ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 34 Information Theory and Coding
Final exam Jan. 25, 2024

4 problems, 54 points
180 minutes
2 sheets (4 pages) of notes allowed.

Good Luck!

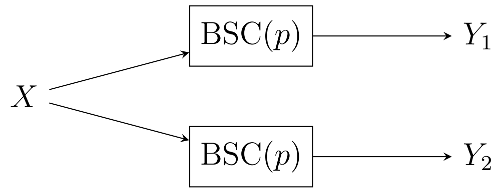
PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS.

PLEASE WRITE THE SOLUTION OF EACH PROBLEM ON A SEPARATE PAGE.

(All logarithms are taken to the base 2.)

Problem 1. (10 points)

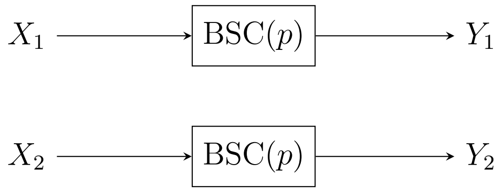
Consider the following binary input channel: the input X is passed through two identical independent BSC's with crossover probability p to produce binary outputs Y_1 and Y_2, as shown in the figure below. The channel's output is $Y = (Y_1, Y_2)$.



(a) (2 points) What is the capacity achieving input distribution?

(b) (4 points) What is the capacity C_1 of this channel (from X to (Y_1, Y_2))?

Consider another channel whose input is (X_1, X_2) with binary X_1 and X_2. X_1 is passed through a BSC(p) to produce Y_1; X_2 is passed through an independent BSC(p) to produce Y_2, as shown in the figure. The channel's output is again $Y = (Y_1, Y_2)$.



(c) (2 points) What is the capacity C_2 of this channel (from (X_1, X_2) to (Y_1, Y_2))?

(d) (2 points) How do C_1 and C_2 compare?

Problem 2. (12 points)

Consider a sequence of binary block codes $C_1, C_2, \ldots$ constructed as follows (with $[x, y]$ denoting the concatenation of x and y and $\mathbf{1}$ denoting the all-1 codeword $1 \ldots 1$ of appropriate length):

- $C_1 = \{0,1\}^2 = \{00, 01, 10, 11\}$,
- $C_{k+1} = \{[u, u] : u \in C_k\} \cup \{[u + \mathbf{1}, u] : u \in C_k\}$, for $k \geq 2$.

Note that C_k is a code of blocklength 2^k.

(a) (2 points) Show that C_k contains the all-1 codeword.
Hint: Use induction.

(b) (2 points) Show that C_k is linear.
Hint: Use induction.

(c) (2 points) Show that C_k has 2^{k+1} codewords, or equivalently, $|C_k| = 2^{k+1}$.
Hint: Use induction.

The Plotkin bound says the blocklength n, number of codewords M and minimum distance d of a binary code C satisfy

$$d \leq \left\lfloor \frac{nM}{2(M-1)} \right\rfloor .$$

(d) (2 points) Show that the minimum distance of C_k, $d_{\min}(C_k)$, satisfies $d_{\min}(C_k) \leq 2^{k-1}$.
Hint: Use the Plotkin bound on C_k.

(e) (4 points) Show that $d_{\min}(C_k) \geq 2^{k-1}$.
Hint: Show that $w_{\min}(C_k) \geq 2w_{\min}(C_{k-1})$, with $w_{\min}(C_k)$ denoting the minimum weight of the nonzero codewords in C_k.

Problem 3. (16 points)

Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be disjoint subsets of $\mathbb{Z}$ with $a_i = |\mathcal{A}_i|$, $i = 1, \ldots, n$, and let $\mathcal{B}_1, \ldots, \mathcal{B}_n$ be disjoint subsets of $\mathbb{Z}$ with $b_i = |\mathcal{B}_i|$, $i = 1, \ldots, n$. Let $\mathcal{C}_i = \mathcal{A}_i \times \mathcal{B}_i \subseteq \mathbb{Z}^2$ for $i = 1, \ldots, n$ (observe that $|\mathcal{C}_i| = a_i b_i$). Pick an index $C \in \{1, \ldots, n\}$ according to the probability distribution

$$\Pr(C = c) = \frac{a_c b_c}{\sum_{k=1}^n a_k b_k}.$$

Now, given $C = c$, pick two points (X_1, Y_1) and (X_2, Y_2) uniformly and independently from $\mathcal{C}_c = \mathcal{A}_c \times \mathcal{B}_c$, i.e.,

$$\Pr\Big((X_1, Y_1) = (x_1, y_1), (X_2, Y_2) = (x_2, y_2) \mid C = c\Big) = \frac{1}{(a_c b_c)^2} \mathbb{1}\Big\{(x_1, y_1) \in \mathcal{C}_c, (x_2, y_2) \in \mathcal{C}_c\Big\}.$$

Observe that this means that X_1, Y_1, X_2, Y_2 are pairwise conditionally independent given $C = c$.

(a) (2 points) Compute $\Pr\big((X_1, Y_1) = (x_1, y_1)\big)$ and $\Pr\big((X_2, Y_2) = (x_2, y_2)\big)$.

