

PRACTICA 11: Eliminación de producciones unitarias en una Gramática Independiente del Contexto

11.1. Introducción

Dada una gramática independiente del contexto (de ahora en adelante, una gramática), $G \equiv (V, \Sigma, S, P)$, una producción se dice que es inútil si tiene la forma $A \rightarrow B$ ($A, B \in V$).

Las producciones inútiles hacen a la gramática innecesariamente compleja: simplemente renombran un símbolo no terminal).

El algoritmo que se presenta en la Figura 11.1 elimina las producciones unitarias de una gramática G . El algoritmo se basa en construir el conjunto H definido como:

$$H = \{(A, B) \in V \times V \mid A \Rightarrow^* B\}.$$

1. $H = \emptyset$
2. **for** (cada producción de la forma $A \rightarrow B$) **do**
3. $H = H \cup \{(A, B)\}$
4. **while** (H cambie) **do**
5. **for** (cada par de parejas $(A, B), (B, C)$) **do**
6. **if** $((A, C) \notin H)$ **then**
7. $H = H \cup \{(A, C)\}$
8. **Eliminar** todas las producciones unitarias
9. **for** (cada producción $A \rightarrow \alpha$) **do**
10. **for** (cada pareja $(B, A) \in H$) **do**
11. Añadir una producción $B \rightarrow \alpha$

Figura 11.1: Eliminación de producciones unitarias

11.2. Ejemplo

Veamos un ejemplo de aplicación del algoritmo. Consideremos para ello la gramática:

$S \rightarrow A|Aa$
 $A \rightarrow B$
 $B \rightarrow C|b$
 $C \rightarrow D|ab$
 $D \rightarrow b$

El primer bucle del algoritmo (líneas 2-3) calcula el conjunto inicial

$H_0 = \{(SA), (AB), (BC), (CD)\}$

El bucle while de las líneas 4-7 calcula el conjunto de pares:

$H = \{(SA), (AB), (BC), (CD), (SB), (AC), (BD), (SC), (SD), (AD)\}$

Al eliminar las producciones unitarias (línea 8 del algoritmo), la gramática resulta:

$S \rightarrow Aa$
 $B \rightarrow b$
 $C \rightarrow ab$
 $D \rightarrow b$

Añadiendo las producciones:

$A \rightarrow b|ab$
 $B \rightarrow ab|b$
 $C \rightarrow b|ab$
 $S \rightarrow b|ab$

La gramática finalmente resulta:

$S \rightarrow Aa|ab|b$
 $A \rightarrow ab|b$
 $B \rightarrow ab|b$
 $C \rightarrow ab|b$
 $D \rightarrow b$

11.3. Práctica

La práctica consiste en desarrollar un programa en C++ que dada una gramática independiente del contexto, le aplique el algoritmo de eliminación de producciones unitarias.

Los ficheros que se utilizarán para representar las gramáticas tanto de entrada como de salida, tendrán el mismo formato que los que ya se han utilizado en prácticas anteriores.