Source Encoder & Decoder

Syllabus: Information, entropy, source coding and decoding: Fixed length, Shanon fano, huffmann, joint & conditional entropy, redundancy, mutual & self informations, Binary Symmetric Channel, Binary Erase channel, Shanon Hartley theorem (or) channel capacity theorem.

Information:

The basic principle in determining the information content of a message is " more the uncertainity (Surprise / unexpected)" of the event. Most the information content is occurance of the event.

Example: Let us consider the message as

* (1) " Sun rises in the East ".

The probability of the above event is the day-to-day phenomenon. no body shows interest in it, so the uncertainity of the above event is less and the information content is less.

* If some body say that " Sun is going to rise in the West ". then uncertainity of the above event is more & the information content is more.

(2) Dog bites a man (E_1)
A man bites a dog (E_2).

If event 'E_1' is most occurance so no information is there whereas in event 'E_2', it is not most occurance hence it has more information probability of occurance is less.

∴ Information is inversly proportional to the probability of occurance of the event (or message) ie

$$\boxed{\begin{aligned} I &\propto \frac{1}{P(E)} \\ I &= \log_x(1/P) = -\log_x(P) \end{aligned}}$$

" $P(E) = P$ "

where '$\log_x$' is proportionality constant.

Units of Information:

$$I = \log_x(1/p) \text{ (or) } -\log_x(p)$$

a) If the base $x=2$ then $I = \log_2(1/p) = -\log_2 p$ Units : **bits**

b) If the base $x=10$ then $I = \log_{10}(1/p) = -\log_{10} p$ Units : Decits (or) Hartley.

c) If the base $x=e$ then $I = \log_e(1/p) = -\log_e p$ Units : **nats**.

* The source is generating 'x' and 'y' are two independent messages with probability $p(x)$ and $p(y)$, then the total information carried by x & y are

$$I(x,y) = \log_2 \frac{1}{p(x,y)} \text{ bits}$$

$$p(x,y) = p(x) \cdot p(y) \quad (\because \text{Independents events})$$

$$I(x,y) = \log_2\left[\frac{1}{p(x)\cdot p(y)}\right]$$

$$= \log_2(1/p(x)) + \log_2(1/p(y))$$

$$= I(x) + I(y)$$

$$\therefore \boxed{I(x,y) = I(x) + I(y)}$$

where $I(x) = \log_2 1/p(x)$, $I(y) = \log_2(1/p(y))$.

Entropy:

"The average amount of information per individual messages in a particular interval is called Entropy".

✓ The entropy (or) Average information transmitted under the following assumptions.

a) The source is stationary ie the probability of occurance of each message remain constant with time.

b) Each message Emitted from the source is independent of previous messages.

✓ Entropy is denoted as 'H' and the Units for entropy

$H \rightarrow$ bits/message.

Symbol → 1010 → Codeword
↘ length = 4

Let us consider 'M' different messages m_1, m_2, m_3 ...... with their respective probabilities of occurance is P_1, P_2, P_3, ...

For along time interval 'L' messages have been generated then (L >> M)

The no. of messages $m_1 = P_1 \times L$

The amount of information in message $m_1 = \log(1/P_1)$ & $-\log(P_1)$

∴ The total amount of information in all m_1 messages

$$P_1 = L \log\left(\frac{1}{P_1}\right)$$

∴ The amount of information in all 'L' messages will be

$$I_t = P_1 \cdot L \cdot \log\left(\frac{1}{P_1}\right) + P_2 \cdot L \log(1/P_2) + \cdots\cdots$$

The average amount of information per message & Entropy is given by

$$H = \frac{I_t}{L} = P_1 \log(1/P_1) + P_2 \log(1/P_2) + \cdots\cdots$$

$$\therefore \boxed{\text{Entropy, } H = \sum_{K=1}^{M} P_K \log\left(\frac{1}{P_K}\right) \text{ (or) } -\sum_{K=1}^{M} P_K \cdot \log(1/P_K)}$$

Properties of Entropy :

① The entropy function is symmetrical

$$H(P_K, P_{K-1}) = H(P_{K-1}, P_K).$$

② The entropy 'H' is continuous in the interval $0 \le P_K \le 1$.

③ Dividing the entropy into different subsets does not effect the value of entropy.

ie $H(M) = H(m_1, m_2, m_3, \ldots) + H(m_{n+1}, m_{n+2}, m_{n+3}, \ldots)$

$\therefore H(M) = H(m_1, m_2, m_3, \ldots\ldots m_{n+1}, m_{n+2}, m_{n+3}, \ldots)$.

④ $0 \le H \le \log(M)$.

⑤ $H = 0$ if all the probabilities are '0' except one which must be unit.

⑥ $H = \log M$, if all the probabilities are equally likely.

⑦ Entropy is maximum when uncertainity is more

Proof :

4 * Let us Examine entropy under different cases for 'M=2'

Case (i) : $P_1 = 0.01$, $P_2 = 0.99$

Case (ii) : $P_1 = 0.4$, $P_2 = 0.6$

Case (iii) : $P_1 = 0.5$, $P_2 = 0.5$.

case (i) :
$$H = P_1 \log(1/P_1) + P_2 \log(1/P_2)$$
$$= 0.01 \log_2 (1/0.01) + 0.99 \log_2 \left(\frac{1}{0.99}\right)$$
$$\left[\because \frac{\log_{10}(1/0.01)}{\log_{10} 2} \times 0.01 = 0.0664 \right]$$
$$= 0.0664 + 0.0143$$
$$\boxed{H = 0.08 \text{ bits/msg}}$$

case (ii) :
$$H = 0.4 \log_2 (1/0.4) + 0.6 \log_2 (1/0.6)$$
$$= 0.528 + 0.442$$
$$\boxed{H = 0.97 \text{ bits/msg}}$$

Case (iii) :
$$H = 0.5 \log_2 (1/0.5) + 0.5 \log_2 (1/0.5)$$
$$= 0.5 + 0.5$$
$$= 1$$
$$\therefore \boxed{H = 1 \text{ bits/msg.}}$$

Hence

Case (i) : The probabilities of messages m_1, m_2 are 0.01 , 0.99 , So uncertainity is less.

Case (ii) : The uncertainity is more compared to case (i), Since $P_1 = 0.4$, $P_2 = 0.6$.

Case (iii) : The uncertainity is maximum because messeges are equally likely.

5 * If there is only single possible message is there ie

$m_1 = 1$ & $P_1 = 1$ then

$$H = P_1 \log(1/P_1) \Rightarrow H = 1 \cdot \log(1/1) = 0 \quad \therefore \boxed{H = 0}$$

* Let there be only one messege out of 'M' messeges having probability '1' and all others are '0' in that case.

$$H = \sum_{K=1}^{M} P_K \log\left(\frac{1}{P_K}\right)$$

$$H = P_1 \log\left(\frac{1}{P_1}\right) + P_2 \log\left(\frac{1}{P_2}\right) + \cdots + P_M \log\left(\frac{1}{P_M}\right)$$

$$= 1 \log_2 (1/1) + 0 + \cdots + 0$$

$$= 0$$

$$\therefore \boxed{H = 0}$$

* For a binary system i.e. $M=2$ then the entropy is 1 $\therefore \boxed{H=1}$

<u>Proof</u>:

$$H = \sum_{k=1}^{2} P_K \log(1/P_K)$$

$$H = P_1 \log(1/P_1) + P_2 \log(1/P_2)$$

let

$$P_1 = P \text{ and } P_2 = 1-P \quad [\text{ equally likely}]$$

$$\therefore H = P \log(1/P) + (1-P) \log(1/1-P)$$

Differentiating above eqn w.r.t. 'P' & equating it to zero.

$$\therefore \frac{dH}{dP} = 0$$

$$\Rightarrow \frac{dH}{dP} = \frac{d}{dP}\left\{-\left[P\log P + (1-P)\log(1-P)\right]\right\}$$

$$0 = -\left\{P \cdot \frac{1}{P} + \log P (1) + (1-P) \cdot \frac{1}{(1-P)} (-1) + \log(1-P) \cdot (-1)\right\}$$

$$= -\left[\cancel{1} + \log P - \cancel{1} - \log(1-P)\right]$$

$$= -\left[\log P - \log(1-P)\right]$$

$$0 = -\log P + \log(1-P)$$

$$\log P = \log(1-P)$$

$$\therefore P = 1-P \rightarrow \text{Equally likely.}$$

The maximum value of entropy at $P = 1/2$, $1-P = 1/2$

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

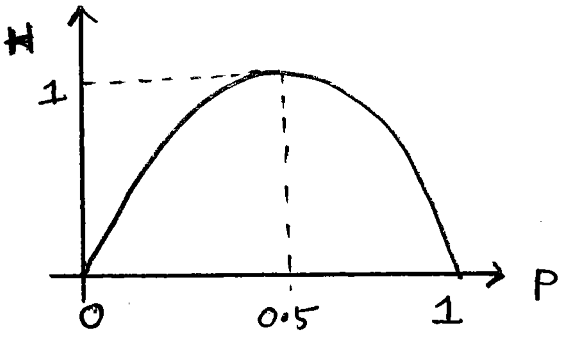
$$= -\left[\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}\right]$$

$$= -\left[1 \log_2 \frac{1}{2}\right]$$

$$= -\log_2 2^{-1}$$

$$H = 1 \qquad \therefore \boxed{H = 1} \text{ bits/msg (or) bits/symbol.}$$

ie The entropy is maximum when both the messages are equally likely ie when $p = \frac{1}{2}$.



6 * Similarly in the case of M messages, the entropy is maximum when all the messages are equally likely.

ie $P_1 = P_2 = P_3 = \cdots = P_M = \frac{1}{M}$

$$\therefore H = \sum_{K=1}^{M} P_K \log(1/P_K)$$

$$= P_1 \log(1/P_1) + P_2 \log\left(\frac{1}{P_2}\right) + \cdots + P_M \log(1/P_M)$$

$$= \frac{1}{M}\log(M) + \frac{1}{M}\log(M) + \cdots + \frac{1}{M}\log(M)$$

$$= \cancel{M} \times \frac{1}{\cancel{M}} \log_2 M$$

$$\therefore \boxed{H_{max} = \log_2 M} \text{ bits/msg.}$$

ie The Entropy ranges from 0 to $\log_2 M$ $\therefore \boxed{0 \le H \le \log_2 M}$

Information rate (R):

The average no. of bits per sec

ie If a message source generate messages at a rate of 'r' messages per sec, the rate of information & Information rate is given by

$$\boxed{R = r \times H} = \text{msg/sec} \times \text{bits/msg}$$
$$= \text{bits/sec}$$

$\therefore$ The unit for information rate $\boxed{R = \text{bits/sec}}$

where R - Information rate
H - Entropy
r - no. of messages per sec.

Let us consider two sources of equal entropy, generating 'r_1 & r_2' messages per sec respectively

a) The first source will transmit its information at a rate

$$R_1 = r_1 \times H$$

b) The Second source will transmit the information at a rate

$$R_2 = r_2 \times H.$$

c) If $r_2 > r_1$ then $R_2 > R_1$ ie more information transmitted from the second source than the first source in a given second.

Hence the source is not described by its entropy alone but also by its rate of information.

Problem: An event has 6 possible outcomes with probabilities $P_1 = 1/2$, $P_2 = \frac{1}{4}$, $P_3 = 1/8$, $P_4 = \frac{1}{16}$, $P_5 = \frac{1}{32}$, $P_6 = \frac{1}{32}$. Find the entropy (H) and the rate of information (R) if there are 16 outcomes per sec.

Sol

P_1	P_2	P_3	P_4	P_5	P_6
$1/2$	$1/4$	$1/8$	$1/16$	$1/32$	$1/32$

$r = 16$ msg/se

Entropy $H = \sum_{K=1}^{M} P_K \log(1/P_K) = \sum_{K=1}^{6} P_K \log(1/P_K)$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{16}\log 16 + \frac{1}{32}\log 32 + \frac{1}{32}\log 32$$

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

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

$$= \frac{16 + 16 + 12 + 8 + 5 + 5}{32}$$

$$= \frac{62}{32}$$

$$= 1.9375 \quad \therefore \boxed{H = 1.9375 \text{ bits/msg}}$$

Information rate $R = r \times H = 16 \times 1.9375 = 31$

$$\therefore \boxed{R = 31 \text{ bits/sec.}}$$

(2) A continuous signal is band limited to 5 KHz, The signal is sampled and quantized into 8 levels of a PCM system, each quantization levels (or) the probability of occurance of each level are 0.25, 0.2, 0.2, 0.1, 0.1, 0.05, 0.05, 0.05. Calculate Entropy and information rate.

Sol

$f_m = 5KHz$, PCM system

Given 8 - quantization levels

ie $\{P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8\} = \{0.25, 0.2, 0.2, 0.1, 0.1, 0.05, 0.5, 0.05\}$

$\therefore$ Entropy $H = \sum_{K=1}^{M} P_K \log_2 (1/P_K) = -\sum_{K=1}^{8} P_K \log P_K$.

$$H = -\Big[0.25\log_2(0.25) + 0.2\log_2(0.2) + 0.2\log_2(0.2) + 0.1\log_2(0.1) + 0.1\log_2(0.1) + 0.05\log_2(0.05) + (0.05)\log(0.05) + (0.05)\log(0.05)\Big] \Big/ \log 2$$

$$= -[-0.5 - 0.464 - 0.464 - 0.332 - 0.332 - 0.216 - 0.216 - 0.216]$$

$$= -[-2.74]$$

$$= 2.74$$

$\therefore$ Entropy $\boxed{H = 2.74 \text{ bits/msg.}}$ Quantization level

Given

$f_m = 5 KHz$

$f_s = 2f_m = 2 \times 5 KHz = 10 KHz$

$\therefore r = 10,000$ Quantization levels / sec.

Information rate $R = r \times H$

$= 10,000 \times 2.74$ × Quantization/sec × bits/Quantization

$= 27.4 \times 10^3$ bits/sec

Information rate $\boxed{R = 27.4K \text{ bits/sec}}$

(3) A given telegraph source having two symbols '•' and '—'. Dot duration is 0.2 sec and dash duration is 3 times the dot duration. The probability of dot occuring is twice that of dash occurance and the time difference between Dot and Dash is 0.2 sec. Find entropy (H) and Information rate (R) of telegraph source.

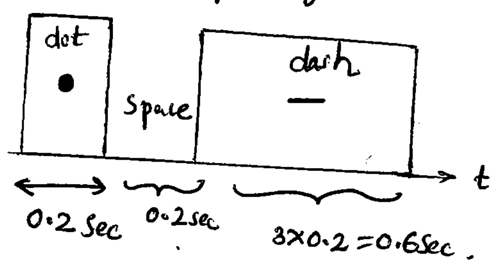
Sol Given

Dot duration $t_{dot} = 0.2\,Sec$

Dash duration $t_{dash} = 3 \times t_{dot} = 0.6\,Sec$

$t_{space} = 0.2\,sec.$

Telegraph system



$$P_{(dot)} = 2 \cdot P_{(dash)}$$

Total probability $= 1$

$$P_{dot} + P_{dash} = 1$$

$$2P_{dash} + P_{dash} = 1$$

$$3P_{dash} = 1$$

$$\boxed{P_{dash} = 1/3} \qquad \boxed{P_{dot} = 2/3}$$

∴ Entropy $H = \sum_{K=1}^{2} P_K \cdot \log(1/P_K)$ & $- \sum_{K=1}^{2} P_K \log(P_K)$.

$$= -\left[P_{dash} \cdot \log(P_{dash}) + P_{dot} \cdot \log(P_{dot})\right].$$

$$= -\left[\frac{1}{3}\log_2(1/3) + 2/3\log_2(2/3)\right]$$

$$= -\left[\frac{-0.1590 - 0.11739}{\log 2}\right]$$

$$= \frac{0.27639}{0.30102} = 0.918 \quad \therefore \boxed{H = 0.918\ bits/msg}$$

Avg time per symbol

Time duration $t_s = P_{dot} \cdot t_{dot} + P_{dash} \cdot t_{dash} + P_{space} \cdot t_{space}$

$$= 2/3 \cdot (0.2) + \frac{1}{3} \times 0.6 + 1 \times 0.2 \qquad (\because \text{let } P_{space} = 1)$$

$$t_s = 0.533\ sec$$

∴ Information rate $\boxed{r = \frac{1}{t_s} = 1.875\ msg/sec}$

$$R = r \times H = 1.875 \times 0.918 \Rightarrow \boxed{R = 1.72125\ bits/sec}$$

④ A high resolution black & white TV picture consists of 2×10^6 pixels/frame and 16 different bright levels, pixel is repeated at a rate of 32 frames/sec. All pixel elements are assumed to be independent & equiprobability. Calculate Entropy (H) and Information rate (R)?

