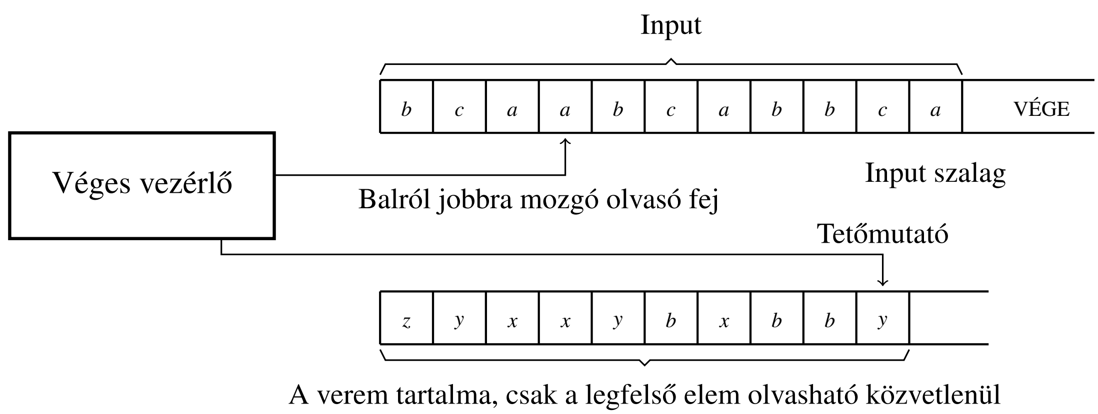
Veremautomaták

A) Elméleti háttér



Jelölés: ha X egy halmaz, jelölje $\mathcal{P}_{\text{véges}}(X)$ az X véges részhalmazainak halmazát.

Definíció A **veremautomata** egy $A = \langle Z, Q, T, \delta, z_0, q_0, F\rangle$, rendezett hetes, ahol

- Z a veremszimbólumok véges halmaza (veremábécé),
- Q az állapotok véges halmaza,
- T az inputszimbólumok véges halmaza (inputábécé),
- $\delta : Z \times Q \times (T \cup \{\varepsilon\}) \to \mathcal{P}_{\text{véges}}(Z^* \times Q)$, az ún. átmeneti függvény,
- $z_0 \in Z$ a kezdeti (kezdő) veremszimbólum,
- $q_0 \in Q$ a kezdeti állapot (kezdőállapot),
- $F \subseteq Q$ az elfogadó állapotok vagy végállapotok halmaza.

Megjegyzések:

- alapértelmezetten nemdeterminisztikus
- Ha $\delta(z, q, \varepsilon)$ nem üres, akkor ún. **ε-átmenet** (ε-lépés, ε-mozgás) hajtható végre, ami lehetővé teszi, hogy a veremautomata anélkül változtassa meg az állapotát, hogy valamilyen szimbólumot olvasson az inputszalagról.
- ε-mozgásra lehetőség van már az első inputszimbólum elolvasása előtt is illetve még az utolsó inputszimbólum elolvasása után is.

Definíció A veremautomata **konfigurációja** alatt egy zqw alakú szót értünk, ahol $z \in Z^*$ a verem aktuális tartalma és $q \in Q$ az aktuális állapot és $w \in T^*$ az input még feldolgozatlan része.

- z első betűje van a verem alján, míg utolsó betűje a verem tetején.
- Az input olvasófeje w első betűjén áll.
- Így a q baloldalán lévő szimbólum van a verem tetején, míg a jobboldalán lévő szimbólum az input következő feldolgozandó betűje.

Definíció Az $A = \langle Z, Q, T, \delta, z_0, q_0, F \rangle$ veremautomata $w \in T^*$ bemenethez tartozó **kezdőkonfigurációja** $z_0 q_0 w$.

