CMR
INSTITUTE OF
TECHNOLOGY

USN | | | | | | | | | |

Internal Assesment Test – I

Sub:	Information Theory and Coding			Sec	*ALL*			Code:	18EC54
Date:	12 / 11 / 2021	Duration:	90 mins	Max Marks:	50	Sem:	V	Branch:	ECE

Answer Any FIVE FULL Questions

	Question	Marks	OBE CO	OBE RBT
1	Define the terms Self information content, average information content, average rate of information and hence derive the expression for average information content in long independent sequences.	10	CO1	L1
2	Mention different properties of entropy and prove *extremal* property.	10	CO1	L2
3	Define self-information, entropy and information rate. Consider transmission of pictures in a black and white television, there are about 3.25 $Megapixels/frame$. For a good reproduction, 20 brightness levels are necessary. Assuming that all the levels are equally likely to occur, find the rate of transmission if one frame is transmitted in every 3 $msec$.	10	CO1	L3
4	Obtain the relationship between Hartley, nats and bits. Consider a discrete memory less source with $S = \{s_1, s_2, s_3, s_4\}$ with $P = \{0.5, 0.2, 0.2, 0.1\}$ show that $H(S^2) = 2H(S)$.	10	CO1	L2
5	For the markov model shown the fig. 5, compute State probabilities, State entropies, Source entropy and Average rate of information if the symbol rate is 1000 symbols/sec.	10	CO1	L4

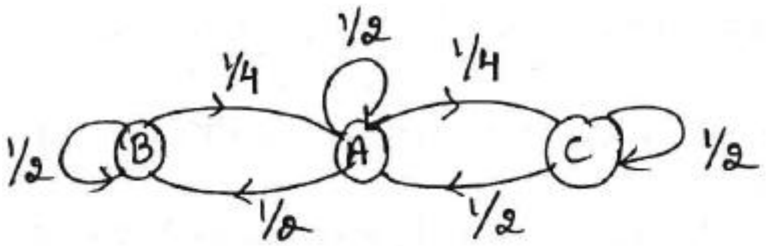


Fig. 5

	Question	Marks	OBE CO	OBE RBT
6	The state diagram of a Markov source is shown in the fig. 6. Show that $G_1 \geq G_2 \geq H$.	10	CO1	L3

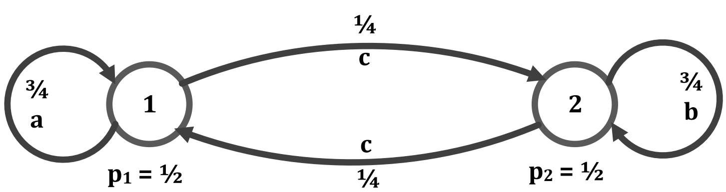


Fig. 6

---------------------***---------------------

IAT-1 SOLUTIONS

A1) **SELF INFORMATION CONTENT :**

Let m_k be a symbol for transmission at any instant of time with probability 'p_k', then amount of information (or) self information of 'm_k' is $I_k = \log_2 \frac{1}{p_k}$ bits or $I_k = \log_{10} \frac{1}{p_k}$ Hartley or $I_k = \log_e \frac{1}{p_k}$ nats

AVERAGE INFORMATION CONTENT :

→ Suppose we have a source that emits one of 'm' symbols $s_1, s_2, \ldots, s_m$ in a statistically independent order.

→ Let $p_1, p_2, p_3, \ldots, p_m$ be the probability of occurance of m symbols respectively, therefore, in a long message of N symbols,

s_1 will occur on an average of $p_1 N$ times

s_2 will occur on an average of $p_2 N$ times

⋮

s_i will occur on an average of $p_i N$ times

→ Treating individual symbols as message of length 1, then the i^{th} symbol information content will be,

$$I(s_i) = \log_2 \frac{1}{p_i} \text{ bits}$$

$$\Rightarrow I_{(total)} = \sum_{i=1}^{i=m} p_i N \log_2 \frac{1}{p_i} \text{ bits}$$

$$I_{(total)} = N \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} \text{ bits}$$

$$\therefore \text{Entropy } (H \text{ or } H(S)) = \frac{I_{(total)}}{N} = \frac{N \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i}}{N} \text{ bits/sym.}$$

$$\boxed{\therefore H = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} \text{ bits/sym.}}$$ → average information content (Entropy) of long independent sequences.

