ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 17 | Information Theory and Coding
Midterm Exam | Nov. 13, 2012

3 problems, 110 points
2 hours 45 minutes
2 sheets of notes allowed

Good Luck!

Please write your name on each sheet of your answers

Please use a separate sheet to answer each question

Problem 1. (35 points) We will consider source codes for the alphabet $\{0,1\}^n$.

(a) (10 pts) Show that there is a prefix-free code $\mathcal{C}_n$ which assigns to each sequence $u^n \in \{0,1\}^n$ a codeword of length

$$\left\lceil \log_2(n+1) + \log_2 \binom{n}{k} \right\rceil$$

where k is the number of 1's in u^n. (Here $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ denotes the binomial coefficient.)

(b) (5 pts) Show that for any $0 \le z \le 1$, and for all $0 \le k \le n$,

$$\binom{n}{k} \le z^{-k}(1-z)^{-(n-k)}.$$

Hint: $1 = (z + (1-z))^n = \sum_{i=0}^{n} \binom{n}{i} z^i (1-z)^{n-i}$.

(c) (10 pts) Show that

$$\frac{1}{n} \log_2 \binom{n}{k} \le h_2\bigl(k/n\bigr)$$

where $h_2(z) = -z \log_2 z - (1-z)\log_2(1-z)$ is the binary entropy function.

(d) (10 pts) Show that if $U_1, \dots, U_n$ are i.i.d. binary random variables, with entropy $\Pr(U_1 = 1) = p$, then

$$\frac{1}{n} E\bigl[\text{length}(\mathcal{C}_n(U^n))\bigr] \le h_2(p) + \frac{1 + \log_2(n+1)}{n}$$

[Hint: h_2 is concave $\cap$.]

PROBLEM 2. (35 points) Denote the output of a stationary memoryless source with alphabet $\mathcal{U}$ by $U_1, U_2, \ldots$. Let p_U denote the common distribution of these random variables.

As always, we will encode the source into a binary representation. However, the bits 0 and 1 have possibly unequal cost; let $c(0)$ and $c(1)$ denote these costs. We will assume $c(0) \geq 0$ and $c(1) \geq 0$.

We will assume that the cost of a binary sequence $x^n = (x_1, \ldots, x_n)$ is given by $c(x^n) = \sum_{i=1}^{n} c(x_i)$.

For $0 \leq q \leq 1$ and $\epsilon > 0$, let $T_2(n, \epsilon, q)$ denote the set of binary sequences of length n that are ϵ-typical with respect to the distribution $P(X = 0) = 1 - q$, $P(X = 1) = q$.

(a) (5 pts) Show that for $x^n \in T_2(n, \epsilon, q)$, we have

$$\frac{1}{n} c(x^n) \leq \big[(1-q)c(0) + qc(1)\big](1+\epsilon).$$

We will consider a fixed-to-fixed encoding strategy. To start with, we only provide representations to typical source sequences of length m.

(b) (5 pts) Denoting by $T(m, \epsilon, p_U)$ the typical source sequences of length m, show that

$$2^{m(1+\epsilon)H(U)} \leq (1-\epsilon)2^{n(1-\epsilon)h_2(q)}$$

is a sufficient condition to assign to each member of $T(m, \epsilon, p_U)$ a unique representative in $T_2(n, \epsilon, q)$, for large enough n.

(c) (5 pts) Show that the cost per source letter of the above scheme is at most

$$[(1-q)c(0) + qc(1)](1+\epsilon)\frac{n}{m}.$$

(d) (10 pts) Show that for any $\delta > 0$, there is a fixed-to-fixed encoding strategy (i.e., a choice of m, n, ϵ) that has cost per source letter at most

$$\frac{(1-q)c(0) + qc(1)}{h_2(q)} H(U)(1+\delta)$$

and negligible probability that U^m is not assigned a representation.

(e) (5 pts) Explain how would you take care of non-typical u^m's?

(f) (5 pts) Conclude that for any $\delta > 0$ we can represent data at an expected cost of at most

$$\min_{0 \leq q \leq 1} \frac{(1-q)c(0) + qc(1)}{h_2(q)}(1+\delta)$$

per information bit.

Problem 3. (40 points) Suppose we are given a discrete memoryless channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and also an encoder enc : $\{1,\dots,M\} \to \mathcal{X}^n$ of blocklength n and rate $\frac{1}{n}\log_2 M$ to transmit information over this channel.

Denote enc(m) by $(c_1(m),\dots,c_n(m))$, and define the following distributions on $\mathcal{X}$:

$$p_i(x) = \frac{1}{M}\#\{m : c_i(m) = x\}, \quad i = 1,\dots,n,$$

and $\bar{p}(x) = \frac{1}{n}\sum_{i=1}^n p_i(x)$.

Let I_i denote the value of the mutual information $I(X;Y)$ when X has distribution p_i, and let $\bar{I}$ denote the mutual information when X had distribution $\bar{p}$.

(a) (5 pts) Suppose W is a message uniformly chosen among $1,\dots,M$. The message W is encoded to $X^n = \text{enc}(W)$ which is sent over the channel and Y^n is received. Show that the distribution of X_i is p_i.

(b) (10 pts) Show that

$$\frac{1}{n}I(W;Y^n) \le \frac{1}{n}\sum_{i=1}^n I_i.$$

[Hint: Consider how $I(W;Y^n)$, $I(X^n;Y^n)$ and $\sum_i I(X_i;Y_i)$ are related to each other.]

(c) (10 pts) Show that

$$\frac{1}{n}I(W;Y^n) \le \bar{I}.$$

(d) (5 pts) Suppose a communications engineer claims that his newest code can reliably communicate over a binary symmetric channel at rates very close to the capacity. Remarkably, his code has the property that each message is encoded into a sequence that is 30% 1's and 70% 0's. Show that the engineer's claim is wrong. What is the highest rate he may hope to communicate with such a design?

(e) (10 pts) Suppose an encoder is constructed by choosing each $c_i(m)$ randomly and independently, according to a distribution p. Since the encoder is the outcome of a random experiment, the distributions p_i, $\bar{p}$, and the quantities $I(W;Y^n)$, I_i, $\bar{I}$ are all random variables. Show that I_i are i.i.d. with

$$E[I_i] \le I(p)$$

where $I(p)$ denote the mutual information $I(X;Y)$ when X has distribution p.