A Note on Symmetric Discrete Memoryless Channels

Ingmar Land

1 Introduction

Discrete memoryless channels (DMCs) are often used to model communication channels. Many practically relevant DMCs show a certain symmetry. Such a symmetry also simplifies the computation of the capacity of these channels. Unfortunately, various definitions of symmetry can be found in literature, e.g., in [1,2]. This note gives an overview of these different kinds of symmetry and relates them to each other.

The next section introduces DMCs. The following three sections address then three kinds of symmetry. Examples illustrate the concepts. (Drawing the example channels may help understanding the concepts.)

2 Discrete Memoryless Channels

A DMC is defined by a discrete input alphabet $\mathbb{X}$, a discrete output alphabet $\mathbb{Y}$, and a probability mass function $p_{Y|X}(y|x)$ for $x \in \mathbb{X}$ and $y \in \mathbb{Y}$. The output of the channel is assumed to depend only on the current input, such that the channel is memoryless. The channel input is denoted by the random variable $X \in \mathbb{X}$, and the channel output is denoted by the random variable $Y \in \mathbb{Y}$. Symbolically, we may denote the DMC by $X \to Y$.

The size (cardinality) of the input alphabet is denoted by $M_X = |\mathbb{X}|$, and the size (cardinality) of the output alphabet is denoted by $M_Y = |\mathbb{Y}|$. Without loss of generality, we may assume

$$\mathbb{X} = \{0, 1, 2, \dots, M_X - 1\}, \qquad \mathbb{Y} = \{0, 1, 2, \dots, M_Y - 1\}.$$

The values of $p_{Y|X}(y|x)$ may be represented in a matrix

$$T = \begin{bmatrix} p_{Y|X}(0|0) & p_{Y|X}(1|0) & \cdots & p_{Y|X}(M_Y - 1|0) \\ p_{Y|X}(0|1) & p_{Y|X}(1|1) & \cdots & p_{Y|X}(M_Y - 1|1) \\ \vdots & & & \\ p_{Y|X}(0|M_X - 1) & p_{Y|X}(1|M_X - 1) & \cdots & p_{Y|X}(M_Y - 1|M_X - 1) \end{bmatrix},$$

called the transition matrix of the DMC. Notice that each row sums up to one, as it represents a probability distribution.

Example 1 (Binary Erasure Channel)
Binary input alphabet $\mathbb{X} = \{0, 1\}$, ternary output alphabet $\mathbb{Y} = \{0, \Delta, 1\}$, transition matrix

$$T = \begin{bmatrix} 1-\delta & \delta & 0 \\ 0 & \delta & 1-\delta \end{bmatrix}.$$

◇

The channel capacity of a DMC (and also of other channels) is defined as the maximal mutual information between the channel input X and the channel output Y, maximized with respect to the input distribution $p_X(x)$:

$$\begin{aligned} C &= \max_{p_X(x)} I(X;Y) \\ &= \max_{p_X(x)} \sum_{x\in\mathbb{X}} \sum_{y\in\mathbb{Y}} p_{X,Y}(x,y) \log_2 \frac{p_{Y|X}(y|x)}{p_Y(y)} \\ &= \max_{p_X(x)} \sum_{x\in\mathbb{X}} \sum_{y\in\mathbb{Y}} p_X(x) p_{Y|X}(y|x) \log_2 \frac{p_{Y|X}(y|x)}{\sum_{x'\in\mathbb{X}} p_X(x) p_{Y|X}(y|x')}. \end{aligned}$$

Notice that the mutual information $I(X;Y)$ is only a function of $p_{Y|X}(y|x)$ (a property of the channel) and of $p_X(x)$ (the maximization parameter), as made explicit in the last line.

The following notation for entropies is used. Consider a random variable $A \in \mathbb{A} = \{0, 1, \ldots, M_A\}$ with distribution $p_A(a)$. Using the vector

$$\boldsymbol{r} = [p_A(0), p_A(1), \ldots, p_A(M_A - 1)],$$

the entropy of A may be written as

$$H(A) = H(\boldsymbol{r}) = -\sum_{a\in\mathbb{A}} p_A(a) \log_2 p_A(a).$$

Furthermore, we define the binary entropy function as

$$e_2(p) := H([p, 1-p]) = -p \log_2 p - (1-p) \log_2 (1-p),$$

which is the entropy of a binary random variable taking its two values with the probabilities p and $1 - p$.

