Multi-User Information Theory | November 2th, 2009

Lecture 3

Lecturer: Haim Permuter | *Scribe: Ohad Lehrer*

I. Channel Coding

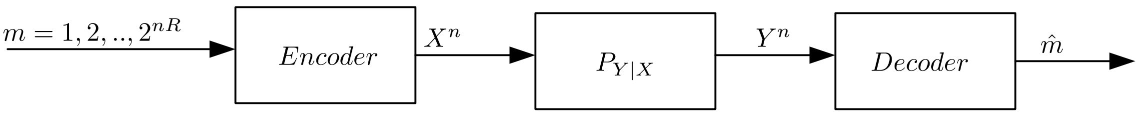

Fig. 1. Communication system

<u>Reminder.</u>

Definition 1 (Channel coding.)

An $(2^{nR}, n)$ *code* for the channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ consist of the following:

- An index set $(2^{nR}, n)$ code.
- An encoding function $f : \mathcal{M} \to X^n$.
- A decoding function $g : \mathcal{Y} \to \mathcal{M}$.

Definition 2 (Probability of error.)

The probability of a discovered message to be incorrect is defined by:

$$P_e^{(n)} = \Pr(\hat{m} \neq m) = \sum_{m=1}^{2^{nR}} P(m) \Pr(\hat{m} \neq m | M = m) = \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \Pr(\hat{m} \neq m | M = m). \tag{1}$$

Definition 3 (Achievable rate)

A Rate R is said to be *achievable*, if there exists a sequence of $(2^{nR}, n)$ codes such that $P_e^{(n)} \xrightarrow{n\to\infty} 0$

Definition 4 (Capacity.)

The *capacity* of a channel C, is the supremum over all achievable rates.

Thus, for rates less than capacity there exists a code which yields arbitrarily small probability of error.

Definition 5 (Memoryless channel.)

A channel $P(y^n||x^n)$ is *memoryless* if:

$P(y^n||x^n) = \prod_{i=1}^{n} P(y_i|x_i)$ or equivalently $P(y_i|y^{i-1}, x^i) = P(y_i|x_i)$

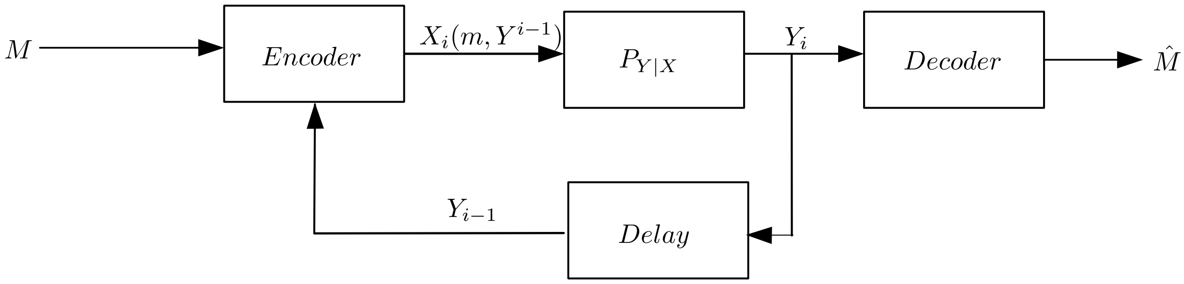

Fig. 2. Communication system with feedback

A. Channel Coding With Feedback [1]

A channel with feedback is illustrated in Figure 2.

We assume that all the received symbols Y_i are sent back with a delay to the transmitter.

Definition 6 (Channel Coding with feedback.) An $(2^{nR}, n)$ feedback code consists of:

- An encoding function $f : \mathcal{M} \times \mathcal{Y}^{i-1} \rightarrow X_i$.
- A decoding function $g : \mathcal{Y} \rightarrow \mathcal{M}$.

