

# **A számításelmélet alapjai I.**

## **3. előadás**

**előadó: Tichler Krisztián**  
**[ktichler@inf.elte.hu](mailto:ktichler@inf.elte.hu)**

# Lineáris grammatikák

## Definíció

Egy  $G = \langle N, \Sigma, P, S \rangle$  környezetfüggetlen grammatikát **lineárisnak** nevezünk, ha minden szabálya vagy

- a)  $A \rightarrow u, A \in N, u \in \Sigma^*$  vagy
- b)  $A \rightarrow u_1 B u_2$  alakú, ahol  $A, B \in N$  és  $u_1, u_2 \in \Sigma^*$ .

A  $G$  lineáris grammatikát **bal-lineárisnak** nevezzük, ha minden b) alakú szabályában  $u_1 = \varepsilon$ .

A  $G$  lineáris grammatikát **jobb-lineárisnak** mondjuk, ha minden b) alakú szabályában  $u_2 = \varepsilon$ .

**Megjegyzés:** A jobb-lineáris grammatika egy másik elnevezés a reguláris (3-típusú) grammatikákra.

## Példa:

- 1.  $S \rightarrow aSb \mid \varepsilon$  lineáris, de se nem bal-, se nem jobb-lineáris
- 2.  $S \rightarrow Saa \mid b$  bal-lineáris
- 3.  $S \rightarrow aS \mid bbS \mid c$  jobb-lineáris

# Lineáris grammatikák és nyelvek

## Definíció

Egy  $L$  nyelvet lineárisnak, bal-lineárisnak, illetve jobb-lineárisnak mondunk, ha van olyan  $G$  lineáris, bal-lineáris, illetve jobb-lineáris grammatika, amelyre  $L = L(G)$  teljesül.

**Megjegyzés:** A jobb-lineáris nyelvek tehát azonosak a reguláris (3-típusú) nyelvekkel.

## Tétel

Minden bal-lineáris grammatika reguláris (3-típusú) nyelvet generál.

# Bal- és jobblineáris nyelvek

**Bizonyítás:** Legyen  $G = \langle N, \Sigma, P, S \rangle$  bal-lineáris grammatika és legyen  $N = \{S, A_1, \dots, A_n\}$ .

Feltehetjük, hogy  $S$  nem fordul elő egyetlen szabály jobboldalán sem. Ha mégis előfordul, akkor minden előfordulását nevezzünk át  $S'$ -re ( $S' \notin N \cup \Sigma$ ) és adjuk hozzá  $P$ -hez az  $S \rightarrow S'$  szabályt.

Megkonstruálunk egy  $G' = \langle N, \Sigma, P', S \rangle$  jobb-lineáris grammatikát, amelyre  $L(G) = L(G')$  teljesül.

1.  $S \rightarrow u : \in P'$  akkor és csak akkor, ha  $S \rightarrow u \in P$ ,
2.  $S \rightarrow uA_k : \in P'$  akkor és csak akkor, ha  $A_k \rightarrow u \in P$ ,
3.  $A_j \rightarrow uA_k : \in P'$  akkor és csak akkor, ha  $A_k \rightarrow A_ju \in P$ ,
4.  $A_j \rightarrow u : \in P'$  akkor és csak akkor, ha  $S \rightarrow A_ju \in P$ .

( $u \in \Sigma^*$  minden a 4 pontban)

Megmutatjuk, hogy  $L(G) = L(G')$ .

# Bal- és jobblineáris nyelvek

$L(G) \subseteq L(G')$ :

Legyen  $w \in L(G)$ .

$S \rightarrow u : \in P'$	$\leftrightarrow$	$S \rightarrow u \in P,$
$S \rightarrow uA_k : \in P'$	$\leftrightarrow$	$A_k \rightarrow u \in P,$
$A_j \rightarrow uA_k : \in P'$	$\leftrightarrow$	$A_k \rightarrow A_j u \in P,$
$A_j \rightarrow u : \in P'$	$\leftrightarrow$	$S \rightarrow A_j u \in P.$

