Pumpáló lemma reguláris nyelvekre

más néven: Bar-Hillel-lemma (Yehoshua Bar-Hillel (1915–1975))

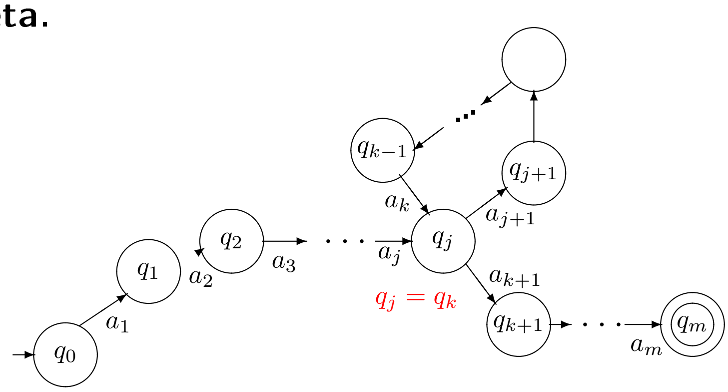
Tétel (pumpáló lemma). Bármely L reguláris nyelv esetében létezik olyan $n \geq 1$ természetes szám (amely csak L-től függ), hogy L bármely legalább n hosszúságú u szava felírható $u = xyz$ alakban úgy, hogy
(1) $|xy| \leq n$,
(2) $|y| \geq 1$,
(3) $xy^iz \in L$ **minden** $i = 0, 1, 2, \ldots$ **értékre.**

Ha L reguláris nyelv, akkor létezik olyan DVA, amely felismeri az L nyelvet): $\mathrm{A} = (Q, \Sigma, E, \{q_0\}, F)$, tehát $L = L(\mathrm{A})$.

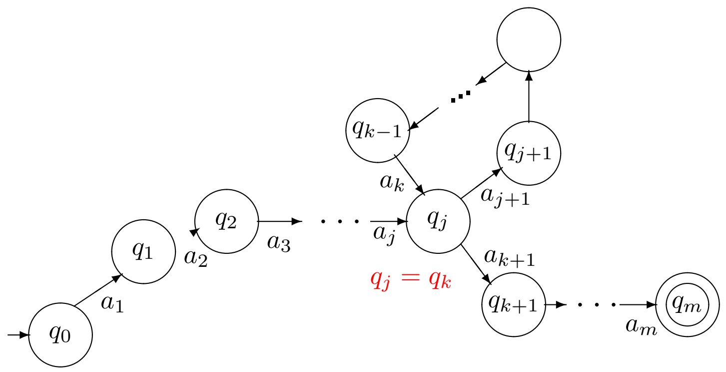
Legyen n az automata állapotainak száma, azaz $|Q| = n$. Legyen $u = a_1a_2\ldots a_m \in L$ és $m \geq n$. A determinisztikus véges automata felismeri az u szót $\Longrightarrow$ léteznek a $q_0, q_1, \ldots, q_m$ állapotok és a

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} \ \cdots \ \xrightarrow{a_{m-1}} q_{m-1} \xrightarrow{a_m} q_m, \ q_m \in F$$

séta.



Mivel csak n állapotunk van, és $m \geq n$, a skatulya-elv alapján a $q_0, q_1, \ldots, q_m$ állapotok között van legalább két megegyező.



Legyen $q_j = q_k$, ahol $j < k$ és k a legkisebb ilyen index. Ekkor $j < k \leq n$. Bontsuk fel az u szót a következőképpen:

$x = a_1 a_2 \dots a_j$
$y = a_{j+1} a_{j+2} \dots a_k$
$z = a_{k+1} a_{k+2} \dots a_m$.

Látszik, hogy $|xy| \leq n$ és $|y| \geq 1$. Bebizonyítjuk, hogy $xy^iz \in L$ tetszőleges i-re.

Mivel $u = xyz \in L$, létezik a

$$q_0 \xrightarrow{x} q_j \xrightarrow{y} q_k \xrightarrow{z} q_m, \quad q_m \in F \qquad \textbf{séta,}$$

és $q_j = q_k$ miatt felírható

$$q_0 \xrightarrow{x} q_j \xrightarrow{y} q_j \xrightarrow{z} q_m, \quad q_m \in F \qquad \textbf{alakban is.}$$

Ebből következik, hogy a $q_j \xrightarrow{y} q_j$ séta elhagyható vagy többször is beilleszthető. Tehát léteznek a következő séták:

$$q_0 \xrightarrow{x} q_j \xrightarrow{z} q_m, \quad q_m \in F,$$

$$q_0 \xrightarrow{x} q_j \xrightarrow{y} q_j \xrightarrow{y} \dots \xrightarrow{y} q_j \xrightarrow{z} q_m, \quad q_m \in F.$$

Ebből következik, hogy $xy^iz \in L$ tetszőleges i-re, és ezzel bebizonyítottuk a lemmát.

1. példa. Bebizonyítjuk, hogy $L_1 = \{a^k b^k \mid k \geq 1\}$ nem reguláris.

Tegyük fel, hogy L_1 reguláris, és legyen n a pumpáló lemma szerint az L_1-hez tartozó természetes szám.

Mivel az $u = a^n b^n$ szó hossza $2n$, ezért ez a szó is felbontható a lemmában megadott módon. Bebizonyítjuk, hogy ez ellentmondáshoz vezet.

Legyen a felbontás $u = xyz$. A lemma szerint ekkor $|xy| \leq n$, tehát x is és y is csak a-t tartalmazhatnak, és mivel $|y| \geq 1$, y legalább egy a-t tartalmaz.

Ekkor xy^iz, $i \neq 1$-re, különböző számú a-t és b-t tartalmaz, tehát $xy^iz \notin L_1$ tetszőleges $i \neq 1$ értékre. Ez ellentmond a lemma állításának, tehát az a feltevésünk, hogy L_1 reguláris, hamis. Tehát $L_1 \notin \mathcal{L}_3$.

