Error control Codes - Salvator Gravano } ~~inform~~ channel coding

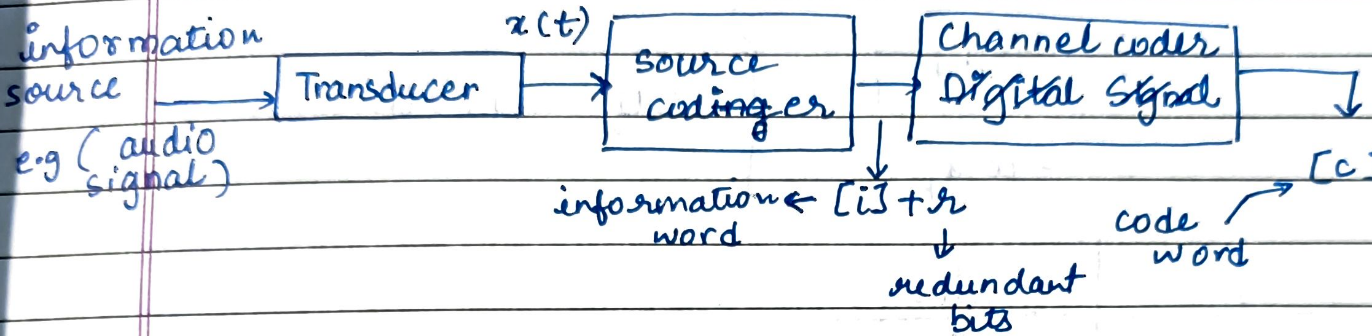
Book - ① Digital Communication By John G. Proakis } Source coding

② Tsernard Sklar

③ Haykin

INFORMATION THEORY AND CODING

Block diagram of digital Communication system



redundant bits are added by the channel coder.

$[c] = [i] + r$

$[i] = [1\ 0\ 1\ 0]$

$[r] = [0\ 0\ 0\ 1]$

$[c] = [1\ 0\ 1\ 0\ 0\ 0\ 1]$

Drawback of adding redundant bits =

bandwidth requirement ↑↑ → Data rate ↓↓

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

information (I) — probability of occurrence (P).

LOGARITHMIC FUNCTION

source
$$X = x_1, x_2, \ldots x_n$$

X is a source that emits symbols x_1 $x_2 \ldots x_n$

$$\left.\begin{aligned} I &= \log_b \frac{1}{P} \\ &= -\log_b P \end{aligned}\right\} \text{self information}$$

b	units
2	bit
e	nat
10	Hartley

Mutual information

it exists for more than one sources that are statiscally dependent

$$X_i = x_1 \quad x_2 \quad \ldots \quad x_n$$
$$Y_j = y_1 \quad y_2 \quad \ldots \quad y_n$$

$$I(x_i, y_j) = \log_b \frac{P(x_i / y_j)}{P(x_i)}$$

Properties of information (I)

i) I is always +ve
$$I > 0$$

ii) I is inversely proportional to probability
$$I \propto \frac{1}{P}$$

iii) Joint information is the sum̃ation of individual I

$$I_{i,j} = I_i + I_j$$

Q Prove that $I(x_i, y_j) = I(y_j, x_i)$

$$I(x_i, y_j) = \log_b \frac{P(x_i/y_j)}{P(x_i)}$$

$$I(x_i, y_j) = \log_b \left(\frac{P(x_i, y_j)}{P(y_j)\,P(x_i)} \right)$$

$$I(x_i, y_j) = \log_b \frac{P(y_j/x_i)}{P(y_j)}$$

$$I(x_i; y_j) = I(y_j, x_i)$$

Conditional self information

$$I(x_i/y_j) = -\log_b P(x_i/y_j)$$

Q In a binary PCM if bit 0 occurs with a probability of (1/4) and 1 occurs with a probabity of (3/4.) Then calculate the info carried by each bit.

$$I_0 = \log_b \frac{1}{P_0} = \log_b \frac{1}{(1/4)} = \log_b(4)$$

$$I_0 = \log_2 4 = 2 \text{ bits}$$

$$I_1 = \log_2 \frac{1}{3/4}$$

$$= \log_2 \frac{4}{3}$$

$$= 0.415 \text{ bits}$$

ENTROPY

information source

With memory

Memoryless

Discrete

<u>entropy</u>: It defines the statistical average of the information corresponding to all the symbols produced by the source. It is used to characterise the information producing capability of the source.

Self entropy → in case of single source

$$H(X) = \sum_{i=1}^{n} P(x_i)\, \underline{I(x_i)}$$

self information

$$I(x_i) = -\log_2 P(x_i)$$

Date
Page

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i)$$

Mutual entropy - more than one source

$$H(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i, y_j)\, I(x_i, y_j)$$

Q Prove that if there are n number of source symbol and source symbols are equiprobable then the entropy will be $\log n$.

$$H(x) = \sum_{i=1}^{n} P(x_i)\, I(x_i)$$

since all are equiprobable $P(x_i) = \frac{1}{n}$ $\forall i$

$$H(x) = -\sum_{i=1}^{n} \frac{1}{n} \log_b \left(\frac{1}{n}\right)$$

$$= \frac{1}{n} \log_b n \sum_{i=1}^{n} (1)$$

$$= \frac{1}{\cancel{n}} \log_b n \times \cancel{n}$$

$$= \log_b n$$

Information rate (R)

$R = r\, H(X)$ unit = bits/sec ↳ if b=2

r ↓ rate of symbol generation

Date
Page

Q. A discrete binary memoryless source X produced four symbols x_1, x_2, x_3, x_4 with probabilities 0.4, 0.3, 0.2, 0.1 respectively. Calculate entropy and amount of info carried by the symbol sets $I(x_1\ x_2\ x_1\ x_3)$ and $I(x_4\ x_3\ x_3\ x_2)$.