3 Strongly Symmetric DMCs

Definition 1
A DMC $X \to Y$ is called strongly symmetric if in the transition matrix, each row is a permutation of each other row, and each column is a permutation of each other column.

—

This definition follows [2] (just called "symmetric" therein) and implicitly in [1] (not explicitely defined but used to define "symmetric" DMCs).

> **Remark 1**
> A permutation of a vector is the same vector with the elements reordered. For example, $[2, 1, 3, 0]$ is a permutation of $[3, 1, 0, 2]$.

The capacity of a strongly symmetric channel is achieved by a uniform input distribution, i.e., $p_X(x) = 1/M_X$ for all $x \in \mathbb{X}$, and it can be computed as

$$C = \log_2 M_Y - H(\boldsymbol{r}), \tag{1}$$

where $\boldsymbol{r}$ denotes an arbitrary row of T. The proof can be found in [2].

> **Example 2 (Binary Symmetric Channel)**
> This DMC has a binary input alphabet $\mathbb{X} = \{0, 1\}$, a binary output alphabet $\mathbb{Y} = \{0, 1\}$, and the transition matrix
>
> $$T = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}.$$
>
> The capacity results as
>
> $$C = \log_2 2 - H([1-\epsilon, \epsilon] = 1 - e_2(\epsilon).$$
>
> ◇

4 Symmetric DMCs

Definition 2
A DMC is called symmetric if the columns of the transition matrix can be grouped into sub-matrices such that for each sub-matrix, each row is a permutation of each other row and if each column is a permutation of each other column. —

This definition is given in [1].

Notice that the sub-matrices are required to have a property which is also required for the transition matrix of strongly symmetric channels.

> **Example 3 (Binary Erasure Channel)**
> We continue Example 1. The transition matrix can be split into the two sub-matrices
>
> $$\begin{bmatrix} 1-\delta & 0 \\ 0 & 1-\delta \end{bmatrix}, \qquad \begin{bmatrix} \delta \\ \delta \end{bmatrix},$$
>
> which fulfill the required property. ◇

Symmetric DMCs achieve the capacity for uniformly distributed input, i.e., for $p_X(x) = 1/M_X$, see [1]. For computing the capacity, some further definitions and concepts are required. The example below may be studied in parallel to the following explanations.

Assume that the DMC is symmetric, and let the sub-matrices of T (fulfilling the required properties) be denoted by T'_j, $j \in \mathbb{J} = \{0, 1, \ldots, M_J\}$. As each column of T corresponds to an output value $y \in \mathbb{Y}$, each T'_j corresponds to a subset of $\mathbb{Y}$, denoted by $\mathbb{Y}_j$. The channel output Y is a random variable, and so the index of the subset $\mathbb{Y}_j$ is also a random variable, denoted by

$$J \in \mathbb{J} = \{0, 1, \ldots, M_J\},$$

The probability distribution of J is given by

$$p_J(j) = \Pr(y \in \mathbb{Y}_j) = \sum_{y \in \mathbb{Y}_j} p_Y(y). \tag{2}$$

Notice that $p_J(j)$ is the sum of the elements of one row of T'_j.

The random variable J imposes a decomposition of the symmetric DMC into sub-channels that are strongly symmetric, and it is called the sub-channel indicator. The sub-channel corresponding to $J = j$ has the input alphabet $\mathbb{X}$ and output alphabet $\mathbb{Y}_j$, and it occurs with probability $p_J(j)$; its transition matrix is given by

$$T_j = \frac{1}{p_J(j)} T'_j;$$

symbolically, this sub-channel is denoted by $X \to Y | J = j$. Notice that the sub-channels are all strongly symmetric.

