Az informatika számítástudományi alapjai
4. előadás

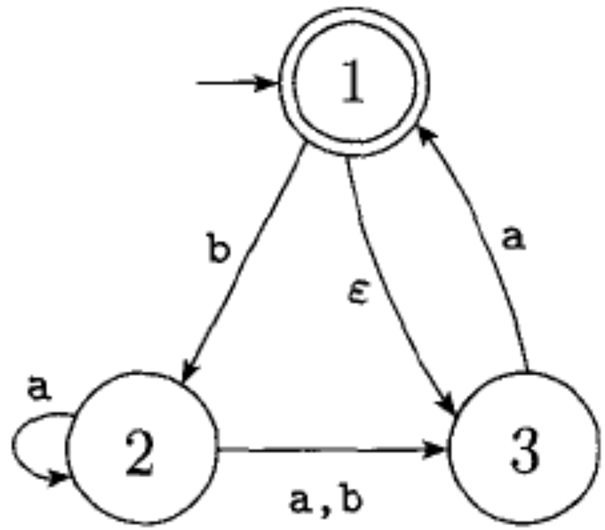
Vaszil György

vaszil.gyorgy@inf.unideb.hu

I. emelet 110-es szoba

A mai órán

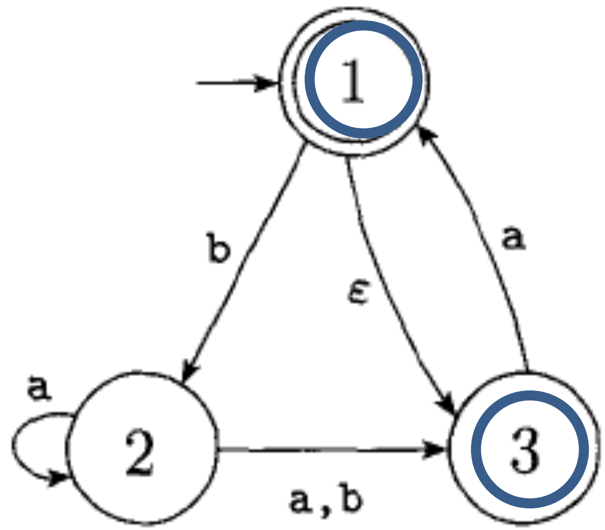
- Reguláris kifejezések, reguláris nyelvek
- Nemdeterminisztikus véges automaták
- A nemdeterminisztikusság kiküszöbölése – nemdeterminisztikus véges automaták determinisztikussá alakítása
- Reguláris műveletek és véges automaták
- Reguláris nyelvek és véges automatával elfogadható nyelvek



A bemenet:

abaa

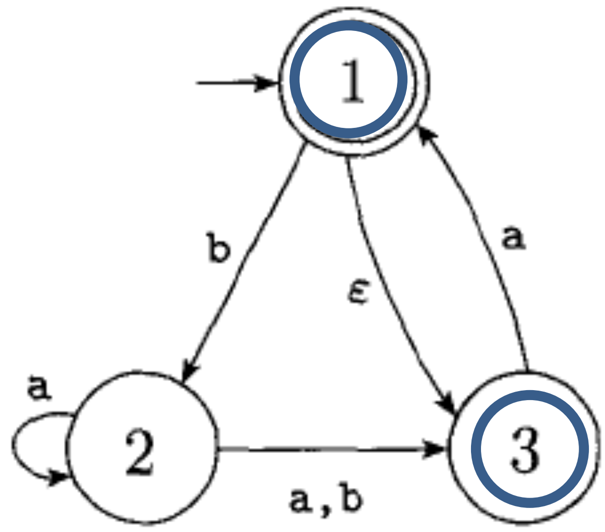
A „közös” kezdőállapot: (1,3)



A bemenet:

abaa

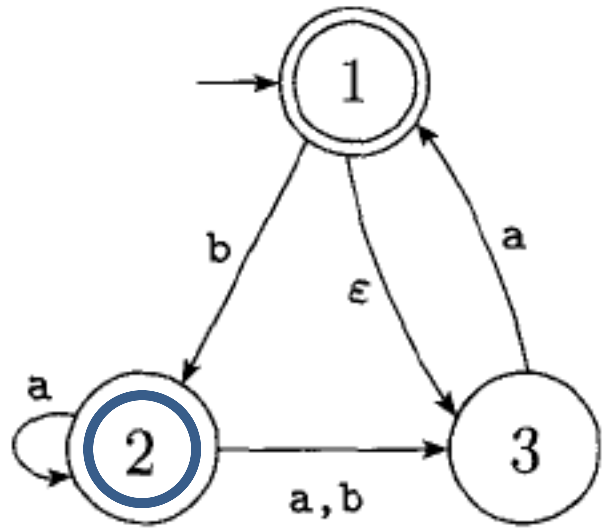
A „közös” kezdőállapot: (1,3)



A bemenet:

