

Kérem e-mail-ben jelezze, ha hibát talál: ([veanna@inf.elte.hu](mailto:veanna@inf.elte.hu), vagy [veanna@elte.hu](mailto:veanna@elte.hu) )

### 1. feladat megoldása

		a	b
→	1	2	3
	2	4	2
	3	2	1
←	4	6	3
	5	10	6
	6	8	7
	7	9	7
←	8	8	9
←	9	8	8
←	10	5	1

#### I. Összefüggőség vizsgálat.

$$H_0 = \{1\}$$

$$H_1 = \{1, 2, 3\}$$

$$H_2 = \{1, 2, 3, 4\}$$

$$H_3 = \{1, 2, 3, 4, 6\}$$

$$H_4 = \{1, 2, 3, 4, 6, 8, 7\}$$

$$H_5 = \{1, 2, 3, 4, 6, 8, 7, 9\}$$

$$H_6 = H_5$$

5 és 10 kiesik, mert nem érhető el a kezdőállapotból.

Az összefüggő automata:

		a	b
→	1	2	3
	2	4	2
	3	2	1
←	4	6	3
	6	8	7
	7	9	7
←	8	8	9
←	9	8	8

A redukálás (ekvivalens állapotok meghatározása):

$$0 \sim: \{1, 2, 3, 6, 7\} = \text{I.} \quad \{4, 8, 9\} = \text{II.}$$

A halmazok vizsgálata:

	a	b
1	I.	I.
2	II.	I.
3	I.	I.
6	II.	I.
7	II.	I.

	a	b
4	I.	I.
8	II.	II.
9	II.	II.

$$1 \sim: \{1, 3\} \quad \{2, 6, 7\} \quad \{4\} \quad \{8, 9\}$$

halmazokat hasonlóan megvizsgálva:

$2\sim: \{1,3\} \{2\} \{6,7\} \{4\} \{8,9\}$

$3\sim = 2\sim$

A minimális automata:

		a	b
→	{1,3}	{2}	{1,3}
	{2}	{4}	{2}
	{6,7}	{8,9}	{6,7}
←	{4}	{6,7}	{1,3}
←	{8,9}	{8,9}	{8,9}

### 3. feladat megoldása

$$L = b(a \mid \varepsilon)(c \mid ac)^* = ba(c \mid ac)^* \mid b(c \mid ac)^*$$

A maradéknyelvek meghatározása:

0 hosszú „előtag” maradéknyelve:

$$L_\varepsilon = L$$

1 hosszú „előtag”-hoz tartozó maradéknyelvek:

(a szó első betűje a, b, vagy c, mi lehet a folytatás)

$$L_a = \emptyset \quad a\text{-vel nem kezdődhet szó}$$

$$L_b = a(c \mid ac)^* \mid (c \mid ac)^* \quad \text{ezt számoljuk tovább}$$

$$L_c = \emptyset \quad c\text{-vel nem kezdődhet szó}$$

2 hosszú „előtag”-hoz tartozó maradéknyelvek:

(csak a b-vel kezdődő nyelvosztályokat kell kiszámolni, tehát meg kell nézni, mi a maradéknyelv ba, bb, és bc kezdet esetén)

$$L_{ba} = (c \mid ac)^* \quad (\text{új, tovább kell számolni})$$

$$L_{bb} = \emptyset$$

$$L_{bc} = (c \mid ac)^* = L_{ba} \quad (\text{ugyanaz, mint } L_{ba})$$

3 hosszú „előtag”-hoz tartozó maradéknyelvek:

(csak az ba-val és bc-vel kezdődők folytatásaival foglalkozunk)

$$L_{baa} = c(c \mid ac)^* \quad (\text{új, tovább kell számolni})$$

$$L_{bab} = \emptyset$$

$$L_{bac} = L_{ba}$$

4 hosszú „előtag”-hoz tartozó maradéknyelvek:

(csak az baa-val kezdődők folytatásaival foglalkozunk)

