Gramáticas LL

Clase 23

IIC 2223

Prof. Cristian Riveros

Definiciones de prefijos (recordatorio)

Definiciones

$$w|_k \;=\; \begin{cases} a_1 \ldots a_n & \text{si } n \leq k \\ a_1 \ldots a_k & \text{si } k < n \end{cases} \qquad\qquad L|_k \;=\; \{w|_k \mid w \in L\}$$

$$u \odot_k v \;=\; (u \cdot v)|_k \qquad\qquad L_1 \odot_k L_2 \;=\; \{w_1 \odot_k w_2 \mid w_1 \in L_1 \text{ y } w_2 \in L_2\}$$

Los operadores $|_k$ y $\odot_k$ "miran" hasta un prefijo k.

Definición de $\texttt{first}_k$ y $\texttt{follow}_k$ (recordatorio)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

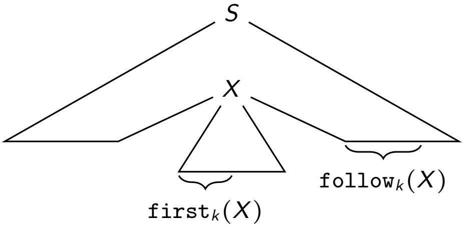
Definiciones

Se define la función $\texttt{first}_k : (V \cup \Sigma)^* \to 2^{\Sigma^{\leq k}}$ tal que, para $\gamma \in (V \cup \Sigma)^*$:

$$\texttt{first}_k(\gamma) \;=\; \{u|_k \mid \gamma \overset{\star}{\Rightarrow} u\}$$

Se define la función $\texttt{follow}_k : V \to 2^{\Sigma_{\#}^{\leq k}}$ como:

$$\texttt{follow}_k(X) \;=\; \{w \mid S \overset{\star}{\Rightarrow} \alpha X \beta \;\texttt{y}\; w \in \texttt{first}_k(\beta\#)\}$$



Propiedades de $\texttt{follow}_k$

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

$$\texttt{follow}_k(X) \;=\; \{w \mid S \overset{\star}{\Rightarrow} \alpha X \beta \text{ y } w \in \texttt{first}_k(\beta\#)\}$$

Si consideramos $X \neq S$:

$$\begin{aligned}
\texttt{follow}_k(X) &= \bigcup_{S \overset{\star}{\Rightarrow} \alpha X \beta} \texttt{first}_k(\beta\#) \\
&= \bigcup_{S \overset{\star}{\Rightarrow} \alpha Y \beta \Rightarrow \alpha\alpha' X \beta' \beta} \texttt{first}_k(\beta'\beta\#) \\
&= \bigcup_{Y \to \alpha' X \beta'} \; \bigcup_{S \overset{\star}{\Rightarrow} \alpha Y \beta} \texttt{first}_k(\beta'\beta\#) \\
&= \bigcup_{Y \to \alpha' X \beta'} \; \bigcup_{S \overset{\star}{\Rightarrow} \alpha Y \beta} \texttt{first}_k(\beta') \odot_k \texttt{first}_k(\beta\#) \\
&= \bigcup_{Y \to \alpha' X \beta'} \texttt{first}_k(\beta') \odot_k \bigcup_{S \overset{\star}{\Rightarrow} \alpha Y \beta} \texttt{first}_k(\beta\#) \\
&= \bigcup_{Y \to \alpha' X \beta'} \texttt{first}_k(\beta') \odot_k \texttt{follow}_k(Y)
\end{aligned}$$

Propiedades de $\texttt{follow}_k$

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

$$\texttt{follow}_k(X) \;=\; \{w \mid S \overset{\star}{\Rightarrow} \alpha X \beta \text{ y } w \in \texttt{first}_k(\beta\#)\}$$