Mivel a $G_1 = (\{S\}, \{a, b\}, \{S \to ab, S \to aSb\}, S)$ környezetfüggetlen nyelvtan L_1-et generálja, így $L_1 \in \mathcal{L}_2$. E két állításból rögtön következik, hogy $\mathcal{L}_3 \subset \mathcal{L}_2$.

2. példa. Bebizonyítjuk, hogy $L_2 = \{u \in \{0, 1\}^* \mid n_0(u) = n_1(u)\}$ nem reguláris. ($n_0(u)$ az u-ban szereplő nullák, $n_1(u)$ pedig az 1-esek számát jelenti).

Az előbbi példához hasonlóan járunk el az $u = 0^n 1^n$ szóval, ahol n most a pumpáló lemmában az L_2-höz tartozó természetes szám.

3. példa. Bebizonyítjuk, hogy $L_3 = \{uu \mid u \in \{a, b\}^*\}$ nem reguláris.

Legyen $w = a^n b a^n b = xyz$, ahol n itt is a pumpáló lemma szerinti L_3-hoz tartozó természetes szám. Mivel $|xy| \leq n$, következik, hogy y csak a betűket tartalmazhat, és legalább egyet tartalmaz is. De ekkor a lemma szerint $xz \in L_3$, ami lehetetlen. Tehát L_3 nem reguláris.

A pumpáló lemmának több érdekes következménye van.

• **Az L reguláris nyelv akkor és csakis akkor nem üres, ha létezik $u \in L$, $|u| < n$, ahol n a pumpáló lemmában az L-hez tartozó természetes szám.**

Az állítás egyik irányba nyilvánvaló: ha létezik n-nél rövidebb szó L-ben, akkor $L \neq \emptyset$. Fordítva, legyen $L \neq \emptyset$, és legyen u a legrövidebb szó L-ben. Megmutatjuk, hogy $|u| < n$. Ha ugyanis $|u| \geq n$, akkor alkalmazzuk a pumpáló lemmát, és azt kapjuk, hogy $u = xyz$, $|y| > 1$ és $xz \in L$. Ellentmondás, mivel $|xz| < |u|$, és u a legrövidebb L-beli szó. Tehát $|u| < n$.

• **Létezik olyan algoritmus, amely eldönti, hogy egy reguláris nyelv üres-e.**

Tegyük fel, hogy $L = L(\mathrm{A})$, ahol $\mathrm{A} = (Q, \Sigma, E, \{q_0\}, F)$ egy determinisztikus véges automata. L akkor és csakis akkor nem üres, ha tartalmaz n-nél rövidebb szót, ahol n az A automata állapotainak a száma. Következésképpen, elegendő azt eldönteni, hogy van-e olyan n-nél rövidebb szó, amelyet A elfogad. Mivel az n-nél rövidebb szavak száma véges, a kérdés algoritmikusan eldönthető.

Amikor a véges automaták elérhetetlen állapotainak meghatározására adtunk eljárást, akkor megjegyeztük, hogy az az eljárás használható annak eldöntésére is, hogy az automata által felismert nyelv üres-e. Mivel a véges automaták reguláris nyelveket ismernek fel, immár két eljárást ismerünk annak eldöntésére, hogy egy reguláris nyelv üres-e vagy sem. Sőt, van egy harmadik eljárásunk is, ha figyelembe vesszük, hogy a nem produktív állapotok kizárására szolgáló algoritmus is alkalmazható arra, hogy eldöntsük egy reguláris nyelvről, hogy üres-e vagy sem.

- Egy L reguláris nyelv akkor és csakis akkor végtelen, ha létezik $u \in L$ úgy, hogy $n \leq |u| < 2n$, ahol n a pumpáló lemmában az L-hez tartozó természetes szám.

Ha L végtelen, akkor tartalmaz $2n$-nél hosszabb szót, és legyen u a legrövidebb, de $2n$-nél hosszabb L-beli szó. Mivel L reguláris, alkalmazható rá a pumpáló lemma, tehát $u = xyz$, ahol $|xy| \leq n$, tehát $|y| \leq n$ is igaz. A lemma szerint $u' = xz \in L$. Mivel $|u'| < |u|$, és a legrövidebb, de $2n$-nél hosszabb L-beli szó u, kapjuk, hogy $|u'| < 2n$. Másrészt $|y| \leq n$ miatt $|u'| \geq n$ is teljesül.

Fordítva, ha létezik $u \in L$ úgy, hogy $n \leq |u| < 2n$, akkor alkalmazva rá a pumpáló lemmát, következik, hogy $u = xyz$, $|y| \geq 1$ és $xy^iz \in L$ tetszőleges i-re, tehát L végtelen.

Feltehetjük a kérdést, hogy alkalmazhatjuk-e a pumpáló lemmát egy véges nyelvre, hisz a pumpálással végtelen sok szót kapunk? A válasz abban rejlik, hogy egy véges L nyelvet felismerő bármely véges automata állapotainak száma nagyobb, mint L leghosszabb szavának a hossza. Ezért L-ben egyetlen szó sincs, amelynek a hossza legalább n, ahol n a pumpáló lemmában az L nyelvhez tartozó természetes szám. Tehát egyetlen L-beli szó sem bontható fel xyz alakban, ahol $|xyz| \geq n$, $|xy| \leq n$, $|y| \geq 1$, és ezért nem kaphatunk végtelen sok további L-beli szót.

Reguláris kifejezések

A következőkben tetszőleges Σ ábécé esetén bevezetjük a Σ feletti reguláris kifejezés és az általa jelölt nyelv fogalmát.

A reguláris kifejezés egy formula, míg az általa jelölt nyelv egy Σ feletti nyelv lesz.