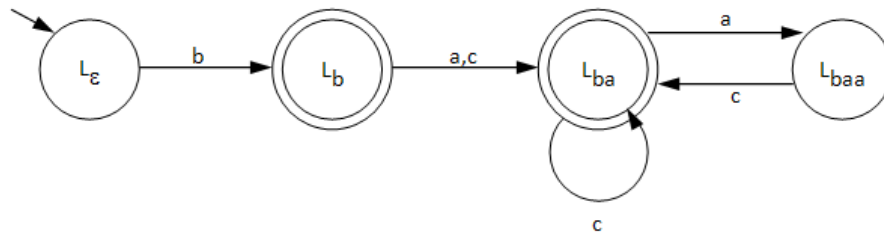
$$L_{baaa} = \emptyset$$

$$L_{baab} = \emptyset$$

$$L_{baac} = L_{ba}$$

Nem keletkezett új maradéknyelv, készen vagyunk. Sárgával és bíborral jelzettek a különböző maradék nyelvek. Bíbor színűek azok, amelyek az  $\varepsilon$ -t tartalmazzák, így végállapotok lesznek.

Az automata gráffal ábrázolva (ez még csak PDA, a hibaállapot felvételével lesz VDA):



### 5. feladat megoldása

Redukálja az alábbi 2-es típusú nyelvtant:

$G = \langle \{a,b,c\}, \{S,A,B,C,D,F,K\}, P, S \rangle$

$P = \{$   
 $S \rightarrow aAS \mid AD$   
 $A \rightarrow aCBb \mid aa \mid bD$   
 $B \rightarrow BA \mid bbF \mid F$   
 $C \rightarrow aA$   
 $D \rightarrow bCB \mid bA \mid KK \mid \varepsilon$   
 $F \rightarrow aFa \mid BS$   
 $K \rightarrow ScS \mid cCBc \}$

1. Aktív nyelvtani jelek meghatározása, majd a nem aktív (zsákutca) nyelvtani jelek kiszűrése.

$A_1 = \{A, D\}$

$A_2 = \{A, D, S, C\}$

$A_3 = \{A, D, S, C, K\}$

$A_4 = A_3$

tehát B és F inaktív, az első lépés után a nyelvtan:

$P = \{$   
 $S \rightarrow aAS \mid AD$   
 $A \rightarrow aa \mid bD$   
 $C \rightarrow aA$   
 $D \rightarrow bA \mid KK \mid \varepsilon$   
 $K \rightarrow ScS \}$

2. Elérhető nyelvtani jelek meghatározása (nyelvtan összefüggőségének vizsgálata).

$R_0 = \{S\}$

$R_1 = \{S, A, D\}$

$R_2 = \{S, A, D, K\}$

$R_3 = R_2$

tehát C nem érhető el a kezdőszimbólumból, így a redukált nyelvtan:

$G = \langle \{a,b,c\}, \{S,A,D,K\}, P, S \rangle$

$P = \{$   
 $S \rightarrow aAS \mid AD$   
 $A \rightarrow aa \mid bD$   
 $D \rightarrow bA \mid KK \mid \varepsilon$   
 $K \rightarrow ScS \}$

**6. feladat megoldása**

Hozza Chomsky normál formára az alábbi  $\varepsilon$  mentes 2-es típusú nyelvtant:

$G = \langle \{a,b,c\}, \{S,A,B,C,D\}, P, S \rangle$

$P = \{$   
 $S \rightarrow aAS \mid D \mid CSC$   
 $A \rightarrow bB \mid aa \mid C$   
 $B \rightarrow bABBa \mid SC$   
 $C \rightarrow AS \mid D \mid c$   
 $D \rightarrow bA \mid b \}$

1. Álterminálisok bevezetése (Ha a szabály nem  $A \rightarrow t$  ( $A \in N, t \in T$ ) alakú, átalakítjuk úgy, hogy csupa nemterminális jelből álljon a jobb oldala):

$P = \{$   
 $S \rightarrow G_aAS \mid D \mid CSC$   
 $A \rightarrow G_bB \mid G_aG_a \mid C$   
 $B \rightarrow G_bABBG_a \mid SC$   
 $C \rightarrow AS \mid D \mid c$   
 $D \rightarrow G_bA \mid b$   
 $G_a \rightarrow a$   
 $G_b \rightarrow b \}$