<u>a</u>*baa*

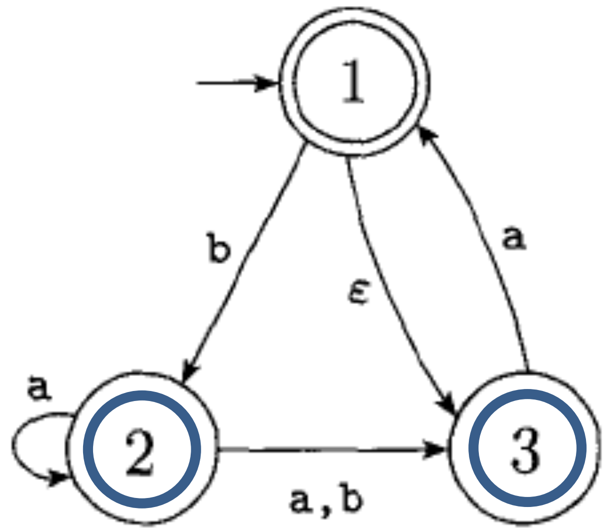
A „közös” állapot: (1,3)



A bemenet:

ab*aa*

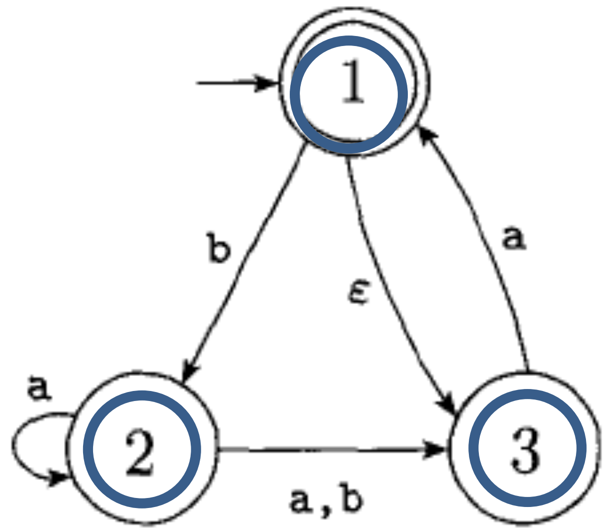
A „közös” állapot: (2)



A bemenet:

***aba**a*

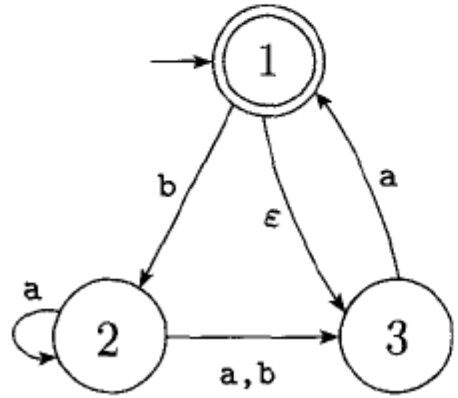
A „közös” állapot: (2,3)



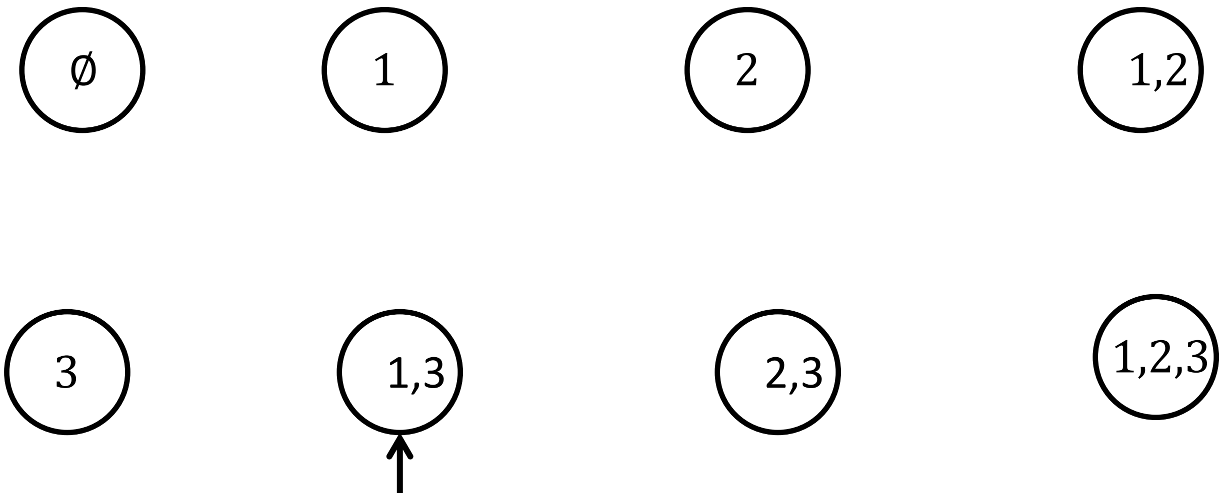
A bemenet:

