ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 23 Information Theory and Coding
Homework 10 (Graded - Due on Dec. 09, 2013 - 3 PM) Nov. 26, 2013

PROBLEM 1. Let P_1 and P_2 be two channels of input alphabet $\mathcal{X}_1$ and $\mathcal{X}_2$ and of output alphabet $\mathcal{Y}_1$ and $\mathcal{Y}_2$ respectively. Consider a communication scheme where the transmitter chooses the channel (P_1 or P_2) to be used and where the receiver knows which channel were used. This scheme can be formalized by the channel P of input alphabet $\mathcal{X} = (\mathcal{X}_1 \times \{1\}) \cup (\mathcal{X}_2 \times \{2\})$ and of output alphabet $\mathcal{Y} = (\mathcal{Y}_1 \times \{1\}) \cup (\mathcal{Y}_2 \times \{2\})$, which is defined as follows:

$$P(y, k'|x, k) = \begin{cases} P_k(y|x) & \text{if } k' = k, \\ 0 & \text{otherwise.} \end{cases}$$

Let $X = (X_k, K)$ be a random variable in $\mathcal{X}$ which will be the input distribution to the channel P, and let $Y = (Y_k, K) \in \mathcal{Y}$ be the output distribution. Define X_1 as being the random variable in $\mathcal{X}_1$ obtained by conditioning X_k on $K = 1$. Similarly define X_2, Y_1 and Y_2. Let α be the probability that $K = 1$.

(a) Show that $I(X; Y) = h_2(\alpha) + \alpha I(X_1; Y_1) + (1 - \alpha) I(X_2; Y_2)$.

(b) What is the input distribution X that achieves the capacity of P?

(c) Show that the capacity C of P satisfies $2^C = 2^{C_1} + 2^{C_2}$, where C_1 and C_2 are the capacities of P_1 and P_2 respectively.

PROBLEM 2. Let $P(y|x)$ be a channel of input alphabet $\mathcal{X}$ and of output alphabet $\mathcal{X}$, and let $p(x)$ be a distribution on $\mathcal{X}$. Let $r(x|y)$ be a conditional distribution on $\mathcal{X}$ given $\mathcal{Y}$, i.e., for each $x \in \mathcal{X}$ and each $y \in \mathcal{Y}$, $r(x|y) \geq 0$ and $\sum_{x' \in \mathcal{X}} r(x'|y) = 1$. Define the functional $F(p, r)$ as follows:

$$F(p, r) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x) P(y|x) \log_2 \frac{r(x|y)}{p(x)}.$$

Now for each input distribution p on $\mathcal{X}$, define the conditional distribution r_p as $r_p(x|y) = \dfrac{p(x)P(y|x)}{\sum_{x' \in \mathcal{X}} p(x')P(y|x')}$. I.e., r_p is the "true" conditional distribution of $\mathcal{X}$ given $\mathcal{Y}$ when p is the input distribution.

(a) Use the positivity of divergence to show that for all conditional distributions r we have $F(p, r) \leq F(p, r_p) = I(X; Y)$, and deduce that $I(X; Y) = \max_r F(p, r)$.

(b) Show that $F(p, r)$ is strictly concave in both p and r.

The fact that the capacity C is equal to $\max_p \max_r F(p, r)$ suggests the following algorithm to compute the capacity of the channel P:

1. Set p_0 to be uniform in $\mathcal{X}$, and set $k = 0$.
2. Set $r_k = \operatorname*{argmax}_r F(p_k, r) = r_{p_k}$.
3. Set $p_{k+1} = \operatorname*{argmax}_p F(p, r_k)$.
4. Set $k = k + 1$.
5. Go to step 2.

(c) Use the Kuhn-Tucker conditions to show that $p_{k+1}(x) = \dfrac{\alpha_k(x)}{\sum_{x' \in \mathcal{X}} \alpha_k(x')}$, where

$$\log_2 \alpha_k(x) = \sum_{y \in \mathcal{Y}} P(y|x) \log_2 r_k(y|x).$$

This shows how to do step 3 of the algorithm.

(d) Show that $C \geq F(p_{k+1}, r_k) = \log_2 \sum_{x \in \mathcal{X}} \alpha_k(x)$.

(e) Show that $\log_2 \dfrac{\alpha_k(x)}{p_k(x)} = \sum_{y \in \mathcal{Y}} P(y|x) \log_2 \dfrac{P(y|x)}{\sum_{x' \in \mathcal{X}} P(y|x') p_k(x')}$.

(f) Let p^* be the input distribution that achieves the capacity C of the channel P. Use the result of homework 8 problem 4 to show that

$$C \leq \sum_x p^*(x) \log_2 \frac{\alpha_k(x)}{p_k(x)}.$$

(g) Show that

$$C - F(p_{k+1}, r_k) \leq \sum_{x \in \mathcal{X}} p^*(x) \log_2 \frac{p_{k+1}(x)}{p_k(x)} \leq \max_{x \in \mathcal{X}} \log_2 \frac{p_{k+1}(x)}{p_k(x)}.$$

This upper bound provides us with a stopping condition for the algorithm. I.e., we can run the algorithm until $\max_{x \in \mathcal{X}} \log_2 \dfrac{p_{k+1}(x)}{p_k(x)} \leq \epsilon$, where ϵ is some desired accuracy.

(h) Show that

$$\sum_{k=0}^{n} (C - F(p_{k+1}, r_k)) \leq \sum_{x \in \mathcal{X}} p^*(x) \log_2 \frac{p_{n+1}(x)}{p_0(x)} \leq \log |\mathcal{X}|.$$

Hint: p_0 was chosen to be uniform.

(i) Deduce that the sequence $F(p_{k+1}, r_k)$ converges to C and that the stopping condition $\max_{x \in \mathcal{X}} \log_2 \dfrac{p_{k+1}(x)}{p_k(x)} \leq \epsilon$ is guaranteed to be met eventually.