Hosszredukció (a fenti, beszínezett szabályok átalakítása):

$P = \{$   
 $S \rightarrow G_aZ_1 \mid D \mid CZ_2$   
 $A \rightarrow G_bB \mid G_aG_a \mid C$   
 $B \rightarrow G_bZ_3 \mid SC$   
 $C \rightarrow AS \mid D \mid c$   
 $D \rightarrow G_bA \mid b$   
 $G_a \rightarrow a$   
 $G_b \rightarrow b$   
 $Z_1 \rightarrow AS$   
 $Z_2 \rightarrow SC$   
 $Z_3 \rightarrow AZ_4 \quad Z_4 \rightarrow BZ_5 \quad Z_5 \rightarrow BG_a \quad \}$

2. Lánementesítés (beszínezett szabályok a láncszabályok):

$P = \{$   
 $S \rightarrow G_aZ_1 \mid D \mid CZ_2$   
 $A \rightarrow G_bB \mid G_aG_a \mid C$   
 $B \rightarrow G_bZ_3 \mid SC$   
 $C \rightarrow AS \mid D \mid c$   
 $D \rightarrow G_bA \mid b$   
 $G_a \rightarrow a$   
 $G_b \rightarrow b$   
 $Z_1 \rightarrow AS$   
 $Z_2 \rightarrow SC$   
 $Z_3 \rightarrow AZ_4 \quad Z_4 \rightarrow BZ_5 \quad Z_5 \rightarrow BG_a \quad \}$

Halmazok felépítése:

$H_1(S) = \{S\}$     $H_2(S) = \{S, D\} = H_3(S)$     $H(S) = \{S, D\}$   
 $H_1(A) = \{A\}$     $H_2(A) = \{A, C\}$     $H_3(A) = \{A, C, D\} = H_4(A)$     $H(A) = \{A, C, D\}$   
 $H(B) = \{B\}$   
 $H_1(C) = \{C\}$     $H_2(C) = \{C, D\} = H_3(C)$     $H(C) = \{C, D\}$

$$H(D) = \{D\}$$

(Az előző lépésekben felvett új nyelvtani jeleknek biztosan nincs láncszabályuk, így azok halmazait nem is számoljuk ki.)

A kapott nyelvtan:

$$G = \langle \{a, b, c\}, \{S, A, B, C, D, G_a, G_b, Z_1, Z_2, Z_3, Z_4, Z_5\}, P, S \rangle$$

