

A digitális számítás elmélete

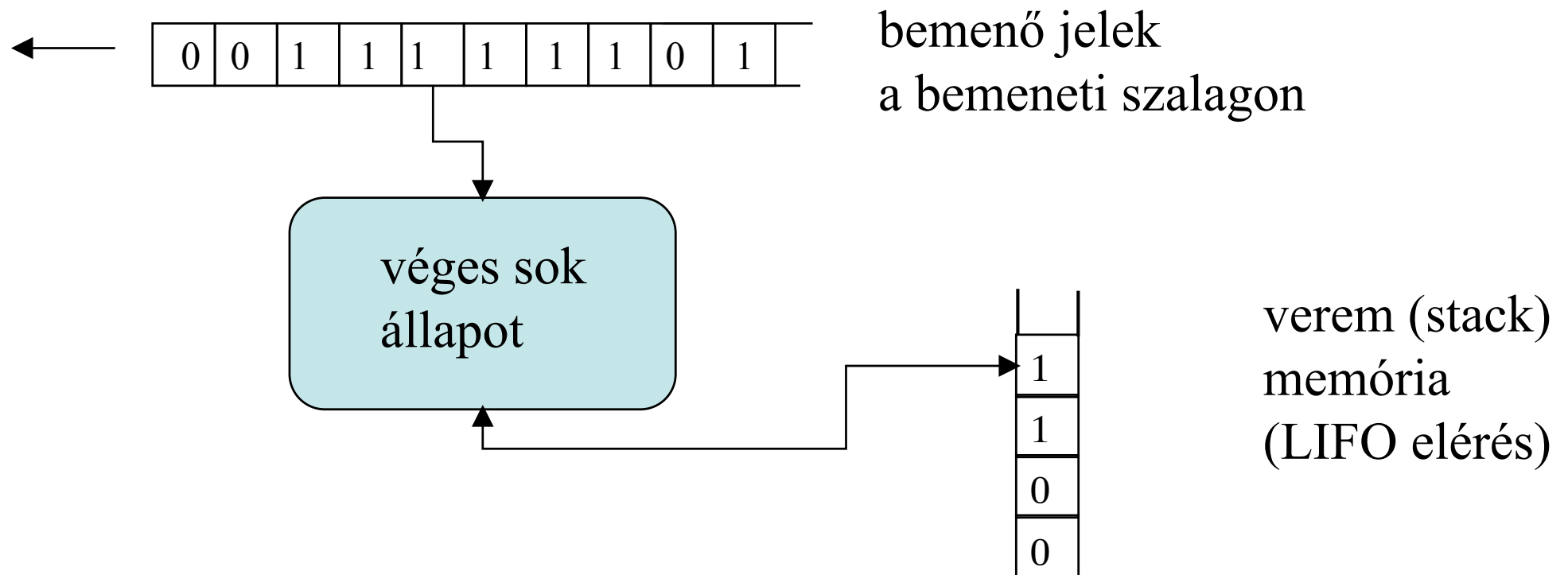
4. előadás márc. 7.

- Verem automata
- Verem automaták által felismert nyelvek éppen a CF nyelvek
- Operátorok CF nyelveken.
- CF nyelv és R (reguláris) nyelv metszetre CF nyelv
- Nem determinisztikus verem automata, nem determinisztikus CF nyelvek.

Verem automaták

- DVA nem alkalmas CF nyelv felismerésére ...
- Milyen módosítás szükséges?

Pl: $\{w w^R \mid w \in \Sigma^*\}$, $\{0^n 1^n \mid 0 \leq n\}$ CF nyelvek felismeréséhez



Verem automaták

Definíció: Verem automata

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

1. Q , egy véges halmaz az **állapotok** halmaza
2. Σ , egy véges halmaz a **bemeneti abc**
3. Γ , egy véges halmaz a **verem abc**
4. $\delta: Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \rightarrow \mathcal{P}(Q \times \Gamma_\varepsilon)$ az **állapotátmenet függvény**
5. $q_0 \in Q$ a **kezdeti állapot**
6. $F \subseteq Q$ a **végállapotok halmaza**

Jelölés: $\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$, $\Gamma_\varepsilon = \Gamma \cup \{\varepsilon\}$ ('nem olvasunk-írunk')

Verem automaták

Definíció: **Felismert (elfogadott) jelsorozat**

$M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ egy verem automata, $w = w_1 w_2 \dots w_n \Sigma_\varepsilon$ -beli jelek sorozata.