Sol

Given Different bright levels $M = 16$ (Independent & equally likely)

Entropy $H = \log_2 M = \log_2 2^4 = 4$

$\therefore \boxed{H = 4 \text{ bits/pixel}}$

Information rate $R = r \times H$

where $r = 2\times10^6$ pixels/frame $\times$ $32 \times$ frames/sec

$r = 64\times10^6$ pixels/sec

$\therefore$ Information rate $R = r \times H$

$= 64\times10^6$ pixels/sec $\times$ 4 bits/pixel

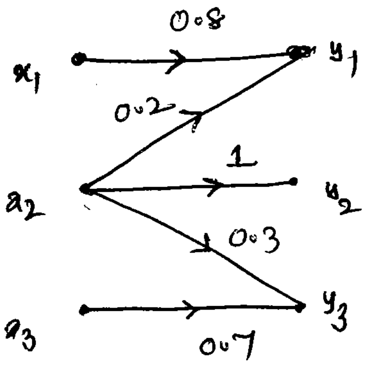
$= 256$ Mbits/sec

$\therefore \boxed{R = 256 \text{ M bits/sec}}$

Problems: →

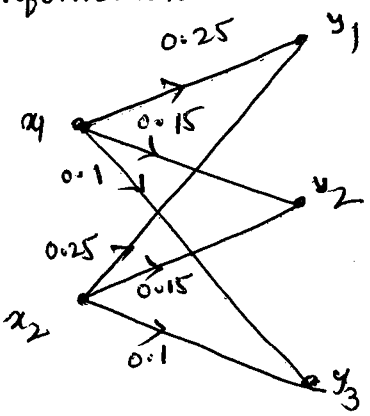
① Find All entropies and Mutual Information.

ⓐ



P(X/Y)

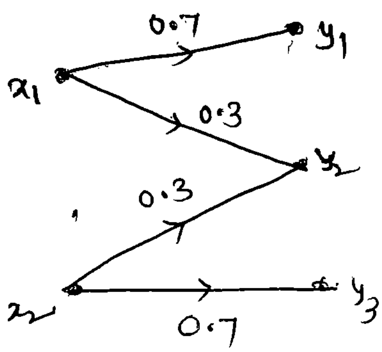
ⓑ



P(XY)

$p(y_1) = 0.2$, $p(y_2) = 0.5$, $p(y_3) = 0.3$.

ⓒ



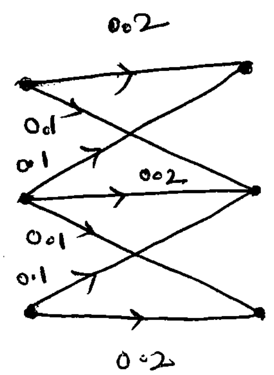
P(Y/X)

$p(x_1) = 0.5$

$p(x_2) = 0.5$

channel capacity $C = ?$

Ⓓ

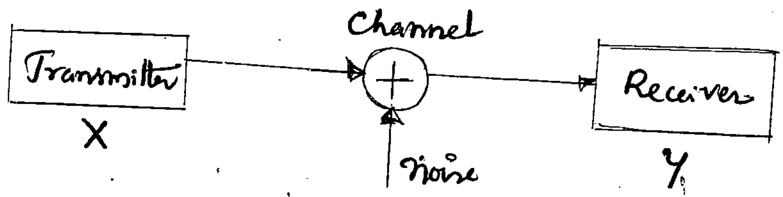


P(XY)

Joint Entropy and Conditional Entropy:

To study the behaviour of a communication system, we must study the behaviour of transmitter and the receiver, this gives rises to the concept of a two dimentional probability scheme.

* Let us consider two finite discrete sample space S_1 & S_2 let their product space $S = S_1 S_2$.



Let us consider transmitter 'x' generating 'm' symbols $\{x_1, x_2, x_3 \dots x_m\}$ with probabilities $\{p(x_1), p(x_2), p(x_3) \dots p(x_m)\}$

Let us consider the receiver 'y' receiving 'n' symbols $\{y_1, y_2, y_3 \dots y_n\}$ with probabilities $\{P(y_1), P(y_2), P(y_3), \dots P(y_n)\}$.

$$[XY] = \begin{bmatrix} x_1 y_1 & x_1 y_2 & x_1 y_3 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & x_2 y_3 & \dots & x_2 y_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_m y_1 & x_m y_2 & x_m y_3 & \dots & x_m y_n \end{bmatrix}$$

Thus we have 3 sets of complete probability schemes.

$$P(X) = [p(x_i)]$$
$$P(Y) = [P(y_j)]$$
$$P(XY) = [P(x_i y_j)]$$

Three associated entropies are

$$H(X) = -\sum_{i=1}^{m} P(x_i) \cdot \log [p(x_i)], \text{ where } p(x_i) = \sum_{j=1}^{n} P(x_i y_j)$$

$$H(Y) = -\sum_{j=1}^{n} P(y_j) \log [P(y_j)], \text{ where } P(y_j) = \sum_{i=1}^{m} p(x_i y_j)$$

∴ The joint entropy

$$\boxed{H(XY) := -\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i y_j) \log [p(x_i y_j)]}$$

$H(X)$ and $H(Y)$ are marginal entropies of X & Y.

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

- The conditional probability, from Baye's theorem

$$P(X/Y) = \frac{P(XY)}{P(Y)} \Rightarrow P(XY) = P(X/Y) \cdot P(Y)$$

(or)

$$P(Y/X) = \frac{P(XY)}{P(X)} \Rightarrow P(XY) = P(Y/X) \cdot P(X).$$

We know that 'y_j' may occur in conjuction with $x_1, x_2, \cdots x_m$.

thus $[X/y_j] = [x_1/y_j \;\; x_2/y_j \;\; x_3/y_j \;\; \cdots \;\; x_m/y_j]$

Probability $P(X/y_j) = [\,p(x_1/y_j) \;\; p(x_2/y_j) \;\; p(x_3/y_j) \;\; \cdots \;\; p(x_m/y_j)\,]$

& $p(x_1 y_j) + p(x_2 y_j) + p(x_3 y_j) + \cdots\cdots p(x_m, y_j) = P(y_j)$

Thus $\sum_{i=1}^{m} p(x_i/y_j) = 1.$

∴ The entropy may be associated with

$$H(X/y_j) = -\sum_{i=1}^{m} \left[\frac{p(x_i\, y_j)}{p(y_j)}\right] \cdot \log\left[\frac{p(x_i, y_j)}{p(y_j)}\right]$$

$$= -\sum_{i=1}^{m} p(x_i/y_j) \log p(x_i/y_j)$$

Consider all the messages of source & 'y_j' of receiver &

The measure of an average conditional entropy of the system.

$$H(X/Y) = \overline{H(X/y_j)}$$

$$= \sum_{j=1}^{n} p(y_j) \cdot H(X/y_j)$$

$$= -\sum_{j=1}^{n} p(y_j) \left[\sum_{i=1}^{m} p(x_i/y_j) \log p(x_i/y_j)\right]$$

Conditional Entropy

$$= -\sum_{i=1}^{m} \sum_{j=1}^{n} p(y_j) \cdot p(x_i/y_j) \cdot \log p(x_i/y_j)$$

$$\therefore \boxed{H(X/Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i y_j) \log p(x_i/y_j)} \quad (\because \text{Bayes theorem})$$

Similarly

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

$\therefore$ There are five entropies associated with two dimentional probability scheme. They are : $H(X), H(Y), H(XY), H(X/Y)$ & $H(Y/X)$.

let

X - represents a transmitter

Y - represents a Receiver.

H(X) : Average information per character at the transmitter (or) Entropy of the transmitter.

H(Y) : Averge information per character at the receiver (or) Entropy of the receiver.

H(XY) : Entropy of a communication system.

H(X/Y) : Measure of information about the transmitter, where it is known that 'Y' is Received.

H(Y/X) : Measure of information about the receiver, where it is known that 'X' is transmitted.

Hence :

$$\left\{\begin{aligned}
H(X) &= -\sum_{i=1}^{m} p(x_i)\log p(x_i) \\
H(Y) &= -\sum_{j=1}^{n} p(y_j)\log p(y_j) \\
H(XY) &= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(x_i y_j) \\
H(X/Y) &= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(x_i/y_j) \\
H(Y/X) &= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(y_j/x_i)
\end{aligned}\right\}$$

Relation b/w Different Entropies :

Case(i): $H(XY) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\cdot\log[p(x_i y_j)]$

Using Baye's theorem

$$p(x_i/y_j) = \frac{p(x_i y_j)}{p(y_j)} \quad (\text{or}) \quad p(y_j/x_i) = \frac{p(x_i y_j)}{p(x)}$$

$$\therefore \quad H(XY) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\cdot\log\left\{ p(x_i/y_j)\cdot p(y_j)\right\}$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\left\{ \log p(x_i/y_j) + \log p(y_j)\right\}$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} \left[p(x_i y_j)\log p(x_i/y_j) + p(x_i y_j)\log p(y_j)\right]$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(x_i/y_j) - \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(y_j)$$

$$H(XY) = H(X/Y) - \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log p(y_j).$$

$$\left\{ \begin{array}{l} \because \; p(y_j) = \sum_{i=1}^{m} p(x_i y_j) \\ \quad\; p(x_i) = \sum_{j=1}^{n} p(x_i y_j). \end{array} \right\}$$

$$\therefore \quad H(XY) = H(X/Y) - \sum_{j=1}^{n} p(y_j)\log p(y_j)$$

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

$$\therefore \quad \boxed{H(X/Y) = H(XY) - H(Y).}$$

Similarly $\quad H(XY) = H(Y/X) + H(X)$

$$\therefore \quad \boxed{H(Y/X) = H(XY) - H(X)}$$

<u>Case (ii)</u>:

If X and Y are independent events then

$$p(x_i y_j) = p(x_i)\cdot p(y_j).$$

$$\therefore \quad H(XY) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\log\left[p(x_i y_j)\right]$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\left[\log p(x_i) + \log p(y_j)\right]$$

$$= -\sum_{i=1}^{m}\left[\sum_{j=1}^{n} p(x_i y_j)\right]\log p(x_i) - \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\cdot\log p(y_j)$$

$$= -\sum_{i=1}^{m} p(x_i)\log[p(x_i)] - \sum_{j=1}^{n} p(y_j)\log[p(y_j)]$$

$$= H(X) + H(Y)$$

$$\therefore \quad \boxed{H(XY) = H(X) + H(Y)}$$

Case iii: $H(X) \geq H(X/Y)$ and $H(Y) \geq H(Y/X)$.

The relation is proved by using the inequality as $\ln\left(\frac{1}{x}\right) \geq 1-x$ by writing the entropy in 'nats'.

$$\therefore H(X) - H(X/Y) \geq -\sum_{i=1}^{m} p(x_i) \log p(x_i) + \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \cdot \log[p(x_i/y_j)]$$

$$\geq -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(x_i) + \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(x_i/y_j)$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \left[\log p(x_i/y_j) - \log p(x_i)\right]$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log\left[\frac{p(x_i/y_j)}{p(x_i)}\right]$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \left(1 - \frac{p(x_i)}{p(x_i/y_j)}\right) \qquad \left(\because \log\frac{1}{x} = 1-x\right)$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) - \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{p(x_i y_j) \cdot p(x_i)}{p(x_i/y_j)}$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i, y_j) - \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{p(\cancel{x_i y_j}) \cdot p(x_i) \cdot p(y_j)}{p(\cancel{x_i y_j})}.$$

$$\geq \underbrace{\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)} - \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i) \cdot p(y_j) \qquad \text{(Bayes theorem)}$$

$$\geq \sum_{i=1}^{m} p(x_i) - \sum_{i=1}^{m} p(x_i) \qquad \left[\because \sum_{j=1}^{n} p(x_i y_j) = p(x_i), \ \sum_{j=1}^{n} p(y_j) = 1.\right.$$

$$\geq 0$$

$$H(X) - H(X/Y) \geq 0$$

$$\Rightarrow \boxed{H(X) \geq H(X/Y).}$$

Similarly $\boxed{H(Y) \geq H(Y/X)}$

$\Rightarrow$ ① $H(X/Y) = H(XY) - H(Y)$, $H(Y/X) = H(XY) - H(X)$

② $H(XY) = H(X) + H(Y)$, X & Y are independent events

③ $H(X) \geq H(X/Y)$, $H(Y) \geq H(Y/X)$

① Consider that the two sources S_1 & S_2 emit messages x_1, x_2, x_3 and y_1, y_2, y_3, with the joint probability $P(XY)$ as shown in below. Calculate $H(X)$, $H(Y)$, $H(XY)$, $X(X/Y)$, $H(Y/X)$.

$$P(XY) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 3/40 & 1/40 & 1/40 \\ x_2 & 1/20 & 3/20 & 1/20 \\ x_3 & 1/8 & 1/8 & 3/8 \end{array}$$

Sol Given joint probability

$$P(XY) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 3/40 & 1/40 & 1/40 \\ x_2 & 1/20 & 3/20 & 1/20 \\ x_3 & 1/8 & 1/8 & 3/8 \end{array}$$

$$P(x_i) = \sum_{j=1}^{n} p(x_i y_j)$$

let $i=1$ $\quad p(x_1) = \sum_{j=1}^{3} p(x_i y_j) = p(x_1 y_1) + p(x_1 y_2) + p(x_1 y_3)$

$= 3/40 + 1/40 + 1/40 = 5/40 = 1/8$. $\quad \therefore \boxed{p(x_1) = 1/8}$

$i=2$ $\quad p(x_2) = \sum_{j=1}^{3} p(x_i y_j) = p(x_2 y_1) + p(x_2 y_2) + p(x_2 y_3)$

$= 1/20 + 3/20 + 1/20 = 5/20 = 1/4$ $\quad \therefore \boxed{p(x_2) = 1/4}$

$i=3$ $\quad p(x_3) = \sum_{j=1}^{3} p(x_i y_j) = p(x_3 y_1) + p(x_3 y_2) + p(x_3 y_3)$

$= 1/8 + 1/8 + 3/8 = 5/8$ $\quad \therefore \boxed{p(x_3) = 5/8}$

$$P(y_j) = \sum_{i=1}^{m} p(x_i y_j)$$

let $j=1$ $\quad p(y_1) = \sum_{i=1}^{3} p(x_i y_j) = p(x_1 y_1) + p(x_2 y_1) + p(x_3 y_1)$

$= 3/40 + \frac{1}{20} + \frac{1}{8} = \frac{3+2+5}{40} = \frac{1}{4}$ $\quad \therefore \boxed{p(y_1) = \frac{1}{4}}$

$j=2$ $\quad p(y_2) = \sum_{i=1}^{3} p(x_i y_j) = p(x_1 y_2) + p(x_2 y_2) + p(x_3 y_2)$

$= 1/40 + 3/20 + \frac{1}{20} = \frac{1+6+5}{40} = 3/10$ $\quad \therefore \boxed{p(y_2) = 3/10}$

$j=3$ $\quad p(y_3) = \sum_{j=1}^{3} p(x_i y_j) = p(x_1 y_3) + p(x_2 y_3) + p(x_3 y_3)$

$= \frac{1}{40} + \frac{1}{20} + \frac{3}{8} = \frac{1+2+15}{40} = \frac{9}{20}$ $\quad \therefore \boxed{p(y_3) = 9/20}$

