Roll No.												

PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING

TEST 1

Sem & AY: Odd Sem. 2019-20

Course Code: ECE 216

Course Name: INFORMATION THEORY AND CODING

Program & Sem: B.Tech (ECE) & VII.

Date: 01.10.2019

Time: 9:30AM to 10:30AM

Max Marks: 40

Weightage: 20%

Instructions:

i. Question paper consists of 3 parts.

ii. Scientific and Non-programmable calculators are permitted.

Part A [Memory Recall Questions]

Answer all the Questions. Each Question carries four marks. (3Qx4M=12M)

1. A binary source is emitting an independent sequence of '0's and '1's with probabilities p and 1-p respectively. Tabulate the source entropy H(S) versus probabilities.
(C.O.NO.1) [Knowledge]

2. A discrete source emits one of six symbols once every milli-sec. The symbol probabilities are 1/2, 1/4, 1/8, 1/16, 1/32 and 1/32 respectively. Find the average information rate and also express H(S) in nats/symbol. (C.O.NO.1) [Comprehension]

3. State Shannon's first theorem. Define average length of the code?
(C.O.NO.1) [Knowledge]

Part B [Thought Provoking Questions]

Answer both the Questions. Each Question carries eight marks. (2Qx8M=16M)

4. A zero memory source has a source alphabet $S = \{s_1, s_2, s_3\}$ with $P= \{1/2, 1/4, 1/4\}$. Find the source entropy. Also determine the entropy of its 2[nd] extension by listing all elements for 2[nd] extension and verify that $H(S^2) = 2H(S)$. (C.O.NO.1) [Application]

5. Apply Shannon's encoding algorithm to the set of messages $\{m_1, m_2, m_3, m_4, m_5\}$ with their respective probabilities as {1/8, 1/16, 3/16, 1/4, 3/8}. Obtain the code table.
(C.O.NO.1) [Application]

Part C [Problem Solving Questions]

Answer the Questions. The Question carries twelve marks. **(1Qx12M=12M)**

6. Consider the state diagram of the Markov source with a source S = {A, B, C, D} as shown in Fig. 1. (C.O.NO.1) [Application]

   i. Compute the state probabilities using state equations. [5M]
   ii. Find the entropy of each state and source entropy. [5M]
   iii. Find the entropy of the adjoint source. [2M]

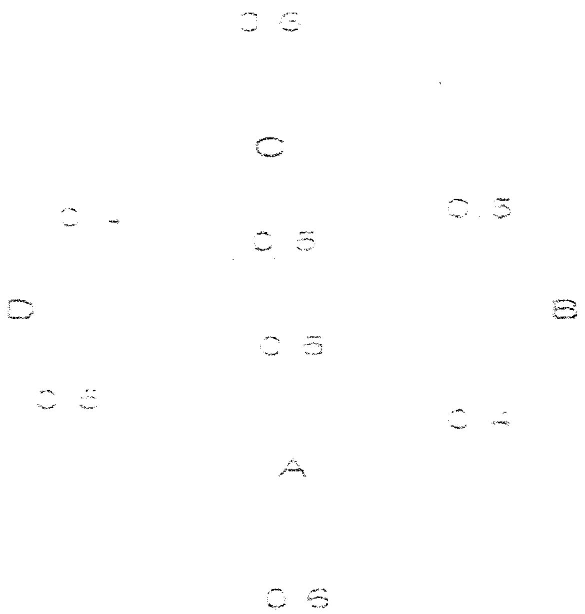

Fig. 1: Markov source

SCHOOL OF ENGINEERING

Semester: VII

Course Code: ECE 216

Course Name: Information Theory and Coding

Date: 01-10-2019

Time: 9.30 AM – 10.30 AM

Max Marks: 40

Weightage: 20%

Extract of question distribution [outcome wise & level wise]

Q.NO.	C.O.NO	Unit/Module Number/Unit /Module Title	Memory recall type [Marks allotted] Bloom's Levels			Thought provoking type [Marks allotted] Bloom's Levels			Problem Solving type [Marks allotted]			Total Marks
			K			C			A			
1	1	1/1/ Fundamentals of Coding Theory and Source Coding	4									4
2	1	1/1/ Fundamentals of Coding Theory and Source Coding				4						4
3	1	1/1/ Fundamentals of Coding Theory and Source Coding	4									4
4	1	1/1/ Fundamentals of Coding Theory and Source Coding							8			8
5	1	1/1/ Fundamentals of Coding Theory and Source Coding							8			8
6	1	1/1/ Fundamentals of Coding Theory and Source Coding							12			12
	Total Marks		8			4			28			40

K =Knowledge Level   C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt.

