Az informatika számítástudományi alapjai
7. előadás

Vaszil György

vaszil.gyorgy@inf.unideb.hu

I. emelet 110-es szoba

A múlt órán

- Chomsky féle normálforma, Cocke-Younger-Kasami algoritmus
- Pumpálási lemma környezetfüggetlen nyelvekre

Chomsky féle normál-forma

- Egy grammatika Chomsky normálfor-mában van, ha ~~[illegible]~~ csak

$$\boxed{A \to BC \text{ és } A \to a} \qquad a \in \Sigma,\ A, B, C \in N$$

alakú szabályokat tartalmaz.

- Minden G-hez van olyan G_1 Chomsky normálformában, hogy $L(G_1) = L(G) - \{\lambda\}$.

- Alapötlet:

$$A \to BCDE \Big\} \iff \begin{cases} A \to BX \\ X \to CY \\ Y \to DE \end{cases}$$

Mire jó a Chomsky normálforma:

- Cocke-Younger-Kasami algoritmus
  - adott egy *G* környezetfüggetlen grammatika Chomsky normálformában
  - adott egy *w* sztring

Az algoritmus eldönti, hogy generálható-e a *w* sztring a *G* nyelvtannal. (Előállítja a lehetséges levezetéseket is.)

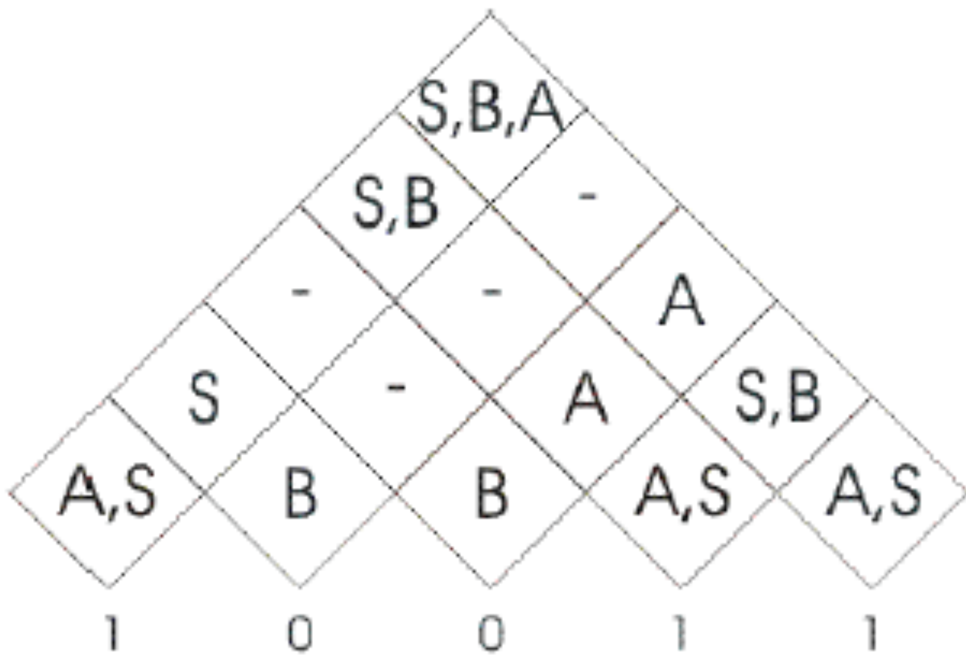
Mire jó a Chomsky normálforma: Cocke-Younger-Kasami algoritmus

Tekintsük a következő grammatikát!

$G=(\{S,A,B\},\{0,1\},S,H)$,ahol H szabályai:

$\{S \to SA,\ S \to AB,\ A \to BS,\ B \to SA,\ A \to 1,\ S \to 1,\ B \to 0\}$

Bizonyítsuk be, hogy az 10011 szó benne van a grammatika által generált nyelvben,



(miért kellett a Chomsky normálforma?)

A múlt órán

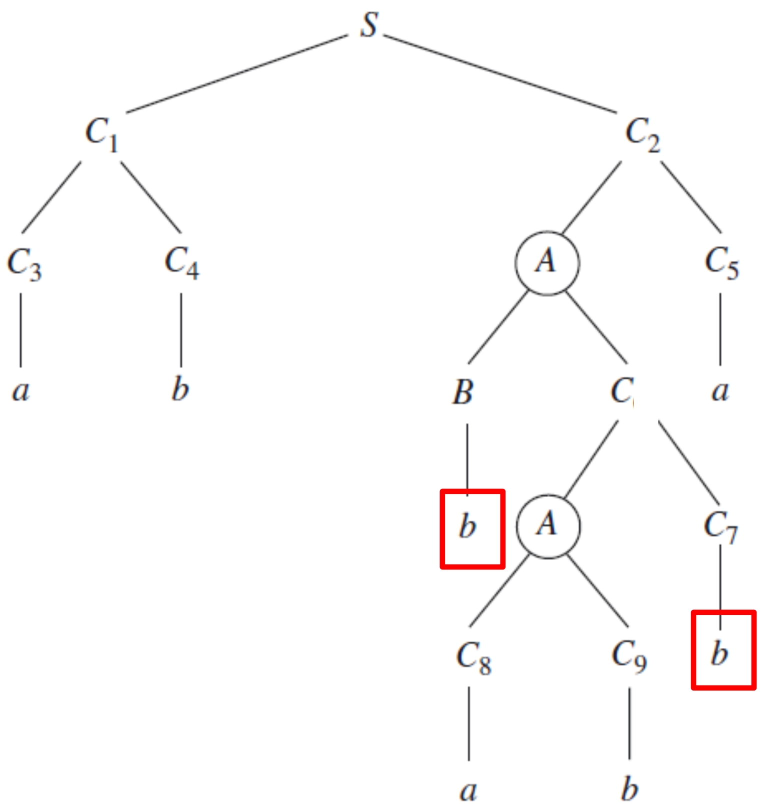
- Chomsky féle normálforma, Cocke-Younger-Kasami algoritmus
- Pumpálási lemma környezetfüggetlen nyelvekre