Example 1

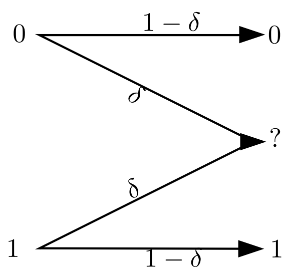

Fig. 3. Erasure Channel

$$
\begin{aligned}
C &= I(X;Y) && (2)\\
&= H(X) - H(X|Y) && (3)\\
&= H(X) - \sum_{i=1}^{3} P(y_i) H(X|Y = y_i) && (4)
\end{aligned}
$$

$$\stackrel{(a)}{=} \quad H(X) - \delta H(X) \tag{5}$$

$$= \quad 1 - \delta. \tag{6}$$

Where: (a) Follows from the fact that $H(X|y = 0) = 0,\ H(X|y = 1) = 0,\ H(X|y = '?') = 1$, and $P(y = '?') = \delta$

Let assume that we use the channel 500 usages (block length n=500), how many bits can we transmit though the channel '?'
Answer: Since δ portion of time we will receive δ, we can transmit $(1-\delta)500$ bit, and this by repeating the last bit, each time that the output is '?'. In this case, the feedback simplify the coding scheme, though is not increasing the capacity.

Theorem 1 Feedback does not increase capacity at memoryless channel.

Using Fano inequality we will show that feedback does not increase capacity.

Lemma 1 (Fano inequality[2].) For any estimator $\hat{X}$, with $P_e = \Pr(\hat{x} \neq x) = \epsilon$, we have:

$$H(P_e) + P_e \log |\mathcal{X}| \geq H(X|\hat{X}). \tag{7}$$

This inequality can be weakened to

$$H(X|\hat{X}) \leq 1 + \epsilon \log |\mathcal{X}| \,. \tag{8}$$

Proof: (*Theorem 1*)

Assumptions:

- $M \in \left\{1, 2, \ldots, 2^{nR}\right\}$ uniformly distributed.
- We have a sequence of codes $(2^{nR}, n)$ with feedback.
- The rate R is achievable.

We need to show that:

$$R \leq \max_{P_X} I(X;Y) \tag{9}$$

Fix a code $(2^{nR}, n)$ with probability of error P_e than:

$$\begin{aligned} nR &= H(M) \qquad (10) \\ &= H(M) + H(M|Y^n) - H(M|Y^n) \\ &= H(M) + H(M|Y^n) - H(M|Y^n) \\ &= I(M;Y^n) + H(M|Y^n, \hat{M}) \end{aligned}$$

$$
\begin{aligned}
&\leq && I(M;Y^n) + H(M|\hat{M}) \\
&\overset{(a)}{\leq} && I(M;Y^n) + n\epsilon_n \\
&\overset{(b)}{=} && \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, M, X^i) + n\epsilon_n \\
&\overset{(c)}{=} && \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|X_i) + n\epsilon_n \\
&\overset{(d)}{\leq} && \sum_{i=1}^{n} H(Y_i) - H(Y_i|X_i) + n\epsilon_n \\
&\leq && \sum_{i=1}^{n} \max_{P_{X_i}} I(X_i;Y_i) + n\epsilon_n \\
&= && nC^{(I)} + n\epsilon_n. \qquad (11)
\end{aligned}
$$

Where:

(a) Follows from fano inequality $H(X|Y) \leq 1 + P_e nR \triangleq n\epsilon_n$.

(b) Follows from the chain rule of mutual information.

(c) Follows from the fact that X_i is a function of Y^{i-1} and M.

(d) Conditioning reduce entropy.

If R is achievable, then $P_e^n \to 0$ which implies $\epsilon_n \to 0$ and from (11) we conclude:

$$
R \leq C^{(I)} = \max_{P_X} I(X;Y). \qquad (12)
$$

■

Theorem 2 (*Capacity of a feedback channel).* For a general channel, possibly with memory

$$
C_{\underline{Feedback}} \leq \liminf_{\max P_{X^n||Y^{n-1}}} \frac{1}{n} I(X^n \to Y^n) \qquad (13)
$$

$$
\begin{aligned}
I(M;Y^n) + n\epsilon_n &\overset{(a)}{=} \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, M, X^i(M, Y^{i-1})) + n\epsilon_n \qquad (14) \\
&\overset{(b)}{=} \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, X^i) \\
&\overset{(c)}{=} H(Y^n) - H(Y^n||X^n) + n\epsilon_n \\
&\overset{(d)}{=} I(X^n \to Y^n) + n\epsilon_n.
\end{aligned}
$$

(a) Follows from the definition and chain rule of mutual information, and by know M and Y^{i-1}, X^i is known.

