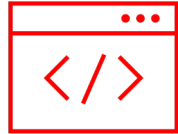




UNIVERSIDAD
NACIONAL
DE COLOMBIA

PROYECTO **CULTURAL, CIENTÍFICO Y COLECTIVO** DE NACIÓN



Generalidades

Prof. Oscar Mauricio Salazar Ospina

Correo: omsalazaro@unal.edu.co

3007743 - Programación Lógica y Funcional

3010426 - Teoría de Lenguajes de Programación

Facultad de Minas

Departamento de ciencias de la computación y la decisión

Agosto 17 de 2022

Universidad Nacional de Colombia

PROYECTO **CULTURAL, CIENTÍFICO Y COLECTIVO** DE NACIÓN

Generalidades

Contenido clase #2



Gramáticas y el análisis sintáctico



BNF – Backus Naur Form ó Backus Normal Form



Gramáticas libres de contexto



Derivaciones (izquierda, derecha)



Gramáticas con notación infija



Gramáticas con notación prefija



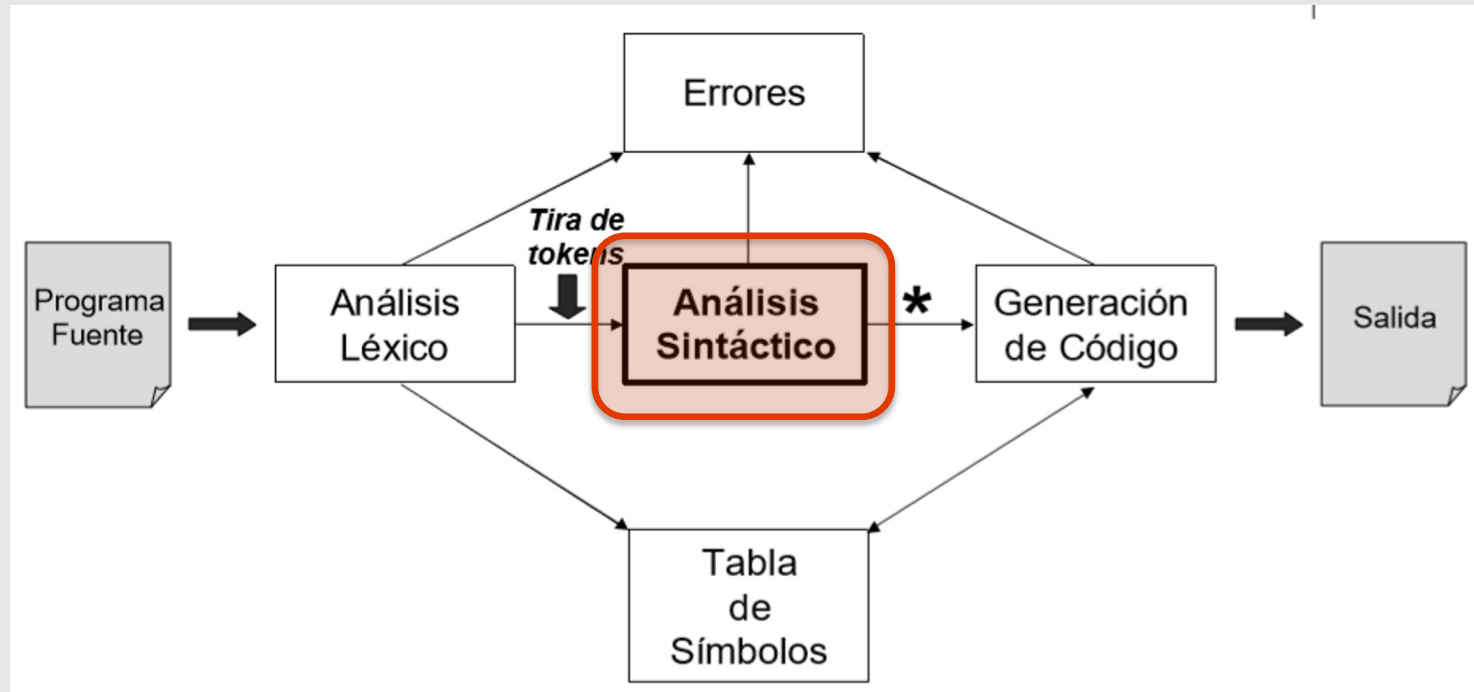
Árboles de análisis sintáctico (*Parser Tree*)