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log_b (x_i)$$

$b = 2$ (binary)

$$H(x) = -[0.4 \log_2 0.4 + 0.3 \log_2 0.3 + 0.2 \log_2 0.2 + 0.1 \log_2 0.1]$$

$$= -[-0.528 - 0.521 - 0.464 - 0.332]$$

$$= 1.845 \text{ bits/symbol.}$$

Memoryless $\rightarrow$ statistically independent

$$P(x_1\ x_2\ x_1\ x_3) = P(x_1) P(x_2) P(x_1) P(x_3)$$

$$= (0.4)(0.3)(0.4)(0.2)$$

$$= 9.6 \times 10^{-3}$$

$$I(x_1\ x_2\ x_1\ x_3) = \log_2 \left(\frac{1}{P(x_1\ x_2\ x_1\ x_3)} \right)$$

$$= \log_2 \left[\frac{1}{9.6 \times 10^{-3}} \right]$$

$$= 6.703$$

Date
Page

$$P(x_4 x_3 x_3 x_2) = P(x_4) P(x_3) P(x_3) P(x_2)$$
$$= (0.1)(0.2)^2 (0.3)$$
$$= 1.2 \times 10^{-3}$$

$$I(x_4 x_3 x_3 x_2) = \log_2 \left[\frac{1}{1.2 \times 10^{-3}} \right]$$
$$= 9.703$$

Q A telegraph source having two symbols '.' and '—'. The '.' duration is 0.2 seconds and '—' duration is 3 times of the ~~dash~~ '.' duration. The probability of '.' occurrence is twice of '—' occurence and the time between the symbols is 0.2 seconds. Calculate the information rate of the telegraph source.

$$P(dot) = 2P(dash)$$
$$t(dot) = 3\, t(dash)$$
$$\cancel{I(dot)} \quad R = r H(x)$$
$$H(x) = \sum_{i=1}^{2} P(x_i) \log_2 \frac{1}{P(x_i)}$$

$$P(dot) + P(dash) = 1$$
$$2P(dash) + P(dash) = 1$$
$$P(dash) = \frac{1}{3} \qquad P(dot) = \frac{2}{3}$$

$$H(x) = \frac{1}{3} \log_2 \frac{1}{1/3} + \frac{2}{3} \log_2 \frac{1}{2/3} = 0.922 \text{ bits/symbol}$$

Date
Page

$R = r\,H(x)$

$r = \frac{10}{\cancel{0.2}} = 8 \quad \frac{1}{T_s}$

T_s = average symbol duration
~~per~~ or
time duration per symbol

$t\{dot\} = 0.2\,s$
$t\{dash\} = 0.6\,s$
$t\{spacing\} = 0.2\,s$

$T_s = P(dot)\,t\{dot\} + P\{dash\}\,t\{dash\} + t\ spacing$

$= \frac{2}{3} \cdot 0.2 + \frac{1}{3} \cdot 0.6 + 0.2$

$= 0.133 + 0.2 + 0.2 = 0.533$

$R = \frac{1}{0.533} \times 0.922 = 1.73$ bits / second

<u>fixed length coding</u>

conditions to apply fixed length coding :-
1) source symbols must be equiprobable.
2) Total no. of symbols (N) = power of 2

If the ~~coo~~ above conditions satisfy then,
average code length (L)

$$L = \log_2 N$$

code efficiency $\eta = \frac{H(x)}{L} = \frac{\log_2 N}{\log_2 N} = 1 = 100\%$

since source symbols are equiprobable.

Variable length Coding

if a symbol is more frequent then it will be coded with a shorter length.

Prefix free code - a code in which none of the codeword can be formed by adding code symbols to another codeword is called a prefix free code. Thus, in pre-fix free code, no codeword is a prefix of another codeword.

distinct code - a code is distinct if each code word is distinguishable from other codeword.

uniquely decodable code - a code is uniquely decodable if the original source

sequence can be reconstructed perfectly from the encoded binary sequence.
Prefix free ~~or~~ codes are always uniquely decodable.

Instantaneous code - A uniquely decodable code is called instantenous if the end of any code ~~symbol~~ word is ~~recogni~~ distinguishable without examining subsequent code symbols. Instantenous code also satisfy ~~uniquely~~ prefix free ~~decodable~~ property

Kraft Inequality - A necessary and sufficient condition for the existance of binary code word with code words having length.

$$n_1 \le n_2 \le n_3 \ldots\ldots \le n_l$$

$$\sum_{K=1}^{l} 2^{-n_k} \le 1$$

Q

x_i	code A	code B	code C	code D
x_1	00	0	0	0
x_2	01	10	11	100
x_3	10	11	100	110
x_4	11	110	110	111

Q ① Prove that all codes except B satisfy Kraft inequality

② Show that code A and D are uniquely decodable and B and C are not uniquely decodable.

① <u>code A</u>

$$\cancel{\sum_{k=1}^{4} 2^{-2}} = \cancel{\left(\frac{1}{4}\right)^{1}} + \cancel{\left(\frac{1}{4}\right)^{2}} + \cancel{\left(\frac{1}{4}\right)^{3}} + \cancel{\left(\frac{1}{4}\right)^{4}}$$

$$\sum_{k=1}^{4} 2^{-n_k} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$$

$$1 \leq 1$$

<u>code B</u>

$$\sum_{k=1}^{4} 2^{-n_k} = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^3}$$

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

$$= 1 + \frac{1}{8} = \frac{9}{8} > 1$$

<u>code C</u>

$$\sum_{k=1}^{4} 2^{-n_k} = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^3}$$

$$\cancel{= \frac{1}{2} + \frac{2}{8} = \frac{4+3}{8} = \frac{7}{8} \not\leq 1}$$

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

code D.

$$\sum_{K=1}^{4} 2^{-n_K} = \frac{1}{2^1} + \frac{1}{2^3} + \frac{1}{2^3} + \frac{1}{2^3}$$

$$= \frac{7}{8} \leq 1$$

(2) B and C are not prefix free because in B x_3 is prefix of x_4 and in C x_2 is prefix is x_4.
∴ Only A and D are uniquely decodable.