$$P = \{ S \rightarrow G_a Z_1 \mid CZ_2 \mid G_b A \mid b$$

$$A \rightarrow G_b B \mid G_a G_a \mid AS \mid c \mid G_b A \mid b$$

$$B \rightarrow G_b Z_3 \mid SC$$

$$C \rightarrow AS \mid c \mid G_b A \mid b$$

$$D \rightarrow G_b A \mid b$$

$$G_a \rightarrow a$$

$$G_b \rightarrow b$$

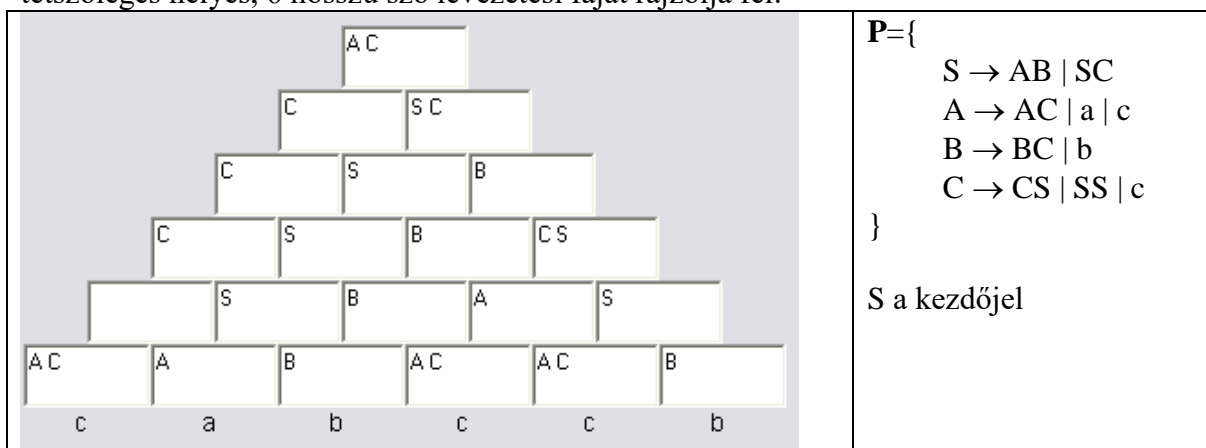
$$Z_1 \rightarrow AS$$

$$Z_2 \rightarrow SC$$

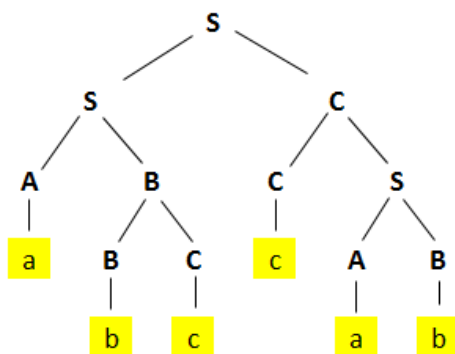
$$Z_3 \rightarrow AZ_4 \quad Z_4 \rightarrow BZ_5 \quad Z_5 \rightarrow BG_a \quad \}$$

### 7. feladat megoldása

A CYK algoritmus segítségével döntse el, hogy a *cabccb* szó eleme-e a nyelvtan által generált nyelvnek. Ha igen, rajzolja fel a szó egy lehetséges levezetési fáját is, ha nem, akkor egy tetszőleges helyes, 6 hosszú szó levezetési fáját rajzolja fel:



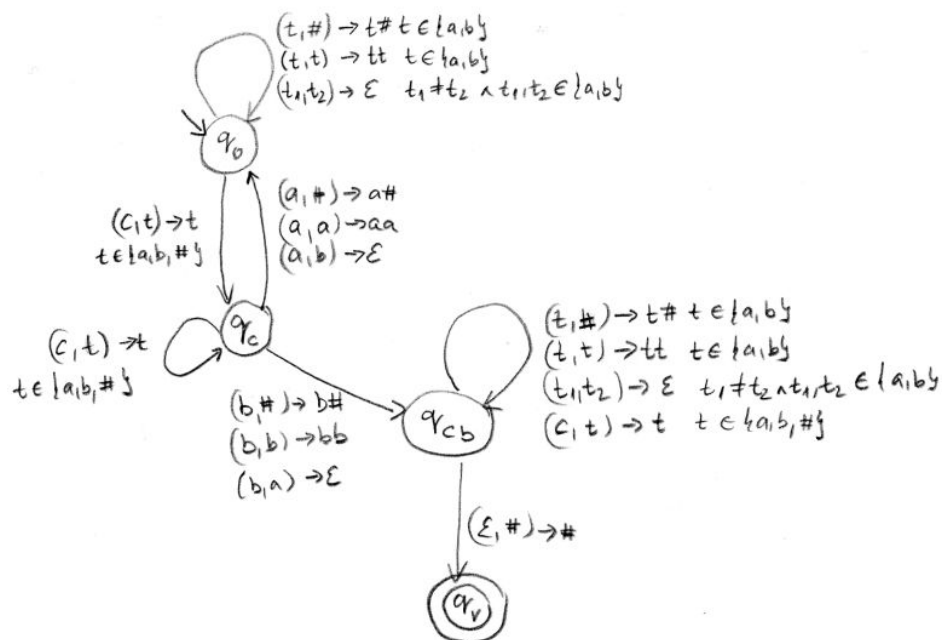
A szó nem eleme a generált nyelvnek. Levezethető lenne az A és C nemterminálisból, de S-ből nem! Példa 6 hosszú helyes szó levezetési fájára:



## 8. feladat megoldása

Az  $L = \{u \in \{a,b,c\}^* \mid |a(u)| = |b(u)| \text{ és } cb \subset u\}$  nyelvhez kell verem automatát készíteni. Az automata gráfos megjelenítéssel:

$$V = \langle \{q_0, q_c, q_{cb}, q_v\}, \{a, b, c\}, \{a, b, \#\}, \delta, q_0, \#, \{q_v\} \rangle$$



