Q1) a) Expression for Entropy

Consider a Source S that emits M possible Symbols say $s_1, s_2, s_3, \ldots, s_M$ which are statistically independent sequence let $p_1, p_2, p_3, \ldots, p_M$ be their respective probability of occurance of N symbols.

In a long message of N symbols

s_1 occur's on a avg of $p_1 N$ times
s_2 " " " " " $p_2 N$ times
⋮
s_M occur's on a avg of $p_M N$ times

Treating individual symbols as message's of length one, for i^{th} symbol $I_{s_i} = \log_2\left(\frac{1}{P_i}\right)$ bits

∴ $P_i N$ Occurance's of s_i Contributes to information content of $P_i N \log_2\left(\frac{1}{P_i}\right)$ bits

Thus, $I_{total} = \sum_{i=1}^{M} P_i N \log_2\left(\frac{1}{P_i}\right)$ bits

Thus, Entropy (Avg information per Symbol),

$$H = \frac{I_{total}}{N} = \sum_{i=1}^{M} P_i \log_2\left(\frac{1}{P_i}\right) \text{ bits/symbol.}$$

Q1 b]

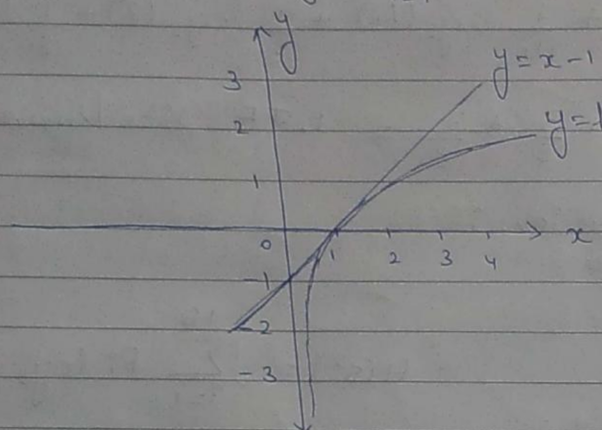
Upper limit of Entropy

Upper limit of Entropy is given by $\log_2 q - H(S)$

$$= \sum_{i=1}^{q} P_i \log_2 q - \sum_{i=1}^{q} P_i \log_2 \left(\frac{1}{P_i}\right)$$

$$= \sum_{i=1}^{q} P_i \log_2 (q P_i) = \sum_{i=1}^{q} P_i \frac{\log_e (q P_i)}{\log_e 2}$$

$$= \log_2 e \sum_{i=1}^{q} P_i \ln (q P_i) \quad \text{——— } ①$$



from curve,

$[\ln(x) \le x - 1] \, (-1)$

$-\ln(x) \ge (1 - x)$

$\ln\left(\frac{1}{x}\right) \ge (1 - x)$

$x = \frac{1}{q P_i}$

$$\ln (q P_i) \ge (1 - 1/q P_i)$$

Multiply both sides ~~sides~~ with $\sum_{\forall i} P_i \quad i = 1, 2, \ldots q$

$$\sum_{i=1}^{q} P_i \ln (q P_i) \ge \sum_{i=1}^{q} P_i (1 - 1/q P_i)$$

multiply both sides with $\log_2 e$.

$$\log_2 q - H(S) = \log_2 e \sum_{i=1}^{q} P_i \ln (q P_i) \ge \left[\sum_{i=1}^{q} P_i \left(1 - \frac{1}{q P_i}\right) \log_2 e\right]$$

$$\log_2 q - H(S) \ge \log_2 e \left(1 - \sum_{i=1}^{q} \frac{1}{q}\right)$$

$$\ge \log_2 e \, (1 - 1)$$

$$\log_2 q \ge H(S)$$

Thus $H(S)_{max} = \log_2 q$

02] messages probabilities

messages	probabilities
S_1	$1/2$
S_2	$1/4$
S_3	$1/8$
S_4	$1/8$

q = no of symbols = 4

i] Entropy of Source, $H(S) = \sum_{i=1}^{4} P_i \log_2\left(\frac{1}{P_i}\right)$ bits/symbol

$$= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8)$$

$$H(S) = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{3}{8} = 1.75 \text{ bits/symbol}$$

ii) Second extension:

