Gramatici Regulate

Seminar 4.

Exemplu (intuitiv!)

$$\left[\begin{array}{l} \text{Expresie} \xrightarrow[\text{poate fi}]{} \text{Expresie} + \text{Termen} \;/\; \text{Termen} \\ \text{Termen} \xrightarrow[\text{poate fi}]{} \text{Termen} * \text{Factor} \;/\; \text{Factor} \\ \text{Factor} \xrightarrow[\text{poate fi}]{} (\text{Expresie}) \;/\; a \end{array}\right.$$

- $a * (a+a)$ respectă aceste reguli de formare:

$\text{Expresie} \xrightarrow[\text{poate fi}]{} \text{Termen} \xrightarrow[\text{poate fi}]{} \text{Termen} * \text{Factor} \xrightarrow[\text{poate fi}]{} a * \text{Factor}$

$\xrightarrow[\text{poate fi}]{} a * (\text{Expresie}) \xrightarrow[\text{poate fi}]{} a * (\text{Expresie} + \text{Termen}) \xrightarrow[\text{poate fi}]{}$

$\xrightarrow[\text{poate fi}]{} a * (\text{Termen} + \text{Termen}) \xrightarrow[\text{poate fi}]{} a * (\text{Factor} + \text{Factor})$

$\xrightarrow[\text{poate fi}]{} a * (a+a)$

unde am înlocuit <u>membrul stâng al regulii</u> cu <u>membrul drept al regulii</u>

def $G = (N, T, S, P)$ unde N, T, P mulțimi <u>finite</u>

N = neterminale (notate cu litere mari ex: $A, B \ldots$)

T = terminale (notate cu litere mici ex: $a, b, \ldots$)

"alfabetul cu care se formează cuvintele"

$S \in N$ simbol de start

P = reguli de producție de forma $\boxed{\alpha \to \beta}$ unde $\alpha, \beta \in (N \cup T)^*$

def derivare $\underbrace{\alpha_1 \boxed{\beta} \alpha_2} \Rightarrow \underbrace{\alpha_1 \boxed{\gamma} \alpha_2}$ dacă $\boxed{\beta \to \gamma \in P}$

[în cuvântul $\alpha_1 \beta \alpha_2$ se înlocuiește membrul stâng al producției (β) cu membrul drept al producției (γ)]

extinderea reflexivă $\alpha \stackrel{*}{\Rightarrow} \alpha$

tranzitivă $\left.\begin{array}{l} \alpha \stackrel{*}{\Rightarrow} \beta \\ \beta \stackrel{*}{\Rightarrow} \gamma \end{array}\right| \to \alpha \stackrel{*}{\Rightarrow} \gamma$

$L(G) = \{ w \in T^* \mid S \stackrel{*}{\Rightarrow} w \}$

"

limbajul generat de gramatica G

pentru exemplul anterior - al expresiilor aritmetice:

$G = (N, T, E, P)$ unde $T = \{a, *, +, (,)\}$

$N = \{E, T, F\}$

iar P: $E \to E + T / T$

$T \to T * F / F$

$F \to (E) / a$

am $E \Rightarrow T \Rightarrow T * F \Rightarrow a * F \Rightarrow \ldots \Rightarrow a*(a+a)$

deci $E \stackrel{*}{\Rightarrow} a*(a+a)$ adică $a*(a+a) \in L(G)$

(cuvântul poate fi generat folosind aceste reguli) def $G=(N,T,S,P)$ e gramatică regulată ($G \in \mathcal{L}_3$)

$\iff$ P are reguli de forma $\boxed{\begin{array}{l} A \to aB \\ A \to b \\ A \to \lambda \end{array}}$ unde

$A, B \in N \quad a, b \in T$

Exemple: găsiți $G \in \mathcal{L}_3$ ai

ex 1) $L(G) = \{a^n \mid n \geq 1\}$ P: $S \to a$, $\boxed{S} \to a\boxed{S}$ $\quad T = \{a\} \quad N = \{S\}$

! • pentru a obține cu un nr finit de neterminale, un număr infinit de cuvinte $\Rightarrow$ există $\boxed{\text{ciclare}}$

ex 2) $L(G) = \{a^n \mid n \geq 0\} = \{\lambda, a, a^2 \dots\}$

$S \to \lambda$

$S \to aS$

Obs - $S \to a$ nu e necesară pot obține $\boxed{S} \Rightarrow aS \Rightarrow a \cdot \lambda$