SHANNON - FANO CODING

Algorithm -

1) List the source symbols in order of decreasing probability
2) Divide the symbol set into two subsets so that each has nearly equal probability.
3) Assign bit zero to the upper side symbols and bit one to the lower side symbols.
4) Continue step 2 and 3 untill a single message remains in each subset.

Q A discreate memoryless source X which is having symbols $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ and probabilities are 1/8, 1/16, 1/16, 1/4, 1/16, 1/4, 1/16 respectively. Encode the message X with variable length binary codes ~~which~~ using Shannon Fano procedure. Find the avg. code length, entropy and coding efficiency.

$X \rightarrow$	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$P(x_i) =$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$

X	P(X)	I	II	III	IV	codeword	codelength
x_4	1/4	0	0			00	2
x_6	1/4	0	1			01	2
x_1	1/8	1	0	0		100	3
x_7	1/8	1	0	1		101	3
x_2	1/16	1	1	0	0	1100	4
x_3	1/16	1	1	0	1	1101	4
x_5	1/16	1	1	1	0	1110	4
x_8	1/16	1	1	1	1	1111	4

$$H(X) = \sum_{i=1}^{8} P(x_i) \log_2 \frac{1}{P(x_i)}$$

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

$$= 1 + \frac{6}{8} + 1 = 2.75$$

$$L = \sum_{i=1}^{8} P(x_i)\, n_i$$

$$= \left(\frac{1}{4} \times \cancel{2}\right) + \left(\frac{1}{4} \times \cancel{2}\right) + \left(\frac{1}{8} \times 3\right) \times 2 + \left(\frac{1}{16} \times 4\right) \times 4$$

$$= 1 + \frac{3}{4} + 1 = 2 + \frac{3}{4} = 2.75$$

$$\eta = \frac{H(X)}{L} = \frac{2.75}{2.75} = 1 = 100\%$$

Q $X_i = x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$

$P(X_i) = 0.4 \quad 0.19 \quad 0.16 \quad 0.15 \quad 0.1$

Construct Shannon table & calculate the coding efficiency.

X_i	$P(X)$	I	II	III	codeword	codelength
x_1	0.4	0	0		00	2
x_2	0.19	0	1		01	2
x_3	0.16	1	0		10	2
x_4	0.15	1	1	0	110	3
x_5	0.1	1	1	1	111	3

$$H(X) = \sum_{i=1}^{5} P(x_i) \log_2\left(\frac{1}{P(x_i)}\right)$$

$$= 0.528 + 0.455 + 0.423 + 0.410 + 0.332$$

$$= 2.148$$

$$L = \sum_{i=1}^{5} P(x_i) n_i$$

$$= 2 \times (0.4 + 0.19 + 0.16) + 3(0.15 + 0.1)$$

$$= 1.5 + 0.75 = 2.25$$

$$\eta = \frac{H(X)}{L} = \frac{2.148}{2.25} = 0.9546 = 95.46\%$$

Huffman Source coding

① Prefix free optimal code

② Efficiency of Huffman $\geq$ Shanon - fano

Algorithm of huffman coding

1) Arrange the source symbols in decreasing order of probabilities.
2) Add the probabilities of last two message symbols that are having least probability values to get a new symbol probability and rearrange the resultant probabilities set.
3) Repeat step 2 untill there are only two message symbols remaining.
4) Start encoding with end reduction symbols and
5) Assign bit 0 to the upper one and bit 1 to the lower one.
5) Go back to the previous iteration stage and take the assigned 0 or 1 as MSB for the

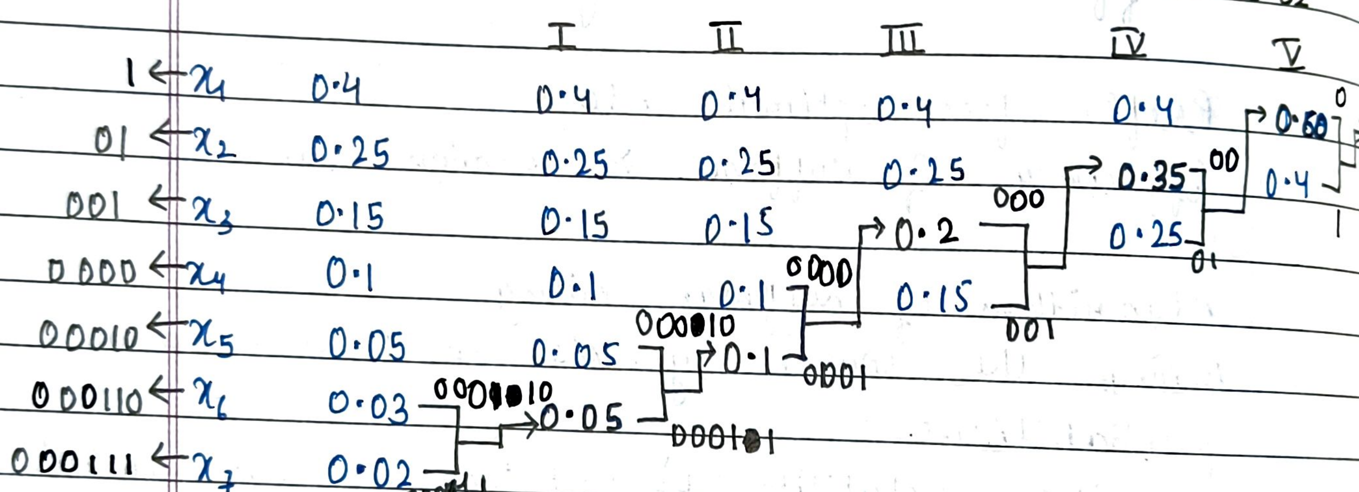
two probabilities that were combined in the previous reduction stage. Again assign bit 0 to the upper one and bit 1 to the lower 1.

6) Repeat step 5 until encoding done for all the message symbols.

