KFUPM-EE DEPT.
EE430: Information Theory and Coding
Problem Set # 2 Dr. Ali Muqaibel

1. Consider a sequence of letters of the English alphabet with their probabilities of occurrence as given here:

Letter	*a*	*i*	*l*	*m*	*n*	*o*	*p*	*y*
Probability	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1

Compute two different Huffman codes for this alphabet. In one case, move a combined symbol in the coding procedure as high as possible, and in the second case, move it as low as possible. Hence, for each of the two codes, find the average code-word length and the variance of the average code-word length over the ensemble of letters.

2. A discrete memoryless source has an alphabet of seven symbols whose probabilities of occurrence are as described here:

Symbol	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Probability	0.25	0.25	0.125	0.125	0.125	0.0625	0.0625

Compute the Huffman code for this source, moving a "combined" symbol as high as possible. Explain why the computed source code has an efficiency of 100 percent.

3. Consider a discrete memoryless source with alphabet $\{s_0, s_1, s_2\}$ and statistics {0.7, 0.15, 0.15} for its output.

(a) Apply the Huffman algorithm to this source. Hence, show that the average code-word length of the Huffman code equals 1.3 bits/symbol.
(b) Let the source be extended to order two. Apply the Huffman algorithm to the resulting extended source, and show that the average code-word length of the new code equals 1.1975 bits/symbol.
(c) Compare the average code-word length calculated in part (b) with the entropy of the original source.

4. Consider the following binary sequence

11101001100010110100...

Use the Lempel-Ziv algorithm to encode this sequence. Assume that the binary symbols 0 and 1 are already in the codebook.

5. A discrete memoryless source with *A* = { *a, b,* c } emits the following string.

bccacbccccccccccccaccca

Using the Lempel-Ziv algorithm, encode this sequence and find the code dictionary and the transmitted sequence.

6. A source with *A* = {*a, b,* c} is encoded using the Lempel-Ziv algorithm. The transmitted code word sequence is 2,3,3,1,3,4,5,10,11,6,10.
Construct the dictionary and decode this sequence.

Note: *not all answers will be posted.* *Questions are selected from different text books (Wells, Haykin).*

Practice is something we do before and not during exams! **Best regards, Dr. Ali Muqaibel**