Pumpálási lemma: Ha L környezetfüggetlen akkor létezik p, hogy ha $s \in L$ és $|s| \geq p$ akkor s felírható $s = uvxyz$ alakban, ahol

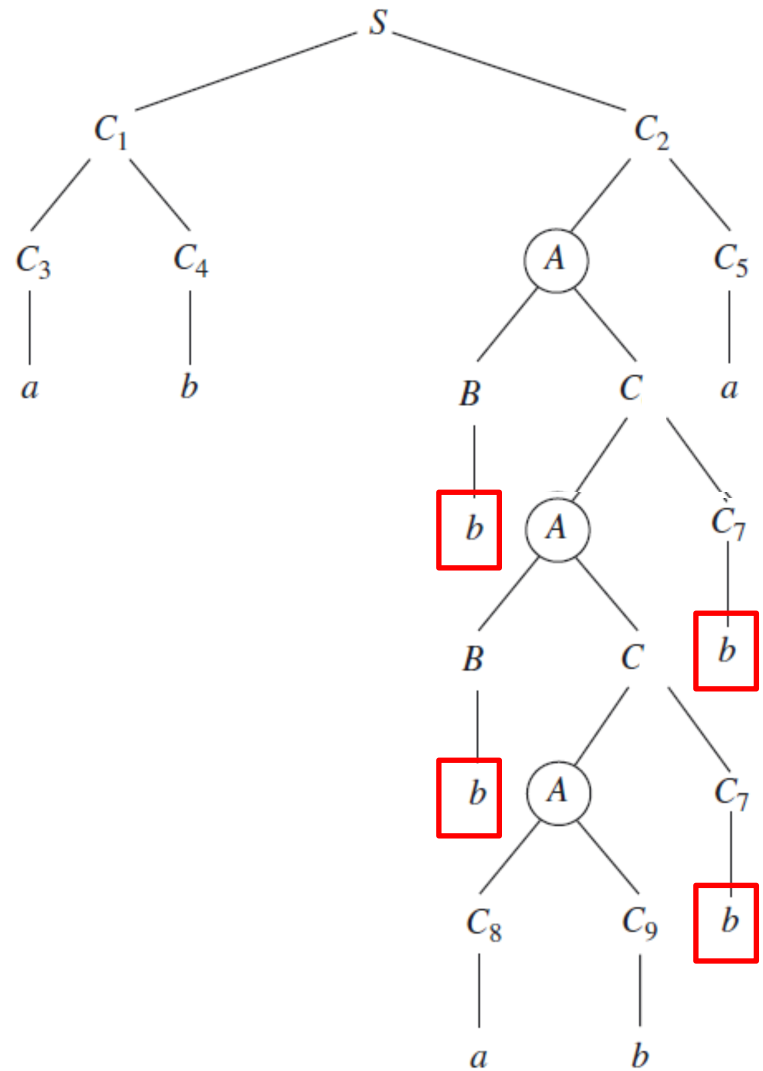
1. $|vxy| \leq p$
2. $|vy| > 0$
3. $uv^ixy^iz \in L$ minden $i \geq 0$-ra

(környezetfüggetlen nyelv = környezetfüggetlen grammatikával generálható nyelv)

