CMR
INSTITUTE OF
TECHNOLOGY

Internal Assessment Test 1– Sep. 2016

Sub:	Information Theory & Coding					**Code:**	10EC55
Date:	06 / 09 /16	**Duration:** 90 mins	**Max Marks:** 50	**Sem:**	V C	**Branch:**	ECE

Note: Answer any five full questions.

1. **(a).** The output of an information source consists of 185 symbols, 50 of which occurs with probability of 1/100 and remaining 135 occur with a probability of 1/270. The source emits 480 symbols/sec. Assume that the symbols are chosen independently, find the rate of information of the source. **[05]**

H(S) (03)

R_S (02)

(b). Prove the ADDITIVITY property of entropy. **[05]**

Proof for H' ≥ H (05)

2. The state diagram of a Markov source is shown in the Fig. 2; find the state probabilities, entropy of each state and entropy of the source. **[10]**

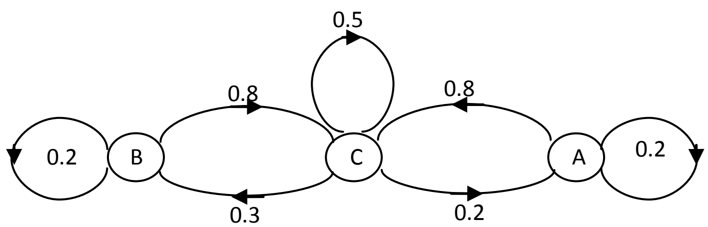


Fig. 2

P(A) (01)

P(B) (01)

P(C) (01)

H_A (02)

H_B (02)

H_C (02)

H(S) (01)

3. Apply Shannon's encoding technique for the message, **CMR**, where each letter is associated with probabilities 0.2, 0.5 and 0.3 respectively. Calculate the code redundancy and if the same is applied for second extension, what is the impact on the redundancy? **[10]**

Coding using I extension (02)

Redundancy (01)

Coding using II extension (04)

Redundancy (02)

Reduction in Redundancy (01)

4. Given the message A to H with respective probabilities 0.2, 0.02, 0.1, 0.12, 0.3, 0.07, 0.04 and 0.15. Construct a binary code using Shannon- Fano encoding procedure and find the efficiency and redundancy. Draw the code tree for the so formed code. **[10]**

Encoding (08)
Code tree (02)

5. Consider a zero memory source with S = {E, O, I, N, S, C, F} appearing with probabilities P = {0.4, 0.2, 0.1, 0.1, 0.1, 0.05, 0.05} respectively. Hence encode **'*SEEN CONFESSION*',** using HUFFMAN coding. **[10]**

Encoding (08)
Coding SEEN CONFESSION (02)

6. **(a).** A CRT terminal is used to enter alphanumeric data into a computer. The CRT is connected to the computer through a voice graded telephone having a usable bandwidth of 3KHz and an output S/N ratio of 10dB. Assume that the terminal has 128 characters and that the data sent from the terminal consist of independent sequences of equiprobable characters.

(i). Find the capacity of the channel.

(ii). Find the maximum rate at which data can be transmitted from the terminal to the computer without errors. **[04]**

Channel capacity (02)
Rsmax (02)

(b) Consider the MORSE code, assume a dash is 3 times as long as a dot and has one third probability of occurrence. Calculate

(i). The information in a dot and a dash

(ii). The entropy of the code.

(iii). The average rate of information if a dot lasts for 10ms and this time is allowed between symbols. **[06]**

I_{dot} (01)
I_{dash} (01)
$H(S)$ (02)
R_S (02)

7. **(a).** State and prove KRAFT Mc-MILLAN inequality. **[05]**

Proof for $\sum_{j=0}^{K-1} r^{-l_j} \leq 1$ (05)

(b). If a binary memory-less source is emitting independent sequence of 0's and 1's with probabilities p and 1-p respectively. Plot entropy of the source versus p. **[05]**

Calculation of $H(S)$ (03)
Plot of $H(S)$ *v/s p* (02)

1a. W.k.T. $H(s) = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i}$

$$= \sum_{i=1}^{185} p_i \log_2 \frac{1}{p_i}$$

$$= 50\left[\frac{1}{100}\log_2 100\right] + 135\left[\frac{1}{270}\log_2 270\right]$$

$$= \underline{\underline{7.360}} \text{ bits/sym}$$

Given $r_s = 480$ sym/sec

The rate of transmission is $R_s = r_s H(s)$

$$= (7.360)(480)$$

$$= \underline{\underline{3532.96}} \text{ bits/sec}$$

b. Consider $S = \{s_1, s_2, s_3 \ldots s_q\}$ with $P = \{p_1, p_2, p_3 \ldots p_q\}$ respectively

