ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 16 Information Theory and Coding
Midterm Exam Nov. 8, 2011

3 problems, 100 points
3 hours
1 sheets of notes allowed

Good Luck!

PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS

PROBLEM 1. (36 points) Suppose $\mathcal{C}$ is a *non-singular* code for an alphabet $\mathcal{U}$. (Recall that non-singularity is the property that $\mathcal{C}(u) \neq \mathcal{C}(v)$ for any $u, v \in \mathcal{U}$, $u \neq v$.)

Let $\ell(u) := \text{length}(\mathcal{C}(u))$ denote the length of the codeword assigned to the letter u. Note that ℓ takes values in the non-negative integers $\{0, 1, 2, \dots\}$.

(a) (8 pts) Suppose $\mathcal{C}_0$ is a prefix-free code for the non-negative integers. Show that the code

$$\tilde{\mathcal{C}}(u) := \mathcal{C}_0(\ell(u))\,\mathcal{C}(u)$$

obtained by concatenating the codeword $\mathcal{C}_0(\ell(u))$ with $\mathcal{C}(u)$ is a uniquely decodable code for the alphabet $\mathcal{U}$. Is $\tilde{\mathcal{C}}$ prefix-free?

(b) (4 pts) Show that there is a prefix-free code for nonnegative integers where the integer n is assigned a binary string of length

$$\ell_0(n) := \lceil 2\log_2(n+1) + 1 \rceil.$$

[Hint: $\sum_{n=0}^{\infty} 1/(n+1)^2 < 2$.]

(c) (8 pts) Suppose $\mathcal{U} = \{1, 2, \dots, K\}$ and the probabilities $p_1, \dots, p_K$ of the letters satisfy $p_1 \geq p_2 \geq \cdots \geq p_K$. How should one choose the *non-singular* code $\mathcal{C}$ to minimize the expected codeword length? Show that letter j is assigned a codeword of length

$$\ell(j) = \lfloor \log_2 j \rfloor.$$

(d) (4 pts) With $p_1, \dots, p_K$ as above show that

$$jp_j \leq 1 \quad \text{for all } j = 1, \dots, K.$$

(e) (4 pts) For the *non-singular* code $\mathcal{C}$ with shortest expected length chosen in part (c), show that there is a uniquely decodeable code $\tilde{\mathcal{C}}$ with lengths

$$\tilde{\ell}(j) \leq \ell(j) + \lceil 2\log_2(1 + \log_2(1/p_j)) + 1 \rceil.$$

(f) (8 pts) Show that any non-singular code satisfies

$$H(U) \leq E[\ell(U)] + 2\log_2(1 + H(U)) + 2.$$

[Hint: $\log_2$ is a concave function, so that $\sum_j p_j \log_2(x_j) \leq \log_2(\sum_j p_j x_j)$.]

PROBLEM 2. (28 points) Suppose $U_1, U_2, \ldots, U_n$ are i.i.d. random variables taking values in the alphabet $\mathcal{U}$ with common distribution p.

Suppose further, that S is a subset of $\mathcal{U}^n$ with the property that

$$\Pr\big(U^n \in S\big) > 1 - \delta$$

where $0 < \delta < 1$.

(a) (4 pts) Let A denote the set of ϵ-typical sequences of length n with respect to the distribution p. Show that

$$\Pr\big(U^n \in S \cap A\big) > 1 - \delta - \epsilon$$

for large enough n.

(b) (8 pts) Show that

$$|S \cap A| > (1 - \delta - \epsilon) 2^{nH(p)(1-\epsilon)}$$

for large enough n.

(c) (8 pts) Suppose $\tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_n$ are i.i.d. random variables taking values in the alphabet $\mathcal{U}$ with common distribution $\tilde{p}$.

Show that

$$\Pr\big(\tilde{U}^n \in S\big) \geq (1 - \delta - \epsilon) 2^{-n(1+\epsilon)D(p\|\tilde{p})}\, 2^{-n2\epsilon H(p)}$$

(d) (8 pts) Suppose we are told that a sequence $U_1, U_2, U_3, \ldots$ is either i.i.d. with distribution p or i.i.d. with distribution $\tilde{p}$. We are asked to design a device that decides between these two alternatives based on the observation $u_1, \ldots, u_n$.

Let α_n be the probability of the event "device decides $\tilde{p}$" when the true distribution is p. Let β_n be the probability of the event "device decides p" when the true distribution is $\tilde{p}$.

Show that for any $0 < \delta < 1$, if $\alpha_n \leq \delta$ then

$$\beta_n \mathrel{\dot{\geq}} 2^{-nD(p\|\tilde{p})}$$

where "$\dot{\geq}$" means "greater than while ignoring $n\epsilon$ terms in the exponents".

[Hint: Define S to be the set of sequences u^n for which the device decides p.]

PROBLEM 3. (36 points) Suppose $U_1, U_2, \ldots$ is a stationary Markov process, and the process $V_1, V_2, \ldots$ is obtained from the U process by setting

$$V_i = f(U_i)$$

where f is a function from $\mathcal{U}$ to $\mathcal{V}$. Note that V is also a stationary process.

Let H_U denote the entropy rate of the process $U_1, U_2, \ldots$ and H_V denote the entropy rate of the process $V_1, V_2, \ldots$

(a) (12 pts) Justify each of the following equalities and inequalities.

$$\begin{aligned} H(V_{n+m}|V_{n+m-1}, \ldots, V_1) &\geq H(V_{n+m}|V_{n+m-1}, \ldots, V_{n+1}, U_n, \ldots, U_1) && \text{(A1)} \\ &= H(V_{n+m}|V_{n+m-1}, \ldots, V_{n+1}, U_n) && \text{(A2)} \\ &= H(V_{m+1}|V_m, \ldots, V_2, U_1) && \text{(A3)} \\ &= H(V_{m+1}|V_m, \ldots, V_1, U_1) && \text{(A4)} \end{aligned}$$

(b) (4 pts) Show that

$$H_V \geq H(V_{m+1}|V_m, \ldots, V_1, U_1)$$

[Hint: $H_V = \lim_{n\to\infty} H(V_{m+n}|V_{m+n-1}, \ldots, V_1)$.]

(c) (4 pts) Show that

$$H(V_{m+1}|V_m, \ldots, V_1) - H(V_{m+1}|V_m, \ldots, V_1, U_1) = I(U_1; V_{m+1}|V_1, \ldots, V_m)$$

(d) (8 pts) Justify

$$\begin{aligned} H(U_1) &\geq I(U_1; V_1, V_2, V_3, \ldots) && \text{(D1)} \\ &= \sum_{m=0}^{\infty} I(U_1; V_{m+1}|V_1, \ldots, V_m) && \text{(D2)} \end{aligned}$$

(e) (4 pts) Show that $\lim_{m\to\infty} I(U_1; V_{m+1}|V_1, \ldots, V_m) = 0$.

(f) (4 pts) Show that the sequence $H(V_{m+1}|V_m, \ldots, V_1, U_1)$ approaches H_V from below.