Árboles sintácticos abstractos (AST)

Generalidades

Fases de compilación



Generalidades

Fases de compilación

Token	Identificación del token
ID	27
CTE	28
IF	59
THEN	60
ELSE	61
+	70
/	73
>=	80
:=	85

precio = costo + impuesto * 60



id1 = id2 + id3 * 60

Generalidades

Fases de compilación

Token	Identificación del token
ID	27
CTE	28
IF	59
THEN	60
ELSE	61
+	70
/	73
>=	80
:=	85

Gramática libre de contexto



Analizador sintáctico

precio = costo + impuesto * 60



id1 = id2 + id3 * 60

Generalidades

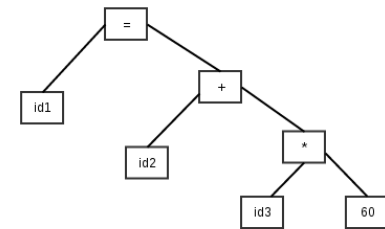
Fases de compilación

Token	Identificación del token
ID	27
CTE	28
IF	59
THEN	60
ELSE	61
+	70
/	73
>=	80
:=	85

Gramática libre de contexto



Analizador sintáctico



Árbol sintáctico

precio = costo + impuesto * 60



id1 = id2 + id3 * 60

Generalidades

BNF y gramáticas libres de contexto



John Backus, científico estadounidense en Ciencias de la computación.



Propuso las **gramáticas libres de contexto** que utilizan los **compiladores**. Dicha **notación** se conoce como **Backus-Naur Form** ó **Backus-Normal Form (BNF)**.



Describió la estructura del lenguaje **Algol** utilizando la **notación BNF**.



Escribió el artículo: ***“Can programming be liberated from the Von Neumann Style?”*** (Ganador del premio Turing).



Generalidades

BNF y gramáticas libres de contexto



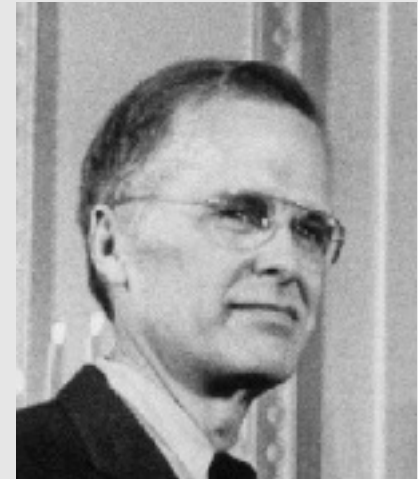
“They don't like thinking in medical school. They memorize - that's all they want you to do. You must not think.”



“Much of my work has come from being lazy.”



“You need the willingness to fail all the time.”



Generalidades

BNF y gramáticas libres de contexto



BNF es un **metalenguaje** formal para **describir sintaxis** de un lenguaje.



Formal == **No ambigua.**



Español, inglés, Alemán son ambiguos.



Solo hay una forma de leerlo.



BNF consta de un conjunto de reglas que tienen la forma:

<categoría sintáctica> ::= **T y N**

::= 'está compuesto de'

T: símbolos terminales

N: Símbolos no terminales



**Gramáticas libres de
contexto**

Generalidades

Gramáticas libres de contexto



Son **fórmulas bien formadas (FBF)** que describen la **gramática** de un **lenguaje**.



No depende del contexto en el que se utiliza.



Se define como una **tupla de cuatro elementos**:

$$G = (N, T, P, E)$$

- N es un conjunto de símbolos llamados no terminales o categorías sintácticas.
- T es un conjunto de símbolos llamados terminales o tokens.
- P es un conjunto de reglas de producción de la forma:

$$n \rightarrow \alpha \text{ donde } n \in N \text{ y } \alpha \in \{N \cup T\}^*.$$

- $E \in N$ es un no terminal llamado **símbolo de inicio** de la gramática.

Generalidades

Derivaciones



Una **expresión** de una gramática es una cadena de **tokens** de la gramática.



