Tema 2: Lenguajes Formales

Informática Teórica I

Teoría de Lenguajes Formales. Bibliografía

I. Alfonseca, J. Sancho y M. Martínez. “Teoría de Lenguajes,

amáticas y Autómatas”, R.A.E.C., Madrid, (1998).

. Isasi, P. Martínez y D. Borrajo. “Lenguajes, Gramáticas y

tómatas, un Enfoque Práctico”. Addison-Wesley, (1997).

Teoría de Lenguajes Formales. Definiciones

nbolo: entidad abstracta, no se define (análogo al punto en geometría).

n letras, dígitos, caracteres, etc. También posible encontrar símbolos

nados por varios caracteres, pej: IF, THEN, ELSE, ...

fabeto (Σ): conjunto **finito no vacío** de letras o símbolos.

Sea “a” una letra y Σ un alfabeto, si a pertenece a ese alfabeto $\Rightarrow$ a $\in \Sigma$

ejemplos:

Σ_1= {A, B, C, ...,Z}

Σ_2= {0, 1}

Σ_3= {IF, THEN, ELSE, BEGIN, END}

Teoría de Lenguajes Formales. Definiciones

bra, cadena, tira: toda secuencia finita de símbolos del alfabeto.

emplos:

alabras sobre Σ_1 JUAN, ISABEL, etc

alabras sobre Σ_2 00011101

alabras sobre Σ_3 IFTHENELSEEND

epresentan las palabras por letras minúsculas del final del alfabeto

z), pej *x*= JUAN, *y*= IFTHENELSEEND

Teoría de Lenguajes Formales. Definiciones

gitud de palabra:

s el número de símbolos que componen una palabra

a longitud de la palabra *x* se representa por $|x|$

jemplos:

$|x| = |\text{JUAN}| = 4$

$|y| = |\text{IFTHENELSEEND}| = 13$

NO, por la definición del alfabeto es 4

bra vacía λ:

aquella palabra cuya longitud es cero

representa por λ, $|\lambda| = 0$

bre cualquier alfabeto es posible construir λ

lidad: es elemento neutro en muchas operaciones con palabras y

nguajes

Teoría de Lenguajes Formales. Definiciones

iverso del discurso, W(Σ):

- Es el conjunto de todas las palabras que se pueden formar con los símbolos de un alfabeto Σ
- También se denomina Lenguaje Universal de Σ
- Se representa como W(Σ)
- Es un conjunto infinito
- Ejemplo: sea $\Sigma_4 = \{A\}$, $W(\Sigma_4) = \{\lambda, A, AA, AAA, ...\}$ con un número $\propto$ de palabras

COROLARIO:

$\forall\ \Sigma,\ \lambda \in W(\Sigma) \Rightarrow$ La palabra vacía pertenece a todos los lenguajes universales de todos los alfabetos posibles

Teoría de Lenguajes Formales. Definiciones

peraciones con palabras: sobre palabras de un universo del discurso dado

1. Concatenación de palabras

2. Monoide Libre

3. Potencia de una palabra

4. Reflexión de una palabra

Teoría de Lenguajes Formales. Definiciones

raciones con palabras: sobre palabras de un universo del discurso dado

Concatenación de palabras: sean dos palabras *x*, *y* tal que $x \in W(\Sigma)$, $y \in W(\Sigma)$, y sea $|x| = i = l\, x_1 x_2 \ldots x_i|$ y $|y| = j = |y_1 y_2 \ldots y_j|$, se llama concatenación de x con y, a:

$x\,.y = lx_1 lx_2 \ldots lx_i\; ly_1 ly_2 \ldots ly_j = \mathbf{z}$, donde $z \in W(\Sigma)$

Propiedades:

- Operación cerrada
- Propiedad Asociativa
- Con elemento neutro
- No conmutativa

Definiciones:

- cabeza
- cola
- longitud de palabra

Teoría de Lenguajes Formales. Definiciones

raciones con palabras:

. **Monoide Libre** sean

- un alfabeto Σ
- cada símbolo de Σ es una palabra de longitud 1
- aplicando la operación concatenación se puede formar cualquier palabra de W(Σ) excepto λ

➔ el alfabeto Σ es un generador de universos del discurso menos la palabra vacía. **Si se añade λ, entonces W(Σ) es el monoide libre generado por Σ**

- Cumple la ley de cancelación izquierda y derecha:

$\forall$ x,y,z$\in$W(Σ) Si se cumple xy=xz $\Rightarrow$ y=z

Si se cumple xy=zy $\Rightarrow$ x=z

Teoría de Lenguajes Formales. Definiciones

raciones con palabras:

