

Algoritmo para el cálculo de siguientes en una gramática incontextual

IG29: Compiladores e intérpretes

Octava sesión de teoría

1. Introducción

El tema de teoría menciona las ecuaciones que cabe aplicar en el cálculo de los siguientes de los símbolos no terminales de una gramática incontextual, pero no especifica ningún orden de aplicación de éstas que garantice la obtención de todos los siguientes de todos los no terminales de la gramática. Es por ello que este documento propone un algoritmo que hace un uso ordenado de tales ecuaciones.

2. El algoritmo

El algoritmo propuesto consta de tres pasos, el último de los cuales puede hacer falta aplicarlo repetidamente:

1. **Inicialización.** Para cada no terminal $\langle A \rangle$ de la gramática, el conjunto $SP[\langle A \rangle]$ se inicializa vacío¹; además, se añade **eof**² a $SP[\langle S \rangle]$, donde con $\langle S \rangle$ se hace referencia al símbolo inicial de la gramática.
2. **Incorporación de primeros.** Se recorren ordenadamente las partes derechas de las reglas buscando no terminales para aplicar que, si $\langle A \rangle \rightarrow \alpha \langle B \rangle \beta$, hay que añadir los elementos de $\text{Primeros}(\beta)$ a $SP[\langle B \rangle]$.
3. **Actualizaciones.** Se recorren ordenadamente las partes derechas de las reglas buscando no terminales para aplicar que, si $\langle A \rangle \rightarrow \alpha \langle B \rangle \beta$ y β es anulable, hay que añadir los elementos de $SP[\langle A \rangle]$ a $SP[\langle B \rangle]$.

El tercer paso hay que aplicarlo hasta que un recorrido completo por la gramática no suponga la modificación de ningún conjunto $SP[\langle A \rangle]$; en ese momento, podrá afirmarse que $\text{Siguintes}(\langle A \rangle) = SP[\langle A \rangle]$ para todo símbolo no terminal de la gramática³.

3. Un ejercicio

He aquí un pequeño ejercicio en el que no se obtiene el resultado correcto si sólo se aplica una vez el tercer paso del algoritmo. Se trata de calcular los primeros y siguientes de cada no terminal de la gramática que se muestra a continuación:

$$\begin{aligned}\langle A \rangle &\rightarrow \langle B \rangle \mathbf{x} \langle C \rangle \mid \mathbf{x} \\ \langle B \rangle &\rightarrow \mathbf{y} \langle A \rangle \mid \mathbf{z} \\ \langle C \rangle &\rightarrow \mathbf{z}\end{aligned}$$

El error típico consiste en no incluir **x** en los siguientes de $\langle C \rangle$.

1. Así, SP será un diccionario de conjuntos, concretamente uno de *siguintes provisionales* donde el algoritmo irá almacenando símbolos a medida que vaya descubriendo que son siguientes de determinados no terminales.

2. En realidad, cabe añadir una representación del final de la entrada que podrá ser **eof** o, si seguimos la notación del tema de teoría, **\$**.

3. Es decir, llegado ese momento, los resultados provisionales del cálculo de siguientes pueden pasar a considerarse definitivos.