- a^2 se obține $\underline{S} \Rightarrow a\underline{S} \Rightarrow aaS \Rightarrow aa \cdot \lambda = a^2$

- a^k se obț! $\underbrace{S \Rightarrow aS \Rightarrow \dots \Rightarrow a^k S}_{k \text{ aplicări ale regulii } S \to aS} \Rightarrow a^k \cdot \lambda$

- deci $L(G) \supseteq \{a^k \mid k \geq 0\}$

- $\forall$ cuvânt se poate obține aplicând regula $\underline{S \to aS}$ de un nr finit (k) ori și apoi regula $\underline{S \to \lambda}$ $\Rightarrow$

$\Rightarrow L(G) \subseteq \{a^k \mid k \geq 0\}$

ex 3 $L(G) = \{a^n b \mid n \geq 0\} = \{b, ab, a^2b, \ldots\}$

$S \to aS$ pt a obține $a^n S$

$S \to b$

(ex: pt a^2b $S \Rightarrow aS \Rightarrow a^2S \Rightarrow a^2b$)

ex 4 $L(G) = \{a^n b \mid n \geq 1\} = \{ab, a^2b, a^3b, \ldots\}$

obs: $S \to b$ nu e corectă pt că s-ar obține $b \notin L(G)$

pt a obține ab:

$$\boxed{\begin{array}{l} S \to aB \\ B \to b \\ S \to aS \end{array}}$$

pentru a cicla ($S \Rightarrow a^n S$)

ex 5 $L(G) = \{\alpha \in \{a, b\}^* \mid N_a(\alpha) = 1\}$

$S \to bS$ generează b^k $k \geq 0$

$S \to aA$ a

$A \to bA/\lambda$ nu poate reveni la S (ar mai putea genera alt a)

ex 6 $L(G) = \{a^{2n} \mid n \geq 0\}$

$S \to \lambda$

$S \to aA$

$A \to aS$

deci $\boxed{S \overset{*}{\Rightarrow} a^2 S} \Rightarrow \ldots \Rightarrow a^{2k} S \Rightarrow a^{2k}$

ex 7 $L(G) = \{a^{2n+1} \mid n \geq 0\}$

$S \to aS$

$A \to aS$

$S \to a$

$S \overset{*}{\Rightarrow} a^2 S \Rightarrow \ldots \Rightarrow a^{2k} S \Rightarrow a^{2k+1}$

Prop. Pentru orice gramatică regulată G, există o gramatică G' echivalentă (L(G) = L(G')) cu reguli de forma: $\begin{cases} A \to \alpha B & \text{unde } \alpha \in T^* \\ A \to \alpha \end{cases}$

adică reguli de forma

$A \to a_1 \dots a_n B$

$\boxed{A \to a_1 B}$

$A \to B$

$A \to a_1 \dots a_n$

$\boxed{\begin{array}{l} A \to a_1 \\ A \to \lambda \end{array}}$

mai ușor de construit !

Concluzie
- în demonstrații – forma redusă (mai puține cazuri)
- la construcția gramaticilor – forma largă (mai ușor de construit)

idee intuitivă :

1) $A \to a_1 \dots a_n B$ se poate descompune în

$A \to a_1 X_1$

$X_1 \to a_2 X_2$

$\vdots$

$X_{n-1} \to a_n B$

unde $X_1 \dots X_{n-1}$ neterminale noi

$\Downarrow$

singura posibilitate de derivare este:

$A \Rightarrow a_1 X_1 \Rightarrow a_1 a_2 X_2 \Rightarrow \dots \Rightarrow a_1 \dots a_n B$

echivalent cu regula

$A \to a_1 \dots a_n B$

2) analog $A \to a_1 \dots a_n$ se poate descompune în

$A \to a_1 X_1$

$X_1 \to a_2 X_2$

$\vdots$

$X_{n-1} \to a_n$

unde $X_1 \dots X_{n-1}$ neterminale noi

$\Downarrow$

singura derivare :

$A \Rightarrow a_1 X_1 \Rightarrow a_1 a_2 X_2 \Rightarrow \dots \Rightarrow a_1 \dots a_n$

echivalent cu regula

$A \to a_1 \dots a_n$

3) eliminarea regulilor de forma $A \to B$:

[ex: dacă am: $A \to B$ — elimin $A \to B$ și adaug $A \to \alpha$
$B \to \alpha$]

Algoritm

P0. $\forall A \in N \leadsto M_A = \{X \in N \mid A \overset{*}{\Rightarrow} X\}$

