ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 35 Information Theory and Coding
Final exam Jan. 29, 2019

4 problems, 85 points
165 minutes
2 sheet (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. (25 points) Suppose a binary code of blocklength n with $M = 2^{nR}$ codewords is constructed by random coding, by choosing each letter of each codeword independently by a fair coin flip. Let $\mathbf{X}(1), \dots, \mathbf{X}(M)$ denote the codewords by this procedure.

(a) (4 pts) For $m \neq m'$, what is $\Pr(\mathbf{X}(m) = \mathbf{X}(m'))$?

(b) (5 pts) Let $G_i = 1$ if $\mathbf{X}(i)$ is different from $\mathbf{X}(1), \dots, \mathbf{X}(i-1)$, and $G_i = 0$ otherwise. Find $\Pr\big(G_i = 1 \mid G_1 = \dots = G_{i-1} = 1\big)$. [Hint: the event $G_1 = ... = G_{i-1} = 1$ is the same as $\mathbf{X}(1), ..., \mathbf{X}(i-1)$ being distinct.]

(c) (4 pts) Find $\Pr(G_1 = \dots = G_M = 1)$.

(d) (4 pts) Let q denote the probability that all codewords are distinct (i.e., for every $m \neq m'$, $\mathbf{X}(m) \neq \mathbf{X}(m')$.) Using (c) and the identity $1 - x \leq \exp(-x)$, show that $q \leq \exp(-\sum_{i=1}^{M}(i-1)/2^n)$.

(e) (4 pts) Show that for $R > 1/2$, $q \to 0$ as n gets large, i.e., for rates larger than $1/2$ and large blocklength a random code will have repeated codewords with high probability.

(f) (4 pts) Suppose now that $\mathbf{X}(1), ..., \mathbf{X}(M)$ are chosen independently (but not necessarily according to the "i.i.d letter"s procedure above. Show that the value of q found above is an upper bound to the probability that $\mathbf{X}(1), ..., \mathbf{X}(M)$ are all distinct. [Hint: show that $\Pr(\mathbf{X}(m) = \mathbf{X}(m'))$ is lower bounded by the value you found in (a).]

Problem 2. (12 points) Consider random variables X_1, X_2, Y_1, Y_2.

(a) (4 pts) Show that

$$I(X_1, X_2\,;\, Y_1, Y_2) \geq I(X_1; Y_1) + I(X_2; Y_2)$$

when X_1 and X_2 are independent.

Consider now two discrete memoryless channels whose outputs Y_1 and Y_2 depend on their inputs x_1 and x_2 as

$$Y_1 = f_1(x_1, Z_1), \quad Y_2 = f_2(x_2, Z_2)$$

where f_1 and f_2 are deterministic functions, and, Z_1 and Z_2 are random variables (perhaps dependent) chosen independently of the inputs (x_1, x_2).

A communication system has access to both channels, i.e., the effective channel between the transmiter and the receiver takes as input the pair (x_1, x_2), and outputs the pair (Y_1, Y_2).

(b) (3 pts) Show that the capacity of the effective channel is larger than the sum of the capacities of the individual channels.

(c) (5 pts) Suppose the inputs x_1, x_2 are binary. Further suppose $Z_1 = Z_2$ and is equally likely to be 0 or 1. Suppose

$$f_1(x_1, z_1) = x_1 + z_1 \mod 2, \quad f_2(x_2, z_2) = x_2 + z_2 \mod 2.$$

What are the capacities of the individual channels? What is the capacity of the effective channel?

Problem 3. (22 points) Consider a linear code defined over the ternary alphabet $\mathbb{F}_3 = \{0, 1, 2\}$ (equipped with modulo-3 addition and multiplication) as follows: $\mathbf{x}$ is a codeword if and only if $H\mathbf{x} = \mathbf{0}$ where

$$H = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

(and all operations are done in modulo-3 arithmetic).

(a) (4 pts) What is the blocklength, the number of codewords, and the rate of this code?

A codeword $\mathbf{x}$ is sent over a channel. It is known that during the transmission either all letters are received correctly, or, one of the letters is changed (to some other element of $\mathbb{F}_3$).

(b) (5 pts) Show that the receiver can detect if a change has happened and correct it if so.

(c) (4 pts) Suppose we are allowed to augment the matrix H by appending to it a fifth column. How will this change the rate of the code?

(d) (4 pts) Which of the following candidate columns (if any) can be appended to H and still preserve the property in (b): $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$?

(e) (5 pts) Suppose it is known that during the transmission all letters are received correctly, or one of the letters is changed in the following restricted way: 0 can be replaced by 1 (but not by 2); 1 can be replaced by 2 (not by 0); 2 can be replaced by 0 (not by 1). Redo part (d) for this channel.

PROBLEM 4. (26 points) Consider a multiple access channel with inputs $X_1 \in \{0,1\}$, $X_2 \in \{0,1\}$ and output $Y \in \{0,1,2\}$ given by $Y = X_1 + X_2$. Note that the channel is noiseless, $Y = 0$ when $(X_1, X_2) = (0,0)$, $Y = 2$ when $(X_1, X_2) = (1,1)$, and $Y = 1$ otherwise.

(a) (5 pts) What is the capacity region of this channel?

Consider now this multiple access channel with feedback: both the encoders get to see the value the past channel outputs $Y_1, \dots, Y_{i-1}$ before transmitting X_{1i} and X_{2i}.

Consider the following transmission scheme. Messages $m_1 = (u_{11}, \dots, u_{1k})$ an $m_2 = (u_{21}, \dots, u_{2k})$ are k-bit sequences, where $u_{11}, \dots, u_{1k}, u_{21}, \dots, u_{2k}$'s are i.i.d and equally likely to be 0 and 1. The transmission takes place in two phases:

Phase 1 (of duration k): the encoders send the messages uncoded, i.e., $X_{1i} = u_{1i}$ and $X_{2i} = u_{2i}$, $i = 1, \dots, k$. Let $T = \sum_{i=1}^{k} \mathbb{1}\{Y_i = 1\}$ be the number of times $Y_i = 1$, and let $i_1, \dots, i_T$ be the values of i for which $Y_i = 1$ in the first phase. Note that T, and $i_1, \dots, i_T$ are known to both the encoders and also to the receiver.

Phase 2: You will design phase 2 below.

(b) (4 pts) $(u_{1i_1}, \dots, u_{1i_T})$ is a T-bit long sequence. Let $Q \in \{0, \dots, 2^T - 1\}$ denote the T bit integer with this binary representation. At the end of phase 1, who (among the encoders and the receiver) knows the value of Q?

(c) (5 pts) Let $S = T \log_3 2$ so that $2^T \leq 3^{\lceil S \rceil}$. Let $(v_1, \dots, v_{\lceil S \rceil})$ be the ternary representation of Q (i.e., Q is radix 3). Show how to design phase 2 of duration $\lceil S \rceil$ so that the receiver, during this phase, receives $v_1, \dots, v_{\lceil S \rceil}$.

(d) (4 pts) Let $N = \lceil k + S \rceil$ denote the total transmission time. Find $E[k + S]$.

(e) (4 pts) What value does $k/E[N]$ approach as k gets large?

(f) (4 pts) Use the law of large numbers to find $\lim_{k\to\infty} T/k$. Using $\log_3(2) < 2/3$, show that $R = \lim_{k\to\infty} k/N > 3/4$.