**1. eset** Ha  $S \rightarrow w \in P$ , akkor  $S \rightarrow w \in P'$ , így  $w \in L(G')$ .

**2. eset** Egyébként  $w$ -hez van  $G$ -ben egy

$$S \Rightarrow A_{i_1} w_1 \Rightarrow A_{i_2} w_2 w_1 \Rightarrow \cdots \Rightarrow A_{i_{m-1}} w_{m-1} \cdots w_1 \Rightarrow w_m w_{m-1} \cdots w_1 = w$$

levezetés ( $m \geq 2$ ). Azonban ekkor  $G'$ -ben létezik egy

$$S \Rightarrow w_m A_{i_{m-1}} \Rightarrow w_m w_{m-1} A_{i_{m-2}} \Rightarrow \cdots \Rightarrow w_m \cdots w_2 A_{i_1} \Rightarrow w_m \cdots w_1 = w$$

levezetés, azaz  $w \in L(G')$ .

$L(G') \subseteq L(G)$ :

Következik a szimmetriából.

# Bal- és jobblineáris nyelvek

**Példa:**

$$L = \{a^n b^k \mid n, k \in \mathbb{N}\}.$$

$S \rightarrow u : \in P'$	$\leftrightarrow$	$S \rightarrow u \in P,$
$S \rightarrow uA_k : \in P'$	$\leftrightarrow$	$A_k \rightarrow u \in P,$
$A_j \rightarrow uA_k : \in P'$	$\leftrightarrow$	$A_k \rightarrow A_j u \in P,$
$A_j \rightarrow u : \in P'$	$\leftrightarrow$	$S \rightarrow A_j u \in P.$

$P$

$S \rightarrow \varepsilon$

$S \rightarrow A_1 b$

$S \rightarrow A_2 a$

$A_1 \rightarrow A_1 b$

$A_1 \rightarrow A_2$

$A_2 \rightarrow A_2 a$

$A_2 \rightarrow \varepsilon$

$P'$

$S \rightarrow \varepsilon$

$A_1 \rightarrow b$

$A_2 \rightarrow a$

$A_1 \rightarrow bA_1$

$A_2 \rightarrow A_1$

$A_2 \rightarrow aA_2$

$S \rightarrow A_2$

$$S \Rightarrow A_1 b \Rightarrow A_2 b \Rightarrow A_2 ab \Rightarrow ab.$$

$$S \Rightarrow A_2 \Rightarrow aA_2 \Rightarrow aA_1 \Rightarrow ab.$$

## $\mathcal{L}_3$ tükrözésre való zártsága

**Állítás:** Minden  $G = \langle N, \Sigma, P, S \rangle$  jobb-lineáris grammatikához meg tudunk konstruálni egy  $G'$  bal-lineáris grammatikát, amely  $L(G)$  tükörképét generálja.

**Bizonyítás:**  $G'$ -ben vegyük a  $G$ -beli  $A \rightarrow u$  és  $A \rightarrow uB$  alakú szabályok helyett az  $A \rightarrow u^R$  illetve  $A \rightarrow Bu^R$  szabályokat ( $A, B \in N, u \in \Sigma^*$ ).

### Következmény:

1.  $\mathcal{L}_3$  zárt a tükrözés (megfordítás) műveletére nézve.
2. Minden reguláris nyelv generálható bal-lineáris grammatikával.

### Bizonyítás:

1. Ha  $L \in \mathcal{L}_3$ , akkor az állítás miatt  $L^R$  generálható egy bal-lineáris  $G$  grammatikával. Az előző tétel alapján  $L^R \in \mathcal{L}_3$ .
2. Ha  $L \in \mathcal{L}_3$ , akkor az épp most bizonyított **1.** miatt  $L^R \in \mathcal{L}_3$  is igaz, így  $L^R$  jobb-lineáris grammatikával generálható. Az állítás miatt  $(L^R)^R = L$  bal-lineáris grammatikával generálható.

# 3-típusú grammatikák normálformája

## Definíció

Egy  $G = \langle N, \Sigma, P, S \rangle$  grammatika **3-as normálformájú**, ha minden  $P$ -beli szabály alakja

- ▶  $X \rightarrow aY$ ,  $X, Y \in N$ ,  $a \in \Sigma$  vagy
- ▶  $X \rightarrow \varepsilon$ , ahol  $X \in N$ .

## Tétel

Minden reguláris (azaz 3-típusú) nyelv generálható 3-as normálformájú grammatikával.

