Lecture 15 Communication over a Noisy Channel

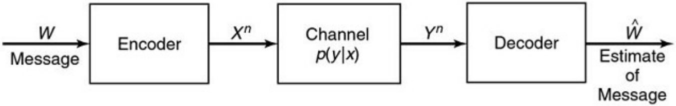
Textbook 7.1-7.4

October 21th, 2022

Outline

1. Discrete Memoryless Channel
2. Examples of channel capacity
3. Symmetric channels
4. Properties of channel capacity
5. Definitions

A discrete memoryless channel Q is characterized by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$, and a set of conditional probability distributions $p(y|x)$, one for each $x \in \mathcal{X}$.



Definition

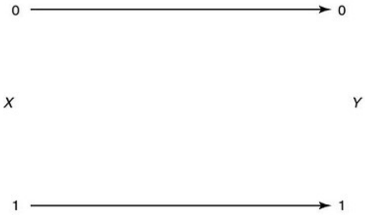
We define the channel capacity of a discrete memoryless channel as

$$C = \max_{p(x)} I(X;Y),$$

where the maximum is taken over all possible input distributions $p(x)$.

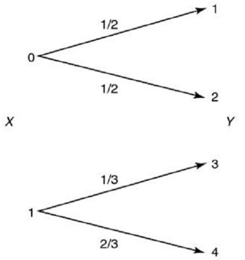
We shall soon give an operational definition of channel capacity as the highest rate in bits per channel use at which information can be sent with arbitrarily low probability of error. Shannon's second theorem establishes that the information channel capacity is equal to the operational channel capacity. Thus we drop the word information in most discussions of channel capacity.

Noiseless binary channel



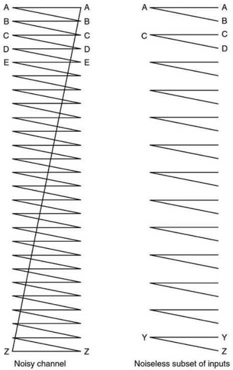
In this case, any transmitted bit is received without error. Hence, one error-free bit can be transmitted per use of the channel, and the capacity is 1 bit. We can also calculate the information capacity $C = \max I(X;Y) = 1$bit, which is achieved by using $p(x) = (\frac{1}{2}, \frac{1}{2})$.

Noisy channel with nonoverlapping outputs



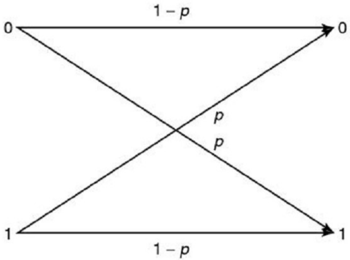
Even though the output of the channel is a random consequence of the input, the input can be determined from the output, and hence every transmitted bit can be recovered without error. The capacity of this channel is also 1 bit per transmission. We can also calculate the information capacity $C = \max I(X;Y) = 1$bits, which is achieved by using $p(x) = (\frac{1}{2}, \frac{1}{2})$.

Noisy typewritter



The input has 26 symbols and we use every alternate input symbol, we can transmit one of 13 symbols without error with each transmission. Hence, the capacity of this channel is $\log 13$ bits per transmission. One can also calculate the information capacity $C = \max I(X;Y) = \max(H(Y) - H(Y|X))$ $= \max H(Y) - 1 = \log 26 - 1 = \log 13$, achieved by using $p(x)$ distributed uniformly over all input.

Binary symmetric channel



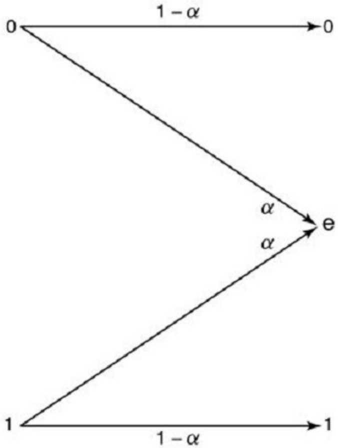
We bound the mutual information by

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) = \sum p(x) H(Y|X=x) \\ &= H(Y) - \sum p(x) H(p) \\ &= H(Y) - H(p) \\ &\leq 1 - H(p). \end{aligned}$$

Equality is achieved when the input distribution is uniform. Hence, the information capacity of a binary symmetric channel with parameter p is

$$C = 1 - H(p) \text{ bits.}$$

Binary erasure channel



We calculate the capacity of the binary erasure channel as follows:

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} (H(Y) - H(Y|X)) \\ &= \max_{p(x)} (H(Y) - H(\alpha)). \end{aligned}$$

Let E be the event $\{Y = e\}$, using the expansion

$$H(Y) = H(Y, E) = H(E) + H(Y|E),$$

and letting $\Pr(X = 1) = \pi$, we have

$$H(Y) = H((1-\pi)(1-\alpha), \alpha, \pi(1-\alpha)) = H(\alpha) + (1-\alpha)H(\pi).$$

Hence

$$C = \max_{p(x)} (H(Y) - H(\alpha)) = \max_{\pi} (1-\alpha)H(\pi) = 1 - \alpha.$$

where capacity is achieved by $\pi = \frac{1}{2}$.

Outline

1. Discrete Memoryless Channel
2. Examples of channel capacity
3. Symmetric channels
4. Properties of channel capacity
5. Definitions

Definition

A channel is said to be symmetric if the rows of the channel transition matrix $p(y|x)$ are permutations of each other and the columns are permutations of each other. A channel is said to be weak symmetric if the every row of the channel transition matrix $p(y|x)$ is a permutation of each other, and all column sums $\sum_x p(y|x)$ are equal.

Letting **r** be a row of the transition matrix, we have

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) - H(\mathbf{r}) \\ &\leq \log|\mathcal{Y}| - H(\mathbf{r}) \end{aligned}$$

with equality if the output distribution is uniform. But $p(x) = \frac{1}{|\mathcal{X}|}$ achieves a uniform distribution on Y, as seen from

$$p(y) = \sum_{x \in \mathcal{X}} p(y|x)p(x) = \frac{1}{\mathcal{X}} \sum_{x \in \mathcal{X}} p(y|x) = c\frac{1}{|\mathcal{X}|} = \frac{1}{|\mathcal{Y}|},$$

where c is the sum of the entries in one column of the probability transition matrix.

Theorem

For a weakly symmetric channel,

$$C = \log |\mathcal{Y}| - H(\textit{row of transition matrix})$$

Outline

1. Discrete Memoryless Channel
2. Examples of channel capacity
3. Symmetric channels
4. Properties of channel capacity
5. Definitions

1. $C \geq 0$ since $I(X;Y) \geq 0$.
2. $C \leq \log|\mathcal{X}|$ since $C = \max I(X;Y) \leq \max h(X) = \log|\mathcal{X}|$.
3. $C \leq \log|\mathcal{Y}|$ for the same reason.
4. $I(X;Y)$ is a continuous function of $p(x)$.
5. $I(X;Y)$ is a concave function of $p(x)$. Since $I(X;Y)$ is a concave function over a closed convex set, a local maximum is global maximum. From Properties 2 and 3, the maximum is finite, and we are justified in using the term maximum rather than supremum in the definition of capacity.