Tema 4: Aplicación de los autómatas: Lenguajes formales

1. Lenguajes formales.

Como se ha indicado en la introducción del tema, el concepto de autómata surgió cuando se modelizó matemáticamente el sistema nueronal humano. Sin embargo, se pueden encontrar conexiones con otras ramas, como por ejemplo los lenguajes formales.

Definición. Un **alfabeto** A es un conjunto finito cuyos elementos se denominan **letras**.

Definición. Un **lenguaje** Ł sobre el alfabeto A es un subconjunto de Ω_A. A los elementos de Ł se les denominan **palabras**.

Definición. Un lenguaje Ł sobre el alfabeto A se dice que es finito si $|Ł|$ es finito.

Definición. Llamaremos **lenguaje vacío** sobre el alfabeto A a $Ł_\emptyset = \emptyset$.

Debemos distinguir entre $Ł_\emptyset$ y $Ł_\Lambda = \{\Lambda\}$, esto es el lenguaje que contiene sólo la palabra vacía.

Como ya sabemos en Ω_A tenemos definida la operación concatenación. Por ello, cuando nos den una palabra $x \in \Omega_A$, podemos verla como la concatenación de varias palabras, alguna de ellas posiblemente vacías. Así, diremos que y_2 es una **subpalabra** de x si $x = y_1y_2y_3$, donde $y_i \in \Omega_A$. Observamos que y_1 ó y_3 pueden ser la palabra vacía. Si $y_1 = \Lambda$, diremos que y_2 es una subpalabra inicial y si $y_3 = \Lambda$, diremos que y_2 es una subpalabra final.

Ejemplos. $Ł_1 = \{\Lambda, a_1a_2, a_1, a_2a_1\}$ es un lenguaje sobre el alfabeto $\{a_1, a_2\}$. $Ł_2 = \{a_1^i | i \in \mathbb{N}\}$ es un lenguaje sobre el alfabeto $\{a_1\}$. $Ł_1$ es un lenguaje finito y $Ł_2$ no.

En el conjunto $\mathbf{L} = \{Ł | Ł \text{ es lenguaje sobre el alfabeto A}\}$, podemos definir las siguientes operaciones:

1. **Suma de dos lenguajes**: Dados $Ł_1, Ł_2 \in \mathbf{L}$,

$$Ł_1 + Ł_2 = \{x \in \Omega_A | x \in Ł_1 \vee x \in Ł_2\}.$$

2. **Intersección de dos lenguajes**: Dados $Ł_1, Ł_2 \in \mathbf{L}$,

$$Ł_1 \cap Ł_2 = \{x \in \Omega_A | x \in Ł_1 \wedge x \in Ł_2\}.$$

3. **Complementario de un lenguaje**: Dado $Ł_1 \in \mathbf{L}$,

$$Ł_1^c = \{x \in \Omega_A | x \notin Ł_1\}.$$

4. **Diferencia de dos lenguajes**: Dados $Ł_1, Ł_2 \in \mathbf{L}$,

$$Ł_1 - Ł_2 = \{x \in \Omega_A | x \in Ł_1 \wedge x \notin Ł_2\}.$$

   Es claro que $Ł_1 - Ł_2 = Ł_1 \cap Ł_2^c$.

5. **Concatenación ó producto de dos lenguajes** Dados $Ł_1, Ł_2 \in \mathbf{L}$,

$$Ł_1 Ł_2 = \{x \in \Omega_A | x = y_1 y_2, \quad y_i \in Ł_i, \ i = 1, 2\}.$$

   Es evidente que la concatenación de lenguajes es asociativo por serlo la concatenación de palabras. Además, $Ł_\emptyset$ y $Ł_\Lambda$ son elementos cero e identidad para la concatenación de lenguajes sobre el mismo alfabeto.

6. **Clausura de un lenguaje**: Dado $Ł_1 \in \mathbf{L}$,

$$Ł_1^* = \sum_{i=0}^{\infty} Ł_1^i,$$

   donde $Ł_1^i$ representa el producto de $Ł_1$ i veces y $Ł_1^0 = \{\Lambda\}$. Esto es, la clausura del lenguaje $Ł_1$ contiene todas las palabras de Ω_A que se pueden obtener como concatenación de palabras de $Ł_1$ más la palabra vacía.