- **Potencia de una palabra:** reducción de la concatenación a los casos que se refieren a una misma palabra
  - potencia *i-ésima* de una palabra al resultado de concatenar esa palabra consigo misma *i* veces
  - concatenación es asociativa $\Rightarrow$ no especificar el orden
  - $x^i = x \,.\, x \,.\, x \,.\, ...x$ *i* veces
  - $|x^i| = i \,.\, |x|$
  - se cumple:

    $x^1 = x$

    $x^{1+i} = x \,.\, x^i = x^i \,.\, x \quad (i>0)$

    $x^{j+i} = x^j \,.\, x^i = x^i \,.\, x^j \quad (i, j>0)$
  - Si se define $x^0 = \lambda$

Teoría de Lenguajes Formales. Definiciones

peraciones con palabras:

4. Reflexión de una palabra Sea la palabra $x= A_1A_2A_3...A_n$, se denomina palabra refleja de x, $x^{-1} = A_n... A_3A_2A_1$

- Formada por los mismos símbolos en distinto orden
- $|x^{-1}| = |x|$

Teoría de Lenguajes Formales. Definiciones

uaje, L: Se denomina lenguaje sobre el alfabeto Σ

todo subconjunto del lenguaje universal de Σ, $L \subset W(\Sigma)$

todo conjunto de palabras sobre un determinado Σ

- son lenguajes especiales:
  - ϕ = Lenguaje vacío, $\phi \subset W(\Sigma)$
  - **{λ}** = Lenguaje de la palabra vacía

    se diferencian en el número de palabras (cardinalidad) que los forman $C(\phi) = 0$ mientras que $C(\{\lambda\})=1$

    se parecen en que ϕ y $\{\lambda\}$ son lenguajes sobre cualquier alfabeto
  - Un alfabeto es uno de los lenguajes generados por el mismo: $\Sigma \subset$ **$W(\Sigma)$**, por ejemplo el chino

Teoría de Lenguajes Formales. Definiciones

peraciones con lenguajes: sobre un alfabeto dado

1. **Unión de lenguajes**
2. **Concatenación de lenguajes**
3. **Binoide Libre**
4. **Potencia de un lenguaje**
5. **Clausura o cierre positivo de un lenguaje**
6. **Iteración, clausura o cierre de un lenguaje**
7. **Reflexión de lenguajes**

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes: sobre un alfabeto dado

Unión de lenguajes

- Sean L_1 y L_2 definidos sobre el mismo alfabeto, L_1, $L_2 \subset W(\Sigma)$, se llama **unión** de dos lenguajes, L_1, L_2 y se representa por $L_1 \cup L_2$ al lenguaje así definido:

$$\mathbf{L_1 \cup L_2 = \{x \,/\, x \in L_1 \text{ O } x \in L_2\}}$$

- Es el conjunto formado indistintamente por palabras de uno u otro de los dos lenguajes (equivale a la suma)

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes:

Unión de lenguajes: Propiedades:

- Operación cerrada
- Propiedad Asociativa
- Con elemento neutro
- Conmutativa
- Idempotente

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes: sobre un alfabeto dado

Concatenación de lenguajes

- Sean L_1 y L_2 definidos sobre el mismo alfabeto, $L_1, L_2 \subset W(\Sigma)$, se llama **concatenación o producto** de dos lenguajes, L_1 y L_2 y se representa por $L_1 \mathbf{.} L_2$ al lenguaje así definido:

  $$\mathbf{L_1.L_2 = \{xy \,/ x \in L_1 \text{ AND } y \in L_2\}}$$

- Es el conjunto de palabras formado por la concatenación de palabras de L_1 con palabras de L_2
- Definición válida para lenguajes con algún elemento.
- Con el lenguaje vacío: $\phi \mathbf{.} L = L \mathbf{.} \phi = \phi$

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes:

2. Concatenación de lenguajes

Propiedades:

- Operación cerrada
- Propiedad Asociativa
- Con elemento neutro
- Propiedad distributiva respecto

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes:

Binoide Libre

La concatenación (monoide) de lenguajes y la unión (monoide) de lenguajes constituyen un binoide

- Los símbolos de Σ se pueden considerar conjuntos de una sola palabra
- Con Σ, la unión y la concatenación se puede formar cualquier lenguaje sobre dicho Σ

➔ el alfabeto Σ es un conjunto de generadores para el conjunto L

⇒ **L es el BINOIDE LIBRE generado por Σ**

Teoría de Lenguajes Formales. Definiciones

eraciones con lenguajes:

