Contenido

3.5 NOTACIONES EXTENDIDAS: EBNF Y DIAGRAMAS DE SINTAXIS

3.5.1 Notación EBNF

Las construcciones repetitivas y opcionales son muy comunes en los lenguajes de programación, lo mismo que las reglas de gramática de la BNF. Por lo tanto, no debería sorprendernos que la notación BNF se extienda en ocasiones con el fin de incluir notaciones especiales para estas dos situaciones. Estas extensiones comprenden una notación que se denomina **BNF extendida**, o **EBNF** (por las siglas del término en inglés).

Considere en primer lugar el caso de la repetición, como el de las secuencias de sentencias. Vimos que la repetición se expresa por la recursión en las reglas gramaticales y que se puede utilizar tanto la recursión por la izquierda como la recursión por la derecha, indicadas por las reglas genéricas

$$A \rightarrow A\ \alpha \mid \beta \quad \text{(recursiva por la izquierda)}$$

y

$$A \rightarrow \alpha\ A \mid \beta \quad \text{(recursiva por la derecha)}$$

donde α y β son cadenas arbitrarias de terminales y no terminales y donde en la primera regla β no comienza con A y en la segunda β no finaliza con A.

Se podría emplear la misma notación para la repetición que la que se utiliza para las expresiones regulares, a saber, el asterisco * (también conocido como cerradura de Kleene en expresiones regulares). Entonces estas dos reglas se escribirían como las reglas no recursivas

$$A \rightarrow \beta\ \alpha^*$$

y

$$A \rightarrow \alpha^*\ \beta$$

En vez de eso EBNF prefiere emplear llaves $\{\ldots\}$ para expresar la repetición (así hace clara la extensión de la cadena que se va a repetir), y escribimos

$$A \rightarrow \beta\ \{\alpha\}$$

y

$$A \rightarrow \{\alpha\}\ \beta$$

para las reglas.

El problema con cualquier notación de repetición es que oculta cómo se construye el árbol de análisis gramatical, pero, como ya se expuso, a menudo no importa. Tomemos por ejemplo el caso de las secuencias de sentencias (ejemplo 3.9). Escribimos la gramática como sigue, en forma recursiva por la derecha:

$$\textit{secuencia-sent} \rightarrow \textit{sent}\ \texttt{;}\ \textit{secuencia-sent} \mid \textit{sent}$$
$$\textit{sent} \rightarrow \texttt{s}$$

Esta regla tiene la forma $A \rightarrow \alpha\ A \mid \beta$, con A = *secuencia-sent*, α = *sent* **;** y β = *sent*. En EBNF esto aparecería como

$$\textit{secuencia-sent} \rightarrow \{\ \textit{sent}\ \texttt{;}\ \}\ \textit{sent} \quad \text{(forma recursiva por la derecha)}$$

También podríamos haber usado una regla recursiva por la izquierda y obtenido la EBNF

$$\textit{secuencia-sent} \rightarrow \textit{sent}\ \{\ \texttt{;}\ \textit{sent}\ \} \quad \text{(forma recursiva por la izquierda)}$$

De hecho, la segunda forma es la que generalmente se utiliza (por razones que analizaremos en el siguiente capítulo).

Un problema más importante se presenta cuando se involucra la asociatividad, como ocurre con las operaciones binarias como la resta y la división. Por ejemplo, consideremos la primera regla gramatical en la gramática de expresión simple de la anterior subsección:

$$\textit{exp} \rightarrow \textit{exp}\ \textit{opsuma}\ \textit{term} \mid \textit{term}$$

Esto tiene la forma $A \rightarrow A\ \alpha \mid \beta$, con A = *exp*, α = *opsuma term* y β = *term*. De este modo, escribimos esta regla en EBNF como

$$\textit{exp} \rightarrow \textit{term}\ \{\ \textit{opsuma}\ \textit{term}\ \}$$

Ahora también debemos suponer que esto implica asociatividad por la izquierda, aunque la regla misma no lo establezca explícitamente. Debemos suponer que una regla asociativa por la derecha estaría implicada al escribir

$$\textit{exp} \rightarrow \{\ \textit{term}\ \textit{opsuma}\ \}\ \textit{term}$$

pero éste no es el caso. En cambio, una regla recursiva por la derecha como

$$\textit{secuencia-sent} \rightarrow \textit{sent}\ \texttt{;}\ \textit{secuencia-sent} \mid \textit{sent}$$

se visualiza como si fuera una *sentencia* seguida opcionalmente por un signo de punto y coma y una *secuencia-sent*.

