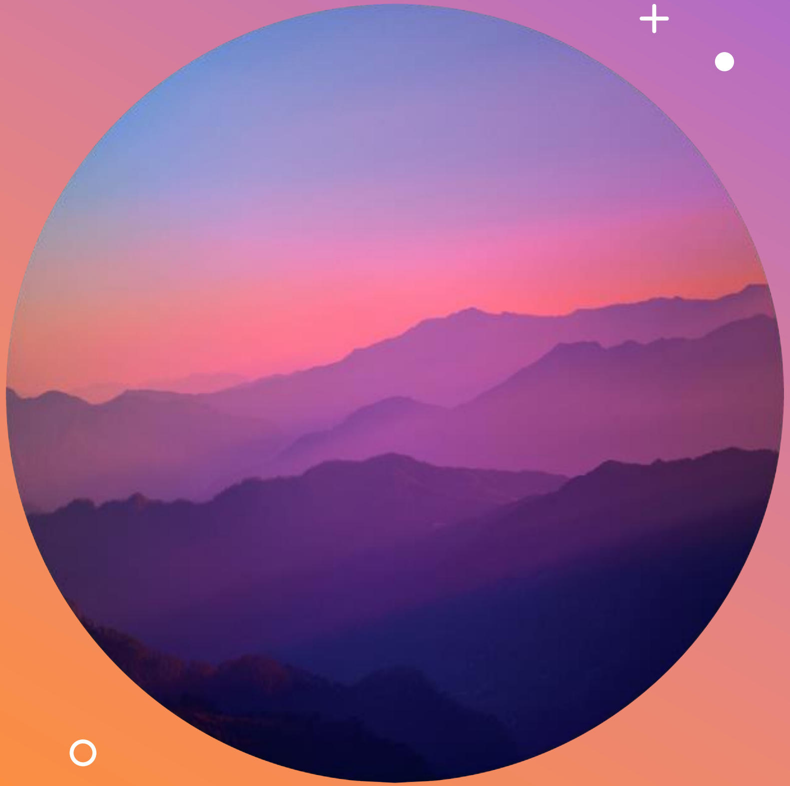
ALFABETOS, SÍMBOLOS Y CADENAS

CONTENIDO



- Alfabetos, símbolos y cadenas.
- Lenguajes
- Operaciones con cadenas.
  - Concatenación de dos cadenas.
  - Prefijos y sufijos de una cadena.
  - Subcadena y subsecuencia.
  - Inversión de una cadena.
  - Potencia de una cadena

Alfabetos, símbolos y cadenas

- Se llama ***alfabeto*** a un conjunto finito, no vacío.
- Los elementos de un alfabeto se llaman ***símbolos***.
- **Un alfabeto se define por la enumeración de los símbolos** que contiene.

$$\Sigma_1 = \{A, B, C, D, E, \ldots, Z\}$$

$$\Sigma_2 = \{0,1\}$$

$$\Sigma_3 = \{0,1,2,3,4,5,6,7,8,9,.\}$$

$$\Sigma_4 = \{/, \backslash\}$$

$$\sum_5 = \{x|x \text{ es una letra minúscula}\}$$

$$\sum_6 = \{x|x \text{ es un dígito decimal}\}$$

Cadena o palabra

- Se llama ***palabra*** o ***cadena*** a aquella formada con los símbolos de un alfabeto.
- *Secuencia finita de símbolos de ese alfabeto.*
- Se utilizarán letras minúsculas ***x*** o ***y*** para representar las cadenas de un alfabeto.

$$x\text{=}JUAN\ (cadena\ sobre\ \Sigma_1 = \{A, B, C, D, E, \ldots, Z\})$$

$$y\text{=}1450\ (cadena\ sobre\ \Sigma_3 = \{0,1,2,3,4,5,6,7,8,9, .\})$$

Longitud de una cadena

- Se llama ***longitud de una cadena*** al número de símbolos que la componen.
- La longitud de la cadena ***x*** se representa con la notación $|x|$.
- La cadena cuya longitud es cero se llama *cadena vacía* y se representa con la letra griega lambda (λ).
- Evidentemente, cualquiera que sea el alfabeto considerado, siempre puede formarse la cadena vacía.