	probabilities
S_1S_1	$P_1 \to 1/4$
S_1S_2	$P_2 \to 1/8$
S_1S_3	$P_3 \to 1/16$
S_1S_4	$P_4 \to 1/16$
S_2S_1	$P_5 \to 1/8$
S_2S_2	$P_6 \to 1/16$
S_2S_3	$P_7 \to 1/32$
S_2S_4	$P_8 \to 1/32$
S_3S_1	$P_9 \to 1/16$
S_3S_2	$P_{10} \to 1/32$
S_3S_3	$P_{11} \to 1/64$
S_3S_4	$P_{12} \to 1/64$
S_4S_1	$P_{13} \to 1/16$
S_4S_2	$P_{14} \to 1/32$
S_4S_3	$P_{15} \to 1/64$
S_4S_4	$P_{16} \to 1/64$

$$H(S^2) = \sum_{i=1}^{16} P_i \log_2\left(\frac{1}{P_i}\right) \text{ bits/symbol}$$

$$= \frac{1}{4}\log_2(4) + \frac{2}{8}\log_2(8) + \frac{5}{16}\log_2(16) + \frac{4}{32}\log_2(32) + \frac{4}{64}\log_2(64)$$

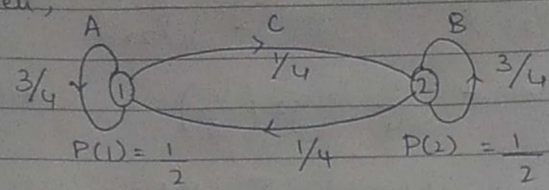
$$= \frac{1}{4}(2) + \frac{6}{8} + \frac{20}{16} + \frac{20}{32} + \frac{24}{64}$$

$= 3.5$ bits/symbol

$= 2 \times 1.75$ bits/symbol

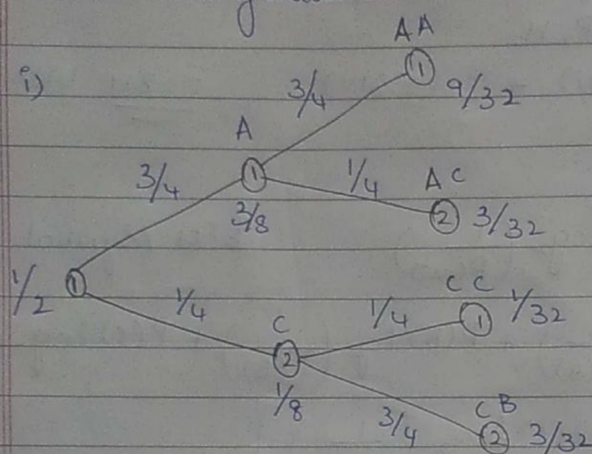
$= 2\,H(S)$

Q4] Given,

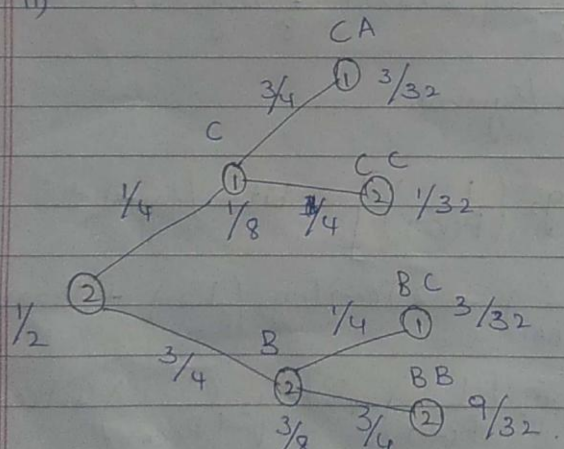


Sol: Tree diagram:

i)



ii)



a) → Entropy of State $(H_i) = \sum_{\forall j} P_{ij} \log_2\left(\frac{1}{P_{ij}}\right)$ bits/Symbol

$$H_1 = P_{11} \log_2\left(\frac{1}{P_{11}}\right) + P_{12} \log_2\left(\frac{1}{P_{12}}\right)$$

$$= \frac{3}{4} \log_2\left(\frac{4}{3}\right) + \frac{1}{4} \log_2(4) = 0.811 \text{ bits/Symbol}$$

$$H_2 = P_{21} \log_2\left(\frac{1}{P_{21}}\right) + P_{22} \log_2\left(\frac{1}{P_{22}}\right) \text{ bits/symbol}$$

$$= \frac{1}{4}\log_2(4) + \frac{3}{4}\log_2\left(\frac{4}{3}\right) = 0.811 \text{ bits/symbol}$$

