Formális nyelvek és a fordítóprogramok alapjai

4. előadás

Előadó: Nagy Sára, mesteroktató Algoritmusok és Alkalmazásaik Tanszék

Véges determinisztikus automata (VDA)

Definíció:

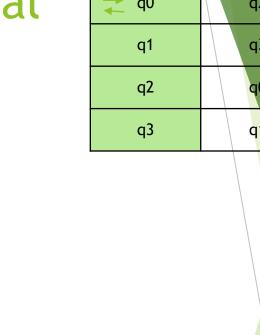
 $A = (Q, T, \delta, q_0, F)$ rendezett ötöst véges determinisztikus automatának nevezzük, ahol

- Q az állapotok nem üres véges halmaza,
- T az input szimbólumok ábécéje,
- δ : Q x T \rightarrow Q leképezés az állapot-átmeneti függvény,
- q₀ ∈ Q a kezdőállapot,
- F ⊆ Q elfogadóállapotok halmaza.

Véges determinisztikus automata (VDA)

Véges determinisztikus automata estén a δ: $Q \times T \rightarrow Q$ állapot-átmeneti függvény minden (q,a) párra értelmezett, ahol $(q,a) \in Q \times T$ és egyetlen olyan $p \in Q$ állapot van, amelyre $\delta(q,a) = p$.

Példa - automata megadása gráffal



q3

	b	
((q0))		$\overline{}$ (q1)
† I	b	<u>† 1</u>
a a		a a
1 •	b	1 *
(q2)		(q3)
	b	

$L(A)=\{u\in T^*\mid$	u-ban páro	s sok .a'	betű és	páros sok	.b' betű var	ıί
	a ball palo	3 301x , a	DC CG C5	pai os son	, b beta vai	ر.

Alternatív jelölés az állapot-átmenetre

 $\delta(q, a) = p$ állapot átmenetet jelölhetjük egy $qa \rightarrow p$ szabállyal.

Ha minden egyes (q, a) párra egyetlen qa → p szabály van, akkor a véges automata determinisztikus, egyébként nemdeterminisztikus.

Véges nemdeterminisztikus automata (VNDA)

Definíció:

 $A = (Q, T, \delta, Q_0, F)$ rendezett ötöst véges nemdeterminisztikus automatának nevezzük, ahol

- Q az állapotok nem üres véges halmaza,
- T az input szimbólumok ábécéje,
- $\delta: Q \times T \to \mathcal{P}(Q)$ (a Q részhalmazaiba képez)
- $Q_0 \subseteq Q$ a kezdőállapotok halmaza,
- F ⊆ Q elfogadó állapotok halmaza.

Megjegyzés: VNDA a VDA általánosítása

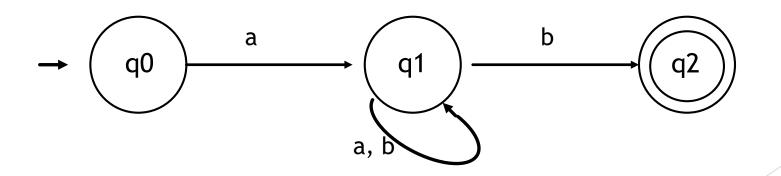
Példa

```
Legyen A = (Q, T, \delta, Q<sub>0</sub>, F) a következő, ahol Q = {q0, q1, q2}, T = {a, b}, Q<sub>0</sub> = {q0}, F = {q2} és \delta(q0, a) = {q1}, \delta(q0, b) = \emptyset \delta(q1, a) = {q1}, \delta(q1, b) = {q1, q2} \delta(q2, a) = \emptyset, \delta(q2, b) = \emptyset
```

L(A)={u ∈ T* | az u szó 'a' betűvel kezdődik és 'b' betűre végződik}

Példa - nemdeterminisztikus automata megadása táblázattal és gráffal

δ	a	b
→ q0	q1	
q1	q1	q1, q2
← q2		



Közvetlen redukció

Legyen A = (Q, T, δ, Q_0, F) egy véges nemdeterminisztikus automata és legyenek $\mathbf{u}, \mathbf{v} \in \mathbf{QT}^*$.

(Konfiguráció: aktuális állapot, input hátralévő része.)

