Veremautomaták és környezetfüggetlen nyelvek

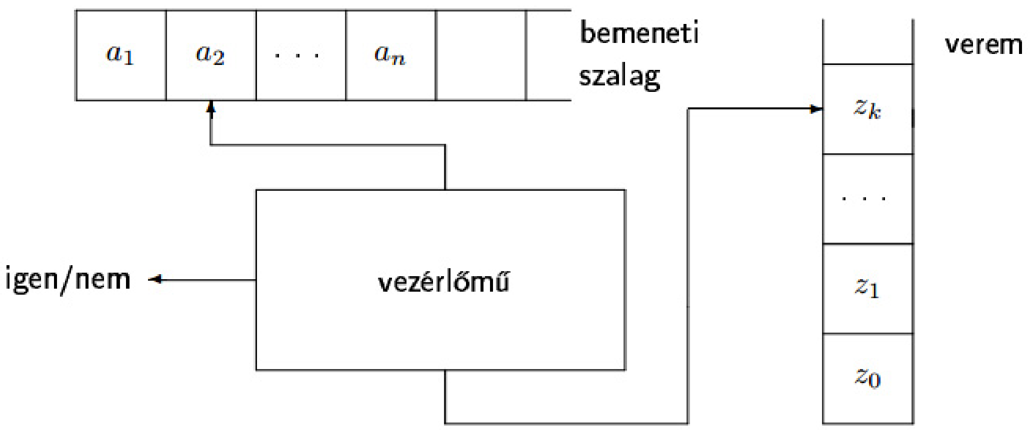
környezetfüggetlen nyelvtan: $G = (N, T, P, S)$,
amelynek szabályai $A \to \beta$ alakúak, $A \in N$, $\beta \in (N \cup T)^+$.
Megengedhető az $S \to \varepsilon$ szabály is, amennyiben S nem szerepel egyetlen szabály jobb oldalán sem.

Az $L(G) = \{u \in T^* \mid S \overset{*}{\underset{G}{\Longrightarrow}} u\}$ nyelv a G nyelvtan által generált környezetfüggetlen nyelv.

Veremautomaták

veremautomaták, amelyek környezetfüggetlen nyelveket ismernek fel

A veremautomaták lényegében abban különböznek a véges automatáktól, hogy egyrészt akkor is válthatnak állapotot, ha nem lépnek tovább a bemeneti szóban (üres jelet olvasnak), másrészt rendelkeznek egy veremmemóriával. A veremben az ún. veremábécé jeleit tároljuk.



A veremautomata egy szót kap bemenetként, a kezdőállapotból indul ki, és a veremben egy speciális szimbólum, a verem kezdőszimbóluma áll. A működése során az aktuális állapot, a következő bemeneti szimbólum (amely üres szó is lehet), és a verem tetején levő szimbólum ismeretében állapotot vált, és a verem tetején levő szimbólum helyére egy szót ír be (amely szintén lehet üres is).

Kétféle felismerési mód lehetséges. Végállapottal való felismerésről beszélünk, ha a veremautomata a bemeneti szó elolvasása után végállapotba kerül. Üres veremmel való felismerésről beszélünk, ha a bemeneti szó elolvasásának pillanatában a verem üres legyen. Meg fogjuk mutatni, hogy a két felismerési mód ekvivalens egymással.

Nemdeterminisztikus veremautomatának nevezzük a

$$\mathrm{V} = (Q, \Sigma, W, E, q_0, z_0, F)$$

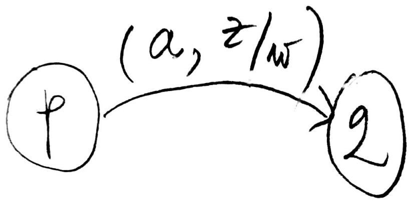
rendezett hetest, ahol

- Q az állapotok véges, nem üres halmaza,
- Σ a bemeneti ábécé,
- W a veremábécé,
- $E \subseteq Q \times \big(\Sigma \cup \{\varepsilon\}\big) \times W \times W^* \times Q$ az átmenetek vagy élek halmaza,
- $q_0 \in Q$ a kezdőállapot,
- $z_0 \in W$ a veremmemória kezdőjele,
- $F \subseteq Q$ a végállapotok halmaza.

Egy (p, a, z, w, q) átmenet azt jelenti, hogy ha a V veremautomata a p állapotban van, a bemeneti szalagról az a jelet olvassa (amely most üres szó is lehet), és a verem tetején levő szimbólum z, akkor q állapotba megy át és verembe a z helyére a w szót írja (betűnként).

A w szó beírása a verembe követi a természetes sorrendet, azaz w betűi balról jobbra, sorrendben kerülnek a verembe (azaz a szó utolsó betűje lesz a verem tetején).

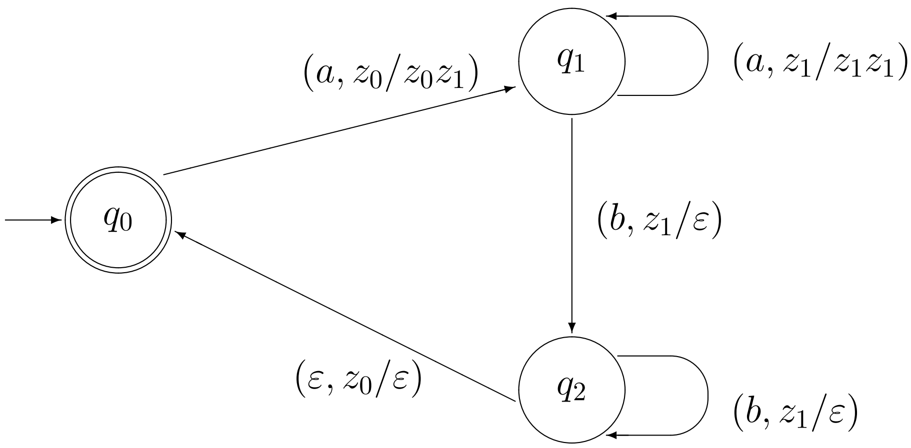
A könnyebb olvashatóság kedvéért a
(p, a, z, w, q) **átmenet helyett a** $\big(p, (a, z/w), q\big)$

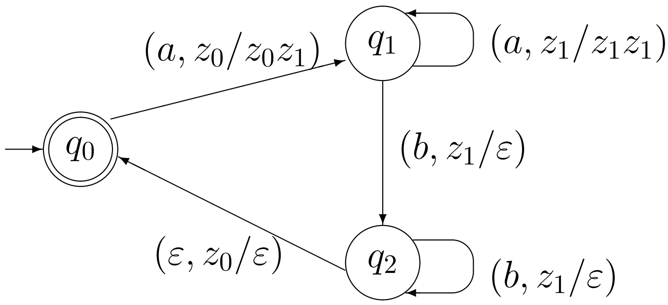


jelölést használjuk, amely utal arra, hogy az a **bemeneti jel elolvasása hatására a veremben kicseréljük** z**-t** w**-re.**