Una **expresión pertenece al lenguaje de una gramática** si puede **derivarse** de la gramática.



Este proceso se llama **construir** una **derivación**.



Una **derivación** es una **secuencia de sentencias** que **comienza** con el **símbolo de inicio** de la gramática y **termina** con la **sentencia** que está tratando de **derivar**.



Una **expresión** es una cadena de **terminales** y **no terminales** de la gramática.



En cada paso de la derivación, un no terminal **A**, es reemplazado por una cadena de terminales y no terminales **β** donde **$A \rightarrow \beta$** es una regla en la gramática.



Para una gramática **G**, el **lenguaje de G** es el conjunto de todas las expresiones que se pueden **derivar** de **G** y se escribe como **$L(G)$**

Generalidades

Tipos de derivación



Derivación izquierda (de izquierda a derecha): siempre reemplace la derivación más a la izquierda cuando se pasa de una expresión gramatical a la siguiente en una derivación.



Derivación derecha (de derecha a izquierda): reemplace siempre el no terminal más a la derecha al pasar de una forma de expresión gramatical a la siguiente en una derivación.

Generalidades

Notaciones (prefija, infija y postfija)

	$(2+(3*4))$	$((2+3)*4)$
Prefija	$+ 2 * 3 4$	$* + 2 3 4$
Infija	$2 + 3 * 4$	$(2 + 3) * 4$
Postfija	$2 3 4 * +$	$2 3 + 4 *$

Nota: la notación **prefija**, conocida también como **notación Polaca**, nos indica que el operador va antes de los operandos:

- No requiere paréntesis para indicar el orden de precedencia de operadores.
- Se evalúa de izquierda a derecha hasta que encontremos el primer operador seguido inmediatamente de un par de operandos.
- Se evalúa la expresión binaria y el resultado se cambia como un nuevo operando. Se repite esto mismo que nos quede un solo resultado.

Generalidades

Ejemplo de derivación

Gramáticas con notación infija

Pueden especificarse como:

$$G = (N, T, P, E)$$

N – No terminal, T – Terminal

P – Producciones, E – Símbolo de inicio

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

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

P se define por el conjunto de producciones (o reglas de producción)

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

¿ $+ * 5 x y$?

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow ???$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E , a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje $L(G)$ de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: **(5 * x) + y** hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que **(5 * x) + y** está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje $L(G)$ de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow$
 $(5 * x) + y$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * x) + T \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje $L(G)$ de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * x) + T \Rightarrow (5 * x) + F \Rightarrow$

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje $L(G)$ de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * x) + T \Rightarrow (5 * x) + F \Rightarrow (5 * x) + y$



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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje $L(G)$ de la **Gramática Infija G**.

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * x) + T \Rightarrow (5 * x) + F \Rightarrow (5 * x) + y$

Derivación derecha:

$E \Rightarrow ???$

Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Generalidades

Gramática con notación infija

¿La expresión: $(5 * x) + y$ hace parte de esta gramática ?

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

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow (E) + T \Rightarrow (T) + T \Rightarrow (T * F) + T \Rightarrow (F * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * F) + T \Rightarrow (5 * x) + T \Rightarrow (5 * x) + F \Rightarrow (5 * x) + y$

Derivación derecha:

$E \Rightarrow E + T \Rightarrow E + F \Rightarrow E + y \Rightarrow T + y \Rightarrow F + y \Rightarrow (E) + y \Rightarrow (T) + y \Rightarrow (T * F) + y \Rightarrow (T * x) + y \Rightarrow (F * x) + y \Rightarrow (5 * x) + y$



Cada **paso** es una **sentencia**.

El **símbolo no terminal izquierdo** se **reemplaza** por el lado **derecho** de una **regla de producción** para ese **no terminal**.

La **derivación procede** del símbolo de **inicio**, E, a la expresión $(5 * x) + y$.

Esto **prueba** que $(5 * x) + y$ está en el lenguaje L (G) de la **Gramática Infija G**.

Generalidades

Ejemplo de derivación

Gramáticas con notación prefija

Pueden especificarse como:

$$G = (N, T, P, E)$$

N – No terminal, T – Terminal

P – Producciones, E – Símbolo de inicio

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

$T = \{\text{identificador, número, +, -, *, /}\}$

P se define por el conjunto de producciones (o reglas de producción)

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

¿La expresión: **+ 4 * - a b x** hace parte de esta gramática ?

