Chapter 1

Languages

While we may think that computers are just machines, there is an incredible work behind that aims to make some kind of language that computers can understand. It is well renewed that computers act depending on streams of binary code, but that same stream of binary code can be considered as a language, something that the computers "speak" and "think" with. Let's define more in detail this "computers' language".

DEFINITION

Alphabeth and Strings

An **alphabet** Σ is a **non-empty**, **finite set**, which contains elements called **symbols** (or **characters**). For instance, the following two sets are considered alphabets:

- $\Sigma = \{0, 1\}$
- $\Sigma = \{a, b, c, \ldots, x, y, z\}$

A **string** w over an alphabet Σ is a **sequence of symbols**, all belonging to Σ, which are all written one after the other and aren't separated by other symbols

Strings also have different properties, such as a **length**, a **reverse** and the possibility to include one or more **substrings**:

- **Length**: defined as $|w|$, it denotes the **number of symbols** contained within w. If a string has length 0, then such string is called **empty string**, and is denoted with ϵ;
- **Reverse**: defined as w^R, the reverse is a string which contains all the symbols of w in the reverse order;
- **Substring**: we say that a string z is a substring of w if z appears consecutively within w.

For instance, the string $w = 00101$ has a length of 5, its reverse is $w^R = 10100$, and has, as possible substrings the following: $001, 10, \epsilon$. Its alphabet is $\Sigma = \{0, 1\}$