I here certify that All the questions are set as per the above lines[Sivakumar]

Sign. Needed

Annexure- II: Format of Answer Scheme

SCHOOL OF ENGINEERING

SOLUTION

Semester: VII

Course Code: ECE 216

Course Name: Information Theory and Coding

Date: 01-10-2019

Time: 9.30 AM – 10.30 AM

Max Marks: 40

Weightage: 20%

Part A (3Q x4 M =12 Marks)

<table>
<tr><th>Q No</th><th>Solution</th><th>Scheme of Marking</th><th>Max. Time required for each Question</th></tr>
<tr><td>1</td><td>

Entropy $H(S) = \sum_{i=1}^{2} pi \log \frac{1}{pi}$

$$= p \log \frac{1}{p} + (1-p) \log 1/(1-p)$$

The table 1 below shows tabulation of entropy values versus probabilities

Table 1: H(S) versus p

<table>
<tr><td>p</td><td>0</td><td>0.1</td><td>0.2</td><td>0.3</td><td>0.4</td><td>0.5</td><td>0.6</td><td>0.7</td><td>0.8</td><td>0.9</td><td>1</td></tr>
<tr><td>H(S)</td><td>0</td><td>0.469</td><td>0.722</td><td>0.881</td><td>0.97</td><td>1</td><td>0.97</td><td>0.881</td><td>0.722</td><td>0.469</td><td>0</td></tr>
</table>

</td><td>1
+ 1
+ 2</td><td>6 mins</td></tr>
<tr><td>2</td><td>

Entropy $H(S) = \sum_{i=1}^{2} pi \log \frac{1}{pi}$

$= 0.5 \log 2 + 0.25 \log 4 + 0.125 \log 8 + 0.0625 \log 16 +$
(2) $0.03125 \log 32$ = **1.9375 bits/message-symbol**

1 nat = 1.443 bits

Hence 1 bit = 0.693 nats, so **H(S)= (1.9375)(0.693) = 1.34268 nats/symbol**

Information rate $R_s = r_s H(S)$

Given r_s = 1 message symbol/m-sec = 10^3 message symbols/sec

Therefore **Information rate R_s = 1937.5 bits/sec**

</td><td>2
+ 1
+ 1</td><td>6 mins</td></tr>
<tr><td>3</td><td>

"Shannon's first theorem" or "Noiseless coding theorem " states that "given a code alphabet with 'r' symbols and source alphabet of 'q' symbols, the average length of code-words can be made as close to Hr(S) as possible by increasing the extension".

The average length L of the code is computed as $L = \sum_{i=1}^{q} Pi\ li$

Pi refers to probabilities of source symbols and li refers to word lengths in binary digits or binits

</td><td>2 + 2</td><td>6 mins</td></tr>
</table>

Part B (2Q x 8M = 16 Marks)

<table>
<tr><th>Q No</th><th>Solution</th><th>Scheme of Marking</th><th>Max. Time required for each Question</th></tr>
<tr><td>4</td><td>
Given S = {s_1, s_2, s_3} with P= {1/2, 1/4, 1/4}.
$$H(S) = \sum_{i=1}^{3} pi\ log \frac{1}{pi}$$
H(S) = 1.5 bits/message-symbol
The 2nd extension of the basic source with 3 symbols will have $3^2 = 9$ symbols which can be listed along with their probability of occurrence as follows in Table 1:
The entropy of the 2nd extended source is
$$H(S^2) = \sum_{j=1}^{9} pj\ log \frac{1}{pj}$$
= ¼ log 4 + 4(1/8) log 8 + 4(1/16) log 16
=3 bits/m-symbol
= 2 x 1.5 bits/m-symbol
Therefore $H(S^2) = 2\ H(S)\quad (proved)$
Table 1:
<table>
<tr><th>Symbol</th><th>Probability</th></tr>
<tr><td>s_1 s_1</td><td>(1/2)(1/2) =1/4</td></tr>
<tr><td>s_1 s_2</td><td>(1/2)(1/4) =1/8</td></tr>
<tr><td>s_1 s_3</td><td>(1/2)(1/4) =1/8</td></tr>
<tr><td>s_2 s_1</td><td>(1/4)(1/2) =1/8</td></tr>
<tr><td>s_2 s_2</td><td>(1/4)(1/4) =1/16</td></tr>
<tr><td>s_2 s_3</td><td>(1/4)(1/4) =1/16</td></tr>
<tr><td>s_3 s_1</td><td>(1/4)(1/2) =1/8</td></tr>
<tr><td>s_3s_2</td><td>(1/4)(1/4) =1/16</td></tr>
<tr><td>s_3s_3</td><td>(1/4)(1/4) =1/16</td></tr>
</table>
</td><td>2 + 2 + 2 + 2</td><td>10 mins</td></tr>
<tr><td>5</td><td>
<u>Shannon's encoding algorithm</u>
Step 1: The symbols are arranged in order of decreasing probabilities as follows:
<table>
<tr><td>m_5</td><td>m_4</td><td>m_3</td><td>m_1</td><td>m_2</td></tr>
<tr><td>6/16</td><td>4/16</td><td>3/16</td><td>2/16</td><td>1/16</td></tr>
<tr><td>p_1</td><td>p_2</td><td>p_3</td><td>p_4</td><td>p_5</td></tr>
</table>
Step 2: The sequences of α's are computed:-
$\alpha_1 = 0$; $\alpha_2 = 6/16 = 0.375$;$\alpha_3 = 6/16 + 4/16 = 0.625$;
$\alpha_4 = 10/16 + 3/16 = 0.8125$; $\alpha_5 = 13/16 + 2/16 = 0.9375$
and $\alpha_6 = 15/16 + 1/16 = 1$
Step 3: The smallest integer value of li is found using
$2^{li} \geq 1/_{pi}$, for all i =1,2,3,4,5
For i=1, $2^{l1} \geq 16/_6 = 2.66$
Smallest value of l1 is 2 , therefore $l1 = 2$
For i=2, $2^{l2} \geq 4$, smallest value of $l2 = 2$
Similarly we get $l3 = 3, l4 = 3$ and $l5 = 4$
Step 4: The decimal numbers α_i are expanded in binary form upto l_i places: $\alpha_1 = 0$; $\alpha_2 = (0.375)_{10} = (.011)_2$
</td><td>2 + 2 + 2 + 2</td><td>11 mins</td></tr>
</table>