M felismeri w -t, ha $\exists$:

$r_0 r_1 \dots r_n$ Q -beli állapotok sorozata, és

$s_0 s_1 \dots s_n \Gamma_\varepsilon$ -feletti szavak sorozata hogy:

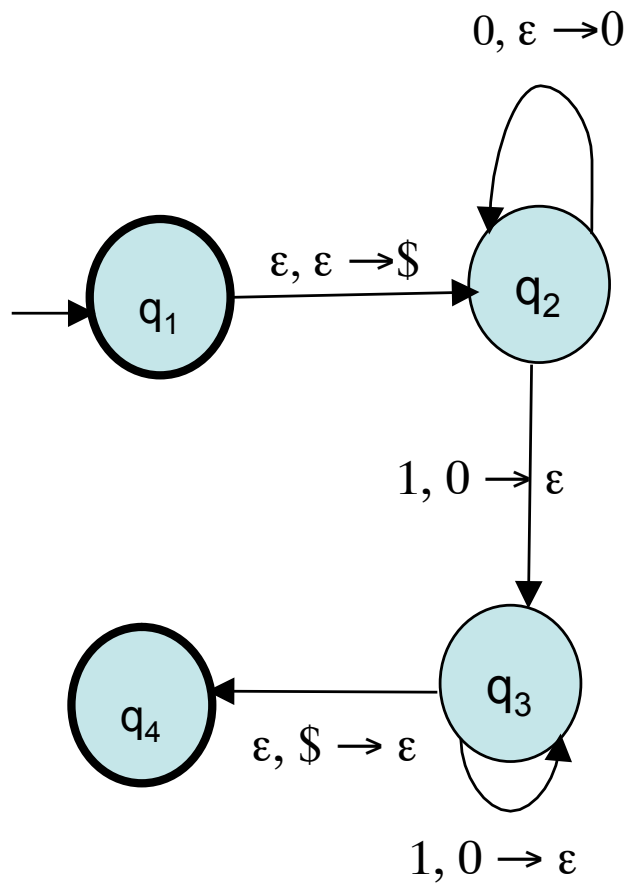
1. $r_0 = q_0$ és $s_0 = \varepsilon$
2. $(r_{i+1}, b) \in \delta(r_i, w_{i+1}, a)$ $i = 0, \dots, n-1$, ahol:
 $s_i = at$ és $s_{i+1} = bt$ valamely $a, b \in \Gamma_\varepsilon$ és $t \in \Gamma^*$
3. $r_n \in F$

Definíció: M verem automata **felismeri az A nyelvet**, ha

$A = \{w \mid M \text{ felismeri } w\text{-t}\}.$

Verem automaták

$$\Sigma = \{0, 1\}, \Gamma = \{0, \$\}$$



Σ_ϵ	0			1			ϵ		
Γ_ϵ	0	\$	ϵ	0	\$	ϵ	0	\$	ϵ
q_1	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(q_2, \$)\}$
q_2	$\emptyset$	$\emptyset$	$\{(q_2, 0)\}$	$\{(q_3, \epsilon)\}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
q_3	$\emptyset$	$\emptyset$	$\emptyset$	$\{(q_3, \epsilon)\}$	$\emptyset$	$\emptyset$	$\emptyset$	$\{(q_4, \epsilon)\}$	$\emptyset$
q_4	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

