EL606: Homework 3 – Solutions

Due date: Oct. 19, 2004

1. *Problem 4, Ch. 5, Cover&Thomas*

   (a) The Huffman tree for this distribution is

Codeword								
1	x_1	0.49	0.49	0.49	0.49	0.49	0.51	1
00	x_2	0.26	0.26	0.26	0.26	0.26	0.49	
011	x_3	0.12	0.12	0.12	0.13	0.25		
01000	x_4	0.04	0.05	0.08	0.12			
01001	x_5	0.04	0.04	0.05				
01010	x_6	0.03	0.04					
01011	x_7	0.02						

   (b) The expected length of the codewords for the binary Huffman code is 2.02 bits. ($H(X) =$ 2.01 bits)

   (c) The ternary Huffman tree is

Codeword					
0	x_1	0.49	0.49	0.49	1.0
1	x_2	0.26	0.26	0.26	
20	x_3	0.12	0.12	0.25	
22	x_4	0.04	0.09		
210	x_5	0.04	0.04		
211	x_6	0.03			
212	x_7	0.02			

   This code has an expected length 1.34 ternary symbols. ($H_3(X) = 1.27$ ternary symbols).

2. *Problem 6, Ch. 5, Cover&Thomas*

   (a) {0,10,11} is a Huffman code for the distribution (1/2,1/4,1/4).

   (b) The code {00,01,10, 110} can be shortened to {00,01,10, 11} without losing its instantaneous property, and therefore is not optimal, so it cannot be a Huffman code. Alternatively, it is not a Huffman code because there is a unique longest codeword.

   (c) The code {01,10} can be shortened to {0,1} without losing its instantaneous property, and therefore is not optimal and not a Huffman code.

3. *Problem 8, Ch. 5, Cover&Thomas*
   It is easy to design an optimal code for each state. A possible solution is

Next state	S_1	S_2	S_3	
Code C_1	0	10	11	$E(L\|C_1) = 1.5$ bits/symbol
code C_2	10	0	11	$E(L\|C_2) = 1.5$ bits/symbol
code C_3	-	0	1	$E(L\|C_3) = 1$ bit/symbol

   The average message lengths of the next symbol conditioned on the previous state being S_i are just the expected lengths of the codes C_i. Note that this code assignment achieves the conditional entropy lower bound.

To find the unconditional average, we have to find the stationary distribution on the states. Let μ be the stationary distribution. Then

$$\mu = \mu \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

We can solve this to find that $\mu = (2/9, 4/9, 1/3)$. Thus the unconditional average number of bits per source symbol

$$\begin{aligned} EL =& \sum_{i=1}^{3} \mu_i E(L|C_i) \\ =& \frac{2}{9} \times 1.5 + \frac{4}{9} \times 1.5 + \frac{1}{3} \times 1 \\ =& \frac{4}{3} \text{ bits/symbol.} \end{aligned}$$

The entropy rate $\mathcal{H}$ of the Markov chain is

$$\begin{aligned} \mathcal{H} =& H(X_2|X_1) \\ =& \sum_{i=1}^{3} \mu_i H(X_2|X_1 = S_i) \\ =& 4/3 \text{ bits/symbol.} \end{aligned}$$

Thus the unconditional average number of bits per source symbol and the entropy rate $\mathcal{H}$ of the Markov chain are equal, because the expected length of each code C_i equals the entropy of the state after state i, $H(X_2|X_1 = S_i)$, and thus maximal compression is obtained.

4. *Problem 9, Ch. 5, Cover&Thomas*
There is a trivial example that requires almost 1 bit above its entropy. Let X be a binary random variable with probability of $X = 1$ close to 1. Then entropy of X is close to 0, but the length of its optimal code is 1 bit, which is almost 1 bit above its entropy.

5. *Problem 25, Ch. 5, Cover&Thomas*

   (a) Since $l_i = \lceil \log \frac{1}{p_i} \rceil$, we have

   $$\log \frac{1}{p_i} \leq l_i < \log \frac{1}{p_i} + 1$$

   which implies that

   $$H(X) \leq L = \sum p_i l_i < H(X) + 1.$$

   The difficult part is to prove that the code is a prefix code. By the choice of l_i, we have

   $$2^{-l_i} \leq p_i < 2^{-(l_i - 1)}.$$

   Thus F_j, $j > i$ differs from F_i by at least 2^{-l_i}, and will therefore differ from F_i is at least one place in the first l_i bits of the binary expansion of F_i. Thus the codeword for F_j, $j > i$, which has length $l_j \geq l_i$, differs from the codeword for F_i at least once in the first l_i places. Thus no codeword is a prefix of any other codeword.

(b) We build the following table

Symbol	Probability	F_i in decimal	F_i in binary	l_i	Codeword
1	0.5	0.0	0.0	1	0
2	0.25	0.5	0.10	2	10
3	0.125	0.75	0.110	3	110
4	0.125	0.875	0.111	3	111

The Shannon code in this case achieves the entropy bound (1.75 bits) and is optimal.