<u>abaa</u>

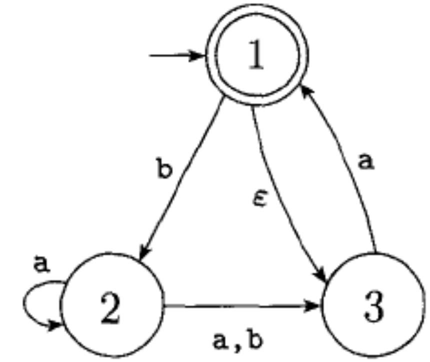
A „közös” állapot: (2,3,1)



Írjuk fel a „közös állapotokból” kapott automatát

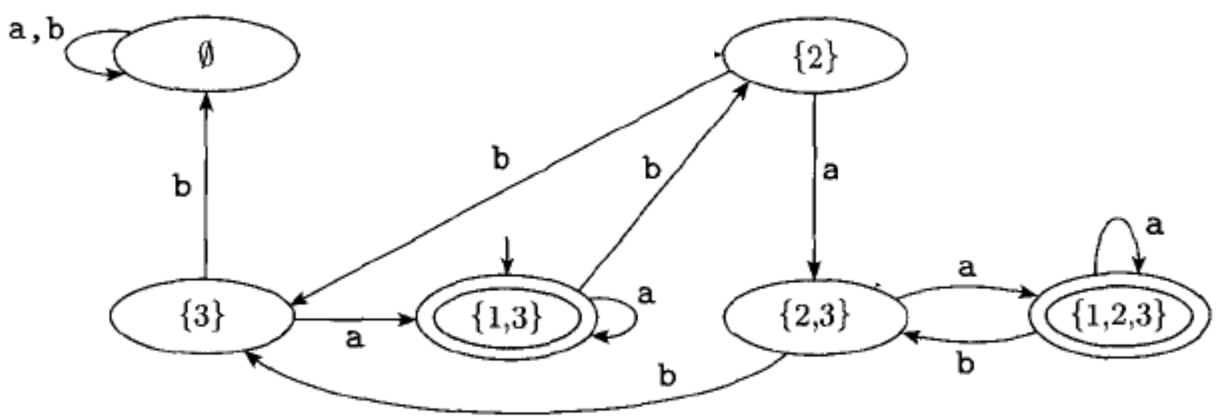


Például



$N = (\{1,2,3\}, \{a,b\}, 1, \{1\}, \delta)$

$M = (2^Q, \{a,b\}, E(\{1\}), \{\{1\}, \{1,2\}, \{1,3\}, \{1,2,3\}\}, \delta')$



($\{1\}$ és $\{1,2\}$ elhagyhatók)

A nemdeterminisztikusság kiküszöbölése

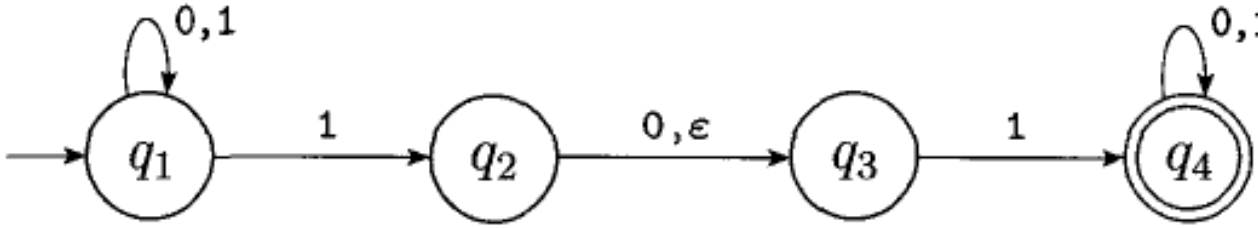
$N=(Q,\Sigma,q_0,A,\delta)$ nemdet $\rightarrow$ $M=(Q',\Sigma,q_0',A',\delta')$ det

- $Q' = 2^{Q}$ (Q részhalmazai)
- $\delta'(R,a) = \bigcup\limits_{r \in R} \delta^{*}(r,a)$, $R \subseteq Q$
- $q_0' = E(\{q_0\})$
- $A' = \{R \subseteq Q \mid R \cap A \neq \emptyset\}$

