7/7/11

PREFIX CODE (OR):- Instantaneous Code

→ Prefix Code is a type of source code which is not only decodabable but also offers the possibility of realising an average code word length i.e ~~may~~ made closed to the source entropy.

→ For the code to be uniquely decodable, we ensure that for each finite sequence of symbols the corresponding sequence of codewords is different from the sequence of code words corresponding to any other source.

→ Any sequence made up of the initial part of the code word is the prefix of any other code word.

→ The Prefix Code are distinguished from other uniquely decodable codes by the fact that the end of code word is always recongnisable. Thus the decoding of prefix can be acomplised as soon as the source

~~Signal~~ symbol,

That's why prefix Code is called instantaneous Code.

PROPERTIES OF PREFIX CODE

* not property

* Kraft-McMillan Inequality :-

⇒ * If a prefix Code is Constructed from a discrete memoryless source, the Code word length should satisfy certain inequality called Kraft-Mc Millan Inequality.

⇒ It states that :-

$$\boxed{\sum_{k=0}^{K-1} 2^{-l_k} \leq 1}$$

length of Code is given by kraft Mc Millan inequality.

Where,

l_k = Codeword length of k^{th} symbol.

K = Total no. of symbols

Properties:-

(i) The prefix Code satisfy ~~that~~ kraft Mc-Millan inequality.

Source Symbol	Probability of occurence	Prefix Code	
S_0	0.5	0	$l_0 = 1$
S_1	0.25	10	$l_1 = 2$
S_2	0.125	110	$l_2 = 3$
S_3	0.125	111	$l_3 = 3$

$$\sum_{k=0}^{k-1} 2^{-l_k} \le 1$$

here, $k = 4$

$$\Rightarrow \sum_{k=0}^{3} 2^{-l_k} = 2^{-l(0)} + 2^{-l_1} + 2^{-l_2} + 2^{-l_3}$$

$$= 2^{-1} + 2^{-2} + 2^{-3} + 2^{-3}$$

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

Thus kraft's Mc-Millan inequality is satisfied.

ii) The probability of k^{th} is related to 2^{-l_k} is given by :-

$$\boxed{P_k = 2^{-l_k}}$$

$$\therefore P_0 = 2^{-l_0} = 2^{-1} = 1/2$$

$$P_1 = 2^{-l_1} = 2^{-2} = 1/4$$

$$P_2 = 2^{-l_2} = 2^{-3} = 1/8$$

$$P_3 = 2^{-l_3} = 2^{-3} = 1/8$$

iii) The average code word length of a prefix code is bounded by

$$\boxed{H(S) \leq \bar{N} \leq H(S)+1}$$

where, $\bar{N} \rightarrow$ Average codeword length

$$\boxed{\bar{N} = \sum_{k=0}^{K-1} P_k l_k}$$

$$\therefore \bar{N} = \sum_{k=0}^{K-1} 2^{-l_k} l_k \qquad [\because P_k = 2^{-l_k}]$$

$$\boxed{\bar{N} = \sum_{k=0}^{K-1} \frac{1}{2^{l_k}} \cdot l_k}$$

Entropy of the source,

$\because p_k = 2^{-l_k}$

$\frac{1}{p_k} = 2^{l_k}$

$$H(S) = \sum_{k=0}^{K-1} p_k \log_2\left(\frac{1}{p_k}\right)$$

$$= \sum_{k=0}^{K-1} \frac{1}{2^{l_k}} \log_2\left(2^{l_k}\right)$$

$$= \sum_{k=0}^{K-1} \frac{1}{2^{l_k}} l_k \log_2 2$$

$\log_2 2 = 1$

$$= \sum_{k=0}^{K-1} \frac{1}{2^{l_k}} l_k (1)$$

$$= \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}}$$

$$\Rightarrow \boxed{H(S) = \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}}}$$

& $\boxed{\text{also } \bar{N} = \sum_{k=0}^{K-1} \frac{l_k}{2^{l_k}}}$

Prop 3

$\Rightarrow$ Thus the avg code word length of a prefix code is equal to the entropy of the source.

$= 2^{-lk}$

$= 2^{lk}$

6/7/11

PROPERTIES OF ENTROPY

(i) <u>To Prove:-</u> Entropy $H = 0$, when $P_k = 0$, & $P_k = 1$.

$$H = \sum_{n=1}^{M} P_k \log_2 \left(\frac{1}{P_k} \right)$$

=)

<u>Case (i)</u> when $P_k = 1$

$$H = \sum_{n=1}^{M} 1 \log_2 (1) \qquad [\because \log_2 1 = 0]$$

$$= 0$$

$\boxed{H = 0}$ when $P_k = 1$.

