Information and Coding Theory

ENEE 5304

Problem Set 2

1. A source generates one of four symbols randomly every one time unit. The probabilities of the symbols are 0.4, 0.3, 0.2, 0.1. Each emitted symbol is independent of the other symbols in the sequence
   a. Construct a Huffman code for the source
   b. Find the average length of the code in bits/codeword
   c. What is the minimum achievable average code length?

2. Consider a source S with the given alphabet and the associated probabilities

Symbol	A	B	C	D	E	F	G	H
Probability	0.4	0.18	0.1	0.1	0.07	0.06	0.05	0.04
codeword	1	001	011	0000	0100	0101	00010	00011

   a. Find a Huffman code and compare it to that given in the table
   b. Find the average length of the code in bits/codeword
   c. How far is the average code length from the Entropy?

3. Perform a Lempel-Ziv compression of the string given below
   000111000011100110001010011100101110111011110001101101110111000010
   001100011

4. What is the shortest possible code length, in bits per average symbol, that could be achieved for a six-letter alphabet whose symbols have the following probability distribution?

   {1/2, ¼, 1/8, 1/16, 1/32, 1/32}

5. Construct an efficient, uniquely decodable binary code, having the prefix-free property and having the shortest possible average code length per symbol, for an alphabet whose five letters appear with these probabilities:

Letter	*Probability*
A	1/2
B	1/4
C	1/8
D	1/16
E	1/16