Teoría de Autómatas y Lenguajes Formales

Capítulo 2: "Lenguajes Formales"

Holger Billhardt
holger.billhardt@urjc.es

Universidad Rey Juan Carlos

Sumario:

- Capítulo 2: "Lenguajes Formales"
  1. Concepto de Lenguaje Formal
  2. Operaciones sobre Lenguajes Formales (u otros conjuntos)

Sumario:

- Capítulo 2: “Lenguajes Formales”

  1. **Concepto de Lenguaje Formal**
  2. Operaciones sobre Lenguajes Formales (u otros conjuntos)

Concepto de Lenguaje Formal

- ¿Qué es un lenguaje?

  - Informalmente: un lenguaje es un conjunto de palabras o sentencias formadas sobre un alfabeto

- Pasaremos a definirlo de manera formal.

Concepto de Lenguaje Formal

- Alfabeto:
  - **Definición (Alfabeto):**
    - Conjunto finito, no vacío, de elementos.

    Generalmente usaremos Σ para especificar alfabetos y los elementos los denominaremos “letras” o “símbolos”.
    - **Ejemplos:**
      - los alfabetos español, inglés, o alemán
      - $\Sigma_1=\{0,...,9\}$, $0 \in \Sigma 1$
      - $\Sigma_2=\{x \mid x$ es un símbolo del código ASCII$\}$
      - $\Sigma_3=\{(,)\}$
      - $\Sigma_4=\{1, A, 2, B\}$
      - $\Sigma_5=\{a, b, c, d\}$
      - $\cancel{\Sigma_6=\{\}}$
      - $\cancel{\Sigma_7=\aleph}$

Concepto de Lenguaje Formal

- Palabras:
  - **Definición (Palabra):**
    - Sea un alfabeto Σ. Una *palabra* sobre Σ es una secuencia finita de las letras de ese alfabeto.

    La *secuencia vacía* representa la *palabra vacía* y la anotamos con λ.
    - **Ejemplos:**

      sobre Σ_5 ={a,b,c,d}:

      λ, a, b, c, d, abc, aab, dcba, ...

      sobre Σ_1 ={0,...,9}:

      λ, 0, 0000, 010, 9980, ...

      sobre Σ_3 ={(,)}

      λ, (,), (), (()()),)())), ...

Concepto de Lenguaje Formal

- Palabras:
  - **Definición (Longitud de una palabra):**
    - Se llama *longitud de una palabra* x, y se representa por |x|, al número de símbolos que la componen.
    - **Ejemplos:**

      sobre Σ_5 ={a,b,c,d}:

      |λ|=0,

      |a|=1,

      |abc|=3

Concepto de Lenguaje Formal

- Operaciones con palabras:
  - **Definición (Concatenación):**
    - Sean dos palabras x e y definidas sobre el alfabeto Σ. La *concatenación* de x e y, denominada "xy", es una palabra que contiene todos los símbolos (de derecha a izquierda) de x seguidos de los símbolos de y (de derecha a izquierda).

      Sean $x=A_1A_2...A_n$ e $y=B_1B_2...B_m$ con $A_i, B_i \in \Sigma$:

      $\Rightarrow xy= A_1A_2...A_nB_1B_2...B_m$
    - **Ejemplos:**

      x =abc, y =da, definidos sobre Σ={a,b,c,d}

      xy=abcda ; |xy|=|x|+|y|=5

Concepto de Lenguaje Formal

- Operaciones con palabras:
  - **Propiedades de la concatenación:**
    - Operación cerrada: sí

      Si x e y están definidos sobre Σ, entonces xy está definido sobre Σ.
    - asociativa: sí

      x(yz)=(xy)z
    - Elemento nulo: λ

      xλ=λx=x
    - Conmutatividad: no

      xy≠yx

Concepto de Lenguaje Formal

- Operaciones con palabras:
  - **Definición (Potencia):**
    - Sea i un número natural, y x una palabra. La *potencia i-ésima* de x, denominada x^i, es la operación que consiste en concatenarla consigo misma i veces.
    - **Ejemplos:**

      x =abc ⇒ x^1=abc

      x^2=abcabc

      x^3=abcabcabc

Concepto de Lenguaje Formal

- Operaciones con palabras:
  - **Propiedades de la potencia:**

    $\forall$ i, j > 0
    - $x^{i+1}=xx^i=x^ix$
    - $x^ix^j=x^{i+j}$
      - Se define $x^0=\lambda$ (*palabra vacía*):