**Bizonyítás:** Legyen  $L$  tetszőleges nyelv. Ekkor létezik  $G = \langle N, \Sigma, P, S \rangle$  3-típusú grammatika, melyre  $L(G) = L$ . Ismeretes, hogy  $G$  szabályai vagy

- a)  $A \rightarrow uB$  vagy
- b)  $A \rightarrow u$  alakúak ( $A, B \in N$ ,  $u \in \Sigma^*$ ).

A  $G$  grammatikát a megfelelő alakra transzformáljuk, figyelve arra hogy az eredmény  $G$ -vel ekvivalens legyen.



# 3-típusú grammatikák normálformája

## 1. lépés: Hosszredukció

### a) típusú szabályok:

Ha  $A \rightarrow uB \in P$  ( $A, B \in N, u \in \Sigma^*$ ),  $|u| > 1$  és  $u = a_1 \cdots a_n, n \geq 2$ , akkor helyettesítsük az  $A \rightarrow a_1 \cdots a_n B$  szabályt az  $\{A \rightarrow a_1 Z_1, Z_1 \rightarrow a_2 Z_2, \dots, Z_{n-1} \rightarrow a_n B\}$  szabályhalmazzal, ahol  $Z_1, \dots, Z_{n-1}$  új, a szabályhoz bevezetett egyedi nemterminálisok.

### b) típusú szabályok:

Hasonlóan, Ha  $A \rightarrow u \in P$  ( $A \in N, u \in \Sigma^*$ ),  $|u| > 0$  és  $u = a_1 \cdots a_m, m \geq 1$ , akkor helyettesítsük az  $A \rightarrow a_1 \cdots a_m$  szabályt az  $\{A \rightarrow a_1 Y_1, Y_1 \rightarrow a_2 Y_2, \dots, Y_{m-1} \rightarrow a_m E, E \rightarrow \varepsilon\}$  szabályhalmazzal, ahol  $Y_1, \dots, Y_{m-1}$  új, a szabályhoz bevezetett egyedi nemterminálisok.  $E$  szintén új nemterminális, de elég a b) típusú szabályok hosszredukciójához egy közös új  $\varepsilon$  szabály.

## 3-típusú grammatikák normálformája

Az eddigi szavakat az átalakítás után is generálni lehet (a hosszú szabály szimulálásához alkalmazzuk a neki megfeleltetett szabályokat a megadott sorrendben).

Másrészt, mivel az új nemterminálisok szabályonként egyediek, így nem lehetséges az eddigieken felül további szavak generálása.

### 2. lépés: Láncmentesítés

Egy már hosszredukált  $G = \langle N, \Sigma, P, S \rangle$  grammatika szabályhalmazának elemei  $X \rightarrow aY$ ,  $X \rightarrow Y$ ,  $X \rightarrow \varepsilon$  alakúak, ahol  $X, Y \in N$  és  $a \in \Sigma$ .

Tehát már csak az  $X \rightarrow Y$  alakú, ú.n. **láncszabályokat** kell eliminálni.

Jelölje  $P_0$  a  $P$ -beli láncszabályok halmazát.

### 3-típusú grammatikák normálformája

Első lépésben meghatározzuk minden  $A \in N$  esetén a  $H(A) := \{B \in N \mid A \Rightarrow^* B\}$  halmazokat.

$H(A)$  hatékony előállításához definiáljuk rekurzívan a  $H_i(A)$  ( $i \geq 0$ ) halmazokat az alábbiak szerint:

$$H_0(A) := \{A\}$$

$$H_{i+1}(A) := H_i(A) \cup \{B \in N \mid \exists C \in H_i(A) \text{ amelyre } C \rightarrow B \in P\}.$$

$$H_0(A) \subseteq H_1(A) \subseteq \dots \subseteq H_k(A) \subseteq \dots \subseteq N$$

$$k := \min\{0 \leq i \leq n - 1 \mid H_i(A) = H_{i+1}(A)\}.$$

$H(A) = H_k(A)$ , hiszen  $H_i(A)$  az  $A$ -ból legfeljebb  $i$  darab láncszabály alkalmazással elérhető nemterminálisok halmaza.

