Teoría de Automatas

Diego Soto - Universidad Austral De Chile

28 de abril de 2025

Índice

1. Introducción

La teoría de autómatas constituye uno de los pilares fundamentales de la computación teórica y el análisis formal de lenguajes. Esta disciplina permite modelar y analizar el comportamiento de sistemas computacionales mediante representaciones abstractas conocidas como autómatas. A través del estudio de estos modelos, es posible comprender mejor cómo funcionan los lenguajes formales, los compiladores, y las máquinas que los procesan.

Para comenzar, definiremos y repasaremos algunos conceptos que son previamente sabidos del ramo Estructuras Discretas y luego de a poco, nos iremos adentrando en materia.

2. Alfabetos y Lenguajes

Cuando hablamos comunmente de un lenguaje, se nos puede venir a la distintas ideas, como palabras, letras, etcétera. Comenzaremos dando algunas definiciones de estas, para poder empezar a contruir otras fokin definiciones.

Definición 1 *Un Alfabeto es un conjunto finito de símbolos.*

$$\Sigma = \{a, b, c, ..., z\}$$

Definición 2 *Una Palabra es una secuencia ordenada de símbolos de un alfabeto.*

$$x = a_1 a_2 a_3 ... a_4$$

Definición 3 *El largo de una palabra es el número de símbolos que los conforman.*

$$|x| = n$$

Definición 4 *La palabra vacía es una palabra que no tiene ningún símbolo.*

$$\textit{Palabra vacía: } \varepsilon \Rightarrow |\varepsilon| = 0$$

Definición 5 *Si Σ es un alfabeto, anotaremos como Σ^* el conjunto de todas las palabras posibles sobre Σ.*

$$\Sigma^* = \bigcup_{k \geq 0} \Sigma^k = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup ...$$

En donde Σ^k es el conjunto de palabras de largo k.

Ejemplo:

$$\text{Si } \Sigma = \{a, b\}, \text{ entonces:}$$

$$\Sigma^0 = \{\varepsilon\}$$

$$\Sigma^1 = \{a, b\}$$

$$\Sigma^2 = \{aa, ab, ba, bb\}$$

$$\cdots$$

$$\Sigma^n = \{\epsilon, a, b, aa, ab, ba, bb, aaa, \dots\}$$

Definición 6 *Un lenguaje es un conjunto de palabras que se pueden armar sobre un alfabeto* Σ*, el cual denotaremos por la letra* L

$$\textit{Con esto, tenemos que } L \subseteq \Sigma^*$$

Ejemplo:

- $L = \{x \in \{a, b\}^* \mid \text{x comienza con 'a'}\}$

  $L = \{a, aa, ab, aab, aba, \dots\}$

- $L = \{x \in \{0, 1\}^* \mid \text{x es palindromo'}\}$

  $L = \{0, 1, 00, 11, 101, 010, 11011, 0110, \dots\}$

- $L = \{x \in 6^+ \mid \text{Los dígitos de x suman 10}\}$

  $L = \{64, 22222, 235, \dots\}$ (Notar que las palabras acá son finitas)

Entonces, podemos decir que existen lenguajes que pueden ser finitos, o infinitos, donde los lenguajes infinitos son aquellos que son compuestos por infinitas palabras que pertenecen al lenguaje, mientras que en los finitos no.

En Σ^* se definen las siguientes operaciones:

- Concatenacion:

$$\text{Sean } x = a_1a_2 \dots a_n,\ y = b_1b_2 \dots b_n, \text{ entonces:}$$

$$x \cdot y = xy = a_1a_2 \dots a_nb_1b_2 \dots b_n$$

- Reflexión:

$$\text{Si } x = a_1a_2 \dots a_n, \text{ entonces:}$$

$$x^r = a_n \dots a_2a_1$$

$$x^r \text{ se denomina "Palabra Refleja"}$$

3. Operaciones Con Lenguajes

4. Recursividad Izquierda no inmediata

Por el proceso anterior se elimina toda recursividad izquierda inmediata, no así la recursividad obtenida con derivaciones de mas de un paso. Por ejemplo en la gramática

$$S \to Aa \mid b$$

$$A \to Ac \mid Sd \mid e$$

Tenemos $S \Rightarrow Aa \Rightarrow Sda$

$$S \Rightarrow^* Sda$$

Lema 1 *Toda L.L.C que no contiene ε puede ser generado por una G.L.X sim recursividad izquierda.*

