Analiza descendenta. Gramatici LL(k) - eliminare recursivitate stanga. factorizare.

Table of Contents

Teorema

4.2.2, 4.2.3 Teorema. O gramatica LL(k) nu poate avea simbol nonterminal recursiv stanga.

Daca $X \Rightarrow X\omega, \omega \neq \varepsilon$ - X nonterminal recursiv stanga

- $E \to E + T|T$
- $T \to T * F|F$
- $F \to (E)|id$

are doua productii cu recursivitate stanga

Teorema

Teorema. Pentru orice gramatica CFG $G = (T, N, P, Z)$ cu simboluri nonterminale recursive stanga exista o gramatica echivalenta $G' = (T, N', P', Z)$ fara nonterminale recursive stanga.

Idee

$$X \to X\alpha|\beta \text{ devine } \begin{cases} X \to \beta X' \\ X' \to \alpha X'|\varepsilon \end{cases}$$

- $E \to E + T|T$
- $T \to T * F|F$
- $F \to (E)|id$

- $E \to TE'$
- $E' \to +TE'|\varepsilon$
- $T \to FT'$
- $T' \to *FT'|\varepsilon$
- $F \to (E)|id$

Dar...

NU intra la examen
vezi Testare online

- $S \to Aa|b$
- $A \to Ac|Sd|\varepsilon$

$$S \Rightarrow Aa \Rightarrow Sda$$

ne trebuie un algoritm care sa elimine toate nonterminalele cu recursivitate stanga

- Consideram ca $N = \{X_1, X_2, ...X_n\}$ - simbolurile nonterminale sunt numerotate consecutiv.
- Daca putem alege indicii a.i. indicii sa respecte $i < j$ pentru toate productiile $X_i \rightarrow X_j\omega$ atunci G nu are recursivitate stanga.
- Daca o astfel de numerotare nu este posibila pentru G, atunci se genereaza G'.

Exemple:

- $S \rightarrow Aa|b$
- $A \rightarrow Ac|Sd|\varepsilon$
- Daca S e 1, A e 2, prima productie respecta $i < j$ dar nu si a doua
- $E \rightarrow E + T$ nu respecta $i < j$

Algoritm de eliminare recursivitate stanga

NU intra la examen Testare online

1. Fie $N' = N, P' = P$. Se executa pasii 2,3 pentru $i = 1, ...n$
2. Pentru $j = 1, ...i - 1$
   $X_i \to X_j\omega \in P'$ se inlocuiesc cu $\{X_i \to \chi_j\omega | X_j \to \chi_j \in P'\}$.
   In consecinta, $X_i \Rightarrow^+ X_j\gamma$ implica $i \leq j$.
3. Se inlocuiesc $X_i \to X_i\omega \in P'$ cu $\{Y_i \to \omega Y_i\} \cup \{Y_i \to \varepsilon\}$
   adaugand un nou simbol Y_i la N'.
   $+$ se inlocuiesc $X_i \to \chi, \chi \neq X_i\gamma$ cu $X_i \to \chi Y_i$.
   Simbolurile noi se numeroteaza cu n+1, n+2,..

- $E \to E + T|T$
- $T \to T * F|F$
- $F \to (E)|id$

presupunem ordinea $E(1) < T(2) < F(3)$

i	pasul2	pasul 3	variabila noua
1	nu se executa	$E \to E + T\|T$ devin $E' \to +TE'\|\varepsilon$ si $E \to TE'$;	$E'(4)$
2	$j = 1$ $T \to E\omega$ nu exista	$T \to T * F\|F$ devin $T' \to *FT'\|\varepsilon$ $T \to FT'$	$T'(5)$
3	$j = 1, 2$ $F \to E\omega$ sau $F \to T\omega$ nu exista	$F \to F\omega$ nu exista	
4,5	nu se modifica nimic		

Rezultat:

- $E \rightarrow E + T|T$
- $T \rightarrow T * F|F$
- $F \rightarrow (E)|id$

- $E \rightarrow TE'$
- $E' \rightarrow +TE'|\varepsilon$
- $T \rightarrow FT'$
- $T' \rightarrow *FT'|\varepsilon$
- $F \rightarrow (E)|id$

- $S \to Aa|b$
- $A \to Ac|Sd|\varepsilon$

...pasul 2 al algoritmului: Pentru $j = 1, ...i - 1$ $X_i \to X_j\omega \in P'$ se inlocuiesc cu $\{X_i \to \chi_j\omega | X_j \to \chi_j \in P'\}$.

