ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 31 Information Theory and Coding
Homework 13 Dec. 19, 2017

Problem 1. Suppose U is $\{0,1\}$ valued with $\mathbb{P}(U=0)=\mathbb{P}(U=1)=1/2$. Suppose we have a distortion measure d given by $d(u,v)=\begin{cases}0 & \text{if } u=v,\\ 1 & \text{if } (u,v)=(1,0),\\ \infty & \text{if } (u,v)=(0,1)\end{cases}$. I.e., we never want to represent a 0 with a 1. Find $R(D)$.

Problem 2. Suppose $\mathcal{U}=\mathcal{V}$ are additive groups with group operation $\oplus$. (E.g., $\mathcal{U}=\mathcal{V}=\{0,\dots,K-1\}$, with modulo K addition.) Suppose the distortion measure $d(u,v)$ depends only on the difference between u and v and is given by $g(u-v)$. Let $\phi(D)$ denote $\max H(Z) : E[g(Z)] <= D$.

a) Show that $\phi(D)$ is convex.

b) Let (U,V) be such that $E[d(U,V)] <= D$. Show that $I(U;V) >= H(U)-\phi(D)$ by justifying

$$I(U;V)=H(U)-H(U|V)=H(U)-H(U-V|V)>=H(U)-H(U-V)>=H(U)-\phi(D).$$

c) Show that $R(D) >= H(U)-\phi(D)$.

d) Assume now that U is uniform on $\mathcal{U}$. Show that $R(D)=H(U)-\phi(D)$.

Problem 3. Suppose $\mathcal{U}=\mathcal{V}=\mathbb{R}$, the set of real numbers, and $d(u,v)=(u-v)^2$. Show that for any U with variance σ^2, $R(D)$ satisfies

$$H(X)-(1/2)\log(2\pi\, eD) <= R(D) <= (1/2)\log(\sigma^2/D).$$

Problem 4. Consider a two-way communication system where two parties communicate via a *common* output they both can observe and influence. Denote the common output by Y, and the signals emitted by the two parties by x_1 and x_2 respectively. Let $p(y|x_1,x_2)$ model the memoryless channel through which the two parties influence the output.

We will consider feedback-free block codes, i.e., we will use encoding and decoding functions of the form

$$\begin{aligned}&\mathrm{enc}_1\colon \{1,\dots,2^{nR_1}\}\to\mathcal{X}_1^n && \mathrm{dec}_1\colon \mathcal{Y}^n\times\{1,\dots,2^{nR_1}\}\to\{1,\dots,2^{nR_2}\}\\ &\mathrm{enc}_2\colon \{1,\dots,2^{nR_2}\}\to\mathcal{X}_2^n && \mathrm{dec}_2\colon \mathcal{Y}^n\times\{1,\dots,2^{nR_2}\}\to\{1,\dots,2^{nR_1}\}\end{aligned}$$

with which the parties encode their own message and decode the other party's messages. (Note that when a party is decoding the other party's message, it can make use of the knowledge of its own message).

We will say that the rate pair (R_1,R_2) is achievable, if for any $\epsilon>0$, there exist encoders and decoders with the above form for which the average error probability is less than ϵ.

Consider the following 'random coding' method to construct the encoders:

(i) Choose probability distributions p_j on $\mathcal{X}_j$, $j=1,2$.

(ii) Choose $\{\mathrm{enc}_1(m_1)_i : m_1=1,\dots,2^{nR_1},\ i=1,\dots,n\}$ i.i.d., each having distribution as p_1. Similarly, choose $\{\mathrm{enc}_2(m_2)_i : m_2=1,\dots,2^{nR_2},\ i=1,\dots,n\}$ i.i.d., each having distribution as p_2, independently of the choices for enc_1.

For the decoders we will use typicality decoders:

(i) Set $p(x_1, x_2, y) = p_1(x_1)p_2(x_2)p(y|x_1, x_2)$. Choose a small $\epsilon > 0$ and consider the set T of ϵ-typical (x_1^n, x_2^n, y^n)'s with respect to p.