Például, ha $\Sigma = \{a, b\}$, akkor az a^*, b^*, $a^* + b^*$ kifejezések Σ feletti reguláris kifejezések lesznek, amelyek rendre az $\{a\}^*$, $\{b\}^*$, $\{a\}^* \cup \{b\}^*$ nyelveket jelölik.

Rekurzívan értelmezzük a Σ feletti reguláris kifejezés és az általa jelölt nyelv fogalmát.
- **$\emptyset$ reguláris kifejezés és az üres nyelvet jelöli.**
- **ε reguláris kifejezés és az $\{\varepsilon\}$ nyelvet jelöli.**
- **Ha $a \in \Sigma$, a reguláris kifejezés és az $\{a\}$ nyelvet jelöli.**
- **Ha x, y reguláris kifejezések és az X, illetve Y nyelveket jelölik, akkor $(x + y)$, (xy), (x^*) is reguláris kifejezések és rendre az $X \cup Y$, XY és X^* nyelveket jelölik.**

Csak azok Σ feletti reguláris kifejezések, amelyeket a fenti szabályok véges sokszori alkalmazásával kapunk.

Egy reguláris kifejezésben bizonyos zárójeleket elhagyhatunk, amennyiben figyelembe véve a műveletek prioritási sorrendjét (iteráció, szorzat, egyesítés), nem változtatjuk meg az általa jelölt nyelvet. Például $((x^*)(x + y))$ helyett $x^*(x + y)$-t is írhatunk.

Két reguláris kifejezés ekvivalens, ha ugyanazt a nyelvet jelöli, azaz $x \equiv y$, ha $X = Y$, ahol X és Y rendre az x és y reguláris kifejezések által jelölt nyelvek.

Ekvivalens kifejezések

$$
\begin{aligned}
x + y &\equiv y + x \\
(x + y) + z &\equiv x + (y + z) \\
(xy)z &\equiv x(yz) \\
(x + y)z &\equiv xz + yz \\
x(y + z) &\equiv xy + xz \\
(x + y)^* \equiv (x^* + y)^* &\equiv (x + y^*)^* \equiv (x^* + y^*)^* \\
(x + y)^* &\equiv (x^*y^*)^* \\
(x^*)^* &\equiv x^* \\
x^*x &\equiv xx^* \\
xx^* + \varepsilon &\equiv x^*
\end{aligned}
$$

Megmutatjuk, hogy minden véges L nyelvhez megadható olyan x reguláris kifejezés, amely L-et jelöli. Ha $L = \emptyset$, akkor $x = \emptyset$. Ha $L = \{w_1, w_2, \ldots, w_n\}$, akkor $x = x_1 + x_2 + \ldots + x_n$, ahol minden $i = 1, 2, \ldots, n$ esetében x_i a $\{w_i\}$ nyelvet jelölő reguláris kifejezés. Ez utóbbit pedig a következőképpen adjuk meg. Ha $w_i = \varepsilon$, akkor $x_i = \varepsilon$. Különben, ha $w_i = a_1a_2 \ldots a_m$, ahol $m \geq 1$ függ i-től, akkor az $x_i = a_1a_2 \ldots a_m$, ahol elhagytuk a zárójeleket.

Stephen Cole Kleene (1909–1994)

Tétel (Kleene tétele). Az $L \subseteq \Sigma^*$ nyelv pontosan akkor reguláris, ha van olyan Σ feletti reguláris kifejezés, amely éppen L-et jelöli.

x reguláris kifejezés $\Longrightarrow$ L nyelv, amelyet x jelöl, szintén reguláris.
indukcióval:

Ha $x = \emptyset$, $x = \varepsilon$, $x = a, \forall a \in \Sigma$, akkor $L = \emptyset$, $L = \{\varepsilon\}$, $L = \{a\}$. Mivel L mindhárom esetben véges, ezért reguláris.

Ha $x = (x_1 + x_2)$, akkor $L = L_1 \cup L_2$, ahol L_1 és L_2 rendre az x_1 és x_2 reguláris kifejezések által jelölt nyelvek. Az indukciós feltevésünk értelmében L_1 és L_2 reguláris nyelvek, így L is az, mivel a reguláris nyelvek osztálya zárt az egyesítésre. Az $x = (x_1 x_2)$ és $x = (x_1^*)$ esetek bizonyítása hasonló.

L reguláris nyelv $\Longrightarrow$ x reguláris kifejezés, amely éppen az L nyelvet jelöli.

Ha L reguláris, akkor létezik egy $\mathrm{A} = (Q, \Sigma, E, \{q_0\}, F)$ determinisztikus véges automata, amelyre $L = L(\mathrm{A})$. Legyenek A állapotai $q_0, q_1, \ldots, q_n$.

Értelmezzük az R_{ij}^k nyelveket, minden $-1 \leq k \leq n$ és $0 \leq i, j \leq n$ értékekre.

R_{ij}^k azon szavak halmaza, amelyek hatására az A véges automata a q_i állapotból a q_j állapotba kerül úgy, hogy közben nem használja a k-nál nagyobb indexű állapotokat.

Az R_{ij}^k halmazokat formálisan is leírhatjuk:

$R_{ij}^{-1} = \{a \in \Sigma \mid (q_i, a, q_j) \in E\}$, **ha** $i \neq j$,
$R_{ii}^{-1} = \{a \in \Sigma \mid (q_i, a, q_i) \in E\} \cup \{\varepsilon\}$,

$R_{ij}^k = R_{ij}^{k-1} \cup R_{ik}^{k-1} \left(R_{kk}^{k-1}\right)^* R_{kj}^{k-1}$ **minden** $i, j, k \in \{0, 1, \ldots, n\}$ **értékre.**

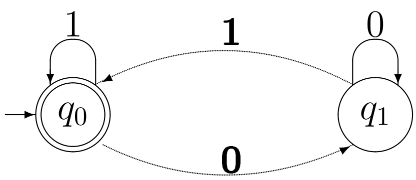
Indukcióval: R_{ij}^k halmazok leírhatók reguláris kifejezésekkel.