AVERAGE RATE OF INFORMATION :

If a source emits symbol at a fixed time rate r_s symbols per second, then the average rate of information is given by R_s and $R_s = r_s H(S)$ bits/sec.

A2) PROPERTIES OF ENTROPY :

PROPERTY ① : The entropy function is continuous for every value of variable p_k over the interval $(0,1)$

PROPERTY ② : The entropy function is symmetrical of its arguments $H[p_k, (1-p_k)] = H[(1-p_k), p_k]$ for every $K = 1, 2, \ldots, m$ where m is number of symbols.

PROPERTY ③ : Splitting a symbol into sub-symbols does not decrease the uncertainity of the source ($\therefore H' \geq H$)

PROPERTY ④ : EXTREMAL PROPERTY :

(To show entropy has boundary)

→ Consider a discrete memoryless source. The entropy $H(S)$ is bounded by $0 \leq H(S) \leq [H(S)_{max}]$

- The lower boundary is obviously 0 for $p_k = 1$
- For upper boundary, let us consider a quantity.

$$\log_2 m - H \quad \text{—①}$$

$$H = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} \text{ bits/sym} \quad \text{—②}$$

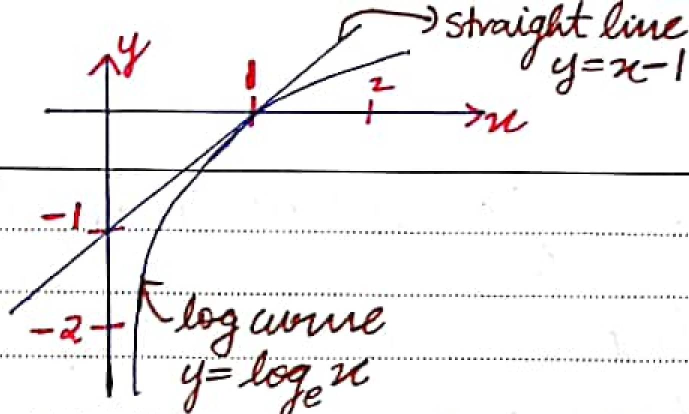
$$\sum_{i=1}^{m} p_i = 1 \quad \text{—③}$$

$$\therefore \log_2 m - H = \sum_{i=1}^{m} p_i \log_2 m - \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i}$$

$$= \sum_{i=1}^{m} p_i \left[\log_2 m - \log_2 \frac{1}{p_i}\right]$$

$$= \sum_{i=1}^{m} p_i \log_2 m p_i = \sum_{i=1}^{m} p_i \frac{\log_e m p_i}{\log_e 2}$$

$$\therefore \log_2 m - H = \log_2 e \sum_{i=1}^{m} p_i \log_e m p_i \quad \text{—④}$$



→ Consider the figure, it is evident that the straight line $y=x-1$ is always above the log curve $y=\log_2 x$ except at $x=1$.

→ $\therefore$ The straight line forms the tangent at $x=1$.

$$\therefore \log_e x \leq x-1 \text{ always.} \quad —⑤$$

→ Multiply by − sign throughout, $-\log_e x \geq 1-x$

$$\log_e \frac{1}{x} \geq 1-x$$

→ If $x=\frac{1}{mp_i} \Rightarrow \log_e mp_i \geq 1-\frac{1}{mp_i}$

→ Multiple the above eq. by p_i and then take $\sum$ for all i,

$$\sum_{i=1}^{m} p_i \log_e mp_i \geq \sum_{i=1}^{m} p_i - \sum_{i=1}^{m} \frac{1}{m}$$