Si consideramos $X \neq S$:

$$\begin{aligned}
\texttt{follow}_k(X) &= \bigcup_{S \overset{\star}{\Rightarrow} \alpha X \beta} \texttt{first}_k(\beta\#) \\
&\vdots \\
&= \bigcup_{Y \to \alpha' X \beta'} \texttt{first}_k(\beta') \odot_k \texttt{follow}_k(Y)
\end{aligned}$$

Si consideramos $X = S$:

$$\begin{aligned}
\texttt{follow}_k(S) &= \{\#\} \cup \bigcup_{S \overset{+}{\Rightarrow} \alpha S \beta} \texttt{first}_k(\beta\#) \\
&= \{\#\} \cup \bigcup_{S \overset{\star}{\Rightarrow} \alpha Y \beta \Rightarrow \alpha\alpha' S \beta' \beta} \texttt{first}_k(\beta'\beta\#) \\
&= \{\#\} \cup \bigcup_{Y \to \alpha' S \beta'} \texttt{first}_k(\beta') \odot_k \texttt{follow}_k(Y)
\end{aligned}$$

Propiedades de $\texttt{follow}_k$

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

Teorema

$$
\begin{aligned}
\text{Para } X \neq S\text{:} \quad \texttt{follow}_k(X) \quad &= \quad \bigcup_{Y \to \alpha X \beta} \texttt{first}_k(\beta) \odot_k \texttt{follow}_k(Y) \\
\texttt{follow}_k(S) \quad &= \quad \{\#\} \cup \bigcup_{Y \to \alpha S \beta} \texttt{first}_k(\beta) \odot_k \texttt{follow}_k(Y)
\end{aligned}
$$

Defina el siguiente **programa recursivo** para todo $X \in V$:

$$
\begin{aligned}
\text{Para } X \neq S\text{:} \quad \texttt{follow}_k^0(X) \quad &:= \quad \varnothing \\
\texttt{follow}_k^0(S) \quad &:= \quad \{\#\} \\
\text{Para } X \neq S\text{:} \quad \texttt{follow}_k^i(X) \quad &:= \quad \bigcup_{Y \to \alpha X \beta} \texttt{first}_k(\beta) \odot_k \texttt{follow}_k^{i-1}(Y) \\
\texttt{follow}_k^i(S) \quad &:= \quad \{\#\} \cup \bigcup_{Y \to \alpha S \beta} \texttt{first}_k(\beta) \odot_k \texttt{follow}_k^{i-1}(Y)
\end{aligned}
$$

¿cómo calcular $\texttt{follow}_k$?

Similar al caso de $\texttt{first}_k$:

- $\texttt{follow}_k^{i-1}(X) \subseteq \texttt{follow}_k^{i}(X)$ para todo $i > 1$.
- Como $\texttt{follow}_k(X) \subseteq \Sigma^{\leq k}$, entonces para algún $i \leq k \cdot |\Sigma|^k \cdot |V|$:

$$\texttt{follow}_k^{j}(X) \;=\; \texttt{follow}_k^{j+1}(X) \quad \text{para todo } j \geq i.$$

Teorema

Sea i^* el menor número tal que $\texttt{follow}_k^{i^*}(X) = \texttt{follow}_k^{i^*+1}(X)$ para todo $X \in V$. Entonces para todo $X \in V$:

$$\texttt{follow}_k^{i^*}(X) \;=\; \texttt{follow}_k(X)$$

Demostración: ejercicio.

...y podemos calcular $\texttt{follow}_k(X)$ con un algoritmo similar que $\texttt{first}_k(X)$.

¿cómo calculamos $\texttt{first}_k$ y $\texttt{follow}_k$ eficientemente?

- Algoritmos toman $\mathcal{O}(k \cdot |\Sigma|^k \cdot |V|)$ repeticiones en el peor caso.
- Si $k = 1$, el número de repeticiones será $\mathcal{O}(|\Sigma| \cdot |V|)$ y tiempo del algoritmo será polinomial en $|\mathcal{G}|$ en el peor caso.
- Para $k = 1$ incluso se puede hacer en tiempo $\mathcal{O}(|V| \cdot |P|)$ en total.