(ii) For decoder 1: given y^n and the correct m_1, dec_1 will declare $\hat{m}_2$ if it is the unique m_2 for which $(\text{enc}_1(m_1), \text{enc}_2(m_2), y^n) \in T$. If there is no such m_2, dec_1 outputs 0. (Similar description applies to Decoder 2.)

(a) Given that m_1 and m_2 are the transmitted messages, show that $(\text{enc}_1(m_1), \text{enc}_2(m_2), Y^n) \in T$ with high probability.

(b) Given that m_1 and m_2 are the transmitted messages, and $\tilde{m}_1 \neq m_1$ what is the probability distribution of $(\text{enc}_1(\tilde{m}_1), \text{enc}(m_2), Y^n)$?

(c) Under the assumptions in (b) show that the

$$\Pr\{(\text{enc}_1(\tilde{m}_1), \text{enc}_2(m_2), Y^n) \in T\} \doteq 2^{-nI(X_1;X_2Y)}.$$

(d) Show that all rate pairs satisfying

$$R_1 \leq I(X_1; YX_2), \quad R_2 \leq I(X_2; YX_1)$$

for some $p(x_1, x_2) = p(x_1)p(x_2)$ are achievable.

(e) For the case when X_1, X_2, Y are all binary and Y is the product of X_1 and X_2, show that the achievable region is strictly larger than what we can obtain by 'half duplex communication' (i.e., the set of rates that satisfy $R_1 + R_2 \leq 1$.)

Problem 5. In class, when proving the 'good news'part of the rate distortion theorem we stated that for any given u^n that is typical with respect to p_U, when V^n has i.i.d. p_V components, the probability that (u^n, V^n) is typical with respect to p_{UV} is approximately $2^{-nI(U;V)}$, but gave a heuristic argument for it. In this problem we will give a proof.

To that end fix p_{UV}, and suppose $u^n \in T(p_U, n, \delta)$. For each $u \in \mathcal{U}$, let $J(u) = \{1 \leq i \leq n : u_i = u\}$ be the indices for which u_i equals u. Suppose V^n has i.i.d. components each distributed according to p_V.

(a) Let $V(u) = (V_i : i \in J(u))$ denote the subvector of V^n by considering only the indices in $J(u)$. Note that $V(u)$ is of dimension $n(u) = |J(u)|$ with i.i.d. components each distributed according to p_V. What is the probability that $V(u)$ belongs to $T(p_{V|U=u}, n(u), \delta)$, expressed in the form $2^{-n(u)(F(\dots)+O(\delta))}$?

(b) Using (a), show that the probability that $V(u) \in T(p_{V|U=u}, n(u), \delta)$ for every $u \in \mathcal{U}$ equals $2^{-n(F+O(\delta))}$ where

$$F = \sum_u \frac{n(u)}{n} D(p_{V|U=u} \| p_V).$$

(c) Using the fact that u^n is in $T(n, p_U, \delta)$, show that F in (b) equals

$$\sum_u p_U(u) \sum_v p_{V|U=u}(v|u) \log \frac{p_{V|U}(v|u)}{p_V(v)} + O(\delta),$$

and conclude that the probability we found in (b) equals $2^{-n(I(U;V)+O(\delta))}$.

(d) Show that when $u^n \in T(n, p_U, \delta)$ and $V(u) \in T(n(u), p_{V|U=u}, \delta)$ for every $u \in \mathcal{U}$, we will necessarily have (u^n, V^n) belonging to $T(n, p_{UV}, 2\delta + \delta^2)$.

(e) Conclude that for any $1 \geq \delta' \geq 3\delta$, for any $u^n \in T(n, p_U, \delta)$, with V^n i.i.d. p_V, the probability that $(u^n, V^n) \in T(n, p_{UV}, \delta')$ is at least $2^{-n(I(U;V)+O(\delta))}$.