A konstrukció praktikusan

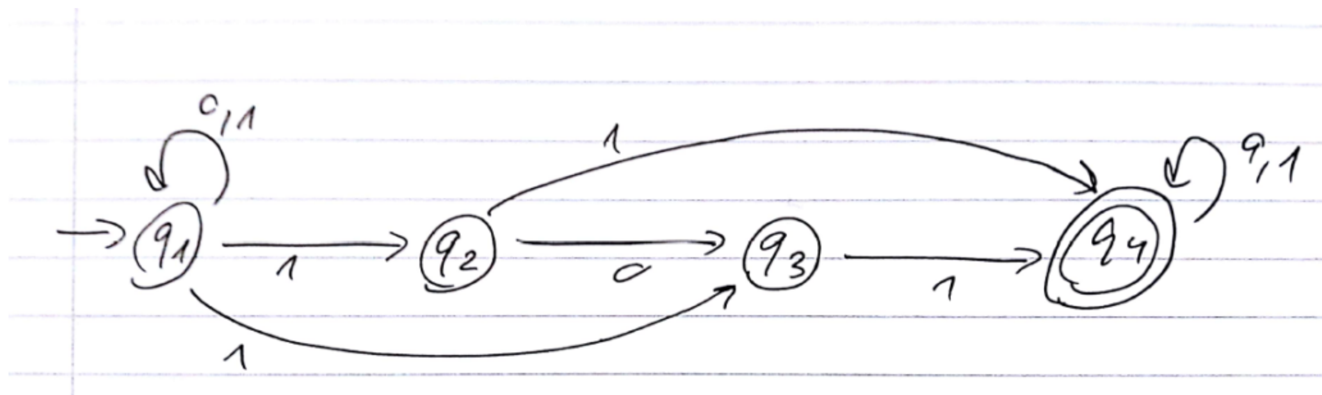
- Először írjuk fel az automatát „üresszó átmenetek” nélkül, és az így kapott (továbbra is nemdeterminisztikus, de üresszó átmenet nélküli) automatában végezzük el a „részhalmazos konstrukciót”.
- Ehhez minden $q \in Q$ és $a \in \Sigma$ értékre meg kell határozni $\delta^*(q,a)$-t.

Az előző példában



$$\delta^*(q_1, 0) = \{q_1\}$$
$$\delta^*(q_1, 1) = \{q_1, q_2, q_3\}$$
$$\delta^*(q_2, 0) = \{q_3\}$$
$$\delta^*(q_2, 1) = \{q_4\}$$
$$\delta^*(q_3, 0) = \emptyset$$
$$\delta^*(q_3, 1) = \{q_4\}$$
$$\delta^*(q_4, 0) = \{q_4\}$$
$$\delta^*(q_4, 1) = \{q_4\}$$

Az így kapott nemdetermininsztikus automata



A mai órán

- Reguláris kifejezések, reguláris nyelvek
- Nemdeterminisztikus véges automaták
- A nemdeterminisztikusság kiküszöbölése – nemdeterminisztikus véges automaták determinisztikussá alakítása
- Reguláris műveletek és véges automaták
- Reguláris nyelvek és véges automatával elfogadható nyelvek

Reguláris nyelvek és véges automaták

Definíció (szóhasználat):
Reguláris nyelvek $\leftrightarrow$ a reguláris kifejezésekkel leírható nyelvek

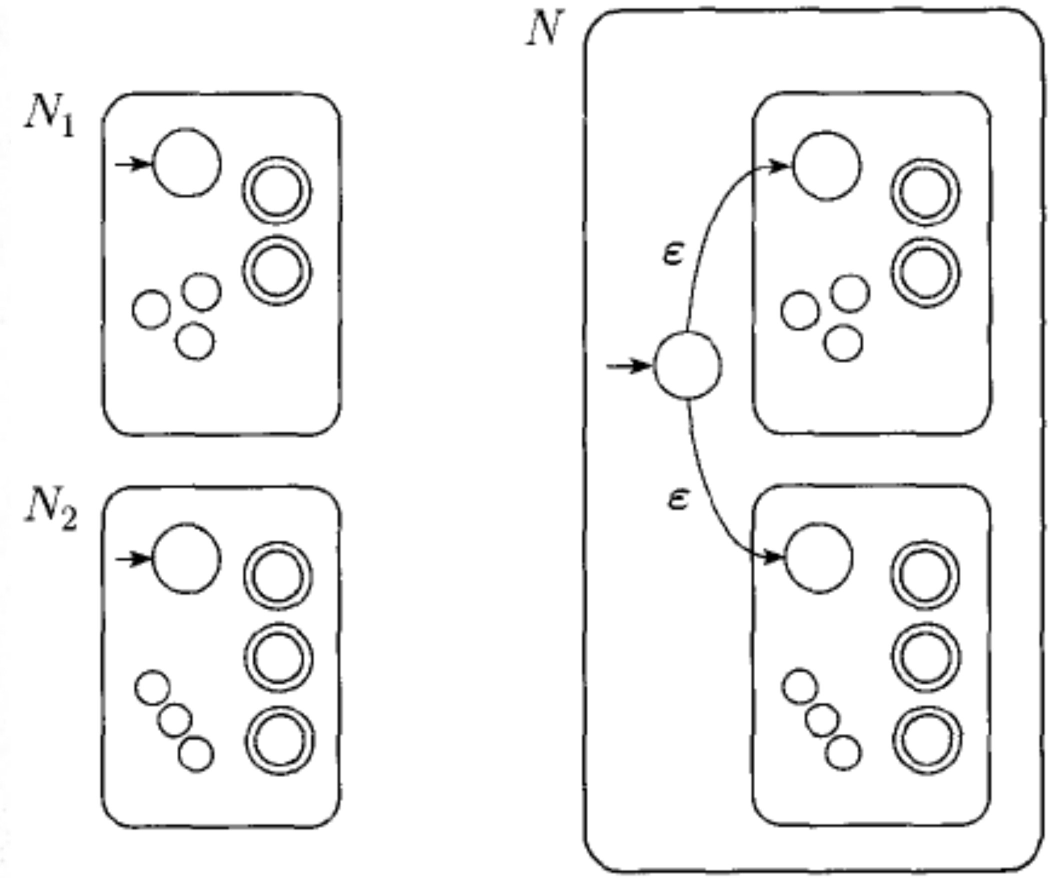
Mi a reguláris nyelvek és a véges automatával megadható nyelvek viszonya?