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow ???$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + E E \Rightarrow$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + E E \Rightarrow + 4 E \Rightarrow$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow + 4 * - EEE \Rightarrow$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a b x hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow + 4 * - EEE \Rightarrow + 4 * - a EE \Rightarrow$

Generalidades

Gramática con notación prefija

¿La expresión: **+** 4 ***** - a **b** **x** hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow + 4 * - EEE \Rightarrow + 4 * - a EE \Rightarrow + 4 * - a b E \Rightarrow + 4 * - a b x$

Generalidades

Gramática con notación prefija

¿La expresión: **+ 4 * - a b x** hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow + 4 * - EEE \Rightarrow + 4 * - a EE \Rightarrow + 4 * - a b E \Rightarrow$
+ 4 * - a b x



Generalidades

Gramática con notación prefija

¿La expresión: **+ 4 * - a b x** hace parte de esta gramática ?

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

Derivación izquierda:

$E \Rightarrow + EE \Rightarrow + 4 E \Rightarrow + 4 * EE \Rightarrow + 4 * - EEE \Rightarrow + 4 * - a EE \Rightarrow + 4 * - a b E \Rightarrow$
+ 4 * - a b x



Derivación derecha:

$E \Rightarrow ???$ (ejercicio)

Generalidades

Teoría de lenguajes de programación – fases de compilación

Ejercicio: Construya las derivaciones derecha e izquierda para la expresión en notación infija **4 + (a - c) * x**.

Gramática para Notación Infija:

$G = (N, T, P, E)$ donde:

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

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

P se define por el conjunto de producciones

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Generalidades

Árboles de análisis sintáctico (Parser Tree)



A partir de una **gramática G** se puede **construir** un **árbol** que **represente** una **expresión** del lenguaje de la gramática **L(G)**.



Un **árbol** de **análisis sintáctico** ó *parser tree* es otra forma de **representar** una **expresión** de un **lenguaje determinado**.



El **árbol** se **construye** con el **símbolo** de **inicio** de la gramática en el **nodo raíz** del árbol.



Los **hijos** de cada nodo en el árbol **deben aparecer en el lado derecho** de una **producción** con el **padre en el lado izquierdo** de la misma producción.



Un **programa** es **sintácticamente válido** si **hay un árbol** de análisis sintáctico para él **usando la gramática** dada.

Generalidades

Árboles de análisis sintáctico (Parser Tree)



Árbol de análisis sintáctico para la expresión prefijo $+ 4 * - a b x$

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$

???

Generalidades

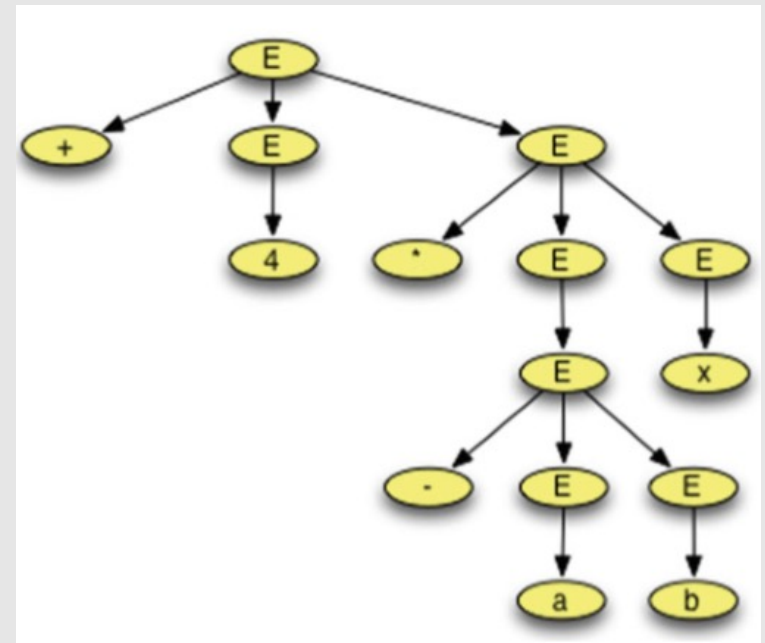
Árboles de análisis sintáctico (Parser Tree)



Árbol de análisis sintáctico para la expresión prefijo $+ 4 * - a b x$

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

$E \rightarrow + EE \mid -EE \mid * EE \mid / EE \mid \text{identificador} \mid \text{número}$



Generalidades

Árboles de análisis sintáctico (Parser Tree)



Árbol de análisis sintáctico para la expresión infija **4 + (a - b) * x**.