Ha $F = \{q_{i_1}, q_{i_2}, \ldots, q_{i_p}\}$ az A véges automata végállapotainak halmaza, akkor $L = L(\mathrm{A}) = R^n_{0i_1} \cup R^n_{0i_2} \cup \ldots \cup R^n_{0i_p}$ is megadható reguláris kifejezéssel az $R^n_{0i_1}, R^n_{0i_2}, \ldots, R^n_{0i_p}$ nyelveket jelölő reguláris kifejezésekből a $+$ művelet segítségével.

Reguláris kifejezés hozzárendelése véges automatához

1. módszer. **Felhasználjuk Kleene tételének az eredményét, azaz megkonstruáljuk az R^k_{ij} halmazokat, és felírjuk az $L = R^n_{0i_1} \cup R^n_{0i_2} \cup \ldots \cup R^n_{0i_p}$ nyelvet jelölő reguláris kifejezést, ahol $F = \{q_{i_1}, q_{i_2}, \ldots, q_{i_p}\}$ az automata végállapotainak halmaza.**

1. példa.



$$L(\mathrm{A}) = R^1_{00} = R^0_{00} \cup R^0_{01} \left(R^0_{11}\right)^* R^0_{10}$$

R^0_{00} : $\quad 1^* + \varepsilon \equiv 1^*$

R^0_{01} : $\quad 1^*0$

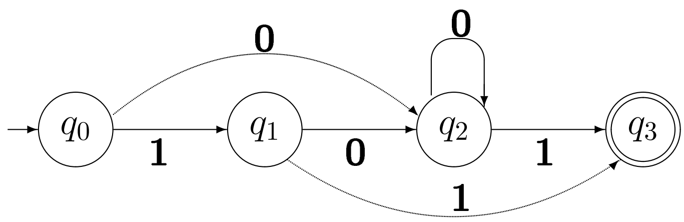
R^0_{11} : $\quad 11^*0 + \varepsilon + 0 \equiv (11^* + \varepsilon)0 + \varepsilon \equiv 1^*0 + \varepsilon$

R^0_{10} : $\quad 11^*$

Ekkor az $L(\mathrm{A})$-nak megfelelő reguláris kifejezés:

$$1^* + 1^*0(1^*0 + \varepsilon)^*11^* \equiv 1^* + 1^*0(1^*0)^*11^*.$$

2. példa.



	$k=-1$	$k=0$	$k=1$	$k=2$	$k=3$
R^k_{00}	ε	ε	ε	ε	
R^k_{01}	**1**	**1**	**1**	1	
R^k_{02}	**0**	0	$0+10$	$(0+10)0^*$	
R^k_{03}	$\emptyset$	$\emptyset$	11	$11+(0+10)0^*1$	$11+(0+10)0^*1$
R^k_{11}	ε	ε	ε	ε	
R^k_{12}	**0**	0	0	00^*	
R^k_{13}	**1**	1	1	$1+00^*1$	
R^k_{22}	$0+\varepsilon$	$0+\varepsilon$	$0+\varepsilon$	0^*	
R^k_{23}	**1**	1	1	0^*1	
R^k_{33}	ε	ε	ε	ε	

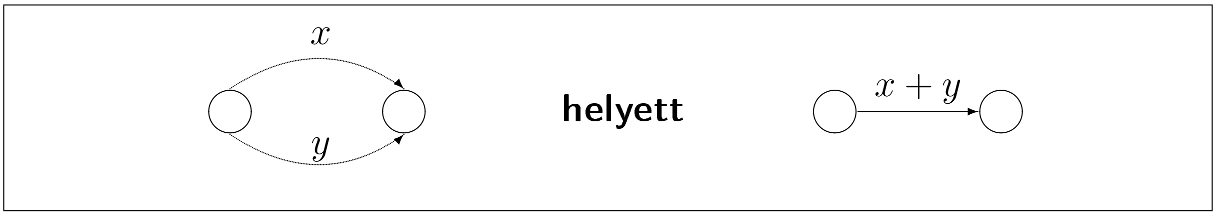
Az R^3_{03}-nak megfelelő reguláris kifejezés: $11+(0+10)0^*1$.

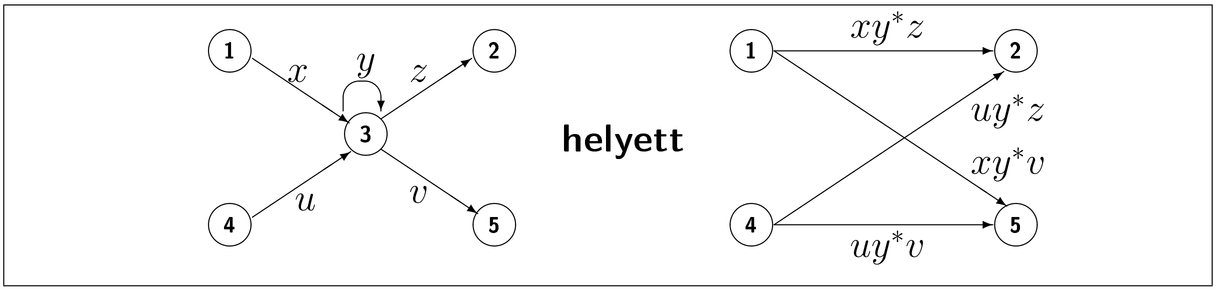
2. módszer. A véges automata fogalmát általánosítjuk úgy, hogy az automata gráfjának éleit nem betűkkel, hanem reguláris kifejezésekkel címkézzük meg.