Kérdés: Mi a **reguláris nyelvek** és a **véges automtával megadható nyelvek** viszonya?

Válasz, 1. lépés: **Reguláris műveletek** és véges automaták

- unió – jele „+” vagy „$\cup$”
- konkatenáció – jele „$\cdot$” vagy semmi
- konkatenáció lezárása – jele „*”

Véges automatával elfogadott nyelvek uniójának elfogadója



$$L(N)=L(N_1) \cup L(N_2)$$

"Unió" nemdeterminisztikus automatával

$L(N_1) = A_1$, $L(N_2) = A_2$, konstruáljunk N-et, hogy $L(N) = A_1 \cup A_2$!

$N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ $\quad$ $N_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$

$$N = (Q_1 \cup Q_2 \cup \{q_0\}, \Sigma, \delta, q_0, F_1 \cup F_2)$$

ahol

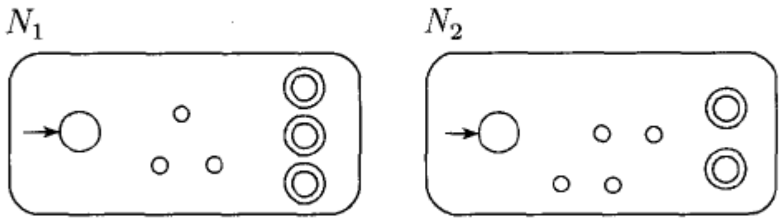
$$\delta(q_0, \varepsilon) = \{q_1, q_2\}$$

$$\delta(q, a) = \delta_1(q, a) \text{ ha } q \in Q_1$$

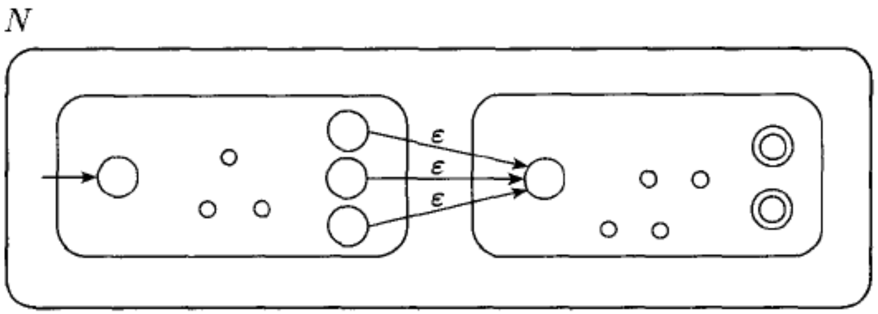
$$\delta(q, a) = \delta_2(q, a) \text{ ha } q \in Q_2$$

konkatenáció

$N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ $\quad$ $N_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$



$L(N)=L(N_1) \cdot L(N_2)$



$N = (Q, \Sigma, \delta, q_1, F_2)$

$Q = Q_1 \cup Q_2$

$\delta(q,a) = \delta_1(q,a)$ ha $q \in Q_1$

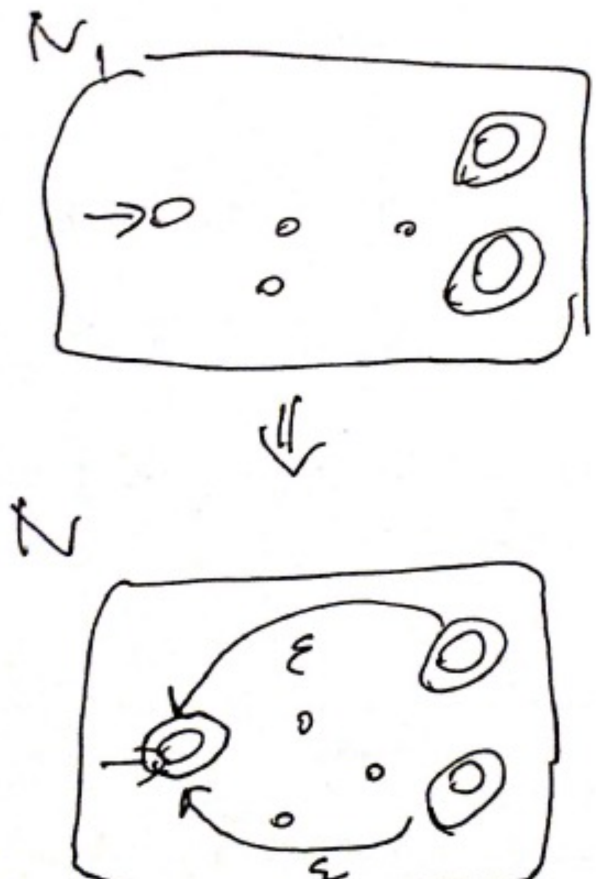
$\delta(q,\varepsilon) = \delta_1(q,\varepsilon) \cup \{q_2\}$ ha $q \in F_1$