Potencia de un lenguaje

- reducción de la concatenación a los casos que se refieren a un mismo lenguaje
- potencia *i-ésima* de un lenguaje al resultado de concatenar ese lenguaje consigo mismo *i* veces
- concatenación es asociativa $\Rightarrow$ no especificar el orden
- $L^i = L \,.\, L \,.\, L \,.\, \ldots.\, L$ *i* veces
- Se define $L^1 = L$
- se cumple:

$$L^{1+i} = L.L^i = L^i.L \quad (i>0)$$

$$L^{j+i} = L^i.L^j \quad (i, j > 0)$$

- Si se define $L^0 = \{\lambda\}$

$(i \geq 0)$

$(i, j \geq 0)$

Teoría de Lenguajes Formales. Definiciones

Operaciones con lenguajes: sobre un alfabeto dado

5. Clausura o cierre positivo de un lenguaje

- Se define como L^{+} y es el lenguaje obtenido uniendo el lenguaje L con todas sus potencias posibles excepto L^{0}

$$\mathbf{L^{+} = U_{i=1}^{i=\infty} L^{i}}$$

- Ninguna clausura positiva contiene a λ, si $\lambda \notin L$
- Como Σ es un lenguaje sobre Σ, la clausura positiva de Σ será:

$$\mathbf{\Sigma^{+} = {}_{i=1}^{i=\infty} \Sigma^{i} = W(\Sigma) - \{\lambda\}}$$

Teoría de Lenguajes Formales. Definiciones

Operaciones con lenguajes:

6. Iteración, clausura o cierre de un lenguaje

- Se define como L^* y es el lenguaje obtenido uniendo el lenguaje L con todas sus potencias posibles.
- *** es el operador unario de Kleene**

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

Toda clausura contiene a λ,

- Se cumple:
  - $L^* = L^+ \cup \{\lambda\}$
  - $L^+ = L^*.L = L.L^*$
- Como Σ es un lenguaje sobre Σ, se le puede aplicar el *:

$\Sigma^* = W(\Sigma)$ → **El lenguaje universal es Σ^***

Teoría de Lenguajes Formales. Definiciones

raciones con lenguajes:

Reflexión de lenguajes

- Se llama lenguaje reflejo o inverso de L y se representa por L^{-1} al lenguaje:

  $$\mathbf{L^{-1} = \{x^{-1} / x \in L\}}$$

  es decir, al lenguaje formado por todas las palabras reflejas de L

Teoría de Lenguajes Formales. Ejercicios

o el alfabeto $\Sigma = \{0,1\}$ escribir el lenguaje formado por palíndromos
Alfonseca: pag 30 ejercicios 1, 2 y 3
nir 2 lenguajes L1 y L2 de cardinalidad 3 y después realizar con
las siguientes operaciones:

- $L_1 \cup L_2$
- $L_1 . L_2$
- L_1^2
- L_1^*
- L_2^+
- $(L_1.L_2)^{-1}$

sasi, Martínez y Borrajo: ejercicios 2.1, 2.2 y 2.3

Teoría de Lenguajes Formales. Definiciones

xto válido, cv: conjunto de prefijos y sufijos que hacen que una pertenezca a un lenguaje.

$L \in \Sigma^*$ y x una palabra cualquiera que no tiene por qué pertenecer

e dice que el par de palabras $u, v \in \Sigma^*$ es un contexto válido de x n L si se cumple: $u . x . v \in L$

as tres palabras por separado no tienen por qué pertenecer a L

i (u,λ) es un contexto válido *de x en L,* se dice que u es un **prefijo álido** de x

i (λ,v) es un contexto válido *de x en L,* se dice que v es un **sufijo álido** de x

Teoría de Lenguajes Formales. Definiciones

xto válido, cv:

nplo: sea Σ = {0,1} y L = {u / $|u|$ = 4}. Sea x = 01 e y= 0101 dos
bras sobre Σ.
ecir cuáles de los siguientes pares son contextos válidos de *x* y de
en L.
0), (0,1), (1,0), (1,1), (λ,00), (λ,01), (λ,10), (λ,11), (00, λ), (01, λ),
0, λ), (11, λ)

Teoría de Lenguajes Formales. Definiciones

exto válido, RELACIONES DE EQUIVALENCIA:

ueden definir dos relaciones de equivalencia entre los elementos , S_L y P_L .

- Se dice **x P_L y** si x e y tienen el mismo conjunto de prefijos válidos en L
- Se dice **x S_L y** si x e y tienen el mismo conjunto de sufijos válidos en L
- En las relaciones de equivalencia **P_L** y **S_L** se cumple:

  Sea **x S_L y,** se cumple **x.z S_L y.z** $\forall z \in \Sigma^*$

  Sea **x P_L y,** se cumple **x.z P_L y.z** $\forall z \in \Sigma^*$

z). u $\in$ L $\Leftrightarrow$ x.(z. u) $\in$ L $\Leftrightarrow$ y .(z. u) $\in$L $\Leftrightarrow$ (y.z).u$\in$ L $\Rightarrow$ **x.z S_L y.z**

