ bits/message.

$H(Y) = -\sum_{K=1}^{2} P(y_K)\log_2 P(y_K) = -[0.5\log_2 0.5 + 0.5\log_2 0.5]$

$= 1$ bit/message.

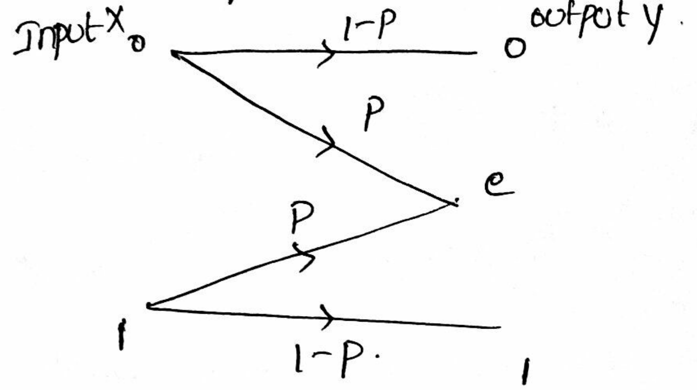
$H(X,Y) = -\sum_{j=1}^{3}\sum_{K=1}^{2} P(x_j, y_K)\log_2 P(x_j, y_K)$

$= 2.485$ bits/message

$I[X;Y] = H(X) + H(Y) - H(X,Y) = 1.485 + 1 - 2.485 = 0$

The channel shown in figure is with an independent input and output and the answer $I(X;Y) = 0$ confirms that the mutual information in such a case is zero.

(11) Figure illustrates a binary erasure channel with transmission probabilities $P(0/0) = P(1/1) = 1-P$ and $P(e/0) = P(e/1) = P$. The probabilities for the input symbols are $P(X=0) = \alpha$ and $P(X=1) = 1-\alpha$



Determine the average mutual information $I[X;Y]$ in bits.

sol. Given $P(X=0) = \alpha \Rightarrow P(x_1)$

$P(X=1) = 1-\alpha = P(x_2)$

the channel matrix is,

$$P(Y/X) = \begin{bmatrix} 1-P & P & 0 \\ 0 & P & 1-P \end{bmatrix}$$

$$P(X) = [\alpha \quad 1-\alpha]$$

$$\therefore P(X,Y) = \begin{bmatrix} \alpha(1-P) & \alpha P & 0 \\ 0 & P(1-\alpha) & (1-P)(1-\alpha) \end{bmatrix}$$

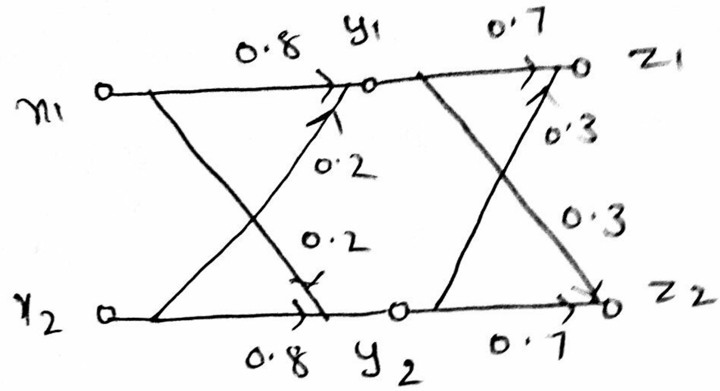
$$P(y_1) = \alpha(1-P), \quad P(y_2) = \alpha P + P(1-\alpha) = \alpha P + P - \alpha P$$
$$= P$$

$$P(y_3) = (1-P)(1-\alpha)$$

$$H(X) = \alpha \log 1/\alpha + (1-\alpha) \log(1/1-\alpha)$$
$$= -[\alpha \log \alpha + (1-\alpha) \log(1-\alpha)].$$

$$H(Y) = -[\alpha(1-P) \log_2 \alpha(1-P) + P \log_2 P + (1-P)(1-\alpha) \log_2 ((1-P)(1-\alpha))]$$

→ Two BSC's are connected in cascade as shown in figure (15)



(i) find the channel matrix of resultant channel

(ii) find $p(z_1)$ and $p(z_2)$ if $p(x_1)=0.6$ and $p(x_2)=0.4$

sol:- $P(Y/X)$, $P(Z/Y)$

channel matrix

$$\rightarrow P(Z/X) = P(Y/X)\, P(Z/Y). = \begin{bmatrix} 0.62 & 0.38 \\ 0.38 & 0.62 \end{bmatrix}.$$

$$P(Z \cdot X) = P(Z/X)\, P(X)$$

$$\begin{array}{c|cc} & z_1 & z_2 \\ \hline x_1 & & \\ x_2 & & \end{array}$$

$\Downarrow$ $\Downarrow$

0.524 0.476

for this problem refer Evaday.