### 3-típusú grammatikák normálformája

$$G' = \langle N, \Sigma, P', S \rangle,$$

$$P' := \{A \rightarrow w \mid \exists B \in H(A), \text{ amelyre } B \rightarrow w \in P\} \setminus P_0.$$

Tekintsünk egy levezetésben láncszabály alkalmazások egy egymást követő sorozatát  $A \in N$ -től  $B \in N$ -ig. Ha a levezetés következő lépése a  $B \rightarrow w$  ( $w \in (N_1 \cup \Sigma)^*$ ) szabály alkalmazása, akkor helyettesítsük a  $G$ -beli  $A \Rightarrow_G^* B \Rightarrow_G w$  levezetés részletet az  $A \Rightarrow_{G'} w$  lépéssel  $G'$ -ben.

Tehát  $L(G) \subseteq L(G')$ .

Fordítva, az új szabályokat helyettesíthetjük láncszabályok és egy eredeti szabály ezen sorrendben történő alkalmazásával, amiből  $L(G') \subseteq L(G)$  adódik.

Tehát a kapott grammatika az eredetivel ekvivalens és 3-típusú.

# 3-típusú grammatikák normálformája

**Példa:**

$$S \rightarrow abS$$

$$S \rightarrow B$$

$$B \rightarrow bB$$

$$B \rightarrow V$$

$$V \rightarrow aa$$

$$V \rightarrow b$$

**1. lépés:** Hosszredukció

$$S \rightarrow aZ_1 \quad Z_1 \rightarrow bS$$

$$S \rightarrow B$$

$$B \rightarrow bB$$

$$B \rightarrow V$$

$$V \rightarrow aZ_2 \quad Z_2 \rightarrow aE$$

$$V \rightarrow bE \quad E \rightarrow \varepsilon$$

**2. lépés:** Láncmentesítés

$$S \rightarrow aZ_1 \quad Z_1 \rightarrow bS$$

$$S \rightarrow B$$

$$B \rightarrow bB$$

$$B \rightarrow V$$

$$V \rightarrow aZ_2 \quad Z_2 \rightarrow aE$$

$$V \rightarrow bE \quad E \rightarrow \varepsilon$$

$$S \rightarrow aZ_1 \quad Z_1 \rightarrow bS$$

$$S \rightarrow bB$$

$$S \rightarrow aZ_2 \quad S \rightarrow bE$$

$$B \rightarrow bB$$

$$B \rightarrow aZ_2 \quad B \rightarrow bE$$

$$V \rightarrow aZ_2 \quad Z_2 \rightarrow aE$$

$$V \rightarrow bE \quad E \rightarrow \varepsilon$$

$$H_0(X) = H_1(X) = H(X) = \{X\} \quad (X \in \{V, Z_1, Z_2, E\})$$

$$H_0(B) = \{B\}, H_1(B) = H_2(B) = H(B) = \{B, V\}.$$

$$H_0(S) = \{S\}, H_1(S) = \{S, B\}, H_2(S) = H_3(S) = H(S) = \{S, B, V\}.$$

# Reguláris kifejezések

## Motiváció:

Vegyük észre, hogy minden véges nyelv reguláris. Tudjuk továbbá, hogy az  $\mathcal{L}_3$  nyelvosztály (a reguláris nyelvek osztálya) zárt az unió, a konkatenáció és az iteratív lezárt műveletekre nézve.

Következésképpen, kiindulva véges számú véges nyelvből és az előzőekben felsorolt, ún. reguláris műveleteket véges sokszor alkalmazva reguláris nyelvet kapunk.

Kérdés az, hogy vajon ezzel az eljárással minden reguláris nyelvet elő tudunk-e állítani, azaz, ez a módszer elégséges-e az  $\mathcal{L}_3$  nyelvosztály leírására?

# Reguláris kifejezések

## Definíció

Legyenek  $V$  és  $V' = \{\emptyset, \varepsilon, \cdot, +, *, (, )\}$  diszjunkt ábécék. A  $V$  ábécé feletti **reguláris kifejezéseket** rekurzív módon a következőképpen definiáljuk:

1.  $\emptyset$  reguláris kifejezés  $V$  felett,
2.  $\varepsilon$  reguláris kifejezés  $V$  felett,
3.  $a$  reguláris kifejezés  $V$  felett, ha  $a \in V$ ,
4. Ha  $R$  reguláris kifejezés  $V$  felett, akkor  $R^*$  is reguláris kifejezés  $V$  felett,
5. Ha  $Q$  és  $R$  reguláris kifejezések  $V$  felett, akkor  $(Q \cdot R)$  is,
6. Ha  $Q$  és  $R$  reguláris kifejezések  $V$  felett, akkor  $(Q + R)$  is.

