

Ciencias de la Computación I

Gramáticas Sensibles al Contexto y Lenguajes Sensibles al Contexto

Octubre 2012

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Gramáticas Formales

Una gramática formal es una cuadrupla $G = \langle N, T, P, S \rangle$

N = conjunto finito de símbolos no terminales

T = conjunto finito de símbolos terminales

$$\left. \begin{array}{l} N \\ T \end{array} \right\} N \cap T = \emptyset$$

S = símbolo distinguido o axioma $S \notin (N \cup T)$

P = conjunto finito de reglas de producción (permiten generar cadenas a partir de S)

$$\left. \begin{array}{l} \alpha \rightarrow \beta \\ \alpha = \varphi A \rho \\ \beta = \varphi \omega \rho \end{array} \right\} \begin{array}{l} A \in N \cup \{S\} \\ \varphi, \omega, \rho \in (N \cup T)^* \end{array}$$

De acuerdo a formato de reglas se pueden definir 4 tipos de gramáticas y sus correspondientes lenguajes

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Gramáticas Sensibles al Contexto (GSC) (Tipo 1)

- Generan los lenguajes sensibles al contexto (reconocidos por ALA)
- Se definen como una cuadrupla $G = \langle N, T, P, S \rangle$

N = conjunto finito de símbolos no terminales

$$N \cap T = \emptyset$$

T = conjunto finito de símbolos terminales

S = símbolo distinguido o axioma $S \notin (N \cup T)$

P = conjunto finito de reglas de producción

$$\gamma A \beta \rightarrow \gamma w \beta$$
$$A \in N \cup \{S\}$$
$$\gamma, \beta \in (\mathbf{N} \cup \mathbf{T})^*$$
$$w \in (N \cup T)^* - \{\varepsilon\}$$
$$S \rightarrow \epsilon$$

Se puede incluir para generar la cadena vacía

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Gramáticas Sensibles al Contexto (Tipo 1)

Ejemplo:

$L = \{a^n b^n c^n / n > 0\}$ Lenguaje Sensible al Contexto

$S \rightarrow A$
 $A \rightarrow aABC$
 $A \rightarrow aBC$

Generar

CB \rightarrow BC } Ordenar

$aB \rightarrow ab$
 $bB \rightarrow bb$
 $bC \rightarrow bc$
 $cC \rightarrow cc$

Reemplazar con contexto

$$G = \langle \{A, B, C\}, \{a, b, c\}, P, S \rangle$$

Ejemplo de Derivación

 $S \Rightarrow$
$$A \Rightarrow$$

aABC \Rightarrow

aaABCBC \Rightarrow

aaaBCBCBC \Rightarrow

aaaBBCBC \Rightarrow

aaaBB**C**CC \Rightarrow

aa**a**BBCCC \Rightarrow

aaabBBCCC \Rightarrow

aaab**bB**CCC \Rightarrow

aaabb**bC**CC \Rightarrow

aaabbb**cCC**⇒

aaabbbccC \Rightarrow

aaabbbccc esta cadena
de símbolos terminales $\in L(G)$

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Gramáticas Sensibles al Contexto (Tipo 1)

Ejemplo:

$L = \{a^n b^n c^n / n > 0\}$ Lenguaje Sensible al Contexto

$S \rightarrow A$
 $A \rightarrow aABC$
 $A \rightarrow aBC$

Generar

$CB \rightarrow BC$

Ordenar

$aB \rightarrow ab$
 $bB \rightarrow bb$
 $bC \rightarrow bc$
 $cC \rightarrow cc$

Reemplazar
con contexto

$G = \langle \{A, B, C\}, \{a, b, c\}, P, S \rangle$

Ejemplo de derivación

$S \Rightarrow$

$A \Rightarrow$

$aABC \Rightarrow$

$aaABCBC \Rightarrow$

$aaaBCBCBC \Rightarrow$

$aaabCBCBC \Rightarrow$

$aaabcBCBC$ (no se puede
llegar de ninguna manera a
reemplazar todos los símbolos
no terminales por símbolos
terminales $\Rightarrow G$ no genera
cadenas incorrectas)

Gramáticas Sensibles al Contexto (Tipo 1)

Formato reglas de tipo 1

Ejemplo: $\gamma A \beta \rightarrow \gamma \omega \beta$ $A \in N \cup \{S\}$ $\gamma, \beta \in (N \cup T)^*$ $\omega \in (N \cup T)^* - \{\epsilon\}$