Levezetési fák közelebbről

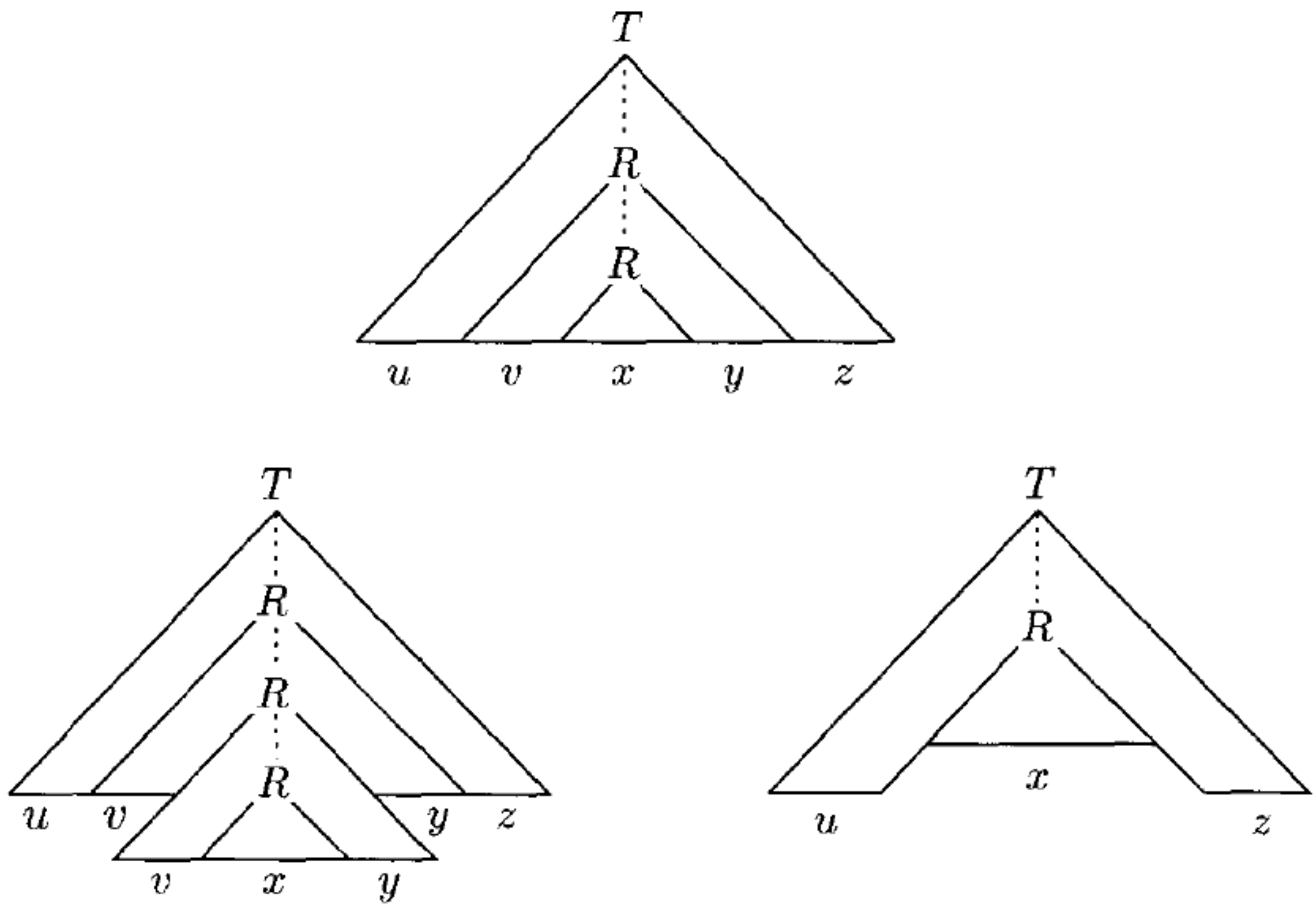


$u = (ab)\boxed{(b)}(ab)\boxed{(b)}(a)$



$u=(ab)\boxed{(b)(b)}(ab)\boxed{(b)(b)}(a)$

van $A\rightarrow BC$, $C\rightarrow AC_7$ és
$A\rightarrow C_8C_9$ szabály
(meg egy csomó más szabály)


T
R
R
u
v
x
y
z
T
R
R
R
u
v
y
z
v
x
y
T
R
x
u
z

Példaként

$L = \{a^n b^n c^n \mid n \geq 0\}$ nem környezet-független. Hiszen:

Ha L környezetfüggetlen volna, létezne p, hogy minden $s \in L$, $|s| > p$ se pumpálható volna. Vegyük $s = a^p b^p c^p$ -t.

Nem lehet pumpálni, tehát L nem lehet környezetfüggetlen.

Nézzük meg az *$\{w \in \{a,b,c\}^*$ ahol $|w|_a = |w|_b = |w|_c\}$* nyelvet is

Nézzük meg az $\{w \in \{a,b,c\}^* \text{ ahol } |w|_a = |w|_b = |w|_c\}$ nyelvet is!

Pumpálási lemma: Ha L környezetfüggetlen akkor létezik p, hogy ha $s \in L$ és $|s| \geq p$ akkor s felírható $s = uvxyz$ alakba, ahol

1. $|vxy| \leq p$
2. $|vy| > 0$
3. $uv^ixy^iz \in L$ minden $i \geq 0$-ra

(környezetfüggetlen nyelv = környezetfüggetlen grammatikával generálható nyelv)

A múlt órán

- Chomsky féle normálforma, Cocke-Younger-Kasami algoritmus
- Pumpálási lemma környezetfüggetlen nyelvekre

MA

- Veremautomaták, környezetfüggetlen nyelvek elfogadása veremautomatával
- Determinisztikus veremautomaták, determinisztikus környezetfüggetlen nyelvek

Legyen G egy környezetfüggetlen nyelvtan, szabályai:

$$S \to aSa \mid bSb \mid c$$

Ekkor

$$L(G) = \{ wcw^R \mid w \in \{a,b\}^* \}$$

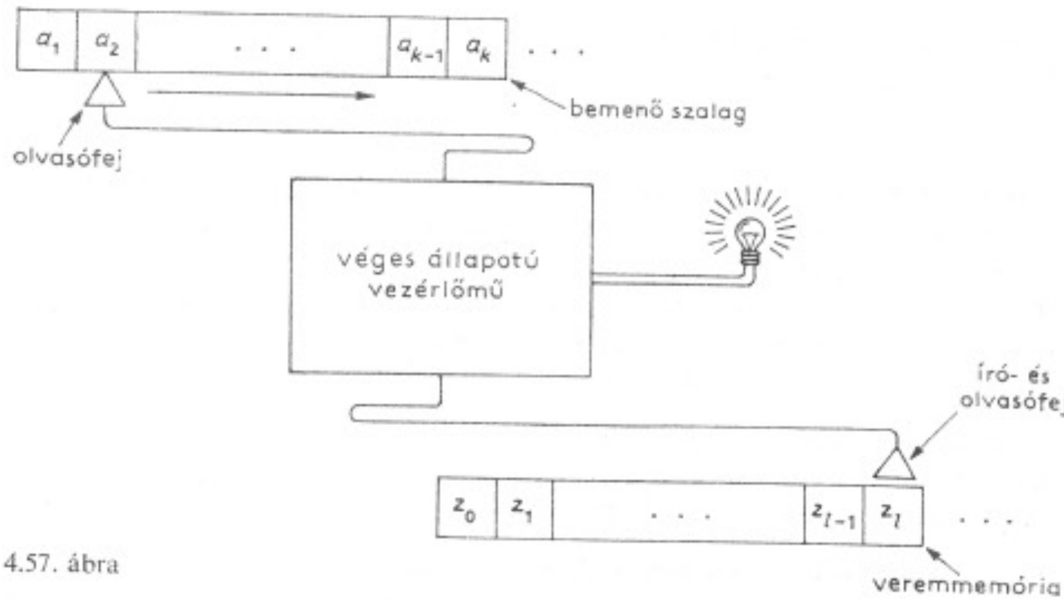
Emlékezzünk rá, hogy L(G) nem reguláris.