Q

$x_i = x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7$

$P(x_i) = 0.4 \quad 0.25 \quad 0.15 \quad 0.1 \quad 0.05 \quad 0.03 \quad 0.02$



$$H(x) = \sum P(x_i)\, I(x_i)$$

$$= -\sum P(x_i) \log_2 P(x_i)$$

$$= -(0.4 \log_2 0.4 + 0.25 \log_2 0.25 + 0.15 \log_2 0.15 + 0.1 \log_2 0.1 + 0.05 \log_2 0.05 + 0.03 \log_2 0.03 + 0.02 \log_2 0.02)$$

$$= 0.528 + 0.5 + 0.4105 + 0.332 + 0.216 + 0.15 + 0.1128$$

$$= 2.25$$

$$L = \sum P(x_i)\, n_i$$

$$= 0.4\times1 + 0.25\times2 + 0.15\times3 + 0.1\times4 + 0.05\times5 + 0.03\times6 + 0.02\times6$$

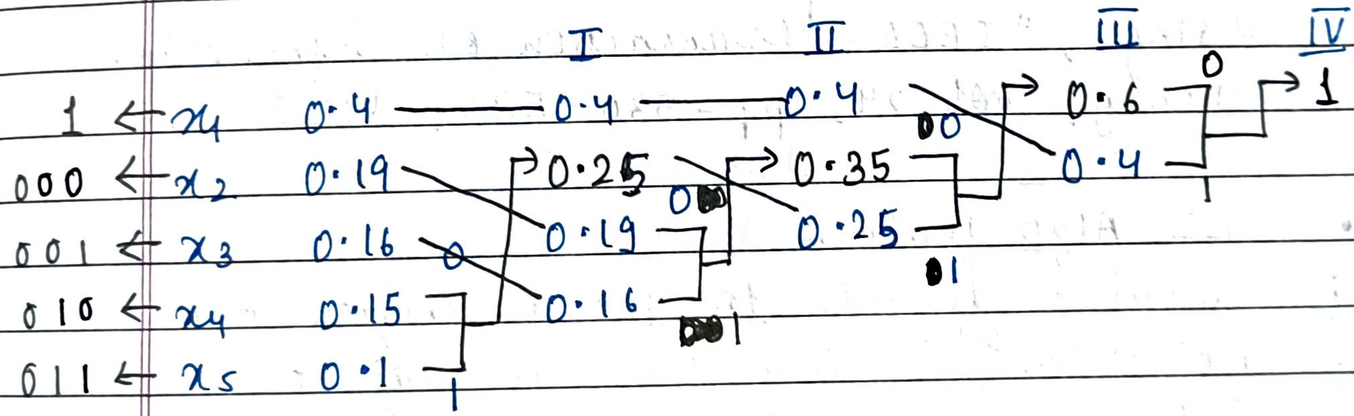
$$= 0.4 + 0.5 + 0.45 + 0.4 + 0.25 + 0.18 + 0.12$$

$$= 2.3$$

$$\eta = \frac{H(x)}{L} = \frac{2.25}{2.3} = 0.978 = 97.8\%$$

Q

X_i =	x_1	x_2	x_3	x_4	x_5
$P(x_i)$ =	0.4	0.19	0.16	0.15	0.1



$$H(x) = -P(x_i)\log_2 P(x_i)$$

$$= -(0.4\log_2 0.4 + 0.19\log_2 0.19 + 0.16\log_2 0.16 + 0.15\log_2 0.15 + 0.1\log_2 0.1)$$

$$= 2.148$$

$$L = 0.4\times1 + 3\times0.19 + 0.16\times3 + 0.15\times3 + 0.1\times3$$

$$= 0.4 + 0.57 + 0.48 + 0.45 + 0.3$$

$$= 2.2$$

$$\eta = \frac{H(X)}{L} = \frac{2.148}{2.2} = 0.976 = 97.6\%$$

LEMPEL - ZIV ALGO (L-Z)

Limitations of shanon-fano & huffmann.

i) Prior probabities are required

ii) Cannot be applied where memory is required

Lempel - Ziv. Algo.

Ziv. J. and Lemple A., "Compression of individual sequences via variable-rate coding," IEEE transaction on Information Theory, vol 24, pp-530-536, 1978.

- LZ Algo features

i) Variable length to fixed length coding

ii) universal source coding technique as the info regarding the probability of source symbols.

Algorithm

i) Divide the source sequence into variable length phrases. that was not occurred earlier. and have the last letter different than any of the previous phrases.

ii) List the phrases serially according to

Date ____
Page ____

their occurrence in a dictionary and give serial number a value whose representation is one bit less than the fixed code word size.

iii) The code word for the new phrase is the serial value of the prefix tree appended with innovation bit. For initialization, the prefix string is taken as bit 0.

iv) The decoder constructs an identical table at the receiver and decodes the received sequence.

1 0 1 0 11 01 00 100 111 010 1000 011 001 11 01 01
1000 11 011

Dictionary } → dictionary location, dictionary content, code word.

Dictionary –

Dictionary location	Dictionary content	code word
0001	1	00001
0010	0	00000
0011	10	00010
0100	11	00011
0101	01	~~000100~~ 00101
0110	00	00100
0111	100	00110
1000	111	01001
1001	010	01010
1010	1000	01110

1011	011	01011
1100	001	01101
1101	110	01000
1110	101	00111
1111	10001	10101
	1011	11101

Where is LZ Algo used?