" mulțimea neterminalelor în care derivă A

P1. - se șterg producțiile $A \to B$ $\forall A, B \in N$

P2. - pt $\forall X \in M_A \leadsto$ adaug $A \to \alpha$
$X \to \alpha$

Ex 8 $L(G) = \{a^{2n} \mid n \geq 0\}$

$$\begin{cases} S \to \lambda \\ S \to a^2 S \end{cases} \text{ echivalent cu } \begin{cases} S \to \lambda \\ S \to aA \\ A \to aS \end{cases}$$

A neterminal nou

Ex 9 $L(G) = \{a^{2n} b^{3m} \mid m, n \geq 0\}$

$S \to a^2 S / B$

$B \to b^3 B / \lambda$

Observăm la ex 9 că din S se produce întâi a^{2n} și apoi B care produce b^{3m}.

T. Familia limbajelor generate de gramaticile regulate este închisă la operațiile de $\cup$, $\cdot$, $*$

$\Downarrow$

Putem construi o gramatică regulată ce generează unde G_1, G_2 gramatici regulate

$L(G_1) \cup L(G_2)$
$L(G_1) \cdot L(G_2)$
$L(G_1)^*$

fie $G_1 = (N_1, T_1, S_1, P_1)$ gramatici regulate

$G_2 = (N_2, T_2, S_2, P_2)$

! considerăm $N_1 \cap N_2 = \emptyset$ pentru a nu se combina regulile gramaticilor

1) $L(G) = L(G_1) \cup L(G_2)$

$$\boxed{G = (N_1 \cup N_2 \cup \{S\}, T_1 \cup T_2, S, P_1 \cup P_2 \cup \{S \to S_1 / S_2\})}$$

ex: $L(G) = \{a^n \mid n \geq 0\} \cup \{b^m \mid m \geq 0\} = \{\lambda, a, a^2, \ldots, b, b^2 \ldots\}$

greșit: $S \to aS / bS / \lambda$ pt că pot obține

$S \Rightarrow aS \Rightarrow abS \Rightarrow ab \notin L(G)$

G: $S_1 \to aS_1 / \lambda$

$S_2 \to bS_2 / \lambda$

$S \to S_1 / S_2$

2) $L(G) = L(G_1) \cdot L(G_2)$

- observăm la derivare ca în cuvânt există cel mult un neterminal
- la fiecare pas de derivare - adaug $\alpha \in T^*$ și (eventual) schimb neterminalul din cuvânt
- la pasul final - <u>neterminalul dispare</u> și adaugă $\alpha \in T^*$

$\Downarrow$

$G = (N_1 \cup N_2, T_1 \cup T_2, P)$

$P = P_1 \setminus \{A \to \alpha \mid \alpha \in T_1^*\} \cup \{A \to \alpha S_2 \mid$ pt $A \to \alpha \in P_1$, $\alpha \in T_1^*\}$

pentru a nu termina cuvântul în G_1

pt a porni derivarea în G_2

$\cup P_2$ → pt a produce cuvânt în G_2

ex: $L(G) = \{a^n \mid n \geq 0\} \cdot \{b^m \mid m \geq 0\}$

$G_1 \begin{cases} S_1 \to aS_1 \\ S_1 \to a \\ S_1 \to \lambda \end{cases}$ se înlocuiesc cu $S_1 \to a\underline{S_2}$, $S_1 \to \underline{S_2}$

$G_2 \; S_2 \to bS_2 / \lambda$

$G: \begin{cases} S_1 \to aS_1 / aS_2 / S_2 \\ S_2 \to bS_2 / \lambda \end{cases}$

3) $L(G) = (L(G_1))^*$

$G = (N_1 \cup \{S\}, T_1, S, P)$ unde

$$\boxed{P = \{S \to S_1 / \lambda\} \cup P_1 \cup \{A \to \alpha S \mid A \to \alpha \in P_1, \alpha \in T^*\}}$$

ex: $L(G) = (\{a^n b \mid n \geq 0\})^* = \{\lambda, ab, a^2b, aba^2b, abab, \dots\}$

greşit $S_1 \to aS_1 / bS_1 / \lambda$ se poate obține

$S_1 \Rightarrow aS_1 \Rightarrow a \notin L(G)$ → obligatoriu $\underline{S \to S_1/\lambda}$

G_1: $S_1 \to aS_1$
$S_1 \to b$
$S_1 \to \lambda$ | se adaugă | $S_1 \to bS$
$S_1 \to S$

