ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 3 Information Theory and Coding
Homework 2 October 2, 2007

PROBLEM 1.

(a) A source has an alphabet of 4 letters, a_1, a_2, a_3, a_4, and we have the condition $P(a_1) > P(a_2) = P(a_3) = P(a_4)$. Find the smallest number q such that $P(a_1) > q$ implies that $n_1 = 1$ where n_1 throughout this problem is the length of the codeword for a_1 in a Huffman code.

(b) Show by example that if $P(a_1) = q$ (your answer in part (a)), then a Huffman code exists with $n_1 > 1$.

(c) Now assume the more general condition, $P(a_1) > P(a_2) \geq P(a_3) \geq P(a_4)$. Does $P(a_1) > q$ still imply that $n_1 = 1$? Why or why not?

(d) Now assume that the source has an arbitrary number K of letters with $P(a_1) > P(a_2) \geq \cdots \geq P(a_K)$. Does $P(a_1) > q$ now imply $n_1 = 1$?

(e) Assume $P(a_1) \geq P(a_2) \geq \cdots \geq P(a_K)$. Find the largest number q' such that $P(a_1) < q'$ implies that $n_1 > 1$.

PROBLEM 2. A random variable takes values on an alphabet of K letters, and each letter has the same probability. These letters are encoded into binary words so as to minimize the average code word length. Define j and x so that $K = x2^j$, where j is an integer and $1 \leq x < 2$.

(a) Do any code words have lengths not equal to j or $j + 1$? Why?

(b) In terms of j and x, how many code words have length j?

(c) What is the average code word length?

PROBLEM 3. Consider the following method for constructing binary code words for a random variable U which takes values $\{a_1, \ldots, a_m\}$ with probabilities $P(a_1), \ldots, P(a_m)$. Assume that $P(a_1) \geq P(a_2) \geq \cdots \geq P(a_m)$. Define

$$Q_i = \sum_{k=1}^{i-1} P(a_k) \quad \text{for } i > 1;\ Q_1 = 0.$$

The code word assigned to the message a_i is formed by finding the binary expansion of $Q_i < 1$ (i.e, $1/2 = 100\ldots$, $1/4 = 0100\ldots$, $5/8 = 1010\ldots$) and then truncating this expansion to the first l_i bits where $l_i = \lceil -\log_2 P(a_i) \rceil$.

(a) Construct binary code words for the probability distribution $\{1/4, 1/4, 1/8, 1/8, 1/16, 1/16, 1/16, 1/16\}$.

(b) Prove that the method described above yields an instantaneous code (i.e., no codeword is a prefix of another) and the average codeword length $\bar{L}$ satisfies

$$H(X) \leq \bar{L} < H(X) + 1.$$

Problem 4. Let $p_{XY}(x,y)$ be given by

$X \backslash Y$	0	1
0	1/3	1/3
1	0	1/3

Find

(a) $H(X)$, $H(Y)$.

(b) $H(X|Y)$, $H(Y|X)$.

(c) $H(X,Y)$.

(d) $H(Y) - H(Y|X)$.

(e) $I(X;Y)$.

Problem 5. (Problem 2.1 from textbook)
Coin flips. A fair coin is flipped until the first head occurs. Let X be the number of flips required.

(a) Find the entropy $H(X)$ in bits. The following expressions may be useful:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}, \quad \sum_{n=1}^{\infty} n r^n = \frac{r}{(1-r)^2}.$$

(b) A random variable X is drawn according to this distribution. Find an "efficient" sequence of yes-no questions to determine X of the form "Is X contained in the set S?" Compare $H(X)$ to the expected number of questions to determine X.

Problem 6. Let X be a random variable taking values in M points $a_1, \ldots, a_M$, and let $P_X(a_M) = \alpha$. Show that

$$H(X) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha} + (1-\alpha) H(Y)$$

where Y is a random variable taking values in $M-1$ points $a_1, \ldots, a_{M-1}$ with probabilities $P_Y(a_j) = P_X(a_j)/(1-\alpha)$; $1 \le j \le M-1$. Show that

$$H(X) \le \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha} + (1-\alpha) \log(M-1)$$

and determine the condition for equality.

Problem 7. Let X, Y, Z be discrete random variables. Prove the validity of the following inequalities and find the conditions for equality:

(a) $I(X,Y;Z) \ge I(X;Z)$.

(b) $H(XY|Z) \ge H(X|Z)$.

(c) $I(X;Z|Y) \ge I(Z;Y|X) - I(Z;Y) + I(X;Z)$.

(d) $H(X,Y,Z) - H(X,Y) \le H(X,Z) - H(X)$.