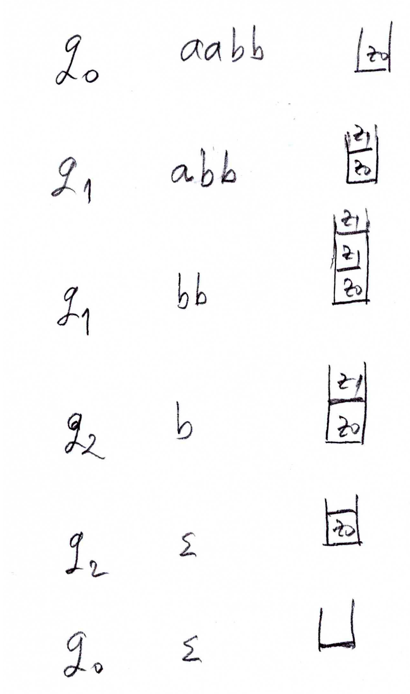
Példa.
$\mathrm{V}_1 = (\{q_0, q_1, q_2\}, \{a, b\}, \{z_0, z_1\}, E, q_0, z_0, \{q_0\})$. **Az** E **halmaz elemei:**

$$\begin{array}{ll} \big(q_0, (a, z_0/z_0z_1), q_1\big) & \\ \big(q_1, (a, z_1/z_1z_1), q_1\big) & \big(q_1, (b, z_1/\varepsilon), q_2\big) \\ \big(q_2, (b, z_1/\varepsilon), q_2\big) & \big(q_2, (\varepsilon, z_0/\varepsilon), q_0\big). \end{array}$$




q_0
q_1
q_2
$(a, z_0/z_0z_1)$
$(a, z_1/z_1z_1)$
$(b, z_1/\varepsilon)$
$(\varepsilon, z_0/\varepsilon)$
$(b, z_1/\varepsilon)$

működése


q_0 aabb z_0
q_1 abb z_1 z_0
q_1 bb z_1 z_1 z_0
q_2 b z_1 z_0
q_2 ε z_0
q_0 ε

átmenetfüggvény:

$$\delta \; : \; Q \times (\Sigma \cup \{\varepsilon\}) \times W \to \mathcal{P}(W^* \times Q),$$

Mivel a veremautomata nemdeterminisztikus, az átmenetfüggvény esetében

$\delta(q, a, z) = \{(w_1, p_1), \ldots, (w_k, p_k)\}$

***(a bemeneti szalagról olvas egy jelet, majd az olvasófej továbblép)*, vagy**

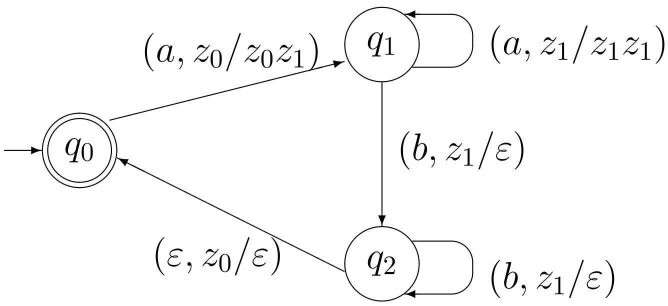
$\delta(q, \varepsilon, z) = \{(w_1, p_1), \ldots, (w_k, p_k)\}$

(ha nem mozdul el az olvasófej a bemeneti szalagon).

A veremautomata determinisztikus, ha $\forall q \in Q$ és $\forall z \in W$ esetében

- $|\delta(q, a, z)| \leq 1$, $\forall a \in \Sigma \cup \{\varepsilon\}$
- **Ha** $\delta(q, \varepsilon, z) \neq \emptyset$, **akkor** $\delta(q, a, z) = \emptyset$, $\forall a \in \Sigma$.

Az átmenetgráf



Az átmenetfüggvény:

$$\begin{array}{ll} \delta(q_0, a, z_0) = \{(z_0z_1, q_1)\} & \\ \delta(q_1, a, z_1) = \{(z_1z_1, q_1)\} & \delta(q_1, b, z_1) = \{(\varepsilon, q_2)\} \\ \delta(q_2, b, z_1) = \{(\varepsilon, q_2)\} & \delta(q_2, \varepsilon, z_0) = \{(\varepsilon, q_0)\}. \end{array}$$

Az átmenettáblázat:

$\Sigma \cup \{\varepsilon\} \rightarrow$	a		b		ε	
$W \downarrow$	z_0	z_1	z_0	z_1	z_0	z_1
q_0	$(z_0 z_1, q_1)$					
q_1		$(z_1 z_1, q_1)$		(ε, q_2)		
q_2				(ε, q_2)	(ε, q_0)	

Az egyszerűség kedvéért, a következőkben az üres oszlopokat kihagyjuk.

$\Sigma \cup \{\varepsilon\}$	a		b	ε
W	z_0	z_1	z_1	z_0
q_0	$(z_0 z_1, q_1)$			
q_1		$(z_1 z_1, q_1)$	(ε, q_2)	
q_2			(ε, q_2)	(ε, q_0)

Az aktuális állapot, a bemeneti szó még el nem olvasott része és a verem tartalma együtt képezik a veremautomata egy konfigurációját, vagyis minden $q \in Q$, $u \in \Sigma^*$, és $v \in W^*$ esetében (q, u, v) egy konfiguráció.

Legyen $u = a_1 a_2 \dots a_k$ **és** $v = x_1 x_2 \dots x_m$.

állapot q

bemeneti szalag

a_{-p}	a_{-p+1}	a_{-p+2}	$\cdots$	a_0	a_1	a_2	$\cdots$	a_k	$\cdots$
					$\uparrow$				

verem

x_1	x_2	$\cdots$	x_m
			$\uparrow$

Ekkor a veremautomata kétféleképpen léphet (azaz konfigurációt válthat):

- $(q,\ a_1 a_2 \dots a_k,\ x_1 x_2 \dots x_{m-1} x_m) \Longrightarrow (p,\ a_2 a_3 \dots a_k,\ x_1, x_2 \dots x_{m-1} w)$,
  ha $\big(q, (a_1, x_m/w), p\big) \in E$
- $(q,\ a_1 a_2 \dots a_k,\ x_1 x_2 \dots x_m) \Longrightarrow (p,\ a_1 a_2 \dots a_k,\ x_1, x_2 \dots x_{m-1} w)$,
  ha $\big(q, (\varepsilon, x_m/w), p\big) \in E$.