Sea $G = \langle \{A, B, C\}, \{a, b, c\}, P, S \rangle$ donde

$P = \{ \underset{A}{\boxed{S}} \rightarrow \underset{\omega}{\boxed{A}} \mid \gamma = \epsilon \text{ y } \beta = \epsilon \}$

$\underset{A}{\boxed{A}} \rightarrow \underset{\omega}{\boxed{aABC}}, \gamma = \epsilon \text{ y } \beta = \epsilon$

$\underset{A}{\boxed{A}} \rightarrow \underset{\omega}{\boxed{abC}}, \gamma = \epsilon \text{ y } \beta = \epsilon$

$\underset{\gamma A}{\boxed{bB}} \rightarrow \underset{\gamma \omega}{\boxed{bb}}, \beta = \epsilon$

$\underset{\gamma A}{\boxed{bC}} \rightarrow \underset{\gamma \omega}{\boxed{bc}}, \beta = \epsilon$

$\underset{\gamma A}{\boxed{cC}} \rightarrow \underset{\gamma \omega}{\boxed{cc}}, \beta = \epsilon$

$\underset{A \beta}{\boxed{CB}} \rightarrow \underset{\gamma A}{\boxed{BC}} (*)$

β cambia
 γA γ cambia

Aclaración:

$CB \rightarrow BC$ no respeta el formato
de reglas de Tipo 1

Se utiliza para "resumir" el efecto
de las siguientes 3 reglas:

$AB \rightarrow XB$

$XB \rightarrow XA$

$XA \rightarrow BA$

donde los símbolos $A, B, X \in N$

(*) En la práctica se puede usar esta regla
aclarando que NO es de Tipo 1

Gramáticas Sensibles al Contexto (Tipo 1)

- Las reglas de producción indican cómo reemplazar un no terminal o el símbolo distinguido teniendo en cuenta el contexto en el que se encuentra

Ejemplo: G sensible al contexto

Sea $G = \langle \{A, B, C\}, \{a, b, c\}, P, S \rangle$ donde

$P = \{S \rightarrow A, A \rightarrow aABC, A \rightarrow abC, CB \rightarrow BC, bB \rightarrow bb, bC \rightarrow bc, cC \rightarrow cc\}$

Del lado izquierdo puede haber símbolos terminales y/o no terminales (también puede aparecer el símbolo distinguido)

$S \Rightarrow A \Rightarrow aABC \Rightarrow aabCBC \Rightarrow aabBCC \Rightarrow$
 $aabbCC \Rightarrow aabbCC \Rightarrow aabbcc$

Se reemplaza el no terminal y su contexto por la cadena del lado derecho de la regla de producción (manteniendo el contexto)

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Gramáticas Sensibles al Contexto (Tipo 1)

Derivación:

Dadas dos cadenas ω_1 , ω_2 y una gramática sensible al contexto $G = \langle N, T, P, S \rangle$ se dice que ω_1 deriva a ω_2 si ω_2 puede obtenerse a partir de ω_1 por aplicación de reglas de G.

Derivación inmediata:

La cadena ω_2 se obtiene de la cadena ω_1 en un paso usando las reglas de P

$\omega_1 \Rightarrow \omega_2$ sí y sólo sí

1) Cuando $\omega_1 = S$ y $\omega_2 = \varepsilon$ y $S \rightarrow \varepsilon \in P$

2) $\omega_1 = \alpha B \beta$

$\omega_2 = \alpha \omega \beta$ $\alpha, \beta \in (N \cup T)^*$ $B \in N \cup \{S\}$ $\omega_1 \neq \omega_2$

$\alpha B \beta \rightarrow \alpha \omega \beta \in P$ y $\omega \in (N \cup T)^* - \{\varepsilon\}$

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Gramáticas Sensibles al Contexto (Tipo 1)

Derivación:

La cadena α_2 se obtiene de la cadena α_1 en cero o más pasos usando las reglas de P.

