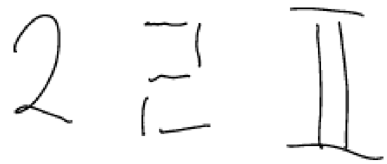
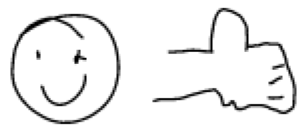
Un símbolo es la representación perceptible de una idea, con rasgos asociados por una convención socialmente aceptada.



Un conjunto no vacío y finito de símbolos se conoce como alfabeto.



Ejemplos:

$\Sigma 1=\{a,b\}=\{b,a\}$
$\Sigma 2=\{0,1\}=\{1,0,0,1\}$

Una secuencia finita de símbolos de un determinado alfabeto se conoce como palabra (cadena).

Ejemplos:

aabbaa es una cadena sobre $\Sigma=\{a,b\}$.

Una cadena puede estar conformada por cualquier cantidad n de símbolos, $n \in \mathbf{N}$.
Si n=0, la cadena se llama cadena vacía y se denota por ε.

Un lenguaje es un conjunto de palabras.

Ejemplos:

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

Son lenguajes:
$L1=\{00,01,10,11\}$
$L2=\{0,1,10,11,100\}$
$L3=\{0,10,100,110,1000,1010,1100,\ldots\}$
$L4=\{\}$
$L5=\{\varepsilon\}$
$L6=\{\varepsilon,0,00,000,0000,00000,\ldots\}$
$L7=\{0,1\}=\Sigma$

Supongamos que tenemos alfabetos $\Sigma 1$ y $\Sigma 2$. Se tiene que $\Sigma 1 \cup \Sigma 2$ también es un alfabeto (y un lenguaje).
$\Sigma 1 \cap \Sigma 2$ es un alfabeto si no es un conjunto vacío. Lo mismo sucede con $\Sigma 1-\Sigma 2$ y $\Sigma 2-\Sigma 1$.

Ejemplo:

Tenemos el lenguaje:

A={ hola, 123, Zócalo}
Σ??
Σ={ h, o, l, a, 1, 2, 3, Z, ó, c }
O también puede ser
Σ= Unicode

B={ a, bb, ccc }
Σ???
Σ={ a, b, c }

C={ 000, 00000, 0000000, …}
Σ???
Σ={0}
Pero es mejor
Σ={0,1}

Operaciones con cadenas

Una cadena w es una subcadena o subpalabra de otra cadena z si existen las cadenas x,y para las cuales z=xwy.

Ejemplos:

z=instituto

w= ns

x= i
y= tituto

xwy=(i)(ns)(tituto)= instituto =z

w= stitu
x= in
y= to
xwy = instituto = z

w= instituto
x= ε
y= ε
xwy = (ε)(instituto)(ε) = instituto

w =insti
x= ε
y= tuto
xwy = (ε)(insti)(tuto)

Toda cadena es subcadena de sí misma

$w=\varepsilon$
x = insti
y = tuto
xwy= (insti)(ε)(tuto) = instituto
La cadena vacía (ε) es subcadena de cualquier cadena.

εiεnεsεtεiεtεuεtεoε

La inversa o transpuesta de una palabra w es la imagen refleja de w.

Ejemplo:

$(escom)^{I}$= mocse

$$w^{I}=\begin{cases} w, \text{ si } w=\varepsilon \\ y^{I}a, \text{ si } w=ay \text{ por tanto } a\in\Sigma \text{ y } y\in\Sigma^* \end{cases}$$

Σ^* es el lenguaje formado por todas las cadenas que se puedan formar usando cualquier número de símbolos de Σ.

$(escom)^{I}= (scom)^{I}e = ((com)^{I}s)e = (((om)^{I}c)s)e = ((((m)^{I}o)c)s)e =(((((\varepsilon)^{I}m)o)c)s)e = (((((\varepsilon)m)o)c)s)e$ = mocse

Si w es una cadena sobre algún alfabeto, su longitud se denota $|w|$ y es el número de símbolos que conforman la cadena.

Ejemplos:
$w=9876$
$|w|=4$

$|\varepsilon|=0$

Si tenemos dos cadenas w,z la concatenación de w con z es la cadena que se obtiene al añadir a la cadena w la cadena z.
Se denota por wz, $(w)(z)$, $w\cdot z$.

Si $w=123$, $z=abc$, $wz=123abc$
$zw=abc123$

$w\varepsilon=\varepsilon w=w$

Obsérvese que $|wz|=|w| + |z|$.

Estrella de Kleene

También conocida como cerradura de Kleene, es el lenguaje formado por todas las cadenas que se pueden formar tomando cualquier número de símbolos de un alfabeto cualquier número de veces. La estrella de Kleene del alfabeto Σ se denota por Σ*.

Ejemplo:

$\Sigma=\{a,b\}$
$\Sigma^*=\{ \varepsilon, a, b, ab, ba, aa, bb, aaa, bbb, aba, abb, aab, baa, bab, bba, \ldots \}$

Potencia de una palabra es una manera de abreviar una concatenación repetida.