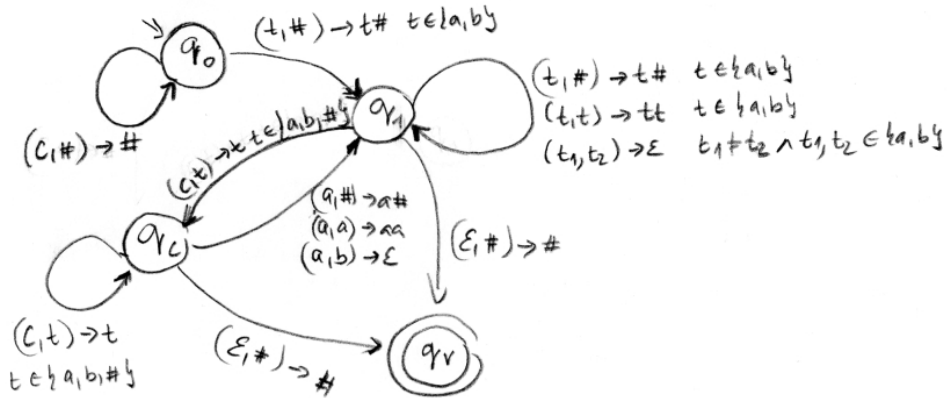
Magyarázat:

- $q_0$  állapotban gyűjti és számlálja az 'a' és 'b' betűket, a veremben mindig a többlet 'a' vagy a többlet 'b' van, vagy #, ha épp egyensúly van.
- Ha jön egy 'c' átmegy  $q_c$  állapotba, ott tetszőleges számú 'c'-t olvashat. De ha jön egy 'a' akkor visszatér  $q_0$ -ba, mert nem érkezett meg a helyes szóhoz szükséges 'cb' szótag.
- $q_c$ -ből  $q_{cb}$ -be lép, ha 'b' érkezik, ilyenkor a helyes szóhoz megérkezett a szükséges 'cb' szótag.  $q_{cb}$  állapotban hasonlóan  $q_0$  állapothoz az 'a' és 'b' betűket számolja, ha 'c' jönne, azzal nem foglalkozik.
- $q_v$ -be léphet, ha a veremben megjelenik a '#'. Ha az input szalagot közben sikerült végigolvasni, akkor helyes volt a szó.

**9. feladat megoldása**

Az  $L = \{u \in \{a,b,c\}^* \mid l_a(u) = l_b(u) > 0 \text{ és } cb \notin u\}$  nyelvhez kell verem automatát készíteni. Az automata gráfos megjelenítéssel:

$$V = \langle \{q_0, q_1, q_c, q_v\}, \{a, b, c\}, \{a, b, \#\}, \delta, q_0, \#, \{q_v\} \rangle$$



Magyarázat:

- $q_0$  állapotban van mindaddig, amíg 'c' betűk vannak az inputon. Ha jön egy 'a' vagy 'b' átmeny  $q_1$ -be (az volt a helyes szó feltétele, hogy legalább egy 'a' és 'b' legyen benne).
- $q_1$  állapotban gyűjti és számlálja az 'a' és 'b' betűket, a veremben mindig a többlet 'a' vagy a többlet 'b' van, vagy #, ha épp egyensúly van.
- Ha jön egy 'c' átmeny  $q_c$  állapotba, ott tetszőleges számú 'c'-t olvashat. Csak 'a'-val térhet vissza  $q_1$ -be, mert nem lehet a helyes szóban 'cb' szótag.
- $q_v$ -be léphet, ha a veremben megjelenik a '#'. Ez előfordulhat úgy, hogy 'c'-re végződik a szó, vagy úgy, hogy 'a' vagy 'b' betűre végződik. Ha az input szalagot közben sikerült végigolvasni, akkor helyes volt a szó.

Ha nincs kikötve, hogy  $l_a(u) > 0$  és  $l_b(u) > 0$ , azaz az üres szó, vagy a csupa 'c' betűből álló szó is helyes, akkor egy egyszerűbb automatát kapunk:

