ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 20 Information Theory and Coding
Homework 9 November 22, 2011

This is a graded homework due to 30 November 2011, 5 pm, INR 036.

PROBLEM 1.

a) Let x^* be the most probable letter of a finite source $\mathcal{X}$, i.e. $P(x^*) \geq P(x)$, for all $x \in \mathcal{X}$. Show that
$$H(X) \geq \log(\frac{1}{P(x^*)}).$$

b) [Fano's Inequality] Assume that $\mathcal{X}$ generates a letter and we want to estimate the outcome of $\mathcal{X}$ by observing random variable Y which is related to X by the conditional distribution $p(y|x)$. From Y, we calculate a function $g(Y) = \hat{X}$, where $\hat{X}$ is an estimate of X. Let P_e be the error probability of estimation defined as $P_e = P\{\hat{X} \neq X\}$. Prove that
$$H(X \mid Y) \leq H(P_e) + P_e \log(|\mathcal{X}| - 1),$$
where $|\mathcal{X}|$ denotes the number of letters in the alphabet $\mathcal{X}$.

c) [Fano's Inverse Inequality] Assume that we use a *Maximum A Posteriori* estimator, i.e. for an observation y,
$$\hat{x} = g(y) = \arg\max_{x \in \mathcal{X}} p(x|y).$$
Prove that
$$P_e \leq 1 - 2^{-H(X|Y)}.$$
Hint: use part (a) and note that $\sum_i p_i 2^{-u_i} \geq 2^{-\sum_i p_i u_i}$.

PROBLEM 2. Consider the following channel. $\mathcal{C}_1$ is a Z-channel with error probability ϵ, $\mathcal{C}_2$ is a BSC with error probability δ. Let us denote the transition matrices of $\mathcal{C}_1$ and $\mathcal{C}_2$ $p_1(y^1|x)$ and $p_2(y^2|x)$ respectively. The transition matrix of channel $\mathcal{C}$ is $p_z(y|x)$, where z is determined by a random variable Z on $\{1, 2\}$. In other words, the output Y of channel $\mathcal{C}$ is either the output Y^1 of channel $\mathcal{C}_1$ or the output Y^2 of channel $\mathcal{C}_2$ depending on Z. Z is selected independently of the channel input and can be observed by the receiver.

(a) Let $\Pr\{Z = 1\} = p$. Show that
$$I(X; YZ) = pI(X; Y^1) + (1 - p)I(X; Y^2)$$

(b) Let C_1, C_2, C be the capacities of channels $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}$ respectively. Note that $C = \max_{p(x)} I(X; YZ)$. Show that
$$C \leq pC_1 + (1 - p)C_2$$
What is the condition for equality (in terms of ϵ, δ, p)?

(c) Now, consider such a channel over n uses. Then, $p(y_1, \ldots, y_n|x_1, \ldots, x_n) = \prod_{i=1}^n p_{z_i}(y_i|x_i)$. Assume the sequence $Z_1, \ldots, Z_n$ is i.i.d. and known in advance by both the encoder and the decoder. Show that the rate $pC_1 + (1 - p)C_2$ is achievable.

PROBLEM 3. We define the set of conditionally typical sequences as

$$A^{\epsilon,n}_{P_{Y|X}}(x_1^n) = \{y_1^n : (x_1^n, y_1^n) \in A^{\epsilon,n}_{P_{XY}}\}$$

where $A^{\epsilon,n}_{P_{XY}}$ is the set of ϵ jointly typical sequences of length n. Show that

$$|A^{\epsilon,n}_{P_{Y|X}}(x_1^n)| \leq 2^{n(1+\epsilon)H(Y|X)}.$$

PROBLEM 4. Let $P(y|x)$ be the transition probability of a binary input discrete memoryless channel with an arbitrary output alphabet $\mathcal{Y}$. We define a quantity

$$Z(P) = \sum_{y \in \mathcal{Y}} \sqrt{P(y|0)P(y|1)}.$$

1. Assume the channel is used only once to transmit an input, and the received channel output is decoded using a maximum-likelihood decoder. Show that $Z(P)$ is an upper bound to the resulting average error probability.

2. Show that $Z(P)$ is a convex function of the channel transition probabilities.

   Hint: You can start by showing that $Z(P) = \frac{1}{2} \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \{0,1\}} \sqrt{P(y|x)} \right)^2 - 1.$