→ Entropy of source, $H(s) = \sum H_i P_i$ bits/symbol

$$H(s) = P_1 H_1 + P_2 H_2$$

$$= \frac{1}{2}(0.811) + \frac{1}{2}(0.811) = \underline{0.811} \text{ bits/symbol.}$$

b) $$G_1 = \frac{1}{1} \sum_{\forall i} P(m_i) \log_2\left(\frac{1}{P(m_i)}\right) \text{ bits/symbol}$$

$$= 1\left[P(A)\log_2\left(\frac{1}{P(A)}\right) + P(B)\log_2\left(\frac{1}{P(B)}\right) + P(C)\log_2\left(\frac{1}{P(C)}\right) + P(C)\log_2\left(\frac{1}{P(C)}\right)\right]$$

$$= \frac{3}{8}\log_2(8/3) + \frac{3}{8}\log_2(8/3) + \frac{2}{8}\log_2(8)$$

$$G_1 = \underline{1.811} \text{ bits/symbol}$$

$$G_2 = \frac{1}{2}\left[\sum_{\forall i} P(m_i) \log_2\left(\frac{1}{P(m_i)}\right)\right] \text{ bits/symbol.}$$

$$= \frac{1}{2}\left[P(AA)\log_2\left(\frac{1}{P(AA)}\right) + 2P(CC)\log_2\left(\frac{1}{P(CC)}\right) + P(BB)\log_2\left(\frac{1}{P(BB)}\right) + P(AC)\log_2\left(\frac{1}{P(AC)}\right) + P(CA)\log_2\left(\frac{1}{P(CA)}\right) + P(CB)\log_2\left(\frac{1}{P(CB)}\right) + P(BC)\log_2\left(\frac{1}{P(BC)}\right)\right]$$

$$= \frac{1}{2}\left[\frac{2\times 9}{32}\log_2(32/9) + \frac{4\times 3}{32}\log_2(32/3) + \frac{2}{32}\log_2(32)\right] = 1.311 \text{ bits/symbol}$$

$$G_1 = 1.811 > G_2 = 1.311 > H(s) = 0.811$$

	Symbols	probabilities	
1	A	1/4	0.25
2	B	1/4	0.25
3	C	1/8	0.125
4	D	1/8	0.125
5	E	1/8	0.125
6	F	1/16	0.0625
7	G	1/16	0.0625

05]

Sol: Shanno - Fano binary Code :

A	1/4	1	1/4	1				
B	1/4	1	1/4	0				
C	1/8	0	1/8	1	1/8	1		
D	1/8	0	1/8	1	1/8	0		
E	1/8	0	1/8	0	1/8	1		
F	1/16	0	1/16	0	1/16	0	1/16	1
G	1/16	0	1/16	0	1/16	0	1/16	0

Symbol	l_i	Code	P_i
A	2	11	1/4
B	2	10	1/4
C	3	011	1/8
D	3	010	1/8
E	3	001	1/8
F	4	0001	1/16
G	4	0000	1/16

$$H(S) = \frac{2}{4}\log_2(4) + \frac{3}{8}\log_2(8) + \frac{1}{16}\log_2(16) = 2.625 \text{ bits/symbol}$$

$$L = \sum_{\forall i} P_i l_i = 2.625 \text{ bits/symbol.}$$

$$\eta = \frac{H(S)}{L} = 1 = 100\% \quad \text{Redundancy, } \underline{R = 0\%}$$

Shanno - Fano Ternary Code.

A	$\boxed{1/4}$	2		
B	1/4	1	$\boxed{1/4}$	2
C	1/8	1	$\boxed{1/8}$	1
D	1/8	1	$\boxed{1/8}$	0
E	1/8	0	$\boxed{1/8}$	2
F	1/16	0	$\boxed{1/16}$	1
G	1/16	0	$\boxed{1/16}$	0

$$\alpha = \frac{7-3}{3-1} = \frac{4}{2} = 2$$

Symbol	P_i	l_i	Code
A	1/4	1	2
B	1/4	2	12
C	1/8	2	11
D	1/8	2	10
E	1/8	2	02
F	1/16	2	01
G	1/16	2	00

$$H(s) = \sum_{\forall i} P_i \log_3 (1/P_i) = \frac{2.625}{\log_2 3} = 1.656 \text{ bits/symbol}$$

$$L = \sum_{\forall i} l_i P_i = 1.75 \text{ bits/symbol}$$

$$\text{Efficiency, } \eta = \frac{H(S)}{L} = 94.64\%$$

$$\text{Redundancy} = (100 - 94.64)\% = \underline{5.36\%}$$

06] Shannon's Encoding Algorithm:

Step 1:

	Symbol	probability	
1	a	9/32	0.28125
2	b	9/32	0.28125
3	c	3/32	0.09375
4	d	3/32	0.09375
5	e	3/32	0.09375
6	f	3/32	0.09375
7	g	2/32	0.0625

Step 2:

