Source Coding Homework #3 (H.-M. Hang, Oct. 2011)

Note that the full score of this homework assignment is 200 points (not 100 points).

[40%] (1) Huffman code: For an alphabet $A=\{a_1, a_2, a_3, a_4\}$ with probabilities $P(a_1)=0.1$, $P(a_2)=0.3$, $P(a_3)=0.25$, $P(a_4)=0.35$, find the following Huffman codes. ((Sayood, pbm5, p.78))

[15%] (1.a) The Huffman Code A designed using the classnotes procedure (the first procedure in the textbook).

[15%] (1.b) The Huffman Code B designed using the minimum variance procedure in the textbook. (In the minimum variance procedure, at each step, the combined symbol (letter) is put as high in the list as possible.)

[10%] (1.c) Calculate the average bits of using (1.a) and (1.b) codes applied to this source. Compute the variances of the length of codewords in each case. Comment on the differences.

[40%] (2) Error in Huffman Coding: We will explore the error effect on two equivalent Huffman codes. ((Sayood, pbm10, p.79))

Huffman Code A

Symbol	Prob.	Codeword
a_2	0.4	1
a_1	0.2	01
a_3	0.2	000
a_4	0.1	0010
a_5	0.1	0011

Huffman Code B

Symbol	Prob.	Codeword
a_1	0.2	10
a_2	0.4	00
a_3	0.2	11
a_4	0.1	010
a_5	0.1	011

[10%] (2.a) Use Code A to encode the following sequence

$$a_2\, a_1\, a_3\, a_2\, a_1\, a_2$$

Suppose there was an error in the channel and the first bit was received incorrectly ("1" → "0"; "0" → "1"). Decode the received binary sequence. How many characters are decoded in error before the first correctly decoded character?

[10%] (2.b) Repeat (2.a) using Code B.

[20%] (2.c) Repeat (2.a) and (2.b) with the error in the third bit.

[30%] (3) Arithmetic code: Given a number a in the interval [0,1) with an n-bit binary representation $[b_1\, b_2 \ldots b_n]$, show that for any other number b to have a binary representation with $[b_1\, b_2 \ldots b_n]$ as the prefix, b must lie in the interval $\left[a, a+\frac{1}{2^n}\right)$.

((Sayood, pbm1, p.114))

[40%] (4) Arithmetic code: For the probability model below, decode a sequence of length 6 with the tag 0.63215699. ((Sayood, pbm6, p.115))

Symbol	Probability
a_1	0.2
a_2	0.3
a_3	0.5

[50%] (5) LZW code: A sequence is encoded using the LZW algorithm and the initial dictionary is given below.

Index	Codeword (Codebook entry)
1	*a*
2	*b*
3	*h*
4	*i*
5	*s*
6	*t*

The output of the LZW encoder is as follows. Decode this sequence. ((Sayood, pbm4, p.139))

6	3	4	5	2	3	1	6	2	9	11	16	12	14

DUE (in class): November 3, 2011