Las construcciones opcionales en EBNF se indican encerrándolas entre corchetes [. . .]. Esto es similar en esencia a la convención de la expresión regular de colocar un signo de interrogación después de una parte opcional, pero tiene la ventaja de encerrar la parte opcional sin requerir paréntesis. Por ejemplo, las reglas gramaticales para las sentencias if con partes opcionales else (ejemplos 3.4 y 3.6) se escribirían en EBNF como se describe a continuación:

sentencia → *sent-if* | `otro`
sent-if → `if` `(` *exp* `)` *sentencia* [`else` *sentencia*]
exp → `0` | `1`

Una regla recursiva por la derecha tal como

secuencia-sent → *sent* `;` *secuencia-sent* | *sent*

también se escribe como

secuencia-sent → *sent* [`;` *secuencia-sent*]

(compare esto con el uso de las llaves anterior para escribir esta regla en forma recursiva por la izquierda).

Si deseáramos escribir una operación aritmética tal como la suma en forma asociativa derecha, escribiríamos

exp → *term* [*opsuma exp*]

en lugar de utilizar llaves.

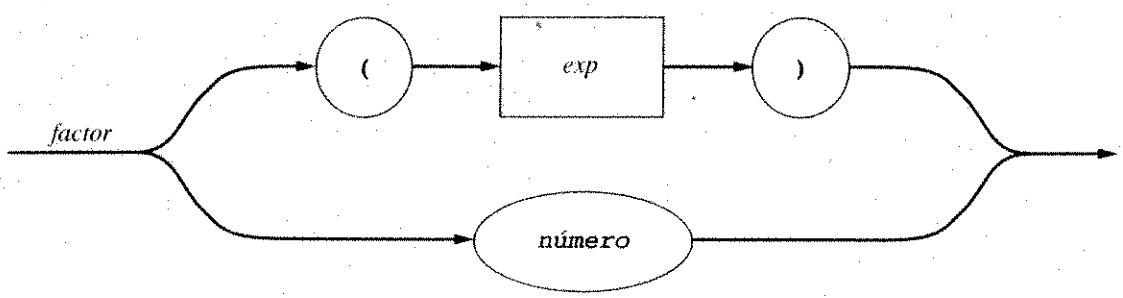
3.5.2 Diagramas de sintaxis

Las representaciones gráficas para simbolizar de manera visual las reglas de la EBNF se denominan **diagramas de sintaxis**. Se componen de cajas que representan terminales y no terminales, líneas con flechas que simbolizan secuencias y selecciones, y etiquetas de no terminales para cada diagrama que representan la regla gramatical que define ese no terminal. Para indicar terminales en un diagrama se emplea una caja de forma oval o redonda, mientras que para indicar no terminales se utiliza una caja rectangular o cuadrada.

Considere como ejemplo la regla gramatical

factor → `(` *exp* `)` | **`número`**

Ésta se escribe como un diagrama sintáctico de la siguiente manera:

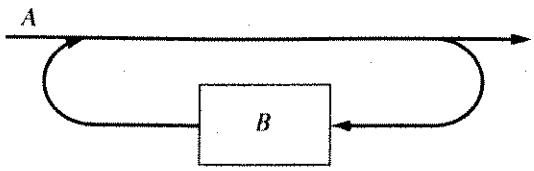


Advierta que *factor* no está situado en una caja, pero se utiliza como una etiqueta para el diagrama de sintaxis, lo cual indica que el diagrama representa la definición de la estructura de ese nombre. Advierta también el uso de las líneas con flechas para indicar la selección y la secuenciación.

Los diagramas de sintaxis se escriben desde la EBNF más que desde la BNF, de modo que necesitamos diagramas que representen construcciones de repetición y opcionales. Dada una repetición tal como

$$A \rightarrow \{\ B\ \}$$

el diagrama de sintaxis correspondiente se dibuja por lo regular como sigue:

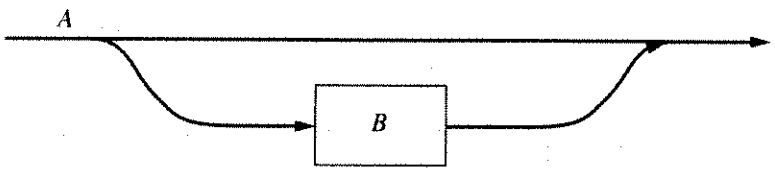


Advierta que el diagrama debe tener en cuenta que no aparezca ninguna *B*.