→ Multiple by $\log_2 e$:

$$\log_2 e \sum_{i=1}^{m} p_i \log_e mp_i \geq \log_2 e \left[\sum_{i=1}^{m} p_i - \sum_{i=1}^{m} \frac{1}{m}\right] \quad —⑥$$

$\therefore$ LHS of the above equation is $\log_e m - H$ and RHS is always 0 for any condition.

$$\therefore \log_2 m - H \geq 0 \text{ always.} \qquad \boxed{\therefore H \leq \log_2 m}$$

→ The equality sign holds good when $(p_i - \frac{1}{m}) = 0$ for every i which means $p_i = \frac{1}{m}$ for every $i=1,2,\ldots,m$.

→ When this condition is satisfied, Eq ⑥ achieves max entropy

$$\boxed{\therefore H(S)_{max} = \log_2 m \text{ bits/sym.}}$$

$\therefore$ entropy attains MAX value when symbols become equi-probable.

A3) SELF INFORMATION :

Let m_k be a symbol for transmission at any instant of time with probability 'p_k', then amount of information (or) self information of m_k is given by I_k.

$$I(m_k) = \log \frac{1}{P_k}$$

ENTROPY :

Suppose we have a source that emits one of 'm' symbols $s_1, s_2, s_3, \ldots, s_m$ in a statistically independent order.

Let $p_1, p_2, p_3, \ldots, p_m$ be the probability of occurrance of m symbols respectively, then average information content (or) entropy of long independent sequences is given by H or H(S)

$$H = \sum_{i=1}^{m} P_i \log_2 \frac{1}{P_i} \text{ bits/sym.}$$

INFORMATION RATE :

If a source emits symbol at a fixed time rate r_s symbols per second, then the average rate of information is R_s and $R_s = r_s H(S)$ bits/sec.

⇒ Total number of pixels per frame is 3.25 Megapixels/frame
$= 3.25 \times 10^6$ pixels/frame

It is given that each element can have 20 brightness levels, then, total number of different frames possible are

$$m = (20)^{3.25 \times 10^6} \text{ frames}$$

⇒ Let us assume all the frames occurs at equal probability, then, w.k.t, the entropy is given by $H(S)_{max}$.

$$H(S)_{max} = \log_2 m$$
$$= \log_2 (20)^{3.25 \times 10^6}$$
$$= (3.25 \times 10^6) \log_2 (20)$$
$$= 14.046 \times 10^6 \text{ bits/frames}$$

$\therefore H(S)_{max} = 14.046 \times 10^6$ bits/frames

$\therefore$ Average rate of information $= R_s = r_s H(S)_{max}$

Given $r_s = \frac{1}{3}$ frames/sec.

$\therefore R_s = r_s H(S) = \frac{1}{3}(14.046 \times 10^6) = 4682000$ bits/sec

$\therefore R_s = 4682000$ bits/sec.

A4) $I = \log_2 \frac{1}{P_i}$ bits, $I = \log_e \frac{1}{P_i}$ nats, $I = \log_{10} \frac{1}{P_i}$ Hartley

Equating them, $\log_2 \frac{1}{P_i} = \log_e \frac{1}{P_i} = \log_{10} \frac{1}{P_i}$

i) Hartley to bits : $\log_2 \frac{1}{P_i}$ bits $= \log_2 \frac{1}{P_i}$ Hartley

$$1 \text{Hartley} = \frac{\log_2 \frac{1}{P_i}}{\log_{10} \frac{1}{P_i}} = \log_2 10 = 3.3219$$

$\therefore$ 1Hartley = 3.3219 bits

ii) Nats to bits : $\log_2 \frac{1}{P_i}$ bits $= \log_e \frac{1}{P_i}$ nats

$$1 \text{Nats} = \frac{\log_2 \frac{1}{P_i}}{\log_e \frac{1}{P_i}} = \log_2 e = 1.4426$$