<u>Case (ii)</u> when $P_k = 0$,

$$\cancel{H = \sum_{k=1}^{M} 0 \log_2 (}$$

$$H = \sum_{k=1}^{M} P_k \log_2 \left(\frac{1}{P_k} \right)$$

$$= \sum_{k=1}^{M} \lim_{k \to 0} P_k \log_2 \left(\frac{1}{P_k} \right)$$

$$= 0$$

ode

$\boxed{H = 0}$, $P_k = 0$

~~Thus~~

Thus Entropy is 0 for both certain and rare message.

ii) <u>To Prove</u> :- $H = \log_2 M$
for 'M' equally likely messages .

$$p = \frac{1}{M}$$

$$\Rightarrow P_1 = P_2 = P_3 = \cdots \cdots P_m = \frac{1}{M}$$

Entropy, $H = \sum_{k=1}^{M} P_k \log_2\left(\frac{1}{P_k}\right)$

$$= P_1 \log_2\left(\frac{1}{P_1}\right) + P_2 \log_2\left(\frac{1}{P_2}\right) + P_3 \log_2\left(\frac{1}{P_3}\right) + \cdots \cdots + P_m \log_2\left(\frac{1}{P_m}\right)$$

$$= \frac{1}{M}\log_2(M) + \frac{1}{M}\log_2 M + \cdots \frac{1}{M}\log_2 M$$

$$\boxed{H = \log_2 M}$$

iii) <u>To prove:-</u> $H_{max} \leq \log_2 M$

Using the property of natural log,

$\ln x \leq x-1$, for $x \geq 0$ — ①

$\boxed{\log_e x = \ln x}$

Let us consider two probability distribution $\{p_1, p_2, \ldots p_M\}$ & $\{q_1, q_2, \ldots, q_n\}$ on the alphabet $X = \{x_1, x_2, \ldots x_m\}$ of discrete memoryless source.

Consider the term,

$$\sum_{K=1}^{M} P_K \log_2 \left(\frac{q_k}{P_k}\right)$$

$$= \sum_{K=1}^{M} P_k \frac{\log_{10}\left(\frac{q_k}{P_k}\right)}{\log_{10}(2)}$$

$\because \log_2 x = \dfrac{\log_{10} x}{\log_{10} 2}$

-M

Multiplying & div ~~by~~ R.H.S by $\log_{10} e$

$$\sum_{K=1}^{M} P_k \log_2 \left(\frac{q_k}{P_k}\right) = \frac{\log_{10} e}{\log_{10}(2)} \frac{\log_{10}\left(\frac{q_k}{P_k}\right)}{\log_{10} e}$$

$$= \log_2 e \cdot \log_e \left(\frac{q_x}{P_x}\right)$$

But $\log_e\left(\frac{q_x}{P_x}\right) = \ln\left(\frac{q_x}{P_x}\right)$

$$\sum_{k=1}^{M} p_k \log_2\left(\frac{q_k}{p_k}\right) = \log_2 e \sum_{k=1}^{M} p_k \ln\left(\frac{q_k}{p_k}\right)$$

from ①.

$$\ln\left(\frac{q_k}{p_k}\right) \leq \frac{q_k}{p_k} - 1$$

$$\sum_{k=1}^{M} p_k \log_2\left(\frac{q_k}{p_k}\right) \leq \log_2 e \sum_{k=1}^{M} p_k \left(\frac{q_k}{p_k} - 1\right)$$

$$\leq \log_2 e \sum_{k=1}^{M} (q_k - p_k)$$

$$\Rightarrow \sum_{k=1}^{M} p_k \log_2\left(\frac{q_k}{p_k}\right) \leq \log_2 e \left[\sum_{k=1}^{M} q_k - \sum_{k=1}^{M} p_k\right]$$

$$\because \sum_{k=1}^{M} q_k = 1$$

$$\sum_{k=1}^{M} p_k = 1$$

$$\Rightarrow \boxed{\sum_{k=1}^{M} p_k \log_2\left(\frac{q_k}{p_k}\right) \leq 0}$$

