

# Capítulo 1

## Lenguajes formales

### 1.1. Alfabetos y palabras

Un **alfabeto** es un conjunto finito no vacío cuyos elementos se llaman **símbolos**. Denotamos un alfabeto arbitrario con la letra  $\Sigma$ .

Una **palabra** o **cadena** sobre un alfabeto  $\Sigma$  es cualquier sucesión finita de elementos de  $\Sigma$ . Admitimos la existencia de una única palabra que no tiene símbolos, la cual se denomina **palabra vacía** y se denota con  $\lambda$ . La palabra vacía desempeña, en la teoría de lenguajes formales, un papel similar al que desempeña el conjunto vacío  $\emptyset$  en la teoría de conjuntos.

**Ejemplo** Sea  $\Sigma = \{a, b\}$  el alfabeto que consta de los dos símbolos  $a$  y  $b$ . Las siguientes son palabras sobre  $\Sigma$ :

$aba$   
 $ababaaa$   
 $aaaab$

Obsérvese que  $aba \neq aab$ . El orden de los símbolos en una palabra es significativo ya que las palabras se definen como *sucesiones*, es decir, conjuntos *secuencialmente ordenados*.

El conjunto de *todas* las palabras sobre un alfabeto  $\Sigma$ , incluyendo la palabra vacía, se denota por  $\Sigma^*$ .

**Ejemplo** Sea  $\Sigma = \{a, b, c\}$ , entonces

$$\Sigma^* = \{\lambda, a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc, aaa, aab, abc, baa, \dots\}.$$

**Observaciones:**

1. Si bien un alfabeto  $\Sigma$  es un conjunto finito,  $\Sigma^*$  es siempre un conjunto

infinito (enumerable). En el caso más simple,  $\Sigma$  contiene solo un símbolo,  $\Sigma = \{a\}$ , y  $\Sigma^* = \{\lambda, a, aa, aaa, aaaa, aaaaa, \dots\}$ .

- Hay que distinguir entre los siguientes objetos, que son todos diferentes entre sí:

$$\emptyset \quad \lambda \quad \{\emptyset\} \quad \{\lambda\}$$

- La teoría de lenguajes se hace con referencia a un alfabeto  $\Sigma$  fijo (pero arbitrario).

Notación usada en la teoría de lenguajes	
$\Sigma$	denota un alfabeto.
$\Sigma^*$	denota el conjunto de todas las palabras que se pueden formar con los símbolos de $\Sigma$ .
$a, b, c, d, e, \dots$	denotan símbolos del alfabeto $\Sigma$ .
$u, v, w, x, y, z, \dots$	denotan palabras, es decir, sucesiones finitas de símbolos de $\Sigma$ .
$\lambda$	denota la palabra vacía, es decir, la única palabra en $\Sigma^*$ que no tiene símbolos.
$A, B, C, \dots, L, M, N, \dots$	denotan lenguajes (definidos más adelante).