Si existen $\alpha_1, \alpha_2, \dots, \alpha_n \in (N \cup T)^*$ $n \geq 1$ y

$$\alpha_1 \Rightarrow \alpha_2 \Rightarrow \dots \Rightarrow \alpha_n \quad \text{decimos que } \alpha_1 \overset{*}{\Rightarrow} \alpha_n$$

(clausura reflexiva y transitiva de \Rightarrow)

Lenguaje generado por una gramática sensible al contexto $G = \langle N, T, P, S \rangle$

$$L(G) = \{ x / x \in T^* y \mid S \overset{*}{\Rightarrow} x \}$$

Es decir, una cadena $\in L(G)$ si:

- 1) La cadena está formada por símbolos terminales únicamente
- 2) La cadena puede ser derivada a partir de S

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Jerarquía de Chomsky

Lenguajes	Máquinas	Gramáticas	Equiv. Det. y No det
R E G U L A R E S (TIPO 3)	Autómata Finito Determinístico $AFD = \langle E, A, \delta, e_0, F \rangle$ E: conjunto finito de estados A: alfabeto de entrada δ: función de transición $\delta: E \times A \rightarrow E$ e_0: estado inicial; $e_0 \in E$ F: conjunto de estados finales; $F \subseteq E$	Regulares o de Tipo 3 $G = \langle N, T, P, S \rangle$ Formato de reglas de Tipo 3: - Lineales a derecha $A \rightarrow aB$ $A \in N \cup \{S\}$ $A \rightarrow a$ $B \in N$ $S \rightarrow \epsilon$ $a \in T$ - Lineales a izquierda $A \rightarrow Ba$ $A \in N \cup \{S\}$ $A \rightarrow a$ $B \in N$ $S \rightarrow \epsilon$ $a \in T$	SI

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Jerarquía de Chomsky

Lenguajes	Máquinas	Gramáticas	EQ. DET. y NO DET
L I B R E S D E L C O N T E N T O (TIPO 2)	Autómata de Pila Determinístico o No Determinístico $AP = \langle E, A, P, \delta, e_0, Z_0, F \rangle$ E: conjunto finito de estados A: alfabeto de entrada P: alfabeto de la Pila; $P \cap A = \emptyset$ δ: función de transición $\delta: E \times (A \cup \{\epsilon\}) \times P \rightarrow E \times P^*$ (determinístico) $\delta: E \times (A \cup \{\epsilon\}) \times P \rightarrow P_f(E \times P^*)$ (no determinístico) (P_f denota los subconjuntos finitos de $E \times P^*$) e_0 : estado inicial; $e_0 \in E$ Z_0 : símbolo distinguido; $Z_0 \in P$ F: conjunto de estados finales; $F \subseteq E$.	Libres del Contexto o de Tipo 2 $G = \langle N, T, P, S \rangle$ Formato reglas de Tipo 2: $A \rightarrow \omega$ donde $A \in N \cup \{S\}$; $\omega \in (N \cup T)^* - \{\epsilon\}$ Se puede incluir $S \rightarrow \epsilon$	NO

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Jerarquía de Chomsky

Lenguajes	Máquinas	Gramáticas	EQ. DET. y NO DET
S E N S I B L E S A L C O N T E N T O (TIPO 1)	Autómata Linealmente Acotado $ALA = \langle E, A, C, \delta, e_0, B, F, \#, \$ \rangle$ E: conjunto finito de estados A: alfabeto de entrada; $A \subseteq C$ C: alfabeto de la cinta; $C = A \cup \{B, \#, \$\} \cup \text{Auxiliares}$ δ: función de transición $\delta: E \times C \rightarrow E \times C \times \{D, I, N\}$ (1 cinta) (*) $\delta: E \times C^k \rightarrow E \times (C \times \{D, I, N\})^k$ (k cintas) (*) e_0 : estado inicial; $e_0 \in E$ B: símbolo blanco; $B \notin A$ y $B \in C$ F: conjunto de estados finales; $F \subseteq E$ $\#$: símbolo de inicio de la/s cinta/s C $\$$: símbolo de fin de la/s cinta/s C (*) En ninguna de las cintas se permiten movimientos a izquierda de $\#$ ni a derecha de $\$$. Tampoco se permite reescribir los símbolos $\#$ y $\$$.	Sensibles al Contexto o de Tipo 1 $G = \langle N, T, P, S \rangle$ Formato reglas Tipo 1: $\gamma A \beta \rightarrow \gamma w \beta$ donde $A \in N \cup \{S\}$; $\gamma, \beta \in (N \cup T)^*$ $w \in (N \cup T)^* - \{\epsilon\}$ Se puede incluir $S \rightarrow \epsilon$	SI