$\log \frac{m}{n}$
$= \log m - \log n$
$= \log m + \log \frac{1}{n}$

Consider $q_k = \frac{1}{p_k}$, for all 'k'

$$\Rightarrow \sum_{k=1}^{M} p_k \left[\log_2 q_k + \log_2\left(\frac{1}{p_k}\right)\right] \leq 0$$

$$= \sum_{k=1}^{M} p_k \left[\log_2\left(\frac{1}{p_k}\right) + \log_2\left(\frac{1}{p_k}\right)\right] \leq 0$$

$$= \sum_{k=1}^{M} p_k \log_2 \frac{1}{p_k} + \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le 0$$

$$= \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le - \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right)$$

$$\le - \sum_{k=1}^{M} p_k \log_2 (q_k)$$

$$\cancel{= \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le - \sum}$$

$$= \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{q_k}\right)$$

Sub $q_k = \frac{1}{M}$

$$\Rightarrow \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le \sum_{k=1}^{M} p_k \log_2 M$$

$$\Rightarrow \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le \log_2 M \sum_{k=1}^{M} p_k$$

$$\because \sum_{k=1}^{M} p_k = 1$$

$$\Rightarrow \sum_{k=1}^{M} p_k \log_2\left(\frac{1}{p_k}\right) \le \log_2 M$$

$$\Rightarrow \boxed{H_{max} \le \log_2 M}$$

PROPERTIES OF MUTUAL INFORMATION

(1) To Prove! $I(X;Y) = I(Y;X)$

From Probability theory, formula's:-

(i) $P(x_i, y_j) = P(x_i/y_j)\, P(y_j)$ — ①

$P(x_i, y_j) = P(y_i/x_j)\, P(x_j)$ — ②

where,

$P(x_i, y_j) \rightarrow$ Joint Probability that x_i is transmitted & y_j is received

$P(x_i/y_j) \longrightarrow$ Conditional Probability that x_i is transmitted & y_j is received.

$P(y_i/x_j) \longrightarrow$ Conditional probability that y_i is received & x_j is transmitted.

$P(x_i) \rightarrow$ Probability of symbol x_i for transmⁿ

$P(y_j) \rightarrow$ Probability of symbol y_j for reception.

UNIT 2

11/7/11

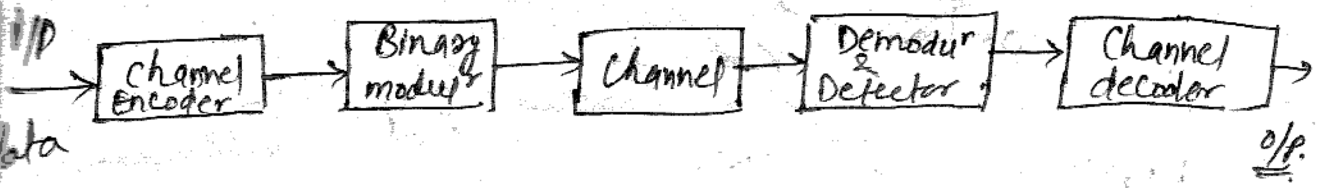
NOISY CODING :-

Types of Channels :-

(i) Binary Symmetric channel or (BSC)
(ii) Discrete Memoryless channel.
(iii) Waveform channel.

Binary Symmetric Channel (BSC)

Block Diagram :-



HARD DECISSION :- If the modulation is binary, the detector decides whether the transmitted bit is bit (0 or 1), this is hard decision.

SOFT DECISION :- If the transmission is M-ary and 'Q' is greter than or equal to M, then detector detects any 1 level out of 8 possible levels.

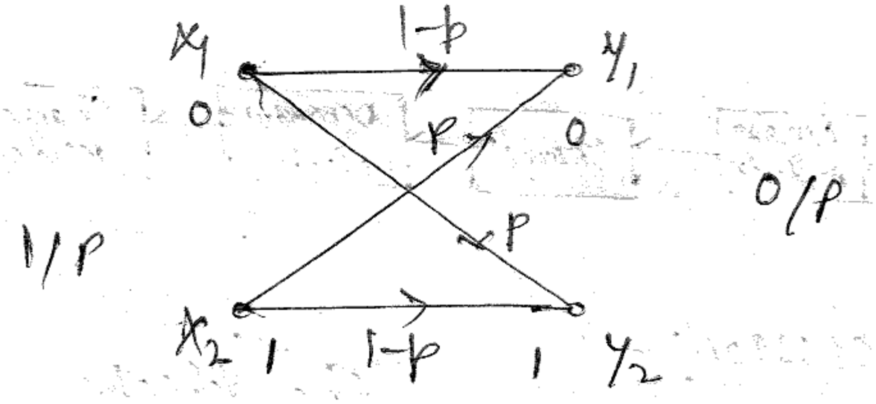
This is similar to quantizing M-ary signal to the nearest Q-level. When $Q \to \infty$, there is no quantization, this is soft decision.

