

Ciencias de la Computación I

Gramáticas Sensibles al Contexto y Lenguajes Sensibles al Contexto

Octubre 2009

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2009

Gramáticas Formales

Una gramática formal es una cuadrupla $G = (N, T, P, S)$

N = conjunto finito de símbolos no terminales

T = conjunto finito de símbolos terminales

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

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

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

$\alpha \rightarrow \beta$

$\alpha = \varphi A \rho$

$\beta = \varphi \omega \rho$

$A \in N \cup \{S\}$

$\varphi, \omega, \rho \in (N \cup T)^*$

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 - 2009

Gramáticas Sensibles al Contexto (GSC) (Tipo 1)

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

N = conjunto finito de símbolos no terminales

T = conjunto finito de símbolos terminales

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

P = conjunto finito de reglas de producción

$\gamma A \beta \rightarrow \gamma w \beta$

$A \in N \cup \{S\}$

$\gamma, \beta \in (N \cup T)^*$

$w \in (N \cup T)^* - \{\epsilon\}$

Se puede agregar $S \rightarrow \epsilon$ (para generar la cadena vacía)

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2009

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

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

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2009

Gramáticas Sensibles al Contexto (Tipo 1)

Formato reglas de tipo 1

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

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

$P = \{S \rightarrow A,$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad A \quad \omega$

$A \rightarrow aABC,$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad A \quad \omega$

$A \rightarrow abC,$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad A \quad \omega$

$bB \rightarrow bp,$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad \gamma A \quad \gamma \omega$

$bC \rightarrow bc,$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad \gamma A \quad \gamma \omega$

$cC \rightarrow cc\}$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad \gamma A \quad \gamma \omega$

$CB \rightarrow BC$
 $\quad \underbrace{\quad} \quad \underbrace{\quad}$
 $\quad A \beta$
 $\quad \underbrace{\quad}$
 $\quad \gamma A$

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 $A, B, X \in N$

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 $G = (\{A, B, C\}, \{a, b, c\}, S, P)$ 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

Gramáticas Sensibles al Contexto (Tipo 1)

Derivación:

Dadas dos cadenas ω_1 , ω_2 y una gramática sensible al contexto $G = (N, T, P, S)$ 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 - 2009

Gramáticas Sensibles al Contexto (Tipo 1)

Derivación:

La cadena α_2 se obtiene de la cadena α_1 en zero 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

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

(clausura reflexiva y transitiva de \Rightarrow)

Lenguaje generado por una gramática sensible al contexto $G = (N, T, P, S)$:

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

Es decir, una cadena $x \in L(G)$ 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 - 2009

Ciencias de la Computación I

Jerarquía de Chomsky

Octubre 2009

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2009

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: conj. finito estados A: alfabeto de entrada δ: función de transición $\delta: E \times A \rightarrow E$ e_0: estado inicial; $e_0 \in E$ F: conj. estados finales; $F \subseteq E$	Regulares o de Tipo 3 $G = (N, T, P, S)$ Formato producciones tipo3: - Lineales a derecha $A \rightarrow aB \quad A \in N \cup \{S\}$ $A \rightarrow a \quad B \in N$ $S \rightarrow \epsilon \quad a \in T$ - Lineales a izquierda $A \rightarrow Ba \quad A \in N \cup \{S\}$ $A \rightarrow a \quad B \in N$ $S \rightarrow \epsilon \quad a \in T$	SI

Ciencias de la Computación I - Filminas de Clase - Facultad Cs. Exactas - UNCPBA - 2009

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 X T O (TIPO 2)	Autómata de Pila Determinístico o No Determinístico $AP = \langle E, A, P, \delta, e_0, Z_0, F \rangle$ E: conj. finito estados A: alfabeto de entrada P: alfabeto de la Pila; $P \cap A = \emptyset$ δ: función de transición $\delta: E \times (A \cup \{\epsilon\}) \times P \rightarrow E \times P^*$ (determin.) $\delta: E \times (A \cup \{\epsilon\}) \times P \rightarrow P_f(E \times P^*)$ (no determ.) (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: conj. estados finales; $F \subseteq E$	Libres del Contexto o de Tipo 2 $G = (N, T, P, S)$ Formato producciones $A \rightarrow \omega$ 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 - 2009

Jerarquía de Máquinas

Lenguajes	Máquinas	Gramáticas	EQ. DET. y NO DET
S E N S I B L E S A L C O N T E X T O (TIPO 1)	Autómata Linealmente Acotado $ALA = \langle E, A, C, d, e_0, B, F, \#, \$ \rangle$ E: conj. finito 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) (idem (*) para cada cinta) e_0: estado inicial; $e_0 \in E$ B: símbolo blanco; $B \notin A \text{ y } B \in C$ F: conj. estados finales; $F \subseteq E$ #: símbolo de inicio de la cinta C \$_: símbolo de fin de la cinta C (*) $\delta: (ei, a) \rightarrow (ej, \#, I)$ no está definida $\delta: (ei, a) \rightarrow (ej, \$, D)$ no está definida $a \in C \text{ y } ei, ej \in E$	Sensibles al Contexto o de Tipo 1 $G = (N, T, P, S)$ Formato producciones $\gamma A \beta \rightarrow \gamma w \beta$ 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 - 2009

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= $\langle E, A, C, d, e_0, B, F \rangle$ E: conj. finito estados A: alfabeto de entrada; $A \subseteq C$ C: alfabeto de 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: conj. estados finales; $F \subseteq E$	Contractivas o de Tipo 0 $G = (N, T, P, S)$ Formato producciones $\gamma A \beta \rightarrow \gamma w \beta$ donde $A \in N \cup \{S\}$; $\gamma, \beta, \omega \in (N \cup T)^*$ (ω puede ser ϵ)	SI

Ciencias de la Computación I - Filminas de Clase – Facultad Cs. Exactas – UNCPBA - 2009

Jerarquía de Lenguajes

