Zewail City of Science and Technology
Nanotechnology Engineering Program
Midterm #1 - Spring 2020

NANENG 461
Communication Theory and Systems
Time: 60 Minutes

Name: Model Answer ID.: ....

This examination paper has 4 pages and 5 questions

Question 1

[1.25 Points]

Choose the suitable answer then write it in the table

Question 1	1	2	3	4	5
Your Answer	b	a	a	c	a

1. Code rate r, k information bits and n as total bits, is defined as:
   (a) $k = n/r$
   (b) $r = k/n$
   (c) $n = r * k$
2. Which of the following statements is not true for arithmetic coding
   (a) Integral number of bits is assigned to each symbol.
   (b) A real number in the interval [0, 1) indicates the entire coding sequence.
   (c) Coding requires a priori knowledge of the probabilities of source symbols.
   (d) Longer sequence of source symbols leads to longer codewords.
3. Which of the following statements is not true for Lempel-Ziv coding
   (a) It is a fixed-length coding for variable-length symbol sequence.
   (b) It does not require a priori knowledge of the source symbol probabilities.
   (c) Larger image/file size leads to poorer compression.
   (d) Decoder dictionary can be derived from the encoded sequence.
4. The coding efficiency of arithmetic coding is expected to improve if
   (a) The interval of real numbers is increased to [0, 10).
   (b) Shorter sequence of source symbols is encoded.
   (c) Longer sequence of source symbols is encoded.
   (d) None of the above.
5. Which of the following sets of codewords could be the Huffman code for some 4 symbol source alphabet?
   (a) 0, 10, 110, 111
   (b) 1, 01, 10, 001
   (c) 0, 110, 111, 101

Question 2

[5 Points]

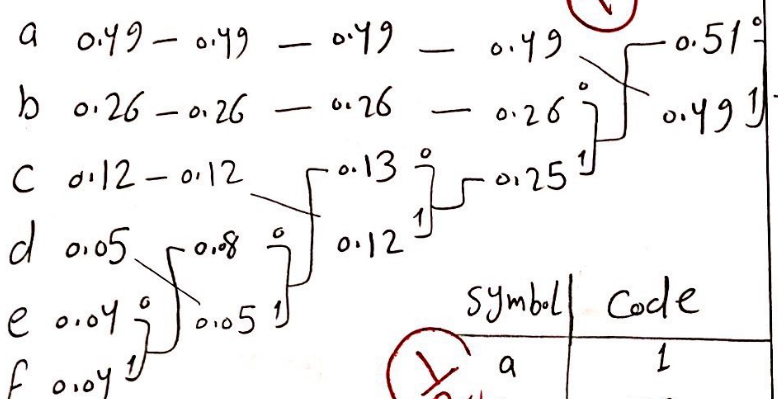
For a source $S = \{a, b, c, d, e, f\}$ with probability distribution $P = \{0.49, 0.26, 0.12, 0.04, 0.04, 0.05\}$, respectively.

- Find a binary Huffman code and compare the average code length with the source entropy.
- Find a binary Shannon-Fano code and compare the average code length with the Huffman one. Comment on your result.
- If the required coding length of each symbol is 3 bits/symbol then for a rate $R_s = 9.6$ kbaud (baud=symbol/second), what is the data rate of the signal after Shannon-Fano coding? What compression factor has been achieved?
- Find ternary Huffman code $\{r = 3$ with code alphabet $(0, 1, 2)\}$. Compare the average code length with the entropy of the source.

Solution

* Huffman:

1



1/2

Symbol	Code
a	1
b	00
c	011
d	01000
e	01001
f	0101

$$\bar{L} = \sum_{K=0}^{6} P_K L_K$$

$$= 1.97$$

$$H(S) = \sum_{K=0}^{6} P_K \log_2 \frac{1}{P_K} = 1.964$$

$$\therefore \bar{L} > H(S).$$

* Shannon-Fano:

1

Symbol	Prob.	Step 1	Step 2	Step 3	Step 4	Step 5
a	0.49	0				
b	0.26	1	0			
c	0.12	1	1	0		
f	0.05	1	1	1	0	
d	0.04	1	1	1	1	0
e	0.04	1	1	1	1	1

$\bar{L} = 1.97$

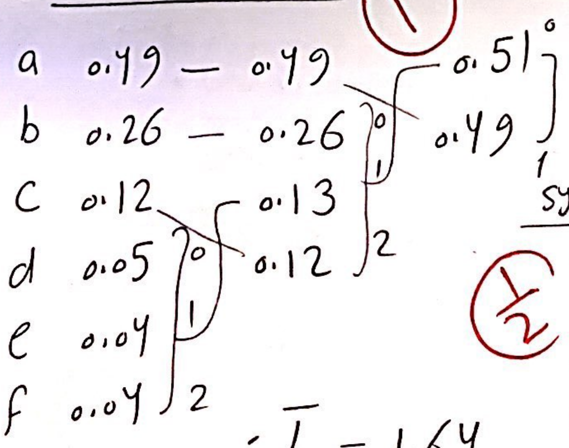
1/2

Same as Huffman

Symbol	Cod
a	0
b	10
c	110
d	1111
e	1111
f	1110

* Ternary Huffman:

1



1/2

Symbol	Code
a	1
b	00
c	02
d	011
e	012
f	010

$$\therefore \bar{L} = 1.64$$

$$\therefore \bar{L} < H(S)$$

* Data Rate:

