

Alfabetos y lenguajes

Alfabetos, palabras y lenguajes

Fabio Martínez Carrillo

Autómatas
Escuela de Ingeniería de Sistemas e Informatica
Universidad Industrial de Santander - UIS

29 de agosto de 2017



Agenda

1 Alfabetos, Palabras y Lenguajes

2 Jerarquia de Máquinas y gramáticas

Autómatas y complejidad

Los autómatas son esenciales para el estudio de los límites de la computación

decibilidad: los problemas que una computadora puede resolver

tratabilidad: los problemas que una computadora puede resolver en un tiempo proporcional a alguna función que crezca lentamente con el tamaño de la entrada (polinómica)

Lenguajes y Gramaticas

Lenguajes Naturales

Existen reglas gramaticas que condicionan de cierto modo la forma de evolucionar un lenguaje. Son flexibles y dependientes de entornos sociales, culturales, etc.

Lenguaje de Maquina o Programación

Tienen estructuras claramente definidas y determinadas por las reglas gramaticales.

- Traductores automaticos (compiladores)
 - **Analizador de léxico:** tranforma el código en una secuencia individual de contenidos
 - **Analizador sintáctico** verifica que las secuencias individuales estén bien escritas
 - **Generador de código**

Alfabeto

Un alfabeto es un conjunto **finito** de símbolos finito y no vacío

- $\Sigma = \{0, 1\}$ alfabeto binario
- $\Sigma = \{a, b, \dots, z\}$ conjunto letras minúsculas
- Conjunto de los caracteres ASCII

Palabra: Cadenas de caracteres

Secuencia finita de símbolos seleccionados de algún alfabeto.

- 01101 es una cadena del alfabeto binario

Palabra: Cadenas de caracteres

Secuencia finita de símbolos seleccionados de algún alfabeto.

- 01101 es una cadena del alfabeto binario

Cadena Vacía ϵ

Cero apariciones de símbolos

Palabra: Cadenas de caracteres

Secuencia finita de símbolos seleccionados de algún alfabeto.

- 01101 es una cadena del alfabeto binario

Longitud de una cadena ω : $|\omega|$

El número de posiciones ocupadas por símbolos dentro de la cadena

- 01101 Longitud de 5

Potencias de un alfabeto

Conjunto de cadenas de una determinada Longitud. Se utiliza la notación exponencial:

- Σ^k Alfabeto Σ con el conjunto de cadenas de longitud k
- Si $\Sigma^k = \{0, 1\}$ entonces $\Sigma^2 = \{00, 01, 10, 11\}$
- Si $\Sigma^k = \{0, 1\}$ entonces $\Sigma^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$

- Todas las cadenas de un alfabeto:

$\Sigma^* = \{\epsilon, 0, 1, 00, 11, 01, 10, 000, 101, \dots\}$ Expresado como:

$$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \dots$$

- Conjunto de cadenas no vacías: $\Sigma^+ = \Sigma^1 \cup \Sigma^2 \cup \dots$
- $\Sigma^* = \Sigma^+ \cup \{\epsilon\}$

Concatenación de cadenas

Concatenación

Sean x y y dos cadenas, entonces xy es el proceso de concatenación. Si x es una cadena de i simbolos y y es una cadena de j simbolos. Entonces $xy = i + j = a_1 a_2 \cdots a_i b_1 b_2 \cdots b_j$

Ejemplo

Sea $x = 01101$ y $y = 110$. Cual es la concatenación xy y yx ?

Concatenación de cadenas

Concatenación

Sean x y y dos cadenas, entonces xy es el proceso de concatenación. Si x es una cadena de i símbolos y y es una cadena de j símbolos. Entonces $xy = i + j = a_1 a_2 \cdots a_i b_1 b_2 \cdots b_j$

Iteración de concatenación de una cadena en Σ^*

- $u^0 = \varepsilon$
- $u^{i+1} = u^i u, \forall i \geq 0$

Concatenación de cadenas

Concatenación

Sean x y y dos cadenas, entonces xy es el proceso de concatenación. Si x es una cadena de i símbolos y y es una cadena de j símbolos. Entonces $xy = i + j = a_1 a_2 \cdots a_i b_1 b_2 \cdots b_j$

Propiedades

- $|u, v| = |u| + |v|, \forall u, v \in \Sigma^*$
- *Asociativa:* $u \cdot (vw) = (uv) \cdot w, \forall u, v, w \in \Sigma^*$
- *Elemento neutro:* $u\varepsilon = \varepsilon u = u, \forall u \in \Sigma^*$

Cadenas

Reflexión o inverso

$$u^R = \begin{cases} \varepsilon & \text{Si } u = \varepsilon \\ a_n \dots a_1 & \text{Si } u = a_1 \dots a_n \end{cases}$$

Subcadenas, Prefijos y Sufijos

- Una cadena v es **subcadena** de u si $u = xvy$. x, y son subcadenas también.
- Un **prefijo** de u es una cadena v tal que $u = vw$. Prefijo propio si $u \neq v$
- Un **sufijo** de u es una cadena v tal que $u = wv$. Sufijo propio si $u \neq v$

Dado $\sigma = \{a, b, c, d\}$ y $u = bcbaadb$. Cuales son los posibles prefijos y sufijos de u

Lenguajes

Un conjunto de cadenas Σ^* , donde Σ es un determinado alfabeto se denomina **lenguaje**.