$\therefore$ Entropy $\quad H(X) = -\sum_{i=1}^{m} p(x_i) \cdot \log_2 p(x_i)$

$= -\left[p(x_1) \log p(x_1) + p(x_2) \log p(x_2) + p(x_3) \log p(x_3) \right]$

$= -\left[\frac{1}{8} \log 1/8 + 1/4 \log 1/4 + 5/8 \log 5/8 \right] / \log 2$

$= 1.2988$

$\therefore \boxed{H(X) = 1.2988 \text{ bits/msg}}$

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

$= -\left[\frac{1}{4}\log 1/4 + 3/10\log 3/10 + 9/20\log 9/20\right] / \log 2$

$= 1.5394$

$\therefore \boxed{H(Y) = 1.5394 \text{ bits/msg.}}$

joint Entropy

$H(XY) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(x_i y_j)$

$= -\left[3/40\log 3/40 + 1/40\log 1/40 + \frac{1}{40}\log 1/40 + \frac{1}{20}\log 1/20 + 3/20\log 3/20 + \frac{1}{20}\log 1/20 + \frac{1}{8}\log 1/8 + \frac{1}{8}\log 1/8 + 3/8\log 3/8\right]$

$= 2.67$ $\therefore \boxed{H(XY) = 2.67 \text{ bits/msg.}}$

H(X/Y) :

Using Baye's theorem $p(x/y) = \frac{p(xy)}{p(y)}$

ie First column of $p(XY)$ is divided by $p(y_1)$, $1/4$
Second column of $p(XY)$ is divided by $p(y_2)$ & $3/10$
Third column of $p(XY)$ is divided by $p(y_3)$. $9/20$

$$\therefore P(X/Y) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 3/10 & 1/12 & 1/18 \\ x_2 & 1/5 & 1/2 & 1/9 \\ x_3 & 1/2 & 5/12 & 5/6 \end{array}$$

$3/40 \times 4 = 3/10$
$3/20 \times 4 = 1/5$
$1/8 \times 4 = 1/2$
$\frac{1}{40} \times \frac{10}{3} = \frac{1}{12}$ $\frac{1}{40} \times \frac{20}{9} = \frac{1}{18}$
$3/20 \times \frac{10}{3} = \frac{1}{2}$ $\frac{1}{20} \times \frac{20}{9} = \frac{1}{9}$
$1/8 \times \frac{10}{3} = 5/12$ $\frac{3}{8} \times \frac{20}{9} = 5/6$

$\therefore H(X/Y) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i y_j) \log[p(x_i/y_j)]$

$= -\left[3/40\log 3/10 + 1/40\log 1/12 + 1/40\log 1/18 + 1/20\log 1/5 + 3/20\log 1/2 + \frac{1}{20}\log 1/9 + 1/8\log 1/2 + 1/8\log 5/12 + 3/8\log 5/6\right]$

$= 1.12$ $\therefore \boxed{H(X/Y) = 1.12 \text{ bits/msg.}}$

H(Y/X) :

Using Baye's theorem $P(Y/X) = \frac{P(XY)}{P(X)}$

ie First row of $p(XY)$ is divided by $p(x_1) \rightarrow 1/8$
Second row of $p(XY)$ is divided by $p(x_2) \Rightarrow 1/4$
Third row of $p(XY)$ is divided by $p(x_3) \rightarrow 5/8$

$$P(Y/X) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 3/5 & 1/5 & 1/5 \\ x_2 & 1/5 & 3/5 & 1/5 \\ x_3 & 1/5 & 1/5 & 3/5 \end{array}$$

$3/40 \times 8 = 3/5$
$1/40 \times 8 = 1/5$, $\frac{1}{40} \times 8 = 1/5$
$1/20 \times 4 = 1/5$
$3/20 \times 4 = 3/5$
$1/20 \times 4 = 1/5$
$\frac{1}{8} \times \frac{8}{5} = 1/5$
$\frac{3}{8} \times \frac{8}{5} = 3/5$

Conditional Entropy

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

$$= -\left[\tfrac{3}{40}\log 3/5 + \tfrac{1}{40}\log 1/5 + \tfrac{1}{40}\log 1/5 + \tfrac{1}{20}\log 1/5 + \tfrac{3}{20}\log 3/5 + \tfrac{1}{20}\log 1/5 + \tfrac{1}{8}\log 1/5 + \tfrac{1}{8}\log 1/5 + \tfrac{3}{8}\log 3/5\right]$$

$$= 1.369$$

$$\therefore \boxed{H(Y/X) = 1.369 \text{ bits/msg.}}$$

(OR)

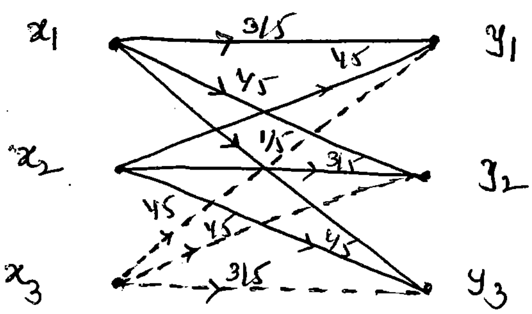
$\therefore$ $H(X) = 1.2988$
$H(Y) = 1.5394$
$H(XY) = 2.67$

$$\therefore H(X/Y) = H(XY) - H(Y)$$
$$= 2.67 - 1.5394$$
$$= 1.13$$
$$\therefore \boxed{H(X/Y) = 1.13 \text{ bits/msg}}$$

Ily

$$H(Y/X) = H(XY) - H(X)$$
$$= 2.67 - 1.2988$$
$$= 1.37$$
$$\therefore \boxed{H(Y/X) = 1.37 \text{ bits/msg}}$$

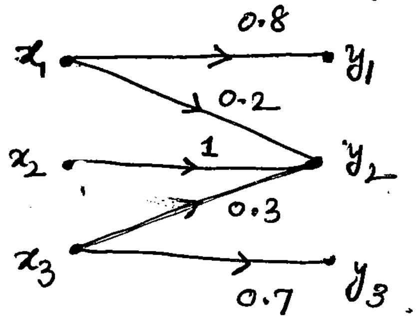
The matrix $p(y/x)$ is also known as channel matrix (or) noise matrix & graphically represented as



Note: Sum of the probabilities $= 1$

$$\sum_{i=1}^{m} P(x_i) = 1 \quad ; \quad \sum_{j=1}^{n} p(y_j) = 1$$

② A discrete source transmitter messages x_1, x_2, x_3 with the probabilities 0.3, 0.4, 0.3 respectively. The source is connected to the channel given in fig. Calculate all the Entropies.



Sol. Given

$$p(x_1) = 0.3$$
$$p(x_2) = 0.4$$
$$p(x_3) = 0.3$$

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

$\therefore$ $P(X) = [0.3 \;\; 0.4 \;\; 0.3]$

Using Baye's theorem $P(Y/X) = \dfrac{P(XY)}{P(X)} \Rightarrow P(XY) = P(Y/X)\cdot P(X)$

$$\therefore \quad P(XY) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 0.24 & 0.06 & 0 \\ x_2 & 0 & 0.4 & 0 \\ x_3 & 0 & 0.09 & 0.21 \end{array} \qquad \therefore \sum P(XY) = 1.$$

Entropy

$$H(X) = -\sum_{i=1}^{m} p(x_i)\log p(x_i)$$
$$= -\left[p(x_1)\log p(x_1) + p(x_2)\log p(x_2) + p(x_3)\log p(x_3)\right]$$
$$= -\left[0.3\log 0.3 + 0.4\log 0.4 + 0.3\log 0.3\right]$$
$$= 1.57 \qquad \therefore \boxed{H(X) = 1.57 \text{ bits/msg}}$$

Entropy

$$H(Y) = -\sum_{j=1}^{n} p(y_j)\log p(y_j)$$

$$\Rightarrow \quad P(y_j) = \sum_{i=1}^{3} P(x_i y_j)$$

let $j=1 \Rightarrow P(y_1) = p(x_1 y_1) + p(x_2 y_1) + p(x_3 y_1)$

$$= 0.24 + 0 + 0 = 0.24 \qquad \therefore \boxed{P(y_1) = 0.24}$$

$j=2 \Rightarrow p(y_2) = p(x_1 y_2) + p(x_2 y_2) + p(x_3 y_2)$

$$= 0.06 + 0.4 + 0.09 = 0.55 \qquad \therefore \boxed{P(y_2) = 0.55}$$

$j=3 \Rightarrow p(y_3) = P(x_1 y_3) + P(x_2 y_3) + P(x_3 y_3)$

$$= 0 + 0 + 0.21 = 0.21 \qquad \therefore \boxed{P(y_3) = 0.21}$$

$$\therefore \quad P(Y) = [0.24 \;\; 0.55 \;\; 0.21]$$

Entropy $H(Y) = -\left[p(y_1)\log p(y_1) + p(y_2)\log p(y_2) + p(y_3)\log p(y_3)\right]$

$$= -\left[0.24\log 0.24 + 0.55\log 0.55 + 0.21\log 0.21\right] / \log 2$$

$$\doteq 1.44 \quad \therefore \boxed{H(Y) = 1.44 \text{ bits/msg}}$$

Joint Entropy

$$H(XY) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\,\log p(x_i y_j)$$

$$= -\left[0.24\log(0.24) + 0.06\log(0.24) + 0.4\log(0.4) + 0.09\log(0.09) + 0.21\log(0.21)\right]$$

$$= 2.051$$

$$\therefore \boxed{H(XY) = 2.051 \text{ bits/msg}}$$

We know

$$H(X/Y) = H(XY) - H(Y)$$

$$= 2.051 - 1.44 = 0.611 \quad \therefore \boxed{H(X/Y) = 0.611 \text{ bits/msg}}$$

$$H(Y/X) = H(XY) - H(X)$$

$$= 2.051 - 1.57 = 0.481 \quad \therefore \boxed{H(Y/X) = 0.481 \text{ bits/msg}}$$

(OR)

$$H(Y/X) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i y_j)\log p(y_j/x_i)$$

$$= -\left[0.24\log 0.8 + 0.06\log 0.2 + 0.4\log 1 + 0.09\log 0.3 + 0.21\log 0.7\right]$$

$$= -\left[-0.077 - 0.139 - 0 - 0.156 - 0.108\right]$$

$$= 0.48 \quad \therefore \boxed{H(Y/X) = 0.48 \text{ bits/msg}}$$

$$H(X/Y) = -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i y_j)\log p(x_i/y_j)$$

$$P(X/Y) = \begin{array}{c|ccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 1 & 0.109 & 0 \\ x_2 & 0 & 0.727 & 0 \\ x_3 & 0 & 0.163 & 1 \end{array}$$

$$= -\left[0.24\log(1) + 0.06\log(0.1) + 0.4\log 0.7 + 0.09\log 0.163 + 0.21\log(1)\right]$$

$$= -\left[0 - 0.199 - 0.206 - 0.235 + 0\right]$$

$$= 0.561 \quad \therefore \boxed{H(X/Y) = 0.612 \text{ bits/msg}}$$

Hence

$H(X) = 1.57$ bits/msg

$H(Y) = 1.44$ bits/msg

$H(XY) = 2.05$ bits/msg

$H(X/Y) = 0.611$ bits/msg

$H(Y/X) = 0.48$ bits/msg.

③ A transmitter has an aphabet of four letters $[x_1\ x_2\ x_3\ x_4]$ and the receiver has an alphabet of three letters $[y_1\ y_2\ y_3]$. The joint probability matrix is

$$P(XY) = \begin{array}{c|c} & \begin{matrix} y_1 & y_2 & y_3 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 0.3 & 0.05 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0.15 & 0.05 \\ 0 & 0.05 & 0.15 \end{bmatrix} \end{array}$$

calculate all the Entropies.

Sol Given joint probability

$$P(XY) = \begin{array}{c|c} & \begin{matrix} y_1 & y_2 & y_3 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 0.3 & 0.05 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0.15 & 0.05 \\ 0 & 0.05 & 0.15 \end{bmatrix} \end{array}$$

$\therefore\ P(X) = [\,0.35 \quad 0.25 \quad 0.2 \quad 0.2\,]$

$P(Y) = [\,0.3 \quad 0.5 \quad 0.2\,]$

$$\therefore\ H(X) = -\sum_{i=1}^{4} p(x_i) \log p(x_i)$$

$$= -[0.35 \log 0.35 + 0.25 \log 0.25 + 0.2 \log 0.2 + 0.2 \log 0.2]$$

$$= 1.96 \qquad \therefore \boxed{H(X) = 1.96 \text{ bits/msg}}$$

$$H(Y) = \sum_{j=1}^{3} p(y_j) \log p(y_j)$$

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

$$= 1.49 \qquad \therefore \boxed{H(Y) = 1.49 \text{ bits/msg}}$$

$$H(XY) = -\sum_{i=1}^{4} \sum_{j=1}^{3} p(x_i y_j) \log p(x_i y_j)$$

$$= -[0.3 \log 0.3 + 0.05 \log 0.05 + 0.25 \log 0.25 + 0.15 \log 0.15 + 0.05 \log 0.05 + 0.05 \log(0.05) + 0.15 \log(0.15)]$$

$$= 2.49$$

$$\therefore \boxed{H(XY) = 2.49 \text{ bits/msg.}}$$

$$\therefore\ H(X/Y) = H(XY) - H(Y)$$

$$= 2.49 - 1.49 = 1.00 \qquad \therefore \boxed{H(X/Y) = 1 \text{ bit/msg}}$$

$$H(Y/X) = H(XY) - H(X)$$

$$= 2.49 - 1.96$$

$$= 0.53 \qquad \therefore \boxed{H(Y/X) = 0.53 \text{ bit/msg.}}$$

Mutual Information :

✓ The transfer of information from a transmitter through a channel to a receiver.

✓ The information provided about the event 'x_i' by the reception of the event 'y_j' is known as mutual Information & it is denoted as $I(x_i y_j) \Rightarrow I(XY)$.

* Before 'y_j' is received the uncertainity is $-\log p(x_i)$. is called <u>priori probability</u>.

* After 'y_j' is received the uncertainity becomes $-\log p(x_i/y_j)$ is called <u>posteriori probability</u>.

ie " The information gained about 'x_i' by the reception of 'y_j' is the net reduction in its uncertainity and is known as <u>mutual Information</u> ".

$$\therefore \; I(x_i y_j) = \text{initial uncertainity} - \text{Final uncertainity}$$

$$= -\log p(x_i) - [-\log p(x_i/y_j)]$$

$$= \log p(x_i/y_j) - \log p(x_i)$$

$$\boxed{I(x_i y_j) = \log\left[\frac{p(x_i/y_j)}{p(x_i)}\right]}$$

$$I(x_i y_j) = \log\left[\frac{p(x_i y_j)}{p(x_i)\, p(y_j)}\right] \longrightarrow (1)$$

$\because$ Baye's theorem $p(x/y) = \dfrac{p(xy)}{p(y)}$.

lly

$$I(y_j x_i) = \log\left[\frac{p(x_i y_j)}{p(x_i)\, p(y_j)}\right] \longrightarrow (2)$$

From eqn (1) & (2) the mutual information is symmetrical in 'x_i', and 'y_j'.

$$\therefore \; I(x_i y_j) = I(y_j x_i).$$

Self Information:

Self Information may be treated as a special case of mutual information when $y_j = x_i$ thus.

$$I(x_i\, x_i) = \log \frac{p(x_i/x_i)}{p(x_i)} = \log \frac{1}{p(x_i)} = I(x_i)$$

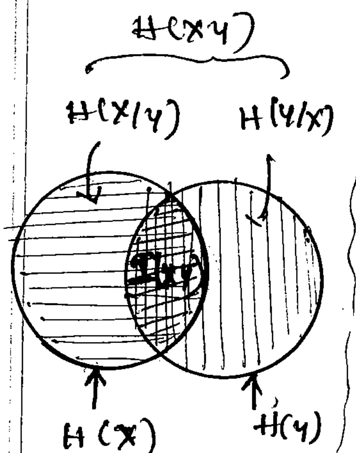
$\therefore$ Self Information $\underline{I(x_i\, x_i) = I(x_i)}$.