Ref $X = \{0, 1\}$

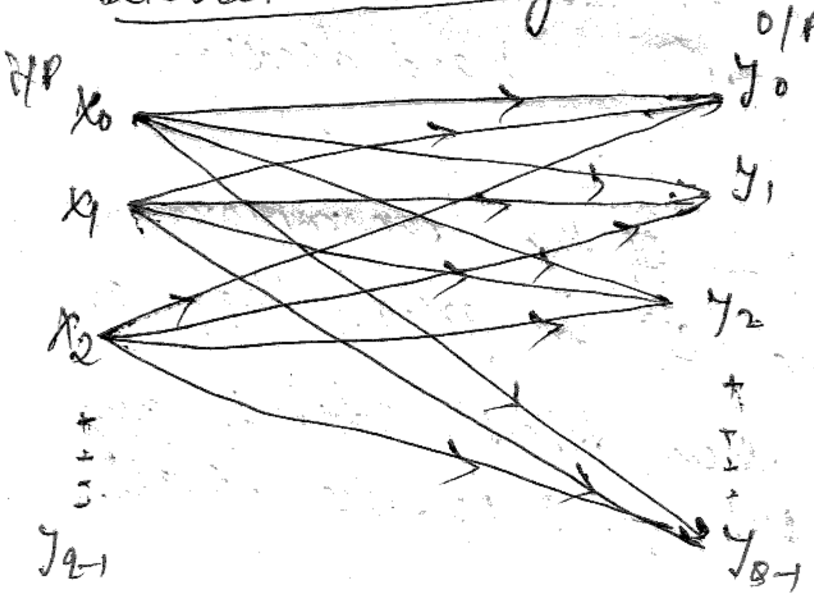
$Y = \{0, 1\}$.

$P(Y=0 / X=1) = P(Y=1 / X=0) = p$.

then, $P(Y=1 / X=1) = P(Y=0 / X=0) = 1-p$.



Discrete Memoryless Channel



Note Here each i/p is connected with all o/p

i.e. x_0 is connected with $y_0, y_1 \ldots y_{Q-1}$

Here we have $q-1$ i/p's & $s-1$ o/p's.

when the i/p to the discrete memoryless channel is sequence of n^{th} symbols & $u_1, u_2, \ldots u_n$ selected from X

then the o/p will be $v_1, v_2, \ldots v_n$.

from Y.

$\Rightarrow$ For the discrete memoryless channel the sets of o/p ~~cha~~ symbols is characterised by ~~joint~~ conditional probability.

$$\boxed{\prod_{k=1}^{n} P(Y = y_k / X = u_k)}$$

13/7/11 :-

*Classification of channels :-

(i) Lossless Channel
(ii) Deterministic Channel
(iii) Noiseless channel.
(iv) Useless Channel.
(v) Symmetric Channel.

~~Loss~~

Let us Consider a discrete memoryless Channel with input alphabets $x_1, x_2 --- x_m$ and o/p alphabets $y_1, y_2, --- y_n$.