Daca $S(1) < A(2)$

- $i = 1$ nimic
- $i = 2$ la pasul 2 $A \to Sd$ se inlocuieste cu $\{A \to Aad|bd\}$

- $S \to Aa|b$
- $A \to Ac|Sd|\varepsilon$

...pasul 3 al algoritmului: Se inlocuiesc $X_i \to X_i\omega \in P'$ cu $\{Y_i \to \omega Y_i\} \cup \{Y_i \to \varepsilon\}$ adaugand un nou simbol Y_i la N'.
$+$ se inlocuiesc $X_i \to \chi, \chi \neq X_i\gamma$ cu $X_i \to \chi Y_i$.

Daca $S(1) < A(2)$

- $i = 1$ nimic
- $i = 2$ la pasul 2 $A \to Sd$ se inlocuieste cu $\{A \to Aad|bd\}$
- $i = 2$ la pasul 3 $A \to Ac|Aad|bd|\varepsilon$ se inlocuieste cu
  $A' \to cA', A' \to adA', A' \to \varepsilon$ si
  $A \to bdA', A \to A'$

Teorema. Daca sirul ω din $X_i \to X_i\omega$ nu incepe cu $X_j, j \le i$ atunci $X_i \to X_i\omega$ se poate inlocui cu $\{Y_i \to \omega, Y_i \to \omega Y_i\}$ si $X_i \to \chi$ cu $\{X_i \to \chi, X_i \to \chi Y_i\}$ la pasul 3.

pasul 3 anterior ...se inlocuiesc $X_i \to X_i\omega \in P'$ cu $\{Y_i \to \omega Y_i\} \cup \{Y_i \to \varepsilon\}$ adaugand un nou simbol Y_i la N'.
$+$ se inlocuiesc $X_i \to \chi, \chi \neq X_i\gamma$ cu $X_i \to \chi Y_i$.

Se evita introducerea productiilor ε.

Table of Contents

Intra la examen

- $E \to E + T | T$
- $T \to T * F | F$
- $F \to (E) | id$

Cu productii ε

- $E \to TE'$
- $E' \to +TE' | \varepsilon$
- $T \to FT'$
- $T' \to *FT' | \varepsilon$
- $F \to (E) | id$

Fara productii ε

- $E \to TE' | T$
- $E' \to +T | +TE'$
- $T \to FT' | F$
- $T' \to *F | *FT'$
- $F \to (E) | id$

Observatii

- Recursivitatea stanga precum $E \rightarrow T|E + T$ - utilizata pentru a reflecta asociativitatea stanga a operatorilor.
- Aceeasi proprietate avem si in $E \rightarrow TE', E' \rightarrow +TE', E' \rightarrow \varepsilon$
- Insa asociativitate dreapta $E \rightarrow T, E \rightarrow T + E$.

Productii ε

Productiile ε se pot elimina intotdeauna dintr-o gramatica LL(k), dar aceasta poate mari valoare lui k.
4.2.3

Teorema

TEOREMA. Pentru orice gramatica LL(k) cu productii ε exista o gramatica LL(k+1) fara productii ε care genereaza limbajul $L(G) - \varepsilon$.

Prin introducerea productiilor ε se poate reduce k.

Teorema

TEOREMA. Pentru orice gramatica LL(k+1), $k > 0$ fara productii ε exista o gramatica LL(k) echivalenta cu productii ε.

Table of Contents

Factorizare stanga

$$Fie\ P = \{ \quad \begin{array}{l} Z \rightarrow X \\ X \rightarrow Yc|Yd \\ Y \rightarrow a|bY\} \end{array}$$

Productiile $X \rightarrow Yc$ si $X \rightarrow Yd$ nu pot fi distinse chiar prin examinarea oricarui numar fix de simboluri din sirul de intrare deoarece din Y se poate deriva un sir de lungime si mai mare.

Solutie: evitarea problemei prin amanarea deciziei. Ambele incep cu Y, nu trebuie facauta distinctie intre ele decat dupa ce Y a fost recunoscut.

Factorizare stanga

$$Fie\ P = \{ \quad \begin{array}{l} Z \rightarrow X \\ X \rightarrow Yc|Yd \\ Y \rightarrow a|bY\} \end{array}$$

devine

$$Fie\ P = \{ \quad \begin{array}{l} Z \rightarrow X \\ X \rightarrow YX' \\ X' \rightarrow c|d \\ Y \rightarrow a|bY\} \end{array}$$

Se poate examina un singur caracter inainte pt a face diferente intre cele doua variante c sau d .

Factorizare stanga

$$
\begin{aligned}
Fie\ P = \{\quad & Z \rightarrow X \\
& X \rightarrow if\ \ E\ \ then\ \ S | if\ \ E\ \ then\ \ S\ \ else\ \ S | a \\
& E \rightarrow b\}
\end{aligned}
$$

devine

$$
\begin{aligned}
Fie\ P = \{\quad & Z \rightarrow X \\
& X \rightarrow if\ \ E\ \ then\ \ S\ \ S' | a \\
& S' \rightarrow else\ \ S | \varepsilon \\
& E \rightarrow b\}
\end{aligned}
$$