Azt mondjuk, hogy az A automata az u konfigurációt a v konfigurációra redukálja közvetlenül (jelölés: $u \Rightarrow_A v$), ha van olyan

 $qa \rightarrow p$ szabály (azaz $\delta(q, a) = p$) és van olyan $w \in T^*$ szó, amelyre

u = qaw és v = pw teljesül.

Redukció

Definíció:

```
Az A = (Q, T, \delta,Q<sub>0</sub>, F) véges automata az u \in QT^* konfigurációt a v \in QT^* konfigurációra redukálja (jelölés: u \underset{A}{\Rightarrow} v), ha vagy u = v, vagy van olyan z \in QT^*, amelyre u \underset{A}{\Rightarrow} z és z \underset{A}{\Rightarrow} v teljesül.
```

Automata által elfogadott nyelv

Definíció:

Az A = (Q, T, δ, Q_0, F) véges automata által elfogadott nyelv alatt az

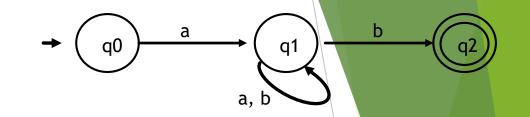
L(A) := $\{u \in T^* | \exists q_0 u \underset{A}{\Rightarrow}^* p, q_0 \in Q_0 \text{ \'es } p \in F\}$ szavak halmazát értjük.

Megjegyzés: Ez azt jelenti, hogy van olyan működése az automatának, hogy egy kezdőállapotból indulva és végig olvasva az inputot elfogadóállapotba jut.

Példa

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

 $\delta: q_0 a \rightarrow q_1, q_1 a \rightarrow q_1, q_1 b \rightarrow q_1, q_1 b \rightarrow q_2$



Redukálás:

Kérdés: $aabb \in L(A)$?

Első próbálkozás:

$$q_0aabb \Rightarrow q_1abb \Rightarrow q_1bb \Rightarrow q_1b \Rightarrow q_1 \qquad q_1 \notin F$$

Második próbálkozás:

$$q_0aabb \Rightarrow q_1abb \Rightarrow q_1bb \Rightarrow q_1b \Rightarrow q_2 \qquad q_2 \in F,$$
 azaz az **aabb** jó szó.

Harmadik próbálkozás:

 $q_0aabb \Rightarrow q_1abb \Rightarrow q_1bb \Rightarrow q_2b$ Nem tudtuk végig olvasni a szót.

Emlékeztető:

3-típusú grammatikák normálformája

Tétel:

Minden 3-as típusú nyelv generálható egy olyan grammatikával, amelynek szabályai

 $A \rightarrow aB$, ahol A, B \in N és a \in T vagy

 $A \rightarrow \epsilon$ alakúak, ahol $A \in N$.

Tétel:

Minden 3-as típusú L nyelvhez megadható egy véges nemdeterminisztikus automata, és fordítva, minden nemdeterminisztikus automata 3-as típusú nyelvet ismer fel.

$$(\mathcal{L}_3 \subseteq \mathcal{L}_{VNDA}, \mathcal{L}_{VNDA} \subseteq \mathcal{L}_3)$$

Megjegyzés: Már láttuk, hogy $\mathcal{L}_{reg} \subseteq \mathcal{L}_3$, így a reguláris kifejezésekhez (lexikális egységekhez) építhető automata.

Bizonyítás vázlat:

Legyen L egy 3-as típusú nyelv. Ez azt jelenti, hogy megadható egy G=(N,T,P,S) 3-as normál formájú grammatikával is.

G alapján megkonstruálható egy

 $A = (Q, T, \delta, Q_0, F)$ egy véges nemdeterminisztikus automata úgy, hogy L=L(G)=L(A).

Bizonyítás vázlat:

Feleltessünk meg minden nemterminálisnak egy állapotot. Jelölje q_A az $A \in N$ -hez rendelt állapotot.