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2012

Jerarquía de Chomsky

Lenguajes	Máquinas	Gramáticas	EQ. DET. y NO DET
E S T R U C T U R A D O S (TIPO 0)	Máquina de Turing Determinística $MTD = \langle E, A, C, \delta, e_0, B, F \rangle$ E: conjunto finito de estados A: alfabeto de entrada; $A \subseteq C$ C: alfabeto de cinta; $C = A \cup \{B\} \cup \text{Auxiliares}$ δ: función de transición δ: $E \times C \rightarrow E \times C \times \{D, L, N\}$ (1 cinta) δ: $E \times C^k \rightarrow E \times (C \times \{D, L, N\})^k$ (k cintas) e_0 : estado inicial; $e_0 \in E$ B: símbolo blanco; $B \notin A$ y $B \in C$ F: conjunto de estados finales; $F \subseteq E$	Contractivas o de Tipo 0 $G = \langle N, T, P, S \rangle$ Formato reglas Tipo 0: $\gamma A \beta \rightarrow \gamma w \beta$ donde $A \in N \cup \{S\}$; $\gamma, \beta, w \in (N \cup T)^*$ (w puede ser ϵ)	SI

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Jerarquía de Lenguajes



Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Ejemplo 2

$L = \{a^n b^n c^n d^n / n > 0\}$ Lenguaje Sensible al Contexto

P: { $S \rightarrow A$
 $A \rightarrow aABCD$
 $A \rightarrow aBCD$
 $DB \rightarrow BD$
 $CB \rightarrow BC$
 $DC \rightarrow CD$
 $aB \rightarrow ab$
 $bB \rightarrow bb$
 $bC \rightarrow bc$
 $cC \rightarrow cc$
 $cD \rightarrow cd$
 $dD \rightarrow dd$ }

$G = \langle \{A, B, C, D\}, \{a, b, c, d\}, P, S \rangle$

Derivación

$S \Rightarrow$
 $A \Rightarrow$
 $aABCD \Rightarrow$
 $aaABCDBCD \Rightarrow$
 $aaaBCDBCBBCD \Rightarrow$
 $aaaBCBDCBDCD \Rightarrow$
 $aaaBCBCDBDCD \Rightarrow$
 $aaaBCBCDBDCD \Rightarrow$
 $aaaBCBCDBDCD \Rightarrow$
 $aaaBCBCDBDCD \Rightarrow$
 $aaaBBCCBCDD \Rightarrow$
 $...$
 $aaaBBBCCCDDD \Rightarrow$
 $...$
 $aaabbbcccd$

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Ejemplo 3

$L = \{a^n g^i h^i c^n d^n / n, i \geq 0\}$ Lenguaje Sensible al Contexto

P: { $S \rightarrow A$
 $A \rightarrow aACD$
 $A \rightarrow B$
 $B \rightarrow gBh$
 $B \rightarrow gh$
 $DC \rightarrow CD$
 $hC \rightarrow hc$
 $cC \rightarrow cC$
 $cD \rightarrow cd$
 $dD \rightarrow dd$
 $A \rightarrow acD$
 $S \rightarrow \varepsilon$ }

si $i=0$
 si $n=i=0$

$G = \langle \{A, B, C, D\}, \{a, g, h, c, d\}, P, S \rangle$

Derivación

$S \Rightarrow$
 $A \Rightarrow$
 $aACD \Rightarrow$
 $aaACDCD \Rightarrow$
 $aaBCDCD \Rightarrow$
 $aagBhCD \Rightarrow$
 $aagghhCD \Rightarrow$
 $aagghhCCDD \Rightarrow$
 $...$
 $aagghhccdd$

Derivación

$S \Rightarrow$
 $A \Rightarrow$
 $aACD \Rightarrow$
 $aacDCD \Rightarrow$
 $aaCDD \Rightarrow$
 $aaccDD \Rightarrow$
 $aaccdd \Rightarrow$
 $aaccdd$

Derivación

$S \Rightarrow$
 $A \Rightarrow$
 $B \Rightarrow$
 $gBh \Rightarrow$
 $gghh$

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Ejemplo 4

$$L = \{a^n c^i b^n d^i / n, i \geq 0\}$$

Lenguaje Sensible al Contexto

P: {S → A
A → aAB
A → H
H → cHD
H → cD
DB → BD
cB → cb
bB → bb
bD → bd
dD → dd
A → ab si i=0
cD → cd si n=0
S → ε} si n=i=0

Derivación
S ⇒
A ⇒
aAB ⇒
aaABB ⇒
aaHBB ⇒
aacHDBB ⇒
aacCHDDBB ⇒
aaccDDDBB ⇒
...
aacc**c**BDDDD ⇒
...
aacc**c**bbddd

