# Gramaticas Independientes del Contexto

Fabio Martínez Carrillo

Autómatas
Escuela de Ingeniería de Sistemas e Informatica
Universidad Industrial de Santander - UIS

11 de octubre de 2017

# Gramaticas libres de contexto

Notación formal para expresar definiciones recursivas de los lenguajes.

- Desarrollo de compiladores
- Procesamiento de lenguajes naturales
- Analizadores sintacticos
- Útil para describir estructuras anidadas
- Describen tipos de formatos como : XML, DTD

# Gramaticas independientes del contextos

1. **terminales (T):** Un conjunto finito de símbolos que forma las cadenas del lenguaje que se está definiendo. $\Sigma$
2. **variables (v):** Un conjunto finito de otros símbolos que representa el lenguaje $V$, $V \cap \Sigma = \emptyset$
3. **símbolo inicial (S).** variable con la que inicia la definición del lenguaje $V$
4. **Conjunto de producciones (P)** $V \times (V \cup \Sigma)^*$

$$G = (V, T, P, S)$$

# Regla de producción

## Definición

*variable(cabecera)* $\rightarrow$ cadena de variables y terminales (cuerpo)

## Convención

- A, B, C ... y S son variables
  - S normalmente es el símbolo inicial
- a, b, c son terminales
- X, Y, Z pueden ser terminales o variables
- w, x, y, z son cadenas de los terminales
- $\alpha, \beta$ cadenas formadas por símbolos terminales y/o variables.

## Gramatica en español

- $< \mathit{frase} > \rightarrow < \mathit{sujeto} > < \mathit{predicado} >$
- $< \mathit{sujeto} > \rightarrow < \mathit{articulo} > < \mathit{sustantivo} >$
- $< \mathit{articulo} > \rightarrow \mathit{el} \mid \mathit{la}$
- $< \mathit{sustantivo} > \rightarrow \mathit{perro} \mid \mathit{luna}$
- $< \mathit{predicado} > \rightarrow < \mathit{verbo} >$
- $< \mathit{verbo} > \rightarrow \mathit{brilla} \mid \mathit{corre}$

- $< \mathit{frase} >$ es el simbolo inicial
- Se aplica recursivamente reglas gramticales hasta que no hayan variables
- La palabra generada hace parte de la gramatica. Por ejemplo:
  - *"el perro corre"*
  - *" el perro brilla"*

# Ejemplo: $\{0^n 1^n \mid n \geq 1\}$

- $S \rightarrow 01$
- $S \rightarrow 0S1$

- **Base** 01 es el lenguaje
- **Inducción** Si $w$ esta en el lenguaje, entonces es $0w1$

# Ejemplo de la Gramatica

$$\{0^n 1^n \mid n \geq 1\}$$

- Terminales: $\{0, 1\}$ constantes
- variables: $\{S\}$
- Simbolo inical $S$
- Reglas de la producción
  - $S \rightarrow 01$
  - $S \rightarrow 0S1$

$$G = (\{S\}, \{0, 1\}\ P, S)$$

# Ejemplo 2: generación de expresiones aritmeticas

## reglas

1. $E \to E + T$
2. $E \to T$
3. $T \to T * F$
4. $T \to F$
5. $F \to CF$
6. $F \to C$
7. $C \to 0|1|2|3|4|5|6|7|8|9$

- Simbolo inicial: $E$
- terminales: $\{+, *, 0, \ldots, 9\}$
- variables: $E, T, F, C$

# Generación de la expresión: $25 + 3 * 12$

| EXPRESION | JUSTIFICACION |
|---|---|
| $E$ | Símbolo inicial, inicia derivación |
| $\Rightarrow E + T$ | Aplicación 1a. regla |
| $\Rightarrow T + T$ | 2a. regla, sobre la $E$ |
| $\Rightarrow F + T$ | 4a. regla, sobre la $T$ izquierda |
| $\Rightarrow CF + T$ | 5a. regla, sobre $F$ |
| $\Rightarrow 2F + T$ | 7a. regla |
| $\Rightarrow 2C + T$ | 6a. regla |
| $\Rightarrow 25 + T$ | 7a. regla |
| $\Rightarrow 25 + T * F$ | 3a. regla |
| $\Rightarrow 25 + F * F$ | 4a. regla |
| $\Rightarrow 25 + C * F$ | 6a. regla, sobre la $F$ izquierda |
| $\Rightarrow 25 + 3 * F$ | 7a. regla |
| $\Rightarrow 25 + 3 * CF$ | 5a. regla |
| $\Rightarrow 25 + 3 * 1F$ | 7a. regla |
| $\Rightarrow 25 + 3 * 1C$ | 6a. regla |
| $\Rightarrow 25 + 3 * 12$ | 7a. regla |

# Derivación utilizando a gramatica

El lenguaje de la gramática son todas las cadenas de terminales que se pueden construir utilizando **derivación.**

- Una cadena $\alpha \in (V \cup \Sigma)^*$ es derivable a partir de la gramatica $(\{V\}, \Sigma, P, S)$, si hay al menos una secuencia de pasos de derivación que la produce a partir del símbolo inicial $S$

$$S \Rightarrow \cdots \alpha$$

**lenguaje $L(G)$**

Palabras construidas por terminales derivables a partir del simbolo inicial

$$\{w \in \Sigma^* \mid S \Rightarrow \cdots w\}$$

# Derivación utilizando a gramatica

## Entonces podemos decir

$$\alpha \boldsymbol{A} \beta \Rightarrow \alpha \gamma \beta$$

Si $A \Rightarrow \gamma$ es una regla de producción

## Ejemplo: $S \Rightarrow 01$; $S \Rightarrow 0S1$

$$S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 000111$$

# Derivación iterativa

Utilizar las reglas de producción para inferir ciertas cadenas que pertenecen al lenguaje de una variable.