$\Longrightarrow$ **reláció reflexív, tranzitív lezártja** $\stackrel{*}{\Longrightarrow}$ **(**$\vdash$ **jel is)**

Az automata működése: elindulunk a $(q_0, b_1 b_2 \dots b_n, z_0)$ **kezdeti konfigurációból, majd meghatározzuk az összes lehetséges következő konfigurációt, majd ezekre a rákövetkezőket, és így tovább, ameddig lehet.**

Azt mondjuk, hogy a V veremautomata végállapottal felismer egy u szót, ha van V-beli konfigurációknak olyan sorozata, amelyre teljesülnek a következők:
- **a sorozat első eleme (q_0, u, z_0),**
- **a sorozat minden eleméből van átmenet a sorozat következő elemébe, kivéve ha csak egy elemből áll,**
- **a sorozat utolsó eleme (p, ε, w), ahol $p \in F$ és $w \in W^*$.**

$(q_0, u, z_0) \stackrel{*}{\Longrightarrow} (p, \varepsilon, w)$, **valamely $w \in W^*$-ra és $p \in F$-re.**

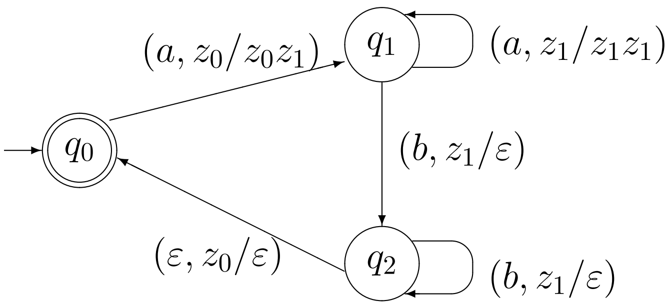
végállapottal felismert nyelv: $L(\mathrm{V})$

Azt mondjuk, hogy a V veremautomata üres veremmel felismer egy u szót, ha van V-beli konfigurációknak olyan sorozata, amelyre teljesülnek a következők:
- **a sorozat első eleme (q_0, u, z_0),**
- **a sorozat minden eleméből van átmenet a sorozat következő elemébe,**
- **a sorozat utolsó eleme $(p, \varepsilon, \varepsilon)$, és p tetszőleges állapot.**

$(q_0, u, z_0) \stackrel{*}{\Longrightarrow} (p, \varepsilon, \varepsilon)$ **valamely $p \in Q$-ra**

üres veremmel felismert nyelv: $L_\varepsilon(\mathrm{V})$

Példa. Ha a V_1 automatát megvizsgáljuk, észrevehetjük, hogy az $\{a^n b^n \mid n \geq 0\}$ nyelvet ismeri fel végállapottal.



Végezzük el a levezetést a következő szavakra: $aaabbb$ és $abab$.

Az a^3b^3 szót felismeri az automata, mivel:
$(q_0, aaabbb, z_0) \Longrightarrow (q_1, aabbb, z_0z_1) \Longrightarrow (q_1, abbb, z_0z_1z_1) \Longrightarrow (q_1, bbb, z_0z_1z_1z_1)$
$\Longrightarrow (q_2, bb, z_0z_1z_1) \Longrightarrow (q_2, b, z_0z_1) \Longrightarrow (q_2, \varepsilon, z_0) \Longrightarrow (q_0, \varepsilon, \varepsilon)$, és mivel q_0 végállapot, a veremautomata felismeri a szót. Ugyanakkor, mivel a verem kiürült, üres veremmel is felismeri.

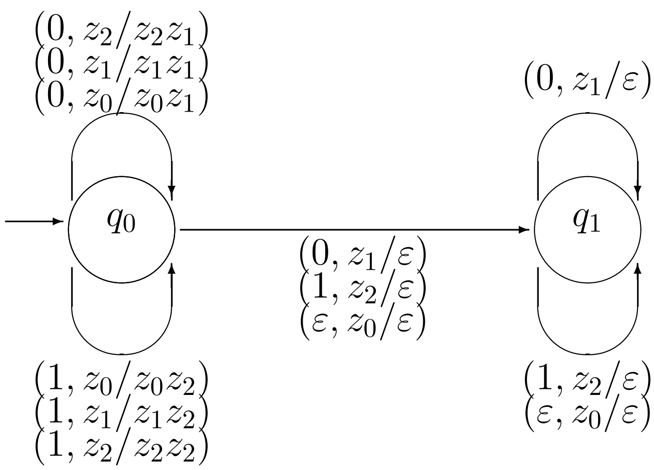
Mivel a kezdőállapot egyben végállapot is, az üres szót is felismeri végállapottal, ellenben üres veremmel nem.

Hogy bebizonyítsuk, hogy az $abab$ szót nem ismeri fel, szükségünk van az összes lehetőség megvizsgálására. Könnyű belátni, hogy ebben az esetben csak egyetlen lehetőség van:

$(q_0, abab, z_0) \Longrightarrow (q_1, bab, z_0z_1) \Longrightarrow (q_2, ab, z_0) \Longrightarrow (q_0, ab, \varepsilon)$, de innen nincs átmenet, tehát nem ismeri fel az $abab$ szót.

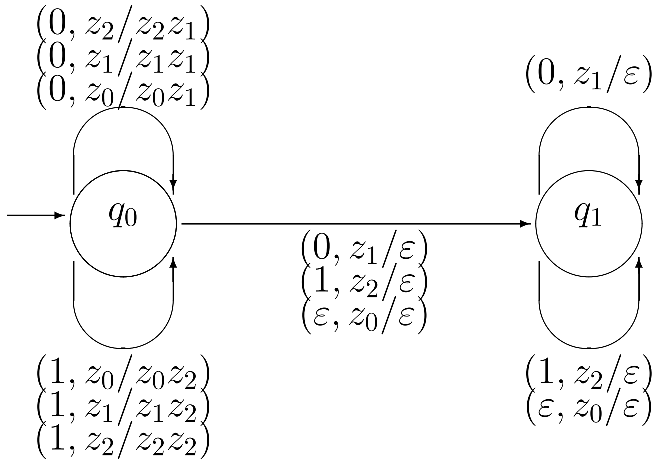
Példa.
A $\mathrm{V}_2 = (\{q_0, q_1\}, \{0, 1\}, \{z_0, z_1, z_2\}, E, q_0, z_0, \emptyset)$ **veremautomata átmenettáblázata:**

