

## Sesión 14

### Gramáticas libres del contexto (GLC o CFG)

Dr. Luis Pineda, IIMAS, 2010

### Hay lenguajes que no son regulares

- $Pal = \{w \mid w = w^R\} \subseteq \{0, 1\}^*$
- $Pal$  no es regular:
  - El lema del bombeo:
    - Sea  $n$  la constante asociada
    - Sea  $w = 0^n 10^n$ :  $|w| = 2n + 1 > n$
    - Si  $Pal$  es regular  $w = xyz$ , tal que  $|xy| \leq n$  &  $|y| > 0$ ;  $y$  es una secuencia al final del primer grupo de 0's:  $x = 0^i$  &  $y = 0^j$ , tal que  $i \geq 0$ ,  $j > 0$ ,  $i + j \leq n$  entonces  $|xy| = |0^i 0^j| \leq n$  &  $|y| = j > 0$ .
    - Sea  $m = 0$ :
      - $xy^m z = xz = 0^i 10^n \notin Pal$  as  $i < n$
- $Pal$  no se puede expresar con una ER o un FA

Dr. Luis Pineda, IIMAS, 2010

### Definición recursiva de lenguajes

- Definición recursiva de un lenguaje:
  - Definir cadenas compuestas de un lenguaje como función de cadenas más simples (en el lenguaje)
- Definición recursiva de  $Pal$ :
  - Base:  $\Lambda, 0$  &  $1 \in Pal$
  - Inducción: Si  $w \in Pal$  entonces  $1w1$  &  $0w0$

Dr. Luis Pineda, IIMAS, 2010

### Definición recursiva de lenguajes

- Una GLC (CFG) es una notación para expresar este tipo de definiciones recursivas
  - Las variables representan clases de cadenas (i.e. categorías gramaticales y lenguajes)
  - Las constantes representan los símbolos léxicos en  $\Sigma$
  - Las reglas de producción tiene la forma:
 
$$\alpha \rightarrow \beta$$
 $\alpha$  puede reescribirse como  $\beta$  en cualquier contexto

Dr. Luis Pineda, IIMAS, 2010

### Definición recursiva de lenguajes

- Definición recursiva de  $Pal$ 
  - Base:  $\Lambda, 0$  and  $1 \in Pal$
  - Inducción: Si  $w \in Pal$  entonces  $1w1$  &  $0w0$
- La gramática de  $Pal$ :
 

1. $P \rightarrow \Lambda$	1. $P \Rightarrow 0P0$	por 4
2. $P \rightarrow 0$	2. $\Rightarrow 01P10$	por 5
3. $P \rightarrow 1$	3. $\Rightarrow 01\Lambda 10$	por 1
4. $P \rightarrow 0P0$	$= 0110$	
5. $P \rightarrow 1P1$		

Dr. Luis Pineda, IIMAS, 2010

### Definición recursiva de lenguajes : Ejemplo 2

- Definición recursiva de  $L_{exp}$  (no-regular)
  - Base:  $a$
  - Inducción: Si  $w \in L$  entonces  $w + w \mid w * w \mid (w)$
- La GLC:
 

1. $E \rightarrow a$	1. $E \Rightarrow E + E$	por 2
2. $E \rightarrow E + E$	2. $\Rightarrow a + E$	por 1
3. $E \rightarrow E * E$	3. $\Rightarrow a + (E)$	por 4
4. $E \rightarrow (E)$	4. $\Rightarrow a + (E * E)$	por 3
	5. $\Rightarrow a + (a * E)$	por 1
	6. $\Rightarrow a + (a * a)$	por 1

Dr. Luis Pineda, IIMAS, 2010

### Definición formal de las GLC

- Una gramática libre de contexto (GLC) es un cuarteto

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

donde:

- $V$  es un conjunto de variables (símbolos no terminales, categorías sintácticas, tipos de cadenas)
- $\Sigma$  es el alfabeto (símbolos léxicos o terminales)
- $S \in V$  el símbolo inicial (oración, programa)
- $P$  es el conjunto de reglas gramáticas o reglas de producción

Dr. Luis Pineda, IIMAS, 2010

### Definición formal de las GLC

- Reglas gramáticas o de producción:

$$A \rightarrow \gamma \quad (\text{Las producciones de } A)$$

donde:

- $A \in V$  es la cabeza (*head*) de la regla
- " $\rightarrow$ " es el símbolo de *producción*
- $\gamma \in \{V \cup \Sigma\}^*$  es el cuerpo de la prod.

