Lenguajes

Alfabeto

Un alfabeto o vocabulario A es un conjunto finito no vacío de símbolos (objetos atómicos o indivisibles).

Ejemplos de alfabetos:

Alfabeto de dígitos decimales $D=\{0,1,2,3,4,5,6,7,8,9\}$;

Alfabeto de dígitos binarios $B=\{0,1\}$

Alfabeto de las caracteres $C=\{a,b,\ldots z,A,\ldots Z, ?,!\ldots,*,\$\}$

Cadena

Una cadena ω es una sucesión finita de símbolos, sobre un alfabeto A.

Una cadena es simplemente representada como $\omega = s_1 s_2 \ldots s_n$ donde $s_1, s_2, \ldots s_n \in A$

El símbolo s_i $1 \le i \le n$, ocurre en la posición i de la cadena.

Por convención, ε denota la cadena vacía (la cadena que no tiene símbolos).

Ejemplo1

Las cadenas $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ sobre $B=\{0,1\}$ se definen como:

$\alpha_1 = 0101$

$\alpha_2 = 1111$

$\alpha_3 = 0111000$

$\alpha_4 = \varepsilon$

Clausura sobre el alfabeto A^*

El conjunto de todas las posibles cadenas sobre un alfabeto A, se describe como A^* también llamado Clausura de Kleene.

$$A^* = \bigcup_{i=0}^{i=\infty} A^i$$

donde A^i es el conjunto de todas las cadenas de longitud i sobre A

En el ejemplo para el alfabeto $B=\{0,1\}$ calcular $B^* =\{0,1\}^*$

$B^0 =\{\varepsilon\}$

$B^1=\{0,1\}$

$B^2=\{ 00,01,10,11\}$

$B^3=\{ 000,001,010,011,100,101,110,111\}$

…

Luego

$B^*=B^0\cup B^1\cup B^2\cup B^3\cup\ldots= \{ \varepsilon, 0,1,00,01,10,11,000,001,010,011,100,101,110,111,\ldots\}$

Nota: Las cadenas $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in B^*$

Operaciones sobre cadenas

Sean dos cadenas sobre el alfabeto A

$\omega_1 = a_1a_2...a_n$ y $\omega_2 = b_1b_2...b_m$ $\omega_1, \omega_2 \in A^*$

- Longitud de la cadena ω_1

$|\omega_1|$ denota la longitud de la cadena ω_1.

$$|\omega_1| = n$$

- Igualdad de cadenas ω_1 y ω_2

$\omega_1 = \omega_2$ si se cumple que $|\omega_1| = |\omega_2|$ y $(\forall i: 1 \leq i \leq n: a_i = b_i)$

- Reversa de la cadena ω_1

ω_1^R denota la reversa de la cadena ω_1

$$\omega_1^R = a_n....a_2a_1$$

- Concatenación de las cadenas ω_1 y ω_2

$\omega_1.\omega_2$ denota la concatenación que consiste de todos los símbolos de ω_1 seguidos por los símbolos de ω_2. (el punto puede omitirse)

$$\omega_1.\omega_2 = a_1a_2...a_nb_1b_2...b_m$$

Propiedades de la concatenación:

1) $\omega_1.\omega_2$ es una cadena sobre A
2) $\omega_1.\varepsilon = \varepsilon.\omega_1 = \omega_1$
3) $|\omega_1.\omega_2| = |\omega_1| + |\omega_2|$
4) No es conmutativa. (puede suceder que para ciertas instancias lo sea)
5) $\omega_1.\omega_1 = \omega_1^2$ (Potencia cuadrada de la cadena ω_1)

- Potencia k-ésima de la cadena ω_1

ω_1^k denota la concatenación de ω_1 con sí misma k-1 veces (o la repetición de ω_1 k veces).

$$\omega_1^k = \omega_1^{k-1}.\ \omega_1 \quad \text{y } \omega_1^0 = \varepsilon \text{ (por convención)}$$

así

$\omega_1^0 = \varepsilon$

$\omega_1^1 = \omega_1$

$\omega_1^2 = \omega_1.\ \omega_1$

…

$\omega_1^k = \omega_1.\ \omega_1.\ \omega_1.......\ \omega_1$ (k-veces)

Ejemplos de operaciones con las cadenas α_1, α_2, α_3, α_4 del ejemplo 1.