	0			1			ε
	z_0	z_1	z_2	z_0	z_1	z_2	z_0
q_0	(z_0z_1, q_0)	(z_1z_1, q_0) (ε, q_1)	(z_2z_1, q_0)	(z_0z_2, q_0)	(z_1z_2, q_0)	(z_2z_2, q_0) (ε, q_1)	(ε, q_1)
q_1		(ε, q_1)				(ε, q_1)	(ε, q_1)

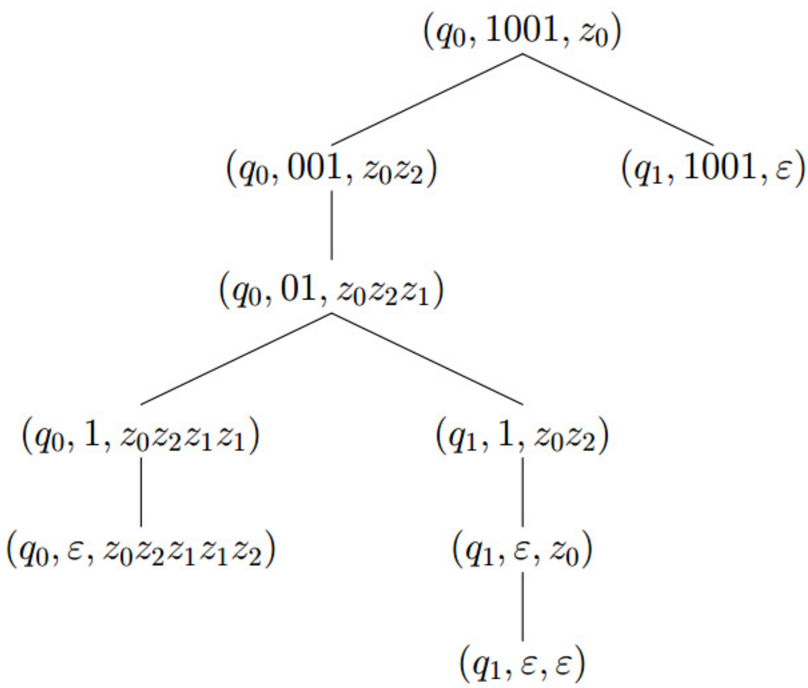


A V_2 **veremautomata az** $\left\{uu^{-1} \mid u \in \{0, 1\}^*\right\}$ **nyelvet ismeri fel.**

Mivel V_2 **nemdeterminisztikus, a** (q_0, u, z_0) **kezdőkonfigurációból elérhető összes konfigurációt egy ún.** **számítási fában** **tudjuk ábrázolni.**

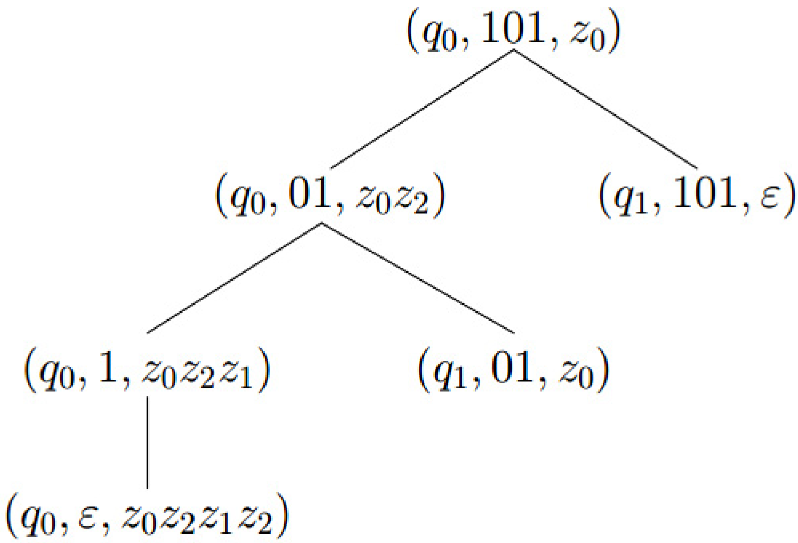


Például a $(q_0, 1001, z_0)$ **kezdőkonfigurációhoz tartozó számítási fa:**



$(q_1, \varepsilon, \varepsilon)$ **a fa egyik levele, a** V_2 **veremautomata üres veremmel felismeri az 1001 szót.**

A következő ábrán látható számítási fa annak bizonyítéka, hogy a V_2 veremautomata nem ismeri fel az 101 szót. A levelekben lévő konfigurációkat nem lehet folytatni, és egyik sem $(q, \varepsilon, \varepsilon)$ alakú.

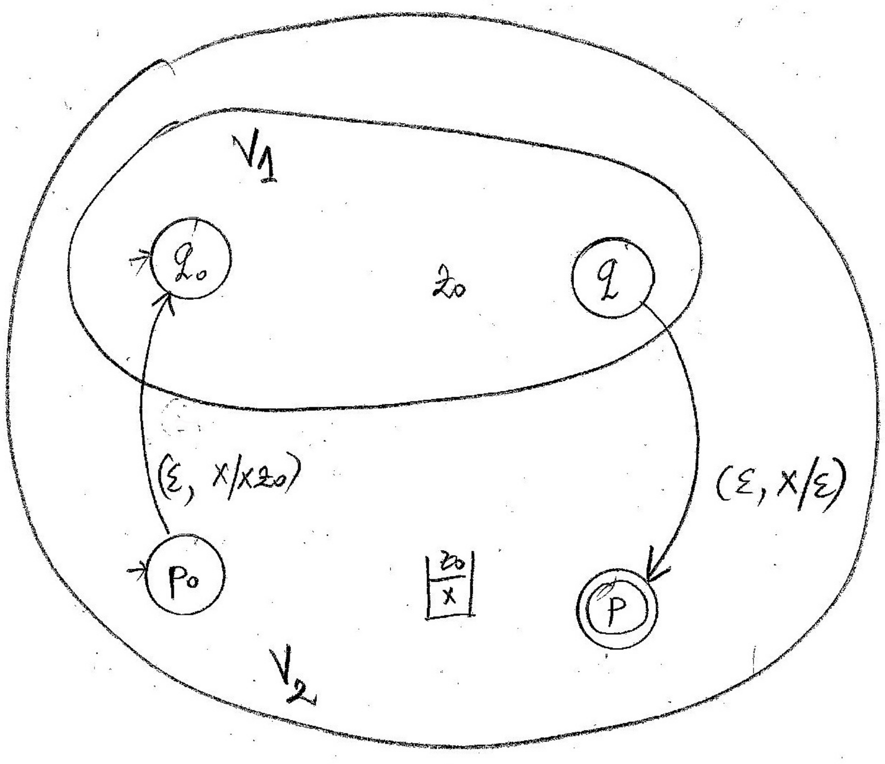


Tétel. **Egy L nyelv akkor és csakis akkor ismerhető fel valamely V_1 nemdeterminisztikus veremautomatával üres veremmel, ha felismerhető valamely V_2 nemdeterminisztikus veremautomatával végállapottal.**

