Teoría de la Computación

Víctor Melchor

Clase 7

Teoría de Autómatas

Definición: Sean $A, B \in \Sigma_N$. Nos referiremos a (A, B) como el par unitario si verifica $A \overset{*}{\to} B$ empleando sólo producciones unitarias.

Método de Eliminación de Producciones Unitarias

Sea $G^1 = (\Sigma_N, \Sigma_T, S, \mathcal{P}^1)$ una GLC con producciones unitarias. Construiremos una GLC $G^2 = (\Sigma_N, \Sigma_T, S, \mathcal{P}^2)$ sin producciones unitarias mediante:

1. Hállese Σ_N
2. Para cada $A \in \Sigma_N$ determine $Unit(A)$
3. Para cada par unitario (A, B) añada a $\mathcal{P}^2$ todas las producciones $A \to w$, donde $B \to w$ es una producción no unitaria de $\mathcal{P}^1$.

Observe que $A = B$ es posible, luego $\mathcal{P}^2$ contiene todas las prod no unitarias de $\mathcal{P}^1$

4. Elimine todas las producciones unitarias de $\mathcal{P}^1$

Ejemplo: Sea G la GLC dada por:

$$
\begin{aligned}
S &\rightarrow A|Aa \\
A &\rightarrow B \\
B &\rightarrow C|b \\
C &\rightarrow D|ab \\
D &\rightarrow b
\end{aligned}
$$

Eliminar todas las producciones unitarias de G y dar la gramática equivalente.

Solución:

1. $\Sigma_N = \{S, A, B, C, D\}$

2. $$\begin{aligned} Unit(S) &= \{S, A, B, C, D\} \\ Unit(A) &= \{A, B, C, D\} \\ Unit(B) &= \{B, C, D\} \\ Unit(C) &= \{C, D\} \\ Unit(D) &= \{D\} \end{aligned}$$

3.

Par Unitario	Producción no Unitaria
(S,S)	$S \to Aa$
(S,A)	——
(S,B)	$S \to b$
(S,C)	$S \to ab$
(S,D)	$S \to b$
(A,A)	——
(A,B)	$A \to b$
(A,C)	$A \to ab$
(A,D)	$A \to b$
(B,B)	$B \to b$
(B,C)	$B \to ab$
(B,D)	$B \to b$
(C,C)	$C \to ab$
(C,D)	$C \to b$
(D,D)	Db

4. Eliminamos las P. Unitarias de $\mathcal{P}^1$

$$S \to A, \quad A \to B, \quad B \to C, \quad C \to D$$

5. La gramática G^2 resultante es:

$$\mathcal{P}^2 \begin{cases} S \to b|ab|Aa \\ A \to b|ab \\ B \to ab|b \\ C \to b|ab \\ D \to b \end{cases}$$

Capítulo 19

Forma Normal de Chomsky

La forma de las gramáticas libres de contexto es extremadamente general. Surge naturalmente una pregunta: ¿hay alguna forma en la que podamos restringir la sintaxis de las gramáticas libres de contexto sin reducir su poder expresivo?.

La forma normal de Chomsky presenta la gramática de una manera muy restringida, lo que facilita considerablemente ciertas pruebas sobre el lenguaje generado.

Definición: Una GLC $G = (\Sigma_T, \Sigma_N, S, P)$ está en la forma Normal de Chomsky (FNC) si

G no contiene variables inútiles.

G no contiene producciones unitarias.

G no contiene producciones ε

Todas las producciones son de la forma:

$$
\begin{aligned}
S &\to \varepsilon && \text{si } \varepsilon \in L(G) \\
A &\to BC && \text{donde } B, C \in \Sigma_N \\
A &\to \sigma && \text{donde } \sigma \in \Sigma_T
\end{aligned}
$$

Teorema: Sea G una GLC. Existe una GLC en la FNC equivalente a G.

Método de Conversión

Para llevar una GLC a su FNC debemos hacer las siguientes simplificaciones:

1. Eliminar las variables inútiles.
2. Eliminar las producciones ε (salvo $S \to \varepsilon$)
3. Eliminar las producciones unitarias

4. Toda producción de tal gramática tendrá la forma:

$$\begin{aligned} A &\to a \qquad \text{con } a \in \Sigma_T \\ A &\to w \qquad \text{con } |w| \geq 2 \end{aligned}$$

Luego la tarea será:

(a) Disponer que todas las cadenas de longitud mayor o igual que 2 consistan solo de variables.

(b) Descomponer todas las cadenas de longitud mayor o igual que 3 en una cascada de producciones, cada una con un cuerpo formado por dos variables.

Para tal efecto, la construcción es como sigue:

Debemos descomponer todas las producciones de la forma:

$$A \rightarrow B_1 B_2 \cdots B_k \quad \text{para} \quad k \geq 3$$

en un grupo de producciones con dos variables en cada cuerpo.

