TCOM 500: Modern Telecommunications
Prof. B.-P. Paris
Homework 4
Solution

Problems 1. **Huffman Coding:** Note, multiple solutions for the Huffman code are possible - they all have the same average code word length.

Symbol X	Probability $P(X)$	$-\log_2(P(X))$	Huffman Code	Code Length
1	0.3	1.737	10	2
2	0.2	2.322	00	2
3	0.2	2.322	01	2
4	0.1	3.322	110	3
5	0.1	3.322	1110	4
6	0.1	3.322	1111	4
		Enropy: 2.44		**Average:** 2.5

2. **Huffman Coding with Multiple Symbols:**

(a) Entropy is 0.97.

(b) Only possible code is to assign one bit to each possibility. Code length is one bit per symbol.

(c) Pertinent probabilities are $P(HH) = 0.36$, $P(HT) = P(TH) = 0.24$, and $P(TT) = 0.16$. Huffman code for these probabilities assigns code words of length 2 to each pair. Therefore, there is no benefit from considering pairs over individual symbols.

(d) Entropy of the source is now only 0.47. For single symbols, the only code uses one bit per symbol for an average code length of 1.
For pairs of symbols, Pertinent probabilities are $P(HH) = 0.81$, $P(HT) = P(TH) = 0.09$, and $P(TT) = 0.01$. A Huffman code for pairs of symbols is: HH=1, HT=10, TH=110, TT=111. This code has an average code length of 1.29 bits or 0.65 bits per symbol.

(e) When the difference between symbols is large, the entropy is small. Therefore, there is much more room for improvement from Huffman coding.

3. **Lempel-Ziv Coding:** The final dictionary contents will be

Index	**Phrase**
1	I
2	M
3	P
4	S
5	MI
6	IS
7	SS
8	SI
9	ISS
10	SIP
11	PP
12	PI

The transmitted sequence of indices is

Message	M	I	S	S	IS	SI	P	P	I
Transmitted	2	1	4	4	6	8	3	3	1

4. **Lempel-Ziv Decoding:** The decoded message is

Received	2	1	3	5	1
Message	B	A	N	AN	A

The following dictionary is built up while symbols are received

Index	**Phrase**
1	A
2	B
3	N
4	BA
5	AN
6	NA
7	ANA