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

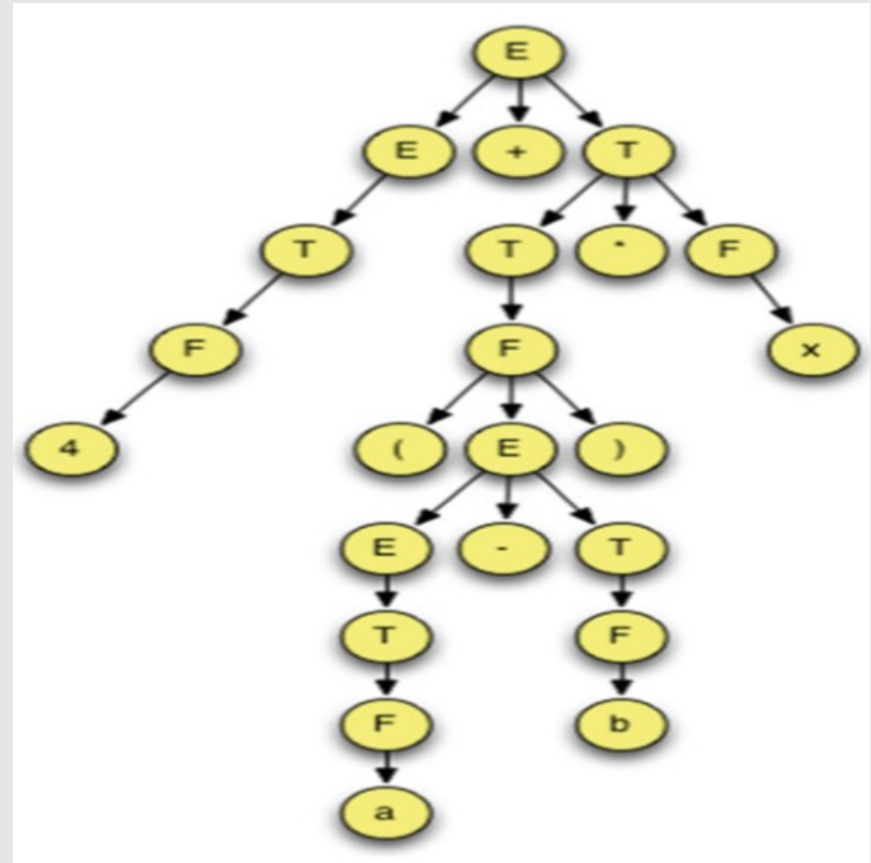
$T = \{\text{identificador, número, +, -, *, /, (,)}\}$

P se define por el conjunto de producciones
(o reglas de producción)

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$



Generalidades

Teoría de lenguajes de programación – fases de compilación

EJERCICIO

¿Cuál es el árbol de análisis sintáctico para la expresión infija **(5 * x) + y** ?

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

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

P se define por el conjunto de producciones
(o reglas de producción)

$E \rightarrow E + T \mid E - T \mid T$

$T \rightarrow T * F \mid T / F \mid F$

$F \rightarrow (E) \mid \text{identificador} \mid \text{número}$

Generalidades

Árboles sintácticos abstractos



Hay **mucha información** en un **árbol de análisis sintáctico** que **NO** es realmente **necesaria** para capturar el **significado** del programa que representa.

