

Analiza sintactica descendenta. Gramatici si analiza LL(K)

.

FIRST_k

- FIRST_k: $(N \cup \Sigma)^* \rightarrow \text{Part}(\Sigma^*)$
- FIRST_k(α) = $\{u \in \Sigma^* \mid (\alpha \overset{*}{>} ux, |u| = k)$
sau $(\alpha \overset{*}{>} u, |u| < k) \}$

FOLLOW_k

$\text{FOLLOW}_k: (\mathbf{N} \cup \Sigma)^* \rightarrow \text{Part}(\Sigma^*)$

$\text{FOLLOW}_k(\alpha) = \{u \in \Sigma^* \mid S \xRightarrow{*} \gamma \alpha \beta, u \in \text{FIRST}_k(\beta)\}$

Operatia \oplus_k

Fie $L1, L2$ – 2 limbaje peste Σ

$L1 \oplus_k L2 = \{w \mid \exists x \in L1, y \in L2 \text{ astfel incat:}$

fie: $|w| = k$ si $\exists z: xy = wz$

fie: $|w| < k$ si $xy = w$

$\}$

Exercitii: $L1 = \{a, ab, abb, abbb, \dots\}$

$L2 = \{c, cc, ccc, cccc, \dots\}$

$L1 \oplus_3 L2 = \dots$

$L1 \oplus_1 L2 = \dots$

Determinarea lui FIRST_1

- pentru fiecare $A \in N$:

$$F_0(A) = \{a \in \Sigma \mid A \rightarrow a\alpha \in P\} \cup \{\varepsilon \mid A \rightarrow \varepsilon\}$$

sf. pentru

$i = 0$

- repeta

$i = i + 1$

pentru fiecare $A \in N$:

$$F_i(A) = \{a \in \Sigma \cup \{\varepsilon\} \mid A \rightarrow X_1 \dots X_k, a \in F_{i-1}(X_1) \oplus_1 \dots \oplus_1 F_{i-1}(X_k)\}$$

sf. pentru

pana cand $F_i = F_{i-1}, \forall A \in N$

- $\text{FIRST}_1 = F_i$

Obs:

$\forall a \in \Sigma : \text{FIRST}_1(a) = \{a\}$

$\forall A \in N : \text{FIRST}_1(A)$ se determina pe baza reg. productie: $A \rightarrow \dots$

FOLLOW₁

- $\text{FOLL}(S) = \{\$ \}$
 $\text{FOLL}(X) = \Phi \quad (X \langle \rangle S)$
 - repeta
 - pentru fiecare $B: A \rightarrow \alpha B \beta$ executa
$$\text{FOLL}(B) = \text{FOLL}(B) \cup (\text{FIRST}_1(\beta) - \{\varepsilon\})$$
Daca $\varepsilon \in \text{FIRST}_1(\beta)$ atunci
$$\text{FOLL}(B) = \text{FOLL}(B) \cup \text{FOLL}(A)$$
sf. daca
 - sf. pentru
- pana cand FOLL nu se mai modifica
- $$\text{FOLLOW}_1 = \text{FOLL}$$

Proprietati:

$$\text{FIRST}_1(\alpha\beta) = \text{FIRST}_1(\alpha) \oplus_1 \text{FIRST}_1(\beta)$$

$$\text{FIRST}_1(X_1 \dots X_n) = \dots \quad ?$$

- $A \rightarrow \alpha$

$$\text{FIRST}_1(A) \supset \text{FIRST}_1(\alpha)$$

- $A \rightarrow \alpha X \beta$

$$(\text{FIRST}_1(\beta) - \{\varepsilon\}) \subset \text{FOLLOW}_1(X)$$

- $A \rightarrow \alpha X$

$$\text{FOLLOW}_1(A) \subset \text{FOLLOW}_1(X)$$

Exercitii: $FIRST_1$ si $FOLLOW_1$

Problema:

- determinati $FIRST_1$ si $FOLLOW_1$ pentru toate terminalele si neterminalele gramaticii cu regulile de productie:

$$S \rightarrow BA \quad (1)$$

$$A \rightarrow +BA \quad (2)$$

$$A \rightarrow \varepsilon \quad (3)$$

$$B \rightarrow DC \quad (4)$$

$$C \rightarrow *DC \quad (5)$$

$$C \rightarrow \varepsilon \quad (6)$$

$$D \rightarrow (S) \quad (7)$$

$$D \rightarrow a \quad (8)$$

Gramatici LL(k)

- analizoare LL(k)
 - analiza sintactica descendenta
 - secv. de intrare este citita de la stanga spre dreapta
 - se folosesc derivari de stanga

Gramatici LL(k)