01

~~010~~

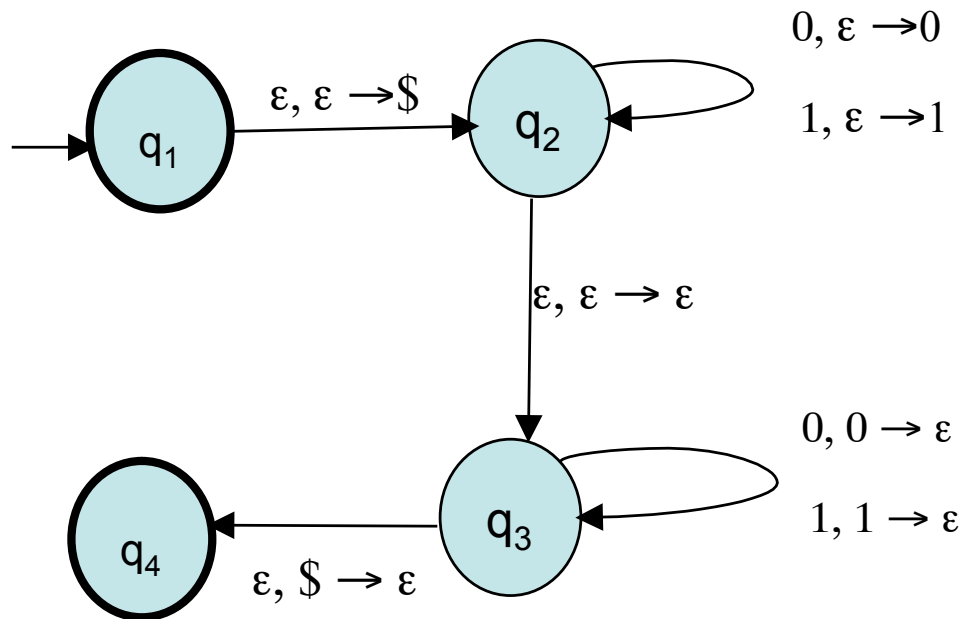
~~1100~~

0011

$\{ 0^n 1^n \mid 0 \leq n \}$

Verem automaták

$$\Sigma = \{0, 1\}, \Gamma = \{0, 1\}$$



00
~~01~~
~~10~~
 11
~~000~~
~~100~~
~~010~~
~~0101~~
 0110

$$\{ ww^R \mid w \in \{0,1\}^* \}$$

Verem automaták

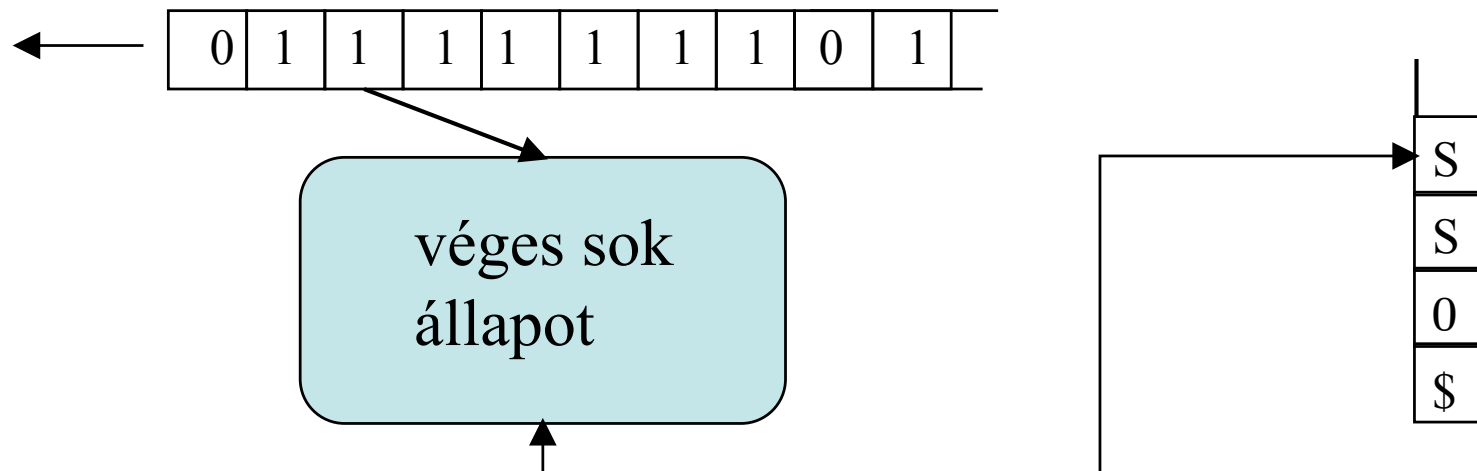
Tétel: Minden L CF nyelvhez létezik őt felismerő verem automata.

Biz: Legyen egy, az L -et generáló CF grammatika G .

Megkonstruálunk egy M verem automatát, amely minden bemeneti jelsorozatról eldönti, hogy levezethető-e a G grammatika szerint.