Calculate Sequence,

$\alpha_1 = 0$

$\alpha_2 = \alpha_1 + P_1 = 9/32$ 0.28125

$\alpha_3 = \alpha_2 + P_2 = 9/16$ 0.5625

$\alpha_4 = \alpha_3 + P_3 = 21/32$ 0.65625

$\alpha_5 = \alpha_4 + P_4 = 24/32$ 0.75

$\alpha_6 = \alpha_5 + P_5 = 27/32$ 0.84375

$\alpha_7 = \alpha_6 + P_6 = 30/32$ 0.9375

$\alpha_8 = \alpha_7 + P_7 = 32/32$ 1

Step 3:

Calculate $l_i = \log_2 (1/P_i)$

$l_1 = \log_2\left(\frac{1}{P_1}\right) = 1.83 \simeq 2$

$l_2 = \log_2\left(\frac{1}{P_2}\right) = 1.83 \simeq 2$

$l_3 = \log_2\left(\frac{1}{P_3}\right) = 3.4 \simeq 4 = l_4 = l_5 = l_6$

$l_7 = \log_2\left(\frac{1}{P_7}\right) = 4$

Date
Page

Step 4 : decimal to binary

*) $\alpha_1 = 0_d$ $\quad c_1 \rightarrow 00_b$

*) $\alpha_2 = 0.28125$

0.28125×2		
0.5625×2	0	01001_b
1.125	1	
0.125×2	0	$l_2 = 2$
0.25×2	0	$c_2 \rightarrow 01_b$
0.5×2	1	
1.0		

*) $\alpha_3 = 0.5625$ $\quad l_3 = 4$

0.5625×2		
1.125	1	
0.125×2		1001_b
0.25×2	0	
0.5×2	0	$c_3 \Rightarrow \underline{\underline{1001_b}}$
1.0	1	

*) $\alpha_4 = 0.65625$ $\quad l_4 = 4$

0.65625×2	$= 1.3125$	1	10101_b
0.3125×2	$= 0.625$	0	
0.625×2	$= 1.25$	1	$c_4 \rightarrow 1010_b$
0.25×2	$= 0.5$	0	
0.5×2	$= 1.0$	1	

*) $\alpha_5 = 0.75$ $\quad l_5 = 4$

0.75×2	$= 1.5$	1	11_b
0.5×2	$= 1.0$	1	$c_5 \rightarrow 1100_b$

*) ~~$\alpha_6 = 0.84375$~~

~~0.84375×2~~

*) $\alpha_6 = 0.84375$ $\quad l_6 = 4$

$0.84375 \times 2 = 1.6875$	1	11011_b
$0.6875 \times 2 = 1.375$	1	
$0.375 \times 2 = 0.75$	0	$C_6 \rightarrow 1101_b$
$0.75 \times 2 = 1.5$	1	
$0.5 \times 2 = 1.0$	1	

*) $\alpha_7 = 0.9375$ $\quad l_7 = 4$

$0.9375 \times 2 = 1.875$	1	1111_b
$0.875 \times 2 = 1.75$	1	
$0.75 \times 2 = 1.5$	1	$C_7 \rightarrow 1111_b$
$0.5 \times 2 = 1.0$	1	

Step 5:
Tabulate :

Symbol	P_i	l_i	code $[C_i]$
a	$9/32$	2	00
b	$9/32$	2	01
c	$3/32$	4	1001
d	$3/32$	4	1010
e	$3/32$	4	1100
f	$3/32$	4	1101
g	$2/32$	4	1111

$$L = \sum_{\forall i} P_i l_i = 2.875$$

$$H(S) = \sum_{\forall i} P_i \log_2 \left(\frac{1}{P_i}\right) = 2.56$$

Efficiency, $\eta = \frac{H(S)}{L} = 0.6606 = 66.06\%$

Redundancy, $R = (1 - \eta)\, 100\% = 33.94\%$

03 a) R/L B/w Hartley, nats and bits.

→ $I_k = \log_e\left(\frac{1}{P_k}\right)$ nats $\longrightarrow$ ①

$I_k = \log_{10}\left(\frac{1}{P_k}\right)$ decit or Hartley $\longrightarrow$ ②

$I_k = \log_2\left(\frac{1}{P_k}\right)$ bits $\longrightarrow$ ③

① = ②

$\log_e\left(\frac{1}{P_k}\right)$ nats $= \log_{10}\left(\frac{1}{P_k}\right)$ decit

$1 \text{ nat} = \frac{\log_{10}(1/P_k)}{\log_e(1/P_k)} = \frac{\log_{10}(1/P_k) \times \log_{10} e}{\log_{10}(1/P_k)}$ decit

