

## ECE 534: Elements of Information Theory

Midterm Exam I (Spring 2007)

### Problem 1 [20 pts]

What are the relations ( $\geq$ ,  $=$ ,  $\leq$ ) between the following pairs of expressions? Explain why.

- 1)  $H(5X)$  and  $H(X)$ ;
- 2)  $H(X_0|X_{-1})$  and  $H(X_0|X_{-1}, X_1)$ ;
- 3)  $H(X_1, X_2, \dots, X_n)$  and  $H(c(X_1, X_2, \dots, X_n))$ , where  $c(x_1, x_2, \dots, x_n)$  is the Huffman codeword assigned to  $(x_1, x_2, \dots, x_n)$ ;
- 4)  $H(X, Y)$  and  $H(X) + H(Y)$ .
- 5)  $H(X_n|X_1, \dots, X_{n-1})$  and  $H(X_{n+1}|X_1, \dots, X_n)$ , where  $(X_1, X_2, \dots, X_n, \dots)$  is a stationary random process.

**Solution:** (Provided by Bhanu Kamandla and Ishita Basu)

- 1)  $H(5X) = H(X)$ , since entropy depends on the probability of the variable not on the value of the variable. Another explanation is that  $f(x) = 5x$  is an invertible function of  $x$ .
- 2)  $H(X_0|X_{-1}) \geq H(X_0|X_{-1}, X_1)$ , since conditioning reduces entropy.
- 3)  $H(X_1, X_2, \dots, X_n) = H(c(X_1, X_2, \dots, X_n))$ , since Huffman code is uniquely decodable, that is  $c(X_1, \dots, X_n)$  is an invertible function.
- 4)  $H(X, Y) = H(X) + H(Y|X) \leq H(X) + H(Y)$  (chain rule and conditioning reduces entropy)
- 5)  $H(X_n|X_1, \dots, X_{n-1}) = H(X_{n+1}|X_2, \dots, X_n)$  Since  $X_1, X_2, \dots, X_n$  is a stationary random process.  $H(X_{n+1}|X_1, X_2, \dots, X_n) \leq H(X_{n+1}|X_2, \dots, X_n)$ , since conditioning reduces entropy. Therefore  $H(X_n|X_1, \dots, X_{n-1}) \geq H(X_{n+1}|X_1, X_2, \dots, X_n)$ .

### Problem 2 [20 pts]

(a) Find a binary Huffman code for the following random variable:

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ 0.35 & 0.26 & 0.19 & 0.07 & 0.04 & 0.04 & 0.03 & 0.02 \end{pmatrix}$$

(b) Find a ternary Huffman code for the above source.

**Solution:** (Provided by Bhanu Kamandla)

(a) The binary Huffman coding tree for this problem can be found in Fig. 1, from which we can see that the codewords are 00, 01, 11, 1000, 1010, 1011, 10010, 10011.

(b) The number of symbols is 8 (not an odd number), therefore a dummy symbol  $x_9$  with zero probability needs to be added. From the ternary coding tree of Fig. 2, we can see that the codewords are 0, 1, 20, 21, 220, 221, 2220, 2221.

