ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 15 | Information Theory and Coding
Midterm | Nov. 1, 2022

4 problems, 70 points
180 minutes
1 sheet (2 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. (10 points) Suppose $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$ are two stochastic processes defined on the same finite alphabet $\mathcal{U}$, with entropy rates H_U and H_V respectively. Let $p_i = \Pr(U_i \neq V_i)$.

(a) (2 pts) Show that $\frac{1}{n}H(U^nV^n) \leq \frac{1}{n}H(U^n) + \frac{1}{n}\sum_{i=1}^n H(V_i|U_i)$.

(b) (4 pts) Show that $\frac{1}{n}\sum_{i=1}^n H(V_i|U_i) \leq h_2\left(\frac{1}{n}\sum_{i=1}^n p_i\right) + \left(\frac{1}{n}\sum_{i=1}^n p_i\right)\log\left(|\mathcal{U}| - 1\right)$.

(c) (4 pts) Show that $|H_U - H_V| \leq h_2(p) + p\log\left(|\mathcal{U}| - 1\right)$, where $p = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n p_i$. (Assume that the limit exists.)

PROBLEM 2. (20 points) Let $\mathcal{U}$ be a finite alphabet. Write $|\mathcal{U}|$ in its binary expansion, $|\mathcal{U}| = \sum_{j=0}^{\infty} m_j 2^j$, with $m_j \in \{0,1\}$. Let $J = \{j : m_j = 1\}$, so that $|\mathcal{U}| = \sum_{j \in J} 2^j$. (E.g., with $|\mathcal{U}| = 13 = 8 + 4 + 1$, $J = \{0, 2, 3\}$.)

(a) (4 pts) Show that there exists an injective function $c : \mathcal{U} \to \{0,1\}^*$ such that $|\{u \in \mathcal{U} : \text{length}(c(u)) = j\}| = 2^j$ for $j \in J$. (With the example above, 8 elements of $\mathcal{U}$ get a three-bit representation by c, 4 elements get a two-bit representation, and one element gets a null representation.)

(b) (4 pts) With c as in (a) and U uniformly distributed on $\mathcal{U}$, let $W = c(U)$ and $L = \text{length}(W)$, so that $W = X_1 X_2 \dots X_L$ is a random binary sequence of (random) length L. Show that, conditioned on $L = i$, $X_1, \dots, X_i$ are i.i.d. with $\Pr(X_1 = 1 | L = i) = \frac{1}{2}$.

(c) (4 pts) Show that $H(W|L) = \mathbb{E}[L]$.

(d) (4 pts) Show that $H(L) \le \log(1 + \log |\mathcal{U}|)$. [Hint: L can take values only in the set J.]

(e) (4 pts) Show that $\mathbb{E}[L] \ge \log |\mathcal{U}| - \log(1 + \log |\mathcal{U}|)$. [Hint: use the chain rule for $H(WL)$.]

Problem 3. (20 points) Suppose U and V are random variables on the same finite alphabet $\mathcal{U}$ with probability distributions p_U and p_V respectively.

(a) (4 pts) Show that for any non-negative function $f : \mathcal{U} \to [0, \infty)$

$$\mathbb{E}[\log f(U)] - D(p_U \| p_V) \le \log \mathbb{E}[f(V)].$$

[Hint: We have shown in class for any probability distribution $p = (p_1, \dots, p_K)$ and non-negative $(x_1, \dots, x_K)$ we have $\sum_i p_i \log x_i \le \log(\sum_i p_i x_i)$.]

(b) (4 pts) For given p_U and p_V find an f for which the inequality in part (a) is an equality, and thus show that

$$D(p_U \| p_V) = \max_f \{\mathbb{E}[\log f(U)] - \log \mathbb{E}[f(V)]\}$$

where the maximization is taken over all non-negative functions defined on $\mathcal{U}$.