a) **Legyen** $\mathrm{V}_1 = (Q, \Sigma, W, E, q_0, z_0, \emptyset)$ **veremautomata, amely üres veremmel ismeri fel az** L **nyelvet. Definiáljuk a** $\mathrm{V}_2 = (Q \cup \{p_0, p\}, \Sigma, W \cup \{x\}, E', p_0, x, \{p\})$ **veremautomatát, ahol** $p, p_0 \notin Q,\ , x \notin W$ **és**

$$E' = E \cup \Big\{ \big(p_0, (\varepsilon, x/xz_0), q_0\big) \Big\} \cup \Big\{ \big(q, (\varepsilon, x/\varepsilon), p\big) \mid q \in Q \Big\}.$$

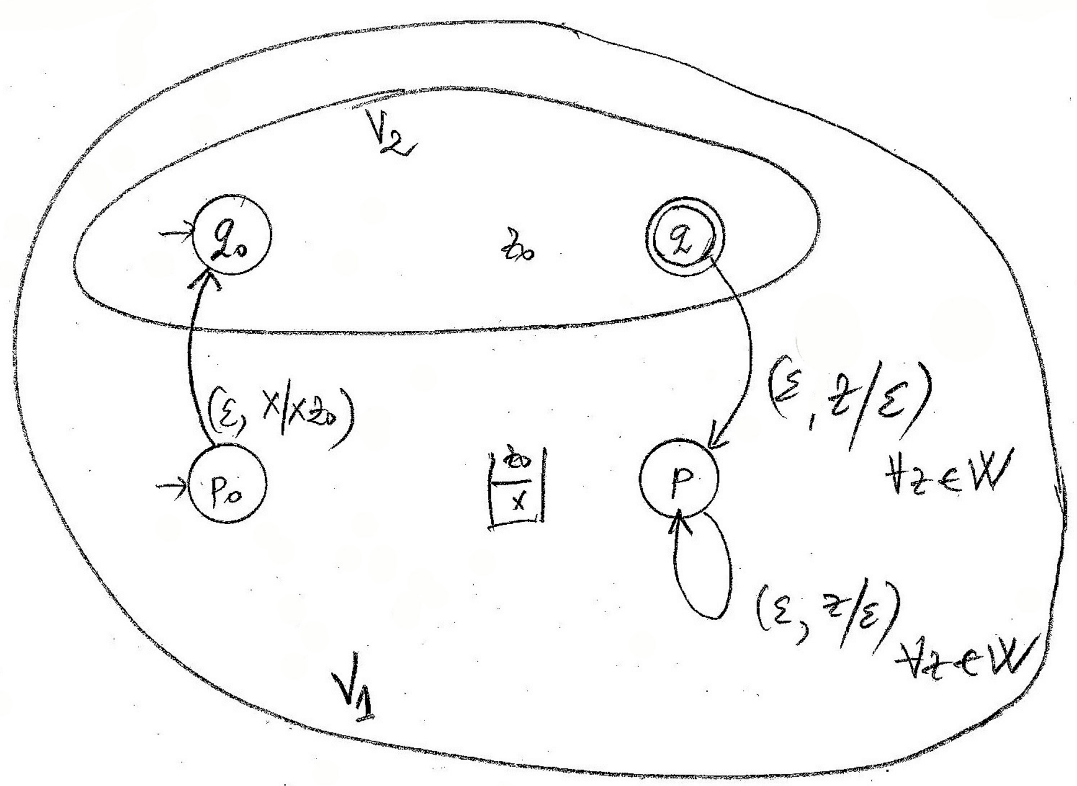
V_2 **működése:** V_2 **először egy** ε**-lépéssel átmegy a** V_1 **kezdőállapotába, beírva a verembe** x **mellé** z_0**-t,** V_1 **kezdőszimbólumát. Ettől kezdve úgy működik, mint** V_1**. Ha** V_1 **egy adott szóra kiüríti a saját vermét, akkor** V_2**-nél még mindig marad egy** x **a veremben, amelyet** V_2 ε**-lépéssel töröl és végállapotba kerül.** V_2 **csak akkor kerülhet végállapotba, ha** V_1 **kiürítette a vermét.**



b) Legyen $\mathrm{V}_2 = (Q, \Sigma, W, E, q_0, z_0, F)$ **egy veremautomata, amely az** L **nyelvet végállapottal ismeri fel. Definiáljuk a** $\mathrm{V}_1 = (Q \cup \{p_0, p\}, \Sigma, W \cup \{x\}, E', p_0, x, \emptyset)$ **veremautomatát, ahol** $p_0, p \notin Q$**,** $x \notin W$ **és**

$$E' = E \cup \Big\{ \big(p_0, (\varepsilon, x/xz_0), q_0\big) \Big\} \cup \Big\{ \big(q, (\varepsilon, z/\varepsilon), p\big) \mid q \in F, p \in Q, z \in W \Big\} \cup \Big\{ \big(p, (\varepsilon, z/\varepsilon), p\big) \mid p \in Q, z \in W \cup \{x\} \Big\}$$

V_1 **működése:** V_1 **először** ε**-lépéssel beírja a verembe** x **mellé** z_0**-t,** V_2 **vermének kezdőszimbólumát, ettől kezdve mint** V_2 **működik, azaz végállapotba jut minden felismert szóra. Innen** V_1 ε**-lépéssel kiüríti a vermet.** V_1 **csak akkor ürítheti ki a vermet, ha** V_2 **végállapotba kerül.**



A következő két tétel azt bizonyítja, hogy a nemdeterminisztikus veremautomaták által felismert nyelvek halmaza éppen a környezetfüggetlen nyelvek halmaza.

Tétel. Ha G környezetfüggetlen nyelvtan, akkor létezik egy olyan V nemdeterminisztikus veremautomata, amely üres veremmel felismeri az $L(G)$ nyelvet, azaz $L_\varepsilon(\mathrm{V}) = L(G)$.