- Si Σ es un alfabeto y $L \subseteq \Sigma^*$. Entonces L es un lenguaje de Σ

Analogía con otros lenguajes

- Lenguaje natural: El español esta conformado por un subconjunto de palabras conformado por las letras
- lenguaje de Programación: un subconjunto de palabras formada por caracteres ASCII

Ejemplos de lenguajes en Automatas

- Conjunto de cadenas binarias con el mismo número de ceros y unos:

Ejemplos de lenguajes en Automatas

- Conjunto de cadenas binarias con el mismo número de ceros y unos:

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

- Conjunto de cadenas binarias cuyo valor es un primo:

Ejemplos de lenguajes en Automatas

- Conjunto de cadenas binarias con el mismo número de ceros y unos:

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

- Conjunto de cadenas binarias cuyo valor es un primo:

$\{10, 11, 101, 111, 1011\}$

- $\{\varepsilon\}$ lenguaje unicamente con una cadena vacia

Lenguajes como conjuntos

$$\{w \mid \text{predicado de } w\}$$

Conjunto de palabras w tal que el predicado

- $\{w \mid w \text{ entero binario primo} \}$

Ejemplos

- $\{0^n 1^n \mid n \geq 1\}$
- $\{0^i 1^j \mid 0 \leq i \leq j\}$
- $\{a^{n^2} \mid n = 1, 2, 3, \dots\}$

Propiedades sobre los lenguajes

Si L_1 y $L_2 \subseteq \Sigma^*$

- $L\emptyset = L\emptyset = \emptyset$
- $\{\varepsilon\}L = L\{\varepsilon\} = L$
- $L_1(L_2L_3) = (L_1L_2)L_3$

Aplican todas las propiedades de conjuntos como por ejemplo la Unión, intersección, diferencia y complemento.

Concatenación de lenguajes

La concatenación de dos lenguajes A y B sobre Σ , es notada como $A \cdot B$:

$$AB = \{uv : u \in A, v \in B\}$$

Dado $\Sigma = \{a, b, c\}$, $A = \{a, ab, ac\}$, $B = \{b, b^2\}$

- Cual es la concatenación AB y BA ?
- Son iguales?
- Cual es la concatenación AB y BA para $A = \{ba, bc\}$ y $B = \{b^n : n \geq 0\}$

Iteración (Potencias) de lenguajes

$$L^0 = \varepsilon$$


$$L^{i+1} = LL^i$$

Clausura de Kleene

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

$$L^+ = \bigcup_{i \geq 1} L^i$$

Algunas propiedades de la Clausura de Kleene

1. $A^+ = A^* \cdot A = A \cdot A^*$.
2. $A^* \cdot A^* = A^*$.
3. $(A^*)^n = A^*$, para todo $n \geq 1$.
4. $(A^*)^* = A^*$.
5. $A^+ \cdot A^+ \subseteq A^+$.
6. $(A^*)^+ = A^*$.
7. $(A^+)^* = A^*$.
8. $(A^+)^+ = A^+$. 

Demuestre la primera propiedad.

Cual es el inverso de un lenguaje?

Algunas propiedades

1. $(A \cdot B)^R = B^R \cdot A^R.$
2. $(A \cup B)^R = A^R \cup B^R.$
3. $(A \cap B)^R = A^R \cap B^R.$
4. $(A^R)^R = A.$
5. $(A^*)^R = (A^R)^*.$
6. $(A^+)^R = (A^R)^+.$

Cual es el inverso de un lenguaje?

$$L^{-1} = \{u \mid u^{-1} \in L\}$$

Algunas propiedades

1. $(A \cdot B)^R = B^R \cdot A^R.$
2. $(A \cup B)^R = A^R \cup B^R.$
3. $(A \cap B)^R = A^R \cap B^R.$
4. $(A^R)^R = A.$
5. $(A^*)^R = (A^R)^*.$
6. $(A^+)^R = (A^R)^+.$

Homorfismo de dos lenguajes

$$h : \Sigma_1^* \rightarrow \Sigma_2^*$$

- $h(a_1, a_2, \dots, a_n) = h(a_1)h(a_2)h(a_3) \dots h(a_n)$

Sea $\Sigma_1^* = \{0, 1, 2, \dots, 9\}$ y $\Sigma_2^* = \{0, 1\}$

- $h(0) =$
- $h(4) =$
- $h(9) =$

Problemas y Lenguajes

En un problema se decide si una determinada cadena pertenece o no a un determinado **lenguaje**

