

Expresiones Regulares

Alan Reyes-Figueroa

Teoría de la Computación

(Aula 03) 09.julio.2025

Operaciones para lenguajes

Operaciones para cadenas

Expresiones regulares

Operaciones para lenguajes

Dados lenguajes L , L_1 y L_2 , tenemos

□ La *concatenación* L_1L_2 es el lenguaje

$$L_1L_2 = \{w_1w_2 : w_1 \in L_1, w_2 \in L_2\}.$$

□ La *unión* $L_1 \cup L_2$ es el lenguaje

$$L_1 \cup L_2 = \{w : w \in L_1 \text{ ó } w \in L_2\}.$$

□ La *intersección* $L_1 \cap L_2$ es el lenguaje

$$L_1 \cap L_2 = \{w : w \in L_1 \text{ y } w \in L_2\}.$$

Operaciones para lenguajes

Dados un lenguaje L , tenemos

□ La *cerradura de Kleene* de L es

$$L^* = \{w_1 w_2 \dots w_n : w_i \in L, n \geq 0\}.$$

□ La *cerradura positiva* de L es el lenguaje

$$L^+ = \{w_1 w_2 \dots w_n : w_i \in L, n \geq 1\}.$$

Operaciones para lenguajes

Ejemplo: $L_1 = \{a, aa\}$ $L_2 = \{\epsilon, a, b, c\}$.

$L_1 \cup L_2 = \{\epsilon, a, aa, b, c\}$. $L_1 \cap L_2 = \{a\}$.

$L_1 L_2 = \{a, aa, ab, ac, aaa, aab, aac\}$.

$L_2^* = \{\epsilon, a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc, aaa, aab, aac, aba, abb, abc, \dots\}$.

Algunas observaciones

Dada una cadena $w = a_1 a_2 \dots a_k$,
su *reversión* es $w^R = a_k \dots a_2 a_1$.

Para un lenguaje L , la reversión del
lenguaje L es

$$L^R = \{w^R: w \in L\}.$$

Observe que $L_2 L_1 = (L_1 L_2)^R$

$$L L L \dots L \text{ (n veces)} = L^n$$

Expresiones regulares

Una *expresión regular* o *regex* es una representación compacta para ciertos lenguajes. (**Obs!** no todos los lenguajes admiten una expresión regular).

Los lenguajes representables mediante expresiones regulares se llaman *lenguajes regulares*.

Dado un alfabeto Σ , para representar un lenguaje regular usamos los símbolos en Σ , y ciertos operadores especiales.

Expresiones regulares

- ab ó $a \cdot b$, para la concatenación
- $a + b$, $a \cup b$, ó $a \mid b$, para la unión
- a^* para la cerradura de Kleene
- a^+ para la cerradura positiva $a^+ = aa^*$
- $(,), [,]$, para definir agrupaciones y jerarquías

También se usan otros símbolos como
abreviaturas: $[a_1, a_2, \dots, a_n]$

$[a_1 - a_n]$

Expresiones regulares

Las expresiones regulares en un alfabeto Σ se construyen siguiendo las

1. ε y cualquier elemento de Σ es una expresión regular.
2. Si α y β son expresiones regulares, también lo es $\alpha\beta$.
3. Si α y β son expresiones regulares, también lo es $\alpha \mid \beta$.
4. Si α es expresión regular, también lo son α^* y α^+ .
5. Sólo las reglas 1-4 generan expresiones regulares.

Jerarquía de operadores

Se siguen las siguientes reglas de precedencia:

1. Si hay $()$ ó $[]$, se ejecutan primero las operaciones dentro de éstos.
2. La estrella de Kleene $*$ (y la suma de Kleene $+$) es un operador unario. Tiene precedencia sobre la concatenación $\cdot$ y sobre la unión $+$.
3. La concatenación $\cdot$ tiene precedencia sobre la unión $+$.

Ejemplo: ¿Qué diferencia hay entre $(a+b)^*$ y $a+b^*$?

Expresiones regulares

Ejemplo: En $\Sigma = \{0,1\}$

$\alpha = (0+1)^*0$ es expresión regular.

Representa todas las cadenas terminadas en 0.

Ejemplo: En $\Sigma = \{a,b,c\}$

$\beta = b^*(abb^*)(a+c)$ es expresión regular.

Cadenas que comienzan con un número cualquiera de b's, luego tiene ab, luego tiene cualquier número de b's luego terminan en a ó en c.

Propiedades

- $r|s = s|r$ la unión es conmutativa
- $r|(s|t) = (r|s)|t$ la unión es asociativa
- $r(st) = (rs)t$ concatenación es asociativa
- $r(s|t) = rs|rt$ la concatenación se distribuye respecto de $|$
- $(s|t)r = sr|tr$
- $\varepsilon r = r\varepsilon = r$ ε elemento neutro de $\cdot$
- $(r|\varepsilon)^* = r^*$ ε siempre está en una cerradura de Kleene
- $r^{**} = r^*$ * es idempotente