

Ciencias de la Computación II

Gramáticas Regulares



Eduardo Contrera Schneider

Universidad de la Frontera

27 de septiembre de 2016

- 1 Introducción
- 2 Reglas de Sustitución
- 3 Gramáticas Regulares
- 4 Gramáticas y Autómatas

Introducción

Dentro de las ciencias de la computación existen distintas formas de especificar o definir un lenguaje. Hasta este punto ya hemos visto como hacerlo a través de expresiones regulares y autómatas finitos, ya sea deterministas o no. En lo que sigue, presentaremos una nueva forma de especificación a través de gramáticas regulares.

Reglas de Sustitución

Consideremos el AFN $M = (\Sigma = \{a, b\}, Q = \{q_1, q_2, q_3, q_4, q_5\}, s = q_1, F = \{q_5\}, \Delta)$ con Δ especificado por la siguiente tabla:

Δ	a	b	ϵ
q_1	$\{q_2\}$	\emptyset	\emptyset
q_2	\emptyset	\emptyset	$\{q_3, q_4\}$
q_3	$\{q_3\}$	$\{q_5\}$	\emptyset
q_4	$\{q_4\}$	$\{q_5\}$	\emptyset
q_5	\emptyset	\emptyset	\emptyset

Podemos formar cadenas recorriendo desde el estado inicial hasta el estado final de aceptación de M . El lenguaje aceptado por M es $a(a^* \cup b^*)b$.

Cuando recorremos el AFN con una secuencia (o camino) de estados obtenemos una cadena como salida. Por ejemplo, si tomamos el camino $q_1 - q_2 - q_3 - q_3 - q_3 - q_5$ obtenemos la cadena aa^2b . Cuando se va de un estado a otro, la salida es el símbolo que etiqueta dicho camino. De esa manera, las cadenas estarán formadas por un símbolo, a en este caso, seguidas de una parte final. Establezcamos un simbolismo para hacer lo anterior.

Simbolismo

Si hacemos que E representa la parte final de la cadena, se puede escribir $S \rightarrow aE$, interpretando la flecha como "se compone de". E estará formado por lista de aes o una lista de bes, de modo que para indicar las diferentes posibilidades para E se escribe $E \rightarrow A$ y $E \rightarrow B$. Cada una de las listas se puede representar como $A \rightarrow aA$ y $A \rightarrow b$ para indicar que siguen más aes o una b . Del mismo modo, podemos escribir $B \rightarrow bB$ y $B \rightarrow b$.

En resumen tenemos:

- $S \rightarrow aE$.
- $E \rightarrow A$.
- $E \rightarrow B$.
- $A \rightarrow b$.
- $A \rightarrow aA$.
- $B \rightarrow b$.
- $B \rightarrow bB$.

Estas expresiones pueden ser consideradas como **reglas de sustitución** para la generación de cadenas, ya que el símbolo que se encuentra a la izquierda de la flecha puede ser sustituido por la cadena de la derecha.

La cadena aab se genera a partir de S de la siguiente forma

$$S \Rightarrow aE \Rightarrow aaA \Rightarrow aab$$

donde la parte final es representada S, A o E e interpretando \Rightarrow como *produce* o *genera*. Si además agregamos el símbolo $|$ que se interpreta como 'o', entonces podemos escribir:

- $S \rightarrow aE$.
- $E \rightarrow A|B$.
- $A \rightarrow aA|b$.
- $A \rightarrow bB|b$.

Gramáticas Regulares

Formalmente damos la siguiente definición:

Gramáticas Regulares

Una **gramática regular** G es una 4-tupla $G = (\Sigma, N, S, P)$, donde Σ es un alfabeto, N es una colección de símbolos no terminales, S es un no terminal llamado *símbolo inicial*, y P es una colección de reglas de sustitución, llamadas *producciones*, y que son de la forma $A \rightarrow w$, donde $A \in N$ y w es una cadena sobre $\Sigma \cup N$ que satisface lo siguiente:

- 1 w contiene un no terminal como máximo.
- 2 Si w contiene un no terminal, entonces es el símbolo que está en el extremo derecho de w .

El lenguaje generado por la gramática regular G se denota por $L(G)$.

Características

- N es una colección de símbolos en mayúscula (por convención) para representar porciones de cadena que aún no han sido generadas.
- Una vez terminada la generación, la cadena estará formada sólo por símbolos de Σ .
- Los pasos intermedios siempre tendrán cadenas del tipo $\Sigma^*(N \cup \epsilon)$.
- Los símbolos de N se llaman no terminales y deben ser sustituidos.
- Los símbolos de Σ se llaman terminales y no es posible sustituirlos.
- El símbolo inicial siempre es un elemento de N .

Ejemplos

- ❶ Obtenga una gramática regular para el lenguaje $a^*b \cup a$.
- ❷ La gramática regular dada por
 - $S \rightarrow bA|aB|\epsilon$.
 - $A \rightarrow abaS$.
 - $B \rightarrow babS$.

genera un lenguaje regular. ¿Qué lenguaje es?