 $A = (\{q_A \mid A \in N\}, T, \delta, q_S, F)$

3-as típusú nyelvek kapcsolata a véges automatákkal Bizonyítás vázlat:

Legyen $\delta(q_A,a)=q_B$,akkor és csak akkor, ha $A{
ightarrow}aB\in P$.

$$q_A a \rightarrow q_B$$



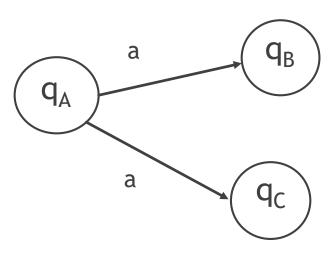
Legyen $q_A \in F$, akkor és csak akkor, ha $A \rightarrow \epsilon \in P$.



Bizonyítás vázlat:

Az így létrehozott automata lehet nemdeterminisztikus.

Ha például $A \rightarrow aB \in P$ és $A \rightarrow aC \in P$ szabály is van, akkor ennek az alábbi automata felel meg.



Bizonyítás vázlat:

A konstrukcióból adódik, ha

 $S \underset{G}{\Rightarrow} *u$, akkor $q_S \underset{A}{\Rightarrow} *u$ és fordítva.

A tétel megfordításának bizonyítása hasonló, csak ott A automatához konstruáljuk meg a G 3-as típusú grammatikát.

Példa

L:= $\{u \in \{a,b\}^* \mid az \ u \ szó 'a' betűvel kezdődik és 'b' betűre végződik}$

$$G=(\{S,A,B\},\{a,b\},P,S) \qquad A=(\{q_S,q_A,q_B\},\{a,b\},\ \delta,\{q_S\},\{q_B\})$$

$$P: S \rightarrow aA$$

$$A \rightarrow aA$$

$$A \rightarrow bA$$

$$A \rightarrow bB$$

$$A \rightarrow$$

Nemdeterminisztikus automaták determinisztikussá tétele

Tétel:

Minden $A=(Q,T,\delta,Q_0,F)$ nemdeterminisztikus automatához megadható egy $A'=(Q',T,\delta',q_0',F')$ véges determinisztikus automata, hogy L(A')=L(A).

$$(\mathcal{L}_{VNDA} \subseteq \mathcal{L}_{VDA})$$

3-as típusú nyelvek és automaták kapcsolata

$$\mathcal{L}_3 \subseteq \mathcal{L}_{VNDA} \text{ \'es } \mathcal{L}_{VNDA} \subseteq \mathcal{L}_3$$

(3-as normál formájú grammatika átírható automatává és fordítva.)

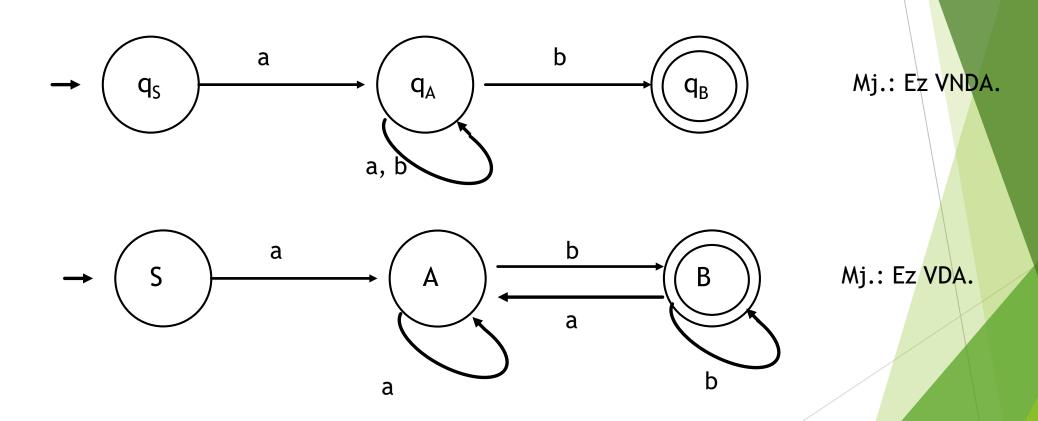
$$\mathcal{L}_{VDA} \subseteq \mathcal{L}_{VNDA}$$
 (Definíció alapján triviális.)

Ha $\mathcal{L}_{VNDA} \subseteq \mathcal{L}_{VDA}$ is teljesülne, akkor

$$\mathcal{L}_3 = \mathcal{L}_{VNDA} = \mathcal{L}_{VDA}$$
.