$\therefore$ 1 Nats = 1.4426 bits

iii) Hartley to nats : $\log_e \frac{1}{P_i}$ nats $= \log_{10} \frac{1}{P_i}$ Hartley

$$1 \text{Hartley} = \frac{\log_e \frac{1}{P_i}}{\log_{10} \frac{1}{P_i}} = \log_e 10 = 2.3025$$

$\therefore$ 1Hartley = 2.3025 nats

$\Rightarrow$ Given, $S = \{S1, S2, S3, S4\}$ and $P = \{0.5, 0.2, 0.2, 0.1\}$

$$H(S) = \sum_{i=1}^{4} P_i \log_2 \frac{1}{P_i} \text{ bits/sym.}$$

$$H(S) = P_{S1} \log_2 \frac{1}{P_{S1}} + P_{S2} \log_2 \frac{1}{P_{S2}} + P_{S3} \log_2 \frac{1}{P_{S3}} + P_{S4} \log_2 \frac{1}{P_{S4}}$$

$$= 0.5 \log_2 \frac{1}{0.5} + 0.2 \log_2 \frac{1}{0.2} + 0.2 \log_2 \frac{1}{0.2} + 0.1 \log_2 \frac{1}{0.1}$$

$$= 0.5 + 0.9287 + 0.3321 = 1.76089 \text{ bits/sym}$$

$\therefore H(S) = 1.76089$ bits/sym.

$$2H(S) = 2 \times 1.76089 \text{ bits/sym}$$

$$2H(S) = 3.52178 \text{ bits/sym.} \quad \text{—} ①$$

IN SECOND EXTENSION $\Rightarrow (4)^2 = 16$

$S_1S_1 \Rightarrow P_1P_1 = 0.25$	$S_3S_1 \Rightarrow P_3P_1 = 0.1$
$S_1S_2 \Rightarrow P_1P_2 = 0.1$	$S_3S_2 \Rightarrow P_3P_2 = 0.04$
$S_1S_3 \Rightarrow P_1P_3 = 0.1$	$S_3S_3 \Rightarrow P_3P_3 = 0.04$
$S_1S_4 \Rightarrow P_1P_4 = 0.05$	$S_3S_4 \Rightarrow P_3P_4 = 0.02$
$S_2S_1 \Rightarrow P_2P_1 = 0.1$	$S_4S_1 \Rightarrow P_4P_1 = 0.05$
$S_2S_2 \Rightarrow P_2P_2 = 0.04$	$S_4S_2 \Rightarrow P_4P_2 = 0.02$
$S_2S_3 \Rightarrow P_2P_3 = 0.04$	$S_4S_3 \Rightarrow P_4P_3 = 0.02$
$S_2S_4 \Rightarrow P_2P_4 = 0.02$	$S_4S_4 \Rightarrow P_4P_4 = 0.01$

$$H(S^2) = \sum_{j=1}^{16} P_j \log_2 \frac{1}{P_j}$$

$$H(S^2) = P_1P_1 \log_2(1/P_1P_1) + P_1P_2 \log_2(1/P_1P_2) + P_1P_3 \log_2(1/P_1P_3) + P_1P_4 \log_2(1/P_1P_4)$$
$$+ P_2P_1 \log_2(1/P_2P_1) + P_2P_2 \log_2(1/P_2P_2) + P_2P_3 \log_2(1/P_2P_3) + P_2P_4 \log_2(1/P_2P_4)$$
$$+ P_3P_1 \log_2(1/P_3P_1) + P_3P_2 \log_2(1/P_3P_2) + P_3P_3 \log_2(1/P_3P_3) + P_3P_4 \log_2(1/P_3P_4)$$
$$+ P_4P_1 \log_2(1/P_4P_1) + P_4P_2 \log_2(1/P_4P_2) + P_4P_3 \log_2(1/P_4P_3) + P_4P_4 \log_2(1/P_4P_4)$$

$$H(S^2) = 0.25 \log_2(1/0.25) + 4[0.1 \log_2(1/0.1)] + 2\left[0.05 \log_2\left(\frac{1}{0.05}\right)\right]$$
$$+ 4[0.04 \log_2(1/0.04)] + 4\left[0.02 \log_2\left(\frac{1}{0.02}\right)\right] + 0.01 \log_2\left(\frac{1}{0.01}\right)$$