Una construcción opcional tal como

$$A \rightarrow [\ B\ ]$$

se dibuja como



Concluiremos nuestro análisis de los diagramas de sintaxis con algunos ejemplos en los que se utilizan ejemplos anteriores de EBNF.

Ejemplo 3.10

Considere nuestro ejemplo de ejecución de expresiones aritméticas simples. Éste tiene la BNF (incluyendo asociatividad y precedencia).

$$\begin{aligned}
&exp \rightarrow exp\ opsuma\ term \mid term\\
&opsuma \rightarrow + \mid -\\
&term \rightarrow term\ opmult\ factor \mid factor\\
&opmult \rightarrow *\\
&factor \rightarrow (\ exp\ ) \mid \texttt{número}
\end{aligned}$$

La EBNF correspondiente es

$$\begin{aligned}
&exp \rightarrow term\ \{\ opsuma\ term\ \}\\
&opsuma \rightarrow + \mid -\\
&term \rightarrow factor\ \{\ opmult\ factor\ \}\\
&opmult \rightarrow *\\
&factor \rightarrow (\ exp\ ) \mid \texttt{número}
\end{aligned}$$

Los diagramas de sintaxis correspondientes se proporcionan en la figura 3.4 (el diagrama de sintaxis para *factor* se dio con anterioridad).

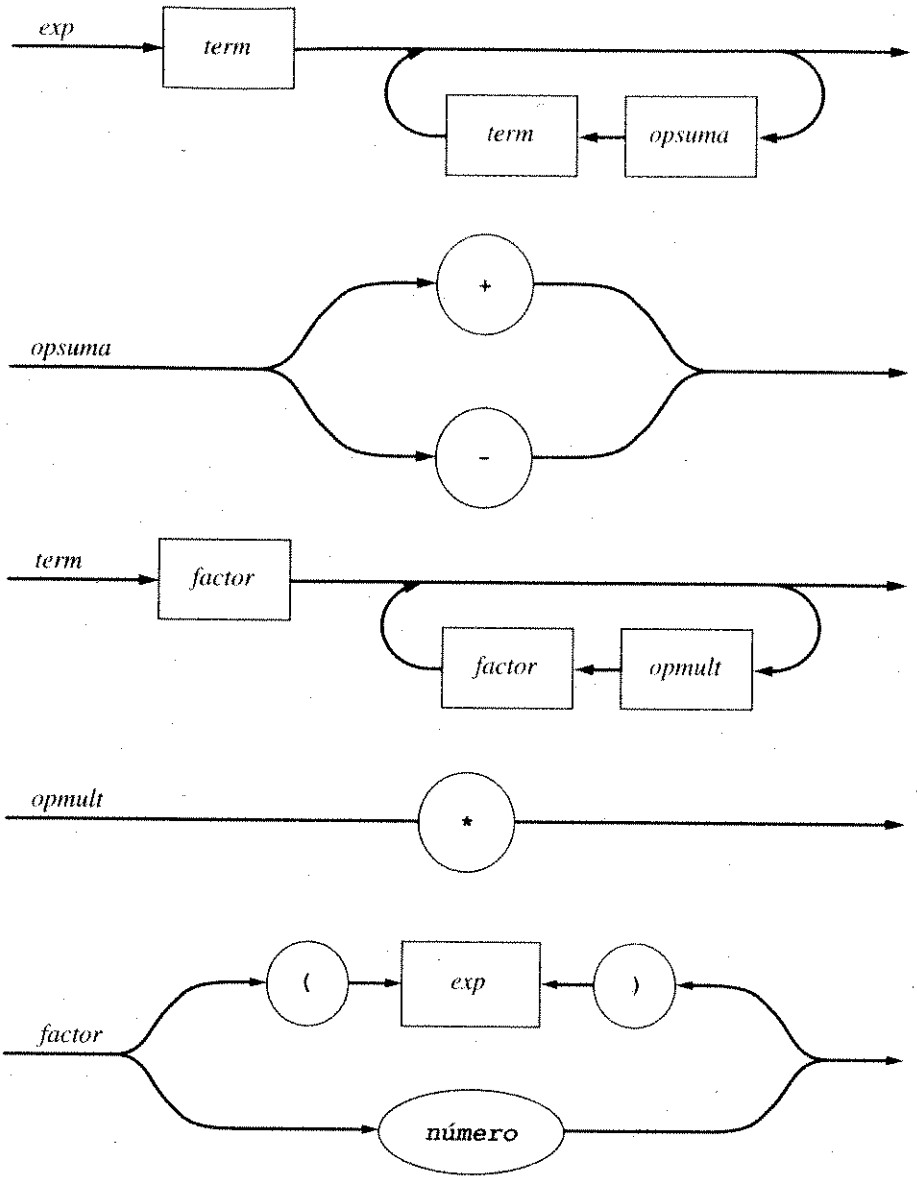


Figura 3.4
Diagramas de sintaxis para la gramática del ejemplo 3.10

§

Ejemplo 3.11 Considere la gramática de las sentencias if simplificadas del ejemplo 3.4 (página 103). Éste tiene la BNF

