ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 30 | Information Theory and Coding
Final Exam | Jan. 11, 2016

4 problems, 60 points
180 minutes
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. (12 points) Let $X_1, X_2, \dots$ be a stationary binary source. An observer tries to guess the current source symbol on the basis of his past observations of the source. For $n = 1, 2, \dots$ let $\hat{X}_n = f_n(X_1, \dots, X_{n-1})$ denote the guess by the observer for X_n after observing $X^{n-1} = (X_1, \dots, X_{n-1})$. Here each $f_n : \{0,1\}^{n-1} \to \{0,1\}$ is a deterministic function, in particular, $\hat{X}_1$ is a constant.

(a) (4 pts) Let Z_n be the indicator variable of the event $\hat{X}_n \neq X_n$, i.e., $Z_n = 0$ if the observer guesses the correctly, $Z_n = 1$ otherwise. Express the entropy rate of the process $Z_1, Z_2, \dots$ in terms of the entropy rate of the source.

(b) (4 pts) Let $p_n = \Pr(Z_n = 1)$ denote the probability that the observer guesses incorrectly. Show that $h_2(p_n) \geq H(X_n|X^{n-1})$, where h_2 is the binary entropy function.

(c) (4 pts) Let $p = \liminf_{n\to\infty} p_n$ denote the 'error rate' of the observer. Show that $h_2(p)$ cannot be smaller than the entropy rate of the source.

PROBLEM 2. (16 points) Consider a two-way communication system where two parties communicate via a *common* output they both can observe and influence. Denote the common output by Y, and the signals emitted by the two parties by x_1 and x_2 respectively. Let $p(y|x_1, x_2)$ model the memoryless channel through which the two parties influence the output.

We will consider feedback-free block codes, i.e., we will use encoding and decoding functions of the form

$$\begin{aligned} &\text{enc}_1 \colon \{1, \dots, 2^{nR_1}\} \to \mathcal{X}_1^n \qquad &&\text{dec}_1 \colon \mathcal{Y}^n \times \{1, \dots, 2^{nR_1}\} \to \{1, \dots, 2^{nR_2}\} \\ &\text{enc}_2 \colon \{1, \dots, 2^{nR_2}\} \to \mathcal{X}_2^n \qquad &&\text{dec}_2 \colon \mathcal{Y}^n \times \{1, \dots, 2^{nR_2}\} \to \{1, \dots, 2^{nR_1}\} \end{aligned}$$

with which the parties encode their own message and decode the other party's messages. (Note that when a party is decoding the other party's message, it can make use of the knowledge of its own message).

We will say that the rate pair (R_1, R_2) is achievable, if for any $\epsilon > 0$, there exist encoders and decoders with the above form for which the average error probability is less than ϵ.

Consider the following 'random coding' method to construct the encoders:

(i) Choose probability distributions p_j on $\mathcal{X}_j$, $j = 1, 2$.

(ii) Choose $\{\text{enc}_1(m_1)_i : m_1 = 1, \dots, 2^{nR_1}, i = 1, \dots, n\}$ i.i.d., each having distribution as p_1. Similarly, choose $\{\text{enc}_2(m_2)_i : m_2 = 1, \dots, 2^{nR_2}, i = 1, \dots, n\}$ i.i.d., each having distribution as p_2, independently of the choices for enc_1.

For the decoders we will use typicality decoders:

(i) Set $p(x_1, x_2, y) = p_1(x_1)p_2(x_2)p(y|x_1, x_2)$. Choose a small $\epsilon > 0$ and consider the set T of ϵ-typical (x_1^n, x_2^n, y^n)'s with respect to p.

(ii) For decoder 1: given y^n and the correct m_1, dec_1 will declare $\hat{m}_2$ if it is the unique m_2 for which $(\text{enc}_1(m_1), \text{enc}_2(m_2), y^n) \in T$. If there is no such m_2, dec_1 outputs 0. (Similar description applies to Decoder 2.)

(a) (3 pts) Given that m_1 and m_2 are the transmitted messages, show that $(\text{enc}_1(m_1), \text{enc}_2(m_2), Y^n) \in T$ with high probability.