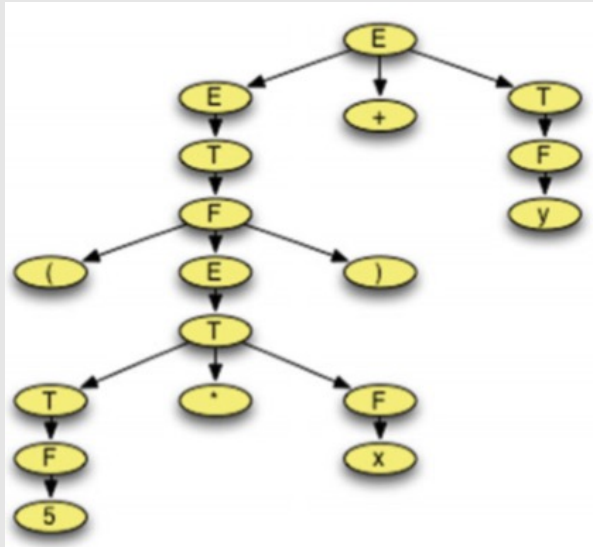
Un **árbol de sintaxis abstracta (AST)** es como un árbol de análisis sintáctico, excepto que se **elimina la información no esencial**.

Más específicamente:

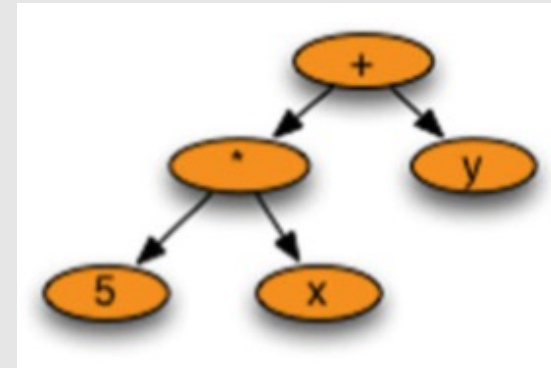
- Los nodos no terminales en el árbol se reemplazan por nodos que reflejan la parte de la expresión que representan.
- Las producciones unitarias (ejm “(“, “)”) en el árbol se eliminan.

Generalidades

Árboles sintácticos abstractos (AST)



Árbol de análisis sintáctico para
 $(5 * x) + y$

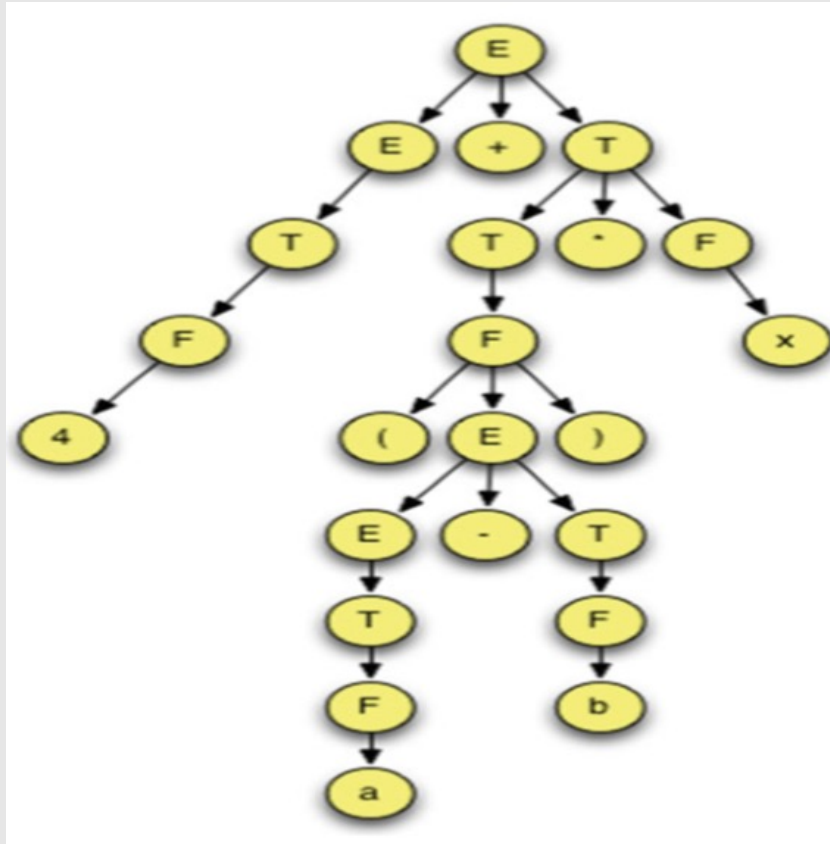


Árbol sintáctico abstracto para
 $(5 * x) + y$

Generalidades

Árboles sintácticos abstractos (AST)