Longitud	$|\alpha_1| = |0101| = 4$
	$|\alpha_4| = |\varepsilon| = 0$
Reversa	$\alpha_1^R = 1010$
Concatenación	$\alpha_1.\alpha_2 = 01011111$
	$\alpha_2.\alpha_4 = 1111$
	$\alpha_2.\alpha_2 = 11111111$
Potencia	$\alpha_2^3 = \alpha_2^2.\alpha_2 = 111111111111$

α_1^4=*0101010101010101*

Lenguaje

Un lenguaje L sobre un alfabeto *A* es un subconjunto de *A**, es decir un conjunto de cadenas sobre *A*. $L \subseteq A^*$.

Por ejemplo los siguientes son Lenguajes sobre $B = \{0,1\}$

$L_a = \emptyset$	Lenguaje finito vacío
$L_b = \{\varepsilon\}$	Lenguaje finito que contiene sólo la cadena vacía
$L_c = \{0,1\}$	Lenguaje finito que contiene sólo las cadenas de longitud 1
$L_d = \{0,00,000,0000,\ldots\}$	Lenguaje infinito que consiste de cadenas con cualquier cantidad de símbolos *0*.
$L_e = \{0^n 1^n \ / \ n \geq 1\} = \{01,0011,000111,00001111,\ldots\}$	Lenguaje infinito que consiste de cadenas que comienzan con una cantidad de símbolos *0*, seguidos por la misma cantidad de símbolos *1*.

Operaciones con Lenguajes

Sean dos lenguajes L_1 y L_2 sobre *A*. $L_1 \subseteq A^*$ y $L_2 \subseteq A^*$

- Unión, intersección, diferencia y complemento entre los lenguajes L_1 y L_2. Ya que los Lenguajes son conjuntos de cadenas estas operaciones están implícitamente definidas:

$$L_1 \cup L_2 = \{\omega \in A^* / \omega \in L_1 \text{ o } \omega \in L_2\}$$
$$L_1 \cap L_2 = \{\omega \in A^* / \omega \in L_1 \text{ y } \omega \in L_2\}$$
$$L_1 - L_2 = \{\omega \in A^* / \omega \in L_1 \text{ y } \omega \notin L_2\}$$
$$\overline{L_1} = \{\omega \in A^* / \omega \notin L_1\} = A^* - L_1$$

- Concatenación de los lenguajes L_1 y L_2

$$L_1 . L_2 = \{\omega_1.\omega_2 \in A^* / \omega_1 \in L_1 \text{ y } \omega_2 \in L_2\}$$

Propiedades

Si L_1, L_2, L_3 son lenguajes definidos sobre *A*

$L_1 . \emptyset = \emptyset = \emptyset . L_1$

La concatenación es asociativa

$(L_1 . L_2) . L_3 = L_1 . (L_2 . L_3)$

La concatenación no es conmutativa

$L_1 . L_2 \neq L_2 . L_1$

Distributiva con respecto a la Unión

$L_1 . (L_2 \cup L_3) = L_1 . L_2 \cup L_1 . L_3$

No Distributiva con respecto a la Intersección

$L_1 .(L_2 \cap L_3) \neq L_1 . L_2 \cap L_1 . L_3$

- Potencia del lenguaje L_1

$$L_1^0 = \{\varepsilon\}$$
$$L_1^1 = L_1$$
$$L_1^2 = L_1.L_1$$
$$\ldots$$
$$L_1^k = L_1^{k-1}. L_1$$

- Clausura del lenguaje L_1

$$L_1^* = \bigcup_{i=0}^{i=\infty} L_1^i = L_1^0 \cup L_1^1 \cup L_1^2 \cup L_1^3 \ldots$$

