ITCT Lecture 4.1: Channel and Channel Capacity

Prof. Ja-Ling Wu

Department of Computer Science and Information Engineering
National Taiwan University

1

1

Channel and Channel Capacity

I.The Discrete Memoryless Channel Model:

Mathematically, we can view the channel as a probabilistic function that transforms a sequence of (usually coded) input symbols, x, into a sequence of channel output symbols, y.

Because of noise and other impairments in the communication system, this transformation is typically not a one-to-one mapping from the set of input symbols, X, to the set of output symbols, Y.

Any particular $x \in X$ may have some probability, $P_{y|x}$, of being transformed to the output symbol $y \in Y$. $P_{y|x}$ is called a Forward Transition Probability.

2

2

In most communication or storage systems, the signal processing associated with modulation, transmission, reception and demodulation are designed so that the sequence of output symbols $\bar{y} = (y_0, y_1, \cdots, y_t)$ are statistically independent if the symbols in the input sequence $\bar{x} = (x_0, x_1, \cdots, x_t)$ are statistically independent.

If the output set Y consists of discrete output symbols, and if the property of statistical independence of the output sequence holds, the channel is called a Discrete Memoryless Channel (DMC).

3

For a DMC, the probability that the received symbol is y is given in terms of transition probabilities by

$$q_y = \sum_{x \in X} P_{y|x} \cdot P_x$$

The probability distribution of the output set Y, denoted by Q_Y, may be easily calculated in matrix form as

$$Q_Y = \begin{pmatrix} q_0 \\ q_1 \\ \vdots \\ q_{|Y|-1} \end{pmatrix} = \begin{pmatrix} P_{y_0|x_0} & P_{y_0|x_1} & \cdots & P_{y_0|x_{|X|-1}} \\ P_{y_1|x_0} & & & \\ \vdots & & & \\ P_{y_{|X|-1}|x_0} & & \cdots & P_{y_{|Y|-1}|x_{|X|-1}} \end{pmatrix} \begin{pmatrix} P_0 \\ P_1 \\ \vdots \\ P_{|X|-1} \end{pmatrix}$$

or more compactly, $Q_Y = P_{Y|X} P_X$.

4

Example 1.1 $X = \{0,\ 1\}$ of equally probable symbols, and let Y be a three-element set $Y = \{y_0, y_1, y_2\}$. Let the channel have transition probability matrix

$$P_{Y|X} = \begin{pmatrix} 0.8 & 0.05 \\ 0.15 & 0.15 \\ 0.05 & 0.8 \end{pmatrix}, \text{ please find } Q_Y.$$

Sol： $Q_Y = P_{Y|X} \cdot P_X = \begin{pmatrix} 0.8 & 0.05 \\ 0.15 & 0.15 \\ 0.05 & 0.8 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 0.425 \\ 0.15 \\ 0.425 \end{pmatrix}$

5

Notice that, in Example 1.1, the columns of $P_{Y|X}$ sum to unity. This means： no matter what symbol x is sent, some output symbol y must result.

In example 1.1, we have $|Y| \neq |X|$.

In many communication systems $|Y| = |X|$, and such systems are typically said to employ Hard-Decision Decoding, i.e., the demodulator makes a firm (or "hard") decision of what symbol x was probably transmitted.

In the case where $|Y| > |X|$, the demodulator is said to make Soft Decision.

6

Example 1.2. Calculate the entropy of Y for the system in example 1. Compare this with the entropy of source X.

$$\text{Sol}: H(Y) = \sum_{y \in Y} q_y \log_2(1/q_y)$$

$$= -[0.425\log_2(0.425)x2 + 0.15\log_2(0.15)]$$

$$= 1.4598$$

and $H(X) = -2 \cdot (0.5)\log_2(0.5) = 1$

$\Rightarrow \ H(Y) > H(X)$!