	α_3 =$(0.625)_{10}$ = $(.101)_2$ α_4 =$(0.8125)_{10}$ = $(.1101)_2$ α_5 =$(0.9375)_{10}$ = $(.1111)_2$ **Step 5:** $\alpha_1 = 0$ and $l1 = 2$, code for s1 is 00 $\alpha_2 = (.011)_2$ and $l2 = 2$, code for s2 is 01 $\alpha_3 = (.101)_2$ and $l3 = 3$, code for s3 is 101 $\alpha_4 = (.1101)_2$ and $l4 = 3$, code for s4 is 110 $\alpha_5 = (.1111)_2$ and $l5 = 4$, code for s5 is 1111		

Symbols	p_i	Code	l1 in binits
m_5	3/8	00	2
m_4	1/4	01	2
m_3	3/16	101	3
m_1	1/8	110	3
m_2	1/16	1111	4

Table 1: Code table for the problem

Part C

(1Q x 12M = 12 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
6	(i) From the state diagram, the state equations are P(A) = 0.6 P(A) + 0.5 P(D) -----(1) P(B) = 0.4 P(A) + 0.5 P(D) -----(2) P(C) = 0.6 P(C) + 0.5 P(B) -----(3) P(D) = 0.4 P(C) + 0.5 P(B) -----(4) From eqn.(1) P(A) = 1.25 P(D) -----(5) From eqn.(2) P(B) = P(D) -----(6) From eqn.(3) P(C) = 1.25 P(D) -----(7) But P(A)+ P(B) + P(C) + P(D) =1 -----(8) Solving we get, P(D) = P(B) = 2/9; P(A) = P(C) =5/18; (ii) Entropy of each state is $H_i = \sum_{i=1}^{n} Pij \log 1/pij$ H_A = 0.971 bits/state Similarly H_B = 1 bits/state H_C = 0.971 bits/state and H_D = 1 bits/state (iii) The entropy of the source is $H(S) = \sum_{i=1}^{n} Pi\ Hi$ H(S) = P(A) H_A + P(B) H_B + P(C) H_C + P(D)H_D H(S) = 0.9839 bits/binary digits Adjoint source is H(S') = $\sum_{i=1}^{4} P(mi) \log 1/p(mi)$ H(S') = (2) 5/18 log 18/5 + (2) 2/9 log 9/2 bit/m-symbol Therefore H(S') > H(S)	5+5+2	18 mins

Roll No.													

PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING

TEST –2

Sem & AY: Odd Sem 2019-20

Course Code: ECE 216

Course Name: INFORMATION THEORY AND CODING

Program & Sem: B.Tech (ECE) & VII

Date: 19.11.2019

Time: 9.30 AM to 10.30 AM

Max Marks: 40

Weightage: 20%

Instructions:

(i) *Read the question properly and answer accordingly.*
(ii) *Question paper consists of 3 parts.*
(iii) *Scientific and Non-programmable calculators are permitted.*

Part A [Memory Recall Questions]

Answer all the Questions. Each Question carries four marks. (3Qx4M=12M)

1. Devise an Instantaneous binary code in two ways for a source producing 5 symbols namely S1 to S5. [4] (C.O.2) [Knowledge]

2. Assuming the received sequence (in the absence of noise) at the receiver as 001100, Check whether the codes given in Table-1 are uniquely decodable and instantaneous?
Table-1 [4] (C.O.2) [Analysis]

Source Symbols	**Code-A**	**Code-B**
S1	0	0
S2	10	01
S3	110	011
S4	1110	0111

3. Define Mutual Information? How mutual information is related to joint entropy of the channel? [4] (C.O.2) [Knowledge]

Part B [Thought Provoking Questions]