(b) X^i is function of M and Y^{i-1}

(c) Is from the definition of causal conditioning entropy (lec.1, definition 4)

(d) Is from the definition of directed information (lec.1, definition 5)

II. MAC - MULTIPLE ACCESS CHANNEL [1]

Until now we assumed one user sends only one message.
A system with many senders and receivers contains many new elements in the communication problem: interference, cooperation and feedback. These are the issues that are the domain of network information theory.

This chapter is based on the review in El Gamal and Cover[3]. The two-way channel was studied by Shannon [4] in 1961. The multiple-access channel capacity region was found by Ahlswede [5] and Liao [6] and was extended to the case of the multiple-access channel with common information by Selpian and Wolf [7].

The first channel that we examine is the multiple access channel, illustrated in Figure 4, in which two transmitters send information to a common receiver.
We can see that the transmitter must contend not only with the receiver noise but with interference from each other as well.

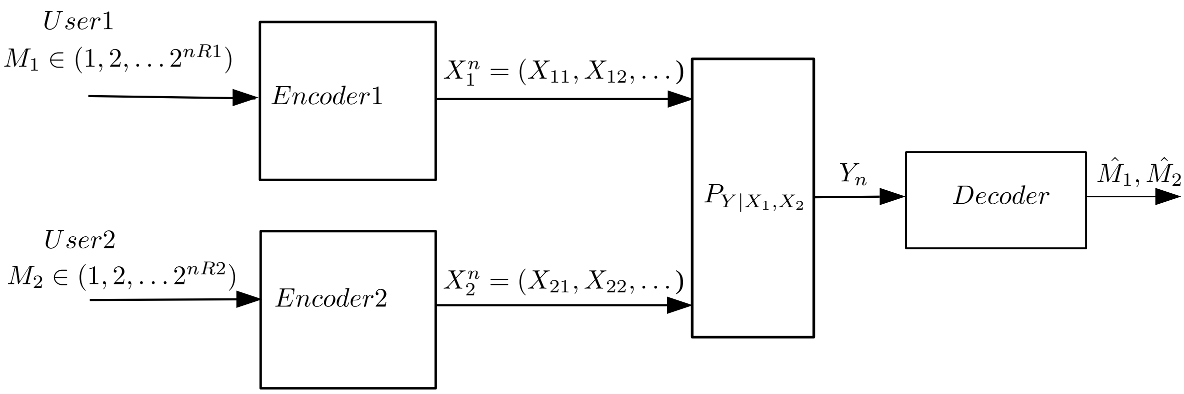

Fig. 4. MAC

Definition 7 (Discrete memoryless multiple-access channel).
A *discrete memoryless multiple-access* channel consist of three alphabets, $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{Y}$, and a probability transition matrix $P(y|x_1, x_2)$.

Definition 8 (Code for a multiple-access channel).
A $(2^{nR_1}, 2^{nR_2}, n)$ code for the multiple-access channel consists of:

- Two sets of integers $\mathcal{M}_1 = \{1, 2, \ldots, 2^{nR_1}\}$ and $\mathcal{M}_2 = \{1, 2, \ldots, 2^{nR_2}\}$ called the message sets.

- $Encoder_1 \quad f_1 : \mathcal{M}_1 \to X_1^n$.
- $Encoder_2 \quad f_2 : \mathcal{M}_2 \to X_2^n$.
- $Decoder \quad g : Y^n \to (\hat{M}_1 \times \hat{M}_2)$.
- $P_e^{(n)} = \Pr(\hat{M}_1 \neq M_1$ or $\hat{M}_2 \neq M_2)$.