Alapvető veremműveletek megvalósítása
Legyen $t \in T \cup \{\varepsilon\}$, $q, r \in Q$ és $z \in Z$

- $(\varepsilon, r) \in \delta(z, q, t)$: a z elemet kivehetjük a veremből (POP művelet)
- $(z, r) \in \delta(z, q, t)$: a verem tartalma változatlan maradhat
- $(z', r) \in \delta(z, q, t)$: z-t lecserélhetjük z'-re a verem tetején ($z' \in Z$)
- $(zz', r) \in \delta(z, q, t)$: z'-t a verem tetejére (z-re rá) tehetjük ($z' \in Z$) (PUSH művelet)
- Egyéb lehetőségek, például $(zz'z'', r) \in \delta(z, q, t)$: $z'z''$-t a verem tetejére tehetjük, z'' lesz a tetején ($z', z'' \in Z$) .
- Általánosan $(w, r) \in \delta(z, q, t)$, ahol $w \in Z^*$ tetszőleges Z feletti szó. A w szó kerül z helyére és w utolsó betűje lesz a verem tetején.

Definíció Az A veremautomata az $\alpha \in Z^*QT^*$ konfigurációt a $\beta \in Z^*QT^*$ konfigurációra **redukálja egy lépésben**, amelyet $\alpha \Rightarrow_A \beta$-val jelölünk, ha létezik olyan $z \in Z, q, p \in Q, a \in T \cup \{\varepsilon\}$, $r, u \in Z^*$ és $w \in T^*$, hogy $(u, p) \in \delta(z, q, a)$ és $\alpha = rzqaw$ és $\beta = rupw$ teljesül.

Példák:

- ha A-ban $\delta(c, q_1, a) = \{(dd, q_2), (\varepsilon, q_4)\}$ és $z_0cddcq_1ababba$ egy konfiguráció, akkor $z_0cddcq_1ababba \Rightarrow_A z_0cddddq_2babba$ és $z_0cddcq_1ababba \Rightarrow_A z_0cddq_4babba$ is teljesül,
- ha A-ban $\delta(c, q_3, \varepsilon) = \{(dd, q_2)\}$ és $z_0cddcq_3ababba$ egy konfiguráció, akkor $z_0cddcq_3ababba \Rightarrow_A z_0cddddq_2ababba$
- ha A-ban $\delta(c, q_5, \varepsilon) = \emptyset$ és $\delta(c, q_5, a) = \emptyset$, akkor nem létezik olyan C konfiguráció, melyre $z_0ccq_5aab \Rightarrow_A C$

Definíció: Az A veremautomata az $\alpha \in Z^*QT^*$ konfigurációt a $\beta \in Z^*QT^*$ konfigurációra **redukálja**, amelyet $\alpha \Rightarrow_A^* \beta$-val jelölünk, ha vagy $\alpha = \beta$, vagy létezik olyan $\alpha_1, \ldots, \alpha_n$ szavakból álló véges sorozat, ahol $\alpha = \alpha_1, \beta = \alpha_n$ és $\alpha_i \Rightarrow_A \alpha_{i+1}$, $1 \le i \le n-1$.

Tehát $\Rightarrow_A^* \subseteq Z^*QT^* \times Z^*QT^*$ a $\Rightarrow_A$ reláció reflexív, tranzitív lezártja.

Példa:
Ha $\delta(d, q_6, b) = \{(\varepsilon, q_5)\}$ és $\delta(d, q_5, \varepsilon) = \{(dd, q_2), (\varepsilon, q_4)\}$ akkor
$\#cddq_6bab \Rightarrow_A \#cdq_5ab \Rightarrow_A \#cddq_2ab$ és
$\#cddq_6bab \Rightarrow_A \#cdq_5ab \Rightarrow_A \#cq_4ab$.
Tehát $\#cddq_6bab \Rightarrow_A^* \#cddq_2ab$ és $\#cddq_6bab \Rightarrow_A^* \#cq_4ab$.

Definíció Az A veremautomata által **elfogadó állapottal (végállapottal) elfogadott nyelv** $L(A) = \{w \in T^* \mid z_0q_0w \Rightarrow_A^* up$, ahol $u \in Z^*, p \in F\}$.