Answer both the Questions. Each Question carries eight marks. (2Qx8M=16M)

4. Consider a source of 8 alphabets A to H with corresponding probabilities as P= {0.22, 0.20, 0.18, 0.15, 0.10, 0.08, 0.05, 0.02} respectively. Compute the Huffman code for this source, moving the composite symbol "as low as possible". Determine the code efficiency and compute the variance of the word-lengths. List any two merits of Huffman coding over Shannon-Fano coding. [8] (C.O.2) [Analysis]

5. (i) State and prove Kraft Inequality relation. [5] (C.O.2) [Analysis]

(ii) Consider a binary block code with 2^n code words of same length n. Verify that Kraft Inequality is satisfied for such a code. [3] (C.O.2) [Analysis]

Part C [Problem Solving Questions]

Answer the Question. The Question carry twelve marks. (1Qx12M=12M)

6. A discrete memory less source has an alphabet of 7 symbols as described below: Source S= {s_1, s_2, s_3, s_4, s_5, s_6, s_7} with its corresponding probabilities as {0.25, 0.25, 0.125, 0.125, 0.125, 0.0625, 0.0625} respectively. Compute Shannon-Fano code for this source and find the coding efficiency. Draw the code-tree. Justify why the computed source code has an efficiency of 100%? Also illustrate the properties of Joint Probability Matrix in a discrete communication channel? [12] (C.O.2) [Analysis]

SCHOOL OF ENGINEERING

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

Semester: VII

Course Code: ECE 216

Course Name: Information Theory and Coding

Date: 19-11-2019

Time: 9.30 AM – 10.30 AM

Max Marks: 40

Weightage: 20%

Extract of question distribution [outcome wise & level wise]

Q.NO.	C.O.NO	Unit/Module Number/Unit /Module Title	Memory recall type [Marks allotted] Bloom's Levels			Thought provoking type [Marks allotted] Bloom's Levels			Problem Solving type [Marks allotted]			Total Marks
			K			C			A			
1	2	2/2/ Channel Coding	4									4
2	2	2/2/ Channel Coding				4						4
3	2	2/2/ Channel Coding	4									4
4	2	2/2/ Channel Coding							8			8
5	2	2/2/ Channel Coding							8			8
6	2	2/2/ Channel Coding	4						8			12
	Total Marks		12			4			24			40

K =Knowledge Level   C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt

Reviewer Comments

1. In scheme of Masking, The mask is just shown in splitted form. Which mark has to be given for what step is not clearly shown.

4. In Thought provoking question, 4 is like problem
5. Normal Theorem (State and prove).

Thought provoking question is it should provoke the thought, instead of giving normal problem or stat & prove type of Questions.

Annexure- II: Format of Answer Scheme

SCHOOL OF ENGINEERING

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

SOLUTION

Semester: VII
Course Code: ECE 216
Course Name: Information Theory and Coding

Date: 19-11-2019
Time: 9.30 AM – 10.30 AM
Max Marks: 40
Weightage: 20%

Part A (3Q x4 M =12 Marks)

<table>
<tr><th>Q No</th><th>Solution</th><th>Scheme of Marking</th><th>Max. Time required for each Question</th></tr>
<tr><td>1</td><td>Start assigning S1→0. Hence all other code-words should start with ‘1’ according to prefix property.
Let us assign for S2→10.
Let us assign for S3→110. (we cannot assign ‘11’ for S3, as no other combinations left for S4 & S5)
Let us assign for S4→1110 and for S5→1111
Code Table-1 of 2 types of instantaneous codes
<table>
<tr><th>Source Symbols</th><th>Code-A</th><th>Code-B</th></tr>
<tr><td>S1</td><td>0</td><td>00</td></tr>
<tr><td>S2</td><td>10</td><td>01</td></tr>
<tr><td>S3</td><td>110</td><td>10</td></tr>
<tr><td>S4</td><td>1110</td><td>110</td></tr>
<tr><td>S5</td><td>1111</td><td>111</td></tr>
</table></td><td>1
+ 1
+ 2</td><td>6 mins</td></tr>
<tr><td>2</td><td>Given the received sequence as 001100.
If Code-A is used, it is decoded as S1 S1S3 S1.
If Code-B is used, it is decoded as S1 S3 S1 S1.
Both Code-A and Code-B are uniquely decodable.
Code-A can be decoded without referring to the succeeding symbols. So Code-A is an instantaneous code.
When Code-B is used for decoding, at every stage we have to wait for the succeeding symbols to</td><td>2 + 2</td><td>6 mins</td></tr>
</table>

	arrive and hence **Code-B is not an instantaneous code.**		
3	When an average information of H(A) is transmitted over the channel, an average amount of information equal to equivocation H(A/B) is lost in the channel due to inter-symbol conversion (due to noise). The balance of information received at the receiver w.r.to an observed output symbol is the mutual information denoted as I (A, B). I (A, B) = H(A) - H(A/B). The mutual information is related to joint entropy of the channel by I (A, B) = H(A) + H(B) - H(A,B). (Venn diagram) Where H(A,B) is the joint entropy.	2 + 1 + 1	6 mins