Dr. Luis Pineda, IIMAS, 2010

### Definición formal: ejemplos

- GLC para *Pal*
  - $G_{pal} = (\{P\}, \{0, 1\}, P, Q)$
  - $Q = \{P \rightarrow \Lambda, P \rightarrow 0, P \rightarrow 1, P \rightarrow 0P0, P \rightarrow 1P1\}$
- Notación compacta para  $Q$ :
$$P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$$

Dr. Luis Pineda, IIMAS, 2010

### Definición formal: ejemplos

- GLC para  $0^n 1^n$ 
  - $G_{pal01} = (\{P\}, \{0, 1\}, P, \{P \rightarrow \Lambda, P \rightarrow 0P1\})$
- Gerenera 000111:
$$\begin{array}{l} 0P1 \\ 00P11 \\ 000P111 \\ 000111 \end{array}$$

Dr. Luis Pineda, IIMAS, 2010

### Definición formal: ejemplos

- GLC para  $L_{exp}$ 
  - $G_{exp} = (\{E\}, \{+, *, (, ), a\}, E, P)$
  - $P = \{E \rightarrow E + E \mid E * E \mid (E) \mid a\}$

Dr. Luis Pineda, IIMAS, 2010

### Producciones

- Si  $\alpha$  es una cadena de forma
$$\alpha_1 A \alpha_2$$
y hay una producción de forma
$$A \rightarrow \gamma$$
entonces  $\alpha$  puede reescribirse como una  $\beta$  de forma:
$$\alpha_1 \gamma \alpha_2$$
- ¿Por qué libre de contexto?
  - La substitución se no depende de la forma de  $\alpha_1$  &  $\alpha_2$

Dr. Luis Pineda, IIMAS, 2010

## Derivaciones

- Se dice que  $\alpha$  genera o deriva a  $\beta$  o que  $\beta$  es generada o derivada a partir de  $\alpha$  en un paso en  $G$ :

$$\alpha \Rightarrow_G \beta$$

- Derivación en *Pal*:

$$P \Rightarrow_G 0P0 \Rightarrow_G 01P10 \Rightarrow_G 01\Lambda 10 = 0110$$

Dr. Luis Pineda, IIMAS, 2010

## Derivaciones en *GLC*

- Derivación en  $L = 0^n 1^n$   
 $P \Rightarrow 0P1 \Rightarrow 00P11 \Rightarrow 00\Lambda 11 = 0011$
- Si es claro cuál es  $G$ , escribimos simplemente " $\Rightarrow$ "

Dr. Luis Pineda, IIMAS, 2010

## Derivaciones-\*

- Derivaciones-\*: derivaciones en cero o más pasos en una gramática  $G$ :

$$\alpha \Rightarrow_G^* \beta$$

- ya sea que  $\alpha = \beta$
- o existe una  $k \geq 1$  y cadenas  $\alpha_0, \alpha_1, \dots, \alpha_k$ , con  $\alpha_0 = \alpha$  &  $\alpha_k = \beta$  tales que  $\alpha_i \Rightarrow_G \alpha_{i+1}$  para toda  $i$  tal que  $0 \leq i \leq k - 1$

- Ejemplos:

- $P \Rightarrow_{Pal}^* 0110$
- $P \Rightarrow_{Pal01}^* 0011$

Dr. Luis Pineda, IIMAS, 2010

## Estructura sintáctica

- Derivación en *Pal*:

-  $P$



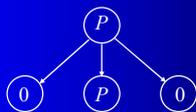
Producciones:  $P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Dr. Luis Pineda, IIMAS, 2010

## Estructura sintáctica

- Derivación en *Pal*:

-  $P \Rightarrow_G 0P0$



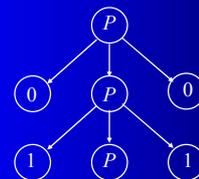
Producciones:  $P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Dr. Luis Pineda, IIMAS, 2010

## Estructura sintáctica

- Derivación en *Pal*:

-  $P \Rightarrow_G 0P0 \Rightarrow_G 01P10$



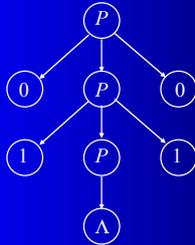
Producciones:  $P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Dr. Luis Pineda, IIMAS, 2010

