

ST0244-032

Clase 7

J.F. Cardona

Universidad EAFIT

12 de agosto de 2015

Agenda

- 1 Capítulo 2. Sintaxis
 - Gramáticas independiente de contexto
 - Derivaciones
 - Tipos de derivación
 - Árboles de análisis
 - Análisis sintáctico

Sintaxis

Gramáticas independiente de contexto

- Otro nombre para una gramática BNF es una gramática independiente de contexto.
- La gran diferencia entre estos dos metalenguajes, es que las gramáticas independientes de contexto es un **método formal**.
- Una gramática independiente de contexto es una cuadrupla:

$$G = (V, \Sigma, P, S)$$

donde

- ▶ V es el conjunto de símbolos que representa los no-terminales.
- ▶ Σ es el conjunto de símbolos llamados *tokens* o terminales.
- ▶ P es el conjunto de producciones de la forma $n \rightarrow \alpha$ donde n es un no-terminal y α es una cadena (o secuencia) de símbolos terminales y no-terminales.
- ▶ S es un símbolo que pertenece al conjunto V y es llamado el axioma o el inicio de la gramática.

Ejemplo de gramática

Una gramática de expresiones puede ser especificada como

$G_{\text{expresión}} = (V, \Sigma, P, S)$. donde

$$V = \{E, T, F\}$$

$$\Sigma = \{\text{identificador}, \text{número}, +, -, *, /, (,)\}$$

$$P = \{E \rightarrow E + T \mid E - T \mid T, \\ T \rightarrow T * F \mid T / F \mid F, \\ F \rightarrow (E) \mid \text{identificador} \mid \text{número}\}$$

$$S = E$$

Sintaxis

Derivaciones

- Un **frase** de una gramática es una cadena de caracteres de una gramática.
- Una **frase** pertenece al lenguaje de una gramática si ésta puede ser derivada de la gramática.
- Una **derivación** es una secuencia de formas sentenciales que inicia con el símbolo inicial de la gramática (o axioma) y termina con un frase que se está tratando de derivar.
- Una **forma sentencial** es una cadena de terminales y no terminales de la gramática.
- En cada paso en la derivación, un no-terminal de una forma sentencial A es reemplazada con un conjunto de terminales y no terminales β , donde $A \rightarrow \beta$ es una producción de la gramática.

Ejemplo de derivación

Prueba que la expresión $(5 * x) + y$ es un miembro del lenguaje definido por la gramática $G_{\text{expresión}}$.

Práctica

Construya una derivación para la expresión $4 + (a - b) * x$.

Sintaxis

Derivación - Tipos de derivación

- Una frase de un gramática es **válida** si existe al menos una derivación de ella utilizando una gramática.
- Existen dos tipos de derivaciones que son de algún interés para entender los lenguajes de programación:
 - ▶ **Derivación más a la izquierda** - Siempre reemplaza el no-terminal más a la izquierda cuando se va de una forma sentencial a otra en una derivación.
 - ▶ **Derivación más a la derecha** - Siempre reemplaza el no-terminal más a la derecha cuando se va de una forma sentencial a otra en una derivación.

Sintaxis

Derivación - Tipos de derivación

Ejemplo tipos derivación

La derivación de la frase $(5 * x) + y$ en el ejemplo anterior es una derivación más a la izquierda.

Práctica

Construya la derivación más a la izquierda y más a la derecha de $(x * y) + z$.

Sintaxis

Árboles de análisis

- Una gramática para un lenguaje puede ser utilizada para construir una representación en forma de árbol (invertido) de una frase (o cadena) de la gramática.
- Este tipo de árbol es llamado **árbol de análisis** (**parser tree**).
- Los árboles de análisis es otra forma de representar una frase (o cadena) de un lenguaje dado.
- Un árbol de análisis es construido a partir del símbolo inicial de la gramática tomando este elemento como la raíz del árbol.
- Los hijos de cada nodo en el árbol deben aparecer en el lado derecho (*rhs*) de una producción con el padre en el lado izquierdo (*lhs*) de la misma producción.

Sintaxis

Árboles de análisis

- Un programa es sintácticamente válido si existe un árbol de análisis para este programa utilizando la gramática dada.
- Mientras existen diferentes tipos de derivación para una frase (o cadena) en un lenguaje, sólo existe un árbol de análisis.
- Lo anterior es cierto si la gramática es no ambigua.
- Un **gramática es ambigua**, sí y sólo sí, si una frase del lenguaje tiene más de un árbol de análisis.

Sintaxis

Árboles de análisis

Ejemplo de un árbol de derivación

Un árbol de análisis para la cadena $(5 * x) + y$, de la gramática de expresiones aritméticas, a través de la derivación más a la izquierda:



Sintaxis

Árboles de análisis

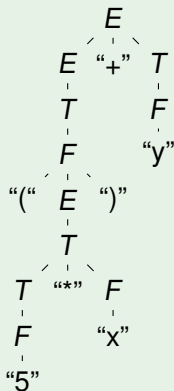
Practica

¿Cómo es el árbol de derivación para la derivación más a la derecha de la cadena $(5 * x) + y$ para la gramática de expresiones aritméticas?

Sintaxis

Árboles de análisis

Respuesta



Sintaxis

Árboles de análisis

Practica

Construya un árbol de análisis para la expresión $4 + (a - b) * x$.

Sintaxis

Análisis sintáctico

- En análisis sintáctico (**Parsing**) es el proceso de detectar si una cadena dada de tokens es una frase (cadena) válida de una gramática.
- Es activada por la compilación explícita de un programa o por la interpretación de uno por parte de un interpretador.
- El resultado básico de un análisis sintáctico es un valor booleano.
- El **analizador sintáctico (parser)** es un programa que dada un frase, examina para ver si la frase es un miembro del lenguaje de una gramática determinada.

Sintaxis

Análisis sintáctico

- Internamente el analizador sintáctico construye el árbol de análisis.—
- Existen dos estrategias para construir este árbol de análisis:
 - ▶ Un analizador sintáctico que comienza a construir el árbol a partir de la raíz.
 - ▶ Un analizador sintáctico que comienza a construir el árbol a partir de las hojas del árbol.
- Ambos analizadores sintácticos evalúan si una frase (cadena) pertenece a una gramática construyendo una derivación de la frase.
- El analizador sintáctico tiene dos opciones al analizar la cadena: la acepta o la rechaza.