$$\begin{aligned} &\textit{sentencia} \rightarrow \textit{sent-if} \mid \texttt{otro} \\ &\textit{sent-if} \rightarrow \texttt{if (}\ \textit{exp}\ \texttt{)}\ \textit{sentencia} \\ &\qquad\qquad \mid \texttt{if (}\ \textit{exp}\ \texttt{)}\ \textit{sentencia}\ \texttt{else}\ \textit{sentencia} \\ &\textit{exp} \rightarrow \texttt{0} \mid \texttt{1} \end{aligned}$$

y la EBNF

$$\begin{aligned} &\textit{sentencia} \rightarrow \textit{sent-if} \mid \texttt{otro} \\ &\textit{sent-if} \rightarrow \texttt{if (}\ \textit{exp}\ \texttt{)}\ \textit{sentencia}\ [\ \texttt{else}\ \textit{sentencia}\ ] \\ &\textit{exp} \rightarrow \texttt{0} \mid \texttt{1} \end{aligned}$$

Los diagramas de sintaxis correspondientes se ofrecen en la figura 3.5.

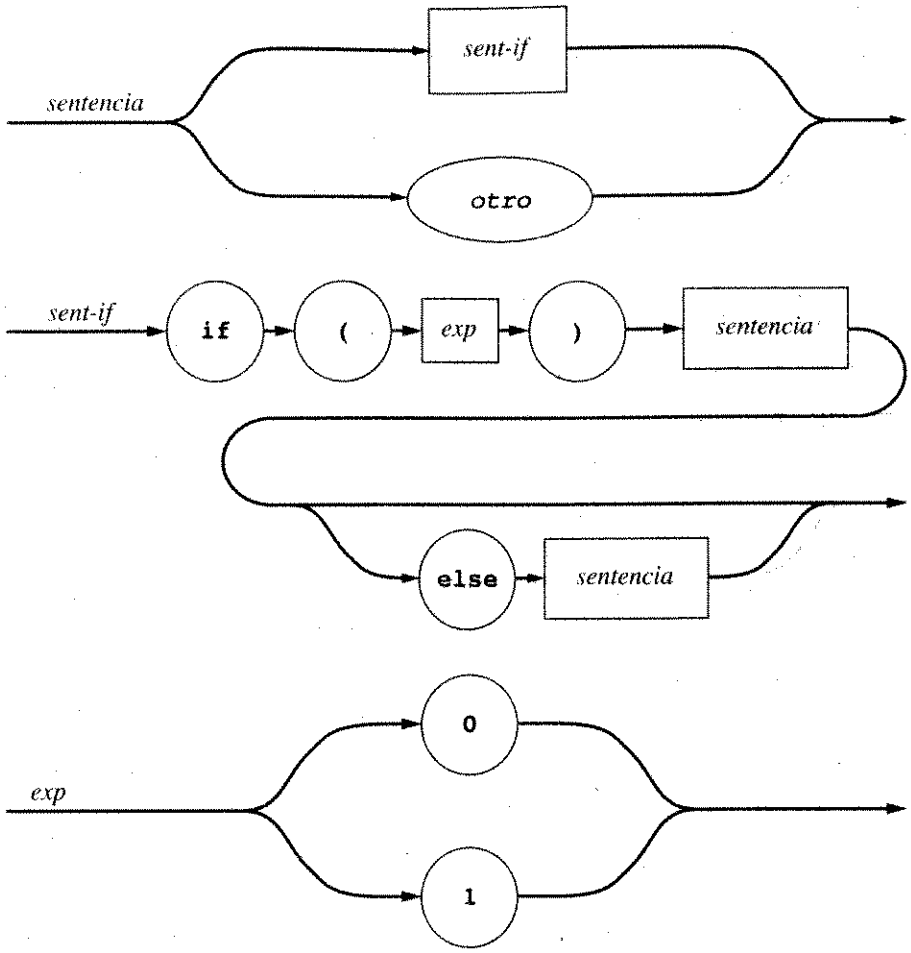


Figura 3.5
Diagramas de sintaxis para la gramática del ejemplo 3.11

§