ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 16 Information Theory and Coding
Midterm exam Oct. 31, 2023

4 problems, 37 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 PAGE.

(All logarithms are taken to the base 2.)

PROBLEM 1. (7 points)

Suppose $X_1, X_2, \ldots$ are i.i.d. random variables with $\Pr(X_1 = 0) = \Pr(X_1 = 1) = 1/2$. Let $Y_n = \sum_{i=1}^{n} X_i$.

Hint: No explicit computation is necessary. For each part you only need to show that the quantity for $n+1$ is at least as large as the quantity for n. "Conditioning reduces entropy" is your friend.

(a) (2 points) Show that $H(Y_n)$ is nondecreasing in n.

(b) (2 points) Show that $H(X^n|Y_n)$ is nondecreasing in n.

(c) (3 points) Show that $H(X_n|Y_n)$ is nondecreasing in n.

Problem 2. (9 points)

Suppose $X_1, X_2, \ldots$ is a binary (i.e., $X_n \in \{0,1\}$) stationary process with entropy rate H. Define the following quantities:

$$\begin{aligned} p &= \Pr(X_1 = 1), \\ \alpha &= \Pr(X_2 = 1 \mid X_1 = 1), \\ \beta &= \Pr(X_2 = 0 \mid X_1 = 0). \end{aligned}$$

(Note that this does not necessarily imply that X is a Markov process.)

(a) (2 points) Show that $p = p\alpha + (1-p)(1-\beta)$.

(b) (2 points) Show that $H \leq ph_2(\alpha) + (1-p)h_2(\beta)$, where h_2 is the binary entropy function given by $h_2(x) = -x\log(x) - (1-x)\log(1-x)$.

(c) (2 points) Show that among all such stationary processes the Markov process has the largest entropy rate. [Recall that for a Markov process, $\Pr(X_n = x_n \mid X^{n-1} = x^{n-1}) = \Pr(X_n = x_n \mid X_{n-1} = x_{n-1})$.]

(d) (3 points) Suppose we have a stationary binary process for which every '1' is immediately followed by a '0'. Show that the entropy rate of this process is at most $\max_{a\in[0,1]} \frac{h_2(a)}{1+a}$.

PROBLEM 3. (9 points)

Suppose $U_1, U_2, \ldots$ are i.i.d. random variables on the alphabet $\mathcal{U}$ with distribution p_U, and define $H := H(U_1)$. Suppose $S_1, S_2, \ldots$ are sets with $S_n \subseteq \mathcal{U}^n$, and define $p_n := \Pr(U^n \in S_n)$. Pick $\epsilon > 0$ and let $T_n = T(n, p_U, \epsilon)$ be the typical sets as defined in class. Let $A_n = T_n \cap S_n$.

(a) (3 points) Show that, for large enough n, $\Pr(U^n \in A_n) \geq p_n - \epsilon$.

(b) (2 points) Show that, for large enough n, $|A_n| \geq (p_n - \epsilon)2^{n(1-\epsilon)H}$.
Hint: Any $u^n \in A_n$ also belongs to T_n.

(c) (2 points) Show that if $p := \lim_{n\to\infty} p_n > 0$, then $\lim_{n\to\infty} \frac{1}{n} \log |S_n| \geq H$.
Hint: Pick $\epsilon < p$, use (b), and note that S_n includes A_n.

(d) (2 points) Fix $\rho \geq 0$ and let $k_n = \lfloor n\rho \rfloor$. Consider assigning k_n-bit representations to n-letter words u^n via a function $f_n : \mathcal{U}^n \to \{0,1\}^{k_n}$ and attempting to recover the n-letter word u^n from the representation via a function $g_n : \{0,1\}^{k_n} \to \mathcal{U}^n$. Suppose $\lim_{n\to\infty} \Pr(U^n = g_n(f_n(U^n))) > 0$. Show that $\rho \geq H$.
Hint: Let $S_n := \{u^n : u^n = g_n(f_n(u^n))\}$.

Problem 4. (12 points)

Consider two binary codes, c_1 and c_2 for the *nonnegative* integers $\{0, 1, 2, \dots\}$. The code c_1 is defined as $c_1(n) = 1^n 0$; e.g, $c_1(3) = 1110$. The code c_2 is given as follows: $c_2(0) =$ null, $c_2(1) = 0, c_2(2) = 1, c_2(3) = 00, c_2(4) = 01, c_2(5) = 10, c_2(6) = 11, c_2(7) = 000$, and so on. Observe that $\text{length}(c_1(n)) = 1 + n$, and $\text{length}(c_2(n)) = \lfloor \log(1 + n) \rfloor$.

(a) (2 points) Is c_1 injective? Is it prefix-free? Is c_2 injective? Is it prefix-free?

(b) (2 points) With $n_1 = \text{length}(c_2(n))$, and $n_2 = \text{length}(c_2(n_1))$, consider the code formed by a concatenation $c(n) = c_1(n_2)c_2(n_1)c_2(n)$. Explain why c is prefix-free.

(c) (3 points) Show that there is a prefix-free code c_3 for *positive* integers $\{1, 2, \dots\}$ with $\text{length}(c_3(n)) \leq \log(n) + 2\log(1 + \log(n)) + 1$.

Suppose that $\dots, U_{-2}, U_{-1}, U_0, U_1, U_2, \dots$ are i.i.d. from an alphabet $\mathcal{U}$, and we observe U_i at time i. Let $N_i = \inf\{j > 0 : U_{i-j} = U_i\}$, i.e., the symbol U_i we observed at time i, was most recently observed at time $i - N_i$.

(d) (2 points) Show that $\Pr(N_i > j \mid U_i = u) = (1 - p_U(u))^j$, $j = 0, 1, \dots$. Conclude that $\mathbb{E}[N_i \mid U_i = u] = 1/p_U(u)$. [Fact: If $X \in \{0, 1, 2, \dots\}$, then $\mathbb{E}[X] = \sum_{i=0}^{\infty} \Pr(X > i)$.]

Suppose that we have already described the "past" $(\dots, U_{-2}, U_{-1}, U_0)$ in binary, we now describe $U_1, U_2, \dots$, by giving a binary descriptions of $N_1, N_2, \dots$, by the code c_3 above.

(e) (3 points) With $H = H(U_i) = H(U_1)$, show that $\mathbb{E}[\text{length}(c_3(N_i))] \leq H + 2\log(1 + H) + 1$.

Hint: First condition on $\{U_i = u\}$.