L(G)-t nem lehet véges automatával elfogadni.

(Miért nem?)

Veremautomata

A véges automatát egészítsük ki egy verem-memóriával → veremautomata



4.57. ábra

Állapotátmenet:

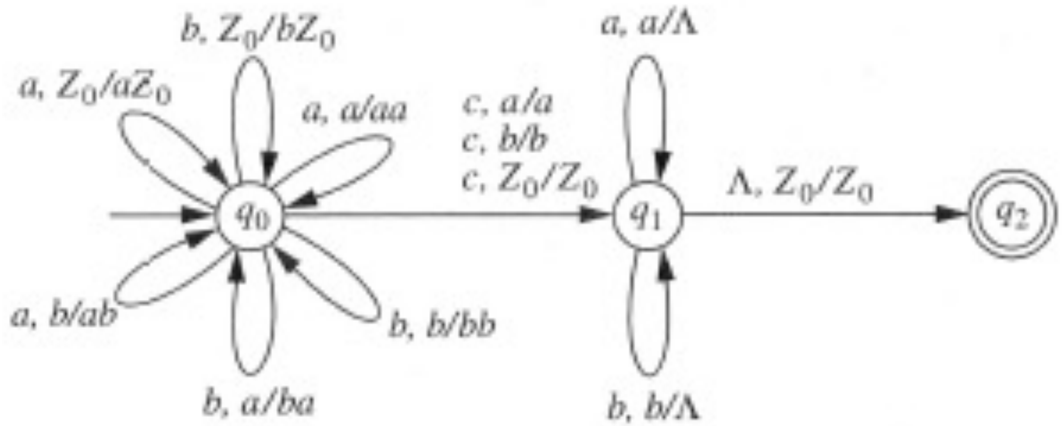
(bemeneti szimbólum, előző állapot, a verem teteje) → (új állapot, a verem tetején lévő szimbólum cseréje)

(Mire cserélhető a verem tetején lévő szimbólum? Üresszóra?)

Példa, veremautomata

Table 7.1 | Transition table for Example 7.1

Move number	State	Input	Stack symbol	Move(s)
1	q_0	a	Z_0	(q_0, aZ_0)
2	q_0	b	Z_0	(q_0, bZ_0)
3	q_0	a	a	(q_0, aa)
4	q_0	b	a	(q_0, ba)
5	q_0	a	b	(q_0, ab)
6	q_0	b	b	(q_0, bb)
7	q_0	c	Z_0	(q_1, Z_0)
8	q_0	c	a	(q_1, a)
9	q_0	c	b	(q_1, b)
10	q_1	a	a	(q_1, Λ)
11	q_1	b	b	(q_1, Λ)
12	q_1	Λ	Z_0	(q_2, Z_0)
	(all other combinations)			none



Kezdő állapot: q_0

Elfogadó állapot: q_2

Kezdetben a verem alján lévő betű: Z_0

Mi történik a abcba, ab, acaa elemekkel?

(Az elfogadás feltételei. Milyen nyelvet is fogad

Veremautomata, definíció

/1

$$M = (Q, \Sigma, \Gamma, q_0, z_0, \delta, F)$$

ahol:

Q - állapothalmaz

Σ - bemeneti ábécé

Γ - veremábécé

$q_0 \in Q$ kezdőállapot

$z_0 \in \Gamma$ kezdeti veremtartalom

δ - állapotátmenet reláció

$F \subseteq Q$ végállapotok halmaza

Veremautomata, definíció /2

Az állapot átmeneti reláció:

$$\delta : (Q \times \Sigma \cup \{\lambda\} \times \Gamma) \to 2^{(Q \times \Gamma^*)}$$