- Reversa del lenguaje L_1

L_1^R denota el Lenguaje reverso de L_1

$$L_1^R = \{ \omega^R \in A^* / \omega \in L_1 \}$$

<u>Ejemplo 2</u>

Dado L_1 y L_2 sobre $A=\{a,b,c\}$

$L_1= \{a^i b^j c^q /q=2i+j \text{ y } i, j \geq 0\}$

$L_2= \{a^i c^{2i} / i \geq 0\}$

Calcular

1) $L_1 \cup L_2$ 2) $L_1 \cap L_2$ 3) $L_1 - L_2$ 4) $L_1.L_2$ 5) $\overline{L_1}$

L_1	$a^i b^j c^{j+2i}$		L_2	$a^i c^{2i}$
i=0 j=0	ε	$\longleftrightarrow$	i=0	ε
i=1 j=0	acc	$\longleftrightarrow$	i=1	acc
i=2 j=0	aacccc	$\longleftrightarrow$	i=2	aacccc
…			…	

i=0 j=1 bc
i=1 j=1 abccc
i=2 j=1 aabccccc
…

i=0 j=2 bbcc
i=1 j=2 abbcccc
i=2 j=2 aabbcccccc
…
…

1) $L_1 \cup L_2 = L_1 = \{a^i b^j c^q /q=2i+j \text{ y } i, j \geq 0\}$

2) $L_1 \cap L_2 = L_2 = \{a^i c^{2i} / i \geq 0\}$

3) $L_1 - L_2 = \{a^i b^j c^q /q=2i+j \text{ y } i \geq 0 \text{ y } j >0\}$

4) $L_1.L_2 = \{a^i\, b^j\, c^q\, a^n\, c^{2n} / q=2i+j \text{ y } i, j, n \geq 0\}$
5) $\overline{L_1} = \{\omega \;/\; \omega \in A^* \text{ y } \omega \neq a^i\, b^j\, c^q \text{ y } q=2i+j \text{ y } i, j \geq 0\}$

Ejemplo 3
Dado L_1 sobre A={*a,b,c,d*} y L_2 sobre A={*c,d,e*}
$L_1 = \{ a^i\, b^i\, c^j\, d^m / i \geq 0 \text{ y } j, m \geq 1\}$
$L_2 = \{ c^k\, d^k\, e^p / k, p \geq 0\}$
Calcular

1) $L_1 \cup L_2$ 2) $L_1 \cap L_2$ 3) $L_1 - L_2$ 4) $L_2 - L_1$ 5) $L_1.L_2$ 6) $\overline{L_1}$

L_1		L_2	
i=0	$c^j\, d^m$	p=0 , k=0	ε
i >0, j>0,m>0	$a^i\, b^i\, c^j\, d^m$	k=0	e^p
		p=0	$c^k\, d^k$
		p >0, k>0	$c^k\, d^k\, e^p$

1) $L_1 \cup L_2 = \{ a^i\, b^i\, c^j\, d^m / i \geq 0 \text{ y } j, m \geq 1\} \cup \{ c^k\, d^k\, e^p / k, p \geq 0\}$
2) $L_1 \cap L_2 = \{c^k\, d^k / k \geq 1\}$
3) $L_1 - L_2 = \{ a^i\, b^i\, c^j\, d^m / i, j, m \geq 1\} \cup \{ c^j\, d^m / j \neq m \text{ y } j, m \geq 1\}$
4) $L_2 - L_1 = \{c^k\, d^k\, e^p / k \geq 0 \text{ y } p>0\} \cup \{\varepsilon\}$
5) $L_1.L_2 = \{ a^i\, b^i\, c^j\, d^m\, c^k\, d^k\, e^p \;/\; i, k, p \geq 0 \text{ y } j, m \geq 1\}$
6) $\overline{L_1} = \{\omega \;/\; \omega \in A^* \text{ y } \omega \neq a^i\, b^i\, c^j\, d^m \text{ y } i \geq 0 \text{ y } j, m \geq 1\}$