Volviendo a la clase de hoy: motivación

Para una gramática $\mathcal{G} = (V, \Sigma, P, S)$
podemos construir un PDA alternativo $\mathcal{D}$ que acepta $\mathcal{L}(\mathcal{G})$:

$$\mathcal{D} \;=\; \big(V \cup \Sigma \cup \{q_0, q_f\}, \Sigma, \Delta, q_0, \{q_f\}\big)$$

La relación de transición Δ se define como:

$$\begin{array}{rcllr} \Delta & = & \{\, (q_0, \epsilon, S \cdot q_f) \,\} & \cup & \\ & & \{\, (X, \epsilon, \gamma) \mid X \to \gamma \in P \,\} & \cup & (\textbf{Expandir}) \\ & & \{\, (a, a, \epsilon) \mid a \in \Sigma \,\} & & (\textbf{Reducir}) \end{array}$$

¿cómo elegir la siguiente producción para expandir?

¿cómo elegir la siguiente producción para expandir?

$$X \rightarrow \alpha \mid \beta$$

¿cómo elegir entre α o β?

Estrategia (intuición)

1. Mirar k símbolos del resto del input v (k-**lookahead**).
2. Usar $v|_k$ y decidir cual regla $X \rightarrow \gamma$ para expandir.

¿cómo **caracterizamos** las gramáticas que cumplen con esta propiedad?

Gramáticas LL(k)

Significado

Primera L: leer el input de izquierda a derecha (**L**eft-right).

Segunda L: producir una derivación por la izquierda (**L**eftmost).

Parámetro k: el número de letras en adelante que utiliza (**lookahead**).

LL(k) son las gramáticas que **caracterizan** la propiedad anterior.

Outline

Outline

Definición Gramáticas LL(k)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

Definición

Decimos que $\mathcal{G}$ es una gramática LL(k) si para todas derivaciones:

- $S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_1$
- $S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces se cumple que $\gamma_1 = \gamma_2$.

Notar que la elección de $Y \to \gamma$ depende de Y, $v|_k$ y u.

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

$$\left.\begin{array}{l} \blacksquare\ S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_1 \\ \blacksquare\ S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_2 \quad \text{y} \\ \blacksquare\ v_1|_k = v_2|_k \end{array}\right\} \quad \text{entonces} \quad \gamma_1 = \gamma_2.$$

Ejemplo 1: Gramática LL(1)

$$\mathcal{G}_1 : \quad S \rightarrow (S) \mid n$$

$$\begin{array}{lllll} S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} \overbrace{(\cdots(}^{u}\overbrace{S}^{Y}\overbrace{)\cdots)}^{\beta} & \underset{\text{lm}}{\Rightarrow} & (\cdots(\gamma_1)\cdots) & \underset{\text{lm}}{\overset{\star}{\Rightarrow}} & \overbrace{(\cdots(}^{u}\overbrace{v_1')\cdots)}^{v_1} \\ & \underset{\text{lm}}{\Rightarrow} & (\cdots(\gamma_2)\cdots) & \underset{\text{lm}}{\overset{\star}{\Rightarrow}} & \underbrace{(\cdots(}_{u}\underbrace{v_2')\cdots)}_{v_2} \end{array}$$

- Si $v_1|_1 = v_2|_1 = n$, entonces $\gamma_1 = \gamma_2 = n$.
- Si $v_1|_1 = v_2|_1 = ($, entonces $\gamma_1 = \gamma_2 = (S)$.

En ambos casos, tenemos que $\gamma_1 = \gamma_2$ y $\mathcal{G}_1$ es una **gramática** LL(1).

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1$
- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces $\gamma_1 = \gamma_2$.