# Reguláris kifejezések

Minden egyes  $R$  reguláris kifejezés egy-egy  $L(R)$ -rel jelölt reguláris nyelvet határoz meg, amelyeket az alábbi rekurzióval definiálunk.

## Definíció

1.  $L(\emptyset) := \emptyset$ ,
2.  $L(\varepsilon) := \{\varepsilon\}$ ,
3.  $L(a) := \{a\} \quad (a \in V)$ ,
4.  $L(R^*) := L(R)^*$ , ha  $R$  reguláris kifejezés,
5.  $L(Q \cdot R) := L(Q)L(R)$ , ha  $Q, R$  reguláris kifejezések,
6.  $L(Q + R) := L(Q) \cup L(R)$ , ha  $Q, R$  reguláris kifejezések.

Azaz a  $*$ ,  $\cdot$  és  $+$  szimbólumok jelölik rendre a lezárt, a konkatenáció és az unió műveleteket jelölik.

**Megjegyzés:**  $(Q + R)$  helyett használatosak még a  $(Q | R)$  illetve a  $(Q \cup R)$  jelölések is.  $(Q \cdot R)$  helyett gyakran röviden  $(QR)$ -t írunk.



# Reguláris kifejezések

**Megjegyzés:**  $R$ -et gyakran azonosítjuk  $L(R)$ -el. Ilyenkor, kicsit pontatlanul, olyat írunk például, hogy  $ab^* = \{ab^n \mid n \in \mathbb{N}\}$ .

**Állítás:** Legyenek  $P$ ,  $Q$  és  $R$  reguláris kifejezések. Ekkor fennállnak a következő azonosságok:

1.  $P + (Q + R) = (P + Q) + R$
2.  $P \cdot (Q \cdot R) = (P \cdot Q) \cdot R$
3.  $P + Q = Q + P$
4.  $P \cdot (Q + R) = P \cdot Q + P \cdot R$
5.  $(P + Q) \cdot R = P \cdot R + Q \cdot R$
6.  $P^* = \varepsilon + P \cdot P^*$
7.  $\varepsilon \cdot P = P \cdot \varepsilon = P$
8.  $P^* = (\varepsilon + P)^*$
9.  $P \cdot \emptyset = \emptyset \cdot P = \emptyset$

Tehát precízebben az állítások így hangzanak:  
 $L(\text{bal oldal}) = L(\text{jobb oldal})$ .

# Reguláris kifejezések

A műveletek precedenciasorrendje  $*$ ,  $\cdot$ ,  $+$ , ennek és az asszociatív szabályok figyelembevételével bizonyos zárójelpárok elhagyhatók.

## Példa:

Az  $(a + b)a^*$  és  $aa^* + ba^*$  reguláris kifejezések által leírt nyelv ugyanaz:  $\{aa^n \mid n \in \mathbb{N}\} \cup \{ba^n \mid n \in \mathbb{N}\}$ .

Viszont  $(a + b)a^* \neq a + ba^*$ , hiszen az utóbbi által leírt nyelv  $\{a, b, ba, ba^2, ba^3, \dots\}$ .

**Megjegyzés:** Néha használjuk a  $Q^+$  jelölést  $(Q \cdot Q^*)$  rövidítésére. Használunk néha **konkrét** pozitív egész kitevőket is, például  $(a^3b + ab^3)b^*$ .

# Reguláris kifejezések – Arden tétele

Az alábbi tételt később, a reguláris kifejezések véges automatákkal való kapcsolatának vizsgálatakor fogjuk használni:

## Arden tétele

Adottak az  $R$  és  $Q$  reguláris kifejezések. A  $P = R + P \cdot Q$  egyenletnek  $P$ -re vonatkozó megoldása  $P = R \cdot Q^*$ . Amennyiben  $\varepsilon \notin L(Q)$ , akkor az egyenletnek ez az egyetlen megoldása.