(b) (2 points) Show that $H(X_1, Y_1) + H(X_2, Y_2) = 2 \log \sum_{k=1}^n a_k b_k$.
Hint: Conclude from (a) that (X_1, Y_1) and (X_2, Y_2) are uniformly distributed on $\cup_{k=1}^n \mathcal{C}_k$.

(c) (4 points) Show that $H(X_1, Y_1) = H(X_1, Y_1, C)$ and $H(X_1, X_2) = H(X_1, X_2, C)$.
Hint: Use the chain rule.

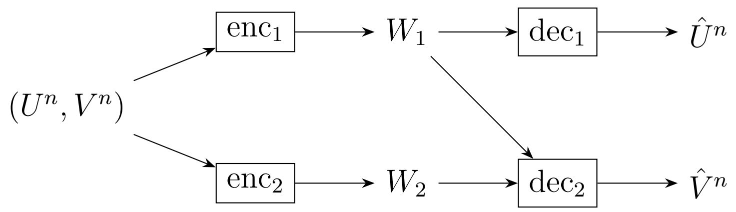
(d) (4 points) Show that $H(X_1, Y_1) + H(X_2, Y_2) = H(X_1, X_2) + H(Y_1, Y_2)$.
Hint: Use the chain rule with (c) and the conditional independence of X_1, Y_1, X_2, Y_2 given C.

(e) (4 points) Show that $H(X_1, X_2) \leq \log\left(\sum_{k=1}^n a_k^2\right)$, and then show that

$$\left(\sum_{k=1}^n a_k b_k\right)^2 \leq \left(\sum_{k=1}^n a_k^2\right)\left(\sum_{k=1}^n b_k^2\right).$$

Problem 4. (16 points)

Suppose $(U_1, V_1), (U_2, V_2), \dots$ are i.i.d. pairs with distribution p_{UV} on $\mathcal{U} \times \mathcal{V}$. An multiple descriptions system is a pair of encoder functions $\text{enc}_1 : (U^n, V^n) \mapsto W_1 \in \{1, \dots, M_1\}$ and $\text{enc}_2 : (U^n, V^n) \mapsto W_2 \in \{1, \dots, M_2\}$ with a pair of decoding functions $\text{dec}_1 : W_1 \mapsto \hat{U}^n$ and $\text{dec}_2 : (W_1, W_2) \mapsto \hat{V}^n$. In other words the encoder gives two descriptions W_1 and W_2; from W_1 we recover U^n, and from the pair (W_1, W_2) we recover the V^n, as shown in the figure. We define $p_e = \Pr\big((\hat{U}^n, \hat{V}^n) \neq (U^n, V^n)\big)$ as the error probability and, $R_1 = \frac{1}{n} \log M_1$, $R_2 = \frac{1}{n} \log M_2$ as the rates of the two descriptions.



(a) (2 points) Show that $R_1 \geq H(U) - p_e \log |\mathcal{U}| - \frac{1}{n} h_2(p_e)$, where $h_2(x) = -x \log x - (1-x)\log(1-x)$ is the binary entropy function.
Hint: $nR_1 \geq H(W_1) \geq I(U^n; W_1)$; and Fano's inequality upper bounds $H(U^n|W_1)$.

(b) (2 points) Show that $R_1 + R_2 \geq H(UV) - p_e \log |\mathcal{U}||\mathcal{V}| - \frac{1}{n} h_2(p_e)$.
Hint: Similar to (a).

Let $\mathcal{C}$ be the set of (r_1, r_2) pairs for which for any $\epsilon > 0$ there is $\text{enc}_1, \text{enc}_2, \text{dec}_1, \text{dec}_2$ with $p_e < \epsilon$, $R_1 < r_1 + \epsilon$, and $R_2 < r_2 + \epsilon$.

(c) (2 points) Show that $\mathcal{C}$ is included in the region

$$\mathcal{R} = \{(r_1, r_2) : r_1 \geq H(U), r_2 \geq 0, r_1 + r_2 \geq H(UV)\}.$$

[For parts (d) and (e), assume that the Huffman code described has expected length equal to the entropy exactly.]

(d) (2 points) Suppose we design a Huffman code c for the pair (U, V). Let W be the concatenation of $c(U_1, V_1), \dots, c(U_n, V_n)$. Show that for $r > H(UV)$, $\lim_{n \to \infty} \Pr(\text{length}(W) > nr) = 0$, and conclude that $(r_1 = H(UV), r_2 = 0)$ belongs to $\mathcal{C}$.
Hint: Use the law of large numbers with $\frac{1}{n}\text{length}(W) = \frac{1}{n}\sum_{i=1}^n \text{length}(c(U_i, V_i))$.

(e) (4 points) Show that $(r_1 = H(U), r_2 = H(V|U))$ belongs to $\mathcal{C}$.
Hint: Follow a similar logic to (d).

(f) (4 points) Show that the region $\mathcal{R}$ is included in $\mathcal{C}$ (and thus, because of (c), $\mathcal{R} = \mathcal{C}$).
Hint: In (d) and (e) you have shown that the extreme points of $\mathcal{R}$ are in $\mathcal{C}$, now use time-sharing. You may find it useful to make a sketch of $\mathcal{R}$ and the two points in (d) and (e).