* UNIX
* GNU
* GIF

$X = \{x_1, x_2, \ldots\ldots, x_m\}$

$Y = \{y_1, y_2, \ldots\ldots, y_n\}$

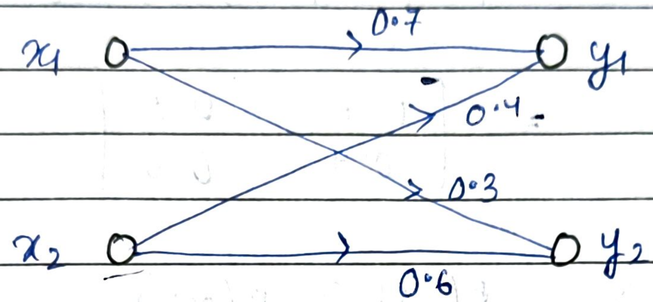
$$\begin{array}{cccc} \begin{matrix} x_1 \\ y_1 \end{matrix} & \begin{matrix} x_1 \\ y_2 \end{matrix} & \cdots\cdots & \begin{matrix} x_1 \\ y_n \end{matrix} \\ \begin{matrix} x_2 \\ y_1 \end{matrix} & \begin{matrix} x_2 \\ y_2 \end{matrix} & \cdots\cdots & \begin{matrix} x_2 \\ y_n \end{matrix} \\ \vdots & \vdots & & \vdots \\ \begin{matrix} x_m \\ y_1 \end{matrix} & \begin{matrix} x_m \\ y_2 \end{matrix} & \cdots\cdots & \begin{matrix} x_m \\ y_n \end{matrix} \end{array}$$

Marginal probability

Marginal entropy $H(X) = -\sum_{i=1}^{m} P(x_i)\log_2 P(x_i)$

$$\underbrace{P(x_i)}_{\text{marginal prob}} = \sum_{j=1}^{n} P(x_i, y_j)$$

Q



Source symbols are equiprobable ~~P(4)~~ calculate recieved ~~sign~~ symbol probability P(y) and joint probability P(XY).

$$P(x_1) = P(x_2) = 0.5$$

$$P(y_j) = \sum_{i=1}^{2} P(x_i, y_j)$$

$$= \sum_{i=1}^{2} P(y_j / x_i) P(x_i)$$

$$P(y_1) = P\left(\frac{y_1}{x_1}\right) P(x_1) + P\left(\frac{y_1}{x_2}\right) P(x_2)$$

$$= (0.7)(0.5) + (0.4)(0.5)$$

$$= 0.35 + 0.20$$

$$= 0.55$$

$$P(y_2) = P\left(\frac{y_2}{x_1}\right) P(x_1) + P\left(\frac{y_2}{x_2}\right) P(x_2)$$

$$= (0.3)(0.5) + (0.6)(0.5)$$

$$= 0.15 + 0.30 = 0.45$$

$$P(XY) = P(x_i, y_j) = P\left(\frac{y_j}{x_i}\right) P(x_i)$$

$$P\left(\frac{y_j}{x_i}\right) = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix}$$

$$= \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

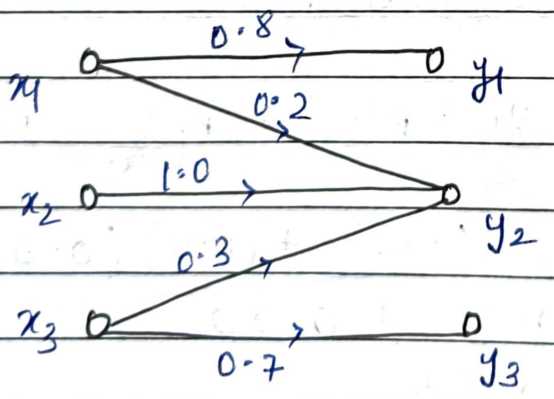
$$P(x_i, y_j) = P\left(\frac{y_j}{x_i}\right) \cdot P(x_i)$$

$$= \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} P(x_i) \end{bmatrix}$$

$$= \begin{bmatrix} 0.7\,P(x_1) & 0.3\,P(x_1) \\ 0.4\,P(x_2) & 0.6\,P(x_2) \end{bmatrix}$$

$$= \begin{bmatrix} 0.35 & 0.15 \\ 0.20 & 0.30 \end{bmatrix}$$

Q



$P(x_1) = 0.3$

$P(x_2) = 0.4$

$P(x_3) = 0.3$

find $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, $H(XY)$

$$H(X) = -\sum_{i=1}^{3} P(x_i) \log_2 P(x_i)$$

$$= 0.521 + 0.5287 + 0.521$$

$$= 1.57 \text{ bits/symbol.}$$

$$P(Y|X) = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0.3 & 0.7 \end{bmatrix}$$

$$P(y_j) = \sum P(x_i, y_j)$$

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

$$P(y_1) = \sum_{i=1}^{3} P(y_1/x_i) P(x_i)$$

$$= P(y_1|x_1) P(x_1) + P(y_1/x_2) P(x_2) + P(y_1|x_3) P(x_3)$$

$$= (0.8)(0.3) + (0)(0.4) + (0)(0.3)$$

$$= 0.24$$

$$P(y_2) = P(y_2/x_1) P(x_1) + P(y_2|x_2) P(x_2) + P(y_2/x_3) P(x_3)$$

$$= (0.2)(0.4) + (1)(0.4) + (0.3)(0.3)$$

$$= 0.06 + 0.4 + 0.09$$

$$= 0.55$$

$$P(y_3) = (0.7)(0.3) = 0.21$$

$$H(Y) = 0.494 + \cancel{0.518}^{\,0.474} + 0.4\cancel{77}28$$

$$= 1.44 \text{ bits/symbol}.$$

$$P(XY) = P(Y/X)\, P(x_i)$$

$$= \begin{bmatrix} 0.8 \times 0.3 & 0.2 \times 0.3 & 0 \\ 0 & 1 \times 0.4 & 0 \\ 0 & 0.3 \times 0.3 & 0.7 \times 0.3 \end{bmatrix} P(x_i)$$

$$= \begin{array}{cccc} y_1 & y_2 & y_3 & \\ \left[\begin{array}{c} 0.24 \\ 0 \\ 0 \end{array}\right. & \begin{array}{c} 0.06 \\ 0.4 \\ 0.09 \end{array} & \left.\begin{array}{c} 0 \\ 0 \\ 0.21 \end{array}\right] & \begin{array}{c} \cancel{x_1}\ X_1 \\ \cancel{x_2}\ X_2 \\ \cancel{x_3}\ X_3 \end{array} \end{array}$$