$\because$ The average of mutual Information ie, the entropy corresponding to <u>mutual information</u> is given by

$$I(XY) = \overline{I(x_i\, y_j)}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \cdot I(x_i\, y_j)$$

$$\boxed{I(XY) = \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \cdot \log \left[\frac{p(x_i/y_j)}{p(x_i)} \right]}$$



$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \left[\log p(x_i/y_j) - \log p(x_i) \right]$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \log p(x_i/y_j) - \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \log p(x_i)$$

$$= -\sum_{i=1}^{m} \Big[\underbrace{\sum_{j=1}^{n} p(x_i\, y_j)}_{p(x_i)} \Big] \log p(x_i) - \left[- \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i\, y_j) \log p(x_i/y_j) \right]$$

$$= -\sum_{i=1}^{m} p(x_i) \log p(x_i) - \left[H(X/Y) \right]$$

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

$\therefore$ $\boxed{I(XY) = H(X) - H(X/Y)} \rightarrow (3)$

lly $I(YX) = H(Y) - H(Y/X)$

$\therefore$ $\boxed{I(YX) = H(Y) - H(Y/X)} \rightarrow (4)$

We know

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

& $H(XY) = H(Y) + H(X/Y)$ $\quad \therefore H(X/Y) = H(XY) - H(Y)$

$\therefore$ $I(XY) = H(X) - H(XY) + H(Y)$

$\therefore$ Mutual Information $\boxed{I(XY) = H(X) + H(Y) - H(XY)} \rightarrow (5)$

Hence $I(XY)$ does not depends on the individual symbols x_i & y_j and it is a property of the entire communication system.

✓ eqn ③, ④ & ⑤ states that the entropy corresponding to mutual information.

✓ $I(XY) = H(X) - H(X/Y)$

where $H(X/Y)$ - the average information loss in the channel. This is also known as '<u>Equivocation</u>'

∴ Mutual Information = Source information − loss of Information.

lly

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

Mutual Information = Receiver Entropy − loss of Information

∴ For <u>noise free</u> channel

ie $\boxed{H(X/Y) = H(Y/X) = 0.}$

$$I(XY) = H(X) = H(Y) \longrightarrow ⑥$$

Eqn ⑥ states that the mutual information is equal to the source entropy and also equal to receiver Entropy. Thus the information transmitted at source is completely received by the receiver through a channel without any loss of information.

Properties of Mutual Information:

① The mutual information of a channel is symmetry

ie $\boxed{I(XY) = I(YX)}$

② Mutual information is always non negative

ie $I(XY) \geq 0$.

<u>Proof</u>: We know that

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

(or) $I(XY) = \sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i y_j) \cdot \log \left[\frac{P(x_i/y_j)}{p(x_i} \right]$

By using log inequality ie $\log x \geq \left(1 - \frac{1}{x}\right)$.

$$I(XY) \geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)\left[1 - \frac{p(x_i)}{p(x_i/y_j)}\right]$$

$$\geq \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) - \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{p(x_i y_j)\cdot p(x_i)}{p(x_i/y_j)}.$$

Apply Baye's theorem

$$p(XY) = p(X)\cdot p(Y/X)$$

$$p(XY) = p(Y)\cdot p(X/Y) \Rightarrow p(XY)/p(X/Y) = p(Y)$$

$$\therefore I(XY) \geq \underbrace{\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j)}_{1} - \sum_{i=1}^{m}\sum_{j=1}^{n} p(y_j)\cdot p(x_i)$$

$$\geq 1 - \underbrace{\sum_{i=1}^{m} p(x_i)}_{1}\cdot \underbrace{\sum_{j=1}^{n} p(y_j)}_{1}.$$

$$\geq 1-1$$

$$\geq 0 \qquad \therefore \boxed{I(XY) \geq 0.}$$

(3) Mutual information is related to entropies.

ie

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

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

$$I(XY) = H(X) + H(Y) - H(XY)$$

(4) Maximum value of mutual information is called channel capacity

ie $\boxed{C = \text{Max.}\,[I(XY)]}$

(5) The channel efficiency (or) Transmission efficiency is defined as The ratio of actual trans information $I(XY)$ to the maximum transformation (C). ie

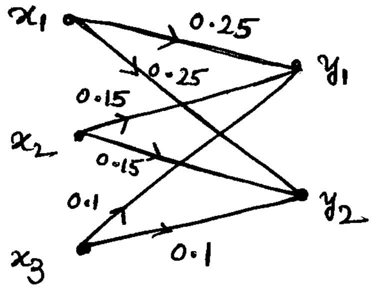
$$\eta = \frac{I(XY)}{\text{Max}\,(I(XY))} \Rightarrow \boxed{\eta = \frac{I(XY)}{C}}$$

(6) Redundancy: The redundancy of the channel is defined as

$$\boxed{\delta = 1-\eta} \Rightarrow \delta = 1 - \frac{I(XY)}{C} \Rightarrow \boxed{\delta = \frac{C - I(XY)}{C}}$$

(4) Find the mutual information for the channel shown in below.

Sol The joint probability matrix for the channel.



$$p(XY) = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \begin{bmatrix} 0.25 & 0.25 \\ 0.15 & 0.15 \\ 0.1 & 0.1 \end{bmatrix} \end{array}$$

$\therefore \; p(x_1) = 0.25 + 0.25 = 0.5$
$p(x_2) = 0.15 + 0.15 = 0.3$
$p(x_3) = 0.1 + 0.1 = 0.2$

$p(y_1) = 0.25 + 0.15 + 0.1 = 0.5$
$p(y_2) = 0.25 + 0.15 + 0.1 = 0.5$

$\therefore \; p(X) = [\,0.5 \quad 0.3 \quad 0.2\,]$ $\qquad p(Y) = [\,0.5 \quad 0.5\,]$

$$H(X) = -\sum_{i=1}^{m} p(x_i) . \log p(x_i)$$
$$= -\left[0.5 \log 0.5 + 0.3 \log 0.3 + 0.2 \log 0.2\right]$$
$$= 1.485 \text{ bits/msg} \qquad \therefore \boxed{H(X) = 1.485 \text{ bits/msg}}$$

$$H(Y) = -\sum_{j=1}^{2} p(y_j) \log p(y_j)$$
$$= -\left[0.5 \log(0.5) + 0.5 \log(0.5)\right]$$
$$= 1 \text{ bit/msg.} \qquad \therefore \boxed{H(Y) = 1 \text{ bit/msg}}$$

$$H(XY) = -\sum_{i=1}^{3} \sum_{j=1}^{2} p(x_i y_j) \log p(x_i y_j)$$
$$= -\left[0.25 \log(0.25) + 0.25 \log(0.25) + 0.15 \log(0.15) + 0.15 \log(0.15) + 0.1 \log(0.1) + 0.1 \log(0.1)\right].$$
$$= 2.485 \text{ bits/msg} \qquad \therefore \boxed{H(XY) = 2.485 \text{ bits/msg}}$$

$$H(X/Y) = H(XY) - H(Y)$$
$$= 2.485 - 1 = 1.485 \qquad \therefore \boxed{H(X/Y) = 1.485 \text{ bits/msg}}$$

$$H(Y/X) = H(XY) - H(X)$$
$$= 2.485 - 1.485 = 1 \qquad \therefore \boxed{H(Y/X) = 1 \text{ bit/msg}}$$

$\therefore$ Mutual information $I(XY) = H(X) + H(Y) - H(XY).$
$$= 1.485 + 1 - 2.485$$
$$= 2.485 - 2.485 = 0 \qquad \therefore \boxed{I(XY) = 0}$$

Hence

The channel shown above is with an independent input & output so mutual information becomes zero.

Note: A channel is said to be with an independent input and output when the joint probability matrix satisfies at least one of the following conditions
a) Each row consists of the same element.
b) Each column consists of the same element.

noisy channel

Discrete Memoryless channel:

* A discrete channel to be a system with input 'x' and output 'y'. The channel accepts input signal 'x' selected from an alphabet 'A'. and in response it delivers and output response 'y' from an alphabet 'B'.
* The probability transition matrix $P(Y/X)$ that express the probability of observing the output symbol 'y' given that the symbol 'x'.

✓ The channel is said to be discrete, both of the alphabets A & B have a finite length.

✓ It is said to be memoryless if the output depends only on the presence input and it is independent of previous inputs and outputs.

Channel Capacity:

Shannon has introduced a significant concept of Channels capacity & defined as the maximum of mutual information

(OR) The maximum rate at which the channel supplies the information to the receiver.

ie $$\boxed{C = \text{Max}.\, I(XY)} \text{ bits/message.}$$

where
$$I(XY) = H(X) - H(X/Y)$$
(or)
$$I(XY) = H(Y) - H(Y/X)$$
(or)
$$I(XY) = H(X) + H(Y) - H(XY)$$

✓ Transmission efficiency & Channel efficiency $\boxed{\eta = \frac{I(XY)}{C}}$

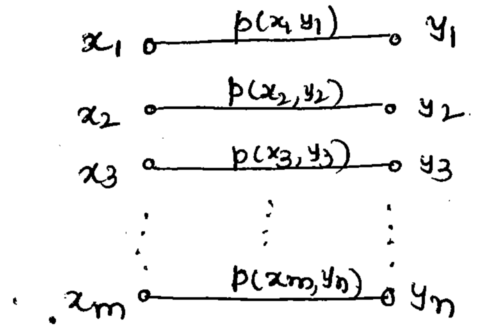
✓ The redundancy of the channel $\boxed{\delta = 1 - \eta}$

Classification of Channels :

① Noisy free Channel :

Noise free channel is also called noiseless binary channel, Noise free channel means there is one to one correspondance between input and output. ie. Each input symbol is received as one and only one output symbol.

ie



$$P(XY) = \begin{bmatrix} p(x_1 y_1) & 0 & 0 & \cdots & 0 \\ 0 & p(x_2 y_2) & 0 & \cdots & 0 \\ 0 & 0 & p(x_3 y_3) & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots 0 & \cdots 0 & \cdots & p(x_m y_n) \end{bmatrix}$$

For a noise free channel

$$H(X/Y) = H(Y/X) = 0.$$

channel capacity

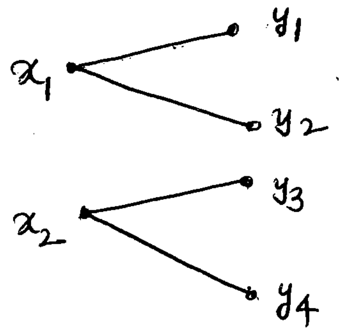
$$C = Max[I(XY)]$$
$$= Max[H(X) - \underbrace{H(X/Y)}_{0}]$$
$$= Max[H(X)]$$
$$C = \log_2 M$$

$$\therefore \boxed{C = \log_2 M} \text{ bits/msg.}$$

where M - No of messages.

② Noisy channel with non overlapping output :

This channel has two possible output corresponds to single input. The channel appears to be noisy but really is not.

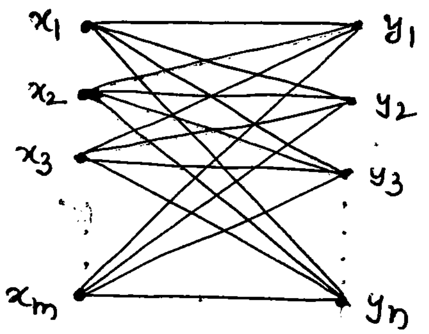


If x_1 is transmitted it may receved as y_1 and y_2.

Even though the output of channel is a random sequence of the input, the input can be determined from the output.

(3) Noisy channel with overlapping :

When the channel has noise it becomes difficult to reconstruct the transmitted signal faithfully.



$$P(XY) = \begin{bmatrix} p(x_1y_1) & p(x_1y_2) & \cdots & p(x_1y_n) \\ p(x_2y_1) & p(x_2y_2) & \cdots & p(x_2y_n) \\ \vdots & \vdots & & \vdots \\ p(x_my_1) & p(x_my_2) & \cdots & p(x_my_n) \end{bmatrix}$$

If the input and output symbol probabilities are statistically independent with each other.

$$p(x_i y_j) = p(x_i) \cdot p(y_j)$$

From Baye's theorem

$$p(x_i y_j) = p(x_i/y_j) \cdot p(y_j)$$

$$p(x_i)\, \cancel{p(y_j)} = p(x_i/y_j) \cdot \cancel{p(y_j)}$$

$$\therefore \quad p(x_i/y_j) = p(x_i)$$

lly

$$p(y_j/x_i) = P(y_j).$$

$$\therefore H(Y/X) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(y_j/x_i)$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \cdot \log p(y_j).$$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i) \cdot p(y_j) \cdot \log P(y_j)$$

$$= -\sum_{j=1}^{n} \underbrace{\left[\sum_{i=1}^{m} p(x_i)\right]}_{1} p(y_j) \cdot \log p(y_j).$$

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

$$\boxed{H(Y/X) = H(Y)} \qquad \therefore \boxed{\begin{aligned} H(Y/X) &= H(Y) \\ H(X/Y) &= H(X) \end{aligned}}$$

Similarly

$\therefore$ Mutual Information $I(XY) = H(X) - H(X/Y) = H(X) - H(X) = 0.$

Channel Capacity $C = Max[I(XY)] = 0$ $\therefore$ $\boxed{C = 0}$

ie No information is transmitted from the channel to the receiver.

④ Symmetric channel :

A symmetric channel is defined as the rows and columns of the channel matrix $P(Y/X)$ are identical except permitations (not in order).

(OR)

It is one for which ⓐ $H(Y/x_i)$ is independent of 'i' ie the entropy corresponding to each row of $P(Y/X)$ is the same.

ⓑ $\sum_{i=1}^{m} P(y_j/x_i)$ is independent of 'j' ie. The sum of all the collums of $P(Y/X)$ is the same.

Example : a) $P(Y/X) = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$

This is a symmetric channel as rows and columns are identical except for permutations. (variation of the order of the series).

b) $P(Y/X) = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \\ 1/2 & 1/4 & 1/4 \end{bmatrix}$ ⇒ This is not a symmetric channel as the rows are identical except for permutations but not the columns.

c) $P(Y/X) = \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \\ 0.6 & 0.4 \\ 0.7 & 0.3 \end{bmatrix}$ ⇒ This is not a symmetric channel as the columns are idetitical except for permutations but not the rows.

d) $P(Y/X) = \begin{bmatrix} 1/3 & 1/6 & 1/3 & 1/6 \\ 1/6 & 1/3 & 1/6 & 1/3 \end{bmatrix}$ ⇒ This is a symmetric channel as the row and columns are identical except for permutations.

e) $P(Y/X) = \begin{bmatrix} 1/3 & 1/6 & 1/3 & 1/6 \\ 1/3 & 1/3 & 1/6 & 1/6 \end{bmatrix}$ ⇒ This is not a symmetric channel as the rows and columns are not identical.

Channel Capacity for Symmetric channel :

Mutual information $I(XY) = H(Y) - H(Y/X)$

where $H(Y/X) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(y_j/x_i)$

$$= -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i) \cdot p(y_j/x_i) \log p(y_j/x_i)$$

($\because$ Bayes's theorem $p(XY) = p(Y/X) \cdot p(X)$)

$$= -\underbrace{\left[\sum_{i=1}^{m} p(x_i)\right]}_{1} \sum_{j=1}^{n} p(y_j/x_i) \log\left[p(y_j/x_i)\right]$$

$$= -\sum_{j=1}^{n} p(y_j/x_i) \log p(y_j/x_i)$$

$H(Y/X) = h$ Consider

where $h = -\sum_{j=1}^{n} p(y_j/x_i) \log p(y_j/x_i)$.