Ejemplo 2: Gramática LL(1)

$$\begin{array}{llcl} \mathcal{G}_2: & S & \rightarrow & (X) \mid n \\ & X & \rightarrow & SX \mid \epsilon \end{array}$$

$$\begin{array}{ccccccc} S & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uX\beta & \underset{\text{lm}}{\Rightarrow} & u\gamma_1\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_1 \\ & & & \underset{\text{lm}}{\Rightarrow} & u\gamma_2\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_2 \end{array}$$

- Si $v_1|_1 = v_2|_1 =$ (o 'n', entonces $\gamma_1 = \gamma_2 = SX$.
- Si $v_1|_1 = v_2|_1 =$), entonces $\gamma_1 = \gamma_2 = \epsilon$.

Por lo tanto, tenemos que $\gamma_1 = \gamma_2$ y $\mathcal{G}_2$ es **también** una gramática LL(1).

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

$$\left.\begin{array}{l} \blacksquare\ S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_1 \\ \blacksquare\ S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \underset{\text{lm}}{\overset{\star}{\Rightarrow}} uv_2 \quad \text{y} \\ \blacksquare\ v_1|_k = v_2|_k \end{array}\right\} \quad \text{entonces} \quad \gamma_1 = \gamma_2.$$

Ejemplo 3: Gramática **NO** LL(1)

$$\begin{array}{rlcl} \mathcal{G}_3 : & S & \rightarrow & (X) \mid n+S \mid n \\ & X & \rightarrow & SX \mid \epsilon \end{array}$$

$$\begin{array}{lllll} S \underset{\text{lm}}{\overset{\star}{\Rightarrow}} \overbrace{(}^{u}\ \overbrace{S}^{Y}\ \overbrace{X)}^{\beta} & \underset{\text{lm}}{\Rightarrow} & (\ \overbrace{n+S}^{\gamma_1}\ X) & \underset{\text{lm}}{\overset{\star}{\Rightarrow}} & \overbrace{(}^{u}\ \overbrace{n+n)}^{v_1} \\ & \underset{\text{lm}}{\Rightarrow} & (\ \underbrace{n}_{\gamma_2}\ X) & \underset{\text{lm}}{\overset{\star}{\Rightarrow}} & \underbrace{(}_{u}\ \underbrace{n)}_{v_2} \end{array}$$

Como $v_1|_1 = v_2|_1 = n$ pero $\gamma_1 \neq \gamma_2$, entonces $\mathcal{G}_3$ **NO** es una gramática LL(1).

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1$
- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces $\gamma_1 = \gamma_2$.

Ejemplo 3: ¿LL(2)?

$$\begin{array}{lllllllll} \mathcal{G}_3: & S & \rightarrow & (X) & \mid & n+S & \mid & n \\ & X & \rightarrow & SX & \mid & \epsilon & & \end{array}$$

$$\begin{array}{lllllll} S & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uS\beta & \underset{\text{lm}}{\Rightarrow} & u\gamma_1\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_1 \\ & & & \underset{\text{lm}}{\Rightarrow} & u\gamma_2\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_2 \end{array}$$

- Si $v_1|_2 = v_2|_2 = n+$, entonces $\gamma_1 = \gamma_2 = n + S$.
- Si $v_1|_1 = v_2|_1 = na$ con $a \neq +$, entonces $\gamma_1 = \gamma_2 = n$.

Por lo tanto, tenemos que $\gamma_1 = \gamma_2$ y $\mathcal{G}_2$ es LL(2) y **no** LL(1).

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1$
- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces $\gamma_1 = \gamma_2$.