$$H(XY) = -\sum_{i=1}^{3}\sum_{j=1}^{3} P(XY) \log_2 P(XY)$$

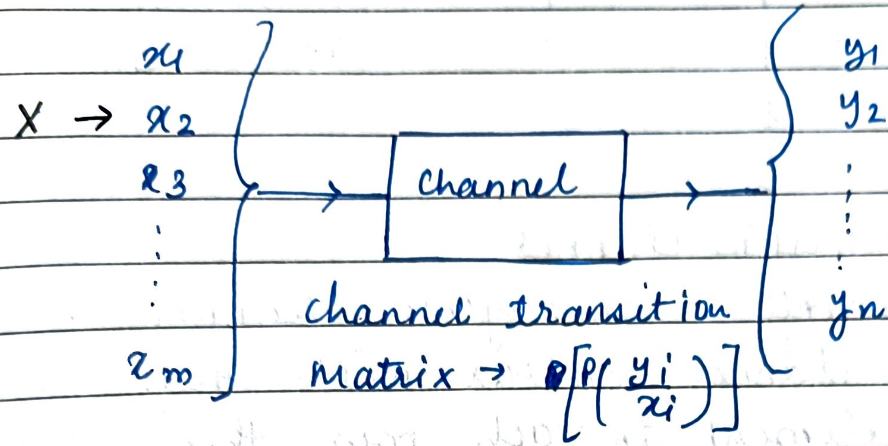
$$= \cancel{2.083} \quad \cancel{2.051}$$

$$= 2.053$$

$$X(Y|X) = -\sum_{j=1}^{3}\sum_{i=1}^{3} P(x_i, y_j) \log_2 \left(\frac{y_j}{x_i}\right)$$

$$= -\sum_{j=1}^{3}\sum_{i=1}^{3} P\left(\frac{x_i}{y_j}\right) P(y_j) \log_2\left(\frac{y_j}{x_i}\right)$$

Discrete Memoryless channel

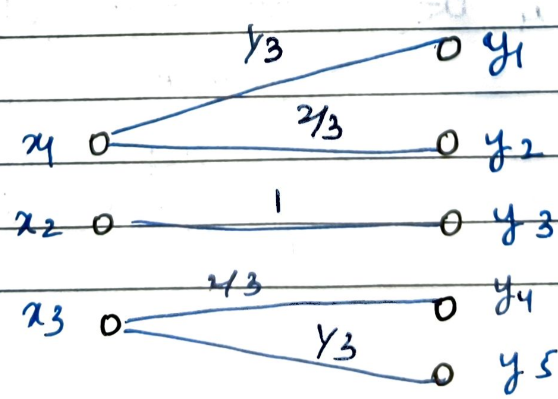


$$\boxed{C = B \log_2(1 + SNR)}$$

$$\left[P\left(\frac{y_i}{x_i}\right)\right] = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & \cdots & P(y_n/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & \cdots & P(y_n/x_2) \\ \vdots & \vdots & & \vdots \\ P(y_1/x_m) & P(y_2/x_m) & \cdots & P(y_n/x_m) \end{bmatrix}$$

1. **Lossless channel**

A ~~the~~ discrete memoryless channel described by the channel matrix with only one non-zero element in each column.



$$P(Y|X) = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2/3 & 1/3 \end{bmatrix}$$

2. **Deterministic Channel**

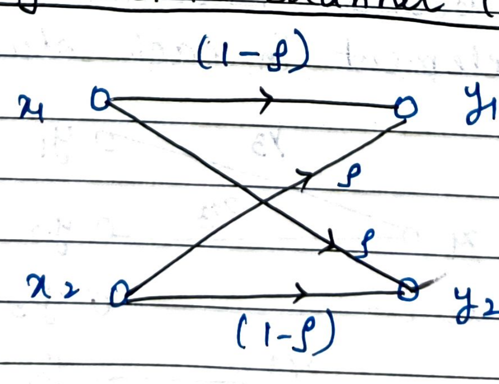
A discrete memoryless channel described by a channel matrix with only one non-zero element in each row then that is called deterministic channel.

3. **Noiseless channel**

A discrete memoryless channel, ~~d~~ ~~it~~ ~~it~~ is called noiseless if the channel is both Lossless and deterministic.

e.g.,
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4. **Binary Symmetric channel (BSC)**



Calculate the average error probability for binary symmetric channel.

$$\cancel{P_{avg}(\varepsilon) = \frac{1}{2}(1-p) + \frac{1}{2}(1-p)}$$

$$\cancel{= 1-p}$$

$$P_e = \cancel{P(y_2|x_1) + P(y_1|x_2)}$$

$$= P(x_1)\,P\left(\frac{y_2}{x_1}\right) + P(x_2)\,P\left(\frac{y_1}{x_2}\right)$$

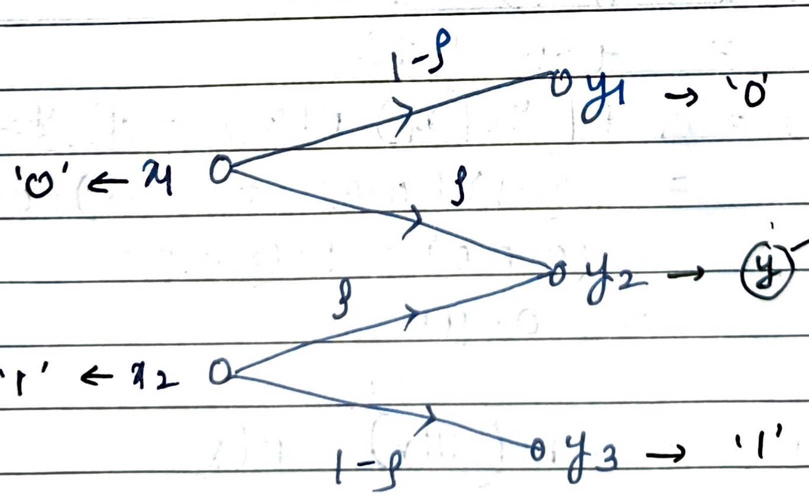
$$= \cancel{\frac{1}{2}p + \frac{1}{2}}$$