Ötlet: M a G baloldali levezetéseit szimulálja.

Pl: $S \rightarrow 01SS0$ szabály alkalmazása



Verem automaták

1. Általánosítás: $(r, u) \in \delta(q, a, s)$, ahol $u = u_1 u_2 \dots u_k$.

Szimulálás az eredeti modellel: beiktatunk új $q_1, q_2, \dots, q_{k-1}$ közbülső állapotokat, melyekre:

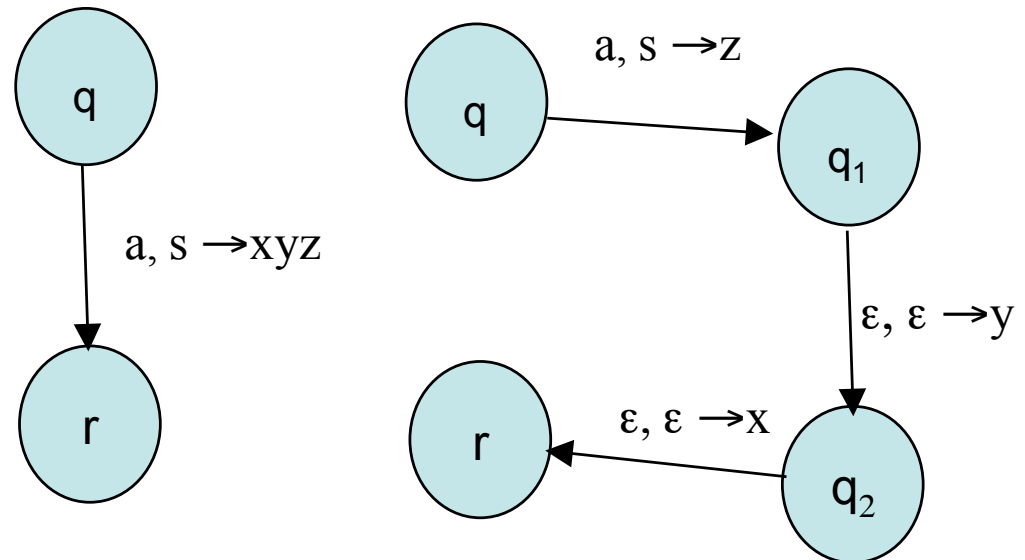
$$\delta(q, a, s) = \{(q_1, u_k)\}$$

$$\delta(q_1, \epsilon, \epsilon) = \{(q_2, u_{k-1})\}$$

$$\delta(q_2, \epsilon, \epsilon) = \{(q_3, u_{k-2})\}$$

...

$$\delta(q_{k-1}, \epsilon, \epsilon) = \{(r, u_1)\}$$



Verem automaták

Tehát használhatjuk azt a rövidítést, hogy $(r, u) \in \delta(q, a, s)$, ami azt jelenti, hogy a q állapotból az a jel és s verembeli jel olvasására az automata az r állapotba kerül, és az u sorozatot betette a verembe.

2. M állapotainak halmaza, $Q = \{q_{\text{kezdő}}, q_{\text{vég}}, q_{\text{ciklus}}\}$.

- $\delta(q_{\text{kezdő}}, \varepsilon, \varepsilon) = \{(q_{\text{ciklus}}, S\$)\}$ **inicializálás**