$\therefore$ $I(XY) = H(Y) - H(Y/X)$

$= H(Y) - h$

$\therefore$ The channel capacity $C = \text{Max}\,\left[I(XY)\right]$

$= \text{Max}\left[H(Y) - h\right]$

$= \text{Max}\left[H(Y)\right] - h$

$C = \log_2 M - h$ ($\because$ $\text{Max}(H(Y)) = \log_2 M$)

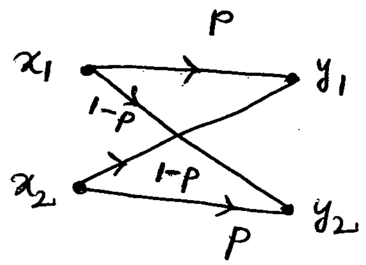
$\therefore$ $\boxed{C = \log_2 M - h}$

where $h = -\sum_{j=1}^{n} p(y_j/x_i) \log p(y_j/x_i)$.

Binary Symmetric Channel (BSC):

The most important case of symmetrical channel is Binary symmetric Channel (BSC).

Here $m = 2$, $n = 2$ (binary)



$\therefore$ The channel matrix $P(Y/X) = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$

The noise characteristics of the channel is

$p(y_1/x_1) = p \quad , \quad p(y_2/x_1) = 1-p, \quad p(y_1/x_2) = 1-p, \quad p(y_2/x_2) = p.$

$\therefore$ Channel capacity of binary symmetric channel

$$C = Max\ [\ I(X;Y)]$$

The channel capacity for symmetric channel is

$$C = \log_2 M - h$$

binary $M=2$

$$C = \log_2 2 - \left[- \sum_{j=1}^{n} p(y_j/x_i) \log p(y_j/x_i)\right]$$

$$= 1 + \sum_{j=1}^{2} p(y_j/x_i) \log p(y_j/x_i)$$

$$= 1 + p(y_1/x_i) \log p(y_1/x_i) + p(y_2/x_i) \log p(y_2/x_i)$$

$x_i = 1$ & 2

for $x=1$

$$= 1 + p\log p + (1-p)\log(1-p)$$

$$= 1 - \left[-\{p\log p + (1-p)\log(1-p)\}\right]$$

$$C = 1 - H(P)$$

$\therefore$ Channel capacity $\boxed{C = 1 - H(p)}$ where $H(p) = -\left[p\log p + (1-p)\log(1-p)\right]$

(5) Find the channel capacity for the Binary Symmetric Channel for the given data: (i) $p = 0.9$ (ii) $p = 0.6$.

Sol Binary symmetric channel $C = 1 - H(p) = 1 [-H(p)]$

(i) $H(p) = -\left[p\log p\ (1-p)\log(1-p)\right]$

$p = 0.9$

$1-p = 0.1$

$$= -\left[0.9\log(0.9) + 0.1\log(0.1)\right]$$

$$= -0.468$$

$\therefore \boxed{C = 0.532 \text{ bits/msg}}$

(ii) $p = 0.6$

$1-p = 0.4$

$$H(p) = -\left[0.6\log(0.6) + (0.4)\log(0.4)\right]$$

$$= -0.97$$

$$C = 1 + [-H(p)]$$

$$= 1 - 0.97$$

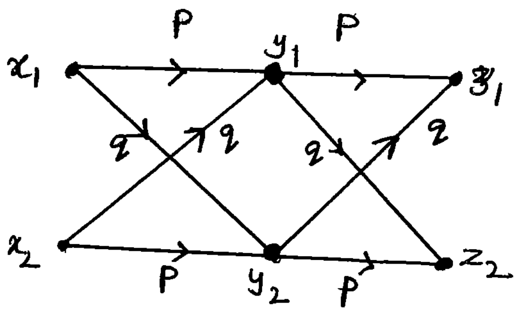
$$C = 0.03$$

$\therefore \boxed{C = 0.03 \text{ bits/msg}}$

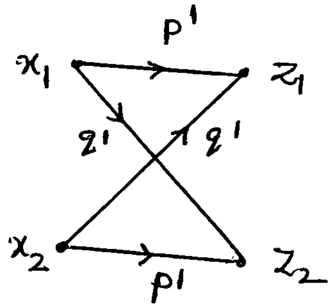
⑤ Cascaded Channel :

Some times channels are to be cascaded for some applications. Let us consider the code of two cascaded identical binary symmetric channel.

$q = 1 - p$



$\cong$



✓ The message from 'x_1' reaches to 'z_1' in two ways

$$x_1 \text{ to } z_1 \Rightarrow \quad x_1 - y_1 - z_1$$
$$x_1 - y_2 - z_1$$

∴ The respective path probabilities are $p \cdot p$ and $q \cdot q$.

$$\therefore \; p' = p \cdot p + q \cdot q = p^2 + q^2$$
$$= (p+q)^2 - 2pq$$
$$= [p + (1-p)]^2 - 2pq$$
$$= 1 - 2pq \qquad \therefore \boxed{p' = 1 - 2pq}$$

✓ The message from 'x_1' reaches to 'z_2' in two ways

$$x_1 \text{ to } z_2 \Rightarrow \quad x_1 - y_2 - z_2$$
$$x_1 - y_1 - z_2$$

The respective path probabilities are qp and $p.q$.

$$\therefore \; q' = q \cdot p + p \cdot q = 2pq \qquad \therefore \boxed{q' = 2pq}$$

$$\Rightarrow \quad p' + q' = 1 - 2pq + 2pq = 1 \qquad \therefore p' + q' = 1.$$

The channel matrix of the cascaded channel is

$$P(Z/X) = \begin{bmatrix} p' & q' \\ q' & p' \end{bmatrix} = \begin{bmatrix} 1-2pq & 2pq \\ 2pq & 1-2pq \end{bmatrix}$$

Thus the cascaded channel is equivalent to a single binary symmetric channel with error probability equal to '<u>2pq</u>'

∴ The channel capacity of a binary symmetric channel is

$$C = 1 - H(p) \qquad \text{where} \quad H(p) = -[p \log p + (1-p) \log(1-p)]$$

∴ The channel capacity of a cascaded channel is

$$C = 1 - H(q') \,, \; q' = 2pq \qquad \therefore \boxed{C = 1 - H(2pq)}.$$

(6) Binary Erasure Channel (BEC):

Binary erasure channel has two inputs (0 & 1) and three outputs (0, n, 1). BEC is also very important.

Here '0' and '1' are transmitted & they are received 0, n, 1.

The symbol 'n' indicates that, due to noise, no deterministic decision can be made as to whether the received symbol is a '0' or a '1'.

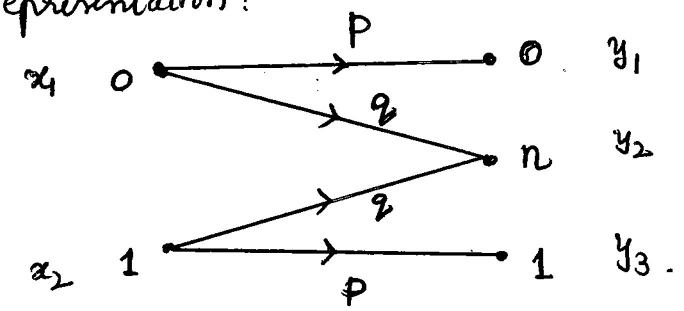
ie The symbol 'n' indicates that the output is erased.

Hence the name <u>Binary Erasure Channel</u>.

✓ If the output is 'n' then the receiver request the transmitter for retransmission till the decision is taken either the symbol '0' or '1'. This can be consider equivallent to error detection and requesting for retransmission.

ie <u>ARQ</u> (Automatic Repeat Request) method of error correction.

Graphical representation:



The channel matrix for binary Erasure channel is

$$p(Y|X) = \begin{array}{c c} & \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} p & q & 0 \\ 0 & q & p \end{bmatrix} \\ & \begin{array}{ccc} 0 & n & 1 \end{array} \end{array} \qquad 'q = 1-p'$$

Let $p(x_1) = \alpha$ and $p(x_2) = 1-\alpha$

∴ Entropy $H(X) = -\sum_{i=1}^{m} p(x_i) \log p(x_i)$

$$= -\left[p(x_1)\log p(x_1) + p(x_2)\log p(x_2)\right]$$

$$\therefore H(X) = -\left[\alpha \log \alpha + (1-\alpha)\log(1-\alpha)\right]$$

The joint probability matrix $p(XY)$ can be obtained by multiplying the rows of $P(Y/X)$ by 'α' and $(1-\alpha)$ respectively.

Since $p(x_1) = \alpha$, $p(x_2) = 1-\alpha$.

[$\because$ Baye's theorem $p(XY) = p(Y/X) \cdot p(X)$]

$$p(XY) = \begin{bmatrix} \alpha p & \alpha q & 0 \\ 0 & (1-\alpha)q & (1-\alpha)p \end{bmatrix}$$

The sum of the columns gives
$$p(y_1) = \alpha p$$
$$p(y_2) = \alpha q + q - \alpha q = q$$
$$p(y_3) = (1-\alpha)p.$$

The conditional probability matrix $p(X/Y)$ can be obtained by dividing the columns of $p(XY)$ by $p(y_1), p(y_2), p(y_3)$ respectively.

$$\therefore \quad P(X/Y) = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1-\alpha & 1 \end{bmatrix} \qquad \left[\because P(X/Y) = \frac{P(XY)}{P(Y)}\right]$$

$$\therefore H(X/Y) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i y_j) \log p(x_i/y_j)$$

$m=2$
$n=3$

$$= -\left[\alpha p \log(1) + \alpha q \log \alpha + 0 + 0 + (1-\alpha) q \log(1-\alpha) + (1-\alpha)p \log(1)\right]$$

$$= -\left[\alpha q \log \alpha + (1-\alpha) q \log(1-\alpha)\right]$$

$$= -q\left[\alpha \log \alpha + (1-\alpha)\log(1-\alpha)\right]$$

$$\therefore H(X/Y) = q \cdot H(X). \quad \text{(or)} \quad \boxed{H(X/Y) = (1-p)H(X)}$$

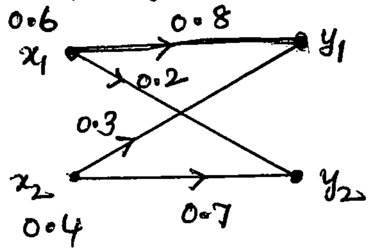
The channel capacity of Binary Erasure channel is

$$C = \text{Max}\left[I(XY)\right]$$
$$= \text{Max}\left[H(X) - H(X/Y)\right]$$
$$= \text{Max}\left[H(X) - (1-p)H(X)\right]$$
$$= \text{Max}\left[\cancel{H(X)} - \cancel{H(X)} + p \cdot H(X)\right] \qquad \boxed{I(XY) = p \cdot H(X)}$$
$$= \text{Max}\left[p \cdot H(X)\right]$$
$$= p. \text{Max}\left[H(X)\right]$$
$$= p. \log_2 M \qquad \left(\because \text{Max}\left[H(X)\right] = \log_2 M\right)$$
$$= p. \log_2 2 \qquad (\because \text{Binary } M=2)$$

$C = p$

$\therefore$ $\boxed{\text{Channel capacity for BEC, } C = p.}$

(6) Find the mutual information and channel capacity of the channel shown in fig below.



Sol Given
$$p(Y/X) = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$$

$p(x_1) = 0.6$, $p(x_2) = 0.4$

$$P(XY) = p(Y/X) \cdot P(X)$$

$$\therefore p(XY) = \begin{bmatrix} 0.48 & 0.12 \\ 0.12 & 0.28 \end{bmatrix}$$

$$\Rightarrow \quad p(y_1) = 0.48 + 0.12 = 0.6 \qquad \therefore p(y_1) = 0.6$$
$$p(y_2) = 0.12 + 0.28 = 0.4 \qquad p(y_2) = 0.4$$

$$P(X/Y) = \frac{P(XY)}{P(Y)} \Rightarrow p(X/Y) = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \quad \begin{bmatrix} 0.48/0.6 = 0.8 \\ 0.12/0.6 = 0.2 \\ 0.12/0.4 = 0.3 \\ 0.28/0.4 = 0.7 \end{bmatrix}$$

$$\therefore H(X) = -\sum_{i=1}^{2} p(x_i) \log p(x_i)$$
$$= -[0.6 \log 0.6 + 0.4 \log 0.4]$$
$$= 0.971 \text{ bits/msg}$$

$$\therefore \boxed{H(X) = 0.970 \text{ bits/msg}}$$

$$H(Y) = -\sum_{i=1}^{2} p(y_j) \log p(y_j)$$
$$= -[0.6 \log 0.6 + 0.4 \log 0.4]$$
$$= -0.970 \text{ bits/msg}$$

$$\therefore \boxed{H(Y) = 0.970 \text{ bits/msg}}$$

$$H(XY) = -\sum_{i=1}^{2} \sum_{j=1}^{2} p(x_i y_j) \log p(x_i y_j).$$
$$= -[0.48 \log(0.48) + 0.12 \log(0.12) + 0.12 \log(0.12) + 0.28 \log(0.28)]$$
$$= -[-0.508 - 0.367 - 0.367 - 0.514]$$
$$= 1.756 \text{ bits/msg}$$

$$\therefore \boxed{H(XY) = 1.756 \text{ bits/msg}}$$

$$H(X/Y) = H(XY) - H(Y) = 1.756 - 0.97 = 0.786 \qquad \therefore \boxed{H(X/Y) = 0.786 \text{ bits/msg}}$$

$$H(Y/X) = H(XY) - H(X) = 1.756 - 0.97 = 0.786 \qquad \therefore \boxed{H(Y/X) = 0.786 \text{ bits/msg}}$$

Mutual Information
$$I(XY) = H(X) + H(Y) - H(XY)$$
$$= 0.97 + 0.97 - 1.756 = 0.184 \qquad \therefore \boxed{I(XY) = 0.184 \text{ bits/msg.}}$$

channel capacity $\quad C = 1 - H(p)$

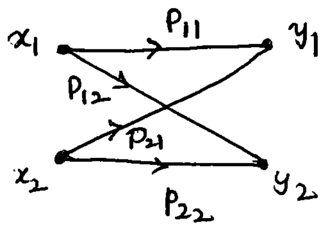
where $H(p) = -\{p \log p + (1-p) \log(1-p)\}$

for symmetric structures in channel ie $\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$

If Not perform as follows.

Note: A binary channel with non symmetric structures

$$D = P(Y/X) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$



To find the channel capacity of a binary channel, the auxillary variables Q_1 & Q_2 are defined by

$$[D][Q] = -[H]$$

$$\Rightarrow \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} P_{11}\log P_{11} + P_{12}\log P_{12} \\ P_{21}\log P_{21} + P_{22}\log P_{22} \end{bmatrix}$$

$$\Rightarrow P_{11}Q_1 + P_{12}Q_2 = P_{11}\log P_{11} + P_{12}\log P_{12} \longrightarrow (1)$$

$$P_{21}Q_1 + P_{22}Q_2 = P_{21}\log 21 + P_{22}\log P_{22} \longrightarrow (2)$$

By solving eqn (1) & (2), we get Q_1 and Q_2.

$\therefore$ The channel capacity $\boxed{C = \log(2^{Q_1} + 2^{Q_2})}$

From problem.

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} 0.8\log 0.8 + 0.2\log 0.2 \\ 0.3\log 0.3 + 0.7\log 0.7 \end{bmatrix}$$

$$\Rightarrow 0.8Q_1 + 0.2Q_2 = -0.721$$
$$0.3Q_1 + 0.7Q_2 = -0.881$$

$\rightarrow$ $a_1 = 0.8$ $b_1 = 0.2$, $c_1 = -0.721$
$a_2 = 0.3$ $b_2 = 0.7$ $c_2 = -0.881$

By solving above eqns $Q_1 = -0.656$
$Q_2 = -0.976$

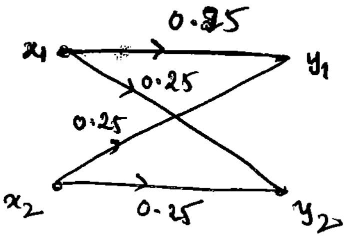
$\therefore$ The channel capacity

$$c = \log(2^{Q_1} + 2^{Q_2})$$
$$= \log\left[2^{-0.656} + 2^{-0.976}\right]$$
$$= \log(0.633 + 0.513)$$
$$= \log_2 1.146 = \log(1.146)/\log 2 = 0.196 \text{ bits/msg}$$
$$= 0.2$$

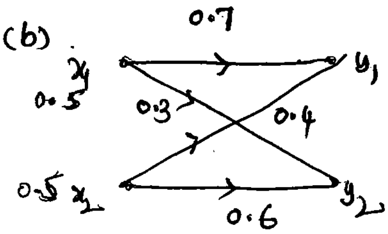
$\therefore$ $\boxed{C = 0.2 \text{ bits/message}}$

Problems.