The Channel matrix $[a_{ij}]$, where a_{ij}

$a_{ij} = P(y_j/x_i)$, where $i = 1, 2, --- M$ & $j = 1, 2, --- N$

The joint distribution of X and Y is given by

$P[X = x_i, Y = y_j]$

$= P(x_i) \cdot P(y_j/x_i)$,

⇒ The distribution of Y is given by

$$P[Y = y_j] = \sum_{i=1}^{M} P(x_i) \, P(y_j/x_i) \quad \text{— ① eq}^n$$

⇒ The information processed by the channel is given by

$$I(X;Y) = H(X) - H(X/Y) \quad \text{— ② eq}^n$$

But we have,

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

$$= H(X) + H(Y) - H(X,Y) \quad \text{— ③ eq}^n$$

→ The channel capacity is given by :-

$$\boxed{C = \max_{P(x_i)} I(X;Y)} \quad \text{— ④ eq}^n$$

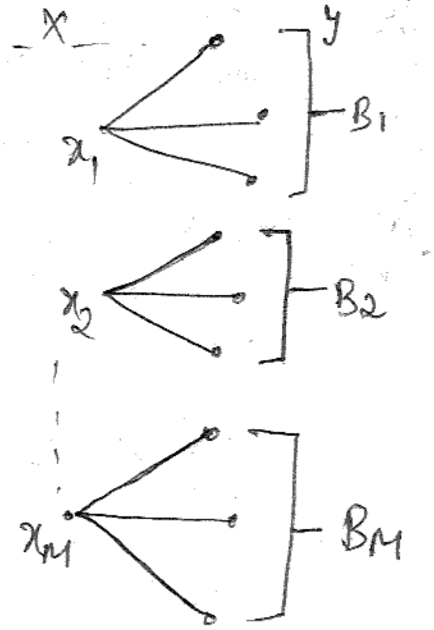
Probability of max. informat^n transmitted

→ A discrete memory channel can be defined by channel matrix 'D' where,

$$D = P(y_j/x_i) = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & \cdots & P(y_N/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & \cdots & P(y_N/x_2) \\ \vdots & \vdots & & \vdots \\ P(y_1/x_M) & P(y_2/x_M) & \cdots & P(y_N/x_M) \end{bmatrix}$$

* **Classificat^n of Channel**

(1) Lossless channel :-



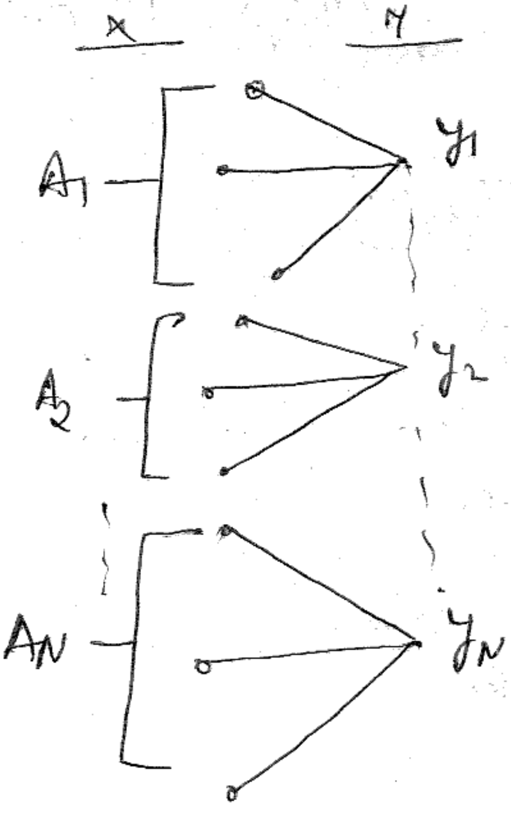
A Channel is lossless if $H(X/Y)$ is equal to zero for all i/p distributions.

The lossless Channel is one in which the i/p is determined by o/p, and hence <u>no transmission error occurs</u>

i.e no Information is lost but Ideally it is not possible.

The amount of inform transmitted = amount of " received

Deterministic Channel :-



A channel is deterministic if $\boxed{P(y_j/x_i) = 1 \text{ or } 0}$

for all values of i, & j

$\Rightarrow$ The conditional probability $P(y_j/x_i)$

can be defined as,

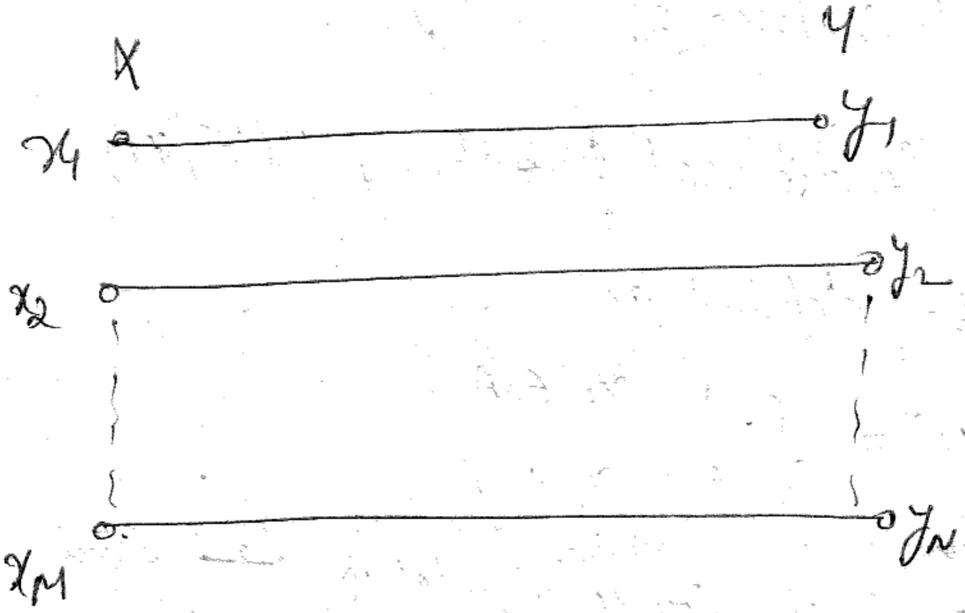
$$P(y_j/x_i) = \begin{cases} 1, & x_i \in A \\ 0, & x_i \notin A \end{cases}$$

