CSC 310 — First Mid-term Test — 2002-02-15

For all questions, show enough of your work to indicate how you obtained your answer. No books or notes allowed. You may use a calculator.
This test is 50 minutes in length.

1. [14 marks total]

   a) [10 marks] For each of the following codes, which are given as a set of codewords for a binary code alphabet, give an example showing that the code is not uniquely decodable. Also show whether or not the code satisfies the Kraft-McMillan inequality.

   $\{00, 001, 100, 1\}$

   $\{10, 11, 0, 01\}$

   b) [4 marks] For each of the following codes, given as a set of codewords for a binary code alphabet, give an example showing that the code is not instantaneously decodable.

   $\{101, 111, 10, 00\}$

   $\{10, 11, 110, 0\}$

2. [26 marks total] These questions concern a binary Huffman code constructed for symbols $s_1, \ldots, s_q$, with probabilities $p_1, \ldots, p_q$.

   a) [15 marks] Suppose that $p_1 > 1/2$. Must the codeword for s_1 be only one bit long? Explain why this must be true, or give an example showing that it is not always true.

   b) [11 marks] Suppose instead that p_1, p_2, and p_3 are all greater than $1/4$. Say what you can about the lengths of the Huffman codewords for s_1, s_2, and s_3. Explain the reasoning behind your statement.

3. [20 marks total] These questions concern a source with symbols s_1, s_2, s_3, with probabilities $p_1 = 1/2$, $p_2 = 1/4$, $p_3 = 1/4$.

   a) [5 marks] Calculate the binary entropy of this source. Show your work.

   b) [4 marks] Calculate the average code length of the code $s_1 \to 00$, $s_2 \to 01$, $s_3 \to 1$. Show your work.

   b) [4 marks] Calculate the average code length of the code $s_1 \to 0$, $s_2 \to 10$, $s_3 \to 11$. Show your work.

   c) [7 marks] Suppose we encoded symbols from this source in blocks of size ten. (In other words, suppose we encode the 10th extension of the source.) What will the average code length be for a binary Huffman code for these blocks of ten symbols? Explain.

4. [20 marks] Find a binary Huffman code for the source alphabet $\{s_1, s_2, s_3, s_4, s_5\}$, with symbol probabilities $p_1 = 0.2$, $p_2 = 0.08$, $p_3 = 0.07$, $p_4 = 0.32$, and $p_5 = 0.33$. Show this code as a table giving the codeword for every source symbol. Show your work.

5. [20 marks] These questions concern arithmetic coding for the source alphabet $\{s_1, s_2, s_3, s_4\}$, with symbol probabilities $p_1 = 1/2$, $p_2 = 1/12$, $p_3 = 1/3$, and $p_4 = 1/12$. Assume that arithmetic coding is done using cumulative symbol probabilities with the symbols ordered as they are numbered here.

   a) [4 marks] If the first symbol encoded is s_3, what will be the coding interval, $[l, u)$, after this symbol has been seen, but before any rescaling of the interval is done? Show your work.

   b) [4 marks] Will the encoder be able to transmit the first encoded bit after s_3 is received, but before the next source symbol is received? Explain.

   c) [4 marks] If the second symbol encoded is s_1, what will be the coding interval, $[l, u)$, after this symbol is encoded, assuming that *no rescaling* was done after the first symbol was encoded. Show your work.

   b) [4 marks] If the second symbol is the last symbol of the message, what further bits will the encoder have to transmit to ensure that the decoder will correctly identify the two symbols that were transmitted?