COMPSCI 650: HOMEWORK 1

INSTRUCTOR: A. MAZUMDAR

All problems carry equal points
Some problems are from Cover and Thomas, 2nd Ed.
Due Date: Feb 16, 2016: Start of Class

(1) For two random variables X and Y, let the joint distribution is given as follows:

$$p(x, y) = \begin{cases} 1/3, & \text{when } x = 0, y = 0 \\ 1/3, & \text{when } x = 0, y = 1 \\ 1/3, & \text{when } x = 1, y = 1 \\ 0, & \text{when } x = 1, y = 0. \end{cases}$$

Find out:
- $H(X)$, $H(Y)$
- $H(X|Y)$, $H(Y|X)$
- $H(X, Y)$
- $H(Y) - H(Y|X)$
- $I(X; Y)$.

(2) Mutual information of heads and tails. (5 points).
- Consider a fair coin flip. What is the mutual information between the top and bottom sides of the coin?
- A six-sided fair die is rolled. What is the mutual information between the top side and the front face (the side most facing you)?

(3) A source emits letters from an alphabet $\mathcal{X} = \{1, 2, 3, 4, 5\}$, with $p(1) = 0.15, p(2) = 0.04, p(3) = 0.26, p(4) = 0.05, p(5) = 0.50$.
- Calculate the entropy of this source.
- Find a Huffman code for this source.
- Find the average length of the Huffman code.

(4) Let X and Y be random variables that take on values $x_1, x_2, \ldots, x_r$ and $y_1, y_2, \ldots, y_s$, respectively. Let $Z = X + Y$.
- Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus, the addition of independent random variables adds uncertainty.
- Give an example of (necessarily dependent) random variables in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.
- Under what conditions does $H(Z) = H(X) + H(Y)$?

(5) Consider two probability distributions p and q on the alphabet $\mathcal{X} = \{a, b, c\}$, given by
$$p(a) = 1/2; p(b) = 1/4; p(c) = 1/4;$$
and
$$q(a) = 1/3; q(b)1/6; q(c) = 1/2.$$
Compute $H(p), H(q)$, $D(p||q)$ and $D(q||p)$.

(6) Answer the following short questions.
- Consider the code $\{0, 01\}$. Is this code uniquely decodable? Why? Is it instantaneous?
- Suppose $\mathcal{X} = \{0, 1\}$. The random variable (source) X takes value in $\mathcal{X}$, with $\Pr(X = 0) = \frac{3}{4}$ and $\Pr(X = 1) = \frac{1}{4}$. What is the probability that the source produce a sequence 0000011111?

(7) Write the Lempel-Ziv parsing for the file
$$\mathcal{F} = 00000101110001000010100001000111001010101 00100.$$
What is the number of bits that you need to write the entire compressed file (with LZ algorithm).

(8) Consider a random variable X that takes on four values with probabilities $(1/3, 1/3, 1/4, 1/12)$.
- Construct a Huffman code for this random variable.
- Show that there exist two different sets of optimal lengths for the codewords; namely, show that codeword length assignments $(1, 2, 3, 3)$ and $(2, 2, 2, 2)$ are both optimal.
- Conclude that there are optimal codes with codeword lengths for some symbols that exceed the Shannon code length $\lceil \log \frac{1}{\log p(a)} \rceil$.

(9) Although the codeword lengths of an optimal variable-length code are complicated functions of the message probabilities $\{p_1, p_2, \ldots, p_m\}$, it can be said that less probable symbols are encoded into longer codewords. Suppose that the message probabilities are given in decreasing order, $p_1 > p_2 \geq \cdots \geq p_m$.
- Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 > 2/5$, that symbol must be assigned a codeword of length 1.
- Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 < 1/3$, that symbol must be assigned a codeword of length ≥ 2.

Bonus question: Chernoff bound.

Suppose a biased coin with probability of heads being equal to p is tossed n times. Suppose the number of heads is H. Using the definition of KL-divergence and the bounds in the class prove that,
$$\Pr(H \geq (p + \epsilon)n) \leq e^{-cnp\epsilon^2},$$
for an appropriate constant c.