Agregamos $(k-2)$ nuevas variables: $\mathcal{Z}_1, \mathcal{Z}_2, \ldots, \mathcal{Z}_{k-2}$

La producción original es reemplazada por las $(k-1)$ producciones:

$$\begin{aligned} A &\rightarrow B_1 \mathcal{Z}_1 \\ \mathcal{Z}_1 &\rightarrow B_2 \mathcal{Z}_2 \\ &\cdots \\ \mathcal{Z}_{k-3} &\rightarrow B_{k-2} \mathcal{Z}_{k-2} \\ \mathcal{Z}_{k-2} &\rightarrow B_{k-1} B_k \end{aligned}$$

Ejemplo: Dada la producción $A \rightarrow abBaC$

reemplazarla con producciones simples y binarias

Solución

Creamos las nuevas variables T_a y T_b y las producciones:

$$T_a \to a$$

$$T_b \to b$$

Así, la producción original se reemplaza por:

$$\begin{cases} A \to T_a T_b B T_a C \\ T_a \to a \\ T_b \to b \end{cases}$$

Definimos las nuevas variables $\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_3$ y las producciones binarias necesarias, obteniendo:

$$\begin{cases} A & \to T_a Z_1 \\ Z_1 & \to T_b Z_2 \\ Z_2 & \to B Z_3 \\ Z_3 & \to T_a C \\ T_a & \to a \\ T_b & \to b \end{cases}$$

Ejemplo: Reemplazar la regla $A \to BAaCbb$ con producciones simples y binarias.

Solución

Agregamos las variables T_a y T_b y las reglas

Luego se obtiene

$$\begin{cases} A & \to BAT_aCT_bT_b \\ T_a & \to a \\ T_b & \to b \end{cases}$$

Definimos las variables Z_1, Z_2, Z_3 y Z_4 y las producciones binarias necesarias

Se obtiene:

$$\begin{cases} A & \to BZ_1 \\ Z_1 & \to AZ_2 \\ Z_2 & \to T_aZ_3 \\ Z_3 & \to CZ_4 \\ Z_4 & \to T_bT_b \\ T_a & \to a \\ T_b & \to b \end{cases}$$

Ejemplo: Sea la GLC G dada por

$$\begin{cases} S & \to AB|aBC|SBS \\ A & \to aA|C \\ B & \to bbB|b \\ C & \to cC|\varepsilon \end{cases}$$

convertirla a su Forma Normal Chomsky.

1. Identificar las variables anulables

$$\mathcal{N} = \{C, A\}$$

C deriva directamente en ε. A deriva en ε

2. Eliminar las producciones ε

Antes de eliminar las producciones ε, agregamos

$$S \to B|aB$$
$$A \to a$$
$$C \to c$$

Se obtiene la gramática equivalente G^1

$$G^1 \begin{cases} S & \rightarrow AB|aBC|SBS|B|aB \\ A & \rightarrow aA|C|a \\ B & \rightarrow bbB|b \\ C & \rightarrow cC|c \end{cases}$$

3. Identificamos las variables derivables de A, para cada A

$$Unit(S) = \{S, B\}$$
$$Unit(A) = \{A, C\}$$
$$Unit(B) = \{B\}$$
$$Unit(C) = \{C\}$$

4. Eliminar las producciones unitarias

   Agregamos las producciones:

$$S \to bbB|b$$
$$A \to cC|c$$

   Al eliminar las producciones unitarias se obtiene la gramática equivalente G^2

$$G^2 \begin{cases} S & \to AB|aBC|SBS|aB|bbB|b \\ A & \to aA|a|cC|c \\ B & \to bbB|b \\ C & \to cC|c \end{cases}$$

5. Convertir a la FNC

   Agregamos las variables nuevas T_a, T_b, T_c y las producciones

$$T_a \to a \qquad T_b \to b \qquad T_c \to c$$

   y obtenemos

$$G^3 \begin{cases} S & \to AB|T_aBC|SBS|T_aB|T_bT_bB|b \\ A & \to T_aA|a|T_cC|c \\ B & \to T_bT_bB|b \\ C & \to T_cC|c \end{cases}$$

Esta gramática aún no está en la FNC.

Nos queda evaluar las producciones

$A \rightarrow w \qquad$ con $|w| > 2$

Agregamos variables Z_1, Z_2, Z_3 para:

$$
\begin{array}{lll}
S \rightarrow T_aBC & Z_1 \rightarrow BC & S \rightarrow T_aZ_1 \\
S \rightarrow SBS & Z_2 \rightarrow BS & S \rightarrow SZ_2 \\
S \rightarrow T_bT_bB & Z_3 \rightarrow T_bB & S \rightarrow T_bZ_3 \\
B \rightarrow T_bT_bB & B \rightarrow T_bZ_3 &
\end{array}
$$

Finalmente, se obtiene la gramática en FNC

$$G^4 \begin{cases} S & \to AB|T_aZ_1|SZ_2|T_aB|T_bZ_3|b \\ A & \to T_aA|a|T_cC|c \\ B & \to T_bZ_3|b \\ C & \to T_cC|c \\ T_a & \to a \\ T_b & \to b \\ T_c & \to c \\ Z_1 & \to BC \\ Z_2 & \to BS \\ Z_3 & \to T_bB \end{cases}$$