Definíció Az $A = \langle Z, Q, T, \delta, z_0, q_0, F\rangle$ veremautomatát **determinisztikusnak** nevezzük, ha minden $(z, q, a) \in Z \times Q \times T$ esetén $|\delta(z, q, a)| + |\delta(z, q, \varepsilon)| = 1$.

Tehát minden $q \in Q$ és $z \in Z$ esetén

- vagy $\delta(z, q, a)$ pontosan egy elemet tartalmaz minden $a \in T$ inputszimbólumra és $\delta(z, q, \varepsilon) = \emptyset$,
- vagy $\delta(z, q, \varepsilon)$ pontosan egy elemet tartalmaz és $\delta(z, q, a) = \emptyset$ minden $a \in T$ inputszimbólumra.

Tétel A determinisztikus veremautomaták számítási ereje kisebb, mint a (nemdeterminisztikus) veremautomatáké, de nagyobb a véges automatáknál, azaz van olyan nyelv, amelyik felismerhető veremautomatával, de nem ismerhető fel determinisztikus veremautomatával.

Definíció Az A veremautomata által **üres veremmel elfogadott nyelv** $N(A) = \{w \in T^* \mid z_0 q_0 w \Rightarrow_A^* p$, ahol $p \in Q\}$.

Tétel Bármely L nyelvre ekivalensek a következő állítások

- L környezetfüggetlen, azaz környezetfüggetlen (2-es típusú) grammatikával generálható
- L (nemdeterminisztikus) veremautomatával végállapottal felismerhető
- L (nemdeterminisztikus) veremautomatával üres veremmel felismerhető

Alternatív reprezentációk

- **Átírási szabályokkal:**

  A δ leképezést szabályok formájában is megadhatjuk. Az így nyert szabályhalmazt M_δ -val jelöljük. Tehát ezzel az alternatív jelöléssel:

  $zqa \to up :\in M_\delta \iff (u, p) \in \delta(z, q, a)$,

  $zq \to up :\in M_\delta \iff (u, p) \in \delta(z, q, \varepsilon)$.

  $(p, q \in Q, a \in T, z \in Z, u \in Z^*)$

- **Átmenetdiagrammal:**

  $p, q \in Q, a \in T \cup \{\varepsilon\}, z \in Z, u \in Z^*$ esetén:

  $$q \xrightarrow{a;\, z \to u} p \qquad \iff (u, p) \in \delta(z, q, a)$$

  A végállapotokat duplán karikázzuk. A kezdőállapotot $\to$ jelöli.

B) Mintapéldák:

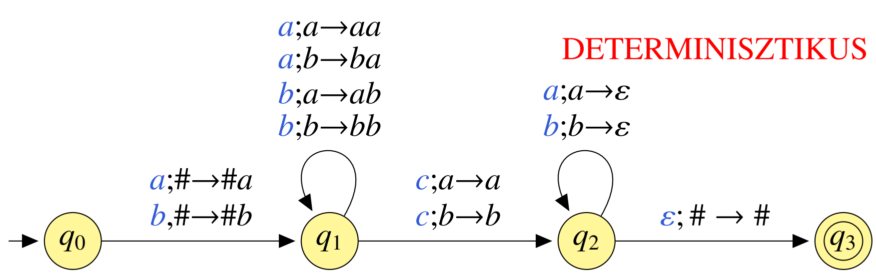
1. Feladat:

Legyen $L_1 = \{wcw^{-1} \mid w \in \{a, b\}^+\}$. Készítsünk egy A veremautomatát, melyre $L(A) = L_1$.