Egy ilyen automatában minden séta meghatároz egy reguláris kifejezést, amely meghatároz egy reguláris nyelvet.

Az általánosított véges automata által felismert nyelven a produktív séták által meghatározott reguláris nyelvek egyesítését értjük. Könnyen belátható, hogy az ilyen általánosított véges automatákkal is éppen a reguláris nyelvek ismerhetők fel.

Az ekvivalens átalakítások a köv. ábrán láthatók. Amennyiben az ábrán látható 1, 2, 4, 5 csúcsok közül bármelyik kettő egybeesik, a végeredményben ezeket összevonjuk, így hurokél is megjelenik.

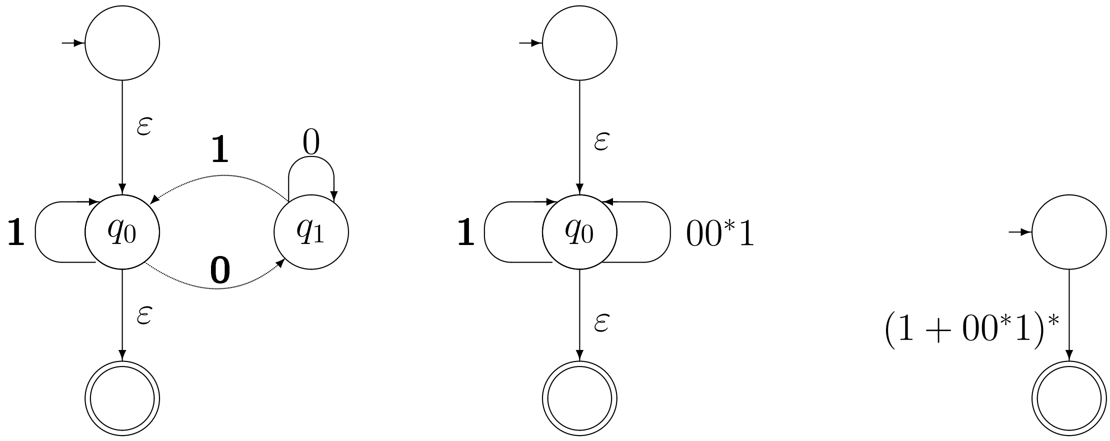




Először átalakítjuk a véges automatát megfelelő ε-átmenetek segítségével úgy, hogy egyetlen kezdő- és végállapota legyen.

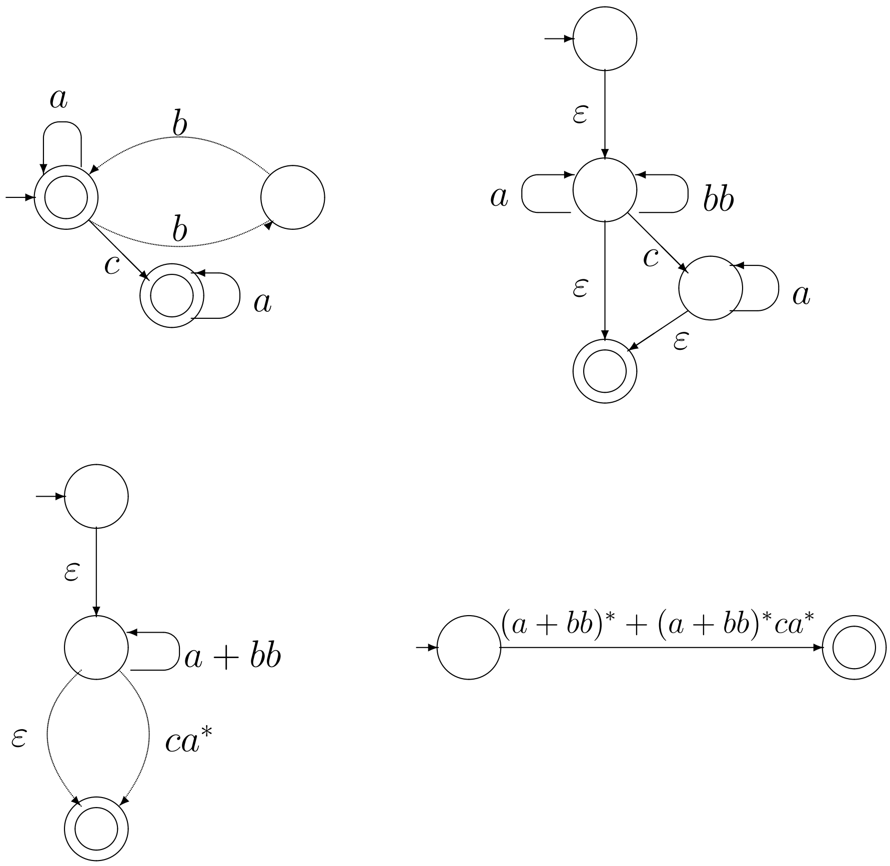
Ezután addig alkalmazzuk rá az ekvivalens átalakításokat, amíg gráfja egyetlen élt tartalmaz, amelynek címkéje az eredeti véges automata által felismert nyelvet jelölő reguláris kifejezés.

1. példa.