### Estructura sintáctica

- Derivación en Pal:

$$- P \Rightarrow_G 0P0 \Rightarrow_G 01P10 \Rightarrow_G 01\Lambda 10$$



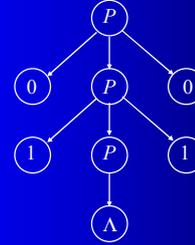
Producciones:  $P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Dr. Luis Pineda, IIMAS, 2010

### Estructura sintáctica

- Derivación en Pal:

$$- P \Rightarrow_G 0P0 \Rightarrow_G 01P10 \Rightarrow_G 01\Lambda 10 = 0110$$



Producciones:  $P \rightarrow \Lambda \mid 0 \mid 1 \mid 0P0 \mid 1P1$

Dr. Luis Pineda, IIMAS, 2010

### Estructura sintáctica

- Derivación en Pal:

$$- P \Rightarrow_G 0P0 \Rightarrow_G 01P10 \Rightarrow_G 01\Lambda 10 = 0110$$

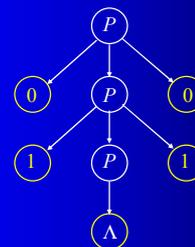


Dr. Luis Pineda, IIMAS, 2010

### Estructura sintáctica

- Derivación en Pal:

$$- P \Rightarrow_G 0P0 \Rightarrow_G 01P10 \Rightarrow_G 01\Lambda 10 = 0110$$



La cosecha (The yield)

Dr. Luis Pineda, IIMAS, 2010

### Derivaciones izquierdas (leftmost)

- Derivación más izquierda:  $\Rightarrow_{lm}$

$$- E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E)$$

$$\Rightarrow a + (a * E) \Rightarrow a + (a * a)$$

$$- E \Rightarrow_{lm}^* a + (a * a)$$

Dr. Luis Pineda, IIMAS, 2010

### Derivación más izquierda

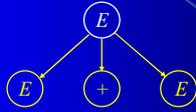
- $E$



Dr. Luis Pineda, IIMAS, 2010

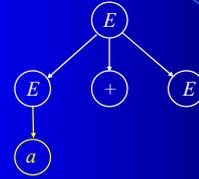
### Derivación más izquierda

- $E \Rightarrow E + E$



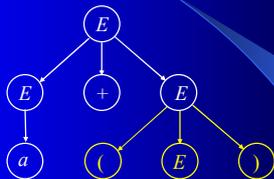
### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E$



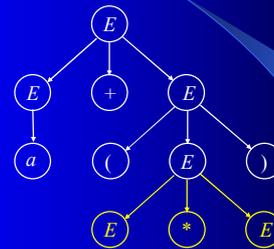
### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E)$



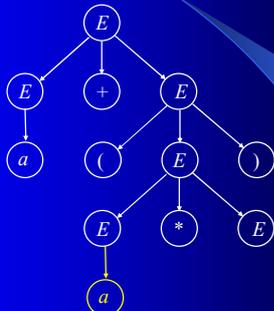
### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E)$



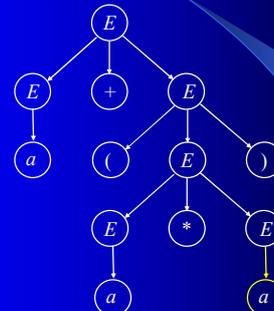
### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E) \Rightarrow a + (a * E)$



### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E) \Rightarrow a + (a * E) \Rightarrow a + (a * a)$



### Derivación más izquierda

- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E)$   
 $\Rightarrow a + (a * E) \Rightarrow a + (a * a)$



Dr. Luis Pineda, IIMAS, 2010

### Derivaciones derechas (rightmost)

- Derivación más derecha:  $\Rightarrow_{rm}$
- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$   
 $\Rightarrow E + (E * a) \Rightarrow E + (a * a) \Rightarrow a + (a * a)$
- $E \Rightarrow_{rm}^* a + (a * a)$

Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

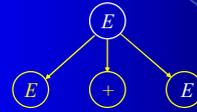
- $E$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

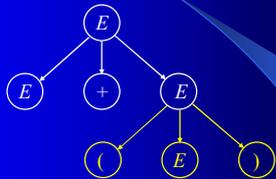
- $E \Rightarrow E + E$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

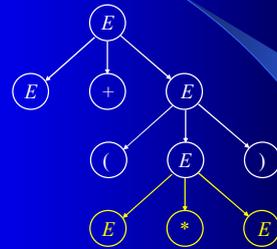
- $E \Rightarrow E + E \Rightarrow E + (E)$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