Derivación
S ⇒
A ⇒
aAB ⇒
aab**B** ⇒
aabb

Derivación
S ⇒
A ⇒
H ⇒
cHD ⇒
cc**D**D ⇒
cc**d**D ⇒
ccdd

$$G = \langle \{A, B, H, D\}, \{a, c, d, b\}, P, S \rangle$$

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Gramática incorrecta

$$L = \{a^n b^n c^n / n > 0\}$$

Lenguaje Sensible al Contexto

Supongamos que construimos una gramática

S → A
A → aAbc
A → abc

S ⇒ A ⇒ aAbc ⇒ aaAbcbcb ⇒ aaabcbcbcb

Genera las cantidades correctas pero desordenadas.
Una vez que se escribieron, los símbolos terminales no se pueden cambiar de lugar; por lo tanto se generan cadenas que no pertenecen al lenguaje que se quiere generar. **No es correcto**

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

REGLAS DE GENERACIÓN: PASO 1

$$L = \{a^n b^n c^n / n > 0\}$$

Lenguaje Sensible al Contexto

$S \rightarrow A$
 $A \rightarrow aABC$
 $A \rightarrow aBC$

$S \Rightarrow A \Rightarrow aABC \Rightarrow aaABCBC \Rightarrow aaaBCBCBC$

Todo lo que está desordenado debería generarse con
símbolos NO TERMINALES para poder ordenarlos

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

REGLAS DE ORDENAMIENTO: PASO 2

$$L = \{a^n b^n c^{2n} / n > 0\}$$

Lenguaje Sensible al Contexto

$S \rightarrow A$
 $A \rightarrow aABC$
 $A \rightarrow aBC$
 $CB \rightarrow BC$

$S \Rightarrow$
 $A \Rightarrow$
 $aABC \Rightarrow$
 $aaABCBC \Rightarrow$
 $aaaB\textcolor{brown}{C}BCBC \Rightarrow$
 $aaaBB\textcolor{brown}{C}C\textcolor{brown}{B}C \Rightarrow$
 $aaaBB\textcolor{brown}{C}BCC \Rightarrow$
 $aaaBBBCCC \Rightarrow$
....

Una vez que los símbolos no terminales están ordenados, se tendrían que
reemplazar por símbolos terminales

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

GRAMÁTICA CON REEMPLAZOS INCORRECTOS

$$L=\{a^n b^n c^n / n>0\}$$

$S \rightarrow A$
 $A \rightarrow aABC$
 $A \rightarrow aBC$
 $CB \rightarrow BC$

Lenguaje Sensible al Contexto

Ejemplo de derivación

$S \Rightarrow$

$A \Rightarrow$

$aABC \Rightarrow$

$aaABCBC \Rightarrow$

$aaaB\textcolor{red}{C}BCBC \Rightarrow$

$aaaBB\textcolor{red}{C}C\textcolor{red}{B}C \Rightarrow$

$aaaBB\textcolor{red}{C}BCC \Rightarrow$

$aaaBBBCCC \Rightarrow$

Una vez que los símbolos no terminales están ordenados se tendrían que reemplazar por símbolos terminales

~~$B \rightarrow b$~~
 ~~$C \rightarrow c$~~

Si ponemos reglas independientes del contexto, se pueden usar en cualquier momento (por ejemplo en el paso 5); por lo tanto también podemos generar cadenas que tienen sus símbolos desordenados

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012

Gramáticas Sensibles al Contexto (Tipo 1)

Formato reglas de tipo 1 $\gamma A \beta \rightarrow \gamma w \beta$ $A \in N \cup \{S\}$ $\gamma, \beta \in (N \cup T)^*$ $w \in (N \cup T)^+ - \{\epsilon\}$

γ	A	β	\rightarrow	γ	w	β
ϵ	S	ϵ	\rightarrow	ϵ	A	ϵ
ϵ	A	ϵ	\rightarrow	ϵ	$aABC$	ϵ
ϵ	A	ϵ	\rightarrow	ϵ	abC	ϵ
b	B	ϵ	\rightarrow	b	b	ϵ
b	C	ϵ	\rightarrow	b	c	ϵ
c	C	ϵ	\rightarrow	c	c	ϵ

Aclaración:

$CB \rightarrow BC$ no respeta el formato de reglas de Tipo 1

Se utiliza para "resumir" el efecto de las siguientes 3 reglas::

$AB \rightarrow XB$

$XB \rightarrow XA$

$XA \rightarrow BA$

donde los símbolos $A, B, X \in N$

(*) En la práctica se puede usar, aclarando que no es tipo 1

γ	A	β	\rightarrow	γ	w	β
ϵ	C	B	\rightarrow	ϵ	BC	cambia
C	B	ϵ	\rightarrow	cambia	BC	ϵ

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2012