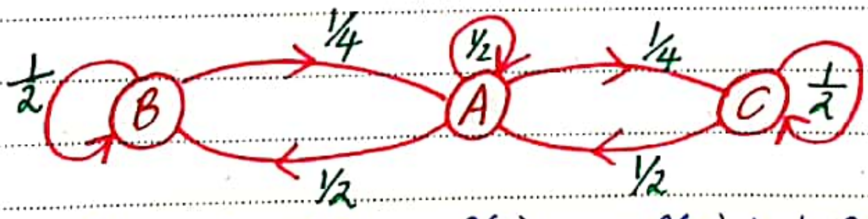
$$H(S^2) = 0.5 + 1.3287 + 0.4321 + 0.74301 + 0.4515 + 0.06643$$

$$H(S^2) = 3.527 \text{ bits/sym.} \quad \text{—} ②$$

$\therefore$ From ① and ② $\Rightarrow$ $H(S^2) = 2H(S)$

Hence Proved.

A5)



STATE PROBABILITIES : $P(A) = \frac{1}{2}P(A) + \frac{1}{4}P(B) + \frac{1}{2}P(C)$

$$P(B) = \frac{1}{2}P(B) + \frac{1}{2}P(A) \Rightarrow P(B) = P(A)$$

$$P(C) = \frac{1}{2}P(C) + \frac{1}{4}P(B) \Rightarrow P(C) = \frac{1}{2}P(A)$$

W.k.t, $P(A) + P(B) + P(C) = 1$

$$P(A) + P(A) + \frac{1}{2}P(A) = 1$$

$$\therefore P(A) = \frac{2}{5}, \quad P(B) = \frac{2}{5}, \quad P(C) = \frac{1}{5}$$

STATE ENTROPIES : $H_i = \sum_{j=1}^{n} P_{ij} \log_2\left(\frac{1}{P_{ij}}\right)$ bits/sym

For A : $H_A = P_{AA}\log_2\left(\frac{1}{P_{AA}}\right) + P_{AB}\log_2\left(\frac{1}{P_{AB}}\right) + P_{AC}\log_2\left(\frac{1}{P_{AC}}\right)$

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

$$H_A = 3/2 \text{ bits/sym.}$$

For B : $H_B = P_{BA}\log_2\left(\frac{1}{P_{BA}}\right) + P_{BB}\log_2\left(\frac{1}{P_{BB}}\right) + P_{BC}\log_2\left(\frac{1}{P_{BC}}\right)$

$$= \left(\frac{1}{4}\right)\log_2(4) + \left(\frac{1}{2}\right)\log_2(2) + 0$$

$$H_B = 1 \text{ bits/sym.}$$

For C : $H_C = P_{CA}\log_2\left(\frac{1}{P_{CA}}\right) + P_{CB}\log_2\left(\frac{1}{P_{CB}}\right) + P_{CC}\log_2\left(\frac{1}{P_{CC}}\right)$

$$= \left(\frac{1}{2}\right)\log_2 2 + 0 + \left(\frac{1}{2}\right)\log_2(2)$$

$$H_C = 1 \text{ bits/sym.}$$

SOURCE ENTROPY : $H = \sum_{i=A}^{C} P_i H_i = P_A H_A + P_B H_B + P_C H_C$

$$= \left(\frac{2}{5} \times \frac{3}{2}\right) + \left(\frac{2}{5} \times 1\right) + \left(\frac{1}{5} \times 1\right)$$