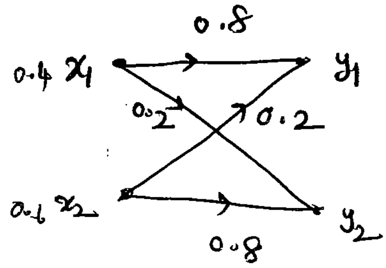
(a)



(b)



(c)



Source Coding :

Coding offers the most significant application of the information theory, the main purpose of the coding is to improve the efficiency of the communication system.

* A conversion of the output of discrete memoryless source (DMS) ie voice, image & data into a sequence of binary symbols (ie binary code word) is called 'Source coding' 1's and 0's.

✓ The device the performs this conversion is called 'Source encoder'

* Source coding is mainly used to minimize the average bit rate required for representation of the source by reducing the redundancy of the information source.

Code length & Coding Efficiency :

✓ Let X be the discrete memoryless source with an alphabet $(x_1, x_2, \cdots x_m)$ with corresponding probabilities $[p(x_1), p(x_2), \cdots p(x_m)]$.

✓ Let binary code word assigned to symbol/alphabet x_i by the encoder have length 'n_i' measured in bits.

"The length of the codeword is the no. of binary digits in the code word"

∴ Code length : The average length of the message (or) average length per code word is given by

$$\boxed{L = \sum_{i=1}^{m} p(x_i) \cdot n_i} \quad \text{Letters/msg.}$$

ie 'L' should be minimum to have efficient transmission.

Coding Efficiency :
$$\eta = \frac{L_{min}}{L}$$

where L_{min} - minimum possible value of 'L' when 'η' approaches unity, the code is said to be efficient.

Let $H(X)$ be the entropy of the source in bits/msg.

Let $\log_2 M$ be the maximum average information associated with each letter in bits/letter.

The minimum average no of letters per message

$$\boxed{L_{min} = \frac{H(X)}{\log_2 M}} \qquad \text{Units} = \frac{\text{bit/msg}}{\text{bits/letter}} = \text{letters/message}$$

$\therefore$ Hence

The coding efficiency $\eta = \frac{L_{min}}{L} = \frac{H(X)}{\log M} \Big/ L$

$$\therefore \boxed{\eta = \frac{H(X)}{L \cdot \log_2 M}}$$

and the redundancy is given by $\boxed{\delta = 1-\eta}$

Example:

Let $M = [m_1\ m_2\ m_3\ m_4]$

$P(M) = [\frac{1}{2}\ \frac{1}{4}\ \frac{1}{8}\ \frac{1}{8}]$

(a) Without coding & considering a one-one correspondance ie a noiseless channel, The efficiency is

$\eta = \frac{I(XY)}{C}$ → mutual information; C → channel capacity.

$I(XY) = H(X) - H(X/Y)$ → for noiseless channel $H(X/Y) = 0$.

$\therefore I(XY) = H(X)$ & $C = \text{Max}[I(XY)] = \text{Max}[H(X)] = \log_2 M$

$\therefore \eta = \frac{H(X)}{\log_2 M}$ $\qquad \therefore C = \log M$

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

$= -[\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{8}\log\frac{1}{8} + \frac{1}{8}\log\frac{1}{8}]$

$= 7/4$

$\log_2 M = \log_2 4 = \log_2 2^2 = 2\log_2 2 = 2$

$\therefore$ Efficiency $\eta = \frac{7/4}{2} = \frac{7}{8} = 0.875$ $\qquad \therefore \boxed{\%\eta = 87.5\%}$

By using Coding

Fixed length coding: A fixed length code is one whose codeword length is fixed. Ex: 8421 code.

b) Let us use binary code for coding

let the code letters be '0' and 1. So, $M = 2$

using fixed length coding.

Probabity	Message	Code	Length of Code
$p(m_1) = 1/2$	m_1	$c_1 = 00$	$n_1 = 2$
$p(m_2) = 1/4$	m_2	$c_2 = 01$	$n_2 = 2$
$p(m_3) = 1/8$	m_3	$c_3 = 10$	$n_3 = 2$
$p(m_4) = 1/8$	m_4	$c_4 = 11$	$n_4 = 2$

$\therefore$ Entropy $H(x) = -\sum_{i=1}^{m} p(x_i) \log p(x_i) = 7/4$ bits/msg.

length $L = \sum_{i=1}^{4} n_i \times p(x_i)$

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

$L = 2$ letters/msg $\qquad \therefore \log_2 M = \log_2 2 = 1$. bits/letter

$\therefore \frac{H(x)}{\log M} = \frac{7/4}{1} = \frac{7}{4}$ letters/msg.

$\therefore$ Efficiency $\eta = \frac{H(x)}{L \cdot \log M} = \frac{7/4}{2} = \frac{7}{8} = 87.5\%$

$\therefore \boxed{\% \eta = 87.5\%}$

By using fixed length coding procedure is not improving the efficieny.

Variable length Coding: A variable length code is one whose codeword is not fixed. ie Length of the code word is variable.

* There are two functional requirements for the source encoder to be Efficient.

a) Codeword should be binary form

b) Code is uniquely decodable.

Uniquely decodable: A distinct code uniquely decodable if the original source sequence can be reconstructed perfectly from the encoded binary sequence.

c) Let us consider in this code

message	Code	length of the Code	Probability
m_1	$c_1 = 0$	$n_1 = 1$	$p(m_1) = 1/2$
m_2	$c_2 = 10$	$n_2 = 2$	$p(m_2) = 1/4$
m_3	$c_3 = 110$	$n_3 = 3$	$p(m_3) = 1/8$
m_4	$c_4 = 111$	$n_4 = 3$	$p(m_4) = 1/8$

$\therefore \quad H(X) = -\sum_{i=1}^{4} p(x_i) \log p(x_i) = 7/4$ bits/msg.

$\log_2 M = \log_2 2 = 1 \qquad (\because M = 2$ uses '0' and '1' letters)

$\therefore \dfrac{H(X)}{\log M} = \dfrac{7}{4} = 7/4$ letters/msg.

Length $\quad L = \sum_{i=1}^{4} n_i \cdot p(x_i)$

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

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

$L = 7/4$ letters/msg.

$\therefore$ Efficiency $\quad \eta = \dfrac{H(X)}{L \cdot \log_2 M} = \dfrac{7/4}{7/4} = 1$

$\therefore \boxed{\% \eta = 100\%}$

Hence
Variable length coding is better than fixed length coding.

Note: Encode a message with high probability in a short-cord word. Only then the average length of the code word 'L' will decrease resulting in an increase in efficiency.

ie

$\longrightarrow$

$1/2 \to 0$
$1/4 \to 10$
$1/8 \to 110$
$1/8 \to 111$

Variable length coding is two types.

ⓐ Shannon Fano Coding

ⓑ Huffman coding.

Shannon-Fano Coding : ***

{Top- Bottom Approach}

- ✓ Shanon fano coding generates an efficient code in which "Word length increases as symbol probability decreases."
- ✓ Shannon fano coding is variable length code.
- ✓ It is a top-bottom approach
- ✓ The transmission of an encoded message is reasonably efficient. ie '1' and '0' appear independently, with almost equal probabilities.

Algorithm / Procedure :

1. List the source symbols / messages in order of decreasing probabilities.
2. The message set then is partitioned into two equi-probable subsets $[X_1]$ and $[X_2]$.
3. A '0' is assigned to each message contained in one subset ie upper set. & A '1' is assigned to each message contained in the other subset ie Lower set.
4. The same procedure is repeated for the subsets of $\overset{0}{[X_1]}$ and $\overset{1}{[X_2]}$. ie $[X_1]$ will be partitioned into two subsets $\overset{00}{[X_{11}]}$ and $\overset{01}{[X_{12}]}$; $[X_2]$ will be partitioned into two subsets $\overset{10}{[X_{21}]}$ and $\overset{11}{[X_{22}]}$.
5. The procedure is continued, each time partitioning the sets with as nearly equal probabilities as possible untill each subset contains only one message / further partitioning is not possible.
6. Calculate Entropy and length of the code to obtain Coding efficiency.

$$H(x) = -\sum_{i=1}^{m} p(x_i) \cdot \log p(x_i)$$

$$L = \sum_{i=1}^{m} p(x_i) \cdot n_i \qquad \therefore \eta = \frac{H(X)}{L \cdot \log M}$$

where $M = 2$ ('0' & '1')

(1) A source is transmitting messages A, B, C, D, E, F with corresponding probabilities $1/2, 1/4, 1/8, 1/16, 1/32, 1/32$. Find the coding efficiency using Shannon fano encoding method.

Sol: Shannon Fano Coding.

	Message (X)	Probability P(X)	Encoded Message I	II	III	IV	V	length of the code (n_i)
1.	A	$1/2$	0					1
2.	B	$1/4$	1	0				2
3.	C	$1/8$	1	1	0			3
4.	D	$1/16$	1	1	1	0		4
5.	E	$1/32$	1	1	1	1	0	5
6.	F	$1/32$	1	1	1	1	1	5

✓ The encoding alphabets 0 and 1 So $M=2$, $\log_2 M = \log_2 2 = 1$.

✓ The entropy $H(X) = -\sum_{i=1}^{6} p(x_i) \log p(x_i)$

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

$$= \frac{62}{32} = \frac{31}{16} = 1.9375$$

$\therefore \boxed{H(X) = 1.9375 \text{ bits/msg.}}$

✓ Coding length $L = \sum_{i=1}^{6} p(x_i) \cdot n_i$

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

$$= \frac{31}{16} = 1.9375 \quad \therefore \boxed{L = 1.9375 \text{ letters/msg}}$$

$$\frac{H(X)}{\log_2 M} = \frac{1.9375}{1} \frac{\text{bits/msg}}{\text{bits/letter}} = \text{letters/msg}$$

Coding efficiency $\% \eta = \dfrac{H(X)}{L \cdot \log_2 M} = \dfrac{1.9375}{1.9375\times 1} \times 100 = 100\%$

$\therefore \boxed{\eta = 100\%}$

Redundancy $(1-\eta) = \delta = 0$

$\therefore \boxed{\delta = 0}$

(2) Apply the Shannon-Fano coding procedure for the following message ensemble:

$$[X] = [x_1 \; x_2 \; x_3 \; x_4 \; x_5 \; x_6 \; x_7 \; x_8]$$
$$[P] = [1/4 \; 1/8 \; 1/16 \; 1/16 \; 1/16 \; 1/4 \; 1/16 \; 1/8]$$

Sol

Shannon-Fano Coding.

SL No	Message X	Probability	Encoded Message I	II	III	IV		Code	Length of the code (n_i)
1	x_1	$\frac{1}{4}=0.25$	0	0			→	00	2
2	x_6	$1/4=0.25$	0	1			→	01	2
3	x_2	$1/8=0.125$	1	0	0		→	100	3
4	x_8	$1/8=0.125$	1	0	1		→	101	3
5	x_3	$1/16=0.0625$	1	1	0	0	→	1100	4
6	x_4	$1/16=0.0625$	1	1	0	1	→	1101	4
7	x_5	$1/16=0.0625$	1	1	1	0	→	1110	4
8	x_7	$1/16=0.0625$	1	1	1	1	→	1111	4

✓ The Encoding alphabets/Letters '0' & '1' So, $M=2$, $\log_2^M = \log_2^2 = 1$ bit/letter.

✓ The Entropy $H(X) = -\sum_{i=1}^{8} p(x_i)\log p(x_i)$

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

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

$$= 22/8 = 2.75 \quad \therefore \boxed{H(X) = 2.75 \text{ bits/msg.}}$$

✓ Code length $L = \sum_{i=1}^{8} p(x_i)\cdot n_i.$

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

$$= 22/8 = 2.75 \quad \therefore \boxed{L = 2.75 \text{ letters/msg}}$$

$$\frac{H(X)}{\log M} = \frac{2.75}{1} = 2.75 \text{ letters/msg}$$

∴ Coding efficiency $\eta = \frac{H(X)}{L\cdot\log_2^M} = \frac{2.75}{2.75\times 1}\times 100 = 100\%$ $\boxed{\therefore \eta = 100\%}$

Redundancy $\delta = 1-\eta = \delta = 0 \Rightarrow \boxed{\delta = 0}$

③ Apply the Shannon-Fano Coding procedure for the following message ensemble.

(a) Take $\boxed{M=2}$ (b) Take $\boxed{M=3}$

$$[X] = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7]$$

$$[P] = [0.4 \quad 0.2 \quad 0.12 \quad 0.08 \quad 0.08 \quad 0.08 \quad 0.04]$$

Sol (a) $\boxed{M=2}$ The encoding alphabets / letters are '0' and '1' _ $\log_2 M = \log_2 2 = 1$. bit / letter

Shannon Fano Coding.

SL No	Message (X_i)	Probability $P(x_i)$	Encoded Message I	II	III	IV		Code	Length of the code (n_i)
1	x_1	0.4	0				→	0	1
2	x_2	0.2	1	0	0		→	100	3
3	x_3	0.12	1	0	1		→	101	3
4	x_4	0.08	1	1	0	0	→	1100	4
5	x_5	0.08	1	1	0	1	→	1101	4
6	x_6	0.08	1	1	1	0	→	1110	4
7	x_7	0.04	1	1	1	1	→	1111	4

Coding length $L = \sum_{i=1}^{7} p(x_i) \cdot n_i$

$$= 0.4\times1 + 0.2\times3 + 0.12\times3 + 0.08\times4 + 0.08\times4 + 0.08\times4 + 0.04\times4$$

$$= 2.48 \text{ letters/msg.} \quad \therefore \boxed{L = 2.48 \text{ letters/msg}}$$

Entropy $H(x) = -\sum_{i=1}^{7} p(x_i) \log p(x_i)$

$$= -[0.4\log 0.4 + 0.2\log 0.2 + 0.12\log 0.12 + 0.08\log 0.08 + 0.08\log 0.08 + 0.08\log 0.08 + 0.04\log 0.04]$$

$$= 2.42 \text{ bits/msg} \quad \therefore \boxed{H(x) = 2.42 \text{ bits/msg}}$$

$$\therefore \frac{H(X)}{\log_2 M} = \frac{2.42}{1} = 2.42 \text{ letters/msg.}$$

$$\therefore \text{Coding efficiency } \eta = \frac{H(X)}{L \cdot \log_2 M} = \frac{2.42}{2.48} \times 100 = 97.6\%$$

$$\boxed{\%\eta = 97.6\%}$$

Redundancy, $\delta = 1 - \eta = \boxed{2.4\%}$

(b) $M=3$. ie let the encoding alphabet be 0, 1, 2.

SL No	Message (X_i)	Probability $P(X_i)$	Encoded Message / Code I	II	III	Code	Length of the code (n_i)
1	x_1	0.4	0			→ 0	1
2	x_2	0.2	1	0		→ 10	2
3	x_3	0.12	1	1		→ 11	2
4	x_4	0.08	2	0		→ 20	2
5	x_5	0.08	2	1		→ 21	2
6	x_6	0.08	2	2	0	→ 220	3
7	x_7	0.04	2	2	1	→ 221	3

Coding Length $L = \sum_{i=1}^{7} p(x_i)\cdot n_i$

$= 0.4\times1 + 0.2\times2 + 0.12\times2 + 0.08\times2 + 0.08\times2 + 0.08\times3 + 0.04\times3$

$= 1.72$ letters/message.