Csak a bizonyítás ötletét adjuk meg. Legyen $G = (N, T, P, S)$ egy környezetfüggetlen nyelvtan. Értelmezzük a $\mathrm{V} = (\{q\}, T, N \cup T, E, q, S, \emptyset)$ veremautomatát, ahol $q \notin N \cup T$, az E átmenethalmaz értelmezése pedig:

- **Ha létezik a G nyelvtan szabályai között $A \to \alpha$ szabály, akkor vegyük be E-be a $\big(q, (\varepsilon, A/\alpha^{-1}), q\big)$ átmenetet,**
- **Minden $a \in T$ jelre vegyük be E-be a $\big(q, (a, a/\varepsilon), q\big)$ átmenetet.**

Ha van $S \to \alpha$ szabály G-ben, a veremautomata egy ε-lépéssel beírja a verembe az α tükörképét. Ha a beolvasott betű egyezik a verem tetején lévővel, akkor törli azt a veremből. Ha a verem tetején az A nemterminális betű van, akkor beviszi a verembe valamelyik A-val kezdődő szabály jobb oldalának a tükörképét. Ha a szó beolvasása végén a verem kiürül, a veremautomata felismerte a szót.

A következő algoritmus egy $G = (N, T, P, S)$ környezetfüggetlen nyelvtanhoz megkonstruálja azt a $\mathrm{V} = (\{q\}, T, N \cup T, E, q, S, \emptyset)$ veremautomatát, amelyik üres veremmel felismeri a G által generált nyelvet.

Környezetfüggetlen-nyelvtanból-veremautomata(G)

```
for minden A → α szabályra
    do vegyük be E-be a (q, (ε, A/α^{-1}), q) átmenetet
for minden a ∈ T terminálisra
    do vegyük be E-be a (q, (a, a/ε), q) átmenetet
return V
```

Ha a G nyelvtan szabályainak száma n, a terminális betűké pedig m, akkor az algoritmus lépésszáma $\Theta(n+m)$.

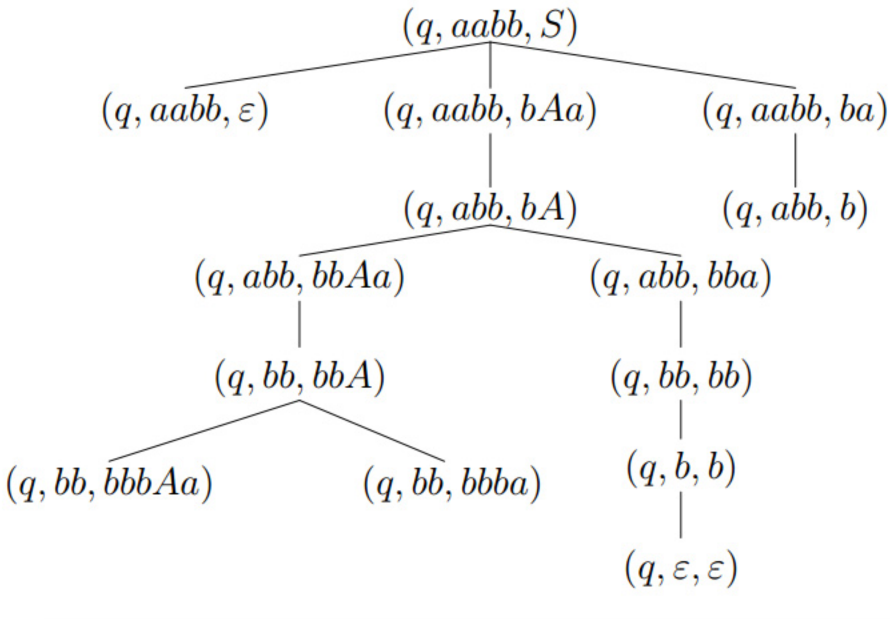
Példa. Legyen $G = (\{S, A\}, \{a, b\}, \{S \to \varepsilon,\ S \to ab,\ S \to aAb,\ A \to aAb,\ A \to ab\}, S)$. **Ekkor** $\mathrm{V} = (\{q\}, \{a, b\}, \{a, b, A, S\}, E, q, S, \emptyset)$, **a következő átmenettáblázattal.**

$\Sigma \cup \{\varepsilon\}$	a	b	ε	
W	a	b	S	A
q	(ε, q)	(ε, q)	(ε, q) (ba, q) (bAa, q)	(bAa, q) (ba, q)

Nézzük meg, hogyan ismeri fel a V **veremautomata az** $aabb$ **szót, amelyet a** G **nyelvtanban a következőképpen lehet levezetni.**

$$S \Longrightarrow aAb \Longrightarrow aabb,$$

ahol alkalmaztuk az $S \to aAb$ **és** $A \to ab$ **szabályokat.**



Tétel. Ha V nemdeterminisztikus veremautomata, akkor létezik egy olyan G környezetfüggetlen nyelvtan, hogy V üres veremmel felismeri az $L(G)$ nyelvet, azaz $L_\varepsilon(\mathrm{V}) = L(G)$.

Bizonyítás helyett megadjuk, hogyan kell definiálni a G nyelvtant. Legyen a nemdeterminisztikus veremautomata $\mathrm{V} = (Q, \Sigma, W, E, q_0, z_0, \emptyset)$.

Ekkor $G = (N, T, P, S)$, ahol
$N = \{S\} \cup \{S_{p,z,q} \mid p, q \in Q, z \in W\}$ és $T = \Sigma$.

A P szabályait pedig a következőképpen kapjuk meg.

• Minden q állapotra vegyük be P-be az $S \to S_{q_0,z_0,q}$ szabályt.
• Ha $\big(q, (a, z/z_k \dots z_2 z_1), p\big) \in E$, ahol $q \in Q$, $z, z_1, z_2, \dots z_k \in W$ ($k \geq 1$) és $a \in \Sigma \cup \{\varepsilon\}$, vegyük be P-be minden lehetséges $p_1, p_2, \dots, p_k$ állapotra az
$S_{q,z,p_k} \to aS_{p,z_1,p_1} S_{p_1,z_2,p_2} \dots S_{p_{k-1},z_k,p_k}$ szabályokat.
• Ha $\big(q, (a, z/\varepsilon), p\big) \in E$, ahol $p, q \in Q$, $z \in W$, és $a \in \Sigma \cup \{\varepsilon\}$, vegyük be P-be az $S_{q,z,p} \to a$ szabályt.

Az így értelmezett nyelvtan kiterjesztett környezetfüggetlen nyelvtan, amelyhez, mint tudjuk, mindig hozzárendelhető egy vele ekvivalens környezetfüggetlen nyelvtan.

A következő algoritmus egy $\mathrm{V} = (Q, \Sigma, W, E, q_0, z_0, \emptyset)$ veremautomatához megkonstruálja azt a $G = (N, T, P, S)$ környezetfüggetlen nyelvtant, amely a V veremautomata által üres veremmel felismert környezetfüggetlen nyelvet generálja.