$$\boxed{H = 1.2 \text{ bits/sym}}$$

$\Rightarrow$ Given, $r_s = 1000$ sym/sec

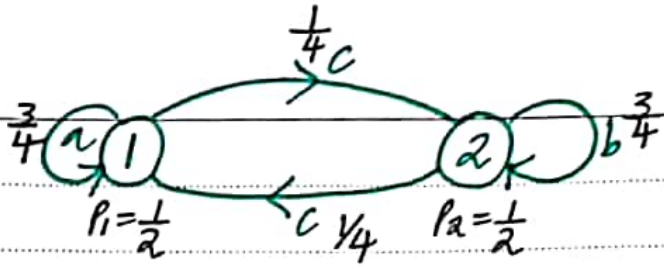
$$R_s = r_s H(S) = 1000(1.2)$$

$$R_s = 1200 \text{ bits/sec}$$

$$\therefore R_s = 1200 \text{ bits/sec}$$

A6) To show that, $G_1 \geq G_2 \geq H$

Scanned with CamScanner



STATE ENTROPY: $H_i = \sum_{j=1}^{2} P_{ij}\log_2 \frac{1}{P_{ij}}$

For $i=1$ $\Rightarrow$ $H_1 = P_{11}\log_2\frac{1}{P_{11}} + P_{12}\log_2\frac{1}{P_{12}}$

$$= \frac{3}{4}\log_2\frac{4}{3} + \frac{1}{4}\log_2(4)$$

$$H_1 = 0.8113 \text{ bits/sym.}$$

For $i=2$ $\Rightarrow$ $H_2 = P_{21}\log_2\frac{1}{P_{21}} + P_{22}\log_2\frac{1}{P_{22}}$