$\therefore \boxed{L = 1.72 \text{ letters/msg}}$

Entropy $H_{\log} = -\sum_{i=1}^{7} p(x_i)\log p(x_i)$

$\therefore \boxed{H(x) = 2.42 \text{ bits/msg.}}$

$\therefore$ Coding Efficiency $\therefore \eta = \dfrac{H(x)}{L\cdot\log_2 M} = \dfrac{2.42}{1.72\times\log_2 3} = 88.7\%$

$\therefore \boxed{\%\eta = 88.7\%}$

Redundancy $\delta = 1-\eta = 0.113$

$\therefore \boxed{\delta = 11.3\%}$

Problems:

① Find the coding efficiency for the following data using coding techniques

(i) $[x] = [x_1\ x_2\ x_3\ x_4\ x_5\ x_6\ x_7\ x_8]$

$[P] = [0.1\ 0.25\ 0.15\ 0.05\ 0.15\ 0.1\ 0.05\ 0.15]$

(ii) $[X] = [x_1\ x_2\ x_3\ x_4\ x_5]$

$[P] = [0.4\ 0.15\ 0.15\ 0.15\ 0.15]$

(a) Fixed length coding.

(b) Shannon-Fano coding procedure for M=2 and M=3

(c) Huffman coding procedure for M=2 and M=3.

Solutions for Page: 30

② $[P] = [\frac{1}{2}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{16}\ \frac{1}{16}\ \frac{1}{16}\ \frac{1}{32}\ \frac{1}{32}]$ for M=2, a) Shannon Fano Coding b) Huffman coding.

8-messages.

③ $[P] = [0.27\ 0.20\ 0.17\ 0.16\ 0.06\ 0.06\ 0.04\ 0.04]$ a) Shanon Fano coding b) Huffman coding.

8 messages

** Huffman Coding: (Bottom-top approach)

✓ An another type of source code is "Huffman code" which leads to the lowest possible value of 'L' for a given Messages 'M', resulting in a maximum efficiency & minimum redundancy. Hence it is also known as "minimum redundancy code" (or) optimum code.

Algorithm / Procedure:

1. Arrange the probabilities in the order of descending.
2. Combine the last two probabilities into one probability by adding their probabilities.
3. Rearrange the probabilities in the descending order.
4. Combine the last two probabilities into one probability.
5. Repeat the procedure until no of messages is reduced to two elements.
6. Assign '0' (up) and '1' (down) to this last two messages as their first digit in the code sequence.
7. Go back to assign the numbers '0' and '1' to second digit for the two messages that were combine in previous steps.
8. Keep this way until first column is reached.
9. The first column coding gives the final Optimal Code.
10. Calculate entropy $H(X)$, lengths of the code (L) to obtain Coding efficiency $\boxed{\eta = \frac{H(X)}{L \cdot \log_2 M}}$ and Redundancy $\boxed{\delta = 1-\eta}$

(1) A source is trasmitting six messages with probabilities 0.3, 0.25, 0.15, 0.12, 0.10, 0.08 respectively. Find the efficiency of the code using Huffman encoding technique.

Sol Given 0.3, 0.25, 0.15, 0.12, 0.10, 0.08.

Huffman encoding:

Length n_i	Probability		First reduction	Second reduction	Third reduction	Fourth reduction
2 ←	0.3	00	0.3 00	0.3 00	0.43 1	0.57 0
2 ←	0.25	10	0.25 10	0.27 01	0.3 00	0.43 1
3 ←	0.15	010	0.18 11	0.25 10	0.27 01	
3 ←	0.12	011	0.15 010	0.18 11		
3 ←	0.10	110	0.12 011			
3 ←	0.08	111				

Length of the code $L = \sum_{i=1}^{6} p(x_i) \cdot n_i$

$= 0.3 \times 2 + 0.25 \times 2 + 0.15 \times 3 + 0.12 \times 3 + 0.10 \times 3 + 0.08 \times 3$

$L = 2.45$ letters/msg.

$\therefore \boxed{L = 2.45 \text{ Letters/msg.}}$

Entropy $H(x) = -\sum_{i=1}^{6} p(x_i) \log p(x_i)$

$= -[0.3 \log 0.3 + 0.25 \log 0.25 + 0.15 \log 0.15 + 0.12 \log 0.12 + 0.10 \log 0.10 + 0.08 \log 0.08]$

$H(x) = 2.422$ bits/msg.

$\therefore \boxed{H(x) = 2.422 \text{ bits/msg.}}$

In encoding alphabets are '0' & '1' so $M = 2$, $\log_2 M = \log_2 2 = 1$ bits/letters.

$\therefore \frac{H(X)}{\log_2 M} = \frac{2.422}{1}$ = letters/msg.

$\therefore$ Coding Efficiency $\eta = \frac{H(X)}{L \cdot \log_2 M} = \frac{2.422}{2.45} = 0.9885$ $\boxed{\%\eta = 98.85\%}$

Redundancy $\delta = 1 - \eta = 1 - 0.9885 \Rightarrow \boxed{\%\delta = 1.15\%}$

② Apply the Huffman coding procedure for the following messages

(a) Take $\boxed{M=2}$ (b) Take $\boxed{M=3}$

$$[X] = [x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7 \ x_8]$$

$$[P] = \left[\frac{1}{4} \ \frac{1}{8} \ \frac{1}{16} \ \frac{1}{16} \ \frac{1}{16} \ \frac{1}{4} \ \frac{1}{16} \ \frac{1}{8}\right]$$

Sol

Huffman encoding:

(a) $\boxed{M=2}$, the encoding alphabets / letters are '0' and '1'. $\therefore \log_2^M = \log_2^2 = 1$ bit/letter

Code length (n_i)	Probability	First reduction	Second Reduction	Third reduction	Fourth reduction	Fifth Reduction
2 ← x_1	$\frac{1}{4}=0.25$ (10)	0.25 10	0.25 10	0.25 01	0.25 00	0.5 1
2 ← x_6	$\frac{1}{4}=0.25$ (11)	0.25 11	0.25 11	0.25 10	0.25 01	0.25 00
3 ← x_2	$1/8=0.125$ (010)	0.125 001	0.125 000	0.25 11	0.25 10	0.25 01
3 ← x_8	$1/8=0.125$ (011)	0.125 010	0.125 001	0.125 000	0.25 11	
4 ← x_3	$1/16=0.0625$ (0000)	0.125 011	0.125 010	0.125 001		
4 ← x_4	$1/16=0.0625$ (0001)	0.0625 0000	0.125 011			
4 ← x_5	$1/16=0.0625$ (0010)	0.0625 0001				
4 ← x_7	$1/16=0.0625$ (0011)					

0.5 0
0.5 1

Length of the Code $L = \sum_{i=1}^{8} P(x_i) \cdot n_i$

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

$$= 22/8 = 11/4 = 2.75$$

$\therefore \boxed{L = 2.75 \text{ letters/message}}$

Entropy $H(X) = -\sum_{i=1}^{8} P(x_i)\log p(x_i)$

$$= +\left[\frac{1}{4}\log 4 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{8}\log 8 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16\right].$$

$$= 22/8 = 11/4 = 2.75$$

$\therefore \boxed{H(X) = 2.75 \text{ bits/message}}$

$\therefore \frac{H(X)}{\log_2^M} = \frac{2.75}{1} = 2.75$ letters/msg.

$\therefore$ Coding efficiency $\eta = \frac{H(X)}{L \cdot \log_2^M} = \frac{2.75}{2.75\times 1} = 1$ $\boxed{\%\eta = 100\%}$

Redundancy $\delta = 1-\eta = 1-1 = 0$ $\therefore \boxed{\delta = 0}$

(b) $\underline{M = 3}$

The encoding alphabets/letters are 0, 1, 2. So, $M = 3$.

$\therefore \log_2 M = \log_2 3 = \frac{\log 3}{\log 2} = 1.585$ letters/msg.

Huffman encoding:

Length (n_i)	Probability $P(x_i)$	First Reduction I	Second Reduction II	Third reduction III
2	$\frac{1}{4}=0.25$ (00)	0.25 00	0.3125 1	0.6875 0
2	$\frac{1}{4}=0.25$ (01)	0.25 01	0.25 00	0.3125 1
2	$\frac{1}{8}=0.125$ (10)	0.1875 02	0.25 01	
2	$\frac{1}{8}=0.125$ (11)	0.125 10	0.1875 02	
2	$\frac{1}{16}=0.0625$ (12)	0.125 11		
3	$\frac{1}{16}=0.0625$ (020)	0.0625 12		
3	$\frac{1}{16}=0.0625$ (021)			
3	$\frac{1}{16}=0.0625$ (022)			

Lengths of the code $L = \sum_{i=1}^{8} P(x_i) \cdot n_i$

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

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

$L = 2.1875$ letters/msg.

$\therefore \boxed{L = 2.1875 \text{ letters/msg.}}$

Entropy $H(x) = \sum_{i=1}^{8} P(x_i) \cdot \log\left[\frac{1}{P(x_i)}\right]$

$$= \frac{1}{4}\log 4 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{8}\log 8 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16 + \frac{1}{16}\log 16]$$

$= 2.75$ bits/msg

$\therefore \boxed{H(x) = 2.75 \text{ bits/msg.}}$

$\therefore \frac{H(x)}{\log_2 M} = \frac{2.75}{1.585} = 1.735$ letters/msg.

$\therefore$ Coding efficiency. $\frac{H(x)}{\log_2 M} \times \frac{1}{L} = \frac{1.735}{2.1875} = 0.793$ $\therefore \boxed{\% \eta = 79.3\%}$

Redundancy $\delta = 1 - \eta = 1 - 0.793 = 0.207$ $\boxed{\% \delta = 20.7\%}$

Continuous Channel :

* A no-of communication systems use continuous sources and thus use the channel continuously.
* AM, FM, PM are examples of systems using continuous channel.

In a similar way the different entropies in continuous distributions

If $p(x)$ probability density function

$$\therefore h(x) = E\left[-\log p(x)\right] = -\int_{-\infty}^{\infty} p(x) \cdot \log p(x)\, dx .$$

$$h(y) = -\int_{-\infty}^{\infty} p(y) \cdot \log p(y)\, dy .$$

$$h(xy) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) \cdot \log p(x,y)\, dx\, dy .$$

$$h(x/y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) \log p(x/y)\, dx\, dy$$

$$h(y/x) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y) \log p(y/x)\, dx\, dy .$$

$$h(x/y) = h(xy) - h(y), \quad i(x,y) = h(x) - h(x/y)$$

$$h(y/x) = h(xy) - h(x) \quad i(x,y) = h(y) - h(y/x)$$

$$\therefore i(x,y) = h(x) + h(y) - h(xy) .$$

where $p(x)$, $p(y)$ are marginal densities.

$p(x,y)$ are joint density

$p(x/y)$, $p(y/x)$ are conditional densities.

$$\Rightarrow \int_{-\infty}^{\infty} p(x)\, dx = 1 , \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\, dx\, dy = 1 .$$

Gaussian PDF : A gaussian random variable is a continuous whose probability density function (PDF) is given by.

$$\boxed{p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2}} \quad -\infty \leq x \leq \infty .$$

where μ and σ^2 are the mean & variance respectively.

Shannon Hartley Theorem : (OR) Channel Capacity of Gaussian Channel : (OR) Channel Capacity theorem.

Statement : The channel capacity of a white-bandlimited Gaussian channel is

$$\boxed{C = B \log\left[1 + \frac{S}{N}\right]} \text{ bits/sec.}$$

where B - the channel bandwidth
S - the average signal power
& N - the average noise power.

It is also called Shannon's Third theorem.

Proof :

Let us consider the channel is assumed to be Gaussian channel i.e. It generates gaussian noise with zero mean and variance $\sigma^2 = N$. (AWGN) Additive White Gaussian Noise. whose probability density function (PDF) is

$$\boxed{P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}} \; ; -\infty \leq x \leq \infty \longrightarrow (1)$$

The entropy of Gaussian noise can be written as

$$h[x] = -\int_{-\infty}^{\infty} p(x) \cdot \log_2 p(x) \, dx \longrightarrow (2)$$

Consider $-\log_2 p(x) = \log_2\left(\frac{1}{p(x)}\right)$

$$= \log_2\left[\frac{1}{\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-x^2/2\sigma^2}}\right] \quad (\because \text{From eqn } (1))$$

$$= \log_2\left[\sqrt{2\pi\sigma^2} \cdot e^{+x^2/2\sigma^2}\right]$$

$$= \log_2 \sqrt{2\pi\sigma^2} + \log_2 e^{x^2/2\sigma^2}$$

$$= \log(2\pi\sigma^2)^{1/2} + \frac{x^2}{2\sigma^2} \cdot \log_2 e$$

$$-\log_2 p(x) = \frac{1}{2}\log(2\pi\sigma^2) + \frac{x^2}{2\sigma^2} \cdot \log_2 e \longrightarrow (3)$$

From eqn (2) & (3)

$$h(x) = \int_{-\infty}^{\infty} p(x) \cdot \left[\frac{1}{2}\log(2\pi\sigma^2) + \frac{x^2}{2\sigma^2} \cdot \log_2 e\right] dx$$

$$h(x) = \frac{1}{2}\log_2(2\pi\sigma^2)\int_{-\infty}^{\infty} p(x)\,dx + \frac{1}{2\sigma^2}\log_2 e\int_{-\infty}^{\infty} x^2 p(x)\,dx$$

$$= \frac{1}{2}\log_2(2\pi\sigma^2)(1) + \frac{\log_2 e}{2\cancel{\sigma^2}}\cdot\cancel{\sigma^2} \qquad \left[\because \int_{-\infty}^{\infty} p(x)\,dx = 1 \quad \int_{-\infty}^{\infty} x^2 p(x)\,dx = \sigma^2 \text{ variance}\right]$$

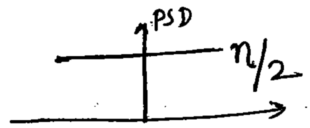
$$= \frac{1}{2}\log_2(2\pi\sigma^2) + \frac{1}{2}\log_2 e$$

$$\boxed{h(x) = \frac{1}{2}\log_2(2\pi e\sigma^2)} \text{ bits/symbol}$$

$$\text{(or)} \quad h(x) = \log_2\sqrt{2\pi e\sigma^2} \text{ bits/symbol.} \longrightarrow (4)$$

Consider white gaussian noise with mean square value N

ie $\sigma^2 = N$ = noise power



$\therefore$ Entropy $h(x) = \frac{1}{2}\log_2(2\pi e N)$

$$h(x) = \log_2\sqrt{2\pi e N} \text{ bits/symbol.}$$

$\therefore$ The Rate of information $\boxed{R(x) = 2B\cdot h(x)}$ bits/sec

$\hookrightarrow$ Sampling theorem $\hookrightarrow$ Channel bandwidth

$$\therefore \quad R(x) = 2B\log_2\sqrt{2\pi e N}\text{ bits/sec.}$$

$\therefore$ The rate of information received by the receiver can be written as

channel capacity $\quad C \geq \max[\underbrace{R(Y)}_{S+N} - \underbrace{R(N)}_{N}]$ bits/sec. $\quad [\because C \geq R]$

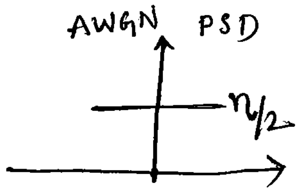
$$\geq \max\left[2B\log_2\sqrt{2\pi e(S+N)} - 2B\log_2\sqrt{2\pi e N}\right]$$

$$\geq \max\left[\cancel{2}B\cdot\frac{1}{\cancel{2}}\log_2(2\pi e(S+N)) - \cancel{2}B\cdot\frac{1}{\cancel{2}}\log_2(2\pi e N)\right]$$

$$\geq \max\left[B\log_2[2\pi e(S+N)] - B\log_2(2\pi e N)\right]$$

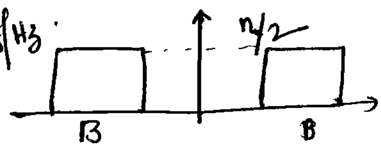
$$\geq \max\left\{B\log\left[\frac{\cancel{2\pi e}(S+N)}{\cancel{2\pi e}N}\right]\right\}$$

$$\geq \max\left\{B\log\left[\frac{S}{N}+1\right]\right\}$$



$\therefore \boxed{C = B\log[1+S/N]}$ bits/sec Hence proved.