Part B (2Q x 8M = 16 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
4	<u>Huffman Coding</u> Given S 'with P= {0.22, 0.20, 0.18, 0.15, 0.10, 0.08, 0.05, 0.02}. $Average\ length\ L = \sum_{i=1}^{8} pi\ \ li$ = (0.22) 2 + (0.20) 2 +(0.18) 3+(0.15) 3+(0.10) 3+(0.08) 4 +(0.05) 5+(0.02) 5 L = 2.8 binits/m-symbol $H(S) = \sum_{i=1}^{3} pi\ log\frac{1}{pi}$ $= 0.22\ log\frac{1}{0.22} + 0.20\ log\frac{1}{0.20} + 0.18\ log\frac{1}{0.18} + 0.15\ log\frac{1}{0.15} + 0.10\ log\frac{1}{0.10} + 0.08\ log\frac{1}{0.08} + 0.05\ log\frac{1}{0.05} + 0.02\ log\frac{1}{0.02}$ H(S) = 2.7535 bits/message-symbol Source efficiency = $\frac{H(S)}{L} = \frac{2.7535}{2.8} = 98.34$ % Variance placing the composite symbol "as low as possible" $\text{Var}\ (li) = E\ [(li - L)^2] = \sum_{i=1}^{8} pi\ \ (li - L)^2$ = (0.22) $(2-2.8)^2$ + (0.20) $(2-2.8)^2$ +(0.18) $(3-2.8)^2$+ (0.15) $(3-2.8)^2$+ (0.10) $(3-2.8)^2$ + (0.08) $(4-2.8)^2$ + (0.05) $(5-2.8)^2$ + (0.02) $(4-2.8)^2$ **Merits of Huffman coding**: 1. Minimum redundancy code compared to shannon-fano codes. 2. Average length is low and small variance	2 + 2 + 2 + 2	10 mins

	3. It is an optimal code but shannon-fano codes are sub-optimal.		
5	**Kraft Inequality Statement**: A necessary and sufficient condition for the existence of an instantaneous code with word lengths $l_1, l_2, \ldots. l_q$ is that $\sum_{i=1}^{q} r^{-li} \leq 1$ Where 'q' is no source symbols, 'r' no of different symbols in code alphabet X and l_i is the word length in binary digits in the code-word corresponding to i^{th} source symbol. **Proof**: The word lengths $l_1, l_2, \ldots. l_q$ are arranged in ascending order so that $l_1 \leq l_2 \leq \ldots. \leq l_q$ Let n_i represent the no. of messages encoded into code-words of length 'i'. For i=1, $n_1 \leq r$ For i=2, for getting an instantaneous code, we must start encoding using ($r- n_1$) symbols only as the 1st digit and the 2nd digit can be any of 'r' symbols of the code alphabet. Therefore for i=2, $n_2 \leq (r - n1)\, r$ (or) $n_2 \leq r^2 - n_1 r$ Similarly for i=3, $n_3 \leq [(r^2 - n_1 r) - n_2]\, r$ Or $n_3 \leq r^3 - n_1 r^2 - n_2 r$. Proceeding this way, we can arrive at $n_i \leq r^i - n_1 r^{(i-1)} - n_2 r^{(i-2)} - \cdots - n_{(i-1)} r$ Multiplying throughout by r^{-i} , we get $n_i\, r^{-i} + n_{(i-1)} r^{-(i-1)} + n_{(i-2)} r^{-(i-2)} + \ldots.+$ $n_1\, r^{-1} \leq 1$ Or $\sum_{m=1}^{i} n_m\ r^{-m} \leq 1$ Since the actual no. of messages n_i *has to be an integer* $$\sum_{m=1}^{i} n_m\ r^{-m} = \sum_{j=1}^{n1} r^{-1} + \sum_{j=1}^{n2} r^{-2} + .. + \sum_{j=1}^{n_i} r^{-i} \leq 1$$ Combining all the groups (since $n_1 + n_2 + \ldots. + n_i = q$), we can write that $\sum_{i=1}^{q} r^{-li} \leq 1$ (proved) (ii) Since all the code words of same length n, we have $l_i = n$ and the no of source symbols is $q = 2^n$ with r=2 for binary codes. $$\sum_{i=1}^{q} r^{-li} = \sum_{i=1}^{q} 2^{-n}$$ $= 2^{-n} + 2^{-n} + \ldots. + 2^{-n}$ (occurs 2^n *terms*) = $[2^{-n}]\, [2^n\}$	2 + 2 + 2 + 2	11 mins

= 1 (Thus Kraft Inequality is satisfied with equality sign).