        Si i=0 $\Rightarrow x^{0+1}=x^1=x=x\lambda=xx^0=\lambda x=x^0x$

        Si i,j=0 $\Rightarrow x^ix^j=x^0x^0=\lambda\lambda=\lambda=x^0=x^{0+0}$
      - **Nota:** $\lambda\lambda=\lambda$; $\lambda x=x$; $\lambda\lambda x\lambda=x$
    - $|x^i|=i\cdot|x|$

Concepto de Lenguaje Formal

- Operaciones con palabras:
  - **Definición (Palabra inversa):**
    - Sea $x=A_1A_2...A_n$ con $A_i\in\Sigma$ una palabra sobre el alfabeto Σ. Se llama *palabra refleja* o *inversa* de x, y se representa por x^{-1}, a la palabra $A_nA_{n-1}...A_1$. Si $x=\lambda$ entonces $x^{-1}=\lambda$.
    - **Ejemplos:**

      x =abc $\Rightarrow$ x^{-1}=cba
  - **Propiedades de la palabra inversa:**
    - $|x^{-1}|=|x|$

Concepto de Lenguaje Formal

- Lenguajes Formales:
  - **Definición (Lenguaje universal):**
    - Sea Σ un alfabeto. El *lenguaje universal* de Σ es el conjunto formado por todas las palabras que se pueden formar con las letras de Σ. Representamos dicho lenguaje con $W(\Sigma)$.
    - **Ejemplos:**
      - $\Sigma_1 = \{a\}$ $\Rightarrow$ $W(\Sigma_1)=\{\lambda, a, aa, aaa, ...\}$
  - **Nota:** La palabra vacía pertenece a todos los lenguajes universales de todos los alfabetos posibles.

Concepto de Lenguaje Formal

- Lenguajes Formales:
  - **Definición (Lenguaje):**
    - Sea un alfabeto Σ. Un *lenguaje* L sobre Σ es cualquier subconjunto del lenguaje universal $W(\Sigma)$.
    - **Ejemplos:**

$\Sigma_1 = \{a\} \Rightarrow W(\Sigma_1)=\{\lambda, a, aa, aaa, ...\}$
$L_1 = \{a\} \subseteq W(\Sigma_1)$
$L_2 = \{\} \subseteq W(\Sigma_1)$ $(L_2 = \emptyset)$
$L_3 = \Sigma_1 \subseteq W(\Sigma_1)$
$L_4 = W(\Sigma_1) \subseteq W(\Sigma_1)$
$L_5 = \{\lambda\} \subseteq W(\Sigma_1)$ (Nota: $L_5 \neq L_2$)
$L_6 = \{\lambda, a, aaa, aaaaa\} \subseteq W(\Sigma_1)$
$L_7 = \{\lambda, a, aaa, aaaaa, ...\} \subseteq W(\Sigma_1)$

  - Hay lenguajes finitos, infinitos y vacíos.

Sumario:

- Capítulo 2: "Lenguajes Formales"

1. Concepto de Lenguaje Formal
2. **Operaciones sobre Lenguajes Formales (u otros conjuntos)**

Operaciones con lenguajes (u otros conjuntos)

- **Unión:**
  - **Definición (Unión de lenguajes):**
    - Sea el alfabeto Σ y dos lenguajes $L_1 \subseteq W(\Sigma)$ y $L_2 \subseteq W(\Sigma)$. La *unión* de L_1 y L_2, $L_1 \cup L_2$, es un lenguaje que se define de la siguiente forma:

      $$L_1 \cup L_2 = \{x | x \in L_1 \text{ o } x \in L_2\}.$$
  - **Propiedades de la unión:**
    - Operación cerrada: $L_1 \subseteq W(\Sigma)$, $L_2 \subseteq W(\Sigma) \Rightarrow L_1 \cup L_2 \subseteq W(\Sigma)$

      (la unión de dos lenguajes sobre el mismo alfabeto es también un lenguaje sobre este alfabeto)
    - Asociativa: $(L_1 \cup L_2) \cup L_3 = L_1 \cup (L_2 \cup L_3)$
    - Elemento neutro: $\forall L_1$, $N \cup L_1 = L_1$ **¿Que es N?**
    - Conmutativa: $L_1 \cup L_2 = L_2 \cup L_1$
    - Idempotencia: $L \cup L = L$

Operaciones con lenguajes (u otros conjuntos)

- **Concatenación:**
  - **Definición (Concatenación de lenguajes):**
    - Sean dos lenguajes L_1, L_2. La *concatenación* de L_1 y L_2, representado por L_1L_2 (a veces por $L_1.L_2$), es un lenguaje que se define de la siguiente forma: $L_1L_2=\{xy \mid x\in L_1, y\in L_2\}$.
  - **Ejemplos:** $\Sigma=\{a,b,c\}$