7

Remark：

It is possible for the output entropy to be greater than the input entropy but that the "additional" information carried in the output was not related to the information from the source. Rather, it is due to the fact that the mapping from input to output is "one-to-many" and the "extra" information in the output is due to the "randomness" of the encoding process itself.

In a similar fashion, the result in example 1.2 comes from the presence of random noise in the channel during transmission and not from the source X.

The "extra" information carried in Y is truly "useless" to us and, in fact, is harmful because it produces uncertainty about what symbols were being transmitted.

8

Can we solve this problem by using only systems which employ hard-decision decoding？

Example 1.3. Let X be the binary source of example 1.1, and let the channel have binary outputs $Y=\{0,1\}$, and transition probability matrix.

$$P_{Y|X}=\begin{pmatrix}0.98 & 0.05\\0.02 & 0.95\end{pmatrix}$$ please find $H(Y)$ and compare it with the source entropy.

Sol： $$Q_Y=P_{Y|X}\cdot P_X=\begin{pmatrix}0.98 & 0.05\\0.02 & 0.95\end{pmatrix}\begin{pmatrix}0.5\\0.5\end{pmatrix}=\begin{pmatrix}0.515\\0.485\end{pmatrix}$$

and $H(Y)=0.99935<1=H(X)$.

9

Example 1.3 shows that Y carries less information than was transmitted by the source. Where did it go？

It was lost during the transmission process.

The channel is information lossy !

So far, we have looked at two examples where the output entropy was either greater than or less than the input entropy. What we have not considered yet is what effect all this has on the ability to "Tell From Observing Y What Original Information was Transmitted". After all, the purpose of the receiver is to recover the original transmitted information. What does our observation of Y tell us about the transmitted information sequence？

10

Recall： Mutual information is defined as a measure of how much our uncertainty of one variable is reduced by knowledge of another !

$$I(X;Y)=\sum_{x\in X}\sum_{y\in Y}P_{x,y}\log_2\frac{P_{x,y}}{q_y\cdot P_x}$$

If $I(X;Y)=0$, then Y tell us nothing at all about X

$$I(X;Y)=H(X)-H(X\mid Y)\le H(X)$$

remark： the conditional entropy is a measure of how much information loss occurs in the channel.

11

Example 1.4. Calculate the mutual information for the system of example 1.1.

Sol： Since $P_{x,y}=P_{y|x}\cdot P_x$

$$I(X;Y)=\sum_{x\in X}P_x\sum_{y\in Y}P_{y|x}\log_2\left(\frac{P_{y|x}}{q_y}\right)=\sum_{x\in X}P_x I(Y;x)$$

where the “partial mutual information” $I(Y;x)$ is

$$I(Y;x)=\sum_{y\in Y}P_{y|x}\log_2\left(\frac{P_{y|x}}{q_y}\right)$$

$$I(X;Y)=2\cdot(0.5)\cdot\left[0.8\log_2\frac{0.8}{0.425}+\overset{0}{\cancel{0.15\log_2\left(\frac{0.15}{0.15}\right)}}+0.05\log_2\left(\frac{0.05}{0.425}\right)\right]$$

$$=0.57566$$

The mutual information for this system is well below the entropy of the source and so this channel has a high level of information loss. $(H(X)=1)$

12

Example 1.5. Calculate the mutual information for the system of example 1.3

Sol：

$$I(X;Y)=0.5\begin{bmatrix}0.98\log_2\left(\frac{0.98}{0.515}\right)+0.02\log_2\left(\frac{0.02}{0.485}\right)\\+0.05\log_2\left(\frac{0.05}{0.515}\right)+0.95\log_2\left(\frac{0.95}{0.485}\right)\end{bmatrix}$$

$$=0.78543<1=H(X)$$

This channel is quite lossy also.

Notice how even though $H(Y)$ was almost equal to $H(X)$ in example 1.3, the mutual information is considerably less than $H(X)$.

remark： We cannot tell how much information loss we are dealing with simply by comparing the input and output entropies.

13