$\delta(q,a) = \delta_2(q,a)$ ha $q \in Q_2$

Konkatenáció

lezárás (* - csillag)

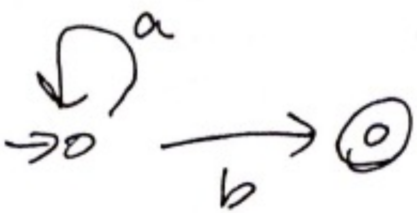
ötlet?



Mi ezzel a hiba?

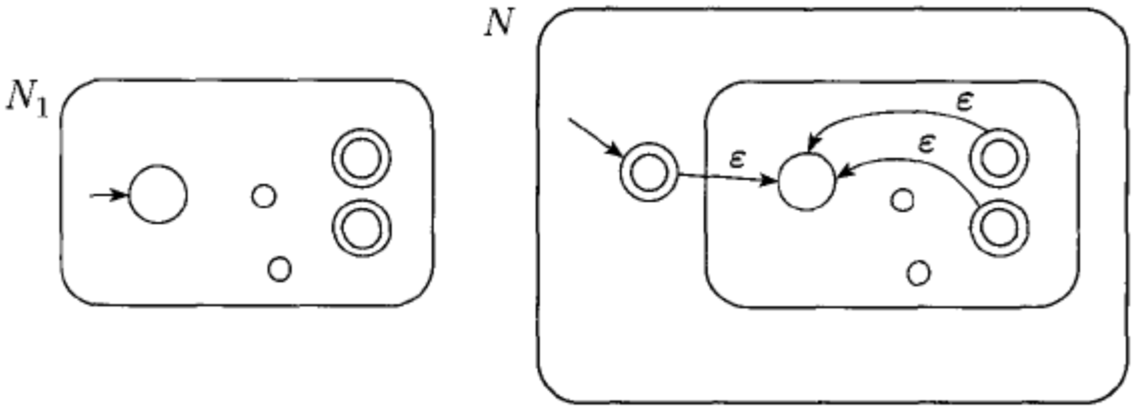
Mi a hiba az előző ötletben?

Például:

N_1 :  $L(N_1) = \{a^n b \mid n \geq 0\}$

Konkatenáció lezárása

$N_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$



$L(N)=L(N_1)^*$

$N = (Q, \Sigma, \delta, q_0, F)$

$Q = \{q_0\} \cup Q_1$

q_0 a kezdőállapot

$F = \{q_0\} \cup F_1$

$\delta(q,a) = \delta_1(q,a) \quad q \in Q_1$

$\delta(q,\varepsilon) = \delta_1(q,\varepsilon) \cup \{q_1\}$ ha $q \in F_1$

$\delta(q_0,\varepsilon) = \{q_1\}$

Tehát:

Ha L_1 nyelvet és L_2 meg lehet adni véges automatával, akkor az

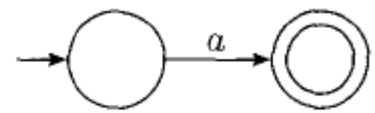
- $L_1 \cup L_2$ nyelvet is meg lehet adni véges automatával, az
- $L_1 \cdot L_2$ nyelvet is meg lehet adni véges automatával, és az
- $(L_1)^*$ és az $(L_2)^*$ nyelvet is meg lehet adni véges automatával.

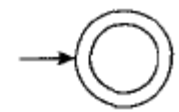
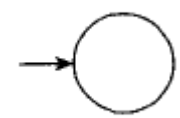
A mai órán

- Reguláris kifejezések, reguláris nyelvek
- Nemdeterminisztikus véges automaták
- A nemdeterminisztikusság kiküszöbölése – nemdeterminisztikus véges automaták determinisztikussá alakítása
- Reguláris műveletek és véges automaták
- Reguláris nyelvek és véges automatával elfogadható nyelvek

A reguláris kifejezések és a véges automaták ekvivalenciája

1. $\boxed{\text{reg. kif} \longrightarrow \text{véges automata}}$

- $R = a\ ,\ a \in \Sigma \longrightarrow$ 

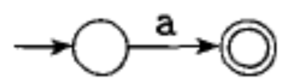
- $R = \varepsilon \longrightarrow$ 
- $R = \emptyset \longrightarrow$ .

Ezekből az előzőek szerint összerakható az automata
$R = R_1 \cup R_2$, $R = R_1 \bullet R_2$, $R = R_1^*$
kifejezésekhez.