Példa

L:= $\{u \in \{a,b\}^* \mid az \ u \ szó 'a' betűvel kezdődik és 'b' betűre végződik}$



Determinisztikus automata megkonstruálása:

Legyen $Q':=\mathcal{P}(Q)$, azaz Q összes részhalmazainak halmaza, azaz hatványhalmaza.

Legyen a δ ': Q' x T \rightarrow Q' a következőképpen definiálva:

 $δ'(q',a):= ∪_{q∈q'}δ(q,a)$, ahol q'∈Q' és a∈T.

Legyen $q_0' := Q_0$ és $F' := \{q' \in Q' \mid q' \cap F \neq \emptyset\}$.

Konstrukció helyességének bizonyítása (L(A) L(A'))

Lemma1:

```
\forall q,p\inQ és q'\inQ' és u,v\inT* esetén, ha qu \Rightarrow* pv és q\inq', akkor \existsp' \inQ' úgy, hogy q'u \Rightarrow* p'v és p \inp'.
```

Ha $u \in L(A)$, akkor $\exists q_0 u \Rightarrow^* p$, ahol $q_0 \in Q_0$ és $p \in F$.

Lemma1 alapján $\exists p' \in Q'$ úgy, hogy $q_0'u \underset{A_t}{\Rightarrow} p'$ és $p \in p'$.

De p∈F, így p'∩F≠∅, azaz p'∈F' azaz u∈L(A').

Konstrukció helyességének bizonyítása (L(A') CL(A))

Lemma2:

```
\forall q',p'\inQ' és p\inQ és u,v\inT* esetén, ha q'u \Rightarrow* p'v és p\inp', akkor \existsq\inQ úgy, hogy qu \Rightarrow* pv és q \inq'. Ha u\inL(A'), akkor \exists q_0'u \Rightarrow* p' és p'\inF'.
```

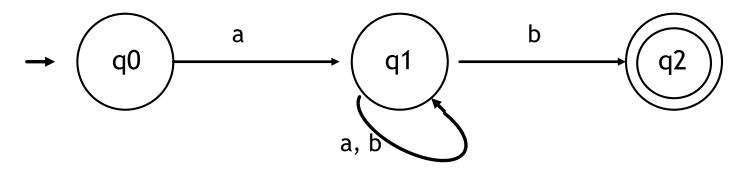
F' definiciója alapján van olyan p∈p', hogy p∈F.

Lemma2 alapján $\exists q_0 \in Q$ úgy, hogy $q_0 u \underset{A}{\Rightarrow} * p$, ahol $q_0 \in q_0' (=Q_0)$. De $p \in F$, azaz $u \in L(A)$.

Példa:

L:={u ∈ {a,b}* | az u szó 'a' betűvel kezdődik és 'b' betűre végződik}

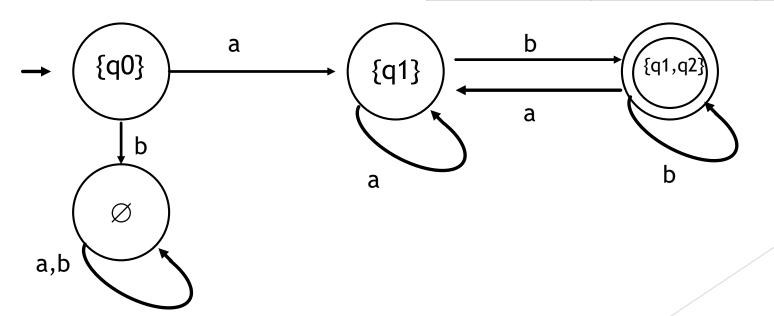
δ	a	b
→ q0	q1	
q1	q1	q1, q2
← q2		



Példa: L:= $\{u \in \{a,b\}^* \mid az \ u \ szó 'a' betűvel kezdődik és 'b' betűre végződik}$

δ	a	b
→ q0	q1	
q1	q1	q1, q2
← q2		

δ'	a	b
→ {q0}	{q1}	Ø
{q1}	{q1}	{q1,q2}
← {q1,q2}	{q1}	{q1,q2}
Ø	Ø	Ø



Kleene tétele

Tétel: $\mathcal{L}_3 = \mathcal{L}_{reg}$

Minden reguláris nyelvhez adható 3-as típusú grammatika, és fordítva minden 3-as típusú nyelv felépíthető az elemi reguláris nyelvekből a reguláris műveletek véges sokszori alkalmazásával.