**Példa:** A  $P = ab^* + Pbc^*$  egyenlet egyértelmű megoldása  $P = ab^*(bc^*)^*$ .

### Bizonyítás:

(1)  $P = R \cdot Q^*$  megoldás:

$$\begin{aligned} R + P \cdot Q &= R + (R \cdot Q^*) \cdot Q = R + R \cdot (Q \cdot Q^*) = \\ &= R \cdot (\varepsilon + Q \cdot Q^*) = R \cdot Q^* = P. \end{aligned}$$

# Reguláris kifejezések – Arden tétele

(2) A megoldás egyértelműsége:

$P = R + P \cdot Q = R + (R + P \cdot Q) \cdot Q = R + RQ + PQ^2 = \dots = R + RQ + \dots RQ^n + PQ^{n+1}$ , ahol  $n$  tetszőleges nagy lehet, azaz

$$P = R(\varepsilon + Q + \dots Q^n) + PQ^{n+1}. \quad (*)$$

Legyen  $w \in P$  és  $n = |w|$  az előző (\*) képletben. Mivel  $\varepsilon \notin Q$  ezért  $w \notin PQ^{n+1}$  (minden szava legalább  $n + 1$  hosszú), tehát  $w \in R(\varepsilon + Q + \dots + Q^n) \subseteq RQ^*$ .

Fordítva, ha  $w \in RQ^*$ , akkor  $\exists n \in \mathbb{N}$ , hogy  $w \in RQ^n$ , azaz benne van (\*) jobboldalában, azaz  $P$ -ben is.

Tehát csak  $P = R \cdot Q^*$  lehetséges.

# Reguláris kifejezések leíró ereje

## Tétel

Minden reguláris kifejezés reguláris (azaz 3-típusú) nyelvet ír le, és megfordítva, minden reguláris nyelvhez megadható egy, ezen nyelvet leíró reguláris kifejezés.

## Bizonyítás:

1. Az, hogy reguláris kifejezések csakis reguláris nyelvet írhatnak le következik az  $\emptyset$ ,  $\{\varepsilon\}$ ,  $\{a\}$  ( $a \in V$ ) nyelvek reguláris voltából és  $\mathcal{L}_3$  reguláris műveletekre való zártságából.
2. Megmutatjuk, hogy minden  $L$  reguláris nyelvhez, amelyet a  $G = \langle N, \Sigma, P, S \rangle$  normálformában adott reguláris grammatika generál, meg tudunk konstruálni egy reguláris kifejezést, amely az  $L$  nyelvet jelöli.

# Reguláris kifejezések leíró ereje

Legyen  $N = \{A_1, \dots, A_n\}$ ,  $n \geq 1$ ,  $S = A_1$ . ( $G$  minden szabálya vagy  $A_i \rightarrow aA_j$  vagy  $A_i \rightarrow \varepsilon$  alakú, ahol  $a \in \Sigma$ ,  $1 \leq i, j \leq n$ .)

Azt mondjuk, hogy az  $A_i \Rightarrow^* uA_j$  ( $u \in \Sigma^*$ ) levezetés **érinti** az  $A_m$  nemterminálist, ha  $A_m$  előfordul valamely közbülső mondatformában az  $A_i \Rightarrow^* uA_j$  levezetésben.

Az  $A_i \Rightarrow^* uA_j$  levezetést  **$\ell$ -megszorítottnak** nevezzük, ha  $1 \leq m \leq \ell$  teljesül minden  $A_m$  nemterminálisra, amelyet a levezetés érint.

Definiáljuk a következő halmazokat  $0 \leq k \leq n$ ,  $1 \leq i, j \leq n$ -re:

$$E_{i,j}^k = \{u \in \Sigma^* \mid \text{létezik } A_i \Rightarrow^* uA_j \text{ } k\text{-megszorított levezetés}\}.$$

Az  $i \neq j$  esetben az  $E_{i,j}^0$  halmaz vagy üres vagy  $\Sigma$ -beli betűkből áll:  $a \in E_{i,j}^0$ , akkor és csak akkor, ha  $A_i \rightarrow aA_j \in P$ .

## Reguláris kifejezések leíró ereje

Ha  $i = j$ , akkor  $E_{i,i}^0$  tartalmazza  $\varepsilon$ -t és nulla vagy több elemét  $\Sigma$ -nak (ha  $A_i \rightarrow aA_i \in P$ ).

$k$ -ra vonatkozó indukcióval bizonyítunk. Látható, hogy  $E_{i,j}^0$  leírható reguláris kifejezéssel.