Construya el árbol sintáctico abstracto (AST) para la expresión $4 + (a - b) * x$

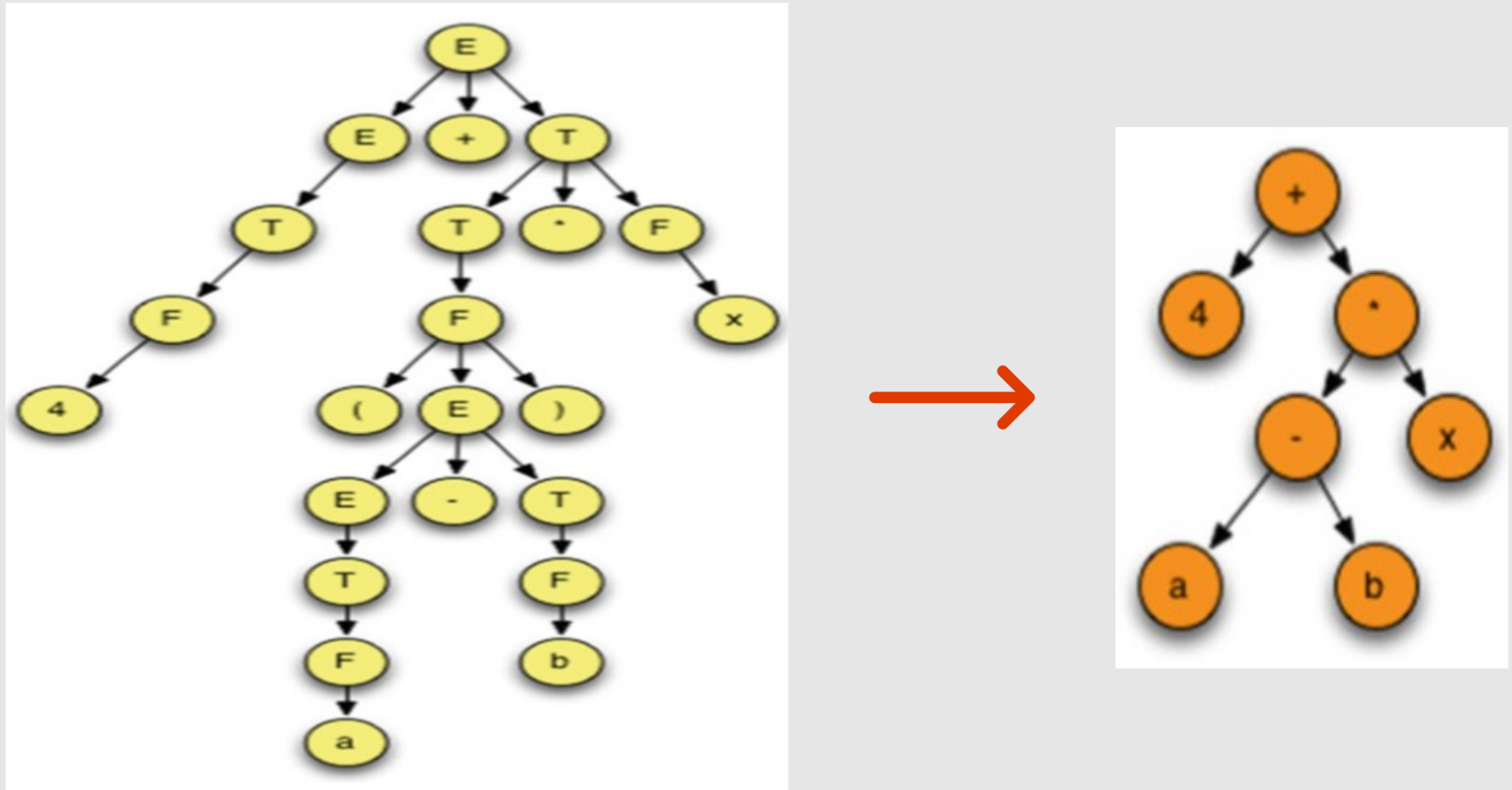


???