- **a) a verem legfelső eleme nemterminális**

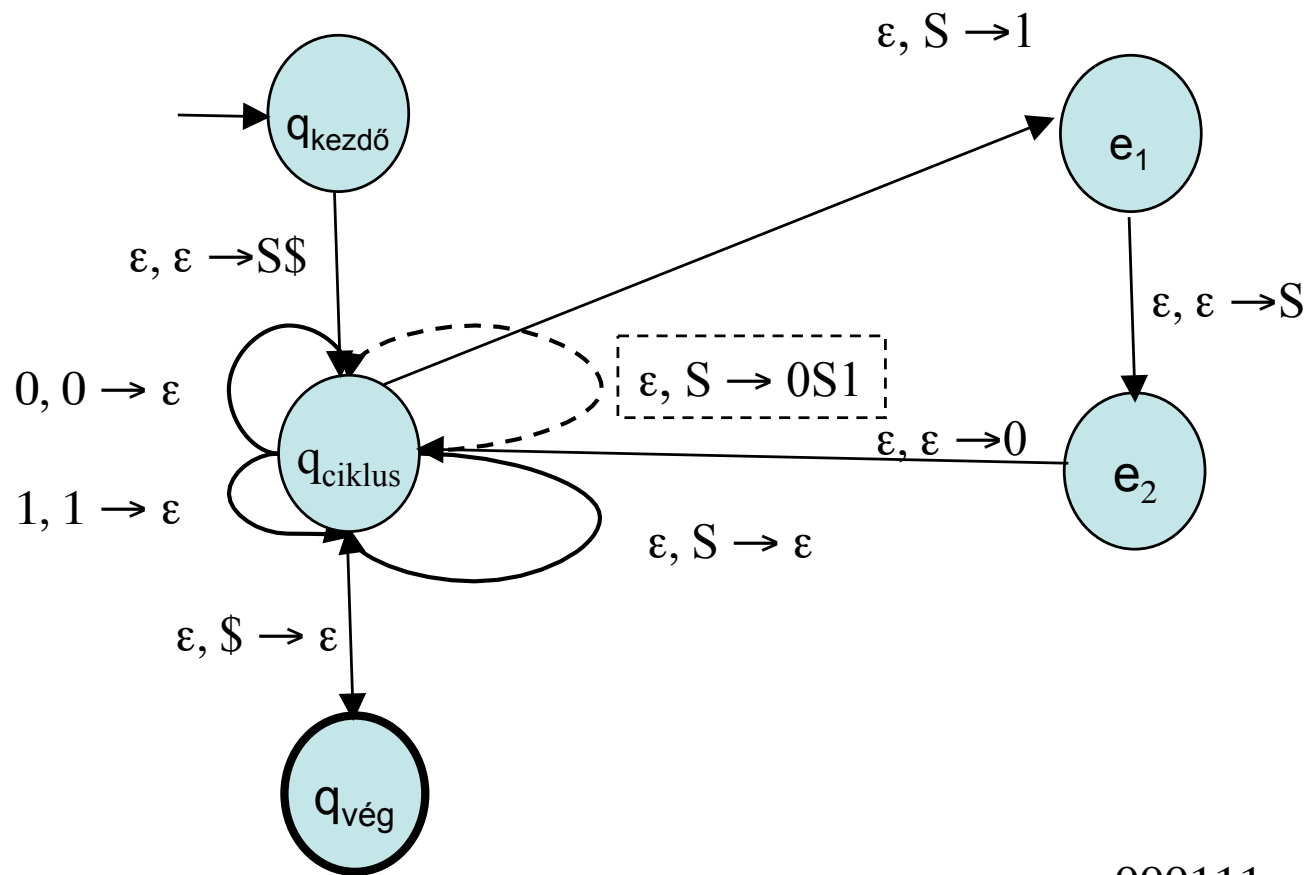
$$\delta(q_{\text{ciklus}}, \varepsilon, A) = \{(q_{\text{ciklus}}, w) \mid A \rightarrow w \text{ szabálya } G\text{-nek}\}$$

- b) a verem legfelső eleme terminális: $\delta(q_{\text{ciklus}}, a, a) = \{(q_{\text{ciklus}}, \varepsilon)\}$**

- c) a verem legfelső eleme $\$$: $\delta(q_{\text{ciklus}}, \varepsilon, \$) = \{(q_{\text{vég}}, \varepsilon)\}$**

Verem automaták

Példa: $S \rightarrow 0S1 \mid \varepsilon$



Verem automaták

Tétel: Minden M verem automatához létezik G CF grammatika, mely éppen az M által felismert nyelvet generálja.

Biz: Feltesszük, hogy M -re igaz, hogy:

1. Egyetlen végállapota van, ez $q_{\text{vég}}$.
2. A végállapotot üres veremmel éri el M .
3. Minden egyes állapotátmenet vagy **betesz** egy szimbólumot a verembe, vagy **kivesz** egyet onnan.

Egy adott M módosítható úgy, hogy a feltételek teljesüljenek.

Verem automaták

Olyan nyelvtant kreálunk, amely minden p és q állapotpárhoz tartalmaz egy A_{pq} változót, melyből minden olyan szó levezethető, amely az M automatát p -ből q -ba viszi úgy, hogy p -ben és q -ban is üres a verem.