Tegyük fel, hogy rögzített  $k$ -ra ( $0 < k \leq n$ ), az  $E_{i,j}^{k-1}$  nyelvek mindegyike leírható reguláris kifejezéssel.

Meggondolható, hogy teljesül

$$E_{i,j}^k = E_{i,j}^{k-1} + E_{i,k}^{k-1} \cdot (E_{k,k}^{k-1})^* \cdot E_{k,j}^{k-1}.$$

Tekintsük ugyanis egy  $u \in E_{i,j}^k$  szónak egy  $k$ -megszorított  $A_i \Rightarrow^* uA_j$  levezetését és tekintsük ebben az  $A_k$  előfordulásait. Vegyük észre, hogy az első előfordulásig levezetett prefix, az egymást követő előfordulások közötti részsavak, valamint az utolsó előfordulás után levezetett suffix egyaránt  $(k - 1)$ -megszorított levezetésből adódik.

## Reguláris kifejezések leíró ereje

Tehát mivel minden  $0 \leq k \leq n$ ,  $1 \leq i, j \leq n$ -re  $E_{i,j}^k$  megkapható  $(k - 1)$ -megszorított halmazokból a reguláris műveletekkel, ezért az indukciós hipotézis alapján  $E_{i,j}^k$  szintén leírható reguláris kifejezéssel.

Legyen  $I_\varepsilon$  azon  $i$  indexek halmaza, amelyekre  $A_i \rightarrow \varepsilon \in P$ .

Ekkor  $L(G) = \bigcup_{i \in I_\varepsilon} E_{1,i}^n$ .

Tehát  $L$  leírható reguláris kifejezéssel.



# Reguláris kifejezések leíró ereje

## 1. Példa:

$R = a(a + ba^*)^*$ . Készítsünk 3-típusú grammatikát  $R$ -hez!

A zártsági tétel konstrukcióit alkalmazzuk.

Legyen  $R_1 = a$ ,  $R_2 = ba^*$ ,  $R_3 = R_1 + R_2$ ,  $R_4 = R_3^*$ .

$R_1 : A \rightarrow a$

$R_2 : B \rightarrow bC, C \rightarrow aC \mid \varepsilon$  (kezdőszimbólum:  $B$ )

$R_3 : D \rightarrow A \mid B, A \rightarrow a, B \rightarrow bC, C \rightarrow aC \mid \varepsilon$   
(kezdőszimbólum:  $D$ )

$R_4 : X \rightarrow D \mid \varepsilon, D \rightarrow A \mid B, A \rightarrow a \mid aD, B \rightarrow bC, C \rightarrow aC \mid \varepsilon \mid D$   
(kezdőszimbólum:  $X$ )

$R : S \rightarrow aX, X \rightarrow D \mid \varepsilon, D \rightarrow A \mid B, A \rightarrow a \mid aD, B \rightarrow bC,$   
 $C \rightarrow aC \mid \varepsilon \mid D$  (kezdőszimbólum:  $S$ )

# Reguláris kifejezések leíró ereje

## 2. Példa:

$$A_1 \rightarrow bA_1 \mid aA_2$$

$$A_2 \rightarrow bA_1 \mid \varepsilon$$

$$E_{i,j}^k = E_{i,j}^{k-1} + E_{i,k}^{k-1} \cdot (E_{k,k}^{k-1})^* \cdot E_{k,j}^{k-1}$$

$E_{i,j}^0$	$j = 1$	$j = 2$	$E_{i,j}^1$	$j = 1$	$j = 2$
$i = 1$	$\varepsilon + b$	$a$	$i = 1$	$b^*$	$b^*a$
$i = 2$	$b$	$\varepsilon$	$i = 2$	$b^+$	$\varepsilon + b^+a$

  

$E_{i,j}^2$	$j = 1$	$j = 2$
$i = 1$	$b^* + b^*a(b^+a)^*b^+$	$b^*a(b^+a)^*$
$i = 2$	$(b^+a)^*b^+$	$(b^+a)^*$

## Például:

$$E_{12}^2 = E_{12}^1 + E_{12}^1(E_{22}^1)^*E_{22}^1 =$$

$$b^*a + b^*a(\varepsilon + b^+a)^*(\varepsilon + b^+a) = b^*a + b^*a(b^+a)^* = b^*a(b^+a)^*.$$

$$L(G) = E_{12}^2 = b^*a(b^+a)^*.$$