$$= \frac{1}{4}\log_2 4 + \frac{3}{4}\log_2\frac{4}{3}$$

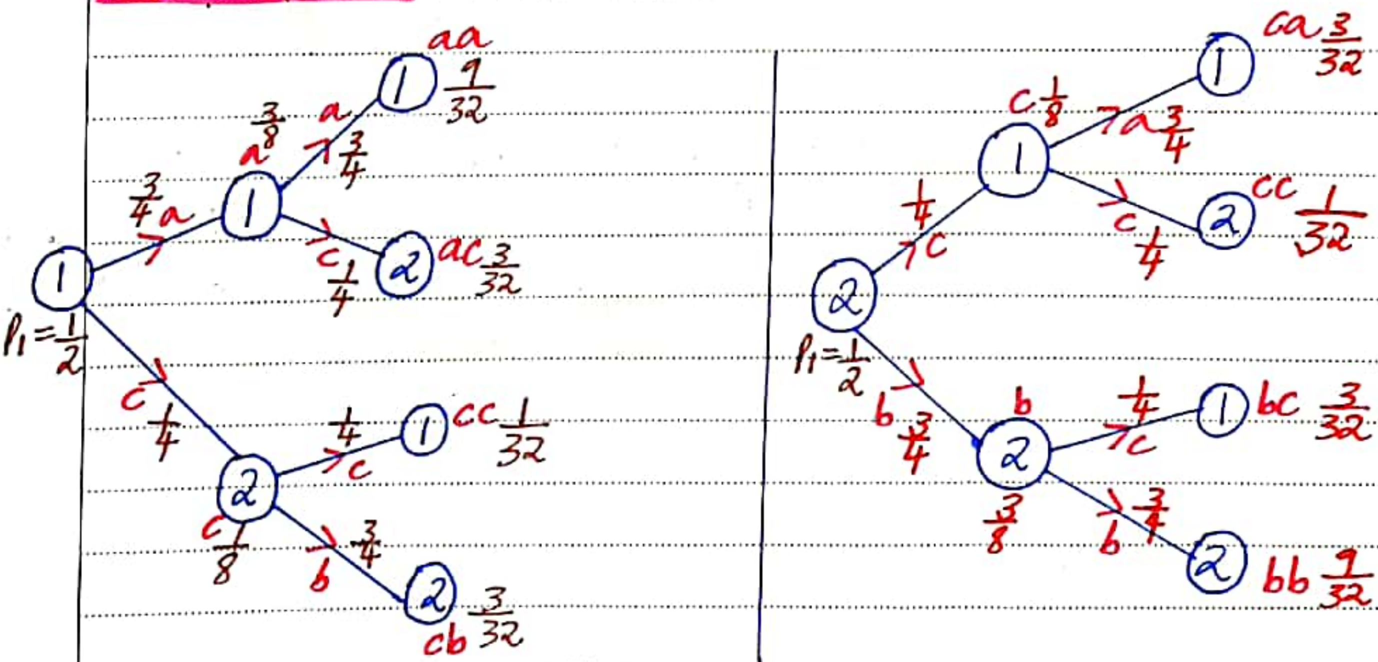
$$H_2 = 0.8113 \text{ bits/sym.}$$

SOURCE ENTROPY: $H = \sum_{i=1}^{2} P_i H_i$ bits/sym

$$H = \frac{1}{2}(0.8113) + \frac{1}{2}(0.8113)$$

$$H = 0.8113 \text{ bits/sym.}$$

TREE DIAGRAM:



I interval

Symbol	Probability
a	3/8
b	3/8
c	$\frac{1}{8}+\frac{1}{8}=\frac{2}{8}$
	$\frac{3}{8}+\frac{3}{8}+\frac{2}{8}=1$

II interval

Sym $(3)^2$	Probability
aa	9/32
$ab=0 \leftarrow$ ac	3/32
bb	9/32
$ba=0 \leftarrow$ bc	3/32
ca	3/32
cb	3/32
cc	$\frac{1}{32}+\frac{1}{32}=\frac{2}{32}$

$$\Rightarrow G_N = \frac{1}{N}\sum_{\forall p} P_{mi} \log_2 \frac{1}{P_{mi}}$$

$$G_1 = \frac{1}{1}\sum_{i=a}^{c} P_{mi}\log_2\frac{1}{P_{mi}} = P_a\log_2\frac{1}{P_a} + P_b\log_2\frac{1}{P_b} + P_c\log_2\frac{1}{P_c}$$

$$= \frac{3}{8}\log_2\frac{8}{3} + \frac{3}{8}\log_2\frac{8}{3} + \frac{2}{8}\log_2\frac{8}{2}$$

$$= 0.5306 + 0.5306 + 0.5$$

$$\therefore G_1 = 1.5612 \text{ bits/sym.}$$

$$G_2 = \frac{1}{2}\sum_{i=a}^{c} P_{mi}\log_2\frac{1}{P_{mi}}$$

$$= \frac{1}{2}\Big[P_{aa}\log_2\frac{1}{P_{aa}} + P_{ab}\log_2\frac{1}{P_{ab}} + P_{ac}\log_2\frac{1}{P_{ac}} + P_{ba}\log_2\frac{1}{P_{ba}} + P_{bb}\log_2\frac{1}{P_{bb}}$$

$$+ P_{bc}\log_2\frac{1}{P_{bc}} + P_{ca}\log_2\frac{1}{P_{ca}} + P_{cb}\log_2\frac{1}{P_{cb}} + P_{cc}\log_2\frac{1}{P_{cc}}\Big]$$

$$= \frac{1}{2}\Big[\left(\left(\tfrac{9}{32}\right)\log_2\left(\tfrac{32}{9}\right)\right)\times 2 + \left(0\log_2(0)\right)\times 2 + \left(\left(\tfrac{3}{32}\right)\log_2\left(\tfrac{32}{3}\right)\right)\times 4 + \left(\tfrac{2}{32}\right)\log_2\left(\tfrac{32}{2}\right)\Big]$$

$$= \frac{1}{2}\left[0.5147 + 0.3201 + 0.5147 + 0.3201 + 0.3201 + 0.3201 + 0.25\right]$$

$$= \frac{1}{2}[2.559] = 1.2799$$

$$\therefore G_2 = 1.2799$$

$$\boxed{\therefore G_1 \geq G_2 \geq H}$$