Kleene tétele: $\mathcal{L}_3 = \mathcal{L}_{reg}$

Bizonyítás vázlat:

- 1. $\mathcal{L}_{reg} \subseteq \mathcal{L}_3$ korábbi előadáson láttuk
- 2. $\mathcal{L}_3 = \mathcal{L}_{VDA}$ előbb láttuk
- 3. $\mathcal{L}_{VDA} \subseteq \mathcal{L}_{reg}$

Kleene tétele: $\mathcal{L}_3 = \mathcal{L}_{reg}$

Bizonyítás vázlat:

3. $\mathcal{L}_{VDA} \subseteq \mathcal{L}_{reg}$ kiszámítható egy reguláris kifejezés Legyen A egy n állapotú VDA, azaz Q= $\{q_1,q_2,...,q_n\}$ és q_1 a kezdőállapot.

Kleene tétele bizonyítás folytatása:

 $\mathcal{L}_{VDA} \subseteq \mathcal{L}_{reg}$

Azt mondjuk, hogy $q_i u \underset{A}{\Rightarrow} * q_j$ redukció érinti a q_m állapotot, ha q_m előfordul a redukciós levezetés valamely közbülső lépésében.

Azt mondjuk, hogy a redukció k-megszorított, ha a redukció csak 1 és k közötti indexű állapotot érint közbülső lépésként.

Kleene tétele bizonyítás folytatása:

$$\mathcal{L}_{VDA} \subseteq \mathcal{L}_{reg}$$

```
E_{i,j}^k := \{ u \in T^* \mid \exists \ q_i u \underset{A}{\Rightarrow^*} \ q_j \ k\text{-megszoritott redukció} \}, ahol 0 \le k \le n és 1 \le i,j \le n. E_{i,j}^0 := \{ a \in T \mid \exists \ q_i a \longrightarrow q_j \ \text{állapotátmenet} \}, ahol i \ne j. E_{i,i}^0 := \{ \epsilon \} \cup \{ \ a \in T \mid \exists \ q_i a \longrightarrow q_i \ \text{állapotátmenet} \}
```

$$E_{i,j}^{k} = E_{i,j}^{k-1} \cup E_{i,k}^{k-1} (E_{k,k}^{k-1})^{*} E_{k,j}^{k-1}$$

Kleene tétele bizonyítás folytatása:

$$\mathcal{L}_{VDA} \subseteq \mathcal{L}_{reg}$$

Legyen I:={elfogadó állapotok indexei}.

$$L(A) = \bigcup_{i \in I} E_{1,i}^n$$

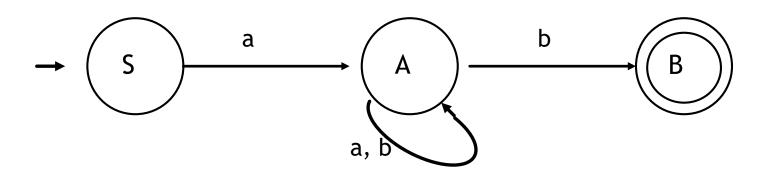
Megjegyzés: A kifejezésekben csak reguláris műveletek (unió, konkatenáció, lezárás) szerepeltek.