Then if q^{th} sym is split into 'n' different symbols i.e.,

$s_q = s_{q1}, \cancel{s_{q2}}, \ldots s_{qn}$ with $\{p_q\} = \{p_{q1}, p_{q2} \ldots p_{qn}\}$ respectively

such that $p_q = p_{q1} + p_{q2} + \ldots + p_{qn} = \sum_{j=1}^{n} p_{qj}$

Then the split symbol entropy is given by

$$H' = H(p_1, p_2, \ldots p_{q-1}, p_{q1}, p_{q2}, \ldots p_{qn})$$

$$= \sum_{i=1}^{q-1} p_i \log_2 \frac{1}{p_i} + \sum_{j=1}^{n} p_{qj} \log_2 \frac{1}{p_{qj}}$$

$$= \sum_{i=1}^{q} p_i \log_2 \frac{1}{p_i} - p_q \log_2 \frac{1}{p_q} + \sum_{j=1}^{n} p_{qj} \log_2 \frac{1}{p_{qj}}$$

$$H' = H + p_q \sum_{j=1}^{n} \frac{p_{qj}}{p_q}\left[\log_2 \frac{p_q}{p_{qj}}\right]$$

$= H +$ a positive quantity since $p_{qj} \le p_q$; for all j

$\therefore \boxed{H' \geq H(s)}$

2. $P(A) = 0.2\,P(A) + 0.2\,P(C) \rightarrow (1)$

$P(B) = 0.2\,P(B) + 0.3\,P(C) \rightarrow (2)$

$P(C) = 0.8\,P(A) + 0.8\,P(B) + 0.5\,P(C) \rightarrow (3)$

From (1) $\quad P(A) = \frac{1}{4} P(C) \rightarrow (4)$

From (2) $\quad P(B) = \frac{3}{8} P(C) \rightarrow (5)$

W.K.T. $\quad P(A) + P(B) + P(C) = 1 \rightarrow (6)$

$$\Rightarrow \frac{1}{4} P(C) + \frac{3}{8} P(C) + P(C) = 1$$

$$\Rightarrow \boxed{P(C) = \frac{8}{13}} \rightarrow (7)$$

(7) in (4) $\Rightarrow \boxed{P(A) = \frac{2}{13}} \rightarrow (8)$

(7) in (5) $\Rightarrow \boxed{P(B) = \frac{3}{13}} \rightarrow (9)$

Entropy each state is given by

$$H_i = \sum_{j=1}^{n} p_{ij} \log_2 \frac{1}{p_{ij}}$$

$$H_A = \sum_{j=A}^{C} p_{Aj} \log_2 \frac{1}{p_{Aj}}$$

$$= p_{AA} \log_2 \frac{1}{p_{AA}} + p_{AB} \log_2 \frac{1}{p_{AB}} + p_{AC} \log_2 \frac{1}{p_{AC}}$$

$$= 0.2 \log_2 \frac{1}{0.2} + 0 + 0.8 \log_2 \frac{1}{0.8}$$

$$= 0.722 \text{ bits/sym}$$

$H_B = 0.722$ bits/sym

$H_C = 1.485$ bits/sym

Entropy of the source is $H = \sum_{i=A}^{C} p_i H_i = 1.192$ bits/symbol

3.

symbols	p_i	Code	l_i
M	0.5	0	1
R	0.3	10	2
C	0.2	110	3

$$H(S) = \sum_{i=C}^{R} p_i \log_2 \frac{1}{p_i}$$

$$= 0.5 \log_2 \frac{1}{0.5} + 0.3 \log_2 \frac{1}{0.3} + 0.2 \log_2 \frac{1}{0.2}$$

$$= \underline{1.4855} \text{ bits/sym}$$

$$L = \sum_{i=C}^{R} p_i l_i$$

$$= 0.5(1) + 2(0.3) + 0.2(3)$$

$$= \underline{1.7 \text{ bits/sym}}$$

$\eta_s = \underline{87.38\%}$

$R_{\eta s} = \underline{12.62\%}$

II extension, total no of symbols $3^2 = 9$

MM	MR	MC	RM	RR	RC	CM	CR	CC
0.25	0.15	0.10	0.15	0.9	0.06	0.10	0.06	0.04

Sym	MM	MR	RM	MC	CM	RR	RC	CR	CC
Prob	0.25	0.15	0.15	0.10	0.10	0.09	0.06	0.06	0.04
CW	00	010	011	1000	1010	1100	11010	11100	11110
	(2)	(3)	(3)	(4)	(4)	(4)	(5)	(5)	(5)

$$L_2 = \sum_{i=1}^{9} p_i l_i = \underline{3.36 \text{ bits/sym}}$$

$H(S^2) = \underline{2.971}$ bits/sym

$\eta_c^{(2)} = \underline{88.42\%}$ $R_{\eta c} = \underline{11.58\%}$

Impact code efficiency is improved by $\underline{1.04\%}$

4.