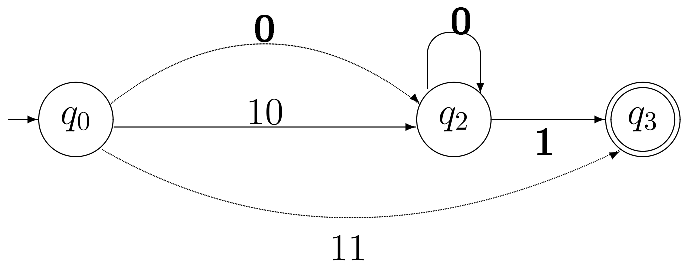


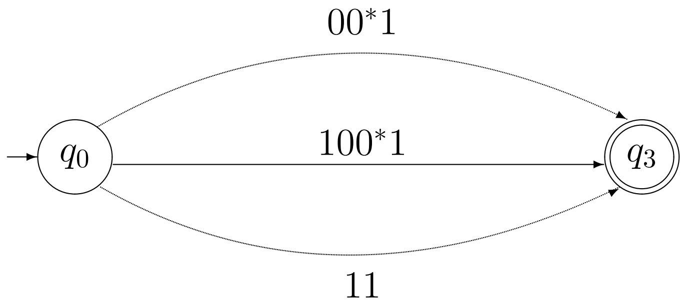
Az eredmény $(1+00^*1)^*$**, amely, habár más alakú, de ugyanazt a nyelvet jelenti, mint az előbbi módszerrel kapott kifejezés**

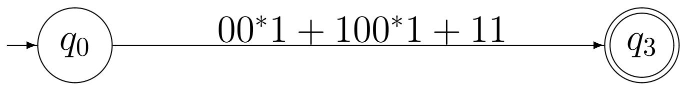
2. példa.



3. példa.







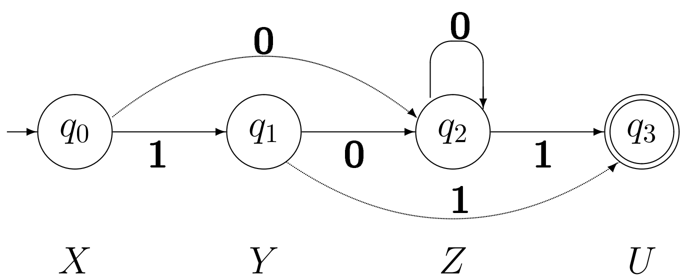
3. módszer.

Formális egyenletek módszere

Minden állapothoz hozzárendelünk egy X változót (különböző állapotokhoz különbözőt) és egy egyenletet, amelynek bal oldalán X, jobb oldalán pedig Ya alakú kifejezések összege vagy ε állhatnak, ahol Y is egy állapothoz rendelt változó, a pedig egy bemeneti szimbólum.

Ha az X változónak megfelelő állapotba nem vezet él, akkor az X baloldalú egyenlet jobb oldalán ε szerepel, különben az összes olyan Ya alakú tag összege, amelyekre teljesül, hogy a véges automata gráfjában az Y változónak megfelelő állapotból egy a-val címkézett él vezet az X változónak megfelelő állapotba. Amennyiben az X változónak megfelelő állapot kezdőállapot és egyben végállapot is, akkor az X baloldalú egyenlet jobb oldalán megjelenik egy ε tag is.

Például a köv. ábra esetében



legyenek ezek a változók X, Y, Z, U, amelyek a q_0, q_1, q_2, q_3 állapotoknak felelnek meg. A megfelelő egyenletek a következők:

$X = \varepsilon$
$Y = X1$
$Z = X0 + Y0 + Z0$
$U = Y1 + Z1.$

Ha egy egyenlet $X = X\alpha + \beta$ alakú, ahol α, β tetszőleges szavak, amelyek nem tartalmazzák az X változót, akkor könnyű ellenőrizni, egyszerű behelyettesítéssel, hogy $X = \beta\alpha^*$ megoldása az egyenletnek.

Mivel az egyenletek lineárisak a változókban, minden egyenlet felírható

$X = X\alpha + \beta$ vagy $X = X\alpha$ alakban,

ahol α nem tartalmaz egyetlen változót sem.

Ezt behelyettesítve a többi egyenletbe, eggyel csökkentjük azok számát. Így a rendszer, megfelelő helyettesítésekkel, megoldható minden változóra.

Az egyenletrendszert megoldva a végállapotoknak megfelelő változók adják a megoldást jelentő reguláris kifejezést úgy, hogy összeadjuk az ezen változóknak megfelelő reguláris kifejezéseket.

Példánkban

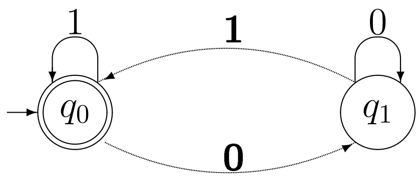
$$X = \varepsilon$$
$$Y = X1$$
$$Z = X0 + Y0 + Z0$$
$$U = Y1 + Z1.$$

a fenti egyenletrendszer első egyenletének segítségével azt kapjuk, hogy $Y = 1$.

Innen $Z = 0 + 10 + Z0$, azaz $Z = Z0 + (0 + 10)$, és ezt megoldva azt kapjuk, hogy $Z = (0 + 10)0^*$.

Innen pedig U egyszerűen megkapható: $U = 11 + (0 + 10)0^*1$.

Ezt a módszert alkalmazva a köv. ábra esetében



a következő egyenletekhez jutunk:

$X = \varepsilon + X1 + Y1$

$Y = X0 + Y0$

Kiemelés után:

$X = \varepsilon + (X + Y)1$

$Y = (X + Y)0.$

A két egyenletet összeadva a következő egyenlethez jutunk:

$X + Y = \varepsilon + (X + Y)(0 + 1)$**, ahonnan (**$\varepsilon$**-t** β**-nak,** $(0 + 1)$**-et** α**-nak tekintve) eredményül kapjuk a következőt:**

$X + Y = (0 + 1)^*$**.**

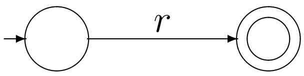
Innen – behelyettesítés után – megkapjuk X **értékét:**

$X = \varepsilon + (0 + 1)^*1$**,**

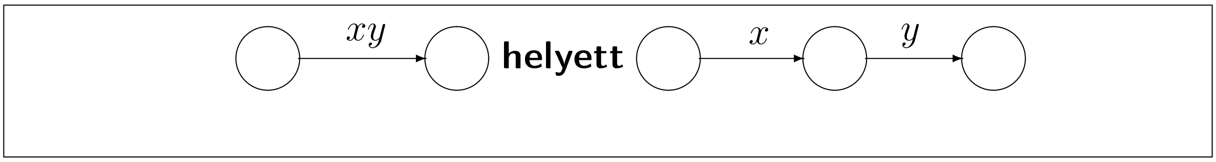
amely ekvivalens a másik módszerrel kapott kifejezéssel.