állapot

bemenő szimbólum vagy üres szó

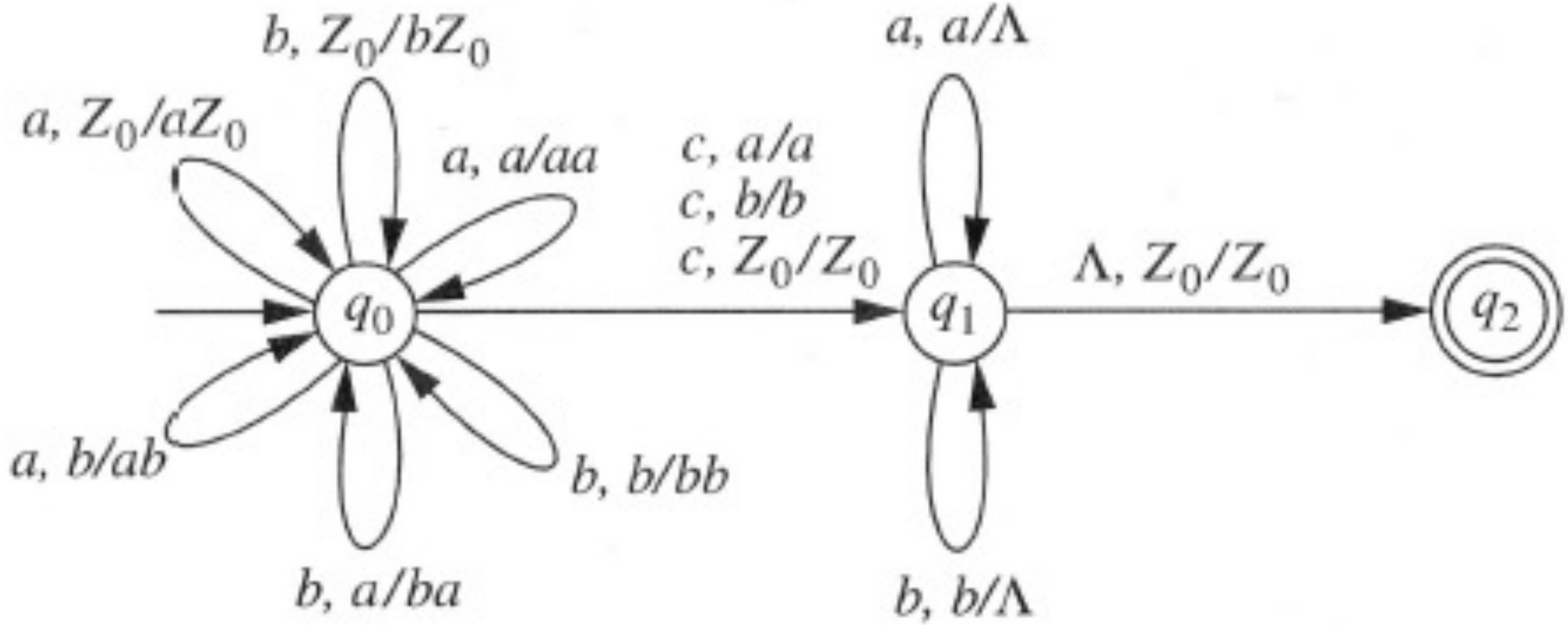
a verem teteje

új állapot

a verem tetejére tett szó

állapot – szó párok halmaza
(nemdtereminisztikus veremautomata)

Például



$\delta(q_0, b, a) = (q_0, ba)$

a verembe teszi b-t

$\delta(q_1, b, b) = (q_1, \lambda)$

a veremből kiveszi b-t

Konfiguráció, konfiguráció sorozat

Konfiguráció: (p, x, α)

$p \in Q$, $x \in \Sigma^*$, $\alpha \in \Gamma^*$

Konfiguráció átmenet: $(p, x, \alpha) \to (q, y, \beta)$

ha

- $x = ay$, $a \in \Sigma \cup \{\lambda\}$
- $\alpha = X\alpha'$, $X \in \Gamma$, $\alpha' \in \Gamma^*$
- $\beta = \gamma\alpha'$, $\gamma \in \Gamma^*$ és
- $(q, \gamma) \in \delta(p, a, X)$

(tábla)

Elfogadott nyelv, szó

Legyen

$$M = (Q, \Sigma, \Gamma, q_0, Z_0, \delta, F).$$

Az M által elfogadott nyelv:

$$L(M) = \{ w \in \Sigma^* \mid (q_0, w, Z_0) \rightarrow \dots \rightarrow (q_f, \lambda, \alpha) \}$$

$q_f \in F$

(**Létezik** olyan konfiguráció átmenet sorozat, ami
a kezdő konfigurációból a w szó elolvasásának hatásár
elfogadó állapotot tartalmazó konfigurációba visz)

A veremautomaták által megadható nyelvek azok, amelyek környezetfüggetlen grammatikával generálhatóak.

Hogyan lehet tetszőleges
környezetfüggetlen
nyelvhez veremautomatát adni?

Például: $S \to [S] \mid SS \mid \lambda$

Move Number	State	Input	Stack Symbol	Move
1	q_0	Λ	Z_0	(q_1, SZ_0)
2	q_1	Λ	S	$(q_1, [S]), (q_1, SS), (q_1, \Lambda)$
3	q_1	[	[	(q_1, Λ)
4	q_1	]	]	(q_1, Λ)
5	q_1	Λ	Z_0	(q_2, Z_0)
	(all other combinations)			none

Kezdőállapot: q_0
Elfogadó állapot: q_2
Kezdeti veremtartalom: Z_0

(mit generál a grammatika, vegyünk egy jó [[]] és egy rossz példát, nemdeterminizmus)

Move Number	State	Input	Stack Symbol	Move
1	q_0	Λ	Z_0	(q_1, SZ_0)
2	q_1	Λ	S	$(q_1, [S]), (q_1, SS), (q_1, \Lambda)$
3	q_1	[	[	(q_1, Λ)
4	q_1	]	]	(q_1, Λ)
5	q_1	Λ	Z_0	(q_2, Z_0)
	(all other combinations)			none

$$
\begin{aligned}
&(q_0, [[][]], Z_0)\\
&\quad\vdash (q_1, [[][]], SZ_0)\\
&\quad\vdash (q_1, [[][]], [S]Z_0)\\
&\quad\vdash (q_1, [][]], S]Z_0)\\
&\quad\vdash (q_1, [][]], SS]Z_0)\\
&\quad\vdash (q_1, [][]], [S]S]Z_0)\\
&\quad\vdash (q_1,][]], S]S]Z_0)\\
&\quad\vdash (q_1,][]],]S]Z_0)\\
&\quad\vdash (q_1, []], S]Z_0)\\
&\quad\vdash (q_1, []], [S]]Z_0)\\
&\quad\vdash (q_1,]], S]]Z_0)\\
&\quad\vdash (q_1,]],]]Z_0)\\
&\quad\vdash (q_1,],]Z_0)\\
&\quad\vdash (q_1, \Lambda, Z_0)\\
&\quad\vdash (q_2, \Lambda, Z_0)
\end{aligned}
$$

$$
\begin{aligned}
&S\\
&\quad\Rightarrow [S]\\
&\quad\Rightarrow [SS]\\
&\quad\Rightarrow [[S]S]\\
&\quad\Rightarrow [[]S]\\
&\quad\Rightarrow [[][S]]\\
&\quad\Rightarrow [[][]]
\end{aligned}
$$

(mi történik a veremben?)

Állítás

Adott $G=(N, \Sigma, S, P)$, legyen M

$$M = (\{q_0, q_1, q_2\}, \Sigma, N \cup \Sigma \cup \{Z_0\}, q_0, \delta, \{q_2\})$$

ahol δ:

- $\delta(q_0, \lambda, Z_0) = \{(q_1, SZ_0)\}$
- $\delta(q_1, \lambda, A) = \{(q_1, \alpha) \mid A \to \alpha \in P\}$ minden $A \in N$-re
- $\delta(q_1, a, a) = \{(q_1, \lambda)\}$ minden $a \in \Sigma$-ra
- $\delta(q_1, \lambda, Z_0) = \{(q_2, Z_0)\}$

Ekkor $L(G) = L(M)$.