Megoldás:

$A = \langle\{q_0, q_1, q_2, q_3\}, \{a, b, c\}, \{\#, a, b\}, \delta, q_0, \#, \{q_3\}\rangle$, ahol:

$$
\begin{array}{ll}
(\#t, q_1) \in \delta(\#, q_0, t) & \forall t \in \{a, b\} \\
(zt, q_1) \in \delta(z, q_1, t) & \forall z, t \in \{a, b\} \\
(z, q_2) \in \delta(z, q_1, c) & \forall z \in \{a, b\} \\
(\varepsilon, q_2) \in \delta(t, q_2, t) & \forall t \in \{a, b\} \\
(\#, q_3) \in \delta(\#, q_2, \varepsilon) &
\end{array}
$$



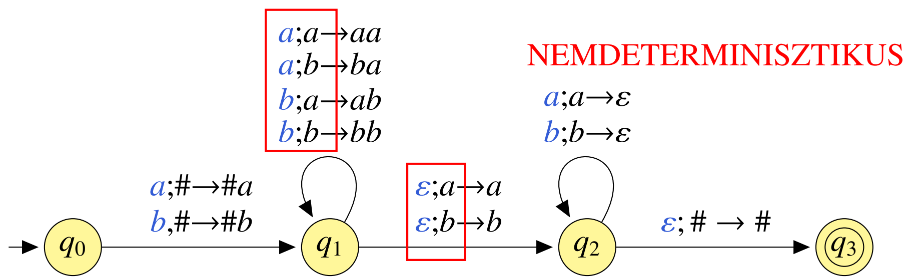
2. Feladat:

Legyen $L_2 = \{ww^{-1} \mid w \in \{a,b\}^+\}$. Készítsünk egy A veremautomatát, melyre $L(A) = L_2$.

Megoldás:

$A = \langle\{q_0, q_1, q_2, q_3\}, \{a, b\}, \{\#, a, b\}, \delta, q_0, \#, \{q_3\}\rangle$, ahol:

$$\begin{array}{rl}
(\#t, q_1) \in \delta(\#, q_0, t) & \forall t \in \{a,b\} \\
(zt, q_1) \in \delta(z, q_1, t) & \forall z, t \in \{a,b\} \\
(z, q_2) \in \delta(z, q_1, \varepsilon) & \forall z \in \{a,b\} \\
(\varepsilon, q_2) \in \delta(t, q_2, t) & \forall t \in \{a,b\} \\
(\#, q_3) \in \delta(\#, q_2, \varepsilon) &
\end{array}$$



3. Feladat: Adjunk meg egy az $L = \{a^n b^n \,|\, n \geq 1\}$ nyelvet üres veremmel felismerő veremautomatát!

Megoldás: Az alábbi $A = \langle\{\$, a\}\{q_0, q_1\}, \{a, b\}, \delta, \$, q_0, \{\,\}\rangle$ veremautomata esetén $N(A) = \{a^n b^n \,|\, n \geq 1\}$.

M_δ :

$$
\begin{aligned}
&\$q_0a \to \$aq_0\\
&aq_0a \to aaq_0\\
&aq_0b \to q_1\\
&aq_1b \to q_1\\
&\$q_1 \quad \to q_1.
\end{aligned}
$$

A determinisztikus, a^2b^3-re:

$\$q_0aabbb \Rightarrow \$aq_0abbb \Rightarrow \$aaq_0bbb \Rightarrow \$aq_1bb \Rightarrow \$q_1b \Rightarrow q_1b$.

A elutasítja $aabbb$-t, mivel hiába lett üres a verem, még volt hátra az inputból.

Jelölje $|u|_t$ az u szó t betűinek a számát.

4. Feladat: Adjunk meg egy az $L = \left\{ u \in \{a,b\}^* \;\middle|\; |u|_a = |u|_b \right\}$ nyelvet felismerő veremautomatát!

1. Megoldás: üres veremmel
$A = \langle \{\#, +, -\}\{q_0\}, \{a, b\}, \delta, \#, q_0, \{\ \}\rangle$

M_δ :

$\# q_0\, a \to \# + q_0$
$\# q_0\, b \to \# - q_0$
$+ q_0\, a \to + + q_0$
$+ q_0\, b \to q_0$
$- q_0\, a \to q_0$
$- q_0\, b \to - - q_0$
$\# q_0 \to q_0.$

2. Megoldás: végállapottal
$A = \langle \{\#, +, -\}\{q_0, q_1\}, \{a, b\}, \delta, \#, q_0, \{q_1\}\rangle$

M_δ :

$\# q_0\, a \to \# + q_0$
$\# q_0\, b \to \# - q_0$
$+ q_0\, a \to + + q_0$
$+ q_0\, b \to q_0$
$- q_0\, a \to q_0$
$- q_0\, b \to - - q_0$
$\# q_0 \to \# q_1.$

$\# q_0$ miatt mindkét verzió nemdeterminisztikus. A veremben a # fölött aktuálisan annyi + illetve - van amennyivel több a-t illetve b-t tartalmazott a már feldolgozott prefix.