Egy ilyen p -t q -ba vivő állapotátmenet-sorozat *betesz* lépéssel kell hogy kezdődjön, és *kivesz* lépéssel kell hogy végződjön.

Legyen t az első lépéssel betett szimbólum. Két esetet lehet:

- a) A t végig a veremben maradt, csak az utolsó *kivesz* távolította el.

Ezt egy $A_{pq} \rightarrow aA_{rs}b$ szabállyal szimuláljuk, ahol a az első, b pedig az utolsó lépésben a bemenetről olvasott szimbólum, r a p -t követő, s a q -t megelőző állapot az állapotátmenet sorozatban.

Verem automaták

b) A t már hamarabb kikerült a veremből valamely r állapotba
érkezve: $A_{pq} \rightarrow A_{pr}A_{rq}$.

Legyen $M=(Q, \Sigma, \Gamma, \delta, q_0, \{q_{\text{vég}}\})$. G grammatika nemterminálisai
legyenek $\{A_{pq} \mid p, q \in Q\}$, és legyen a kezdőszimbólum $A_{q_0q_{\text{vég}}}$

Megadjuk G szabályait:

- a) $A_{pq} \rightarrow aA_{rs}b$ ahol $a, b \in \Sigma_\varepsilon$, $p, q, r, s \in Q$, és $\exists t \in \Gamma$, hogy:
 $(r, t) \in \delta(p, a, \varepsilon)$ és $(q, \varepsilon) \in \delta(s, b, t)$.
- b) $A_{pq} \rightarrow A_{pr}A_{rq}$ minden $p, q, r \in Q$ esetén.
- c) $A_{pp} \rightarrow \varepsilon$ minden $p \in Q$ esetén.

Megmutatjuk, hogy $x \in L(G) \Leftrightarrow x \in L(M)$.

Verem automaták

1. Lemma: Ha x levezethető A_{pq} -ból, akkor M a p állapotból üres veremmel indulva, az x -et beolvasva el tud jutni a q állapotba úgy, hogy a verem ismét üres.

Biz: a levezetések lépésszáma, n szerinti teljes indukcióval.

Alap eset: $n=1$

Az egyetlen levezetési szabály jobb oldalán csak terminális szerepel.

Ilyen szabály csak $A_{pp} \rightarrow \varepsilon$ lehet.

Ekkor valóban, a p állapotból a p állapot elérhető, és üres kiindulási veremmel üres vermet eredményez.

Verem automaták

Indukciós feltevés: legfeljebb k levezetési szabály alkalmazása esetén igaz a lemma állítása.

Megmutatjuk, hogy ekkor $k+1$ -re is igaz.

Legyen $A_{pq} \Rightarrow^*_G x$ egy $k+1$ hosszú levezetés. Az első lépés, a szabályok alakja alapján, kétféle lehet:

a) $A_{pq} \rightarrow aA_{rs}b$, ekkor $x=ayb$, ahol y levezethető A_{rs} -ből k lépésben.

Ekkor az ind. feltevés szerint r -ből s állapotot elérhető az y inputot beolvasva úgy, hogy az r és az s állapotban is a verem üres.

Másrészt a szabály definíciója szerint $\exists t \in \Gamma$, hogy:

$(r,t) \in \delta(p, a, \varepsilon)$ és $(q, \varepsilon) \in \delta(s, b, t)$. Azaz t -t betéve, majd az utolsó lépésben kivéve, a lemma állítása igaz $A_{pq} \Rightarrow^*_G x$ -re.

Verem automaták

b) $A_{pq} \rightarrow A_{pr}A_{rq}$, ekkor $x=yz$, ahol y levezethető A_{pr} -ből és z levezethető A_{rq} -ből legfeljebb k lépésben.

Ekkor az ind. feltevés szerint p -ből az r állapot elérhető az y inputot beolvasva, illetve r -ből a q állapot elérhető az z inputot beolvasva úgy, hogy a p, r és q állapotban is a verem üres.

Azaz p -ből a q állapot elérhető az x inputot beolvasva úgy, hogy a p és q állapotban a verem üres, a lemma állítása igaz $A_{pq} \Rightarrow^*_G x$ -re.