Példa: L= $\{a\}\{a,b\}^*\{b\}$ / reguláris kifejezéssel: a(a|b)*b /

 $S \rightarrow aA$

 $A \rightarrow aA \mid bA \mid bB$

 $B \rightarrow \epsilon$

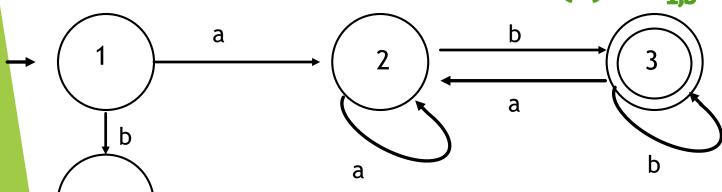


Példa: L={a}{a,b}*{b}

a,b

δ'	a	b
→ 1	2	4
2	2	3
← 3	2	3
4	4	4

$$\mathsf{L}(\mathsf{A}) = E_{1,3}^4 \\ E_{1,3}^4 = E_{1,3}^3 \cup E_{1,4}^3 (E_{4,4}^3)^* E_{4,3}^3; \ E_{4,3}^3 = \varnothing \\ E_{1,3}^3 = E_{1,3}^2 \cup E_{1,3}^2 (E_{3,3}^2)^* E_{3,3}^2; \\ E_{1,3}^2 = E_{1,3}^1 \cup E_{1,2}^1 (E_{1,2}^1)^* E_{1,3}^1; \\ E_{1,3}^1 = E_{1,3}^0 \cup E_{1,1}^0 (E_{1,1}^0)^* E_{1,3}^0; \ E_{1,3}^0 = \varnothing \\ E_{1,3}^2 = \varnothing \cup \{a\}\{a\}^* \{b\}; \\ E_{3,3}^2 = E_{3,3}^1 \cup E_{3,2}^1 (E_{2,2}^1)^* E_{2,3}^1; \\ E_{3,3}^2 = \{\varepsilon, b\} \cup \{a\}\{a\}^* \{b\}; \\ \mathsf{L}(\mathsf{A}) = E_{1,3}^4 = E_{1,3}^3 = \{a\}\{a\}^* \{b\} \{\{a\}^* \{b\}\}^* \}$$



R=aa*b(a*b)*

Minimális véges determinisztikus automata

Definíció:

Az A véges determinisztikus automata minimális állapotszámú, ha nincs olyan A' véges determinisztikus automata, amely ugyanazt a nyelvet ismeri fel, mint A, de A' állapotainak száma kisebb, mint A állapotainak száma.

Tétel:

Az L reguláris nyelvet felismerő minimális véges determinisztikus automata az izomorfizmus erejéig egyértelmű.

Minimális véges determinisztikus automata

Tétel:

Az L reguláris nyelvet felismerő minimális véges determinisztikus automata (VDA) az izomorfizmus erejéig egyértelmű.

Bizonyítás lépései:

- 1. Automata összefüggővé tétele
- 2. Ekvivalens állapotok meghatározása

Összefüggő véges determinisztikus automata

Definíció:

Az A = (Q, T, δ , q₀, F) véges determinisztikus automata q állapotát **elérhetőnek** mondjuk, ha $\exists u \in T^*$ szó, hogy q₀u \Rightarrow_A^* q.

(Gráfos ábrázolásban ez azt jelenti, hogy van irányított út q_0 -ból q-ba.)

Definíció:

Az A = (Q, T, δ , q₀, F) véges determinisztikus automatát **összefüggőnek** mondjuk, ha minden állapota elérhető a kezdőállapotból.

Összefüggő véges determinisztikus automata

Elérhető állapotok meghatározása:

Legyen $H_0 = \{q_0\}$,

 $H_{i+1}=H_i \cup \{r \in \mathbb{Q} \mid \delta(q,a)=r, q \in H_i, a \in \mathbb{T}\} \text{ \'es } i \ge 0.$

 $\exists \ k \ge 0 : H_k = H_m$, ahol $m \ge k$. Legyen $H = H_k$.

H halmaz tartalmazza az elérhető állapotokat.

A' legyen az A azon részautomatája, ahol Q'=H.

(A Q\H nemelérhető állapotok elhagyhatók.)

Definíció:

 $\mathbf{q} \sim \mathbf{p}$ (q és r ekvivalens állapotok), ha \forall u \in T* szóra igaz, hogy qu \Rightarrow * r és pu \Rightarrow * r' esetén r \in F akkor és csak akkor, ha r' \in F.

(Minden szóra igaz, hogy az automatát q-ból indítva vagy p-ből indítva, vagy mind kettő esetben elfogadja a szót, vagy mind kettő esetben elutasítja.)

Állítás: Ha q és p ekvivalens, akkor qa \rightarrow s és pa \rightarrow t esetén s és t is ekvivalens állapotok \forall a \in T betűre.