$S \to [S] \mid SS \mid \lambda$

State	Input	Stack Symbol	Move
q_0	Λ	Z_0	(q_1, SZ_0)
q_1	Λ	S	$(q_1, [S]), (q_1, SS), (q_1, \Lambda)$
q_1	[	[	(q_1, Λ)
q_1	]	]	(q_1, Λ)
q_1	Λ	Z_0	(q_2, Z_0)

Azaz:

Tétel:

L környezetfüggetlen $\Longrightarrow$ Van olyan M veremautomata, amire
$$L = L(M)$$

Bizonyítási ötlet:

Lásd az előző konstrukciót.

Azaz:

Tétel:

L környezetfüggetlen $\Longrightarrow$ Van olyan M veremautomata, amire
$$L = L(M)$$

Bizonyítási ötlet:

Lásd az előző konstrukciót.

Igaz-e ez „visszafelé" is? Azaz igaz-e, hogy bármely veremautomata elfogadott nyelve környezetfüggetlen? Igen, de ezt nem bizonyítjuk.

Azaz : /2

L környezetfüggetlen $\Longleftrightarrow$ Van olyan veremautomata M, hogy

$$L = L(M)$$

— * —

Vagyis a környezetfüggetlen nyelveknél a veremautomata hasonló „szerepet játszik" mint a reguláris nyelveknél a véges automata.

Van-e további analógia → pl. determinizmus

MA

- Veremautomaták, környezetfüggetlen nyelvek elfogadása veremautomatával
- Determinisztikus veremautomaták, determinisztikus környezetfüggetlen nyelvek