(c) (4 pts) Suppose $\tilde{\mathcal{U}}$ is a finite set, and $g : \mathcal{U} \to \tilde{\mathcal{U}}$ is a (deterministic) function. Let $\tilde{U} = g(U)$ and $\tilde{V} = g(V)$. Let $p_{\tilde{U}}$ and $p_{\tilde{V}}$ be their distributions. Show that

$$D(p_{\tilde{U}} \| p_{\tilde{V}}) \le D(p_U \| p_V).$$

[Hint: Note that for any non-negative function $\tilde{f}$ on $\tilde{\mathcal{U}}$, $f(u) = \tilde{f}(g(u))$ is a non-negative function on $\mathcal{U}$.]

Suppose X and Y are random variables defined on the same finite set $\mathcal{X}$ with joint distribution p_{XY}. Let X', Y' be random variables with joint distribution $p_X p_Y$. (I.e., X' and Y' are independent with the same marginal distributions as X and Y.) Let $p_e = \Pr(X \ne Y)$ and $q_e = \Pr(X' \ne Y')$.

(d) (4 pts) Show that

$$I(X;Y) \ge p_e \log \tfrac{p_e}{q_e} + (1 - p_e) \log \tfrac{1-p_e}{1-q_e}$$

[Hint: Both sides of the inequality are divergences.]

(e) (4 pts) Suppose now that X is uniformly distributed on $\mathcal{X}$. Use (d) to find an upper bound on $H(X|Y)$ in terms of p_e. [Hint: First find q_e.]

PROBLEM 4. (20 points) For an encoder $c : \mathcal{U}^* \to \{0,1\}^*$, we define the compressibility of the sequence as

$$\rho_c(u^\infty) := \lim_{n\to\infty} \frac{1}{n} \operatorname{length}(c(u^n)).$$

We also define the finite state compressibility of a sequence as the smallest $\rho_{c_M}(u^\infty)$, where c_M gives the output of an information lossless finite state machine M,

$$\rho_{FSM-IL}(u^\infty) := \inf_{c_M:\ M \text{ is FSM-IL}} \rho_{c_M}(u^\infty).$$

Let $\mathcal{U} = \{a, b\}$. Consider the sequence $u^\infty = abaaabbabbaaaaab\ldots$ which is formed by the concatenation of $null, a, b, aa, ab, ba, bb, aaa, aab, \ldots$ (listing the elements of $\mathcal{U}^*$ in non-decreasing length).

(a) (4 pts) Show that $\rho_{LZ}(u^\infty) = 1$. [Hint: How many bits does LZ produce at each iteration?]

(b) (4 pts) Let $U^\infty = U_1U_2\ldots$ be a stochastic process with a finite entropy rate H. Can Lempel-Ziv compress the sequence to H bits/letter in average? [Hint: What is the entropy rate of a stochastic process that produces a deterministic sequence?]

With u^∞ defined as before, let $v^\infty = u_1u_1u_2u_1u_2u_3u_1u_2u_3u_4\ldots$. In other words $v^\infty = x_1x_2x_3x_4\ldots$ where $x_n = u^n$.

(c) (4 pts) Find $\rho_{LZ}(v^\infty)$.

(d) (4 pts) Let M be an s-state information lossless finite-state machine. We give v^∞ above as input to M. Let y_n be the output the machine produces while processing the segment x_n. Show that

$$\forall \epsilon > 0\ \exists n_0(\epsilon)\ \forall n \geq n_0(\epsilon), \quad \frac{1}{n} \operatorname{length}(y_n) \geq 1 - \epsilon.$$

That is, for large enough n, the length of the output M produces while processing the segment x_n, divided by $n = \operatorname{length}(x_n)$, is arbitrarily close to 1.

[Hint: For any parsing of x_n into m distict words, $\operatorname{length}(y_n) \geq m \log \frac{m}{8s^2}$.]

(e) (4 pts) What is $\rho_{FSM-IL}(v^\infty)$? [Hint: Show that $\dfrac{\operatorname{length}(y_1y_2\ldots y_n)}{\operatorname{length}(x_1x_2\ldots x_n)} \geq 1 - 2\epsilon$ for n large enough.]