Ejemplo 4: Gramática **NO** LL(k)

$$\begin{array}{llllllllll} \mathcal{G}_4: & S & \rightarrow & (X) & | & (X)\hat{}e & | & n+S & | & n \\ & X & \rightarrow & SX & | & \epsilon & & & & \end{array}$$

$$\begin{array}{llllll} S & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & \overbrace{(}^{u}\ \overbrace{S}^{Y}\ \overbrace{X)}^{\beta} & \underset{\text{lm}}{\Rightarrow} & (\ \overbrace{(S)\hat{}e}^{\gamma_1}\ X) & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & \overbrace{(}^{u}\ \overbrace{(\overset{k}{\cdots}(n)\overset{k}{\cdots})\hat{}e)}^{v_1} \\ & & & \underset{\text{lm}}{\Rightarrow} & (\ \underbrace{(S)}_{\gamma_2}\ X) & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & \underbrace{(}_{u}\ \underbrace{(\overset{k}{\cdots}(n)\overset{k}{\cdots}))}_{v_2} \end{array}$$

Como $v_1|_k = v_2|_k = (\overset{k}{\cdots}($ pero $\gamma_1 \neq \gamma_2$, entonces

$\mathcal{G}_3$ **NO** es una gramática LL(k) para todo k.

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1$
- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces $\gamma_1 = \gamma_2$.

Ejemplo 4: Gramática **NO** LL(k)

$$\begin{array}{llllllllll} \mathcal{G}_4 : & S & \rightarrow & (X) & | & (X)\hat{\ }e & | & n+S & | & n \\ & X & \rightarrow & SX & | & \epsilon & & & & \end{array}$$

$$S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} \overbrace{(}^{u}\,\overbrace{S}^{Y}\,\overbrace{X)}^{\beta} \begin{array}{l} \underset{\text{lm}}{\Rightarrow} (\,\overbrace{(S)\hat{\ }e}^{\gamma_1}\; X) \overset{\star}{\underset{\text{lm}}{\Rightarrow}} \overbrace{(}^{u}\,\overbrace{(\overset{k}{\cdots}(n)\overset{k}{\cdots})\hat{\ }e)}^{v_1} \\ \underset{\text{lm}}{\Rightarrow} (\,\underbrace{(S)}_{\gamma_2}\; X) \overset{\star}{\underset{\text{lm}}{\Rightarrow}} \underbrace{(}_{u}\,\underbrace{(\overset{k}{\cdots}(n)\overset{k}{\cdots}))}_{v_2} \end{array}$$

¿es posible **transformar** $\mathcal{G}_4$ para que si sea LL(k)?

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

$$\left.\begin{array}{l}
\blacksquare\ S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1 \\
\blacksquare\ S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2 \quad \text{y} \\
\blacksquare\ v_1|_k = v_2|_k
\end{array}\right\} \quad \text{entonces} \quad \gamma_1 = \gamma_2.$$

Ejemplo 4: Gramática **NO** LL(k) **transformada** en LL(2)

$$\begin{array}{llll}
\mathcal{G}_4' : & S & \rightarrow & (XY \mid n + S \mid n \\
 & X & \rightarrow & SX \mid \epsilon \\
 & Y & \rightarrow &) \mid)\hat{}e
\end{array}$$

...

Demuestre que $\mathcal{G}_4'$ es una gramática LL(2).

Ejemplos y contra-ejemplos de gramáticas LL(k)

Definición LL(k)

- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_1\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_1$
- $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta \underset{\text{lm}}{\Rightarrow} u\gamma_2\beta \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uv_2$ y
- $v_1|_k = v_2|_k$

entonces $\gamma_1 = \gamma_2$.

Ejemplo 5: Lenguaje **NO** LL(k)

$$\mathcal{G}_5 : \quad \begin{array}{lcl} S & \to & X \mid Y \\ X & \to & aXb \mid 0 \\ Y & \to & aYbb \mid 1 \end{array} \qquad \mathcal{L}(\mathcal{G}_5) = \{a^n 0 b^n \mid n \geq 0\} \cup \{a^n 1 b^{2n} \mid n \geq 0\}$$

$$\begin{array}{lllll} S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} S & \underset{\text{lm}}{\Rightarrow} & X & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & a^k 0 b^k \\ & \underset{\text{lm}}{\Rightarrow} & Y & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & a^k 1 b^{2k} \end{array}$$

Para todo $k \geq 1$, se tiene que $\mathcal{G}_5$ **NO** es una gramática LL(k).

