Information Theory II

Spring Semester 2017

Prof. Dr. A. Lapidoth / Prof. Dr. S. Moser

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Model Answers to Exercise 10 of May 4, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it2.html

Problem 1 ***Feedback Capacity of a Channel with Noncausal State Information at the Transmitter***

We wish to show that if a rate R is achievable on a channel with feedback and noncausal state information at the transmitter, then

$$\mathsf{R} \leq \max_{U \multimap\!\!- (X,S) \multimap\!\!- Y} \bigl\{ I(U;Y) - I(U;S) \bigr\}.$$

The derivation of the converse in the Gel'fand–Pinsker theorem still holds with feedback: we have

$$n\mathsf{R} \leq \sum_{i=1}^{n} \bigl(I(U_i;Y_i) - I(U_i;S_i) \bigr) + n\epsilon_n,$$

where $U_i = (M, Y^{i-1}, S_{i+1}^n)$. Because the Markov chain $U_i \multimap\!\!- (X_i,S_i) \multimap\!\!- Y_i$ is also valid with feedback, the converse in the Gel'fand–Pinsker theorem continues to hold with feedback. We conclude that feedback does not increase the channel capacity if the state is known noncausally to the transmitter.

Problem 2 ***Memory with Defects***

a) The capacity in this case can be expressed as

$$\mathsf{C} = \max_{P_T} I(T;Y),$$

where the maximum is over PMFs on the set of all mappings $t\colon \mathcal{S} \to \mathcal{X}$. For the given channel, the number of such mappings is $|\mathcal{X}|^{|\mathcal{S}|} = 8$. However, since the channel output is deterministically 0 if $S = 0$ and deterministically 1 if $S = 1$, it does not matter what input symbol is used in these cases. Thus, we only have to consider two mappings: $t_0(\text{“good”}) = 0$ and $t_1(\text{“good”}) = 1$. The channel from $T \in \{t_0, t_1\}$ to Y is a binary symmetric channel with crossover probability $\frac{p}{2}$, so the channel capacity C is $1 - H_{\mathrm{b}}\bigl(\frac{p}{2}\bigr)$ bits per channel use.

(Note that if no state information is available at the transmitter, the channel from X to Y is also a binary symmetric channel with crossover probability $\frac{p}{2}$ and capacity $1 - H_{\mathrm{b}}\bigl(\frac{p}{2}\bigr)$, so causal state information at the transmitter does not increase the capacity of this channel.

It is also possible to work with the formula

$$\mathsf{C} = \max_{P_U,\, f\colon \mathcal{U}\times\mathcal{S}\to\mathcal{X}} I(U;Y).$$

Because for all $u \in \mathcal{U}$, $P_{Y|U=u}(0) \geq P_S(0) = \frac{p}{2}$ and $P_{Y|U=u}(1) \geq P_S(1) = \frac{p}{2}$ hold, we obtain $H(Y|U=u) \geq H_b\bigl(\frac{p}{2}\bigr)$ and thus $H(Y|U) \geq H_b\bigl(\frac{p}{2}\bigr)$. Consequently,

$$\mathsf{C} \leq I(U;Y) = H(Y) - H(Y|U) \leq 1 - H(Y|U) \leq 1 - H_b\Bigl(\frac{p}{2}\Bigr)$$

holds. Since any rate $\mathsf{R} < 1 - H_b\bigl(\frac{p}{2}\bigr)$ can be achieved by a scheme that ignores the state information, we conclude $\mathsf{C} = 1 - H_b\bigl(\frac{p}{2}\bigr)$.)

b) If the receiver also has state information, then the capacity is

$$\begin{aligned}
\mathsf{C}' &= \max_{P_{X|S}} I(X;Y|S) \\
&= P_S(\text{“good”}) \cdot \max_{P_{X|S=\text{“good”}}} I(X;Y|S=\text{“good”}) \\
&\quad + P_S(0) \cdot \max_{P_{X|S=0}} I(X;Y|S=0) + P_S(1) \cdot \max_{P_{X|S=1}} I(X;Y|S=1) \\
&= (1-p) \cdot 1 + \frac{p}{2} \cdot 0 + \frac{p}{2} \cdot 0 \\
&= 1 - p
\end{aligned}$$

bits per channel use. Since the receiver is free to ignore the state information, we obtain the upper bound $\mathsf{C} \leq \mathsf{C}'$. We now show that this capacity can be achieved if only the transmitter has state information. To this end, we use the formula of Gel'fand and Pinsker,

$$\mathsf{C} = \max_{P_{U|S},\, f\colon \mathcal{U}\times\mathcal{S}\to\mathcal{X}} \bigl\{ I(U;Y) - I(U;S) \bigr\},$$