Sea w una palabra; para n∊**N** se define:

$$w^n=\begin{cases} \varepsilon, & \text{si } n=0 \\ ww^{n-1}, & \text{si } n>0 \end{cases}$$

Ejemplo:

w=123

$w^0=\varepsilon$
$w^1=ww^0=w\varepsilon= w = 123$
$w^2=ww^1=ww= 123123$

$w^n= \underbrace{wwww\ldots w}_{n}$

Se dice que w^i es la potencia i-ésima de w.

$L=\{$ cadenas de la forma $a^n\} = \{ \varepsilon, a, aa, aaa, aaaa, \ldots \}$
$X=\{$ cadenas de la forma $b^{2n} \} = \{ \varepsilon, bb, bbbb, bbbbbb, \ldots \}$
$Y=\{$ cadenas de la forma $10^{2n+1} \}= \{$ cadenas de la forma $1(0^{2n+1})\} =\{ 10, 1000, 100000, \ldots \}$

Si tenemos dos cadenas w,z, decimos que w es igual a z si tienen la misma longitud y los mismos símbolos en la misma posición. Se denota por w = z.

Si w,x son palabras, se dice que x es prefijo de w, si para alguna cadena z se obtiene que w=xz. Un prefijo propio es un prefijo que no es igual a la cadena completa.

w=escuela

Son prefijos:
x=ε, z=escuela
x=e, z=scuela
x=es, z=cuela

x=esc, z=uela
x=escu, z=ela   prefijos propios
x=escue, z=la
x=escuel, z=a
x=escuela, z=ε   No es prefijo propio porque x=w

Si w,t son palabras, se dice que t es sufijo de w, si para alguna cadena r se obtiene que w=rt. Un sufijo propio es un sufijo que no es igual a la cadena completa.

w=escuela

Son sufijos
t=ε, r=escuela
t=a, r=escuel
t=la, r=escue
t=ela, r=escu   sufijos propios
t=uela, r=esc
t=cuela, r=es
t=scuela, r=e
t=escuela, r=ε   No es sufijo propio porque t=w

Ejercicios:

Sea $\Sigma=\{1\}$.
¿Se puede decir que para todo número natural n hay alguna palabra $w\epsilon\Sigma^*$ para la cual $|w|=n$?
Respuesta: Sí. Esto es porque Σ^* contiene todas las cadenas que se puedan formar con el símbolo 1 con cualquier número de repeticiones, por lo que para cualquier número natural n, se puede formar la cadena $w=1^n\epsilon\Sigma^*$ y $|w|=n$.
¿es única?
Respuesta: Sí, $w=1^n$.
¿Qué ocurriría si $\Sigma=\{1,2\}$?
Respuesta: $\forall n\epsilon N$, $\exists w\epsilon\Sigma^*$ y $|w|=n$, por ejemplo $w=1^n$. Pero, w ya no es única, por ejemplo, también $w=2^n$ cumple. Caso especial, si n=0, sí es única porque solo puede ser ε.

Para una palabra w, ¿se puede decir que

$|w^{i+j}| = |w^i|+|w^j|$?

Respuesta: Sí porque $w^{i+j}= (w^i)(w^j)$ y sabemos que $|(w^i)(w^j)| = |w^i|+|w^j|$
Ejemplo:

$|(ab)^{2+3}| = |(ab)^5|=|$ ababababab $| = 10$
$|(ab)^2|+|(ab)^3|=|$ abab$|$ + $|$ ababab $| = 4 + 6 = 10$

Encontrar una expresión para $|w^{i+j}|$ en términos de i,j y $|w|$.
$|w^{i+j}|= (i+j)\ |w|$

¿La cadena ε es un prefijo propio de sí misma?
Respuesta: No es prefijo propio porque como prefijo es igual a la cadena completa.
w=ε, x=ε, z=ε, x es un prefijo porque xz=w, pero no es prefijo propio porque x=w.

Se definen w=aab, z=babb

Obtener

$wzwaz^2$= (aab)(babb)(aab)a$(babb)^2$ = aabbabbaababababbbabb

$(zbaw)^I$= $(babbbaaab)^I$= baaabbbab

Obtener las 5 cadenas de menor longitud de los siguientes lenguajes:
Obtener las 5 cadenas de menor longitud del complemento de los siguientes lenguajes:

Σ={a,b}
L1={ Cadenas de la forma a^nb^n }
L1_cinco_cadenas_más_pequeñas= { ε, ab, aabb, aaabbb, aaaabbbb }
(Σ*-L1)_cinco_cadenas_más_pequeñas= { a, b, ba, bb, aa }

L2={ Cadenas de la forma a^nb^m }
L2_cinco_cadenas_más_pequeñas= { ε, a, b, ab, aa }
(Σ*-L2)_cinco_cadenas_más_pequeñas= { ba, aba, baa, bab, bba }

L3={ Cadenas que tengan n a's y n b's }
L3_cinco_cadenas_más_pequeñas= { ε, ab, ba, aabb, bbaa }
(Σ*-L3)_cinco_cadenas_más_pequeñas= { a, b, abb, bab, bba }