Es posible demostrar que, para toda gramática $\mathcal{G}$ con $\mathcal{L}(\mathcal{G}_5) = \mathcal{L}(\mathcal{G})$, $\mathcal{G}$ **NO** es una gramática LL(k) para todo $k \geq 1$.

Outline

Definiciones de prefijos (recordatorio)

Definiciones

$$w|_k = \begin{cases} a_1 \dots a_n & \text{si } n \leq k \\ a_1 \dots a_k & \text{si } k < n \end{cases} \qquad L|_k = \{w|_k \mid w \in L\}$$

$$u \odot_k v = (u \cdot v)|_k \qquad L_1 \odot_k L_2 = \{w_1 \odot_k w_2 \mid w_1 \in L_1 \text{ y } w_2 \in L_2\}$$

Los operadores $|_k$ y $\odot_k$ "miran" hasta un prefijo k.

Definición de $\texttt{first}_k$ y $\texttt{follow}_k$ (recordatorio)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

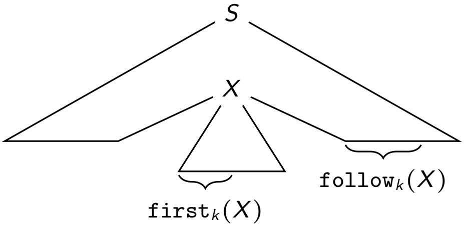
Definiciones

Se define la función $\texttt{first}_k : (V \cup \Sigma)^* \to 2^{\Sigma^{\leq k}}$ tal que, para $\gamma \in (V \cup \Sigma)^*$:

$$\texttt{first}_k(\gamma) \;=\; \{u|_k \mid \gamma \overset{\star}{\Rightarrow} u\}$$

Se define la función $\texttt{follow}_k : V \to 2^{\Sigma_{\#}^{\leq k}}$ como:

$$\texttt{follow}_k(X) \;=\; \{w \mid S \overset{\star}{\Rightarrow} \alpha X \beta \;\text{ y }\; w \in \texttt{first}_k(\beta\#)\}$$



Caracterización de gramáticas LL(k)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto **reducida** y $k \geq 1$.

Teorema

$\mathcal{G}$ es una gramática LL(k) si, y solo si, para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$ y para todo $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta$, se tiene que:

$$\texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta) \;=\; \varnothing$$

Demostración ($\Rightarrow$)

Por contra-positivo, supongamos que $v \in \texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta)$.
Como $\mathcal{G}$ es reducida (sin variables inútiles), entonces:

$$\begin{array}{ccccccc} S & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uY\beta & \underset{\text{lm}}{\Rightarrow} & u\gamma_1\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uvv_1 \\ & & & \underset{\text{lm}}{\Rightarrow} & u\gamma_2\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uvv_2 \end{array}$$

para algún $v_1, v_2 \in \Sigma^*$. Como $\gamma_1 \neq \gamma_2$, entonces $\mathcal{G}$ **NO** es LL(k).

Caracterización de gramáticas LL(k)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto **reducida** y $k \geq 1$.

Teorema

$\mathcal{G}$ es una gramática LL(k) si, y solo si, para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$ y para todo $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta$, se tiene que:

$$\texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta) \;=\; \varnothing$$

Demostración ($\Leftarrow$)

Por contra-positivo (de nuevo), supongamos que $\mathcal{G}$ no es LL(k).
Como $\mathcal{G}$ no es LL(k), entonces tenemos derivaciones de la forma:

$$\begin{array}{ccccccc} S & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uY\beta & \underset{\text{lm}}{\Rightarrow} & u\gamma_1\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_1 \\ & & & \underset{\text{lm}}{\Rightarrow} & u\gamma_2\beta & \overset{\star}{\underset{\text{lm}}{\Rightarrow}} & uv_2 \end{array}$$

$v_1|_k = v_2|k = v$, pero $\gamma_1 \neq \gamma_2$. Por lo tanto, $v \in \texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta)$. ■

Caracterización de gramáticas LL(k)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto **reducida** y $k \geq 1$.

