ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 17 — Information Theory and Coding

Homework 7 — Nov. 4, 2019

PROBLEM 1. A source produces independent, equally probable symbols from an alphabet (a_1, a_2) at a rate of one symbol every 3 seconds. These symbols are transmitted over a binary symmetric channel which is used once each second by encoding the source symbol a_1 as 000 and the source symbol a_2 as 111. If in the corresponding 3 second interval of the channel output, any of the sequences 000,001,010,100 is received, a_1 is decoded; otherwise, a_2 is decoded. Let $\epsilon < 1/2$ be the channel crossover probability.

(a) For each possible received 3-bit sequence in the interval corresponding to a given source letter, find the probability that a_1 came out of the source given that received sequence.

(b) Using part (a), show that the above decoding rule minimizes the probability of an incorrect decision.

(c) Find the probability of an incorrect decision (using part (a) is not the easy way here).

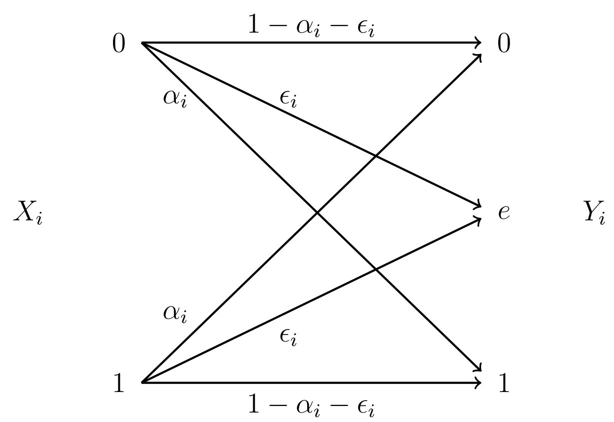
(d) If the source is slowed down to produce one letter every $2n + 1$ seconds, a_1 being encoded by $2n + 1$ 0's and a_2 being encoded by $2n + 1$ 1's. What decision rule minimizes the probability of error at the decoder? Find the probability of error as $n \to \infty$.

PROBLEM 2. Show that a cascade of n identical binary symmetric channels,

$$X_0 \to \boxed{\text{BSC \#1}} \to X_1 \to \cdots \to X_{n-1} \to \boxed{\text{BSC \#n}} \to X_n$$

each with raw error probability p, is equivalent to a single BSC with error probability $\frac{1}{2}\big(1 - (1 - 2p)^n\big)$ and hence that $\lim_{n\to\infty} I(X_0; X_n) = 0$ if $p \neq 0, 1$. Thus, if no processing is allowed at the intermediate terminals, the capacity of the cascade tends to zero.

PROBLEM 3. Consider the following symmetric channel with binary input that maps to a ternary output. (A channel that may either flip or erase the transmitted symbol.)



In other words,

$$p_i(y_i|0) = \begin{cases} 1-\alpha_i-\epsilon_i, & y_i = 0 \\ \epsilon_i, & y_i = e \\ \alpha_i, & y_i = 1 \end{cases} \qquad \alpha_i, \epsilon_i \in [0,1], \quad \alpha_i + \epsilon_i \leq 1$$

and vice versa for $p_i(y_i|1)$. Also, Y_i's are independent of each other given X_i's. (i.e. $p(y_1^n|x_1^n) = \prod_{i=1}^n p_i(y_i|x_i)$ for any $n \geq 1$).

(a) Suppose the channel is not time varying, that is $\alpha_i = \alpha$ and $\epsilon_i = \epsilon$. Find the capacity $C = \max_{p(x)} I(X;Y)$

(b) What are the special cases when $\alpha = 0$, $\epsilon \neq 0$ and $\alpha \neq 0$, $\epsilon = 0$? What happens when $\alpha + \epsilon = 1$?

(c) Now, suppose that the channel is time varying, that is, for each channel use α_i's and ϵ_i's differ. Find $\max_{p(x_1^n)} I(X_1^n; Y_1^n)$.

Problem 4.

In the lectures, we have seen that Shannon's proof of the achievability bound relies on probabilistic methods, i.e., random coding in this case. In this problem, we will derive a similar achievability bound based on a greedy construction of the codebook.

