GEORGIA INSTITUTE OF TECHNOLOGY
School of Electrical and Computer Engineering

ECE 6605
Information Theory

Assigned: Thursday, Oct 4, 2018
Due: Thursday, Oct. 18, 2018
Due: Sections Q (Video Student): Thursday,, Oct 25, 2018

Problem Set #4

This HWK has 8 problems.

Problem 1-2: Solve the questions 5.24, and 5.45 from Chapter 5 of the textbook (the second edition).

Problem 3: Let $\ell_1, \ell_2, \ldots, \ell_{15}$ be the binary huffman codeword lengths for the probabilities $p_1 \geq p_2 \ \ldots \ \geq p_{14} \geq p_{15}$. Suppose we get a new distribution by splitting the last probability mass, p_{15}. What can you say about the optimal binary codeword lengths $\overline{\ell_1}, \overline{\ell_2}, \ldots, \overline{\ell_{15}}, \overline{\ell_{16}}$ for the probabilities $p_1, p_2, \ldots, p_{14}, \alpha p_{15}, (1-\alpha)p_{15}$, where $0 \leq \alpha \leq 1$.

Problem 4: The problem deals with binary prefix codes.

(a) Show that the two distinct binary prefix codes: namely $(00, 01, 10, 11)$ and $(0, 10, 110, 111)$ are both the optimal prefix codes for the source with four symbols having probability distribution $[p_1 = 1/3, p_2 = 1/3, p_3 = 1/6, p_4 = 1/6]$.

(b) Characterize the set of all probability vectors $[p_1, p_2, p_3, p_4]$ with $p_1 \geq p_2 \geq p_3 \geq p_4$ for which these same two codes, in part (a), are still optimal.

(c) This part is a bonus problem. Show that the set of probability vectors you found in part (b) is the convex hull of three fixed probability vectors; and find these three vectors. Clarification: if $\underline{x_1}, \ldots, \underline{x_n}$ are n vectors, their convex hull is the set of all vectors $\underline{x}$ of the form:

$$\underline{x} = a_1\underline{x_1} + a_2\underline{x_2} + \cdots + a_n\underline{x_n}$$

where $(a_1, \ldots, a_n)$ satisfy: $a_1 + a_2 + \cdots + a_n = 1$, $a_i \geq 0$ for all i.

Problem 5: Solve the question 5.19 from Chapter 5 of the textbook (the second edition).

Problem 6: Suppose that $X = \{1, 2, 3, 4\}$ and that the probabilities of the four possible outcomes are $p = \{\frac{1}{2}, \frac{1}{8}, \frac{1}{4}, \frac{1}{8}\}$

(a) Determine $H(X)$.

(b) Let $q = \{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{1}{8}\}$ be probabilities associated with a random variable Y also defined on the set $\{1, 2, 3, 4\}$. Compute $H(Y)$.

(c) Find the relative entropy between p and q, (i.e., $D(p\|q)$. Also find $D(q\|p)$.

(d) Find a Huffman code for X.

(e) Find the expected codeword length for the Huffman code.

(f) Now suppose that q had been the true distribution, but the Huffman code was designed using p as in part d. Find the expected codeword length. What is the cost for not using the true distribution q to design the code?

Problem 7: Suppose that X is a ternary random variable, taking values from $\{0, 1, 2\}$ with a probability distribution as $p(0) = 1/16$, $p(1) = 1/32$, and $p(2) = 1 - 1/16 - 1/32$.

(a) Give a Huffman code for the random variable X. (find the codewords and show your derivation of the code)

(b) Give a Shannon code for the random variable X. (find the codewords and show your derivation of the code)

(c) To represent sequence of length $n = 50$, i.i.d random variables $X_1, X_2, \ldots, X_{50}$, what compression scheme (Huffman, Shannon, or something else) would you use that is practically feasible? State the name of the scheme and justify your choice (non-quantitative explanation suffices). Give good upper and lower bounds on the expected length of your coding scheme?

Problem 8: The problem deals with optimum compression of a Markov source. Consider a 3-state Markov process $X_1, X_2, \ldots$ having transition matrix:

$$[p_{ij}] = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}$$

where $p_{ij} = Prob\{X_{n+1} = j \mid X_n = i\}$. Thus the probability that state 1 follows state 2 is equal to zero. Design three codes $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ (one for each state, 1, 2, and 3). Each code would map elements of the set of states into sequences of 0's and 1's such that this Marlov process can be sent with maximum compression by the following scheme:

(a) Note the present symbol $X_n = i$ where $i \in \{1, 2, 3\}$.

(b) Select code $\mathcal{C}_i$.

(c) Note the next symbol $X_{n+1} = j$ and send the codeword in $\mathcal{C}_i$ corresponding to j.

(d) Repeat the above procedure for the next symbol.

Answer the followings:

(Q1) What is the average message length of the next symbol conditioned on the previous state $X_n = i$? (Hint: Note that every code $\mathcal{C}_i$ has three codewords except the code $\mathcal{C}_2$ which has only two codewords. First find those codewords using Huffman coding. Note also that, due to specific scheme we use above, it is fine if a codeword in the code $\mathcal{C}_i$ also appear in the code $\mathcal{C}_j$).

(Q2) What is the unconditional average number of bits per source symbol? Show that this is equal to the entropy rate of the Markov chain.