Teorema

$\mathcal{G}$ es una gramática LL(k) si, y solo si, para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$ y para todo $S \overset{\star}{\underset{\text{lm}}{\Rightarrow}} uY\beta$, se tiene que:

$$\texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta) \;=\; \varnothing$$

¿cómo usamos esta **caracterización**
para demostrar que una gramática es LL(k)?

... buscaremos condiciones más simples
para verificar si una gramática es LL(k).

Gramáticas LL(k) fuerte

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una gramática libre de contexto y $k \geq 1$.

Definición

$\mathcal{G}$ es una gramática LL(k) **fuerte** si para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$ se tiene que:

$$\texttt{first}_k(\gamma_1) \odot_k \texttt{follow}_k(Y) \;\cap\; \texttt{first}_k(\gamma_2) \odot_k \texttt{follow}_k(Y) \;=\; \varnothing$$

¿Si $\mathcal{G}$ es LL(k) **fuerte**, entonces es $\mathcal{G}$ un gramática LL(k)?

$$\begin{aligned}
\texttt{first}_k(\gamma_1\beta) \cap \texttt{first}_k(\gamma_2\beta) \quad &= \\
\texttt{first}_k(\gamma_1) \odot_k \texttt{first}_k(\beta) \cap \texttt{first}_k(\gamma_2) \odot_k \texttt{first}_k(\beta) \quad &\subseteq \\
\texttt{first}_k(\gamma_1) \odot_k \texttt{follow}_k(Y) \cap \texttt{first}_k(\gamma_2) \odot_k \texttt{follow}_k(Y) \quad &= \quad \varnothing
\end{aligned}$$

¿si $\mathcal{G}$ es LL(k), entonces es LL(k) fuerte?

¿si $\mathcal{G}$ es LL(k), entonces es LL(k) fuerte?

Contra-ejemplo

$$\begin{array}{llll} \mathcal{G}: & S & \to & aXaa \mid bXba \\ & X & \to & b \mid \epsilon \end{array}$$

Recordatorio: $\mathcal{G}$ es LL(k) si para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$ y para todo $S \underset{\mathrm{lm}}{\overset{\star}{\Rightarrow}} uY\beta$, se tiene que:

$$\mathtt{first}_k(\gamma_1\beta) \cap \mathtt{first}_k(\gamma_2\beta) = \varnothing$$

- Si $S \underset{\mathrm{lm}}{\overset{\star}{\Rightarrow}} aXaa$, entonces $\mathtt{first}_2(baa) \cap \mathtt{first}_2(aa) = \varnothing$.
- Si $S \underset{\mathrm{lm}}{\overset{\star}{\Rightarrow}} bXba$, entonces $\mathtt{first}_2(bba) \cap \mathtt{first}_2(ba) = \varnothing$.

Por lo tanto, $\mathcal{G}$ es LL(2).

¿si $\mathcal{G}$ es LL(k), entonces es LL(k) fuerte?

Contra-ejemplo

$$\begin{array}{rrcl} \mathcal{G}: & S & \rightarrow & aXaa \mid bXba \\ & X & \rightarrow & b \mid \epsilon \end{array}$$

Recordatorio: $\mathcal{G}$ es una gramática LL(k) **fuerte** si para todas dos reglas distintas $Y \rightarrow \gamma_1, Y \rightarrow \gamma_2 \in P$ se tiene que:

$$\texttt{first}_k(\gamma_1) \odot_k \texttt{follow}_k(Y) \ \cap \ \texttt{first}_k(\gamma_2) \odot_k \texttt{follow}_k(Y) \ = \ \varnothing$$

Si vemos $X \rightarrow b$ y $X \rightarrow \epsilon$:

$$\begin{aligned} & \texttt{first}_2(b) \odot_2 \texttt{follow}_2(X) \cap \texttt{first}_2(\epsilon) \odot_2 \texttt{follow}_2(X) \\ & \quad = \{b\} \odot_2 \{aa, ba\} \ \cap \ \{\epsilon\} \odot_2 \{aa, ba\} \\ & \quad = \{ba, bb\} \ \cap \ \{aa, ba\} \\ & \quad = \{ba\} \end{aligned}$$

...y $\mathcal{G}$ no es LL(2) fuerte.