Ejemplos

- Buscar si una determinada cadena pertenece al lenguaje de primos L_p
- Analizador sintactico en el lenguaje C. Decidir sobre una cadena ASCII

Lenguajes Regulares

Son todos los lenguajes que se pueden formar a partir de los lenguajes básicos: $\{\emptyset\}$, $\{\varepsilon\}$, $\{a\}$, $a \in \Sigma$

Si A y B son lenguajes regulares:

También son lenguajes regulares:

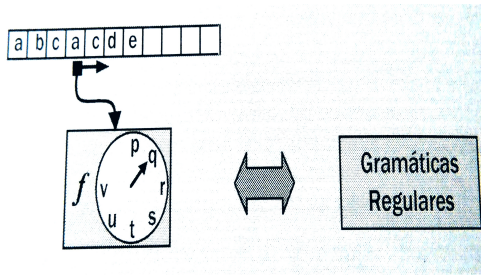
- $A \cup B$
- $A \cdot B$
- A^*

Agenda

1 Alfabetos, Palabras y Lenguajes

2 Jerarquía de Máquinas y gramáticas

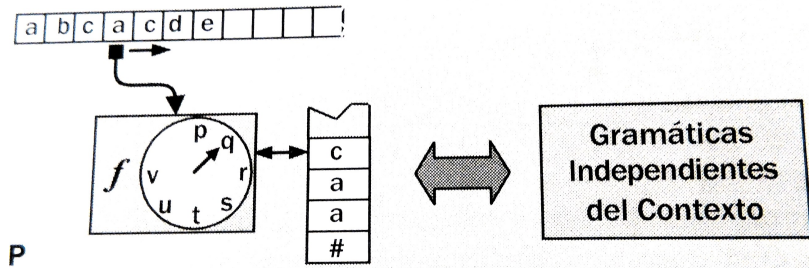
Automata finito y su gramatica asociada



Automata

- un conjunto de estados p, q, r, \dots
- función de transición
- Alfabeto de entrada a, b, c
- cabezal de lectura que se mueve en un sentido para leer secuencialmente

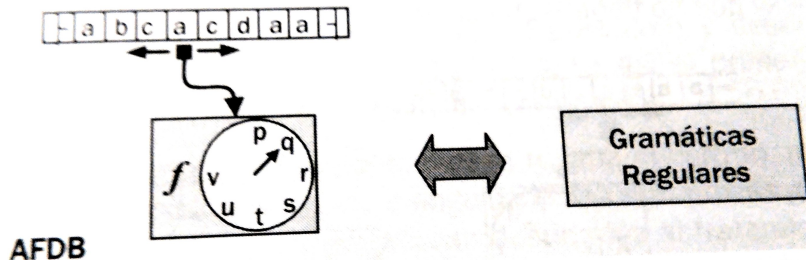
Automata con pila y su gramatica independiente del contexto



Automata

- incorpora una memoria LIFO
- simbolo para definir el fondo de la pila #

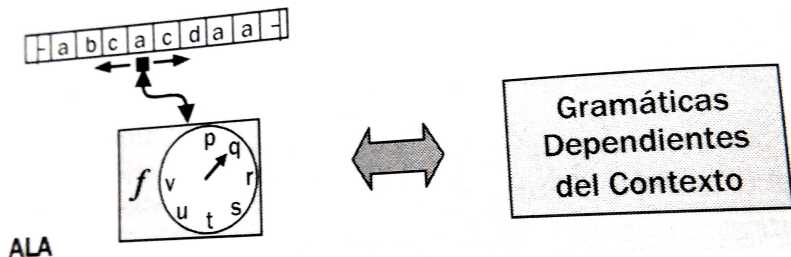
Automata finito birideccional y gramaticas regulares



Automata

- Símbolos secuenciales que limitan los caracteres de entrada

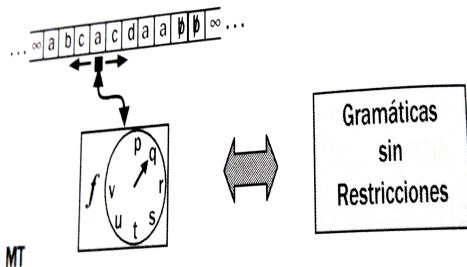
Automata linealmente acotado y gramaticas dependiente del contexto



Automata

- Graba sobre la misma cinta de entrada
- la función de estado define el próximo paso

Maquina de Turing y Gramatica sin restricciones



Automata

- Graba sobre la misma cinta de entrada
- la función de estado define el próximo paso

Muchas gracias por su atención