Code Table-1 instantaneous codes (for n=2)

Source Symbols	Code-A
S1	0
S2	00
S3	01
S4	10
S5	11

Code Table-2 instantaneous codes (for n=3)

Q= 9 symbols

Symbols	Code-A	Symbols	Code-A
S1	000	S6	101
S2	001	S7	110
S3	010	S8	111
S4	011		
S5	100		

Part C (1Q x 12M = 12 Marks)

Q No	Solution	Scheme of Marking	Max. Time require Question
6	Shannon Fano coding	6+2+4	18 mins

Code table-1

Symbols	Code	li	Pi	Verify $log\frac{1}{pi} = li$
S1	11	2	0.25	2 $= l1$
S2	10	2	0.25	2 $= l2$
S3	011	3	0.125	3 $= l3$
S4	010	3	0.125	3 $= l4$

S5	001	3	0.125	$3 = l5$
S6	0001	4	0.0625	$4 = l6$
S7	0000	4	0.0625	$4 = l7$

$Average\ length\ L = \sum_{i=1}^{8} pi\ \ li$

= 2 (0.25) (2) + 3(0.125) (3)+ 2 (0.0625) 4

L = 2.625 binits/m-symbol

$$H(S) = \sum_{i=1}^{3} pi\ log\frac{1}{pi}$$

$= \left[0.25\ log\frac{1}{0.25}\right](2) + \left[0.125\ log\frac{1}{0.125}\right](3) + \left[0.0625\ log\frac{1}{0.0625}\right](2) = 2.625\frac{bits}{m}symbol$

coding efficiency = $\frac{H(S)}{L} = \frac{2.625}{2.625} = 100\ \%$

The reason for the coding efficiency to be 100 % is as follows:

If $li = log\frac{1}{pi}\ \ for\ all\ i = 1,2,..7\ \ then\ efficiency\ becomes\ 100\%$

We observe that from coding table calculation, $li = log\frac{1}{pi}\ \ for\ all\ i = 1,2,..7$, hence the coding efficiency is 100%.

Joint Probability Matrix (JPM) in a discrete communication channel:

"The sum of all the elements in any row of the channel matrix is equal to unity".

JPM Property-1:

"By adding the elements of JPM column wise, we can obtain the probability of output symbols"

JPM Property-2:

"By adding the elements of JPM row wise, we can obtain the probability of input symbols"

JPM Property-3:

"The sum of all the elements of JPM is equal to unity".

Roll No											

PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING

END TERM FINAL EXAMINATION

Semester: Odd Semester: 2019 - 20
Course Code: ECE 216
Course Name: INFORMATION THEORY AND CODING
Program & Sem: B.Tech (ECE) & VII

Date: 27 December 2019
Time: 9:30 AM to 12:30 PM
Max Marks: 80
Weightage: 40%

Instructions:
(i) Read the question properly and answer accordingly. Question paper consists of 3 parts.
(ii) Scientific and Non-programmable calculators are permitted.
(iii) Exchange of calculators is not allowed.

Part A [Memory Recall Questions]

Answer all the sub Questions. Each sub Question carries 2 marks. (10Qx2M=20M)

1.

a. Define entropy. State the maximum value of entropy when all the source symbols become "equi-probable". (C.O.No.1) [Knowledge]

b. State Shannon's first theorem. (C.O.No.1) [Knowledge]

c. When will a non-singular code is said to be "uniquely decodable"? Give one example to illustrate such codes. (C.O.No.2) [Knowledge]

d. When will an Instantaneous code is said to be an "optimal code"? Give one suitable example. (C.O.No.2) [Knowledge]

e. How mutual information is related to joint entropy of the channel? Is mutual information is symmetric? (C.O.No.2) [Knowledge]

f. Justify with reasons why error control coding is important? (C.O.No.3) [Knowledge]

g. A Gaussian channel has a 10 MHz bandwidth. If SNR is 100, calculate the channel capacity and the maximum information rate. (C.O.No.3) [Knowledge]

h. Find the error detecting and error correcting capabilities of (6, 3) code, given the minimum distance d_{min} =3? (C.O.No.4) [Knowledge]

i. When a (n,k) linear block code is said to be cyclic code? Give an example. (C.O.No.4) [Knowledge]

j. How are convolutional codes different from block codes? (C.O.No.4) [Knowledge]

8. Consider a zero-memory source with S= {s_1, s_2, s_3, s_4, s_5, s_6, s_7} whose corresponding probabilities P= {0.4, 0.2, 0.1, 0.1, 0.1, 0.05, 0.05} respectively. Construct a binary Huffman code by (i) placing the composite symbol as low as possible (ii) by moving the composite symbol as high as possible. Compute the variances of the word-lengths in both the cases, find the coding efficiency and also redundancy. Comment on the result.