$1 \text{ nat} = 0.434$ decit

① = ③

$\log_e\left(\frac{1}{P_k}\right)$ nats $= \log_2\left(\frac{1}{P_k}\right)$ bits

$1 \text{ nat} = \frac{\log_2(1/P_k)}{\log_2(1/P_k)} \times \log_2 e = 1.44$ bits

② = ③

$\log_{10}(1/P_k)$ decit $= \log_2(1/P_k)$ bits

$1 \text{ decit} = \frac{\log_2(1/P_k)}{\log_2(1/P_k)} \log_2 10$

$1 \text{ decit} = 3.32$ bits

03] b] Additivity property of Entropy

Suppose a source S emits symbols s_1 to s_q with probability P_1 to P_q.
Symbol s_q is divided into n sub symbols

$$s_q \rightarrow s_{q1}, s_{q2}, \ldots, s_{qn}$$
$$p_q \rightarrow p_{q1}, p_{q2}, \ldots, p_{qn}$$

$$\sum_{i=1}^{q} P_i = 1 \qquad \sum_{j=1}^{n} P_{qj} = P_q$$

$$H'(S) = \sum_{i=1}^{q-1} p_i \log_2\left(\frac{1}{p_i}\right) + \sum_{j=1}^{n} P_{qj} \log_2\left(\frac{1}{P_{qj}}\right)$$

$$= \sum_{i=1}^{q} p_i \log_2\left(\frac{1}{p_i}\right) - P_q \log_2\left(\frac{1}{P_q}\right) + \sum_{j=1}^{n} P_{qj} \log_2\left(\frac{1}{P_{qj}}\right)$$

$$H'(S) = H(S) + \underbrace{\sum_{j=1}^{n} P_{qj} \log_2\left(\frac{P_q}{P_{qj}}\right)}_{+\text{ive quantity}}$$

Thus $\underline{H'(S) \geq H(S)}$

7.

a. State shannon's first theorem and prove that $\lim_{n\to\infty}\left(\frac{L_n}{n}\right) = H_r(S)$ where L_n is average length of the code word, n is number of symbols & $H_r(S)$ is entropy of a r-ary source.

Ans :- Shannon's 1st theorem

Suggested that the length l_i can be known using

$$l_i = \log_r \frac{1}{P_i} \longrightarrow (1)$$

If l_i happens to be a fraction grounded upto next upper Integer

$$\log_r \frac{1}{P_i} < l_i < 1 + \log_r \frac{1}{P_i}$$

$$\Rightarrow \frac{\log_2 1/P_i}{\log_2 r} \le l_i \le 1 + \frac{\log_2 1/P_i}{\log_2 r}$$

$$\Rightarrow \frac{1}{\log_2 r} \sum_{\forall i} P_i \log_2 \frac{1}{P_i} \le \sum_{\forall i} P_i l_i \le \sum_{\forall i} P_i + \frac{1}{\log_2 r} \sum_{\forall i} P_i \log_2 \frac{1}{P_i}$$

$$H_r(S) \le L \le \sum_{\forall i} P_i + H_r(S)$$

$$L \geq H_r(S) \longrightarrow (2)$$

Shannon's suggest that eq^n (2) holds good for n extension of 'S' for better efficiency

n^{th} extension $\rightarrow H_r(S^n) \le L_n \le \sum_{\forall i} P_i$

where L_n is the average length of the code n^{th} extension

$$H(S^n) = n H(S)$$

$$\sum_{\forall i} P_i = 1$$

$$n H_r(S) \le L_n \le 1 + n H_r(S)$$

$$H_r(S) \le \frac{L_n}{n} \le \frac{1}{n} + H_r(S)$$

$$\boxed{\lim_{n \to \infty} \frac{L_n}{n} = H_r(S)}$$

Statement : Given a code alphabet with 'r' symbol & source alphabet of 'q' symbol the average length of code words can be made as close $H_r(S)$ as by increasing the extension

b. Define the following with an example for each

i) Block codes

The code which match each of the symbol of the source as a bit into some finite sequence of code from the code alphabet 'X' and each of these final sequence is called code word

Ex :-

S	S_1	S_2	S_3	S_4
X	00	01	10	11

ii) Non singular code : A block code is said to be non singular if & only if all the code words distict & easily distinguish from one another for any extension

iii) Uniquelly decodable code :- A non singular code is said to be uniquly decode or decodable. If every code word recieved in a long sequence can be uniqully Identified.

iv) Instaneous code :- A uniquely decodable code is instaneous. If it satisfy prefix property

v) Optimal code

An instaneous code is said to be optimal code. If it has minimum average length 'l' for a source with given probability for the source symbol.