(b) (3 pts) Given that m_1 and m_2 are the transmitted messages, and $\tilde{m}_1 \neq m_1$ what is the probability distribution of $(\text{enc}_1(\tilde{m}_1), \text{enc}(m_2), Y^n)$?

(c) (3 pts) Under the assumptions in (b) show that the

$$\Pr\{(\text{enc}_1(\tilde{m}_1), \text{enc}_2(m_2), Y^n) \in T\} \doteq 2^{-nI(X_1;X_2Y)}.$$

(d) (3 pts) Show that all rate pairs satisfying

$$R_1 \leq I(X_1; YX_2), \quad R_2 \leq I(X_2; YX_1)$$

for some $p(x_1, x_2) = p(x_1)p(x_2)$ are achievable.

(e) (4 pts) For the case when X_1, X_2, Y are all binary and Y is the product of X_1 and X_2, show that the achievable region is strictly larger than what we can obtain by 'half duplex communication' (i.e., the set of rates that satisfy $R_1 + R_2 \leq 1$.)

Problem 3. (16 pts) Suppose $\{(X_i, Y_i) : i = 1, 2, \dots\}$ is an i.i.d. sequence of pairs of discrete random variables. Let $p(x, y)$ denote the probability mass function of each pair. Suppose $X_1, X_2, \dots$ is observed by Alice and $Y_1, Y_2, \dots$ is observed by Bob. Alice needs to inform Bob of the sequence she has seen. Consider the following method to accomplish this:

(i) To each ϵ-typical X sequence of length n assign a 'label' randomly and uniformly chosen from $\{1, \dots, 2^{nR}\}$. The assignments are made independently. Let $\text{label}(x^n)$ denote the label assigned to the sequence x^n by this process.

(ii) Upon observing X^n, Alice checks if it is typical and if so, sends $\text{label}(X^n)$ to Bob.

(iii) Upon observing Y^n and receiving the label ℓ from Alice, Bob makes a list of all X sequences $\hat{x}^n$ for which $(\hat{x}^n, Y^n)$ is jointly typical and $\text{label}(\hat{x}^n) = \ell$. If the list contains a single sequence, Bob decides that it is what Alice observed.

(a) (4 pts) As n gets large, what is the chance that the true sequence X^n does not appear on Bob's list?

(b) (4 pts) For a given typical sequence y^n, find an upper bound on the number of x^n sequences that are jointly typical with y^n.

[Hint: mimic the proof for bounding the size of the typical set, noting that for such sequences $p(x^n, y^n) \approx 2^{-nH(X,Y)}$ and $p(y^n) \approx 2^{-nH(Y)}$.]

(c) (4 pts) For a typical y^n, upper bound the expected number of wrong sequences that appear on Bob's list.

(d) (4 pts) Find a condition of the form $R > R_0$ that guarantees that Bob will decide correctly with high probability.

PROBLEM 4. (16 points) Suppose $\mathcal{C}$ is a Reed–Solomon code defined on a field $\mathbb{F}$ with blocklength n, $|\mathbb{F}|^k$ codewords. Let $\alpha_1 \in \mathbb{F}, \ldots, \alpha_n \in \mathbb{F}$ denote the evaluation points that define this code — recall that the Reed–Solomon code maps k information symbols $(u_0, \ldots, u_{k-1}) \in \mathbb{F}^k$ to the codeword $(x_1, \ldots, x_n) \in \mathbb{F}^n$ by setting $x_i = u(\alpha_i)$ where $u(D) = u_0 + u_1 D + \cdots + u_{k-1} D^{k-1}$.

Consider now the code $\mathcal{C}'$ of blocklength $n+1$ that assigns to the information sequence $(u_0, \ldots, u_{k-1})$ the codeword $\mathbf{x}' = (u_{k-1}, x_1, \ldots, x_n)$, where the x_i's are as above.

(a) (4 pts) Show that $\mathcal{C}'$ is linear.

(b) (4 pts) Suppose $u_0, \ldots, u_{k-1}$ are not all zero, but $u_{k-1} = 0$. Show that weight$(\mathbf{x}') \geq n + 2 - k$.

(c) (4 pts) Suppose $u_{k-1} \neq 0$. Show that weight$(\mathbf{x}') \geq n + 2 - k$.

(d) (4 pts) Show that the code $\mathcal{C}'$ satisfies the Singleton bound with equality.