(C.O.No.2) [Analysis]

9. For a systematic (6,3) linear block code, the parity matrix P is given by

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

(i) Find all possible code-vectors and construct the corresponding encoding circuit. [6M]

(ii) If the received code-vector R = [r_1 r_2 r_3 r_4 r_5 r_6], then construct the corresponding syndrome calculation circuit? [4M] (C.O.No.4) [Analysis]

SCHOOL OF ENGINEERING

END TERM FINAL EXAMINATION

Extract of question distribution [outcome wise & level wise]

Q.NO.	C.O.NO (% age of CO)	Unit/Module Number/Unit /Module Title	Memory recall type [Marks allotted] Bloom's Levels	Thought provoking type [Marks allotted] Bloom's Levels	Problem Solving type [Marks allotted]	Total Marks
			K	C	A	
PART A Q.NO.1	C.O.NO 1	MODULE 1 Fundamentals of Coding Theory and Source Coding	2			2
PART A Q.NO.2	C.O.NO 1	MODULE 1 Fundamentals of Coding Theory and Source Coding	2			2
PART A Q.NO.3	C.O.NO 2	MODULE 2 Channel Coding	2			2
PART A Q.NO.4	C.O.NO 2	MODULE 2 Channel Coding	2			2
PART A Q.NO.5	C.O.NO 2	MODULE 2 Channel Coding	2			2
PART A Q.NO.6	C.O.NO 3	MODULE 3 Error control Coding	2			2
PART A Q.NO.7	C.O.NO 3	MODULE 3 Error control Coding	2			2

Q.NO.17		Channel Coding				
PART C Q.NO.18	C.O.NO 4	MODULE 4 Block codes and convolutional codes			10	10
	Total Marks		20	30	30	80

K =Knowledge Level C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt.

I hereby certify that all the questions are set as per the above guidelines.

Faculty Signature: Sign Needed.

Reviewer Commend: → Part B – Thought Provoking

Format of Answer Scheme

6	Error control coding is nothing but calculated use of redundancy. Error control coding is important because (i) It improves the data quality to a great extent. (ii) Great advantage is the reduction in (Eb/N) for a fixed BER, which in turn reduces the transmitted power and hence the hardware costs.	2 marks for 2 merits	2 mins
7	C= B log (1+S/N) = 66.59×10^6 bits/sec Rmax =C∞=(S/N) B log e maximum information rate = 1.44×10^9 bits/sec	1 mark for sampling rate+ 1 mark for information rate	4 mins
8	minimum distance d_{min} =3 error detecting capability = $d_{min} - 1 = 2$ (double errors) error correcting capability of (6,3) code =$[(d_{min}-1)/2] = 1$ only single errors can be corrected	1 mark for definition + 1 mark for error correction	2 mins
9	A (n,k) linear block code is said to be cyclic code if every cyclic shifts of the code is also a code-vector of C. Let C1=110110 be a code-vector of C. If C2 =011011 (the last bit of C1 has moved into the 1st position) is also a code-vector of C, then it is called a "cyclic code".	1 mark for definition + 1 mark for example	2 mins
10	In "convolutional codes", a block of 'n' code digits generated by encoder depends not only on the 'k' message digits, but also on the preceding (m-1) blocks of message digits. (Presence of memory). In "Block codes", a block of 'n' code digits generated by encoder depends only on the 'k' message digits within that time unit. (Absence of memory). Block codes well suited for error detection and convolutional codes are preferred for error correction.	2 marks for 2 differences	2 mins

Part B (5Q x 6M =30 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
11	<u>Shannon-Fano encoding process is as shown in table (i) below:</u> <u>Table (i)</u>	3 for table + 3 for average length and efficiency	15 mins