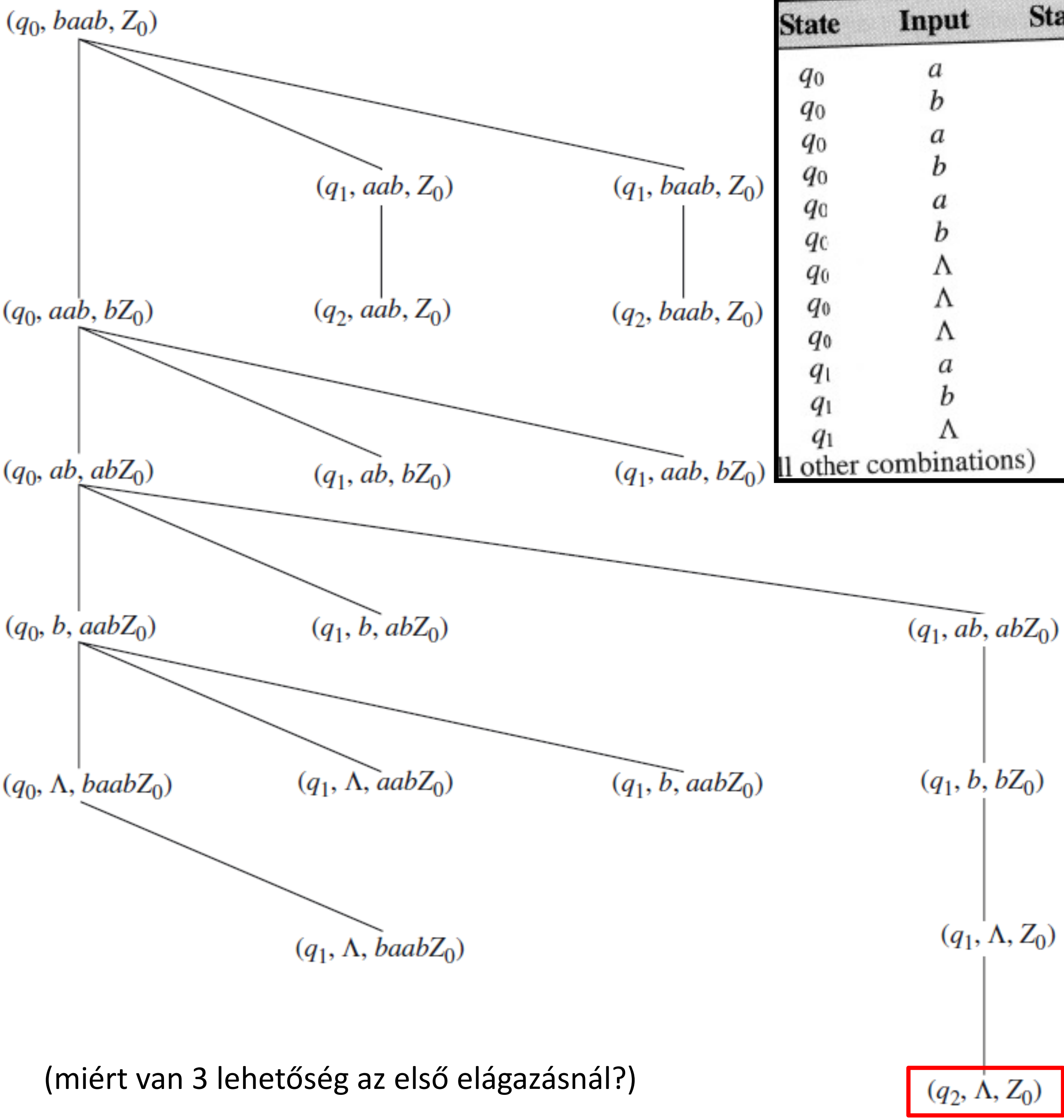
Példa, palindrómák

$$M = (\{q_0, q_1, q_2\}, \{a, b\}, \{a, b, Z_0\}, q_0, Z_0, \delta, \{q_2\})$$

Table 7.2 Transition table for

Move number	State	Input	Stack symbol	Move(s)
1	q_0	a	Z_0	$(q_0, aZ_0), (q_1, Z_0)$
2	q_0	b	Z_0	$(q_0, bZ_0), (q_1, Z_0)$
3	q_0	a	a	$(q_0, aa), (q_1, a)$
4	q_0	b	a	$(q_0, ba), (q_1, a)$
5	q_0	a	b	$(q_0, ab), (q_1, b)$
6	q_0	b	b	$(q_0, bb), (q_1, b)$
7	q_0	Λ	Z_0	(q_1, Z_0)
8	q_0	Λ	a	(q_1, a)
9	q_0	Λ	b	(q_1, b)
10	q_1	a	a	(q_1, Λ)
11	q_1	b	b	(q_1, Λ)
12	q_1	Λ	Z_0	(q_2, Z_0)
	(all other combinations)			none

Legyen a bemenet baab. (Írjuk fel a lehetséges konfigurációkat!)



State	Input	Stack symbol	Move(s)
q_0	a	Z_0	$(q_0, aZ_0), (q_1, Z_0)$
q_0	b	Z_0	$(q_0, bZ_0), (q_1, Z_0)$
q_0	a	a	$(q_0, aa), (q_1, a)$
q_0	b	a	$(q_0, ba), (q_1, a)$
q_0	a	b	$(q_0, ab), (q_1, b)$
q_0	b	b	$(q_0, bb), (q_1, b)$
q_0	Λ	Z_0	(q_1, Z_0)
q_0	Λ	a	(q_1, a)
q_0	Λ	b	(q_1, b)
q_1	a	a	(q_1, Λ)
q_1	b	b	(q_1, Λ)
q_1	Λ	Z_0	(q_2, Z_0)
ll other combinations)			none

(miért van 3 lehetőség az első elágazásnál?)

Determinisztikus veremautomata

1. Bármely $q \in Q$, $a \in \Sigma \cup \{\lambda\}$, $X \in \Gamma$ hármasra, $\delta(q,a,X)$ legfeljebb egy elemű.

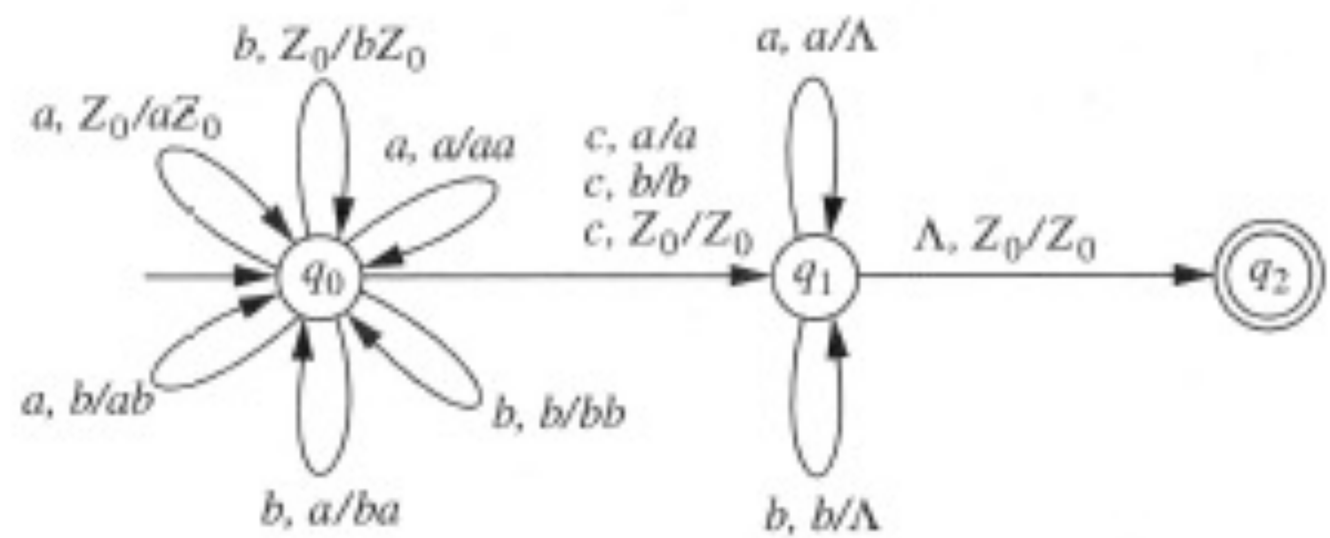
2. Ha $\delta(q,\lambda,X) \neq \emptyset$, akkor $\delta(q,a,X) = \emptyset$

minden $a \in \Sigma$-ra!
$(a \neq \lambda)$

Ha egy állapot-veremszimbólum
párhoz van „üresszó” (input olvasás nélküli)
átmenet, akkor csak „üresszó” átmenet van.

Korábbi determinisztikus példa

Move number	State	Input	Stack symbol	Move(s)
1	q_0	a	Z_0	(q_0, aZ_0)
2	q_0	b	Z_0	(q_0, bZ_0)
3	q_0	a	a	(q_0, aa)
4	q_0	b	a	(q_0, ba)
5	q_0	a	b	(q_0, ab)
6	q_0	b	b	(q_0, bb)
7	q_0	c	Z_0	(q_1, Z_0)
8	q_0	c	a	(q_1, a)
9	q_0	c	b	(q_1, b)
10	q_1	a	a	(q_1, Λ)
11	q_1	b	b	(q_1, Λ)
12	q_1	Λ	Z_0	(q_2, Z_0)
	(all other combinations)			none



Mi történik a abcba, ab, acaa elemzésekor?