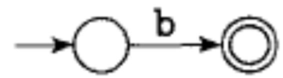
Például :

$$R = (ab \cup a)^*$$

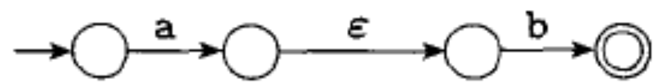
a



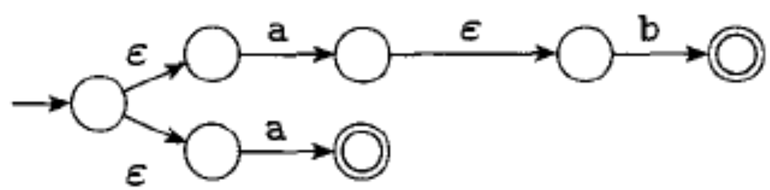
b



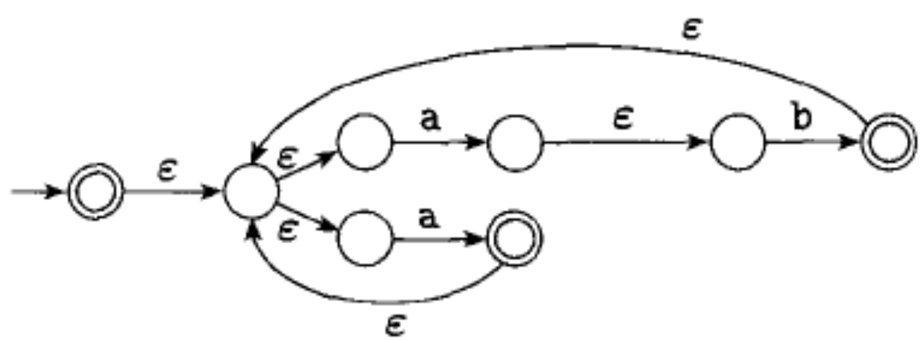
ab



ab ∪ a



(ab ∪ a)*

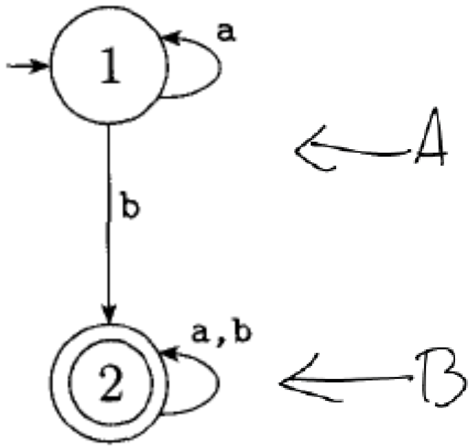


A reguláris kifejezések és a véges automaták ekvivalenciája

1. $\boxed{\text{reg. kif} \rightarrow \text{véges automata}}$
2. véges automata → reguláris kifejezésekkel felírt egyenletrendszer →

→reguláris kifejezés az elfogadott nyelvre

Milyen egyenletrendszer?



- Írjuk fel reguláris kifejezésüket az egyes állapotokból elfogadható nyelvekre
- Használjunk változókat

Például:

$$A = aA + bB$$
$$B = (a+b)B + \lambda$$

Ha A-t ki tudnánk változók nélkül fejezni, kész volnánk. (Miért?)

Hogyan oldjuk meg az egyenletrendszert?

észrevétel:

$$X = \alpha X + \beta \quad \text{megoldása} \quad X = \alpha^* \beta$$

Ha β-ban nem szerepel X, akkor eggyel kevesebb változónk van, hiszen $\alpha^*\beta$ behelyettesíthető X helyére a többi egyenletbe.

Az ... (lásd tábla)

A reguláris kifejezések és a véges automaták ekvivalenciája

1. $\boxed{\text{reg. kif} \longrightarrow \text{véges automata}}$

2. $\boxed{\text{véges automata}}$ ⟶ reguláris kifejezésekkel felírt egyenletrendszer ⟶

$\boxed{\longrightarrow \text{reg. kif}}$

A mai órán

- Reguláris kifejezések, reguláris nyelvek
- Nemdeterminisztikus véges automaták
- A nemdeterminisztikusság kiküszöbölése – nemdeterminisztikus véges automaták determinisztikussá alakítása
- Reguláris műveletek és véges automaták
- Reguláris nyelvek és véges automatával elfogadható nyelvek

A mai órán 2.

- Generatív grammatikák általában, végtelen nyelvek megadása generatív grammatikával
- Környezetfüggetlen grammatikák
- Reguláris nyelvek megadása környezetfüggetlen grammatikával, reguláris grammatikák
- Környezetfüggetlen grammatikák: lényegesen különböző levezetések, levezetési fák, grammatikák egyértelműsége
- Környezetfüggetlen grammatikák szabályainak egyszerűsítése, normálformák
  - törlő szabályok kiküszöbölése
  - láncszabályok kiküszöbölése

Nyelvek rekurzív definíciója

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

1. $\lambda \in L$,
2. Ha $S \in L$, akkor $aSb \in L$.