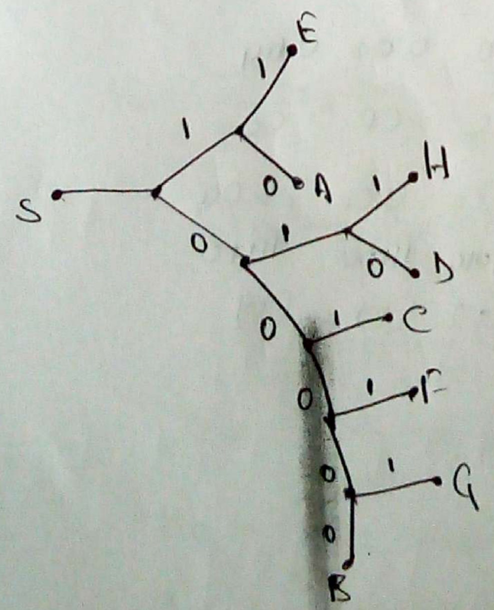
E	0.3	1	0.3	1						
A	0.2	1	0.2	0						
H	0.15	0	0.15	1	0.15	1				
D	0.12	0	0.12	1	0.12	0				
C	0.1	0	0.1	0	0.1	1				
F	0.07	0	0.07	0	0.07	0	0.07	1		
G	0.04	0	0.04	0	0.04	0	0.04	0	0.04	1
B	0.02	0	0.02	0	0.02	0	0.02	0	0.02	0

$H(S) = 2.66$ bits/sym

$L = 2.69$ bits/sym

$\eta_s = \frac{H(S)}{L} = \underline{\underline{98.88\%}}$

$R_{\eta} = \underline{\underline{1.12}}\%$



Scanned by CamScanner

ALAP

5.

sym	prob	cw	prob	cw	prob	cw	prob	cw	prob	cw	prob	cw
E	0.4	0	0.4	0	0.4	0	0.4	0	0.4	0	0.6	1
O	0.2	10	0.2	10	0.2	10	0.2	10	0.4	11	0.4	0
I	0.1	1101	0.1	1101	0.2	111	0.2	111	0.2	10		
N	0.1	1100	0.1	1100	0.1	1101	0.2	110				
S	0.1	1111	0.1	1111	0.1	1100						
C	0.05	11101	0.1	1110								
F	0.05	11100										

SEEN CONFESSION is encoded as.

1111	0	0	1100	11101	10	1100	11100	0	1111	1111	1101	10	1100
S	E	E	N	C	O	N	F	E	S	S	I	O	N

6.a. i Channel capacity

$$C = B\log_2(1+S/N)$$

$$= 3000\log_2(1+10) = 10378.49 \simeq 10378 \text{ bits/sec}$$

ii w.k.t. $R = r_s H < C$

$$H = \log_2(128) = 7 \text{ bits/sym}$$

$$\Rightarrow 7r_s < C$$

$$7r_s < 10378$$

$$r_s < 1482$$

if data is transmitted below 1482 sym/sec then errorless.

Scanned by CamScanner

CMR

6.6 W.K.T

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

Given $P_{dash} = \frac{1}{3} P_{dot}$

$$\therefore P_{dot} + \frac{1}{3} P_{dot} = 1$$

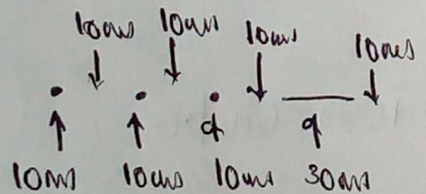
$$\Rightarrow P_{dot} = \frac{3}{4} \quad \& \quad P_{dash} = \frac{1}{4}$$

(1) $I_{dot} = \log_2 \frac{1}{P_{dot}} = 0.415$ bits

$I_{dash} = \log_2 \frac{1}{P_{dash}} = 2$ bits

(2) $H(s) = P_{dot} \log_2 \frac{1}{P_{dot}} + P_{dash} \log_2 \frac{1}{P_{dash}} = 0.8113$ bits/sym

(3)



symbol rate = 4 sym/100ms

= 40 sym/sec

$\therefore R_s = r_s H(s)$

$= (40)(0.8113)$

$R_s = 32.452$ bits/sec