Veremautomatából-környezetfüggetlen-nyelvtan(V)

for minden $q \in Q$
    do vegyük be P-be az $S \to S_{q_0,z_0,q}$ szabályt
for minden $\big(q,(a,z/z_k \dots z_2 z_1),p\big) \in E$
        $\triangleright\ q \in Q,\ \ z, z_1, z_2, \dots z_k \in W\ (k \geq 1),\ a \in \Sigma \cup \{\varepsilon\}$
    do for minden $p_1, p_2, \dots, p_k$ állapotra
        do vegyük be P-be az $S_{q,z,p_k} \to aS_{p,z_1,p_1}S_{p_1,z_2,p_2} \dots S_{p_{k-1},z_k,p_k}$ szabályokat
for minden $\big(q(a,z/\varepsilon),p\big) \in E$ $\triangleright\ p,q \in Q, z \in W,\ a \in \Sigma \cup \{\varepsilon\}$
    do vegyük be P-be az $S_{q,z,p} \to a$ szabályt
return G

Ha az automata állapotainak száma n, átmeneteinek száma pedig m, akkor a fenti algoritmus legfeljebb $n + mn + m$ lépést hajt végre, tehát a lépésszáma legrosszabb esetben $O(nm)$.

Végül bizonyítás nélkül megemlítjük, hogy a determinisztikus veremautomatákkal felismerhető nyelvek osztálya valódi része a nemdeterminisztikus veremautomatákkal felismerhető nyelvek osztályának. Ebből a szempontból a veremautomaták a véges automatáktól eltérő módon viselkednek.

Példa. $\mathrm{V} = (\{q\}, \{a, b\}, \{a, b, A, S\}, E, q, S, \emptyset)$.

$\Sigma \cup \{\varepsilon\}$	a	b	ε	
W	a	b	S	A
q	(ε, q)	(ε, q)	(ε, q) (ba, q) (bAa, q)	(bAa, q) (ba, q)

A G nyelvtan a következő:

$$G = (\{S, S_a, S_b, S_S, S_A, \}, \{a, b\}, P, S),$$

ahol minden $z \in \{a, b, S, A\}$ esetén S_z az $S_{q,z,q}$ jelölés rövidítése. Felírjuk a veremautomata átmeneteit:

$\big(q, (a, a/\varepsilon), q\big)$, $\big(q, (b, b/\varepsilon), q\big)$,
$\big(q, (\varepsilon, S/\varepsilon), q\big)$, $\big(q, (\varepsilon, S/ba), q\big)$, $\big(q, (\varepsilon, S/bAa), q\big)$,
$\big(q, (\varepsilon, A/ba), q\big)$, $\big(q, (\varepsilon, A/bAa), q\big)$.

Ezek alapján a következő szabályokat definiálhatjuk:

$S \rightarrow S_S$
$S_a \rightarrow a$
$S_b \rightarrow b$
$S_S \rightarrow \varepsilon \mid S_aS_b \mid S_aS_AS_b$
$S_A \rightarrow S_aS_AS_b \mid S_aS_b$. S_S **elhagyható**

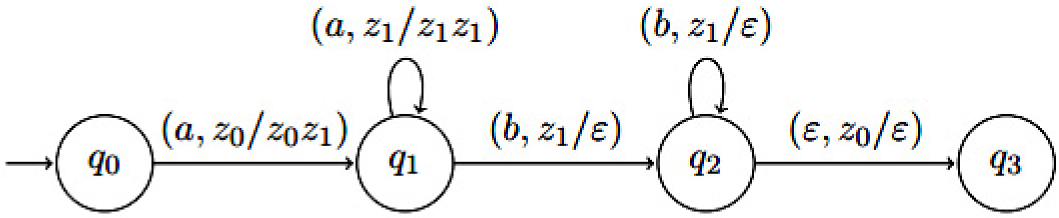
$S \rightarrow \varepsilon \mid S_aS_b \mid S_aS_AS_b,$
$S_A \rightarrow S_aS_AS_b \mid S_aS_b,$
$S_a \rightarrow a$, $S_b \rightarrow b$,

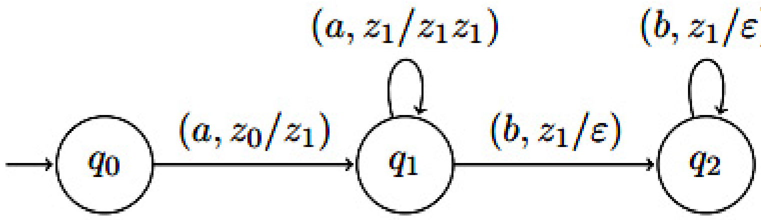
ezek a szabályok pedig helyettesíthetők a következőkkel:

$S \rightarrow \varepsilon \mid ab \mid aAb,$
$A \rightarrow aAb \mid ab$.

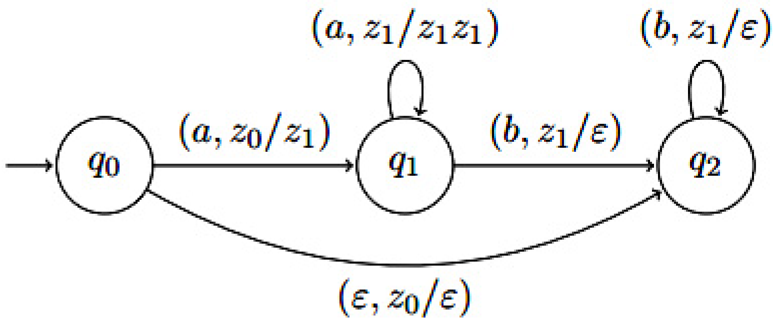
Példák.