Definíció:

```
\mathbf{q} \sim^i \mathbf{p} (q és r i-ekvivalens állapotok),
ha \forall u\inT* szóra,ahol \ell(\mathbf{u}) \leq \mathbf{i} igaz, hogy
\mathbf{q} u \underset{A}{\Rightarrow} * r és pu \underset{A}{\Rightarrow} * r' esetén r\inF akkor és csak akkor, ha r'\inF.
```

(Legfeljebb i hosszú szavak esetén a két állapot nem megkülönböztethető.)

Lemma: $q \sim^{i+1} p$ akkor és csak akkor, ha $\forall a \in T$ -re $qa \rightarrow r$ és $pa \rightarrow t$ esetén $r \sim^i t$.

Tegyük fel, hogy az A automata összefüggő.

 $q \sim^0 p$ akkor és csak akkor, ha $q,p \in F$ vagy $q,p \in Q \setminus F$,

Azaz első lépésben két partícióra osztjuk az állapotokat, az elfogadó és a nem elfogadó állapotokra.

 $Q = B_1 \cup B_2$, ahol $B_1 := F$ és $B_2 := Q \setminus F$.

Csak az ϵ szó hossza nulla. A B_1 -ben szereplő állapotok elfogadják az üres szót, B_2 -ben szereplők pedig elutasítják.

Finomítsuk a szavak hossza szerint a partíciókat:

ha q \sim^i p és az állapotok k \geq 2 partícióra vannak osztva, Q = B₁U ... U B_k, akkor p és q pontosan akkor maradnak együtt, ha \forall a \in T-re qa \rightarrow r és pa \rightarrow t esetén r és t ugyanabba a B_i partícióba tartoznak, egyébként B_i-t szétbontjuk.

Ha p és q \forall a \in T-re az előző lépés partíciói szerint mindig ugyanoda képeznek, akkor q \sim^{i+1} p.

Ezt az eljárást addig ismételjük, amíg változás van.

Minimális automata megadása

```
A' = (Q', T, \delta', q_0', F')

Q'= {az előző eljárással nyert B_i partíciók}

q_0'= a q_0-t tartalmazó partíció.

F' =az F-ből keletkezett partíciók.

\delta'(B_i, a)= B_j, ha \delta(q, a)=p és q \in B_i és p \in B_j.
```

 $A = <\{1,2,3,4,5,6,7,8,9\}, \{a,b\}, \delta, 1, \{2,4,9\} >$

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

Elérhető állapotok:

$$H_0 = \{1\}$$

$$H_1 = \{1\} \cup \{3,6\} = \{1,3,6\}$$

$$H_2 = \{1,3,6\} \cup \{3,6\} \cup \{3,5\} \cup \{8,7\} = \{1,3,5,6,7,8\}$$

$$H_3 = \{1,3,5,6,7,8\} \cup \{2,9\} = \{1,2,3,5,6,7,8,9\}$$

$$H_4 = H_3$$

Nem elérhető állapot: 4. (Elhagyható.)

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

Particiónálás:

$$A = \{1,3,5,6,7,8\}, B=\{2,9\}$$

	a	b
1	Α	Α
3	Α	Α
5	В	Α
6	Α	Α
7	Α	Α
8	В	А

	a	b
2	В	Α
9	В	Α

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

$$C=\{1,3,6,7\}, D=\{5,8\}$$

Particiónálás:

$$B=\{2,9\}, C=\{1,3,6,7\}, D=\{5,8\}$$

	a	b
1	С	С
3	С	D
6	D	С
7	D	С

	a	b
2	В	D
9	В	D

	a	b
5	В	D
8	В	С

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

Particiónálás:

	a	b
6	I	G
7	I	Е

	a	b
2	В	Н
9	В	Н

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

$$J={6}, K={7}$$

Particiónálás:

	a	b
2	В	Н
9	В	Н

Nincs változás.

A 2,9 állapotok összevonhatók.

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

A' =< $\{1,29,3,5,6,7,8\}$, $\{a,b\}$, δ ', 1, $\{29\}$ > a minimális automata.

δ'	a	b
→ 1	3	6
←29	29	5
3	29 3	5
5	29	5
6	8	7
7		1
8	29	3

Köszönöm a figyelmet!