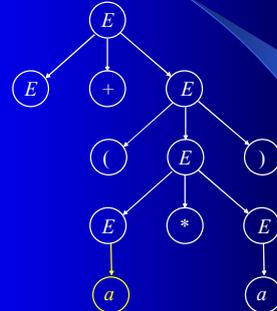
- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$   
 $\Rightarrow E + (E * a)$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$   
 $\Rightarrow E + (E * a) \Rightarrow E + (a * a)$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$   
 $\Rightarrow E + (E * a) \Rightarrow E + (a * a) \Rightarrow a + (a * a)$



Dr. Luis Pineda, IIMAS, 2010

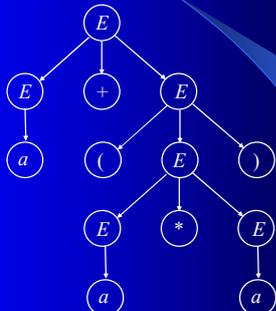
### ¿Cuántas derivaciones hay?

- Tipos de derivaciones:
  - Más izquierdas
  - Más derechas
  - Combinadas!
- Son equivalentes si generan la misma estructura global

Dr. Luis Pineda, IIMAS, 2010

### Derivación más izquierda

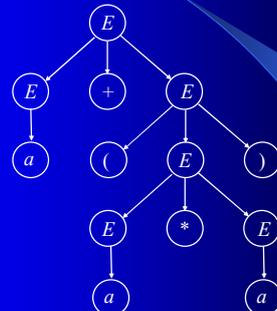
- $E \Rightarrow E + E \Rightarrow a + E \Rightarrow a + (E) \Rightarrow a + (E * E)$   
 $\Rightarrow a + (a * E) \Rightarrow a + (a * a)$



Dr. Luis Pineda, IIMAS, 2010

### Derivación más derecha

- $E \Rightarrow E + E \Rightarrow E + (E) \Rightarrow E + (E * E)$   
 $\Rightarrow E + (E * a) \Rightarrow E + (a * a) \Rightarrow a + (a * a)$



Dr. Luis Pineda, IIMAS, 2010

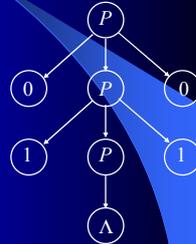
## Parseo

- Parseo: Inferencia recursiva para obtener la estructura sintáctica de una cadena en el lenguaje

0110



Parser



Dr. Luis Pineda, IIMAS, 2010

## Derivación & árbol de parseo

- Si hay una derivación hay también una inferencia recursiva
- Si hay una inferencia recursiva hay también un árbol de parseo
- Si hay un árbol de parseo hay derivaciones más derecha y más izquierda
- Si hay derivaciones más izquierdas y más derechas, hay una derivación!

Dr. Luis Pineda, IIMAS, 2010

## El lenguaje de una gramática

- Sea  $G = (V, \Sigma, S, P)$  una *GLC*. El lenguaje generado por  $G$  es:

$$L(G) = \{x \in \Sigma^* \mid S \Rightarrow_G^* x\}$$

- Si  $x \in \Sigma^*$  entonces  $x$  no tiene variables!
- El lenguaje generado por una gramática es un conjunto de cadenas!
- Un lenguaje  $L$  es *libre de contexto* (*LLC* o context-free *CFL*) si existe una *CFG*  $G$  tal que  $L = L(G)$

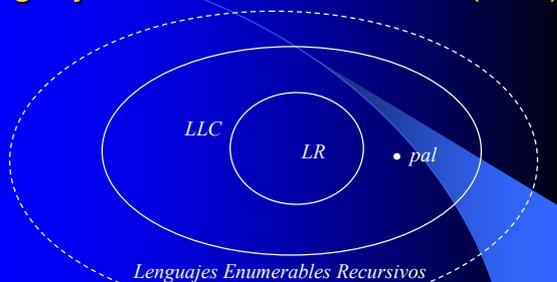
Dr. Luis Pineda, IIMAS, 2010

## El lenguaje de una gramática

- Formas sentenciales u oracionales:
  - generadas por derivaciones que empiezan con el símbolo inicial &
  - sin variables (i.e. cadenas del lenguaje).
- $L(G)$  consiste de las formas sentenciales u oracionales en  $\Sigma^*$

Dr. Luis Pineda, IIMAS, 2010

## Lenguajes Libres del Contexto (LLC)



- Todos los *LR* son *LLC* pero no todos los *LLC* son *LR*
- Hay también lenguajes que no son *LLC*

Dr. Luis Pineda, IIMAS, 2010