3. Másfajta szó nincs a nyelvben

Nyelvek rekurzív definíciója

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

1. $\lambda \in L$,
2. Ha $S \in L$, akkor $aSb \in L$.
3. Másfajta szó nincs a nyelvben

Ennek alapján átírási szabályokat adhatunk meg:

$$S \to \lambda \mid aSb$$

Másik példa: palindrómák

$\{a,b\}$ felett

abaabb | bbaba nem abbabba

1. λ palindróma
2. a, b palindróma
3. Ha x palindróma, akkor axa és bxb is palindróma
4. Másfajta szó nem palindróma

Milyen átírási szabályzatot implikál?

Generatív grammatika

- $G = (N, \Sigma, S, P)$
  - N: nemterminális ábécé
  - Σ terminális ábécé
  - $S \in N$ kezdőszimbólum
  - P . helyettesítési szabályok
- $L(G)$ a G által generált nyelv – szavak halmaza

Generatív grammatika

az első példa

- $G = (N, \Sigma, S, P)$
  - $N = \{S\}$ nemterminális ábécé $N = \{S\}$
  - $\Sigma = \{a, b\}$ terminális ábécé
  - $S \in N$ kezdőszimbólum
  - $P = \{S \Rightarrow \lambda \mid S \to aSb\}$ helyettesítési szabályok
- $L(G)$ az G által generált nyelv - szavak halmaza

Végtelen nyelvek megadása, generatív grammatikák

Egy Generatív grammatika alkotóelemei:

- terminális ábécé, a generálandó nyelv ábécéje
- nemterminális ábécé, segéd szimbólumok a generálás során
- kezdő nemterminális szimbólum
- helyettesítési szabályok, mely szavak helyettesíthetők mely más szavakkal a generálás során

Hogyan generálhatunk egy nyelvet?

$P = \{ S \to ab,\ S \to aSb \}$

(G=({S},{a,b},S,P) mint két diával korábban)

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow \dots \Rightarrow a^k S b^k \Rightarrow a^k ab b^k$$

kezdő szimbólum

helyettesítési lépések

addig folytatjuk a helyettesítést, amíg terminális betűkből álló szót kapunk

(mi a különbség az első példa és e között?)

A levezetés fogalma

formális leírásban

Adott $G=(N,\Sigma,S,P)$. Egy u szó közvetlenül levezethető egy v szóból (ahol $u,v \in (N \cup \Sigma)^*$)

ha
- $v = v_1 \alpha v_2$
- $u = v_1 \beta v_2$
- létezik $\alpha \to \beta \in P$ szabály

Jelölés: $v \Rightarrow u$

A levezetés fogalma

$G=(N,\Sigma,S,P)$ formális nyelvtan adott. Szavak egy $w_1, w_2, \ldots w_n$ sorozatára azt mondjuk, hogy <u>w_n levezethető w_1-ből</u>, ha:

$$w_1 \Rightarrow w_2 \Rightarrow \ldots \Rightarrow w_n$$

az előbbiek szerint.

<u>Jelölés:</u> $w_1 \Rightarrow^* w_n$

Például: $S \Rightarrow^* a^{21}b^{21}$ az előző levezetésben

$G=(N,\Sigma,S,P)$ <u>generálja</u> a w szót, ha

$$S \Rightarrow^* w$$

A $G=(N,\Sigma,S,P)$ grammatika által <u>generált nyelv</u>:

$$L(G)=\{w \in \Sigma^* \mid S \Rightarrow^* w\}$$

A mai órán 2.

- Generatív grammatikák általában, végtelen nyelvek megadása generatív grammatikával
- Környezetfüggetlen grammatikák
- Reguláris nyelvek megadása környezetfüggetlen grammatikával, reguláris grammatikák
- Környezetfüggetlen grammatikák: lényegesen különböző levezetések, levezetési fák, grammatikák egyértelműsége
- Környezetfüggetlen grammatikák szabályainak egyszerűsítése, normálformák
  - törlő szabályok kiküszöbölése
  - láncszabályok kiküszöbölése

A könyvekben

Nemdeterminisztikus véges automaták determinisztikussá alakítása:

- **J. Martin 3.3 fejezet, 104-110. oldal**
- **M. Sipser 1.2 fejezet, 54-58. oldal**
- **Bach I. 2.2 fejezet, 38-43. oldal**
- **Dömösi et al.: 5.4.1 fejezet, 87-92. oldal**

Reguláris műveletek és véges automaták, reguláris nyelvek és véges automatával elfogadható nyelvek

- **J. Martin: 3.4 fejezet („Kleene tétele"), 110-117. oldal**
- **M. Sipser: 66-76. oldal**
- **Dömösi et al.: 5.6 - 5.7 fejezet, 106-113. old**
- **Bach I. 2.6 fejezet, 71-78. oldal**

Generatív grammatikák általában, végtelen nyelvek megadása generatív grammatikával

- **J. Martin 4.1 fejezet**, 130 - 134. oldal
- **Bach I. 1.2 1 3 fejezet**, 14 - 20. oldal