Let $\mathcal{X}$, $\mathcal{Y}$ be the discrete input and output alphabets with $\mathcal{X}$ being endowed with distribution p_X. Let $W(y|x)$ be the channel transition probabilities which describe the discrete memoryless channel (DMC) $\mathcal{W}$. Let M be size of the codebook.

As usual, there is a decoding function $d: \mathcal{Y} \to [M]$ such that $d(y) = m$, $\forall y \in D_m$, $\forall m$, where $\{D_m\}_{m=1}^M$ are the disjoint decoding regions for messages in $[M]$. Our objective is to construct the codebook and the decoding regions associated with each codeword.

Now for every $x \in \mathcal{X}$, and for $\gamma > 0$, define preliminary decoding regions as follows:

$$E_x = \{y \in \mathcal{Y} : W(y|x) \geq \gamma W(y)\}$$

where $W(y) = \sum_x W(y|x)p(x)$. (Note that these regions might overlap, so these are not the actual decoding regions.)

(a) Fix an $\epsilon > 0$ satisfying $\Pr\big(W(Y|X) < \gamma W(Y)\big) \leq \epsilon$. Show that there exists a $c \in \mathcal{X}$ such that $\Pr\big(E_c|X = c\big) \geq 1 - \epsilon$.

Now, we devise the following algorithm.

1. Add c_1 to the codebook such that $\Pr\big(E_{c_1}|X = c_1\big) \geq 1-\epsilon$, set $D_1 = E_{c_1}$.
2. Add c_2 to the codebook such that $\Pr\big(E_{c_2} \setminus D_1|X = c_2\big) \geq 1-\epsilon$, set $D_2 = E_{c_2} \setminus D_1$.

   ...
3. Add c_M to the codebook such that $\Pr\big(E_{c_M} \setminus \bigcup_{j=1}^{M-1} D_j|X = c_M\big) \geq 1-\epsilon$, set $D_M = E_{c_M} \setminus \bigcup_{j=1}^{M-1} D_j$.
4. Terminate the algorithm if there is no other $c_{m'} \in \mathcal{X}\setminus\{c_1, \ldots, c_M\}$ satisfying $\Pr\big(E_{c_{m'}} \setminus \bigcup_{j=1}^{M} D_j|X = c_{m'}\big) \geq 1-\epsilon$.

Observe that the algorithm assigns disjoint decoding regions $\{D_m\}_{m=1}^M$ for each codeword. Moreover, for these decoding regions the maximal error probability over the generated set of codewords is ϵ.

(b) Suppose the algorithm terminates when size of the codebook is M. Show that

$$\Pr\big(\{x, y : W(y|x) \geq \gamma W(y)\} \setminus \{x, y : \; y \in \cup_{j=1}^M D_j\}\big) < 1 - \epsilon.$$

Furthermore, show that

$$\Pr\big(W(Y|X) \geq \gamma W(Y)\big) - \sum_{m=1}^{M} \Pr\big(\{y \in D_m\}\big) < 1 - \epsilon.$$

(c) Show that

$$\epsilon \leq \Pr\big(W(Y|X) < \gamma W(Y)\big) + \frac{M}{\gamma}.$$

Now, consider i.i.d random variables $X_1, \ldots, X_n$, each distributed with p_X and fed into $\mathcal{W}$. (Recall that in this case $W(y_1^n|x_1^n) = \prod_{i=1}^n W(y_i|x_i)$). Set $M = 2^{nR}$.

(d) Fix an $\epsilon_n > 0$ satisfying $\Pr\big(W(Y_1^n|X_1^n) < \gamma W(Y_1^n)\big) \leq \epsilon_n$ and for some $\delta > 0$, show that

$$\epsilon_n \leq \Pr\big(i(X_1^n; Y_1^n) < n(R + \delta)\big) + 2^{-n\delta}$$

where $i(x; y) = \log \frac{W(y|x)}{W(y)}$.

(e) For any selection of $R + \delta < I(X; Y)$, show that the maximal error probability ϵ_n can be made arbitrarily small as n tends to infinity.