↓

pt a continua cu un nou cuvânt

$G: \; S \to S_1 / \lambda$

$S_1 \to aS_1 / bS / S / \underbrace{b / \lambda}$

nu mai sunt necesare pt ca $S \to \lambda$

T. Familia limbajelor generate de gramaticile regulate este Reg (familia limbajelor recunoscute de automatele finite)

Concluzie 1. pt $L = L(A)$, $A \in A.F \Rightarrow \exists\, G \in \mathcal{G}_3$ (regulate), $\underline{L(A) = L(G)}$

2. pt $L = L(G)$, $G \in \mathcal{G}_3 \Rightarrow \exists\, A \in A.F$ (automate finite), $L(A) = L(G)$

1) $L = L(A)$ $\quad A = (\Sigma, Q, q_0, \delta, F)$

$G = (N, T, S, P)$, $N = Q$, $T = \Sigma$, $S = \boxed{q_0}$

$$P = \{\boxed{p} \to a\boxed{q} \mid \delta(p,a) = q\} \cup \{\boxed{p} \to \lambda \mid \boxed{p} \in F\}$$

2) $L = L(G)$ $\quad G = (N, T, S, P)$

P cu reguli de tip $A \to aB$, $\boxed{A \to a}$, $A \to \lambda$

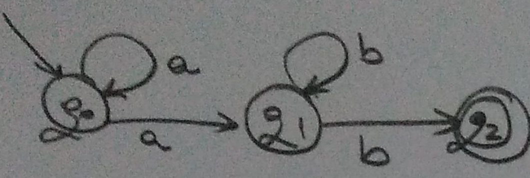
$A \to a$ se transformă în $A \to aX$, $X \to \lambda$

! adaug un singur X neterminal nou

$A = (\Sigma, Q, q_0, \delta, F)$, $Q = N \cup \{X\}$, $\Sigma = \Sigma$, $q_0 = \boxed{S}$, $F = \{\boxed{B} \in N \mid \underline{B \to \lambda}\}$

$\delta(\boxed{A}, a) = \boxed{B}$ pt $A \to aB$

ex:

pentru  construiți $G \in \mathcal{G}_3$

$G = (N, T, S, P)$

$N = \{ \boxed{q_0}, \boxed{q_1}, \boxed{q_2} \}$

$T = \{a, b\}$

$S = \boxed{q_0}$

$P: \boxed{q_0} \to a\boxed{q_0} \,/\, a\boxed{q_1}$

$\boxed{q_1} \to b\boxed{q_1} \,/\, b\boxed{q_2}$

$q_2 \in F \overset{\text{adaug}}{\rightsquigarrow} \boxed{q_2} \to \lambda$

---

ex: pentru

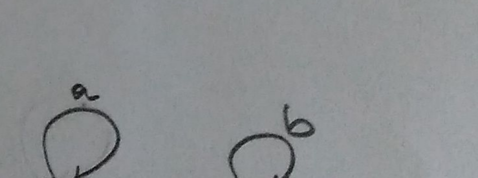
$S \to aS \,/\, aA$

$A \to bA$ construiți A A.F.

$A \to b$ — se transformă în $A \to bX$, $X \to \lambda$

$A \to \lambda$

$A = (Q, \Sigma, q_0, \delta, F)$



$Q = \{S, A, X\}$

$\Sigma = \{a, b\}$

$q_0 = S$

$S \to aS \rightsquigarrow \delta(S, a) \ni S$

$S \to aA \rightsquigarrow \delta(S, a) \ni A$

$A \to bA \rightsquigarrow \delta(A, b) \ni A$

$A \to bX \rightsquigarrow \delta(A, b) \ni X$

$\left.\begin{array}{l} A \to \lambda \\ X \to \lambda \end{array}\right| \Rightarrow \underline{F = \{A, X\}}$

ex: $L(G) = \{ \alpha \in \{a, b\}^* \mid N_a(\alpha) = \text{par}\ \ N_b(\alpha) = \text{impar} \}$

$\underline{S = S_{pp}}$

$\begin{cases} S_{pp} \to a S_{ip} \\ S_{pp} \to b S_{pi} \end{cases}$

$\begin{cases} S_{ip} \to a S_{pp} \\ S_{ip} \to b S_{ii} \end{cases}$

$\begin{cases} S_{ii} \to a S_{pi} \\ S_{ii} \to b S_{ip} \end{cases}$

$S_{pi} \to a S_{ii}$

$S_{pi} \to b S_{pp}$

$\boxed{S_{pi} \to \lambda}$