Dem. 1 *Sea $G = (V, T, P, S)$ una G.L.C que genera L.*

1. *Elimina los ciclos con el procedimiento visto para verificar una gramática.*
2. *Ordenar las variables $V = \{A_1, A_2, \ldots, A_n\}$ de tal forma que $A_i \to A_j\alpha$ es una producción solo si $i < j$.*

   Si $A_k \to A_j\alpha$ con $j < k$, con el procedimiento de sustitución se reemplaza en $A_k \to A_j\alpha$ por las $A_j - producciones$. Se repite este proceso.
3. *Eliminar recursividad izquierda inmediata.*

Se repite 1, 2 y 3 hasta que no quede recursividad izquierda.

Ejemplo:

$$S \to Aa \mid b$$

$$A \to Ac \mid Sd \mid e$$

1. No hay ciclos
2. $V = \{S, A\}$, $S < A$ y $A \to Sd$ con $S < A$

   Se reemplaza por las $S - producciones$

   $S \to Aa \mid b$

   $A \to Ac \mid Aad \mid bd \mid e$
3. $S \to Aa \mid b$

   $A \to bdA' \mid bd \mid eA' \mid e$

   $A' \to cA' \mid c \mid adA' \mid ad$

   $V = S, A, A'$

Ejercicio: Elimina la recursividad izquierda.

$S \to aRb \mid Rb$

$R \to ca \mid Taa$

$T \to Ra$

1. No hay ciclos
2. $V = S, R, T$

   $S \to aRb \mid Rb$

   $R \to ca \mid Taa$

   $T \to caa \mid Taaa$
3. $S \to aRb \mid Rb$

   $R \to ca \mid Taa$

   $T \to caaT' \mid caa$

   $T' \to aaaT' \mid aaa$

   $V = S, R, T, T'$, $T \to cRST$

5. Forma normal de Greibach

Una G.L.C $G = (V, T, P, S)$ se dice qie está en Forma Normal de greibach (F.N.G) si sus producciones toenen la misma forma:

$$A \to a\alpha \text{ con } A \in V,\ a \in T,\ \alpha \in V^*$$

Teorema 1 *Toda L.L.C que no contiene ε puede ser generado por una G.L.C. en su F.N.G.*

Dem. 2 *El algoritmo para trnasformar una G.L.C a F.N.G es:*

1. *Ramifica la gramática y eliminar toda recursividad izquierda.*
2. *Para cada producción de la forma $A \to \alpha_1\alpha_2 \ldots \alpha_k$ con $k \geq 2$, si α_i, $i \geq 2$ es un terminal, se reemplaza por una variable auxiliar X_{α_i} y se agrega la producción $X_{\alpha_i} \to \alpha_i$*
3. *Se ordenan las variables $V = A_1, A_2, \ldots, A_m$ de modo que necesariamente $A_m \to a\beta$ con $a \in T$. Por el procedimiento de sustitución, se puede sustituir en las producciones $A_i \to A_j\beta$ con $i < j$ por las $A_j -$producciones, de modo que hacer aparecer un terminal al comienzo del lado derecho, partiendo de $j = m$ hacia atrás.*

Ejemplo:

$$S \to abA \mid Sa$$

$$A \to Bb \mid cc$$

$$B \to ac \mid ba$$

1. Es positiva

2. Está ramificada

   $V = S, A, B$ (no hay Recursividad izquierda no inmediata")

   Eliminamos Recursividad izquierda inmediata"

   $S \to abA \mid Sa$

   $S' \to aS' \mid a$

   $A \to Bb \mid cc$

   $B \to ac \mid ba$

3. $S \to aX_bAS' \mid aX_bA$

   $S' \to aS' \mid a$

   $A \to BX_b \mid cX_c$

   $B \to aX_c \mid bX_a$

   $X_a \to a$

   $X_b \to b$

   $X_c \to c$

4. $A \to aX_cX_b \mid bX_aX_b \mid cX_c$