Caso LL(1)

Supongamos que $\mathcal{G}$ es LL(1) y $Y \to \gamma_1, Y \to \gamma_2 \in P$ son reglas distintas.

1. Si $\epsilon \notin \mathtt{first}_1(\gamma_1)$ y $\epsilon \notin \mathtt{first}_1(\gamma_2)$, entonces (por caract. de LL(1)):

$$\begin{aligned}
\varnothing &= \mathtt{first}_1(\gamma_1\beta) \cap \mathtt{first}_1(\gamma_2\beta) \\
&= \mathtt{first}_1(\gamma_1) \cap \mathtt{first}_1(\gamma_2) \\
&= \mathtt{first}_1(\gamma_1) \odot_1 \mathtt{follow}_1(Y) \cap \mathtt{first}_1(\gamma_2) \odot_1 \mathtt{follow}_1(Y)
\end{aligned}$$

Caso LL(1)

Supongamos que $\mathcal{G}$ es LL(1) y $Y \to \gamma_1, Y \to \gamma_2 \in P$ son reglas distintas.

1. Si $\epsilon \notin \mathtt{first}_1(\gamma_1)$ y $\epsilon \notin \mathtt{first}_1(\gamma_2)$, entonces (por caract. de LL(1)):

$$\varnothing \quad = \quad \mathtt{first}_1(\gamma_1) \odot_1 \mathtt{follow}_1(Y) \cap \mathtt{first}_1(\gamma_2) \odot_1 \mathtt{follow}_1(Y)$$

2. Si $\epsilon \in \mathtt{first}_1(\gamma_1)$ y $\epsilon \notin \mathtt{first}_1(\gamma_2)$, entonces (por caract. de LL(1)):

$$\begin{aligned} \varnothing &= \mathtt{first}_1(\gamma_1\beta) \cap \mathtt{first}_1(\gamma_2\beta) \\ &= \mathtt{first}_1(\gamma_1\beta) \cap \mathtt{first}_1(\gamma_2) \\ &= \mathtt{first}_1(\gamma_1\beta) \cap \mathtt{first}_1(\gamma_2\beta') \end{aligned}$$

para todo $\beta' \in (V \cup \Sigma)^*$. Por lo tanto:

$$\mathtt{first}_1(\gamma_1) \odot_1 \mathtt{follow}_1(Y) \cap \mathtt{first}_1(\gamma_2) \odot_1 \mathtt{follow}_1(Y)$$

$$= \bigcup_{S \underset{\mathrm{lm}}{\overset{\star}{\Rightarrow}} uY\beta} \mathtt{first}_1(\gamma_1\beta) \cap \bigcup_{S \underset{\mathrm{lm}}{\overset{\star}{\Rightarrow}} uY\beta'} \mathtt{first}_1(\gamma_2\beta') = \varnothing$$

Caso LL(1)

Teorema

Una gramática $\mathcal{G}$ es LL(1) si, y solo si, $\mathcal{G}$ es LL(1) **fuerte**,
esto es, para todas dos reglas distintas $Y \to \gamma_1, Y \to \gamma_2 \in P$:

$$\texttt{first}_1(\gamma_1) \odot_1 \texttt{follow}_1(Y) \ \cap \ \texttt{first}_1(\gamma_2) \odot_1 \texttt{follow}_1(Y) \ = \ \varnothing$$

Podemos verificar esta condición en **tiempo polinomial** en $\mathcal{G}$.