$\Rightarrow$ ~~In~~ ~~then~~ when a given source symbol is ~~said~~ send through a deterministic channel, ~~who~~ it is clear that which destinatⁿ symbol will be received.

$\Rightarrow$ For Deterministic Channel, the channel matrix has only one non-zero element in each row and this element is unity (1)

ex: $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ $\leftarrow$ Deterministic channel

~~Noisy~~ Noiseless Channel :-

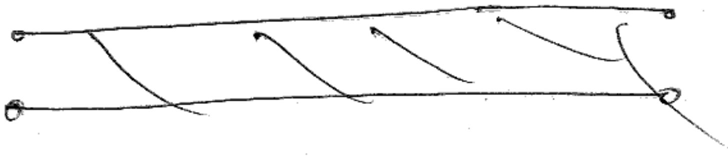


A channel is said to be Noiseless if it is both lossless and deterministic.

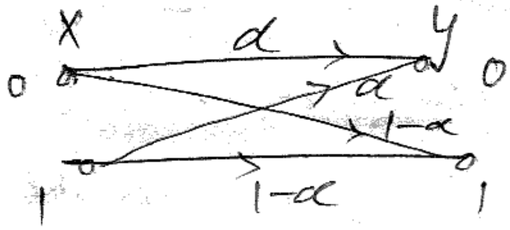
$$\boxed{I(X;Y) = H(X) = H(Y)}$$

⇒ The channel matrix has only one non zero element in each row and in each coloumn, and this element is unit

USELESS CHANNEL :-



USELESS CHANNEL



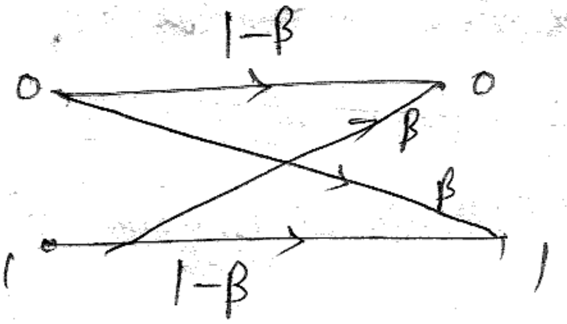
A channel is useless or zero capacity if

$\boxed{I(X;Y) = 0}$ for all i/p distribution

$\Rightarrow$ No informatⁿ can be transmitted

$\Rightarrow$ A channel is useless if & only if its channel matrix has identical rows.

Symmetric Channel



A channel is symmetric if each row of the channel matrix and each column of the channel matrix contains the same set of numbers. i.e rows and columns are identical.

⇒ For Symmetric channel
$H(Y/X)$ is independent of $P(x_i)$, and it depends only on the channel probability $P(y_j/x_i)$.

14/7/14

Calculation of Channel Capacity :-

(i) **Lossless Channel :-**

The channel capacity is obtained when the source entropy is max^m. i.e

$$\boxed{C = \max_{P(x_i)} H(X) = \log_2 M}$$

$P(x_i)$ = uniform distribution

This max^m is obtained when all x_i are equally likely or when $P(x_i)$ is a uniform distribution.

(ii) Deterministic Channel :-

Channel capacity is obtained when the destination Entropy is max^m, i.e

$$\boxed{C = \max_{P(y_j)} H(X) = \log_2 N}$$

This max^m is obtained when all y_j are equally likely, or $P(y_j)$ is a uniform distribution

iii) Noiseless Channel

Channel capacity is given by :-

$$\boxed{C = \log_2 M = \log_2 N}$$

⇒ It has both lossless & deterministic.

