ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 35 | Information Theory and Coding

Final exam | Feb. 1, 2022

4 problems, 80 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. (22 points) Consider a discrete memoryless channel W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. Let $W_{Y|X=x}$ denote the distribution of the channel output when the channel input is x.

Each letter $x \in \mathcal{X}$ has a cost $b(x)$. Assume that there is an $x_0 \in \mathcal{X}$ such that $b(x_0) = 0$, and for all $x \neq x_0$ we have $b(x) > 0$. (I.e., x_0 is the only 'free' input symbol).

(a) Show that for any probability distribution p_X on the input, we have

$$I(X;Y) = \sum_x p_X(x) D(W_{Y|X=x} \| p_Y),$$

where $p_Y(y) = \sum_x p_X(x) W_{Y|X=x}(y)$ is the distribution on Y induced by p_X.

(b) Show that for any probability distribution p_X on the input, we have

$$I(X;Y) \leq \sum_x p_X(x) D(W_{Y|X=x} \| W_{Y|X=x_0}).$$

Define $C_1 := \max_{x:x\neq x_0} \frac{D(W_{Y|X=x} \| W_{Y|X=x_0})}{b(x)}$.

(c) Show that for any probability distribution p_X on the input, we have $I(X;Y) \leq C_1 E[b(X)]$. [Hint: Use (b).]

(d) Show that for any probability distribution p_X on the input, and for any $x \in \mathcal{X}$, we have

$$I(X;Y) \geq p_X(x) D(W_{Y|X=x} \| p_Y),$$

where p_Y is again the distribution on Y imposed by p_X.

(e) For any $x_1 \neq x_0$ and $0 < \delta < 1$, consider the probability distribution p_X with $p_X(x_1) = \delta$, $p_X(x_0) = 1 - \delta$. Show that

$$I(X;Y)/E[b(X)] \geq b(x_1)^{-1} D\big(W_{Y|X=x_1} \| \delta W_{Y|X=x_1} + (1-\delta) W_{Y|X=x_0}\big).$$

[Hint: Use (d).]

(f) Show that $\sup_{p_X} \frac{I(X;Y)}{E[b(X)]} = C_1$.

Problem 2. (24 points) Consider a sequence of 1-bit memory locations, used by a "writer" to store data, and a"reader" later recovers it. However, a fraction p of the memory locations are faulty, these locations are stuck to a value that can't be changed by writing.

Our model for this setup is as follows: For each location $i = 1, 2, \dots,$ let $F_i = \mathbb{1}\{\text{location } i \text{ is faulty}\}$, and let $S_i \in \{0, 1\}$ denote the stuck value if $F_i = 1$. We assume F_i are i.i.d. with $\Pr(F_i = 1) = p$. With $X_i \in \{0, 1\}$ denoting what is written by the writer to location i, and $Y_i \in \{0, 1\}$ denoting what is read by the reader,

$$Y_i = \begin{cases} X_i & F_i = 0 \\ S_i & F_i = 1 \end{cases}.$$

We assume $F_1, F_2, \dots$ and $S_1, S_2, \dots$ are known to the writer in advance. So the writer knows which are the faulty locations and their stuck values; the reader however, is unaware of these.

We wish to design functions write : $(W_n, F^n, S^n) \mapsto X^n$ and read : $Y^n \mapsto \hat{W}_n$ so that an nR bits of data W_n (uniformly distributed in $\{1, ..., 2^{nR}\}$) can be written to n locations and read back as $\hat{W}_n$ with a small $\Pr(\hat{W}_n \neq W_n)$. Consider a "randomly constructed" read() function by setting $\{\text{read}(y^n) : y^n \in \{0, 1\}^n\}$ to be i.i.d., each uniform in $\{1, \dots, 2^{nR}\}$.

(a) For a fault vector f^n and stuck vector s^n, let $\mathcal{Y}(f^n, s^n) = \{y^n : y_i = s_i \text{ for all } i \text{ s.t } f_i = 1\}$. What is $|\mathcal{Y}(f^n, s^n)|$ in terms of $k = \sum_i f_i$?

(b) Fix $w \in \{1, .., 2^{nR}\}$. Conditional on $\sum_i F_i = k$, find the probability that for all $y^n \in \mathcal{Y}(F^n, S^n)$, $\text{read}(y^n) \neq w$.

(c) Show that the probability above can be upper bounded by $\exp(-2^{n(1-q-R)})$ where $q = k/n$ is the fraction of faulty locations. [Hint: $1 - x \leq \exp(-x)$.]

(d) Show that for $R < 1 - p$, $\Pr\big(\text{read}(y^n) \neq w \text{ for all } y^n \in \mathcal{Y}(F^n, S^n)\big) \to 0$ as n gets large. [Hint: Fix $q_0 \in (p, 1 - R)$. Let $K = \sum_i F_i$. Treat the cases $K/n < q_0$ and $K/n \geq q_0$ separately.]