    $L_1=\{ab, ac, cb\}$; $L_2=\{b, bba\} \Rightarrow L_1L_2=\{abb,abbba,acb,acbba,cbb,cbbba\}$

    $L_1=\{a, aa, aaa, ...\}$; $L_2=\{\lambda, b, bb, bbb, ...\} \Rightarrow L_1L_2=$**¿?**

    ¿Qué pasa si L_1 o L_2 es $\varnothing$?
  - **Propiedades de la concatenación**
    - Cerrada: $L_1\subseteq W(\Sigma), L_2\subseteq W(\Sigma) \Rightarrow L_1L_2\subseteq W(\Sigma)$
    - Asociativa: $(L_1L_2)L_3 = L_1(L_2L_3)$
    - No es conmutativa: $\neg(\forall L_1, L_2: L_1L_2=L_2L_1)$
    - Elemento neutro($\{\lambda\}$): $\forall L_1: L_1\{\lambda\}=\{\lambda\}L_1=L_1$
    - No es idempotente: $\neg(\forall L: LL=L)$

Operaciones con lenguajes (u otros conjuntos)

- **Potencia de un lenguaje:**
  - **Definición (Potencia de un lenguaje):**
    - La *potencia i-ésima* de un lenguaje L consiste en el lenguaje resultante de concatenar el lenguaje consigo mismo i veces.

      $L^i = LLL...L$ (i veces)
  - **Propiedades de la potencia**
    - Cerrada: $L \subseteq W(\Sigma) \Rightarrow L^i \subseteq W(\Sigma)$
    - $L^{i+1} = L^iL = LL^i$ (i>0)
    - $L^iL^j = L^{i+j}$ (i,j>0)

  ¿Que pasa si i, j = 0?
    - Se define $L^0 = \{\lambda\}$

      $L^{0+1} = L^1 = L = \{\lambda\}L=L^0L$

      $L^0L^0= \{\lambda\}\{\lambda\} =\{\lambda\}=L^0 = L^{0+0}$

Operaciones con lenguajes (u otros conjuntos)

- **Potencia de un lenguaje:**
  - **Ejemplos:**

$L_1 = \{\lambda, ab, ac\}$
$\Rightarrow L_1^2=\{\lambda,ab,ac,abab,abac,acab,acac\}$
$\Rightarrow L_1^3=\{\lambda,ab,ac,abab,abac,acab,acac,ababab,ababac,$
$abacab,abacac,acabab,acabac,acacab,acacac\}$

$L_2 = \{a, aa, aaa, ...\}$
$\Rightarrow L_2^2=$¿?
$\Rightarrow L_2^3=$¿?

Operaciones con lenguajes (u otros conjuntos)

- **Clausura de un lenguaje**
  - **Definición (Clausura positiva):**
    - La *clausura positiva* de un lenguaje L se define por:
      $L^+= \bigcup_{i=1}^{\infty} L^i$
  - **Ejemplos:**
    - $L =\{a,aa,aaa,aaaa,...\} = \{a^n \mid n\geq 1\}$
      $\Rightarrow L_2=\{aa,aaa,aaaa,...\} = \{a^n a^m \mid n,m\geq 1\} = \{a^n \mid n\geq 2\}$
      $\Rightarrow L_3=\{aaa,aaaa,...\} = \{a^n a^m \mid n\geq 1, m\geq 2\} = \{a^n \mid n\geq 3\}$
      $\Rightarrow L^+= \bigcup_{i=1}^{\infty} L^i=\{a,aa,aaa,aaaa,...\} = L$
    - $\Sigma=\{a,b\}$, Σ es un lenguaje sobre Σ, ya que $\Sigma\subseteq W(\Sigma)$
      $\Sigma^+= \bigcup_{i=1}^{\infty} \Sigma^i=\{a,b,aa,ab,ba,bb,aaa,...\} = W(\Sigma) - \{\lambda\}$
  - **Nota:** Si $\lambda\notin L$, entonces $\lambda\notin L^+$.

Operaciones con lenguajes (u otros conjuntos)

- **Definición (Clausura, Iteración o cierre):**
  - La *clausura* de un lenguaje L se define por:
    $L^* = \bigcup_{i=0}^{\infty} L^i$
- **Nota:** $\forall$ L: $\lambda \in L^*$, ya que $\{\lambda\} = L^0$.