(iv) Symmetric Channel :-

For symbol x_i with marginal probability, a_i, we have

$$P(y_j/x_i) = \alpha_{ij} \quad \text{--- (1)}$$

$$P(x_i/y_j) = \beta_{ij} \quad \text{--- (2)}$$

The max^{m} of $H(Y)$ occurs when all the received symbols have same probability, i.e

$$\boxed{C = \log m - h} \quad \text{--- (3)}$$

m = no. of symbols
h = Constant

For symmetric channel,

$$\boxed{P(x_i)\, \alpha_{ij} = P(y_j)\, \beta_{ij}}$$

Channel capacity,

$$C = \max [H(X) - H(X/Y)]$$

$$= \max [H(X)] - h'$$

$$\boxed{C = \log X - h'}$$

where h' = Conditional entropy.
$= H(X/Y)$

(v) Unsymmetric Channel

⇒ Channel capacity is given by:-

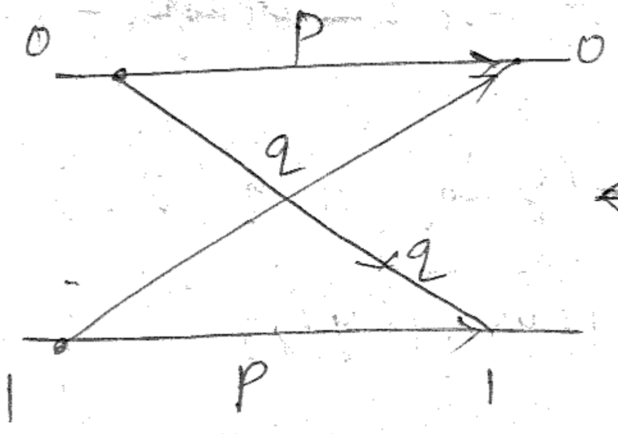
$$\boxed{C = \frac{\ln\left[e^{Q_1} + e^{Q_2}\right]}{\log_2\left[2^{Q_1} + 2^{Q_2}\right]}}$$

when Q_1 & Q_2 = variables.

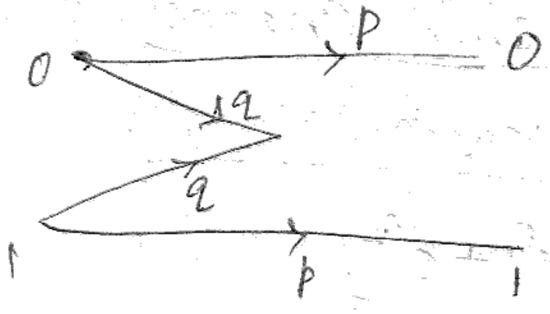
(vi) ~~Useless Channel~~ :-

Binary Symmetric Channel (BSC) :-

and Binary Erasure Channel (BEC) :-



← BSC



← BEC.

we assume that o/p of binary source is transmitted through BSC or BEC

<u>For BSC,</u> Let

$$P(0) = \alpha, \quad P(1) = 1-\alpha,$$
$$P(0/0) = P(1/1) = p$$
$$P(0/1) = P(1/0) = q$$

$$H(X) = H(\alpha, 1-\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$$

$$H(Y/X) = -(p \log p + q \log q)$$

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

$$= H(Y) + (p \log p + q \log q)$$

But

$$C = \max_{P(x_k)} [I(X;Y)]$$

$$\boxed{C = 1 + p \log p + q \log q}$$

For BEC

There are 2 i/p symbols $\{0,1\}$, and three o/p symbols $\{0, y, 1\}$.

The letter y indicates the fact that o/p is erased and no deterministic decision can be made as together the transmitted symbol ~~of~~ for 0 & 1.

Let $P(0) = \alpha$, $P(1) = 1-\alpha$

$$P(0/0) = P(1/1) = p$$
$$P(0/1) = P(1/0) = q$$

$$H(X) = H(\alpha, 1-\alpha) =$$

$$H(X/Y) = (1-p)H(X)$$

$$I(X;Y) = H(X) - H(X/Y)$$
$$= H(X) - (1-p)H(X)$$
$$= \cancel{H(X)} - \cancel{H(X)} + pH(X)$$
$$= pH(X)$$

$$\boxed{C = \max I[x;y] = p}$$