Verem automaták

2. Lemma: Ha M valamely p állapotából és üres veremből x -et beolvasva, van olyan állapotátmenetek sorozata, hogy az automata q állapotba kerül, és a verem ismét üres, akkor x levezethető a grammatika A_{pq} nemterminálisából.

Biz: az állapotátmenetek száma, n szerinti teljes indukcióval.

Alap eset: $n=0$

Ekkor nincs beolvasás, tehát $x = \epsilon$ kell hogy legyen. Valóban szerepel

G szabályai között A_{pp} $\rightarrow \epsilon$, ekkor az állítás igaz.

Verem automaták

Indukciós feltevés: legfeljebb k hosszú állapotátmenet sorozatra igaz a lemma állítása.

Megmutatjuk, hogy ekkor $k+1$ -re is igaz.

Legyen x -hez egy megfelelő $k+1$ hosszú állapotátmenet-sorozat első ill. utolsó állapota p ill. q . E sorozathoz tartozó vermeket tekintve, két eset lehetséges:

a) üres verem nem szerepel közbülső helyen.

Ekkor az első betett elem azonos kell hogy legyen a legvégén kivett elemmel, legyen ez t . Legyen a az elsőnek, b az utolsónak beolvasott elem, r a második, s az utolsó előtti állapot. Mivel $(r,t) \in \delta(p, a, \varepsilon)$ és $(q, \varepsilon) \in \delta(s, b, t)$, így a $A_{pq} \rightarrow aA_{rs}b$ szabály szerepel G -ben. Legyen $x=ayb$. Ekkor az r -ből s -be jutó, y -t beolvasó átmenet-sorozatra alkalmazható az ind. feltevés, azaz y levezethető A_{rs} -ből. De ekkor $A_{pq} \Rightarrow^*_G x$ is igaz.

Verem automaták

- b) Legyen r az a közbülső állapot, amikor a verem kiürül. Ekkor a p -ből r -be illetve r -ből q -ba vivő állapotátmenetek sorozata legfeljebb k hosszú. Legyen $x=yz$, ahol y az r állapotig beolvasott, z pedig a megmaradó részsorozata x -nek. Az ind. feltevés miatt $A_{pr} \Rightarrow^*_G y$ és $A_{rq} \Rightarrow^*_G z$. Mivel az $A_{pq} \rightarrow A_{pr}A_{rq}$ szabály szerepel G -ben, így ennek alkalmazásával x levezethető A_{pq} -ből.

Verem automaták

-

A fenti eredményeket alkalmazva az $A_{q_0q_{\text{vég}}}$ kezdőszimbólumra azt kapjuk, hogy x levezethető $A_{q_0q_{\text{vég}}}$ -ből csakkor, ha M a kezdő állapotából, üres veremmel x -et beolvasva, van állapotátmenetek olyan sorozata, hogy az automata a vég állapotba kerül, és a verem ismét üres.

Tehát x -et felismeri M csakkor, ha x -et generálja a G grammatika.

Ezzel a tétel állítását bebizonyítottuk.

Műveletek CF nyelveken

Tétel: Legyen $\mathcal{L}$ a CF nyelvek halmaza.

$\mathcal{L}$ zárt a reguláris operátorokra nézve.

Azaz ha $A_1 \in \mathcal{L}$ és $A_2 \in \mathcal{L}$, akkor:

1. $A_1 \cup A_2 \in \mathcal{L}$.
2. $A_1 \circ A_2 \in \mathcal{L}$.
3. $A_1^* \in \mathcal{L}$.

Műveletek CF nyelveken

Bizonyítás:

Legyen $G_1 = (V_1, \Sigma_1, R_1, S_1)$ és $G_2 = (V_2, \Sigma_2, R_2, S_2)$ A_1 ill. A_2 $\mathcal{L}$ –beli nyelveket generáló grammatikák. Az általánosság megszorítása nélkül feltehetjük, hogy V_1 és V_2 diszjunktak.