S1	**0.5 (1)**	-	-	-	
S2	0.125 (0)	**0.125 (1)**	**0.125 (1)**	-	

13	Kraft Inequality Statement: A necessary and sufficient condition for the existence of an instantaneous code with word lengths $l_1, l_2, \ldots, l_q$ is that $\sum_{i=1}^{q} r^{-li} \le 1$ Where 'q' is no source symbols, 'r' no of different symbols in code alphabet X and l_i is the word length in binary digits in the code-word corresponding to i^{th} source symbol. **Proof**: The word lengths $l_1, l_2, \ldots, l_q$ are arranged in ascending order so that $l_1 \le l_2 \le \ldots \le l_q$ Let n_i represent the no. of messages encoded into code-words of length 'i'. For i=1, $n_1 \le r$ For i=2, for getting an instantaneous code, we must start encoding using (r- n_1) symbols only as the 1st digit and the 2nd digit can be any of 'r' symbols of the code alphabet. Therefore for i=2, $n_2 \le (r - n1)\, r$ (or) $n_2 \le r^2 - n_1 r$ Similarly for i=3, $n_3 \le [(r^2 - n_1 r) - n_2]\, r$ Or $n_3 \le r^3 - n_1 r^2 - n_2 r$. Proceeding this way, we can arrive at $n_i \le r^i - n_1 r^{(i-1)} - n_2 r^{(i-2)} - \cdots - n_{(i-1)} r$ Multiplying throughout by r^{-i} , we get $n_i r^{-i} + n_{(i-1)} r^{-(i-1)} + n_{(i-2)} r^{-(i-2)} + \ldots + n_1 r^{-1} \le 1$ Or $\sum_{m=1}^{i} n_m r^{-m} \le 1$ Since the actual no. of messages n_i *has to be an integer* $\sum_{m=1}^{i} n_m r^{-m} = \sum_{j=1}^{n1} r^{-1} + \sum_{j=1}^{n2} r^{-2} + .. + \sum_{j=1}^{n_i} r^{-i} \le 1$ Combining all the groups (since $n_1 + n_2 + \ldots + n_i = q$), we can write that $\sum_{i=1}^{q} r^{-li} \le 1$ (proved)	2 marks for statement + 4 marks for derivation	15 mins
14	Shannon-Hartley Law states that the capacity of a band-limited Gaussian channel with AWGN is given by C= B log (1+S/N) bits/sec **First Implication**: When B is increased, Channel capacity C also increases, since Rmax =C, the max.frate of information can be enhanced to as large value as possible. C does not become infinite when B is infinite. C∞ is not becoming infinite even though B increases beyond "Shannon's limit" (-1.6 db), it becomes constant **Second Implication** The exchange of bandwidth with SNR and viceversa. A trade-off exists between bandwidth-to-S/N can be illustrated by a trade-off curve. B/C = 1 / log (1+S/N) There is a 25% reduction in bandwidth for an approximate 60% increase in signal power.	2 + 2+ 2	15 mins
15	Generator matrix consists of [L] rows and [n(L+m)] columns C= d G Given d = 1 0 1 1 1 $g^{(1)}$ = 1011 and $g^{(2)}$ = 1111 Generator matrix consists of [L] = 5 rows and [n(L+m)] = 16 columns	2 marks for encoder diagram + 2 marks for G matrix + 2 marks for C matrix	15 mins

<table>
<tr>
<td></td>
<td>
Source efficiency = $\frac{H(S)}{L}$
Variance placing the composite symbol "as low as possible"
Var $(li) = E\ [(li - L)^2] = \sum_{i=1}^{8} pi\ \ (li - L)^2$=2.25

Code Table arrived after Huffman Coding process by (i) placing the composite symbol as high as possible is, the code table becomes

<table>
<tr><th>Symbols</th><th>Probabilities Pi</th><th>Code Table</th></tr>
<tr><td>S1</td><td>0.4</td><td>00</td></tr>
<tr><td>S2</td><td>0.2</td><td>11</td></tr>
<tr><td>S3</td><td>0.1</td><td>011</td></tr>
<tr><td>S4</td><td>0.1</td><td>100</td></tr>
<tr><td>S5</td><td>0.1</td><td>101</td></tr>
<tr><td>S6</td><td>0.05</td><td>0100</td></tr>
<tr><td>S7</td><td>0.05</td><td>0101</td></tr>
</table>

Variance placing the composite symbol "as high as possible"
Var $(li) = E\ [(li - L)^2] = \sum_{i=1}^{8} pi\ \ (li - L)^2$=0.45

COMMENT: By placing the composite symbol "as high as possible", the variance of word lengths becomes smaller, which is desirable.
</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>
Code vector of (6.3) linear block code
Check bits is n-k = 3.

Message Vectors = 8 namely (000), (001), (010),(011),(100),(101),(110) and (111).
[C] = [D] {G}
[G] = [I_3 P] since k=3
[C] = [D] [G] = [d_1,d_2,d_3, (d_1+d_3),(d_2+d_3),(d_1+d_2)]
Draw the code vector table corresponding to 8 message vectors.
i) No. of Modulo-2 adders used =3
A 3-bit shift register and a 6 bit commutator segment are needed to realize LBC encoder circuit
ii) Find Transpose of Parity check matrix
Parity check matrix H =[P^T I_3]
Now take Transpose of H and multiply with R
[S] =[s_1 s_2 s_3] =R H^T
S= [s_1 s_2 s_3] = [r_1 r_2 r_3 r_4 r_5 r_6] H^T
= [(r_1+r_3+ r_4), (r_2+r_3+ r_5),(r_1+r_2+ r_6)]
Syndrome bits are s_1 = r_1+r_3+ r_4
s_2 = r_2+r_3+ r_5 and s_3 = r_1+r_2+ r_6
syndrome calculation circuit contains:
6 bits of received vector in shift register
3 no. of modulo-2 adders
Syndrome bits s_1 s_2 and s_3 as output from adders.
</td>
<td>3 marks for code vector table+3 marks for circuit + 4 marks for syndrome error</td>
<td>18 mins</td>
</tr>
</table>