- El simbolo inicial se expande utilizando producciones iterativamente hasta obeter una cadena de terminales
- El lenguaje de la gramática son todas las cadenas de terminales que se pueden obtener de esta forma

- Podemos extender $\Rightarrow$ hacia $\Rightarrow^*$ para indicar cero o mas derivaciones
- Operación equivalente a pasar $\delta$ hasta $\hat{\delta}$

# Derivación iterativas

Podemos extender $\Rightarrow$ hacia $\Rightarrow^*$ para indicar cero o mas derivaciones

- **Caso Base:** $\alpha \Rightarrow^* \alpha$ . Cualquier cadena se deriva a sí misma
- **paso Inductivo:** Si $\alpha \Rightarrow^* \beta$ y $\beta \Rightarrow \gamma$ entonces $\alpha \Rightarrow^* \gamma$. Indica que hay una secuencia de cadenas $\gamma_1, \gamma_2, \ldots, \gamma_n$ para $n \geq 1$
  - $\alpha = \gamma_1$
  - $\beta = \gamma_n$
  - Para $i = 1, 2, \ldots, n - 1$ tenemos $\gamma_i \Rightarrow \gamma_{i+1}$

# Ejemplo derivación iterativa

## $S \rightarrow 01$; $S \rightarrow 0S1$

- $S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 000111$
- Entonces:
  - $S \Rightarrow^* S$
  - $S \Rightarrow^* 0S1$
  - $S \Rightarrow^* 00S11$
  - $S \Rightarrow^* 000111$

# Derivación interativa

Considere la expresión: $a * (a + b00)$.

- Defina la expresión de la gramatica
- Pruebe si la cadena es una palabra de la gramatica

| | | | |
|---|---|---|---|
| 1. | $E$ | $\rightarrow$ | $I$ |
| 2. | $E$ | $\rightarrow$ | $E + E$ |
| 3. | $E$ | $\rightarrow$ | $E * E$ |
| 4. | $E$ | $\rightarrow$ | $(E)$ |
| 5. | $I$ | $\rightarrow$ | $a$ |
| 6. | $I$ | $\rightarrow$ | $b$ |
| 7. | $I$ | $\rightarrow$ | $Ia$ |
| 8. | $I$ | $\rightarrow$ | $Ib$ |
| 9. | $I$ | $\rightarrow$ | $I0$ |
| 10. | $I$ | $\rightarrow$ | $I1$ |

# Derivación iterativa

$$E \Rightarrow E * E \Rightarrow I * E \Rightarrow a * E \Rightarrow$$

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

$$a * (a + I) \Rightarrow a * (a + I0) \Rightarrow a * (a + I00) \Rightarrow a * (a + b00)$$

## Formas sentenciales

$$E \Rightarrow^* a * (a + b00)$$

# Derivaciones de Izquierda y Derecha

La derivación nos permite reemplazar cualquier variable en una cadena.

- Esto conduce a diferentes derivaciones para la misma cadena

Podemos forzar para reemplazar la variable más a la izquierda/derecha

# Derivación hacia la izquierda

## Derivación hacia la izquierda

$$wA\alpha \Rightarrow_{lm} w\beta\alpha$$

Si $w$ es una cadena terminal y existe la regla de producción $A \rightarrow \beta$

## Derivación hacia la derecha

$$A\alpha Aw \Rightarrow_{rm} w\beta\alpha$$

Si $w$ es una cadena terminal y existe la regla de producción $A \rightarrow \beta$

## Gramatica de parentesis balanceados (LM)

- $S \rightarrow SS \mid (S) \mid ()$

$$S \Rightarrow^*_{lm} SS \Rightarrow^*_{lm} (S)S \Rightarrow^*_{lm} (())S \Rightarrow^*_{lm} (())()$$

# Ejercicio

Describa la expresión $a * (a + b00)$ como una derivación por izquierda y por derecha

# Ejercicio

Describa la expresión $a * (a + b00)$ como una derivación por izquierda y por derecha

$$E \underset{lm}{\Rightarrow} E*E \underset{lm}{\Rightarrow} I*E \underset{lm}{\Rightarrow} a*E \underset{lm}{\Rightarrow}$$

$$a*(E) \underset{lm}{\Rightarrow} a*(E+E) \underset{lm}{\Rightarrow} a*(I+E) \underset{lm}{\Rightarrow} a*(a+E) \underset{lm}{\Rightarrow}$$

$$a*(a+I) \underset{lm}{\Rightarrow} a*(a+I0) \underset{lm}{\Rightarrow} a*(a+I00) \underset{lm}{\Rightarrow} a*(a+b00)$$

Figura 1: Por izquierda

# Ejercicio

Describa la expresión $a * (a + b00)$ como una derivación por izquierda y por derecha

$$E \underset{rm}{\Rightarrow} E*E \underset{rm}{\Rightarrow} E*(E) \underset{rm}{\Rightarrow} E*(E+E) \underset{rm}{\Rightarrow}$$

$$E*(E+I) \underset{rm}{\Rightarrow} E*(E+I0) \underset{rm}{\Rightarrow} E*(E+I00) \underset{rm}{\Rightarrow} E*(E+b00) \underset{rm}{\Rightarrow}$$

$$E*(I+b00) \underset{rm}{\Rightarrow} E*(a+b00) \underset{rm}{\Rightarrow} I*(a+b00) \underset{rm}{\Rightarrow} a*(a+b00)$$

Figura 1: Por derecha

Muchas gracias por su atención

Universidad Industrial de Santander