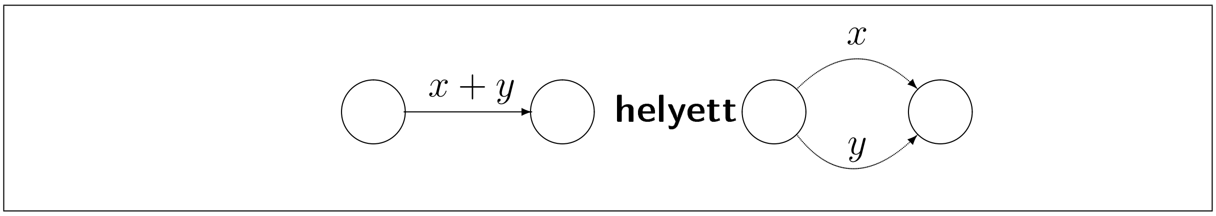
Véges automata hozzárendelése reguláris kifejezéshez

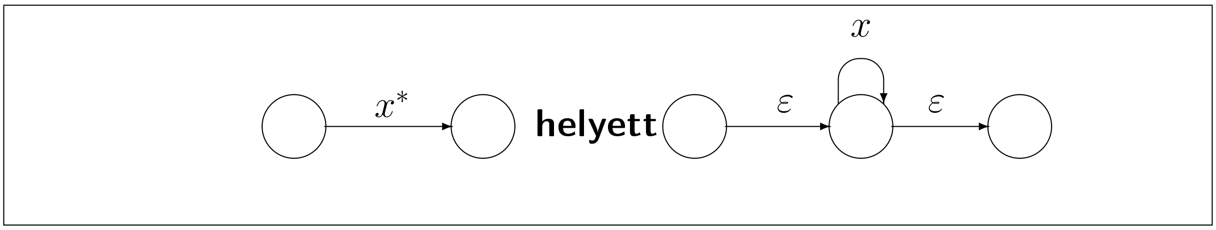
A r reguláris kifejezéshez hozzárendelünk egy általánosított véges automatát:



Azután lépésenként alkalmazzuk a köv. ábrán látható átalakításokat mindaddig, amíg a véges automata élei Σ elemeivel vagy ε-nal lesznek címkézve.

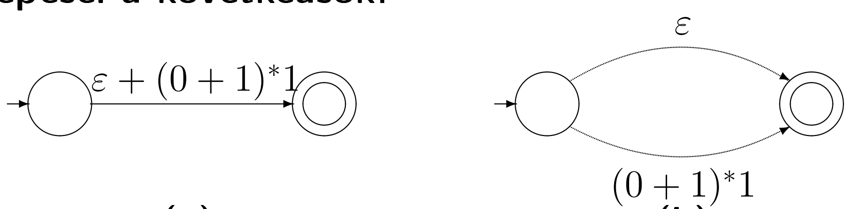




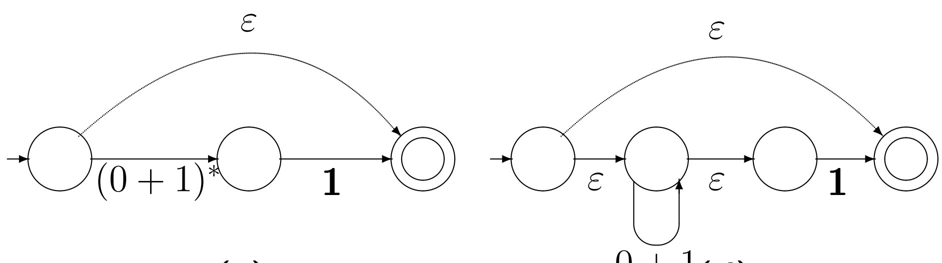


Példa.

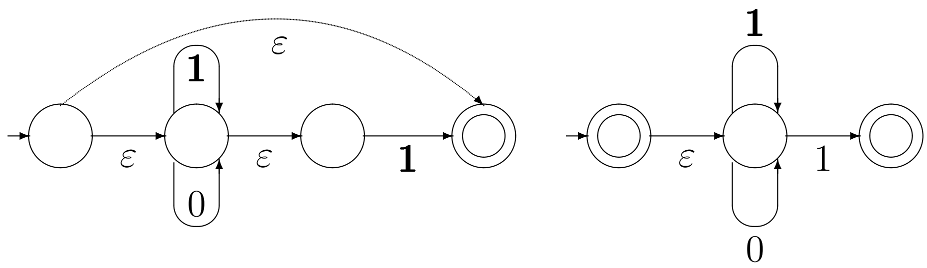
Induljunk el az $\varepsilon + (0+1)^*1$ reguláris kifejezésből. Az átalakítás lépései a követkeasők:



(a) **(b)**



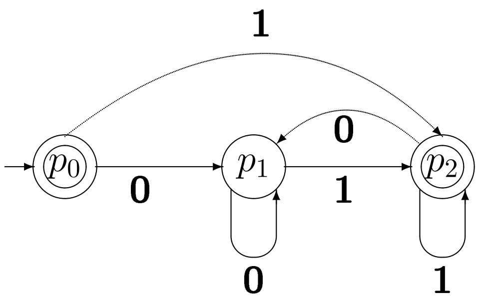
(c) **(d)**



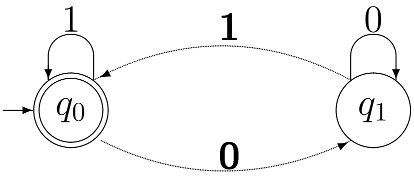
(e) **(f)**

Az utolsó előtti véges automata egyszerűbb alakban is megadható, ez az utolsó ábrán látható.