1. Unió esete

Legyen $G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R, S)$, ahol

$R = R_1 \cup R_2 \cup \{S \rightarrow S_1, S \rightarrow S_2\}$ Ekkor $L(G) = L(G_1) \cup L(G_2)$

2. Szorzat esete

Legyen $G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R, S)$, ahol

$R = R_1 \cup R_2 \cup \{S \rightarrow S_1 S_2\}$ Ekkor $L(G) = L(G_1) \circ L(G_2) = A_1 \circ A_2$

Műveletek CF nyelveken

3. Iterált esete

Legyen $G = (V_1 \cup \{S\}, \Sigma_1, R, S)$, ahol

$R = R_1 \cup \{S \rightarrow \varepsilon, S \rightarrow SS_1\}$ Ekkor $L(G) = L(G_1)^* = A_1^*$

Műveletek CF nyelveken

Tétel: Legyen $\mathcal{L}$ a CF nyelvek halmaza.

$\mathcal{L}$ **nem** zárt a metszetre nézve.

$\mathcal{L}$ **nem** zárt a komplementerre nézve.

Bizonyítás:

$A_1 = \{a^n b^n c^m \mid n, m \geq 0\}$, $A_2 = \{a^n b^m c^m \mid n, m \geq 0\}$ mindketten CF nyelvek. $A_1 \cap A_2 = \{a^n b^n c^n \mid n \geq 0\}$ viszont nem CF (beláttuk).

A De Morgan azonosság szerint: $A_1 \cap A_2 = (A_1^c \cup A_2^c)^c$, így a komplementerre zártság maga után vonná a $\cap$ -re zárttságot.

Műveletek CF nyelveken

Tétel: Ha A_1 CF nyelv, A_2 reguláris nyelv, akkor $A_1 \cap A_2$ CF nyelv.

Bizonyítás: Hasonlít a reguláris nyelvek esetében használtra. Legyen

$M_1 = (Q_1, \Sigma, \Gamma, \delta_1, q_1, F_1)$ verem automata, hogy $A_1 = L(M_1)$ és

$M_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ dva, hogy $A_2 = L(M_2)$. Legyen

$M = (Q_1 \times Q_2, \Sigma, \Gamma, \delta, (q_1, q_2), F_1 \times F_2)$ verem automata, ahol:

$\delta: Q_1 \times Q_2 \times \Sigma_\varepsilon \times \Gamma_\varepsilon \rightarrow \mathcal{P}(Q_1 \times Q_2 \times \Gamma_\varepsilon)$ állapotátmenet fv legyen:

$$\delta(p, q, a, b) = \{ r, (\delta_2(q, a), s) \mid (r, s) \in \delta_1(p, a, b) \}$$

M valóban $A_1 \cap A_2$ -t ismeri fel (hf).

Determinisztikus CF nyelvek

Def: Az $M=(Q, \Sigma, \Gamma, \delta, q, F)$ verem automata determinisztikus, ha a

$\delta: Q \times \Sigma_{\varepsilon} \times \Gamma_{\varepsilon} \rightarrow \mathcal{P}(Q \times \Gamma_{\varepsilon})$ állapotátmenet fv olyan, hogy minden q, a, b esetén $\delta(q, a, b)$ maximum egyelemű halmaz.

Def: Az $L \subseteq \Sigma^*$ nyelv determinisztikus CF, ha az $L\{\$$ nyelvhez van

őt generáló determinisztikus verem automata, ahol

$\$$ nem Σ -beli szimbólum, mely a nyelv szavainak végét jelöli.

Pl: $L = a^* \cup \{a^n b^n \mid n \geq 1\}$

Determinisztikus CF nyelvek

Tétel: A determinisztikus CF nyelvek osztálya zárt a reguláris operátorokra, valamint a komplementerképzésre, és a metszetképzésre.

Biz: nem tárgyaljuk

Tétel: A determinisztikus CF nyelvek osztálya szűkebb, mint CF nyelvek osztálya, azaz van CF nyelv, ami nem determinisztikus CF nyelv.

Pl: $L = \{a^k b^m c^n \mid k \neq m \text{ vagy } m \neq n\}$ CF, de nem determinisztikus CF.

Biz: A fenti tételt alkalmazzuk és keresünk egy alkalmas reguláris nyelvet.

Miről volt szó ?

- Verem automata
- Verem automaták által felismert nyelvek = CF nyelvek
- Operátorok CF nyelveken.
- CF nyelv és R (reguláris) nyelv metszetre CF nyelv
- Nem determinisztikus verem automata, nem determinisztikus CF nyelvek.