TEMA 1. LENGUAJES FORMALES

Alfabeto (∑): Conjunto finito de símbolos gráficos. $\Sigma_1 = \{ 0, 1 \}$, $\Sigma_2 = \{a, b, c, d, \dots , z \}$

Subalfabeto (∑): Subconjunto de un alfabeto. $\Sigma_4 = \{c, d\}$ es un subalfabeto de Σ_2

Palabra (x): dado un alfabeto ∑. Conjunto finito y ordenado de símbolos del alfabeto con o sin repetición.

x,y ⟶ palabras / a,b ⟶ símbolos   x = ababa (palabra)

Palabra vacía (λ): La palabra que no contiene símbolos. λ no ϵ ∑

Longitud de una palabra (| x |): es igual al nº de símbolos de una palabra.

$x = aabb \rightarrow | x | = 4$, $z = \lambda \rightarrow | z | = 0$

Palabra inversa (x^{-1}): es la palabra con todos sus símbolos a la inversa. $x^{-1} = bbaa$

Definición Recursiva: Si $| x | = 0 \rightarrow x = \lambda , x^{-1} = \lambda$ // $| x | > 0 \rightarrow \exists$ w palabra y $\exists$ a ϵ ∑ tal que $x = wa$, $x^{-1} = aw^{-1}$

Palabra simétrica: Una palabra es simétrica si $x = x^{-1}$. aba es simétrica

Lenguajes universal (∑*): El conjunto de todas las palabras que se pueden formar sobre ∑, incluyendo λ. ∑ es finito pero ∑* es infinito.

OPERACIONES CON PALABRAS

Concatenación de palabras (·): Es una operación mediante la cual a partir de dos palabras se obtiene otra al poner una detrás de la otra. x = ab, y = cd; x · y = abcd

Propiedades:

- Operación bien definida: $x, y \in \Sigma^* \rightarrow x \cdot y \in \Sigma^*$
- Asociativa: $x , y , z \in \Sigma^*; \; x \cdot (y \cdot (z)) = ((x \cdot y) \cdot z)$
- Elemento neutro: $\forall x \in \Sigma^* \; \exists \; \lambda$ tal que $x \cdot \lambda = \lambda \cdot x = x$
- No es conmutativa
- (∑*, ·) Tiene estructura de semigrupo.
- $| x \cdot y | = | x |+| y |$
- $\forall x , y \in \Sigma^* \; (x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$
- $x^i = x \cdot x \cdot x \,(i\ veces)\;\; x^0 = \lambda$
- $| x^i | = i \cdot | x |$ // $x^{i+j} = xi \cdot xj$

- **Simplificativa**

  $xy = xz \Rightarrow y = z$

  $xy = zy \Rightarrow x = z$

$x^{i+1} = x^i \cdot x = x \cdot x^i$

$x^{i+j} = x^i \cdot x^j$

$|x^i| = i \cdot |x|$

Potencia: se llama potencia i-ésima de una palabra a la concatenación consigo misma i veces. Propiedades de la potencia:
$x = 01$, $x^3 = 010101$

<u>Posible Ejercicio</u>: *Comprobar que $(wx)^{-1} = x^{-1}w^{-1}$*

$|x| = 0 \longrightarrow x = \lambda \longrightarrow (w\lambda)^{-1} = w^{-1} = \lambda w^{-1} = \lambda^{-1}w^{-1} = x^{-1}w^{-1}$
$|x| = n + 1 \longrightarrow x = ya \quad | \quad y \in \Sigma^* \quad |y| = n \qquad a \in \Sigma$

Hipótesis de inducción: $(wy)^{-1} = y^{-1}w^{-1}$ Suponemos cierto (para a = 1)

$$(wx)^{-1} = (w(ya))^{-1} \underset{\text{Asociativa}}{=} ((wy)a)^{-1} \underset{\text{Definición (inversa)}}{=} a(wy)^{-1} \underset{\text{Hip. Ind}}{=} ay^{-1}w^{-1} \underset{\text{Asociativa}}{=} (ay^{-1})w^{-1} \underset{\text{Definición}}{=} (ya)^{-1}w^{-1} = \mathbf{x^{-1}w^{-1}}$$

LENGUAJES FORMALES

Dado un alfabeto Σ , L es un lenguaje formal sobre Σ cuando sus palabras cumplen una propiedad. $L = \{ x \in \Sigma^* \mid x \text{ cumple } P \}$

$\mathfrak{T}$ = Todos los lenguajes sobre Σ^* (todos los lenguajes posibles sobre el lenguaje universal)

- Tanto $\emptyset$ como $\{ \lambda \}$ son lenguajes sobre cualquier alfabeto.
- El alfabeto Σ puede considerarse como uno de los lenguajes generados por él mismo

Ejemplo 1:

- $\Sigma = \{0, 1\}$
- $L_1 = \{1, 10, 100, 1000 \ldots\}$ L es el conjunto de las palabras que empiezan por 1 seguidos de cero o más ceros.