Assuming that the distribution of the messages over the product set $\mathcal{M}_1 \times \mathcal{M}_2$ is uniform, we define the *average probability of error* for the $(2^{nR_1}, 2^{nR_2}, n)$ code as follows:

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2 \in \mathcal{M}_1 \times \mathcal{M}_2} \Pr\left\{g(Y^n) \neq (m_1, m_2) | (m_1, m_2) sent\right\}. \tag{15}$$

Definition 9 (Achievable pair rate). A pair rate (R_1, R_2) is *achievable* if there exists a sequence of codes $(2^{nR_1}, 2^{nR_2}, n)$ such that $P_e^{(n)} \to 0$.

Definition 10 (Capacity region). The *capacity region* is the closure of all achievable rates.

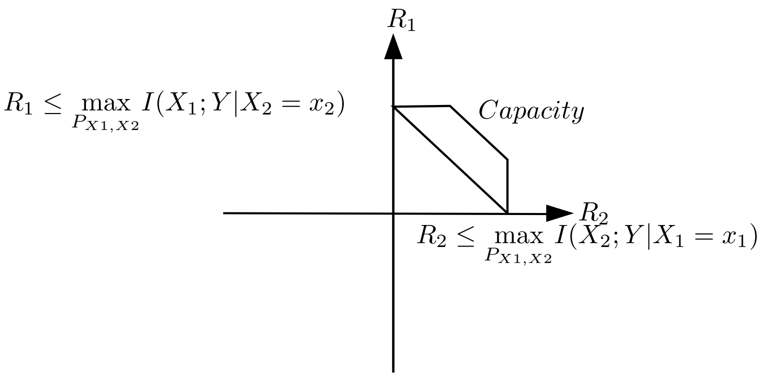

Fig. 5. Capacity of multi user channel

Theorem 3 The *capacity region* of the memoryless multiple-access channel is the closure of the convex hull of:

$$R_1 \leq I(X_1; Y | X_2) \tag{16}$$

$$R_2 \leq I(X_2; Y | X_1) \tag{17}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y) \tag{18}$$

For all $P(x_1)\ P(x_2)$

Equivalently the capacity region is:

$$\mathcal{R} = cl. \bigcup_{P(q)P(X_1|q)P(X_2|q)} \begin{cases} R_1 \leq I(X_1;Y|X_2,Q) \\ R_2 \leq I(X_2;Y|X_1,Q) \\ R_1 + R_2 \leq I(X_1,X_2;Y|Q) = I(X_1;Y) + I(X_2;Y|X_1) = I(X_2;Y) + I(X_1;Y|X_2) \end{cases}$$

Where q is time sharing argument.

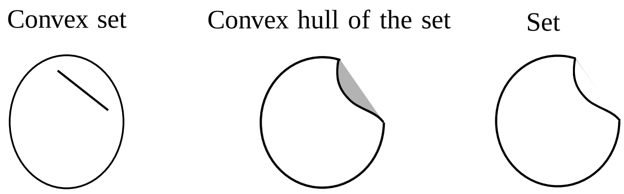

Fig. 6. Set, Convex set and Convex hull

An object is convex if it contains line segment between any two points in the set as illustrated at figure (6). Convex hull of S is the smallest convex set that contains S,in other words convex hull is the set of all convex combinations of points in S.

In Figure (6) we can see that the grey area is the convex hull of the set which is not convex without the grey fill.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New-York: Wiley, 2006.

[2] R. M. Fano. Class note for Transmission of Information, course 6.574 (Technical report). MIT, Cambridge, MA, 1952.

[3] A. El Gamal and T. M. Cover. Multiple user information theory. Proc.IEEE, 68:1466-1483, 1980.

[4] C. E. Shannon. Two-way communication channels. In Proc. 4th *Berkeley Symp. Math. Stat. Prob.*, Vol. 1, pages 611-644. University of California.

[5] R. Ahlswede. Multi-way communication channels. In Proc. 2nd Int. Symp. Inf. Theory (Tsahkadsor, Armenian S.S.R.), pages 23-52.Hungarian Academy of Sciences, Budapest, 1971.

[6] H. Liao. Multiple access channels. Ph.D. thesis, Department of Electrical Engineering, University of Hawaii, Honolulu, 1972.

[7] D. Slepian and J. K. Wolf. A coding theorem for multiple access channels with correlated sources. Bell Syst. Tech. J., 52:1037-1076, 1973.