1/2

data rate $= \bar{L} * R_s = 18912$ bit

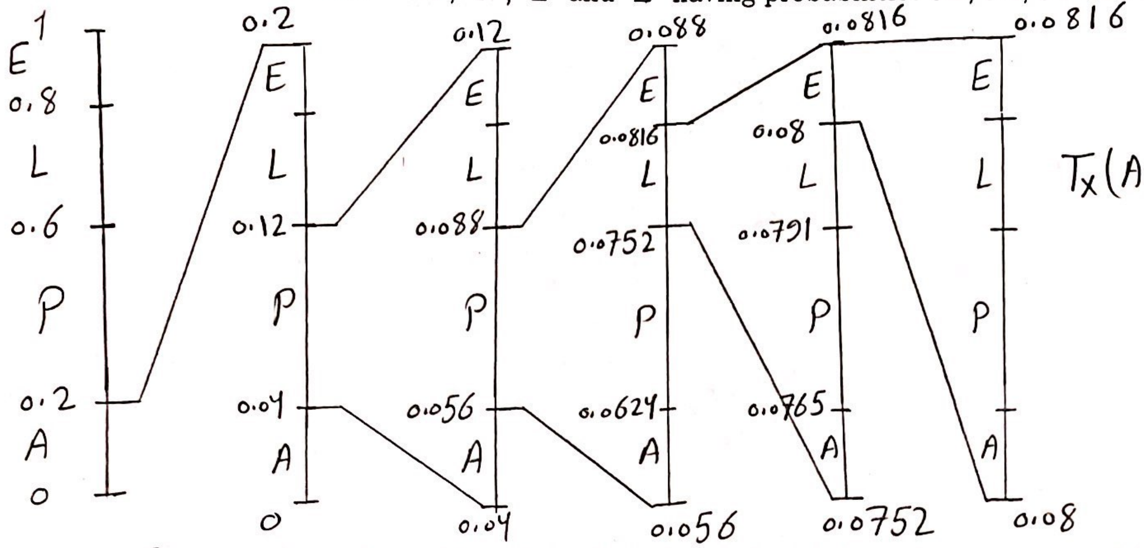
Comp. factor $= \dfrac{3 \text{ bit}}{\bar{L}} = 1.52$

#

Question 3

[4 Points]

- Construct an arithmetic code in real decimal number for the word "APPLE" formed out of a 4-symbol alphabet – "A", "P", "L" and "E" having probabilities 0.2, 0.4, 0.2 and 0.2 respectively.



$$T_X(APPLE) = \frac{0.0816+0.08}{2}$$

$$= 0.0808$$

- Compare the codeword length in digits/symbol with its entropy expressed in the same units. Calculate its coding efficiency.

$$H(S) = \sum P_K \log_2 \frac{1}{P_K} = 1.793$$

$$\bar{L} = -\log_2(0.0816) = 3.6 \simeq 4 \text{ bits} \quad \therefore \eta = 44.82\%$$

Question 4

[2 Points]

- Give an example of a prefix code on $A_1; A_2; \ldots; A_6$ with codeword lengths 1, 3, 3, 3, 4, 4 respectively.

$C(A_1) = 0$ $\qquad C(A_3) = 101$ $\qquad C(A_5) = 1110$

$C(A_2) = 100$ $\qquad C(A_4) = 110$ $\qquad C(A_6) = 1111$

- For the same code, what probabilities $p(A_i)$ would make the average codeword length equal to the entropy? Verify your answer.

For Prefix Code $P(A_i) \simeq 2^{-\lambda_i}$

$P(A_1) = 2^{-1}$ $\qquad P(A_3) = 2^{-3}$ $\qquad P(A_5) = 2^{-4}$

$P(A_2) = 2^{-3}$ $\qquad P(A_4) = 2^{-3}$ $\qquad P(A_6) = 2^{-4}$

- Give an example of probabilities $p(A_i)$ for which your code is an optimal prefix code, but the average codeword length is greater than the entropy, and the probabilities of each symbol are different from the probabilities you calculated before. Verify your answer.

To make $\bar{L} > H(S)$ $\therefore$ Probabilities must be changed (increa

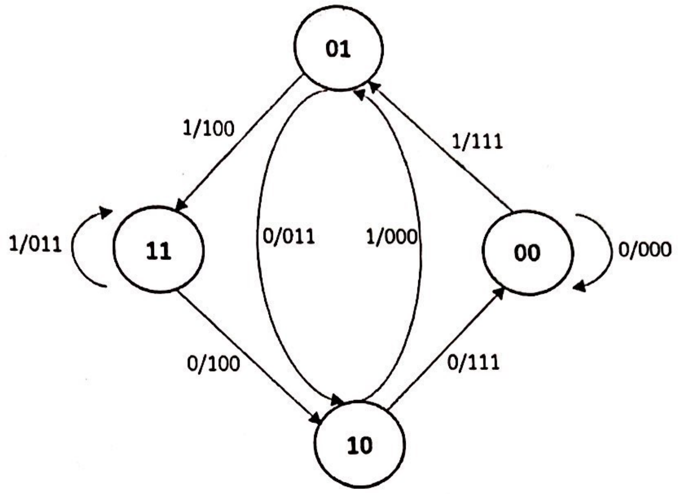
$P(A_1) = 2^{-1} + \epsilon$ $\qquad P(A_3) = 2^{-3} + \epsilon$ $\qquad P(A_5) = 2^{-4} + \epsilon$

$P(A_2) = 2^{-3} + \epsilon$ $\qquad P(A_4) = 2^{-3} + \epsilon$ $\qquad P(A_6) = 2^{-4} + \epsilon$

Question 5

[4 Points]

Assume the encoder specified by the following "State Diagram"



1. What is the code rate?
2. What is the constraint length?
3. Using the Hard Viterbi decoding algorithm, decode this sequence [1 1 1 0 0 1 1 1 1].

Solution

① Code rate $= \frac{1}{3}$

② Constraint length = 3

③ seq. = 1 1 1 0 0 1 1 1 1 $\longrightarrow$ 1 1 1