EE6340: Information Theory
Problem Set 5

1. *Uniquely decodable and instantaneous codes.* Let $L = \sum_{i=1}^{m} p_i l_i^{100}$ be the expected value of the 100th power of the word lengths associated with an encoding of the random variable. Let $L_1 = \min L$ over all instantaneous codes; and let $L_2 = \min L$ over all uniquely decodable codes. What inequality relationship exists between L_1 and L_2?

2. *Slackness in the Kraft inequality.* An instantaneous code has word lengths $l_1, l_2, ..., l_m$ which satisfy the strict inequality

$$\sum_{i=1}^{m} D^{-l_i} < 1. \tag{1}$$

The code alphabet is $\mathcal{D} = \{0, 1, 2, ..., D-1\}$. Show that there exist sequences of code symbols in $\mathcal{D}^*$ which cannot be decoded into sequences of codewords.

3. *Simple optimum compression of a Markov source.* Consider the 3-state Markov process having transition matrix

$U_{n-1} \backslash U_n$	S_1	S_2	S_3
S_1	1/2	1/4	1/4
S_1	1/4	1/2	1/4
S_3	0	1/2	1/2

Thus the probability that S_1 follows S_3 is equal to zero. Design 3 codes C_1, C_2, C_3 (one for each state S_1, S_2, S_3), each code mapping elements of the set of S_i's into sequences of 0's and 1's, such that this Markov process can be sent with maximal compression by the following scheme:

   (a) Note the present symbol S_i.

   (b) Select code C_i.

   (c) Note the next symbol S_j and the send the codeword in C_i corresponding to S_j.

   (d) Repeat for the next symbol.
   What is the average message length of the next symbol conditioned on the previous state $S = S_i$ using this coding scheme? What is the unconditional average number of bits per source symbol? Relate this to the entropy rate $\mathcal{H}$ of the Markov chain.

4. *Huffman code.* Find the (a)*binary* and (b)*ternary* Huffman codes for the random variable X with probabilities

$$p = \left(\frac{1}{21}, \frac{2}{21}, \frac{3}{21}, \frac{4}{21}, \frac{5}{21}, \frac{6}{21}\right)$$

(c) Calculate $L = \sum p_i l_i$ in each case.

5. *The game of Hi-Lo.*

   (a) A computer generates a number X according to a known probability mass function $p(x), x \in \{1, 2, ...100\}$. The player asks a question, "Is $X = i$?" and is told "Yes","You're too high" or "You're too low." He continues for a total of six questions. If he is right (i.e. he receives the answer "yes") during this sequence, he receives a prize of value $v(X)$. How should he proceed to maximize his expected winnings?

   (b) The above doesn't have much to do with information theory. Consider the following variation: $X \sim p(x)$, prize $=v(x), p(x)$ known, as before. But *arbitrary* Yes-No questions are asked sequentially until X is determined. ("Determined" doesn't mean that a "yes" answer is received.) Questions cost one unit each. How should the player proceed? What is the expected return?

   (c) Continuing 2, what if $v(x)$ is fixed, but $p(x)$ can be chosen by the computer (and then announced to the player)? The computer wishes to minimize the player's expected return. What should $p(x)$ be? What is the expected return to the player?

6. *Huffman codes with costs.* Words like Run! Help! and Fire! are short, not because they are frequently used, but perhaps because time is precious in the situations in which these words are required. Suppose that $X = i$ with probability $p_i, i = 1, 2, ..., m$. Let l_i be the number of binary symbols in the codeword associated with $X = i$, and let c_i denote the cost per letter of the codeword when $X = i$. Thus the average cost C of the description of X is $C = \sum_{i=1}^{m} p_i c_i l_i$.

   (a) Minimize C over all $l_1, l_2, ..., l_m$ such that $\sum 2^{-l_i} \leq 1$. Ignore any implied integer constraints on l_i. Exhibit the minimizing $l_1^*, l_2^*, ...l_m^*$ and the associated minimum value C^*.

   (b) How would you use the Huffman code procedure to minimize C over all uniquely decodable codes? let $C_{Huffman}$ denote this minimum.

   (c) Can you show that

$$C^* \leq C_{Huffman} \leq C^* + \sum_{i=1}^{m} p_i c_i$$

7. *Shannon code.* Consider the following method for generating a code for a random variable X which takes on m values $\{1, 2, ...m\}$ with probabilities $p_1, p_2, ...p_m$. Assume that the probabilities are ordered so that $p_1 \geq p_2 \geq .... \geq p_m$. Define

$$F_i = \sum_{k=1}^{i-1} p_i$$

   the sum of the probabilities of all symbols less than i. Then the codeword for i is the number $F_i \in [0, 1]$ rounded off to l_i bits, where $l_i = \lceil \log \frac{1}{p_i} \rceil$.

   (a) Show that the code constructed by this process is prefix-free and the average length satisfies

$$H(X) \leq L \leq H(X) + 1$$

   (b) Construct the code for the probability distribution (0.5,0.25,0.125,0.125).