where we choose $\mathcal{U} = \{0,1\}$ and $X = f(U,S) = U$. The conditional distribution of U given S is chosen as follows: when $S = 0$, U is chosen to be deterministically 0; when $S = 1$, U is chosen to be deterministically 1; when $S =$ “good”, U is drawn randomly according to $\text{Ber}\bigl(\frac{1}{2}\bigr)$. With this choice, $Y = U$ and $Y \sim \text{Ber}\bigl(\frac{1}{2}\bigr)$, so

$$I(U;Y) = H(Y) - H(Y|U) = 1 \text{ bit}.$$

On the other hand, $U \sim \text{Ber}\bigl(\frac{1}{2}\bigr)$ and

$$\begin{aligned}
I(U;S) &= H(U) - H(U|S) \\
&= 1 - P_S(0)H(U|S=0) - P_S(1)H(U|S=1) - P_S(\text{“good”})H(U|S=\text{“good”}) \\
&= 1 - 0 - 0 - (1-p)\cdot 1 \\
&= p \text{ bits}.
\end{aligned}$$

Hence, any rate smaller than

$$I(U;Y) - I(U;S) = 1 - p \quad \text{bits/channel use}$$

is achievable, and we have shown before that even when the state is revealed to the receiver, rates larger than $1-p$ are not achievable. We conclude that the capacity of this channel with noncausal state information at the transmitter is equal to $1-p$ bits per channel use.

Problem 3 — *BSC with Noncausal State Information*

If the receiver also has state information, then the capacity is

$$\mathsf{C}' = \max_{P_{X|S}} I(X;Y|S)$$

$$
\begin{aligned}
&= P_S(\text{“good”}) \cdot \max_{P_{X|S=\text{“good”}}} I(X;Y|S=\text{“good”}) \\
&\quad + P_S(0) \cdot \max_{P_{X|S=0}} I(X;Y|S=0) + P_S(1) \cdot \max_{P_{X|S=1}} I(X;Y|S=1) \\
&= (1-p) \cdot \bigl(1 - H_{\mathrm{b}}(\epsilon)\bigr) + \frac{p}{2} \cdot 0 + \frac{p}{2} \cdot 0 \\
&= (1-p)\bigl(1 - H_{\mathrm{b}}(\epsilon)\bigr),
\end{aligned}
$$

which gives us an upper bound on the Gel'fand–Pinsker capacity of this channel. We next show that this *is* the Gel'fand–Pinsker capacity. Namely, we show that any rate below $(1-p)\bigl(1-H_{\mathrm{b}}(\epsilon)\bigr)$ is achievable even if only the transmitter knows the realization of the state sequence.

Let U be a binary auxiliary random variable whose distribution conditional on s is

$$
P_{U|S}(u|s) = \begin{cases} 1-\epsilon & s = u, \\ \epsilon & s = u \oplus 1, \\ \frac{1}{2} & s = \text{“good”}, \end{cases}
$$

and let $x = u$. Using the Gel'fand–Pinsker theorem we know that any rate below $I(U;Y) - I(U;S)$ is achievable. We first compute $I(U;S)$. Note that U is uniform binary, so

$$
\begin{aligned}
I(U;S) &= H(U) - H(U|S) \\
&= 1 - P_S(\text{“good”})\, H(U|S=\text{“good”}) - P_S(0)\, H(U|S=0) - P_S(1)\, H(U|S=1) \\
&= 1 - (1-p) \cdot 1 - \frac{p}{2} H_{\mathrm{b}}(\epsilon) - \frac{p}{2} H_{\mathrm{b}}(\epsilon) \\
&= p\bigl(1 - H_{\mathrm{b}}(\epsilon)\bigr) \text{ bits.}
\end{aligned}
$$

To determine $I(U;Y)$, we compute the joint distribution of U and Y:

$$
\begin{aligned}
P_{UY}(u,y) &= \sum_s P_{SUY}(s,u,y) \\
&= \sum_s P_S(s) P_{U|S}(u|s) P_{Y|US}(y|u,s),
\end{aligned}
$$

which leads to

$$
\begin{aligned}
P_{UY}(0,1) = P_{UY}(1,0) &= \frac{\epsilon}{2}, \\
P_{UY}(0,0) = P_{UY}(1,1) &= \frac{1-\epsilon}{2}.
\end{aligned}
$$

From this we obtain $I(U;Y) = 1 - H_{\mathrm{b}}(\epsilon)$. Since

$$
I(U;Y) - I(U;S) = (1-p)\bigl(1 - H_{\mathrm{b}}(\epsilon)\bigr) \text{ bits}
$$

coincides with the upper bound, it is the Gel'fand–Pinsker capacity of this channel.