$$P(x_1) = \alpha \qquad P(x_2) = 1-\alpha$$

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

$$= \cancel{\alpha p} + p - \cancel{\alpha p} = p$$

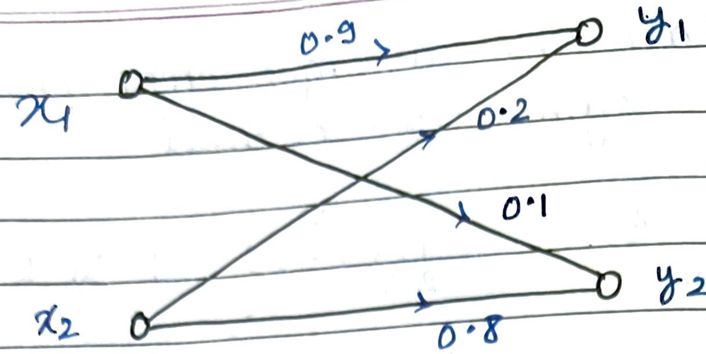
5. <u>Binary Erasure Channel</u> (BEC)



↓ TTL

0-0.8V → '0'
graystate
3.2-5V → '1'

Q



Calculate → (1) channel matrix
(2) $P(y_1)$ & $P(y_2)$
(3) $P(x,y)$
(4) $H(X)$ & $H(Y)$
(5) $H(X/Y)$, $H(Y/X)$, $H(X,Y)$

$P(x_1) = P(x_2) = 0.5$

(1) channel matrix $= \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$

(2) $P(y_1) = P(y_1/x_1) \cdot P(x_1) + P(y_1/x_2)\, P(x_2)$
$= (0.9)(0.5) + (0.2)(0.5)$
$= 0.45 + 0.10$
$= \cancel{0.46}\ 0.55$

$P(y_2) = P(y_2/x_1)\, P(x_1) + P(y_2/x_2)\, P(x_2)$
$= (0.1)(0.5) + (0.8)(0.5)$
$= (0.05) + (0.40)$
$= 0.45$

$P(x,y) = P(y/x) \cdot P(x)$

$$= \begin{bmatrix} 0.9 \times 0.5 & 0.1 \times 0.5 \\ 0.2 \times 0.5 & 0.8 \times 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} 0.45 & 0.05 \\ 0.10 & 0.40 \end{bmatrix}$$

(4) $H(\cancel{X}\,Y) = 0.9927$

$\cancel{H(Y)} = H(X) = 1$

(5) $H(Y|X) = -\sum_{j=1}^{2} \sum_{i=1}^{2} P\left(\cancel{\frac{y_j}{x_i}}\right) \log_2 \left(\frac{y_j}{x_i}\right)$

$P(x_i, y_j)$

BLOCK CODE

error → decoder input

$[c] + [e] = [v]$

Tx —— Rx

$[\hat{e}] + [v] = [\hat{c}]$ → estimated code word

source
↳ $[i] + [p] \rightarrow [c]$

parity bits ($[p]$)

↳ code word ($[c]$)

$[i] + [p] = [c]$ → (n, k)

n : length of codeword
k : length of information word

code rate $= \frac{k}{n}$

$\boxed{\text{no. of codeword in codeset} = 2^k}$

Parity Code

i	Even		Odd ($i_0\, i_1\, i_2\, p$)
0 0 0	0	$\rightarrow C_0$	0 0 0 1
0 0 1	1	$\rightarrow C_1$	0 0 1 0
0 1 0	1	$\rightarrow C_2$	0 1 0 0
0 1 1	0	$\rightarrow C_3$	0 1 1 1
1 0 0	1	$\rightarrow C_4$	1 0 0 0
1 0 1	0	$\rightarrow C_5$	1 0 1 1
1 1 0	0	$\rightarrow C_6$	1 1 0 1
1 1 1	1	$\rightarrow C_7$	1 1 1 0

$s = i_0 + i_1 + i_2 + p$

All additions will be modulo 2 addition

Hamming code

r(parity bits) = $n - k$

$n = \cancel{2^k - 1}\ \ 2^r - 1$

$k = n - r = 2^r - 1 - r$

$r = 0$ X $\quad r = 1$ X $\quad r = 2$ X

$\underbrace{n = 3 ; k = 1}_{\text{very ineffient}}$

$r = 3$

$n = 7 ; k = 4 \Rightarrow$ first possible combination for hamming code.

Q

$p_0 = i_0 \oplus i_1 \oplus i_2$

$p_1 = i_1 \oplus i_2 \oplus i_3$

$p_2 = i_0 \oplus i_1 \oplus i_3$

Find out all the codewords for (7,4) hamming code.

$2^k = 2^4 = 16$ codewords in codeset

hamming weight
↳ No. of non-zero bits

Hamming distance

W	i_0	i_1	i_2	i_3	p_0	p_1	p_2	d
0	0	0	0	0	0	0	0	
3	0	0	0	1	0	1	1	
3	0	0	1	0	1	1	0	
4	0	0	1	1	1	0	1	
4	0	1	0	0	1	1	1	
3	0	1	0	1	1	0	0	
3	0	1	1	0	0	0	1	
4	0	1	1	1	0	1	0	
3	1	0	0	0	1	0	1	
4	1	0	0	1	1	1	0	
4	1	0	1	0	0	1	1	
3	1	0	1	1	0	0	0	
3	1	1	0	0	0	1	0	
4	1	1	0	1	0	0	1	
4	1	1	1	0	1	0	0	
7	1	1	1	1	1	1	1	

d : Hamming distance
↳ How many bits the value differs.

$d(c_0, c_1) = 3$ (include parity bits too)

Minimum Hamming distance (d_{min}) :

Error detection capability $= d_{min} - 1$

Error correction capability = $\frac{1}{2}(d_{min} - 1)$

Linear Block Code

A block code will be linear if sum of any two code words results another code word of the same codeset.