Nemdeterminisztikus veremautomata ugyanarra a nyelvre (az előző konstrukció alapján)

- $G = (N, \Sigma, S, P)$

$S \to aSa$
$S \to bSb$
$S \to c$

$\longleftrightarrow$

$M = (Q, \Sigma, \Gamma, q_0, \delta, \{q_2\})$

$\delta(q_0, \lambda, Z_0) = \{(q_1, SZ_0)\}$

$\boxed{\delta(q_1, \lambda, S) = \{(q_1, aSa), (q_1, bSb)\}}$

$\delta(q_1, a, a) = \{(q_1, \lambda)\}$

$\delta(q_1, b, b) = \{(q_1, \lambda)\}$

$\delta(q_1, c, c) = \{(q_1, \lambda)\},$

$\delta(q_1, \lambda, Z_0) = \{(q_2, Z_0)\}$

5

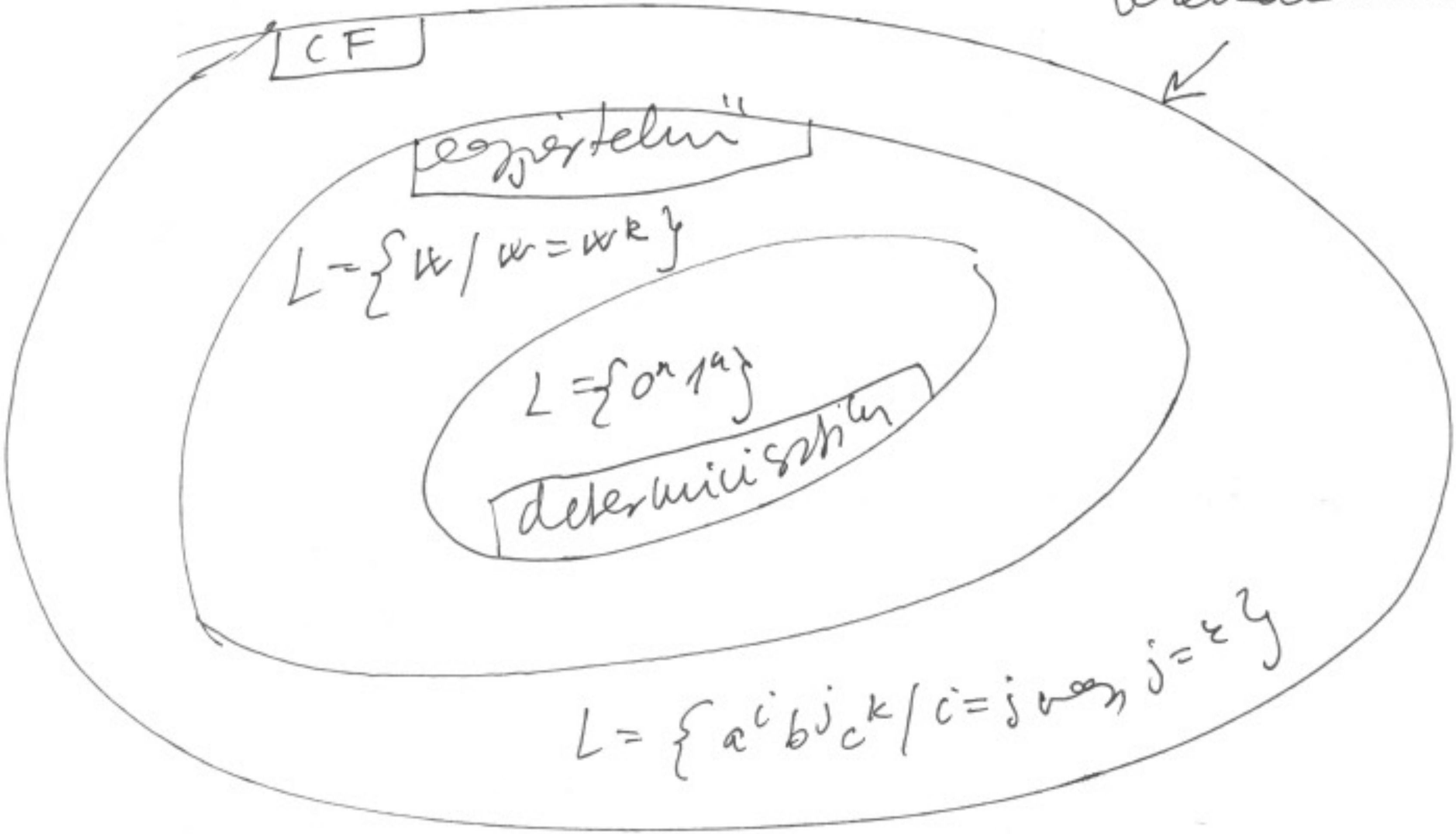
Ellenben:

Általában nem igaz, hogy minden környezet-független nyelvhez lehet determinisztikus veremautomatát csinálni.

Például $L = \{ w \in \{a,b\}^* \mid w = w^R \}$ palindrómák, a „közepük” nincs megjelölve.

(bizonyítás nélkül)

Környezetfüggetlen = van rá nemdeterminisztikus veremautomata



A példák

- $L = \{0^n 1^n \mid n \geq 1\}$ = determinisztikus (miért?)
- $L = \{w \in \{a,b\}^* \mid w = w^R\}$ = nem determinisztikus, de egyértelmű: $S \to aSa \mid bSb$, $S \to a \mid b \mid \lambda$
- $L = \{a^i b^j c^k \mid i = j \text{ vagy } j = k\}$ = nem egyértelmű

Ez az egyik egyszerű példáink azért:

- A programozási nyelvek szintaxisa leírható környezetfüggetlen grammatikákkal
- A compilereknek ellenőrizni kell fordításkor, hogy a program szintaktikailag helyes-e

MA

- Veremautomaták, környezetfüggetlen nyelvek elfogadása veremautomatával
- Determinisztikus veremautomaták, determinisztikus környezetfüggetlen nyelvek