$\Sigma_1 = \{a,b,\ldots,z\}$
$\Sigma_2 = \{la, ba, ca, da\}$

"camisa" tiene longitud 6 sobre Σ_1.

Con símbolos sería:
ω = camisa, $|\omega| = 6$

"cada" tiene longitud 4 sobre Σ_1.

Sin embargo:
"cada" tiene longitud 2, si la consideramos sobre Σ_2.

Lenguaje universal

- El conjunto de todas las cadenas que se pueden formar con las letras de un alfabeto se llama ***lenguaje universal*** de $\boldsymbol{\Sigma}$ y se denota como $\boldsymbol{W(\Sigma)}$.
- Es evidente que $W(\Sigma)$ es un *conjunto infinito*. Incluso en el peor caso, si el alfabeto sólo tiene una letra.

$$\Sigma = \{a\}$$

$$W(\Sigma) = \{\lambda, a, aa, aaa, \ldots\}$$

Lenguaje

- Un lenguaje es un conjunto de palabras.
- Por tanto, el conjunto {1, 12, 123, 1234, 12345, 123456} es un lenguaje sobre el alfabeto compuesto por dígitos.

1. El lenguaje de todas las cadenas que constan de n ceros seguidos de n unos para cualquier $n \geq 0$: $\{\varepsilon, 01, 0011, 000111, \ldots\}$.
2. El conjunto de cadenas formadas por el mismo número de ceros que de unos:

$$\{\varepsilon, 01, 10, 0011, 0101, 1001, \ldots\}$$

OPERACIONES CON CADENAS

Concatenación de dos cadenas

- Sean ***u*** y ***v*** dos cadenas sobre el mismo alfabeto $\boldsymbol{\Sigma}$, la concatenación de ***u*** y ***v*** es una nueva cadena **ω** que se obtiene yuxtaponiendo primero ***u*** y detrás ***v***, escribimos:

$$\omega = uv$$

Ejemplos:

- Sea $u = 01$, $v = 100$ la concatenación de ambas es $\omega = uv = 01100$
- Sea u = mari, v = posa, la concatenación es $\omega = uv$ = mariposa

Propiedades de la concatenación de dos cadenas

1. ***No es conmutativa***, *en general no es lo mismo* ***uv*** *que* ***vu****.*

2. ***Es asociativa***, *es decir cualesquiera que sean las cadenas* ***u****,* ***v*** *y* ***w*** *sobre el mismo alfabeto, se tiene que* ***(uv)w = u(vw)****.*

   - *Esta propiedad nos permite concatenar cualquier número finito de cadenas sin tener que poner los paréntesis. Escribiremos uvw.*

3. ***|uv|=|u|+|v|*** *es decir la longitud de la cadena formada por la concatenación de dos cadenas, es la suma de las longitudes de cada una de ellas.*

4. *La cadena vacía es el elemento neutro de la concatenación.*

   En efecto ***uλ=λu =u****.*

Prefijos y sufijos de una cadena

- Sea $\boldsymbol{\omega}$ una cadena sobre cierto alfabeto $\boldsymbol{\Sigma}$.

- Sean $\boldsymbol{u}$ y $\boldsymbol{v}$ dos cadenas sobre $\boldsymbol{\Sigma}$ tales que $\boldsymbol{\omega=uv}$.

- Decimos que $\boldsymbol{u}$ es un **prefijo** y que $\boldsymbol{v}$ es un **sufijo** de $\boldsymbol{\omega}$.

Un **prefijo de la cadena *S*** es cualquier cadena que se obtiene al eliminar cero o más símbolos del final de ***S***.

Ejemplo:

velo, velocidad y **λ** son prefijos de **ω**=velocidad.

