Tema 4: Aplicación de los autómatas: Lenguajes formales

2. Lenguajes regulares.

Introducimos ahora el concepto de expresión regular .

Definición. Una **expresión regular** sobre el alfabeto A es una palabra de $\Omega_{A\cup I}$, donde $I = \{+, *, \emptyset, (,)\}$, satisfaciendo las siguientes condiciones:

1. Cada letra de A y $\emptyset$ son expresiones regualares.

2. Si α y β son expresiones regulares sobre A, también lo son $(\alpha + \beta)$, $(\alpha\beta)$ y α^*.

3. Nada es expresión regular sobre A, salvo que se obtenga tras la aplicación de un número finito de veces de 1. y 2.

Ejemplo. $\emptyset^*$, $((a_1 + a_2)(a_1 + a_2))^*$ y $(a_1 + a_2^*)a_3$ son expresiones regulares sobre el alfabeto A, cuando $a_1, a_2, a_3 \in A$.

Definición. Sea α una expresión regular sobre el alfabeto A. se llama lenguaje asociado a α, y se denota por $|\alpha|$, al lenguaje de Ω_A que se obtiene de acuerdo a las siguientes condiciones:

1. $|\emptyset| = L_\emptyset$.

2. $\forall a \in A$, $|a| = \{a\}$.

3. $\forall \alpha, \beta$ expresiones regulares sobre A,

$$|(\alpha + \beta)| = |\alpha| + |\beta|, \qquad |(\alpha\beta)| = |\alpha||\beta|, \qquad |\alpha^*| = |\alpha|^*.$$

Ejemplo. $|\emptyset^*| = \{\Lambda\} = Ł_\Lambda$ es el lenguaje asociado a la expresión regular $\emptyset^*$. Por otro lado, el lenguaje asociado a la expresión regular a^*, donde $a \in A$, viene dado por $|a^*| = |a|^* = \{a^i | i \in \mathbb{N}\}$, entendiendo que $a^0 = \Lambda$.

Definición. Un lenguaje Ł sobre el alfabeto A se dice que es **regular** si existe una expresión regular α tal que Ł $= |\alpha|$.

Es fácil comprobar que todo lenguaje finito es regular. Puede suceder también que dos expresiones regulares distintas tengan asociado el mismo lenguaje regular. Por ejemplo, las expresiones regulares $(a_1 + a_2)^*$ y $(a_1^* a_2^*)^*$ tienen por lenguaje asociado a Ω_A, donde $A = \{a_1, a_2\}$.