$\therefore \boxed{C = B\log\left[1+\frac{S}{\eta B}\right]}$ bits/sec



where η - noise pdf - watts/Hz

S - signal power

B - Channel Bandwidth

$$N = 2\times\frac{\eta}{2}\times B = \eta B$$

= x =

Problem :

(1) Calculate the bandwidth of the picture (video) signal in a television. The following are the available data.

(i) No. of distinguishable brightness levels = 10.

(ii) The no. of elements per picture frame = 3,00,000

(iii) Picture frames transmitted per second = 30

& (iv) S/N required = 30 dB.

Sol

∴ No. of distinguishable brightness levels = 10

∴ The information per picture element = $\log_2 10$ = 3.32 bits/element

The information per picture frame = 3,00,000 × 3.32 = 9,96,000 bit/picture frame.

Since no. of elements per picture frame = 3,00,000.

∴ The no. of picture transmitted per second = 30

∴ The 30 picture frames are transmitted per second ie

the information rate $R = 9{,}96{,}000 \times 30$

$R = 29.9 \times 10^6$ bits/sec

∴ To transmit information at this rate, the channel capacity 'C' is

$$C \geq R$$

where

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

Given

$$S/N \text{ in dB} = 30\,\text{dB}$$

$$10 \log_{10}(S/N) = 30$$

$$S/N = 10^3 = 1000$$

∴ Hence

$$C = B \cdot \log(1 + 1000)$$

$$R \leq C$$

$$29.9 \times 10^6 \leq B \cdot \log(1001)$$

$$\therefore \quad B \geq \frac{29.9 \times 10^6}{\log(1001)} = 3.02 \times 10^6 \text{ Hz}.$$

∴ $\boxed{\text{Bandwidth } B \cong 3\,\text{MHz}.}$

Thus the minimum bandwidth required to transmit the picture (video) signal of the given television is approximately 3MHz.

Apr/May 2012 Q. Imp.

Find the efficiency of transmission using Shannon-Fano Coding and Huffman coding for the following messages sequence

$[X] = [A \quad B \quad C \quad D \quad E \quad F \quad G \quad H]$ with probabilities.

$[P] = [0.50 \quad 0.15 \quad 0.15 \quad 0.08 \quad 0.08 \quad 0.02 \quad 0.01 \quad 0.01]$.

Sol Given $[X] = [A \quad B \quad C \quad D \quad E \quad F \quad G \quad H]$

$[P] = [0.50 \quad 0.15 \quad 0.15 \quad 0.08 \quad 0.08 \quad 0.02 \quad 0.01 \quad 0.01]$

(a) Shannon Fano coding:

Sl No	Message	Probability	Encoded Messages/code I	II	III	IV	V	VI	code	Length n_i
1	A	0.50	0						→ 0	1
2	B	0.15	1	0	0				→ 100	3
3	C	0.15	1	0	1				→ 101	3
4	D	0.08	1	1	0				→ 110	3
5	E	0.08	1	1	1	0			→ 1110	4
6	F	0.02	1	1	1	1	0		→ 11110	5
7	G	0.01	1	1	1	1	1	0	→ 111110	6
8	H	0.01	1	1	1	1	1	1	→ 111111	6

The alphabets in encoded message/code are 1 & 0 So $M=2$

$\therefore \log_2 M = \log_2 2 = 1$ bit/letter.

The length $L = \sum_{i=1}^{8} p(x_i) \cdot n_i$

$= 0.50\times1 + 0.15\times3 + 0.15\times3 + 0.08\times3 + 0.08\times4 + 0.02\times5 + 0.01\times6 + 0.01\times6$

$L = 2.18$ letters/msg. $\therefore \boxed{L = 2.18 \text{ letters/msg}}$

The Entropy $H(x) = -\sum_{i=1}^{8} p(x_i) \cdot \log p(x_i)$.

$= 2.15$ bits/msg $\therefore \boxed{H(X) = 2.15 \text{ bits/msg}}$

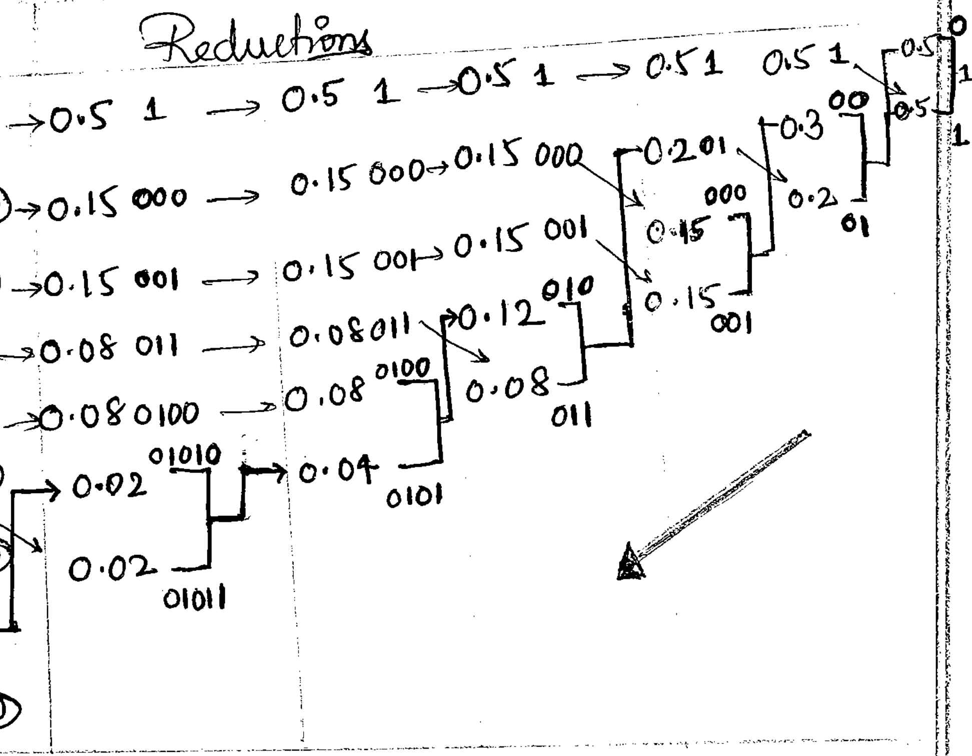
$\therefore$ Efficiency $\eta = \frac{H(X)}{L \cdot \log_2 M} = \frac{2.15}{1\times2.18}\times100 = 98.6\%$ $\therefore \boxed{\eta = 98.6\%}$

$\therefore$ Redundancy $\delta = 1-\eta = 1-0.986 = 0.014$ $\therefore \boxed{\%\delta = 1.4\%}$

Huffman coding:

length n_i	Msg	Probability	Code
1	A	0.5	1
3	B	0.15	000
3	C	0.15	001
3	D	0.08	011
4	E	0.08	0100
5	F	0.02	01011
6	G	0.01	010100
6	H	0.01	010100

Reductions



$\therefore$ The encoded alphabets are '0' & '1', so $M = 2$ $\therefore \log_2 M = 1$ bit/letter

The length $L = \sum_{i=1}^{\infty} p(x_i) \cdot n_i$

$$= 0.5 \times 1 + 0.15 \times 3 + 0.15 \times 3 + 0.08 \times 3 + 0.08 \times 4 + 0.02 \times 5 + 0.01 \times 6 + 0.01 \times 6 .$$

$L = 2.18$ letter/msg. $\therefore \boxed{L = 2.18 \text{ letter/msg.}}$

Entropy $H(X) = -\sum_{i=1}^{8} p(x_i) \log p(x_i)$

$= 2.15$ bits/msg. $\therefore \boxed{H(X) = 2.15 \text{ bits/msg}}$

$$\therefore \frac{H(X)}{\log_2 M} = \frac{2.15}{1} = 2.15 \text{ letter/msg.}$$

$\therefore$ Effeciency $\eta = \dfrac{H(X)}{L \cdot \log_2 M} = \dfrac{2.15}{2.18} \times 100 = 98.6\%$.

$\therefore \boxed{\eta = 98.6\%.}$

Redundancy $\delta = 1 - \eta = 100 - 98.6\% = 1.4\%$, $\boxed{\therefore \delta = 1.4\%.}$

Ans: In both coding techniques i.e Shannon fano coding & Huffman coding the coding effciency is same, $\eta = 98.6\%$

Sollutions for Problems:

(1)

$$[X] = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8]$$

$$[P] = [0.1 \quad 0.25 \quad 0.15 \quad 0.05 \quad 0.15 \quad 0.1 \quad 0.05 \quad 0.15]$$

Fixed length Coding.

Msg	Probability	Coding	length (n_i)
x_2	0.25	000	3
x_3	0.15	001	3
x_5	0.15	010	3
x_8	0.15	011	3
x_1	0.1	100	3
x_6	0.1	101	3
x_4	0.05	110	3
x_7	0.05	111	3

$$L = \sum_{i=1}^{8} p(x_i) \cdot n_i$$

$$= 0.25\times3 + 0.15\times3 + 0.15\times3 + 0.15\times3 + 0.1\times3 + 0.1\times3 + 0.05\times3 + 0.05\times3$$

$L = 3 \quad \therefore \boxed{L = 3 \text{ letters/msg}}$

$$H(X) = -\sum_{i=1}^{8} p(x_i) \log p(x_i)$$

$\Rightarrow \boxed{H(X) = 2.83 \text{ bits/msg}}$

Efficiency $\eta = \dfrac{H(X)}{L \log M_2} = 94.33\%$

$M = 2$

$\therefore$ 0, & 1

$\boxed{\therefore \eta = 94.33\%}$

Redundancy $1-\eta \Rightarrow \boxed{\delta = 5.67\%}$

Shannon Fano Coding:

Msg	Probability	Encoded Msg/code	Length n_i
x_2	0.25	0 0	2
x_3	0.15	0 1 0	3
x_5	0.15	0 1 1	3
x_8	0.15	1 0 0	3
x_1	0.1	1 0 1	3
x_6	0.1	1 1 0	3
x_4	0.05	1 1 1 0	4
x_7	0.05	1 1 1 1	4

Length $L = \sum_{i=1}^{8} p(x_i) \cdot n_i$

$\therefore \boxed{L = 2.85 \text{ letters/msg}}$

Entropy $H(X) = -\sum_{i=1}^{m} p(x_i) \log p(x_i)$

$\therefore \boxed{H(X) = 2.83 \text{ bits/msg}}$

$\therefore$ Efficiency $\eta = \dfrac{H(X)}{L \cdot \log_2 M}$

$\boxed{\eta = 99.3\%}$

Redundancy $1-\eta \Rightarrow \boxed{\delta = 0.7\%}$

Huffman Coding:

n_i								
2	x_2	0.25 (10)	0.25 10	0.25 10	0.25 01	0.3 00	0.45 1	0.55 0
3	x_3	0.15 (000)	0.15 000	0.2 11	0.25 10	0.25 01	0.3 00	0.45 1
3	x_5	0.15 (001)	0.15 001	0.15 000	0.2 11	0.25 10	0.25 01	
3	x_8	0.15 (010)	0.15 010	0.15 001	0.15 000	0.2 11		
3	x_1	0.1 (110)	0.1 011	0.15 010	0.15 001			
3	x_6	0.1 (111)	0.1 110	0.1 011				
4	x_4	0.05 (0110)	0.1 111					
4	x_7	0.05 (0111)						

$$L = + \sum_{i=1}^{m} p(x_i) \cdot n_i$$

$\therefore \boxed{L = 2.85 \text{ letters/msg}}$

Entropy $\boxed{H = 2.83 \text{ bits/msg}}$

Efficiency $\boxed{\eta = 99.3\%}$

Redundancy $\boxed{\delta = 0.7\%}$

(2) Given : $[P] = \left[\frac{1}{2} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{16} \quad \frac{1}{16} \quad \frac{1}{16} \quad \frac{1}{32} \quad \frac{1}{32}\right]$

Shannon Fano Coding:

1.	$\frac{1}{2}$	0					(1)
2.	1/8	1	0	0			(3)
3.	1/8	1	0	1			(3)
4.	1/16	1	1	0	0		(4)
5.	1/16	1	1	0	1		(4)
6.	1/16	1	1	1	0		(4)
7.	1/32	1	1	1	1	0	(5)
8.	1/32	1	1	1	1	1	(5)

Length $L = 2.3125$ letters/msg

Entropy $H = 2.3125$ bits/msg

Efficiency $\eta = 100\%$

Redundancy $\delta = 0\% \Rightarrow \delta = 0\%$

Huffman coding

n_i								
1	1/2 = 0.5 (1)	0.5 1	0.5 1	0.5 1	0.5 1	0.5 1	0.5 0
3	1/8 = 0.125 (010)	0.125 010	0.125 001	0.125 000	0.250 01	0.250 00	0.5 1
3	1/8 = 0.125 (011)	0.125 011	0.125 010	0.125 001	0.125 000	0.250 01	
4	1/16 = 0.0625 (0001)	0.0625 0000	0.125 011	0.125 010	0.125 001		
4	1/16 = 0.0625 (0010)	0.0625 0001	0.0625 0000	0.125 011			
4	1/16 = 0.0625 (0011)	0.0625 0010	0.0625 0001				
5	1/32 = 0.03125 (00000)	0.0625 0011					
5	1/32 = 0.03125 (00001)						

Length = 2.3125 letters/msg

Entropy = 2.3125 bits/msg

Efficiency = 100%

Redundancy = 0%

③ $p = [0.4 \quad 0.15 \quad 0.15 \quad 0.15 \quad 0.15]$

Shannon Fano coding

0.4	0			→ 0	(1)
0.15	1	0	0	→ 100	(3)
0.15	1	0	1	→ 101	(3)
0.15	1	1	0	→ 110	(3)
0.15	1	1	1	→ 111	(3)

Length (L)

$$L = \sum_{i=1}^{5} p(x_i) \cdot n_i$$

$\boxed{L = 2.2 \text{ letters/msg}}$

(1)	0.4 1	0.4 1	0.4 1	0.6 0
(3)	0.15 000	0.3 01	0.3 00	0.4 1
(3)	0.15 001	0.15 000	0.3 01	
(3)	0.15 010	0.15 001		
(3)	0.15 011			

Huffman coding

Entropy H

$$H = -\sum_{i=1}^{5} p(x_i) \log p(x_i)$$

$\boxed{H = 2.17 \text{ bits/msg}}$

Efficiency $\eta = \frac{H(X)}{L \log_2 M} = \frac{2.17}{2.2}$

$\boxed{\eta = 98.64\%}$

Redundancy $\boxed{\delta = 1.4\%}$

④

Shannon fano coding

0.27	0	0			→ 00	(2)
0.2	0	1			→ 01	(2)
0.17	1	0	0		→ 100	(3)
0.16	1	0	1		→ 101	(3)
0.06	1	1	0	0	→ 1100	(4)
0.06	1	1	0	1	→ 1101	(4)
0.04	1	1	1	0	→ 1110	(4)
0.04	1	1	1	1	→ 1111	(4)

Length L

$\boxed{L = 2.73 \text{ letters/msg}}$

(2)	0.27 01	0.27 01	0.27 01	0.27 01	0.33 00	0.4 1	0.6 0
(2)	0.2 11	0.2 11	0.2 11	0.2 10	0.27 01	0.33 00	0.4 1
(3)	0.17 000	0.17 000	0.17 000	0.2 11	0.2 10	0.27 01	
(3)	0.16 001	0.16 001	0.16 001	0.17 000	0.2 11		
(4)	0.06 1000	0.08 101	0.12 100	0.16 001			
(4)	0.06 1001	0.06 1000	0.08 101				
(4)	0.04 1010	0.06 1001					
(4)	0.04 1011						

Huffman coding

Entropy H

$\boxed{H = 2.69 \text{ bits/msg}}$

∴ Efficiency $\boxed{\eta = 98.5\%}$

Redundancy $1 - \eta \Rightarrow \boxed{\delta = 1.5\%}$