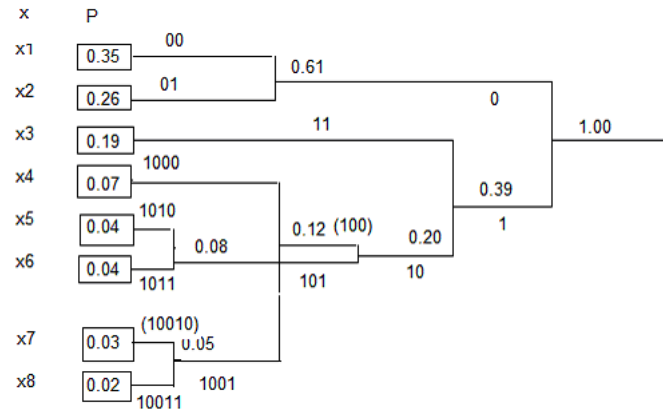


Fig. 1. Binary Huffman Coding Tree

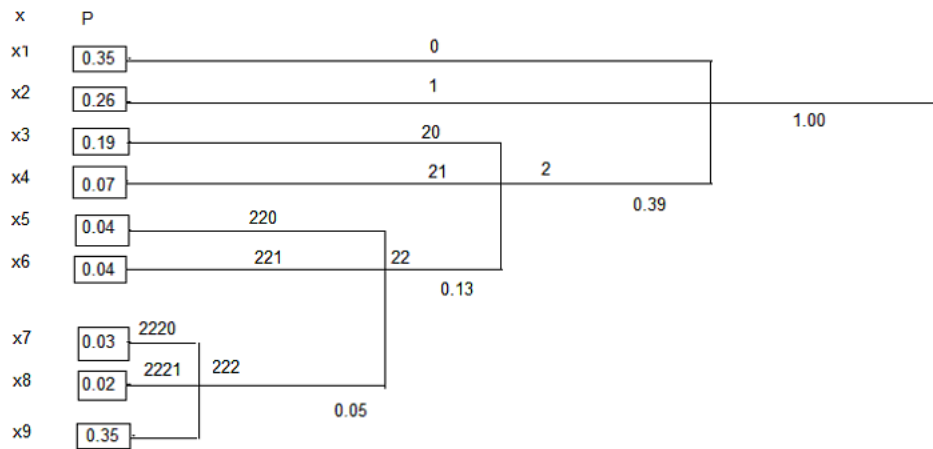


Fig. 2. Ternary Huffman Coding Tree

**Problem 3** [20 pts]

A random variable  $X$  has the following probability mass function

$$P(x) = \begin{cases} \frac{1}{4} & x = 0, 1 \\ \frac{1}{2} & x = 2 \end{cases}$$

Random variables  $Y$  and  $Z$  are generated as follows. If  $X = 0$ , then  $Y = Z = 0$ ; if  $X = 1$ , then  $Y = 1$  and  $Z = 0$ ; if  $X = 2$ , then  $Z = 1$  while  $Y$  is randomly chosen from 0 and 1 with equal probability. Find the values of the following quantities:  $H(X)$ ,  $H(Y)$ ,  $H(Z)$ ,  $H(Y|X)$ ,  $H(X, Y)$ ,  $H(X|Y)$ ,  $H(X, Z)$ ,  $H(X|Z)$ ,  $H(Y, Z)$ ,  $H(Z|Y)$ .

**Solution:** (Provided by Ishita Basu)

The probability distribution of  $Y$  is  $P(Y = 0) = P(X = 0) + 0.5P(X = 2) = 0.5$  and  $P(Y = 1) = P(X = 1) + 0.5P(X = 2) = 0.5$ . The distribution of  $Z$  is  $P(Z = 0) = P(X = 0) + P(X = 1) = 0.5$  and  $P(Z = 1) = P(X = 2) = 0.5$ . We can also see that  $P(Y, Z) = P(Y)P(Z) = 0.25$  for  $Y, Z = 0, 1$ , meaning that  $Z$  and  $Y$  are independent. Therefore,

$$H(X) = 0.25 \log 4 + 0.25 \log 4 + 0.5 \log 2 = 1.5$$

$$\begin{aligned}
H(Y) &= H(0.5) = 1 \\
H(Z) &= H(0.5) = 1 \\
H(Y|X) &= P(X = 0)H(Y|X = 0) + P(X = 1)H(Y|X = 1) + P(X = 2)H(Y|X = 2) \\
&= 0 + 0 + 0.5H(0.5) = 0.5 \\
H(X, Y) &= H(X) + H(Y|X) = 1.5 + 0.5 = 2 \\
H(X|Y) &= H(X, Y) - H(Y) = 2 - 1 = 1 \\
H(Z|X) &= 0 \quad (Z \text{ is completely determined by } X) \\
H(X, Z) &= H(X) + H(Z|X) = 1.5 \\
H(X|Z) &= H(X, Z) - H(Z) = 1.5 - 1 = 0.5 \\
H(Y, Z) &= H(Y) + H(Z) = 1 + 1 = 2 \\
H(Z|Y) &= H(Y, Z) - H(Y) = 2 - 1 = 1
\end{aligned}$$

**Problem 4** [20 pts]

$(X_1, X_2, \dots, X_n, X_{n+1}, \dots, X_{2n})$  are i.i.d. random variables with  $P(X_i = 0) = p$  and  $P(X_i = 1) = 1 - p$ . For  $i = 1, \dots, n$ , let  $U_i = X_i X_{n+i}$  and  $V_i = X_i + X_{n+i}$  (modulo 2 addition, that is,  $1 + 1 = 0$ ).

- (a) Find the achievable rate region for distributed encoding of  $(U_1, \dots, U_n)$  and  $(V_1, \dots, V_n)$ .  
(b) Suppose  $(U_1, \dots, U_n)$  and  $(V_1, \dots, V_n)$  are successfully decoded, what is the amount of remaining uncertainty about  $(X_1, \dots, X_{2n})$ ?

**Solution:** (Provided by Ishita Basu)

The distribution of  $U$  is  $P(U = 0) = 1 - (1 - p)^2$  and  $P(U = 1) = (1 - p)^2$ . The distribution of  $V$  is  $P(V = 0) = p^2 + (1 - p)^2$  and  $P(V = 1) = 2p(1 - p)$ . Therefore,

$$\begin{aligned}
H(U) &= H((1 - p)^2) \\
H(V) &= H(2p(1 - p)).
\end{aligned}$$

The joint distribution of  $U$  and  $V$  is given by  $P[U, V = 00] = p^2$ ,  $P[U, V = 01] = 2p(1 - p)$ , and  $P[U, V = 10] = (1 - p)^2$ . Therefore,

$$H(U, V) = -[p^2 \log(p^2) + p(1 - p) \log(p(1 - p)) + (1 - p)^2 \log((1 - p)^2)]$$