(e) Show that for $R < 1 - p$ and for any $\epsilon > 0$, for n large enough, there exists read/write functions with $\Pr(\hat{W}_n \neq W_n) < \epsilon$.

(f) A colleague claims that he can design write/read functions with a value of R that is strictly larger than $1 - p$ with $\Pr(\hat{W}_n \neq W_n) \to 0$. Can he be right?

Problem 3. (18 points) Let $\mathbb{F}_2$ denote the binary field $\{0,1\}$ equipped with modulo 2 arithmetic. Recall that a binary linear encoder $\text{enc} : \mathbb{F}_2^k \to \mathbb{F}_2^n$ is described by an $n \times k$ generator matrix G so that for $u \in \mathbb{F}_2^k$, $\text{enc}(u) = Gu$.

Suppose $\text{enc}_1 : \mathbb{F}_2^{k_1} \to \mathbb{F}_2^n$ and $\text{enc}_2 : \mathbb{F}_2^{k_2} \to \mathbb{F}_2^n$ are binary linear encoders described by generator matrices G_1 and G_2. (Note that G_1 and G_2 both have n rows.) Consider a new binary linear encoder enc described by the generator matrix $G = \begin{bmatrix} G_1 & 0 \\ G_1 & G_2 \end{bmatrix}$.

(a) What is the blocklength of enc? What is the rate R of enc in terms of the rates R_1 and R_2 of enc_1 and enc_2?

Suppose $x \in \mathbb{F}_2^{2n}$. Write $x = \begin{bmatrix} x_1 \\ x_1 + x_2 \end{bmatrix}$ with x_1, $x_2 \in \mathbb{F}_2^n$. Let $w_H(.)$ denote the Hamming weight.

(b) Show that $w_H(x) \geq w_H(x_2)$.

(c) Show that $w_H(x) \geq 2w_H(x_1)\mathbb{1}(x_2 = 0) + w_H(x_2)\mathbb{1}(x_2 \neq 0)$.

[Hint: Consider the cases $x_2 = 0$ and $x_2 \neq 0$ separately.]

(d) With d denoting the minimum distance of enc, and d_i denoting the minimum distance of enc_i, $i = 1, 2$; show that $d = \min\{2d_1, d_2\}$.

[Hint: Show $d \geq \min\{2d_1, d_2\}$ using (c), then show $d \leq \min\{2d_1, d_2\}$.]

(e) Let $P(G_1, G_2)$ denote the generator matrix G constructed by the procedure above. For $i = 1, 2, \ldots$ let M_i be an all-1 column vector of dimension 2^i. Let $G_1 = I_2$, the 2×2 identity matrix. Let $G_{i+1} = P(G_i, M_i)$, $i = 1, 2, \ldots$. What is the blocklength n_i, rate R_i and minimum distance d_i of the linear encoder with generator matrix G_i as a function of i?

PROBLEM 4. (16 points) Let $X_1, \ldots, X_n$ be an i.i.d. source with p_X taking values in a finite alphabet $\mathcal{X}$, which is to be reconstructed in a finite alphabet $\mathcal{Y}$. Let $d(x,y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be such that $d(x,y) > 0$, and $d(X^n, Y^n) = \prod_i d(X_i, Y_i)^{1/n}$.

(a) For any enc : $\mathcal{X}^n \to \{1, \ldots, 2^{nR}\}$ and dec : $\{1, \ldots, 2^{nR}\} \to \mathcal{Y}^n$ such that $E[d(X^n; Y^n)] \leq D$ (here $Y^n = \text{dec}(\text{enc}(X^n))$), show that $R \geq R(D)$ where

$$R(D) := \inf_{p_{y|x} : E[\log d(X,Y)] \leq \log D} I(X;Y).$$

(b) Show that $R(D)$ is convex function of D.

Suppose we now construct a random quantizer exactly as what we did in class for rate-distortion. I.e., given D and $R > R(D)$, we choose a distribution $p_{Y|X}$ so that $I(X;Y) = R(D)$ and $E[\log d(X,Y)] = \log D$. Choose $\text{dec}(1), \ldots, \text{dec}(2^{nR})$ i.i.d. $\sim p_{Y^n}$. The encoder sets $\text{enc}(x^n) = m$ if there is an m for which $(x^n, \text{dec}(m))$ is ϵ-typical with respect to p_{XY}; otherwise $\text{enc}(x^n)$ is uniformly chosen among $\{1, ..., 2^{nR}\}$. Recall that "(x^n, y^n) is ϵ-typical" is the statement that for all x, y, $np_{XY}(x,y)(1-\epsilon) \leq \sum_i \mathbb{1}(x_i = x, y_i = y) \leq np_{XY}(x,y)(1+\epsilon)$. For the rest of the problem assume $1 \leq d(x,y) \leq d_{\max}$.

(c) For an ϵ-typical (x^n, y^n) pair, what can be the maximum value of $d(x^n, y^n)$?

(d) Show that $E[d(X^n, Y^n)] \leq \epsilon' + D^{1+\epsilon}$, where $\epsilon' \to 0$ as n gets large.