- $G = (N, \Sigma, P, S)$
? $w = a_1 a_2 \dots a_n \in L(G)$
 $S \Rightarrow_{st} \dots \Rightarrow_{st} a_1 a_2 \dots a_i A \alpha$
(derivari de stanga)
- LL(k) : alegerea r.p. pentru a-l retranscrie pe A este unic determinata de: $a_{i+1} a_{i+2} \dots a_{i+k}$
- terminologie:
 - parte inchisa: $a_1 a_2 \dots a_i$
 - predictia (de lungime k): $a_{i+1} a_{i+2} \dots a_{i+k}$
 - simbol de retranscris: A

Gramatici LL(k)

Definitie:

$G = (N, \Sigma, P, S)$ este LL(K):

- Daca:
 - $S \overset{*}{\Rightarrow}_{st} wA\alpha \Rightarrow_{st} w\beta\alpha \overset{*}{\Rightarrow} wx$
 - $S \overset{*}{\Rightarrow}_{st} wA\alpha \Rightarrow_{st} w\gamma\alpha \overset{*}{\Rightarrow} wy$
 - $FIRST_k(x) = FIRST_k(y)$
- atunci:
 $\beta = \gamma$

k=1

•

Gramatici LL(1)

- ? $w = a_1 a_2 \dots a_i a_{i+1} \dots a_n \in L(G)$
 $S \Rightarrow_{st} \dots \Rightarrow_{st} a_1 a_2 \dots a_i A \alpha$
- LL(1): rescrierea lui A este unic determin. de **a_{i+1}**

De exemplu:

Aleg:

$$A \rightarrow \gamma$$

$$A \rightarrow \varepsilon$$

$$A \rightarrow \delta \quad (=^* > \varepsilon)$$

atunci cand:

$$a_{i+1} \in \text{FIRST}_1(\gamma)$$

$$a_{i+1} \in \text{FOLLOW}_1(A)$$

$$a_{i+1} \in \text{FOLLOW}_1(A)$$

Gramatici LL(1)

Teorema:

G – este de tip LL(1)

ddaca:

$$\forall A \in N$$

$$A \rightarrow \alpha_1 \mid \alpha_2 \mid \alpha_3 \mid \dots \mid \alpha_m$$

- $\text{FIRST}_1(\alpha_i) \cap \text{FIRST}_1(\alpha_j) = \Phi$, $i <> j$
- daca $\exists i$ a.i. $\alpha_i \overset{*}{>} \varepsilon$ atunci: $\varepsilon \in \text{FIRST}_1(\alpha_i)$

$$\text{FIRST}_1(\alpha_j) \cap \text{FOLLOW}_1(A) = \Phi$$
 , $i <> j$

Analiza sintactica LL(1)

- analizorul sintactic LL(1)
 - ~ modelare ca un **automat**
 - se construiește tabelul de analiza LL(1)
 - alegerea lui $A \rightarrow \alpha_i$ este indicată în tabel
-

Convenție:

\$: la orice cuvânt din limbaj se adaugă \$:
marcător de sfârșit de cuvânt

Observație:

O gramatică este de tip LL(1) dacă tabelul de analiză nu conține conflicte (nu există mai mult de o valoare într-o celulă din tabel)

Tabelul de analiza LL(1)

indica actiunea posibila la un moment dat

- coloane: $\Sigma \cup \{\$ \}$
- linii: $N \cup \Sigma \cup \{\$ \}$
- celula \Rightarrow (membrul drept al reg.prod , nr. reg.prod.)

Tabelul de analiza LL(1)

$M(X,a) =$

- (α,i) daca $X \rightarrow \alpha \in P$, $a \in \text{FIRST}_1(\alpha)$
 $X \rightarrow \alpha$ – a i-a regula de productie
- (α,i) daca $\varepsilon \in \text{FIRST}_1(\alpha)$, $a \in \text{FOLLOW}_1(X)$
 $X \rightarrow \alpha$ – a i-a regula de productie
- pop $X = a$
- acc $X = \$$, $a = \$$
- err in toate celelalte cazuri

Analizorul LL(1)

- Automat: (α, β, Π)
 - banda de intrare: α (stiva de intrare)
 - stiva β (stiva de lucru)
 - banda de iesire $\Pi \Rightarrow$ sirul regulilor de productie
- config. initiala: $(w$, S, $\varepsilon)$$$
- config. finala: $($, $$, \Pi)$$
- tranzitii
 - push $(\mathbf{ax}$, $\mathbf{A}\beta, \Pi) \vdash (\mathbf{ax}$, $\alpha\beta, \Pi\mathbf{i})$ dc.: $M(A, a) = (\alpha, i)$$$
 - pop $(\mathbf{ax}$, $\mathbf{a}\beta, \Pi) \vdash (\mathbf{x}$, $\beta, \Pi)$$$
 - acc $($, $$, \Pi) \vdash \text{acc}$$
 - err in celelalte cazuri

Transformari echivalente

- \Rightarrow gram. LL(1) echiv. ?
- factorizare la stanga
- ...