Generalidades

Árboles sintácticos abstractos (AST)

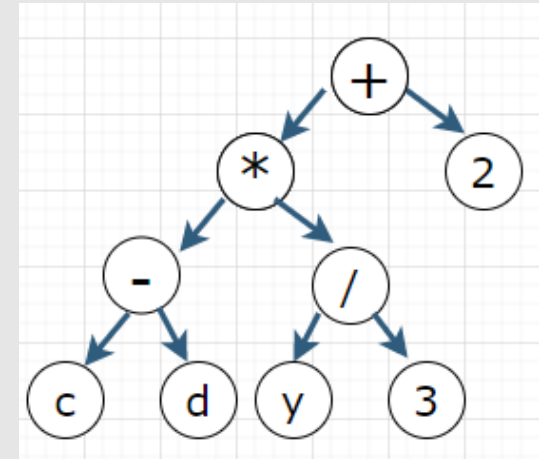
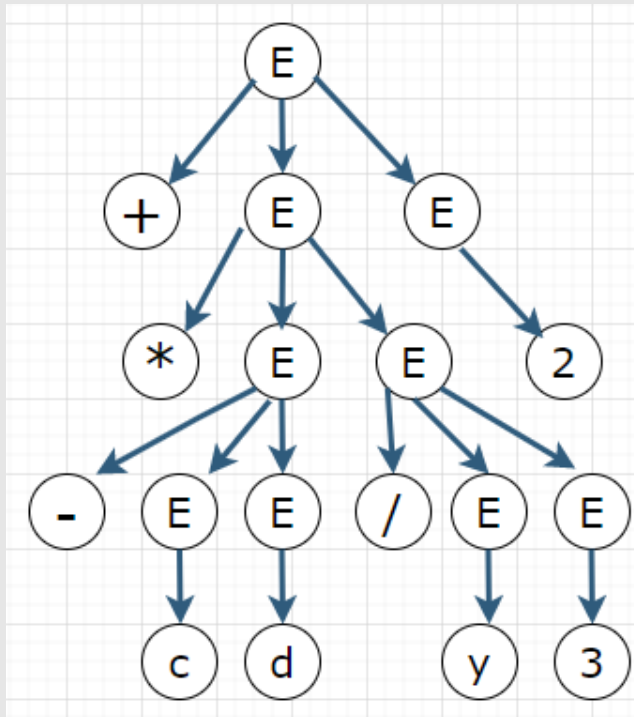
Construya el árbol sintáctico abstracto (AST) para la expresión $4 + (a - b) * x$



Generalidades

Árboles sintácticos abstractos (AST)

Construya el árbol sintáctico abstracto (AST) para la expresión $+ * - c d / y 3 2$



Generalidades

Referencias

- Bratko, Ivan. Prolog Programming for Artificial Intelligence (4th Edition) (International Computer Science Series) Aug 31, 2011
- Programming in Prolog: Using the ISO Standard Oct 4, 2013 by William F. Clocksin and Christopher S. Mellish
- Thinking as Computation (The MIT Press), Jan 6, 2012. by Hector J. Levesque
- Clause and Effect: Prolog Programming for the Working Programmer Apr 29, 2003. by William F. Clocksin
- Programming in Haskell Sep 12, 2016 by Graham Hutton
- Learn You a Haskell for Great Good!: A Beginner's Guide Apr 15, 2011 by Miran Lipovaca
- Learning Haskell Data Analysis, May 28, 2015 by James Church
- Advanced Computer Programming in Python, Advanced Computer Programming in Python Mar 22, 2017, by Karim Pichara and Christian Pieringer
- Functional Python Programming - Create Succinct and Expressive Implementations with Python Jan 31, 2015 by Steven Lott
- Functional Python Programming: Discover the power of functional programming, generator functions, lazy evaluation, the built-in itertools library, and monads, 2nd Edition Apr 13, 2018 by Steven F. Lott

Gracias 

Universidad Nacional de Colombia

PROYECTO **CULTURAL, CIENTÍFICO Y COLECTIVO** DE NACIÓN