ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 15 | Information Theory and Coding

Midterm Exam | Oct. 29, 2013

3 problems, 110 points
3 hours
2 sheets (4 pages) of notes allowed

Good Luck!

PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS

PLEASE WRITE THE SOLUTION OF EACH PROBLEM ON A SEPARATE SHEET

PROBLEM 1. (35 points) Suppose we are given a source alphabet $\mathcal{U}$ and a set of distributions $\{p_\alpha : \alpha \in A\}$ on $\mathcal{U}$. (I.e., for each $\alpha \in A$, p_α is a probability distribution on $\mathcal{U}$.)

Let $q(u) = \max_{\alpha \in A} p_\alpha(u)$ and $Q = \sum_{u \in \mathcal{U}} q(u)$.

(a) (5 pts) Show that there is a prefix-free code $\mathcal{C}$ for the alphabet $\mathcal{U}$ for which

$$\text{length}(\mathcal{C}(u)) = \left\lceil \log \frac{Q}{q(u)} \right\rceil.$$

(b) (5 pts) For the code $\mathcal{C}$ in (a), show that no matter which p_α is the distribution of U,

$$E[\text{length}(\mathcal{C}(U))] - H(U) \leq 1 + \log Q$$

(c) (5 pts) Show that $Q \leq |A|$.

(d) (5 pts) Suppose there is a subset B of A such that for each $u \in \mathcal{U}$

$$\max_{\alpha \in A} p_\alpha(u) = \max_{\alpha \in B} p_\alpha(u).$$

Show that $Q \leq |B|$.

(e) (10 pts) Suppose $\mathcal{U} = \{0,1\}^n$, $A = [0,1]$, and for $(u_1, \ldots, u_n) \in \{0,1\}^n$,

$$p_\alpha(u_1, \ldots, u_n) = \alpha^k (1-\alpha)^{n-k} \text{ where } k \text{ is the number of 1's in } u_1, \ldots, u_n.$$

Show that $B = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$ has the property described in (d).

(f) (5 pts) Show that there is a prefix-free code $\mathcal{C} : \{0,1\}^n \to \{0,1\}^*$ for which for any i.i.d. binary random variables $U_1, \ldots, U_n$

$$\frac{1}{n} E[\text{length}(\mathcal{C}(U_1, \ldots, U_n))] - H(U_1) \leq \frac{1}{n}[1 + \log(n+1)].$$

PROBLEM 2. (30 points) Suppose we have a distribution p on an alphabet $\mathcal{U}$ for which

$$\max_u p(u) < 2 \min_u p(u). \tag{$*$}$$

(a) (5 pts) Show that every Huffman code for $\mathcal{U}$ satisfies

$$\max_u \operatorname{length}(\mathcal{C}(u)) - \min_u \operatorname{length}(\mathcal{C}(u)) \leq 1.$$

(b) (10 pts) Replace the strict inequality in $(*)$ by $\max_u p(u) \leq 2 \min_u p(u)$. Show that there exists a Huffman code for $\mathcal{U}$ with

$$\max_u \operatorname{length}(\mathcal{C}(u)) - \min_u \operatorname{length}(\mathcal{C}(u)) \leq 1.$$

(c) (10 pts) Suppose we express the cardinality $|\mathcal{U}|$ of the source alphabet in the form $|\mathcal{U}| = 2^j + r$ with $0 \leq r < 2^j$. Show that the Huffman code for $\mathcal{U}$ will have $2^j - r$ codewords of length j and $2r$ codewords of length $j + 1$.

(d) (5 pts) Show that the expected codeword length for the Huffman code equals $j + \alpha$ where α is the sum of the probabilities of the $2r$ least likely codewords.

PROBLEM 3. (45 pts) Suppose X, Y are random variables with joint distribution p_{XY}. Suppose Alice knows (X, Y) and needs to communicate X to Bob, who already knows Y.

Consider the following method. For each y, design a Huffman code $\mathcal{C}_y$ for X using the distribution p_y where $p_y(x) = p_{X|Y}(x|y)$. Alice sends Bob $\mathcal{C}_Y(X)$ based on her knowledge of Y and X. Note that which code she uses depends on Y.

(a) (10 pts) Show that the expected codeword length (averaged over both X and Y) satisfies
$$H(X|Y) \le E[\text{length}(\mathcal{C}_Y(X))] \le H(X|Y) + 1.$$

Suppose $U_1, U_2, \ldots$ is a stationary source. We will encode this source by the following means.

1. Fix integers $m \ge 1$ and $k \ge 1$.
2. Use the method described above with $Y = (U_1, \ldots, U_m)$ and $X = (U_{m+1}, \ldots, U_{m+k})$ to construct codes $\mathcal{C}_y$ for each $y \in \mathcal{U}^m$.
3. Describe U_1^m by using a trivial code using $\lceil m \log |\mathcal{U}| \rceil$ bits.
4. Describe $X_1 = U_{m+1}^{m+k}$ using $\mathcal{C}_{Y_1}$ with $Y_1 = U_1^m$; describe $X_2 = U_{m+k+1}^{m+2k}$ using $\mathcal{C}_{Y_2}$ with $Y_2 = U_{k+1}^{m+k}$; describe $X_3 = U_{m+2k+1}^{m+3k}$ using $\mathcal{C}_{Y_3}$ with $Y_3 = U_{k+1}^{m+k}$; $\ldots$.

(b) (5 pts) Explain how we can recover the source sequence from the output of this source code.

(c) (10 pts) Let L_n be the number of bits produced while the source coder processes the first $m + nk$ letters. Show that the expected number of bits per source letter $\rho = \lim_{n\to\infty} \frac{1}{m+nk} E[L_n]$ satisfies
$$\frac{1}{k} H(U_{m+1}^{m+k} | U_1^m) \le \rho \le \frac{1}{k}[H(U_{m+1}^{m+k} | U_1^m) + 1].$$

(d) (5 pts) Show that for a given k, $\frac{1}{k} H(U_{m+1}^{m+k} | U_1^m)$ is nonincreasing in m.

(e) (10 pts) Show that for a given m, $\frac{1}{k} H(U_{m+1}^{m+k} | U_1^m)$ is nonincreasing in k.

(f) (5 pts) Find the limit of $\frac{1}{m} H(U_{m+1}^{2m} | U_1^m)$ in terms of the entropy rate of the process $U_1, U_2, \ldots$.