Un **sufijo de la cadena *S*** es cualquier cadena que se obtiene al eliminar cero o más símbolos del principio de ***S***.

Ejemplo:

cidad, velocidad y **λ** son sufijos de **ω**=velocidad.

Subcadena y subsecuencia de una cadena

Una ***subcadena*** de ***S*** se obtiene al eliminar cualquier prefijo y cualquier sufijo de ***S***.

Ejemplo:

velocidad, loci y **λ** son subcadenas de velocidad.

Los **prefijos, sufijos y subcadenas *propios*** de una cadena ***S*** son esos prefijos, sufijos y subcadenas, respectivamente, de ***S*** que no son ***λ*** ni son iguales a la misma ***S***.

Subcadena y subsecuencia de una cadena

Una ***subsecuencia*** de ***S*** es cualquier cadena que se forma mediante la eliminación de cero o más posiciones no necesariamente consecutivas de ***S.***

Ejemplo:

veoci es una subsecuencia de velocidad.

Inversión de una cadena

- Sea $\boldsymbol{\omega}$ una cadena sobre cierto alfabeto $\boldsymbol{\Sigma}$. Llamamos **inversa** (o reflejada) de la cadena $\boldsymbol{\omega}$, y la representamos por $\boldsymbol{\omega}^{-1}$, a la cadena obtenida al escribir los símbolos que constituyen la cadena $\boldsymbol{\omega}$ en orden inverso.
- Si $\boldsymbol{\omega = a_1, a_2, \ldots, a_n}$ su reflejada sería $\boldsymbol{\omega^{-1} = a_n, \ldots, a_2, a_1}$

Ejemplo:

Si, $\boldsymbol{\omega}$ = camisa, entonces $\boldsymbol{\omega}^{-1}$ = asimac

Puede ocurrir que una cadena coincida con su inversa como es el caso de $\boldsymbol{\omega}$=ana; tales cadenas reciben el nombre de ***palíndromos***.

Propiedades de la inversión y la concatenación de cadenas

1. $(\mathbf{uv})^{-1} = \mathbf{v}^{-1}\mathbf{u}^{-1}$ es decir la cadena inversa (o reflejada) de la concatenación de dos cadenas es la concatenación de las cadenas inversas (o reflejadas) en orden contrario

u = hola
v = mundo
uv = holamundo

$\mathbf{u}^{-1}$ = aloh
$\mathbf{v}^{-1}$ = odnum

$(\mathbf{uv})^{-1}$ = odnumaloh
$\mathbf{v}^{-1}\mathbf{u}^{-1}$ = odnumaloh

2. $|\omega^{-1}| = |\omega|$ es decir, la longitud de una cadena y su inversa coinciden siempre.

ω = holamundo
ω^{-1} = odnumaloh

$|\omega^{-1}| = 9$
$|\omega| = 9$

Potencia de una cadena

Sea $\boldsymbol{\omega}$ una cadena y $\boldsymbol{k}$ un número entero, definimos:

$$\omega^k = \begin{cases} \omega \ \ldots^{k)} \ldots \ \omega & si\ k>0 \\ \lambda & si\ k=0 \\ \omega^{-1} \ \ldots^{-k)} \ldots \ \omega^{-1} & si\ k<0 \end{cases}$$

Sea $\boldsymbol{\omega}$ = 91 sobre el alfabeto $\boldsymbol{\Sigma_1}$ **={0,1,...9}**, entonces será:

$\boldsymbol{\omega^3}$ = 919191

$\boldsymbol{\omega^{-1}}$ = 19

$\boldsymbol{\omega^{-2}}$ = 1919

$\boldsymbol{\omega^0} = \lambda$

Sea $\boldsymbol{\omega}$ = *camisa* sobre el alfabeto $\boldsymbol{\Sigma_2}$**={a,b,c,...,z}**, entonces será:

$\boldsymbol{\omega^{-3} = (\omega^{-1})^3} = (asimac)^3 = asimacasimacasimac$