- **Propiedades de la clausura:**

  - Cerrada: $L \subseteq W(\Sigma) \Rightarrow L^+ \subseteq W(\Sigma)$, $L^* \subseteq W(\Sigma)$
  - $L^* = L^0 \cup (\bigcup_{i=1}^{\infty} L^i) = L^0 \cup L^+ = \{\lambda\} \cup L^+$
  - $L^+ = LL^* = L^*L$
    - Demostración¿?

Operaciones con lenguajes (u otros conjuntos)

- Reflexión de un lenguaje

- **Definición (Reflexión):**
  - Sea L un lenguaje. Se llama *lenguaje inverso (lenguaje reflejo)* de L, y se representa por L^{-1} al lenguaje: $L^{-1} = \{x^{-1} | x \in L\}$.

- **Ejemplos:**
  - L ={ana,julio,jesus,norma} $\Rightarrow$ L^{-1}={ana, oiluj,susej,amron}
  - L ={a,aa,aaa,...} $\Rightarrow$ ¿L^{-1}?

- **Propiedades de la reflexión:**
  - Cerrada: $L \subseteq W(\Sigma) \Rightarrow L^{-1} \subseteq W(\Sigma)$

Operaciones con lenguajes (u otros conjuntos)

- Otras operaciones clásicas de conjuntos
  - **Definición (Intersección):**
    - Sean dos lenguajes L_1 y L_2. La *intersección* de L_1 y L_2, $L_1 \cap L_2$, es el lenguaje que se define por:
    - $L_1 \cap L_2 = \{x | x \in L_1 \text{ y } x \in L_2\}$.
  - **Propiedades de la intersección**
    - Cerrada: $L_1 \subseteq W(\Sigma)$, $L_2 \subseteq W(\Sigma) \Rightarrow L_1 \cap L_2 \subseteq W(\Sigma)$
    - Asociativa: $(L_1 \cap L_2) \cap L_3 = L_1 \cap (L_2 \cap L_3)$
    - Conmutativa: $L_1 \cap L_2 = L_2 \cap L_1$
    - Idempotencia: $L \cap L = L$
    - $L \cap \varnothing = \varnothing$

Operaciones con lenguajes (u otros conjuntos)

- Otras operaciones clásicas de conjuntos
  - **Definición (Complemento):**
    - Sea L un lenguaje sobre el alfabeto Σ. El *complemento* de L, denotado con $\overline{L}$ (o con c(L)) es el siguiente lenguaje: $\overline{L} = \{x | x \in W(\Sigma) \text{ y } x \notin L\}$
  - **Propiedades del complemento**
    - Cerrada: $L \subseteq W(\Sigma) \Rightarrow \overline{L} \subseteq W(\Sigma)$
    - $\overline{W(\Sigma)} = \varnothing$
    - $\overline{\overline{L}} = L$

Operaciones con lenguajes (u otros conjuntos)

- Otras operaciones clásicas de conjuntos
  - **Definición (Diferencia):**
    - Sean dos lenguajes L_1 y L_2. La *diferencia* de L_1 y L_2, $L_1 - L_2$ (o $L_1 \backslash L_2$) es el lenguaje que se define por:

      $L_1 - L_2 = \{x | x \in L_1 \text{ y } x \notin L_2\}$.
  - **Propiedades de la diferencia**
    - Cerrada: $L_1 \subseteq W(\Sigma)$, $L_2 \subseteq W(\Sigma) \Rightarrow L_1 - L_2 \subseteq W(\Sigma)$
    - No es asociativa: $\neg(\forall L_1, L_2: (L_1-L_2)-L_3 = L_1-(L_2-L_3))$
    - No es conmutativa: $\neg(\forall L_1, L_2: L_1-L_2 = L_2-L_1)$
    - No es idempotente: $\forall L: L-L=\varnothing$
    - $A-\varnothing=A$

Operaciones con lenguajes (u otros conjuntos)

- Otras leyes de las operaciones sobre conjuntos
  - **Leyes de De Morgan:**
    - $\overline{L_1 \cap L_2} = \overline{L_1} \cup \overline{L_2}$
    - $\overline{L_1 \cup L_2} = \overline{L_1} \cap \overline{L_2}$
  - **Leyes de complemento:**
    - $L \cap \overline{L} = \varnothing$
    - $L \cup \overline{L} = W(\Sigma)$
  - **Distributividad:**
    - $L_1 \cup (L_2 \cap L_3) = (L_1 \cup L_2) \cap (L_1 \cup L_3)$
    - $L_1 \cap (L_2 \cup L_3) = (L_1 \cap L_2) \cup (L_1 \cap L_3)$
  - $L_1 - L_2 = L_1 \cap \overline{L_2} = \overline{\overline{L_1} \cup L_2}$
  - $\overline{L} = W(\Sigma) - L$