ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 16 Information Theory and Coding
Midterm Exam Oct. 27, 2015

3 problems, 40 points
165 minutes
1 sheet (2 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. (13 points) Suppose $\mathcal{U} = \{1, \dots, K\}$ is an alphabet with K symbols. Let $S_{n_1,\dots,n_K}$ denote the set of sequences $(u_1, \dots, u_n)$ of length $n = n_1 + \cdots + n_K$ that contain exactly n_i occurrences of symbol i for each $i = 1, \dots, K$. Note that the size of $S_{n_1,\dots,n_K}$ is given by the multinomial coefficient,

$$|S_{n_1,\dots,n_K}| = \binom{n}{n_1, \dots, n_K} = \frac{n!}{n_1! \dots n_K!}.$$

Also note that the multinomial coefficient is also encountered when we express

$$(x_1 + \cdots + x_K)^n = \sum_{\substack{n_1,\dots,n_K:\\ n_1+\cdots+n_K=n\\ n_i \ge 0}} \binom{n}{n_1, \dots, n_K} x_1^{n_1} \dots x_K^{n_K}.$$

(a) (2 pts) Show that for any non-negative $x_1, \dots, x_K$ that sum to 1,

$$\binom{n}{n_1, \dots, n_K} \le x_1^{-n_1} \dots x_K^{-n_K}.$$

(b) (2 pts) Show that

$$\log |S_{n_1,\dots,n_K}| \le n h(p_1, \dots, p_K)$$

where $p_i = n_i/n$ and $h(p_1, \dots, p_K) = -\sum_i p_i \log p_i$.

(c) (3 pts) Show that there is a prefix free code $\mathcal{C}_n : \mathcal{U}^n = \{0,1\}^*$ such that

$$\text{length}\big(\mathcal{C}_n(u_1, \dots, u_n)\big) = (K-1)\lceil \log(1+n) \rceil + \lceil \log |S_{n_1,\dots,n_K}| \rceil$$

where n_i is the number of occurrences of i in $(u_1, \dots, u_n)$.

(d) (3 pts) Suppose $(X_1, \dots, X_K)$ are $[0,1]$ valued random variables such that $\sum_i X_i = 1$ and $E[X_i] = \mu_i$. Show that $E\big[h(X_1, \dots, X_K)\big] \le h(\mu_1, \dots, \mu_K)$.

[*Hint:* We know from class that $D\big((X_1, \dots, X_K)\big\|(\mu_1, \dots, \mu_K)\big) := \sum_i X_i \log(X_i/\mu_i)$ is non-negative.]

(e) (3 pts) Suppose $U_1, U_2, \dots$ are i.i.d., show that for the code in (c),

$$\frac{1}{n} E[\text{length}(\mathcal{C}_n(U_1, \dots, U_n))] \le H(U) + \frac{K + (K-1)\log(1+n)}{n}.$$

PROBLEM 2. (17 points) Suppose $\ldots, X_{-1}, X_0, X_1, X_2, \ldots$ is a stationary Markov process, and $U_i = f(X_i)$ where $f : \mathcal{X} \to \mathcal{U}$ is a deterministic function.

For $i \geq 1$ let $a_i = H(U_i|U_{i-1}, \ldots, U_1)$ and $b_i = H(U_i|U_{i-1}, \ldots, U_1, X_0)$.

(a) (2 pts) The process $\{U_i : i \in \mathbb{Z}\}$ is not necessarily Markov. Is it stationary?

(b) (2 pts) What is the value of $I(X_0; U_2, \ldots, U_{i+1}|X_1)$?

(c) (3 pts) Show that

$$b_i = H(U_{i+1}|U_i, \ldots, U_2, X_1, X_0)$$

(d) (3 pts) Show that $b_{i+1} \geq b_i$.

(e) (3 pts) Let $d_i = a_i - b_i$. Show that d_i is non-negative and

$$\sum_{i=1}^{n} d_i = I(X_0; U_1, \ldots, U_n).$$

(f) (2 pts) Show that $d_{i+1} \leq d_i$.

(g) (2 pts) Show that $d_n \leq (\log |\mathcal{X}|)/n$ and conclude that $\lim_{n\to\infty} b_n$ exists and is equal to the entropy rate of the process $\{U_i : i \in \mathbb{Z}\}$.

PROBLEM 3. (10 points) Suppose $U_1, U_2, \dots$ is a source producing an i.i.d. sequence of letters, and $\mathcal{D}$ is a valid and prefix free dictionary for it.

Suppose w_0 is a word in $\mathcal{D}$ and let the dictionary $\mathcal{D}'$ be obtained from the dictionary $\mathcal{D}$ by replacing the word $w_0 \in \mathcal{D}$ by its single letter extensions.

Let W denote the first word in the parsing of $U_1, U_2, \dots$ with the dictionary $\mathcal{D}$. Similarly, let W' denote the first word in the parsing of $U_1, U_2, \dots$ with the dictionary $\mathcal{D}'$. Let $p_0 = \Pr(W = w_0)$.

(a) (2 pts) Express $E[\text{length}(W')] - E[\text{length}(W)]$ in terms of p_0.

(b) (3 pts) By explicit evaluation of $H(W)$ and $H(W')$, express $H(W') - H(W)$ in terms of p_0 and $H(U)$.

For $k = 1, 2, \dots$, let S_k be the statement that "for any valid and prefix-free dictionary with k interior nodes $H(W) = H(U)E[\text{length}(W)]$."

(c) (2 pts) Show that the statement S_1 is true.

(d) (3 pts) Using (a) and (b) show that S_k implies S_{k+1}, and conclude that for any valid and prefix-free dictionary $H(W) = H(U)E[\text{length}(W)]$.