Teoría de Lenguajes Formales. Definiciones

exto válido, RELACIONES DE EQUIVALENCIA:

plos **x.z S_L y.z**

Los verbos regulares cantar y saltar:

están en relación **S_L**: cantar **S_L** saltar

si les añadimos un sufijo siguen en relación **S_L**:

cantara **S_L** saltara,

cantaramos **S_L** saltaramos, etc

Teoría de Lenguajes Formales. Definiciones

ducciones, reglas de escritura o reglas de derivación:

Sea Σ un alfabeto

Se llama producción a un **par ordenado** (*x*,*y*) donde *x*,*y* ∈ Σ*

Se dice que *x* es la parte izquierda de la producción e *y* la parte derecha de la producción

Se representa como: x::= y

Teoría de Lenguajes Formales. Definiciones

ación directa:

- Sea Σ un alfabeto
- Sea (*x*,*y*) una producción sobre palabras de ese Σ, x::=y
- Sean v y w dos palabras sobre Σ (v, w $\in\Sigma^*$)

Se dice que
- "w es derivación directa de v"
- "v produce directamente w"
- "w se reduce directamente a v"

$v \to w$

si $\exists$ dos palabras z, u $\in\Sigma^*$ tales que v = z.x.u y w = z.y.u

COROLARIO: Si x::= y es una producción sobre Σ:

$x ::= y \Rightarrow x \to y$ (una regla de escritura es una derivación directa)

Teoría de Lenguajes Formales. Definiciones

ción directa, ejemplos:

- Sea Σ el alfabeto castellano de las letras mayúsculas y
- ME::=BA una producción sobre Σ
- CABALLO es derivación directa de CAMELLO (CAMELLO
- produce directamente CABALLO)
- on la producción CA::=PE sobre Σ PERA es derivación
- lirecta de CARA

En el castellano no son así las cosas, existen las raíces.

Teoría de Lenguajes Formales. Definiciones

ción directa en un conjunto de producciones:

Sea Σ un alfabeto y P un conjunto de producciones sobre Σ

Sean v y w dos palabras sobre Σ, v, w $\in \Sigma^*$

Se dice que
- "w es derivación directa de v"
- "v produce directamente w"
- "w se reduce directamente a v"

$v \rightarrow w$

dos palabras z, u $\in \Sigma^*$ tales que v = z.x.u y w = z.y.u y se

umple **(x::=y)** $\in$ **P**

Teoría de Lenguajes Formales. Definiciones

vación directa en un conjunto de producciones, ejemplo:

Alfonseca página 32: sea el alfabeto $\Sigma = \{0,1,2,N,C\}$ y el conjunto de producciones sobre dicho alfabeto,

P = {N::CN, N::=C, C::=0, C::=1, C::=2}. Escribir todas las derivaciones directas asociadas a ese conjunto de producciones.

Teoría de Lenguajes Formales. Definiciones

ación:

Sea Σ un alfabeto y P un conjunto de producciones sobre Σ

Sean v y w dos palabras sobre Σ, v, w $\in\Sigma^*$

Se dice que $\left\{ \begin{array}{l} \text{“w es derivación de v”} \\ \text{“v produce w”} \\ \text{“w se reduce a v”} \end{array} \right\}$ $v \overset{+}{\rightarrow} w$

si $\exists$ una secuencia finita de palabras, $u_0, u_1, u_2, \ldots u_n$ (n>0) $\in\Sigma^*$ tales que $v = u_0$

$$\left. \begin{array}{l} u_0 \rightarrow u_1 \\ u_1 \rightarrow u_2 \\ \ldots\ldots\ldots \\ u_n \rightarrow w \end{array} \right\}$$

Derivación de longitud n

Teoría de Lenguajes Formales. Definiciones

ción, ejemplo:

ea el alfabeto $\Sigma = \{0,1,2,N,C\}$ y el conjunto de producciones obre dicho alfabeto, P = {N::CN, N::=C, C::=0, C::=1, C::=2}

omprobar e indicar la longitud de la derivación

$N \overset{+}{\rightarrow} 210$

COROLARIO: si $v \rightarrow w$ entonces $v \overset{+}{\rightarrow} w$ mediante una derivación de longitud 1

Teoría de Lenguajes Formales. Definiciones

ación de Thue:

Sea Σ un alfabeto y P un conjunto de producciones sobre Σ

Sean v y w dos palabras sobre Σ, v, w ∈Σ*

Se dice que existe una relación de Thue entre v y w y se representa por **v *→ w** si se verifica que:

- **v +→ w**
- **v = w**

o ∃ una derivación de longitud n o son iguales

Cumple las propiedades:

- Reflexiva
- Simétrica (en general NO se cumple)
- Transitiva