The mutual information of the symmetric DMC can then be computed by averaging over the mutual informations of the subchannels:

$$I(X;Y) = \sum_{j \in \mathbb{J}} p_J(j)\ I(X;Y|J = j).$$

Accordingly, the capacity of the symmetric DMC is the average of the capacities C_j of the (strongly symmetric) sub-channels,

$$C = \sum_{j \in \mathbb{J}} p_J(j)\ C_j. \tag{3}$$

The probabilities $p_J(j)$ can be computed according to (2), and the capacities may be computed according to (1).

Example 4 (Binary Erasure Channel)
We continue Example 3. The two sub-matrices correspond to the subsets $\mathbb{Y}_0 = \{0, 1\}$ and $\mathbb{Y}_1 = \{\Delta\}$ of the output alphabet $\mathbb{Y}$. The sub-channel indicator $J \in \mathbb{J} = \{0, 1\}$ is distributed as

$$p_J(j) = \begin{cases} 1 - \delta & \text{for } j = 0\ (y \in \mathbb{Y}_0), \\ \delta & \text{for } j = 1\ (y \in \mathbb{Y}_1). \end{cases}$$

These two probabilities are also the probabilities of the two sub-channels. The transition matrices of the two sub-channels are

$$T_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad T_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The capacities of the two sub-channels are $C_0 = 1$ and $C_1 = 0$. The capacity of the BEC results as

$$\begin{aligned} C &= p_J(0) \cdot C_0 + p_J(1) \cdot C_1 \\ &= (1-\delta) \cdot 1 + \delta \cdot 0 = 1 - \delta. \end{aligned}$$

◇

Remark 2
An additive white Gaussian noise (AWGN) channel with input alphabet $\mathbb{X} = \{-1, +1\}$ can also be considered to be symmetric. In a first step, the output can be quantized with quantization intervals being symmetric with respect to $y = 0$; the sub-channels are all BSCs, and the resulting channel is symmetric according to the above definition. In a second step, the quantization intervals can be made arbitrarily small; the sub-channels are (still) BSCs, but the number of sub-channels tends to infinity.

5 Less-Strongly Symmetric Channels

The third definition of symmetry is given in [2], and these channels are called “weakly symmetric” therein. It turns out, however, to be less strict than the one for “symmetric channels” given in the previous section. Therefore, the notion “less-strongly symmetric channels” is used in the following (with a twinkle in one’s eye).

Definition 3
A DMC is called less-strongly symmetric if each row is a permutation of each other row and if all the column sums are equal.

The capacity of a less-strongly symmetric channel is achieved by a uniform input distribution, i.e., $p_X(x) = 1/M_X$ for all $x \in \mathbb{X}$, and it can be computed as

$$C = \log_2 M_Y - H(\boldsymbol{r}), \tag{4}$$

where $\boldsymbol{r}$ denotes an arbitrary row of T. The proof can be found in [2].

Example 5 (A Less-Strongly Symmetric Channel)
Consider the DMC with a ternary input alphabet $\mathbb{X} = \{0, 1, 2\}$, a ternary output alphabet $\mathbb{Y} = \{0, 1, 2\}$, and the transition matrix

$$T = \begin{bmatrix} 1/3 & 1/6 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{bmatrix}.$$

This channel is not strongly symmetric, but it is less-strongly symmetric. The capacity results as

$$C = \log_2 3 - H([1/3, 1/6, 1/2].$$

Notice that this channel is also symmetric. ◇

The condition is stricter than that for symmetric channels. Unfortunately, a BEC does not fulfill the requirements for "less-strongly symmetric DMCs", as can easily be seen. Since it should be included in a definition for symmetry, Definition 3 is obviously not general enough. It may nevertheless be useful in certain cases.

6 Summary

This note has summarized three definitions of symmetry for DMCs which can be found in [1, 2]. The strongest condition is that of "strongly symmetric channels". A bit weaker is the one for "less-strongly symmetric channels" (called "weakly symmetric channels" in [2]). And the least strong condition is that for "symmetric channels".

The notion of "less strongly symmetric channels" has shown to be not general enough as the BEC is not included. Therefore, the notions of "strongly symmetric channels" and "symmetric channels" seem to be sufficient.

References

[1] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1991.