<u>OPERACIONES CON LENGUAJES:</u>

Unión de lenguajes: Dado un alfabeto Σ y L_1 y L_2 lenguajes sobre Σ, la unión de dichos lenguajes esta definida como:

$$L_1 \cup L_2 = \{ x \in \Sigma^* \mid x \in L_1 \text{ o } x \in L_2 \}$$

<u>Propiedades:</u>

- Operación bien definida y cerrada: la unión de dos lenguajes sobre el mismo alfabeto es también un lenguaje sobre dicho alfabeto.
- Asociativa: $\forall L_1, L_2, L_3 \longrightarrow (L_1 \cup L_2) \cup L_3 = L_1 \cup (L_2 \cup L_3)$
- Elemento neutro: $L1 \cup \emptyset = \emptyset \cup L1 = L1$
- Es conmutativa: $L_1 \cup L_2 = L_2 \cup L_1$
- Idempotente: $L \cup L = L$
- $(\Sigma^*, \cdot)$ Tiene estructura de semigrupo con elemento neutro y conmutativa.

Concatenación de lenguajes: Dado un alfabeto $\sum$ y L_1 y L_2 lenguajes sobre $\sum$, la concatenación de dichos lenguajes:

$$L_1 \cdot L_2 = \{ x \in \sum* \mid \exists \ u \in L_1, v \in L_2 , x = uv \}$$

Propiedades:

- Operación bien definida: $x, y \in \sum* \rightarrow x \cdot y \in \sum*$
- Asociativa: $\forall L_1, L_2, L_3 : (L_1 \cdot L_2) \cdot L_3 = L_1 \cdot (L_2 \cdot L_3)$
- Elemento neutro: $\forall L \in \mathfrak{T} \exists \{ \lambda \} , L \cdot \{ \lambda \} = \{ \lambda \} \cdot L = L$
- No es conmutativa
- Tiene estructura de semigrupo con elemento neutro

Potencia de un lenguaje: es la operación que consiste en concatenarlo consigo mismo i veces. $L^1 = L$ $L^i = L \cdot L \cdot L \cdot L \cdot ... \cdot L$ (i veces)

Se verifica que: 1) $L^i + 1 = L^i \cdot L = L \cdot L^i$ $(i > 0)$ 2) $L^i \cdot L^j = L^{i+j}$ $(i, j > 0)$

$L^0 = \{ \lambda \}$ cualquiera que sea L.

Cierre o Clausura (Estrella de Kleene): dado un lenguaje es la unión de todas las potencias del lenguaje, incluida la 0.

$L* = \bigcup_{i=0}^{\infty} L^i = L^0 \cup L^1 \cup L^2 \cup L^3 ...;$ $L^0 = \{ \lambda \}$

- Si $L = \sum$ entonces $L^* = \sum^*$ (El cierre de un lenguaje es igual al lenguaje universal si el lenguaje es igual que su alfabeto)

Clausura Positiva: es el lenguaje que se forma por la unión de todas las potencias de L menos L^0. $L^+ = \bigcup_{i=1}^{\infty} L^i = L^i = L^1 \cup L^2 \cup L^3 ...;$

- Si $L = \sum$ entonces $L^+ = \sum^* - \{ \lambda \}$

Se puede llegar a las siguientes conclusiones:

- $L^* = L^+ \cup \{ \lambda \}$ - $L^+ = L \cdot L^* = L^* \cdot L$

Lenguaje Inverso: dado un lenguaje L está formado por la aplicación de inversión a cada una de las palabras de L. $L^{-1} = \{w \in \sum* \mid w = u^{-1} \text{ donde } u \in L\}$

Ejemplo: $\sum = \{0, 1\}$ $L(\sum) = \{0, 1, 00, 10\}$ $L(\sum)^{-1} = \{0, 1, 00, 01\}$

Intersección de lenguajes: dado un alfabeto ∑ y los lenguajes L_1 y L_2, su intersección será el conjunto que contenga las palabras que pertenezcan a los dos lenguajes.

$$L_1 \cap L_2 = \{x \in \Sigma^* \ / \ x \in L_1 \cap x \in L_2\}$$

Propiedades:

- Operación bien definida:
- Asociativa: $\forall L_1, L_2, L_3 : (L_1 \cap L_2) \cap L_3 = L_1 \cap (L_2 \cap L_3)$
- Elemento neutro: $\forall L \in \Sigma^* \rightarrow L \cap \Sigma^* = \Sigma^* \cap L = L$
- Es conmutativa: $\forall L_1, L_2, L_3 : (L_1 \cap L_2) \cap L_3$
- Tiene estructura de semigrupo con elemento neutro y conmutativa.