$L_1 = \{a^n b^n \mid n \geq 1\}$

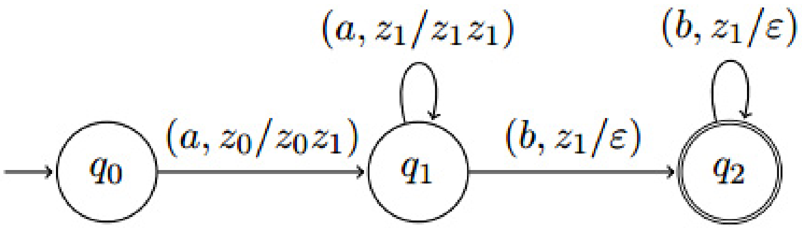




$L_2 = \{a^n b^n \mid n \geq 0\}$



$L_3 = \{a^n b^m \mid n \geq m \geq 1\}$



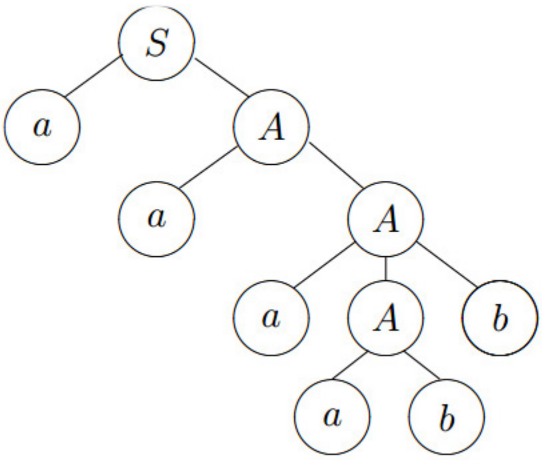
Környezetfüggetlen nyelvek

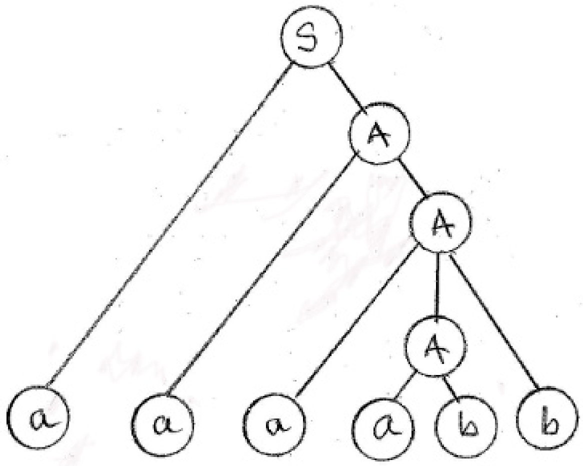
Vegyünk egy $G = (N, T, P, S)$ környezetfüggetlen nyelvtant. G egy levezetési fájának egy olyan véges, rendezett, címkézett fát nevezünk, amelynek gyökere az S kezdőszimbólummal van címkézve, minden belső csúcs címkéje egy nemterminális szimbólum, és levelei terminális szimbólumokkal vannak címkézve. Teljesül továbbá az a feltétel, hogy ha egy belső csúcs címkéje az A nemterminális, és a csúcsnak k közvetlen leszármazottja van, akkor P-ben van egy olyan $A \to a_1a_2 \ldots a_k$ szabály, hogy ezek a közvetlen leszármazottak rendre az a_1, a_2, $\ldots a_k$ szimbólumokkal vannak címkézve. A levezetési fa eredménye az a T feletti szó, amelyet úgy kapunk, hogy a fa leveleinek címkéjét balról jobbra haladva összeolvassuk. A levezetési fát még szintaxisfának is nevezzük.

Tekintsük a $G = (\{S, A\}, \{a, b\}, \{S \to aA,\ S \to a,\ S \to \varepsilon,\ A \to aA,\ A \to aAb,\ A \to ab,\ A \to b\}, S)$ környezetfüggetlen nyelvtant. Ez a nyelvtan az $L(G) = \{a^n b^m \mid n \geq m \geq 0\}$ nyelvet generálja. Az $a^4b^2 \in L(G)$ szó levezetése a következő:

$$S \Longrightarrow aA \Longrightarrow aaA \Longrightarrow aaaAb \Longrightarrow aaaabb.$$

Ezt a levezetést ábrázolhatjuk levezetési (szintaxis-) fával, amelynek eredménye az $aaaabb$ szó.





Minden levezetéshez hozzárendelhető egy levezetési fa.

Az $\alpha_0 \Longrightarrow \alpha_1 \Longrightarrow \ldots \Longrightarrow \alpha_n$ **levezetést legbaloldalibb levezetésnek nevezzük, ha minden** $i = 1, 2, \ldots, n-1$ **esetén van olyan** $u_i \in T^*$, $\beta_i \in (N \cup T)^*$ **szó és** $(A_i \to \gamma_i) \in P$ **szabály, hogy teljesülnek az**

$$\alpha_i = u_i A_i \beta_i \quad \text{és} \quad \alpha_{i+1} = u_i \gamma_i \beta_i$$

összefüggések.

Tekintsük a következő nyelvtant:
$G = (\{S, A\}, \{a, b, c\}, \{S \to bA,\ S \to bAS,\ S \to a,\ A \to cS,\ A \to a\}, S)$.

A $bcbaa$ **szónak két, egymástól különböző legbaloldalibb levezetése van:**
$S \Longrightarrow bA \Longrightarrow bcS \Longrightarrow bcbAS \Longrightarrow bcbaS \Longrightarrow bcbaa$,
$S \Longrightarrow bAS \Longrightarrow bcSS \Longrightarrow bcbAS \Longrightarrow bcbaS \Longrightarrow bcbaa$.

Egy G **környezetfüggetlen nyelvtan nem egyértelmű, ha** $L(G)$**-ben van olyan szó, amelynek egynél több legbaloldalibb levezetése van. Különben a** G **nyelvtan egyértelmű.**

Az előbbi G nyelvtan nem egyértelmű, mert a $bcbaa$ szónak találtunk két legbaloldalibb levezetését. Egy nyelvet több nyelvtan is generálhat, és ezek között lehetnek egyértelműek és nem egyértelműek is. Egy környezetfüggetlen nyelv alapvetően nem egyértelmű, ha nem létezik egyetlen egyértelmű nyelvtan sem, amely generálja.

Példa. Vizsgáljuk meg a következő két nyelvtant.

A $G_1 = (\{S\}, \{a, +, *\}, \{S \to S + S,\ S \to S * S,\ S \to a\}, S)$ **nem egyértelmű, mert**

$$S \Longrightarrow S + S \Longrightarrow a + S \Longrightarrow a + S * S \Longrightarrow a + a * S \Longrightarrow a + a * S + S$$
$$\Longrightarrow a + a * a + S \Longrightarrow a + a * a + a$$

és

$$S \Longrightarrow S * S \Longrightarrow S + S * S \Longrightarrow a + S * S \Longrightarrow a + a * S \Longrightarrow a + a * S + S$$
$$\Longrightarrow a + a * a + S \Longrightarrow a + a * a + a.$$

A $G_2 = (\{S, A\}, \{a, *, +\}, \{S \to A + S \mid A,\ A \to A * A \mid a\}, S)$ **nyelvtan egyértelmű.**

$$S \Longrightarrow A + S \Longrightarrow a + S \Longrightarrow a + A + S \Longrightarrow a + A * A + S \Longrightarrow a + a * A + S$$
$$\Longrightarrow a + a * a + S \Longrightarrow a + a * a + A \Longrightarrow a + a * a + a$$

Be lehet bizonyítani, hogy $L(G_1) = L(G_2)$.