Properties of linear block code

- All zero word is a codeword of linear block code.
- $C_i + C_j = C_k$

  $d(C_i, C_j) = w(C_k)$

- The minimum hamming distance is the minimum hamming weight

  $d_{min} = w_{min}$. (excluding all zero word)

↳ $d_{min} = 3$

error detection capability = $d_{min} - 1 = 2$

Generator Matrix [G]

$C = i[G]$

$G = [\underbrace{I_k}_{\text{identity matrix with dimension } (K \times K)} : P]$ → systematic form

P → parity matrix with dimensions $K \times (n-K)$

This is not identity matrix

$$G = \begin{bmatrix} \boxed{\begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{matrix} \end{bmatrix}$$

This is in non-systematic form

1. $R_2 \oplus R_3 \rightarrow R_3$

$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

2. $R_1 \oplus R_3 \rightarrow R_1$

$$G = \begin{bmatrix} \boxed{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}} & \begin{matrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{matrix} \end{bmatrix}$$

identity matrix

Q For (6x3) linear block code, the generator matrix is given. Calculate the minimum error detection & correction for this.

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$C = i[G]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ & \\ & \end{bmatrix}$$

$2^3 = 8$: codewords.

	i			p			w
	0	0	0	0	0	0	~~3~~ 0
	0	0	1	1	1	0	~~3~~ 3
	0	1	0	1	0	1	~~4~~ 3
$R_2 + R_3$	0	1	1	0	1	1	~~3~~ 4
	1	0	0	0	1	1	~~4~~ 3
	1	0	1	1	0	1	4 4
	1	1	0	1	1	0	4
	1	1	1	0	0	0	3

$\boxed{\text{Alternate}}$

$C_1 = [0\ 0\ 0]\, G$ $\qquad$ $C_5 = [1\ 0\ 0]\, G$

$C_2 = [0\ 0\ 1]\, G$ $\qquad$ $C_6 = [1\ 0\ 1]\, G$

$C_3 = [0\ 1\ 0]\, G$ $\qquad$ $C_7 = [1\ 1\ 0]\, G$

$C_4 = [0\ 1\ 1]\, G$ $\qquad$ $C_8 = [1\ 1\ 1]\, G$

$w_{min} = 3$

$d_{min} = 3$

$d_{min} - 1 = 2$

Date
Page

Decoder for LBC

$[i][G] = C$ $\quad$ $H \rightarrow$ parity check matrix

$$H = \left| P^T \; I_{n-k} \right|$$

$$G = [I_K, P]$$

Q for a (7,4) hamming code,

$$[G] = \begin{bmatrix} \underbrace{\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}}_{\text{identity matrix}} & \underbrace{\begin{matrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{matrix}}_{\text{parity matrix}} \end{bmatrix}$$

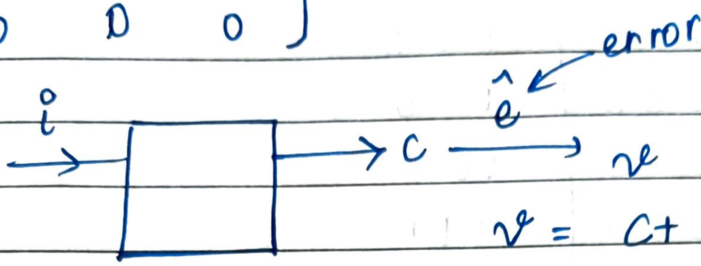
$$[P] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$P^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$H^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$GH^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



$$v = C + \hat{e}$$

parity check sum

$$S = vH^T = (C + \hat{e})H^T$$
$$= CH^T + \hat{e}H^T$$
$$= iGH^T + \hat{e}H^T$$
$$vH^T = \hat{e}H^T$$

$$vH^T = \hat{e}H^T$$

if $\hat{e}$ is 0 then $(s = 0)$

Q. Determine the syndrome table for (6,3)

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

identity matrix

$$P = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad P^T = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

$$H = [P^T, I_{n-k}]$$
$$= \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

code word length.

$\boxed{n=6}$

e	$S = eH^T$
000001	001
000010	010
000100	100
001000	110
010000	101
100000	011

given,

c = 110110   find values of V, S for above question

values of error:
- (a) → 001000
- (b) → 011011
- (c) → 000111

(a) ~~v+e~~ $V = c + e$

$= [1\ 1\ 0\ 1\ 1\ 0] + [0\ 0\ 1\ 0\ 0\ 0]$

$= [1\ 1\ 1\ 1\ 1\ 0]$

$S = V \times H^T$

$$= [1\ 1\ 1\ 1\ 1\ 0]\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$= [1\ 1\ 0]$

is this available in the syndrome table ⇒ YES!

find the corresponding error pattern

$e = [0\ 0\ 1\ 0\ 0\ 0]$

add this to v $[1\ 1\ 1\ 1\ 1\ 0]$

$e + v = [1\ 1\ 0\ 1\ 1\ 0]$

$\hookrightarrow$ This is matching with the code word.

(b) $v = c + e$

$= [1\ 1\ 0\ 1\ 1\ 0] + [0\ 1\ 1\ 0\ 1\ 1]$

$= [1\ 0\ 1\ 1\ 0\ 1]$

$S = v \times H^T$

$$[1\ 0\ 1\ 1\ 0\ 1]\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$= [0\ 0\ 0]$

if $s = 0$ then O/P $= v$

$v = [1\ 0\ 1\ 1\ 0\ 1] \neq c$

error in detection

because it is a one-bit error corrector and here we have 1 in 4 bits (b)

(c) $v = c + e$

$[1\ 1\ 0\ 1\ 1\ 0] + [0\ 0\ 0\ 1\ 1\ 1]$

$= [1\ 1\ 0\ 0\ 0\ 0]$

$$S = v \times H^T$$

$$= [1\,1\,0\,0\,0\,1] \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= [1\ 1\ 1]$$

Not found in syndrome table

$\therefore$ Decoding will fail