Complementación de lenguajes: es el conjunto de palabras que no pertenecen al lenguaje.

$$\overline{L} = \{w \in \Sigma^* \ / \ w \notin L\} = \Sigma^* - L$$

Propiedades de la complementación, Leyes de Morgan:

- $\forall L_1, L_2 \in \mathfrak{T}$; $\quad \overline{L_1} \cap \overline{L_2} = \overline{L_1 \cup L_2} \qquad \overline{L_1 \cap L_2} = \overline{L_1} \cup \overline{L_2}$

EXPRESIONES REGULARES

Permiten representar lenguajes de forma que resumen la descripción de dicho lenguaje. Así, un lenguaje infinito tiene una representación finita.

Dado un alfabeto ∑, se construye a partir de él un nuevo alfabeto $B = \Sigma \cup \{\emptyset, \lambda, +, *, \cdot, (\ )\}$

Definimos expresión regular α sobre ∑.

$\boldsymbol{\alpha \in B^*}$ **si:**

$\boxed{\alpha = ER(\Sigma) \subset B^+}$

$a \in \Sigma$	$\Rightarrow a, \emptyset, \lambda$ son expresiones regulares sobre ∑
$\alpha, \beta \in ER(\Sigma)$	$\Rightarrow \alpha + \beta \in ER(\Sigma)$
$\alpha, \beta \in ER(\Sigma)$	$\Rightarrow \alpha \cdot \beta \in ER(\Sigma)$
$\alpha \in ER(\Sigma)$	$\Rightarrow \alpha^* \in ER(\Sigma)$ que se define $\alpha^* = \bigcup_{i=0}^{\infty} \alpha^i$ $(\alpha^0 = \lambda)$
$\alpha \in ER(\Sigma)$	$\Rightarrow (\alpha) \in ER(\Sigma)$

Son expresiones regulares de ∑ únicamente las obtenidas por la aplicación de un número finito de veces de las 5 reglas anteriores.

El orden de prioridad de las operaciones en una ER es: * (cierre); · (concatenación) y + (unión). Los paréntesis dan prioridad.

Toda ER definida sobre $\sum$, representa un lenguaje regular que se define recursivamente así:

- Si $\alpha = \emptyset$ $\quad L(\alpha) = L(\emptyset) = \emptyset$
- Si $\alpha = \lambda$ $\quad L(\alpha) = L(\lambda) = \lambda$
- Si $\alpha = a$ $\quad L(\alpha) = L(a) = \{a\}$
- Si α y β son ER $\quad L(\alpha + \beta) = L(\alpha) + L(\beta)$
- Si α y β son ER $\quad L(\alpha \cdot \beta) = L(\alpha) \cdot L(\beta)$
- Si α es ER $\quad L(\alpha^*) = (L(\alpha))^*$

Un lenguaje L es un lenguaje regular si $\exists \alpha$ tal que $L(\alpha) = L$. Es decir, si se puede representar mediante una ER. Además, todo lenguaje finito es regular.

Ejemplo:

$L(\alpha) = L(abca) = L(a)L(b)L(c)L(a) = \{a\}\{b\}\{c\}\{a\} = \{abca\}$

EQUIVALENCIAS DE EXPRESIONES REGULARES

Dos expresiones regulares α y β son equivalentes si describen el mismo lenguaje regular.

$$\alpha, \beta \in ER(\sum) \qquad \alpha \equiv \beta \Leftrightarrow L(\alpha) = L(\beta)$$

Tenemos las siguientes propiedades:

1) + es asociativa:
   - $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
2) + es conmutativa:
   - $\alpha + \beta = \beta + \alpha$
3) · es asociativa:
   - $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$
4) · es distributiva respecto a la + (unión):
   - $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$
   - $(\beta + \gamma) \cdot \alpha = \beta \cdot \alpha + \gamma \cdot \alpha$
5) · tiene elemento neutro λ:
   - $\alpha \cdot \lambda = \lambda \cdot \alpha = \alpha$
6) + tiene elemento neutro ∅:
   - $\alpha + \emptyset = \emptyset + \alpha = \alpha$

7) $\lambda^* = \lambda$

8) $\emptyset \alpha = \alpha \emptyset = \emptyset$

9) $\emptyset^* = \lambda$

10) $\alpha^* \cdot \alpha^* = \alpha^*$

11) $\alpha \cdot \alpha^* = \alpha^* \cdot \alpha$

12) $(\alpha^*)^* = \alpha^*$

13) $\boldsymbol{\alpha^* = \lambda + \alpha \cdot \alpha^*}$

14) $(\alpha^* + \beta^*)^* = (\alpha^* \beta^*)^* = (\alpha + \beta)^*$

15) $(\alpha \beta)^* \alpha = \alpha (\beta \alpha)^*$

16) $(\alpha^* \beta)^* \alpha^* = (\alpha + \beta)^*$

Ejemplo: $aa^* + \lambda = a^*$

- $aa^* + \lambda = a(\lambda + a + a^2 + \ldots) + \lambda = (a + a^2 + \ldots) + \lambda$
- $a^* = (\lambda + a + a^2 + \ldots)$

Por tanto, si son ER equivalentes