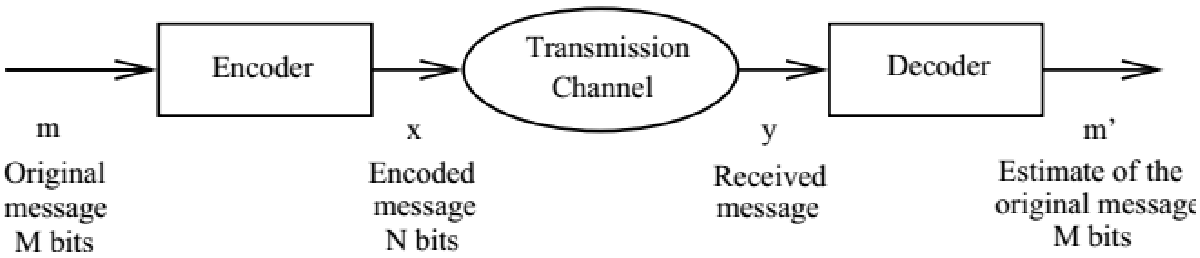
Information transmission:

The typical flowchart of a communication system is shown in Figure below. It applies to situations as diverse as communication between the earth and a satellite, the cellular phones, or storage within the hard disk of your computer.

Alice wants to send a message m to Bob. Let us assume that m is a M bit sequence. This message is first encoded into a longer one. The encoded message is sent through the communication channel. The output of the channel is a message y.

In a realistic channel, y is in general a string of symbols different from x.



When dealing with "information", one of the basic goals is to transmit it reliably. In this context, "transmit" both means "transmitting some information from one point to another", as we usually understand it, but also to "transmit" it through time؛ for instance to store it somewhere (to memorize it) and then retrieve it later on. In both cases however, transmission of information can, in real life, hardly be achieved in a fully reliable manner.

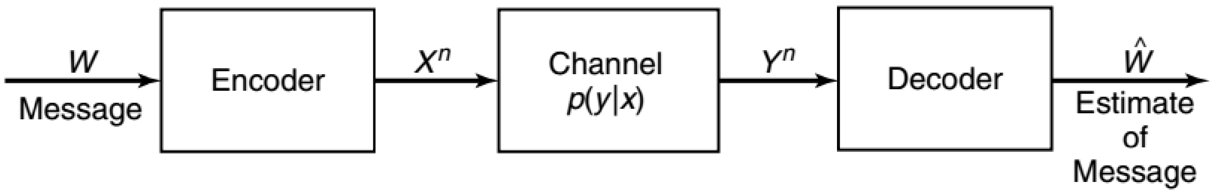
There always exists a risk of distortion of the transmitted information: some noise on the line, some leak of the memory or the hard disk storing the information, etc.

What effect does noise have on the transmission of messages? Several situations could be possible:

1- it is never possible to transmit any messages reliably (too much noise);
2- it is possible to transmit messages with a "reasonable" error probability;
3- it is possible to transmit messages with an error probability which is as small as we can wish for (using error correcting codes).

Channel Capacity

What do we mean when we say that A communicates with B? Source symbols from some finite alphabet are mapped into some sequence of channel symbols, which then produces the output sequence of the channel. The output sequence is random but has a distribution that depends on the input sequence. From the output sequence, we attempt to recover the transmitted message. Each of the possible input sequences brings a probability distribution on the output sequences.



Definition We define a discrete channel to be a system consisting of an input alphabet X and output alphabet Y and a probability transition matrix p(y|x) that expresses the probability of observing the output symbol y given that we send the symbol x.

The channel is said to be memoryless if the probability distribution of the output depends only on the input at that time and is conditionally independent of previous channel inputs or outputs.

Definition 4.1 (Discrete Memoryless Channel) The *discrete memoryless channel* (*DMC**) is the simplest kind of communication channel. Formally, DMC consists of three quantities:

1. a discrete *input alphabet*, $\mathcal{V}_X$, the elements of which represent the possible emitted symbols for all input messages (the source X);
2. a discrete *output alphabet*, $\mathcal{V}_Y$, the elements of which represent the possible received symbols (output sequence); and
3. for each $x \in \mathcal{V}_X$, the *conditional probability distributions* $p_{Y|X=x}$ over $\mathcal{V}_Y$ which describe the channel behavior in the manner that, for all $n = 1, 2, 3, \ldots$:

$$P(Y_n = y_n | X_1 = x_1, \ldots, X_n = x_n, Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1}) = P(Y = y_n | X = x_n), \quad (4.1)$$

These are called the *transmission probabilities** of the channel.

The purpose of a channel is to transmit messages ("information") from one point (the input) to another (the output). The *channel capacity** precisely measures this ability: it is the maximum average amount of information the output of the channel can bring on the input.

Recall that a DMC if fully specified by the conditional probability distributions $p_{Y|X=x}$ (where X stands for the input of the channel and Y for the output). The input probability distribution $p_X(x)$ is not part of the channel, but only of the input source used. The *capacity* of a channel is thus defined as the maximum mutual information $I(X;Y)$ that can be obtained among all possible choice of $p_X(x)$. More formally:

Definition 4.2 (Channel Capacity) The capacity C of a Discrete Memoryless Channel is defined as

$$C = \max_{p_X} I(X;Y), \tag{4.4}$$

where X stands for the input of the channel and Y for the output. ♦

The ***channel capacity*** C is the maximum amount of information that a channel can provide at its output about the input . The rate at which information is transmitted through the channel depends on the entropies of three variables:

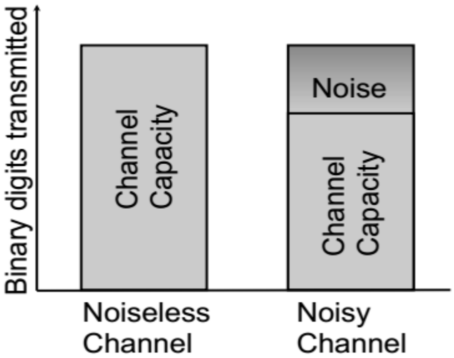
1) the entropy $H(x)$ of the input,

2) the entropy $H(y)$ of the output,

3) the entropy $H(\eta)$ of the noise in the channel.

If the output entropy is high then this provides a large potential for information transmission, and the extent to which this potential is realized depends on the input entropy and the level of noise.

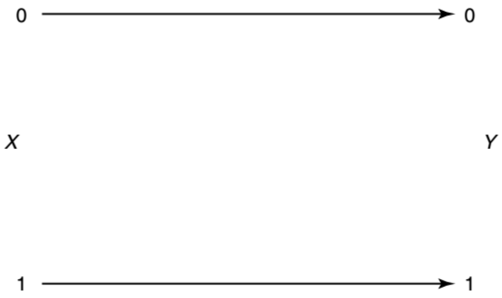
If the noise is low then the output entropy can be close to the channel capacity. However, channel capacity gets smaller as the noise increases. Capacity is usually expressed in bits per usage (i.e. bits per output), or bits per second (bits/s).



EXAMPLES OF CHANNEL CAPACITY

1- Noiseless Binary Channel

Suppose that we have a channel whose the binary input is reproduced exactly at the output . 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 $p(x) = (1/2, 1/2)$.



2- Noisy Channel with Nonoverlapping Outputs

This channel has two possible outputs corresponding to each of the two inputs. The channel appears to be noisy, but really is not. 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 $C = \max I(X;Y) = 1$ bit, which is achieved by using $p(x) = (1/2, 1/2)$.