3. Megoldás:
„Még nemdeterminisztikusabb” megoldást kapunk ha a veremtartalomtól függetlenül írhatunk a verembe + -t illetve - -t a-ra illetve b-re. A következő 2 szabályt hozzáadhatjuk:

$+ q_0\, b \to + - q_0$
$- q_0\, a \to - + q_0$

5. Feladat: Készítsünk üres veremmel elfogadó veremautomatát a csak az a változót tartalmazó helyes kifejezések nyelvéhez!

Megoldás:

Példák jó szavakra: $a + (a - a)$, $((a + a) - a * a) + a$, $(((a))) + a$.

Példák rossz szavakra: $(a + a)a$, $a(($, $)a + a$, $((a - a)) + a)$, $a + +$

$A = \langle\{\#, (\}, \{q_0, q_1\}, \{a, +, -, *, /, (,)\}, \delta, q_0, \#, \{\ \}\rangle$.

$\delta(\sigma, q_0, a) \ni (\sigma, q_1) \quad \forall\sigma \in \{\#, (\}$
$\delta(\sigma, q_0, () \ni (\sigma(, q_0) \quad \forall\sigma \in \{\#, (\}$
$\delta(\sigma, q_1, t) \ni (\sigma, q_0) \quad \forall\sigma \in \{\#, (\}, t \in \{+, -, *, /\}$
$\delta((, q_1,)) \ni (\varepsilon, q_1)$
$\delta(\#, q_1, \varepsilon) \ni (\varepsilon, q_1)$

Mivel $\delta(\#, q_1, \varepsilon)$ és $\delta(\#, q_1, -)$ nem üres a veremautomata nemdeterminisztikus.

A helyes zárójelezést a veremben, a helyes rákövetkezést (pl. műveleti jelet nem követhet műveleti jel) az állapotokkal ellenőrizzük.

C) Gyakorló feladatok:

1. $L = \{a^n b^n \mid n \geq 0\}$
2. $L = \{a^n b^{2n} \mid n \geq 0\}$
3. $L = \{a^{2n} b^n \mid n \geq 0\}$
4. $L = \{a^n b^k \mid n \geq k\}$
5. $L = \{a^n b^k \mid n < k\}$
6. $L = \{a^n b^k \mid n < 2k\}$
7. $L = \{a^n b^k \mid n > 2k\}$
8. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a < |u|_b \right\}$
9. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a \leq |u|_b \right\}$
10. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a = 2 \cdot |u|_b \right\}$
11. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a < 2 \cdot |u|_b \right\}$
12. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a \leq 2 \cdot |u|_b \right\}$
13. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a > 2 \cdot |u|_b \right\}$
14. $L = \left\{ u \in \{a, b\}^* \;\middle|\; |u|_a \geq 2 \cdot |u|_b \right\}$
15. $L = \{a^{2n} b a^k b^{n+1} \mid n, k \geq 0\}$
16. $L = \{a^n b^k c^k d^n \mid n, k \geq 0\}$
17. $L = \{w \in \{a, b\}^* \mid w = w^{-1} \ \wedge \ |w| \equiv 0 \pmod 3\}$
18. $L = \{w \in \{a, b\}^* \mid w = w^{-1} \ \wedge \ |w| \equiv 1 \pmod 3\}$
19. $L = \{w \in \{a, b\}^* \mid w = w^{-1} \ \wedge \ |w|_a \equiv 0 \pmod 3\}$