Say we send  $U_i$  at a rate  $R_u$  and  $V_i$  at a rate  $R_v$ . According to Slepian-Wolf Theorem, the achievable rate region is given by:

$$\begin{aligned}
R_u &\geq H(U|V) = H(U, V) - H(V) \\
R_v &\geq H(V|U) = H(U, V) - H(U) \\
R_u + R_v &\geq H(U, V)
\end{aligned}$$

b) When  $U_i, V_i = 00$  or  $10$ , we can uniquely find the values of  $X_i$  and  $X_{n+i}$ . But if  $U_i, V_i = 01$ , the  $X_i, X_{n+i}$  can be  $01$  or  $10$  with equal probability. Therefore,

$$H(X_i, X_{i+n}|U_i, V_i) = P(U_i = 0, V_i = 1)H(0.5) = 2p(1 - p)$$

and,

$$H(X_1, \dots, X_{2n}|U_1, \dots, U_n, V_1, \dots, V_n) = 2np(1 - p).$$

**Problem 5** [20 pts]

(a) Show that  $h(p) \triangleq -p \log p - (1-p) \log(1-p)$  is maximized when  $p = 0.5$ .

(b) Suppose  $X$  has the following distribution  $P(X = x_i) = p_i, i = 1, \dots, M$ ;  $Y$  has the following distribution

$$P(Y = x_i) = \begin{cases} p_i & i \neq m, n \\ \frac{p_m + p_n}{2} & i = m, n \end{cases}$$

Show that  $H(Y) \geq H(X)$ .

(c) Prove that entropy  $H(X)$  is a concave function of  $P_X$ .

**Solution:** (Provided by Stephano Rini)

(a)

$$\begin{aligned} h(0.5) - h(p) &= p \log 2 + (1-p) \log 2 - p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p} \\ &= p \log \frac{p}{0.5} + (1-p) \log \frac{1-p}{0.5} \geq 0 \end{aligned}$$

where the last inequality comes from the fact that divergence is always nonnegative.

(b) Using the definition of entropy:

$$\begin{aligned} H(Y) - H(X) &= \sum_{i=0, i \neq n, m}^M p_i \log \left( \frac{1}{p_i} \right) + (p_m + p_n) \log \left( \frac{2}{p_n + p_m} \right) + \sum_{i=0}^M p_i \log(p_i) \\ &= p_m \log p_m + p_n \log p_n + \log(p_m + p_n) \log \left( \frac{2}{p_n + p_m} \right) \\ &= p_m \left( \log(p_m) + \log \left( \frac{2}{p_n + p_m} \right) \right) + p_n \left( \log(p_n) + \log \left( \frac{2}{p_n + p_m} \right) \right) \\ &= p_m \log \left( \frac{2p_m}{p_n + p_m} \right) + p_n \log \left( \frac{2p_n}{p_n + p_m} \right) \\ &= (p_n + p_m) D(q(x) \| l(x)) \\ &\geq 0 \end{aligned}$$

By definition of relative entropy, where:

$$q(x) = \begin{cases} p_n / (p_n + p_m) & x = n \\ p_m / (p_n + p_m) & x = m \end{cases}$$

$$l(x) = \begin{cases} 1/2 & x = n \\ 1/2 & x = m \end{cases}$$

*Part c*

Let  $P_X = [p_1, p_2, \dots]$ ,  $H(P_X)$  is a continuous function, being a composition of continuous functions.

Moreover:

$$\begin{aligned} \frac{\partial H(P_X)}{\partial p_x} &= -\log(p_x) - 1 \\ \frac{\partial^2 H(P_X)}{\partial p_x^2} &= -\frac{1}{p_x} \\ \frac{\partial^2 H(P_X)}{\partial p_x \partial p_j, j \neq x} &= 0 \end{aligned}$$

Therefore  $\nabla_{P_X}^2 H(X) \leq 0$  and  $H(X)$  is concave.

(Alternative proof): Given any two distributions  $P_X$  and  $Q_X$ , we have for  $\alpha \in [0, 1]$ ,

$$\begin{aligned}
H(\alpha P_X + (1 - \alpha)Q_X) - \alpha H(P_X) - (1 - \alpha)H(Q_X) &= E_{\sim \alpha P_X + (1 - \alpha)Q_X} \log \frac{1}{\alpha P_X + (1 - \alpha)Q_X} \\
&\quad - \alpha E_{\sim P_X} \log \frac{1}{P_X} - (1 - \alpha) E_{\sim Q_X} \log \frac{1}{Q_X} \\
&= \alpha E_{\sim P_X} \log \frac{P_X}{\alpha P_X + (1 - \alpha)Q_X} \\
&\quad + (1 - \alpha) E_{\sim Q_X} \log \frac{Q_X}{\alpha P_X + (1 - \alpha)Q_X} \\
&= \alpha D(P_X \| \alpha P_X + (1 - \alpha)Q_X) \\
&\quad + (1 - \alpha) D(Q_X \| \alpha P_X + (1 - \alpha)Q_X) \\
&\geq 0
\end{aligned}$$