Ha ebből kiküszöböljük a ε-lépést, átalakítjuk determinisztikussá, akkor a köv. ábrán látható véges automatát kapjuk eredményül

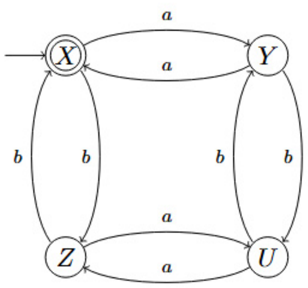


amelyről be lehet bizonyítani, hogy ekvivalens a köv. véges automatával



Példa.

Rendeljünk egy reguláris kifejezést az alábbi automatához.



$$X = \varepsilon + Ya + Zb$$
$$Y = Xa + Ub$$
$$Z = Xb + Ua$$
$$U = Yb + Za$$

Behelyettesítjük az Y és Z változót az első és a negyedik egyenletbe:

$$X = \varepsilon + (Xaa + Uba) + (Xbb + Uab)$$
$$U = (Xab + Ubb) + (Xba + Uaa)$$

Átrendezés után:

$$X = \varepsilon + X(aa + bb) + U(ab + ba)$$
$$U = X(ab + ba) + U(aa + bb)$$

A második egyenletet (ahol U-t tekintve változónak $\alpha = aa + bb$ és $\beta = X(ab + ba)$) megoldjuk U szerint:

$$U = X(ab + ba)(aa + bb)^*$$

Ezt behelyettesítve az első egyenletbe:

$$X = \varepsilon + X(aa + bb) + X(ab + ba)(aa + bb)^*(ab + ba)$$

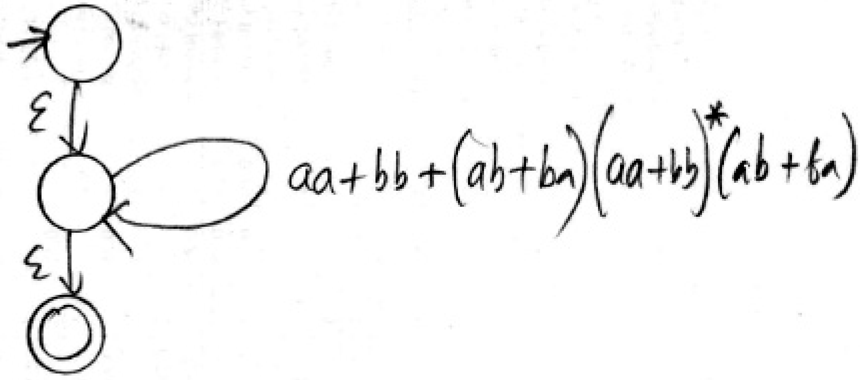
Innen

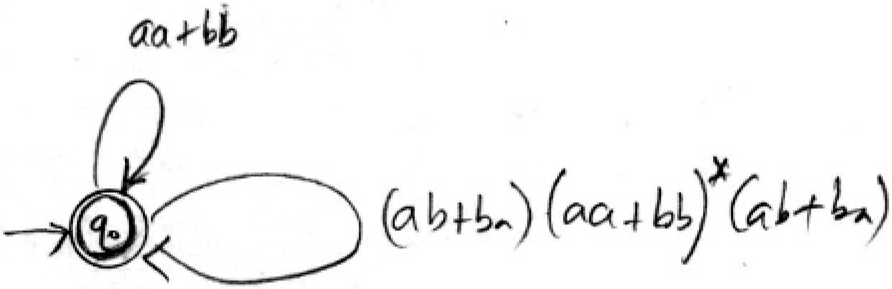
$$X = \varepsilon + X\Big(aa + bb + (ab + ba)(aa + bb)^*(ab + ba)\Big)$$

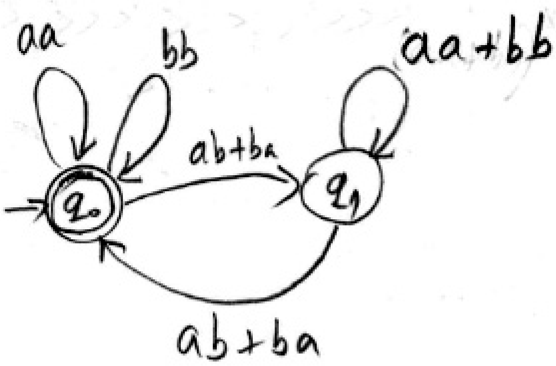
Megoldva:

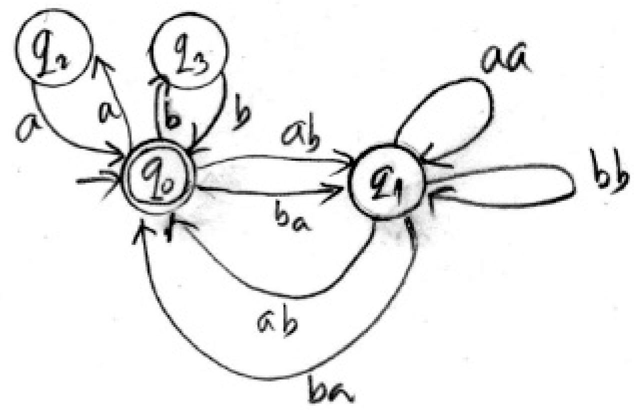
$$X = \Big(aa + bb + (ab + ba)(aa + bb)^*(ab + ba)\Big)^*$$

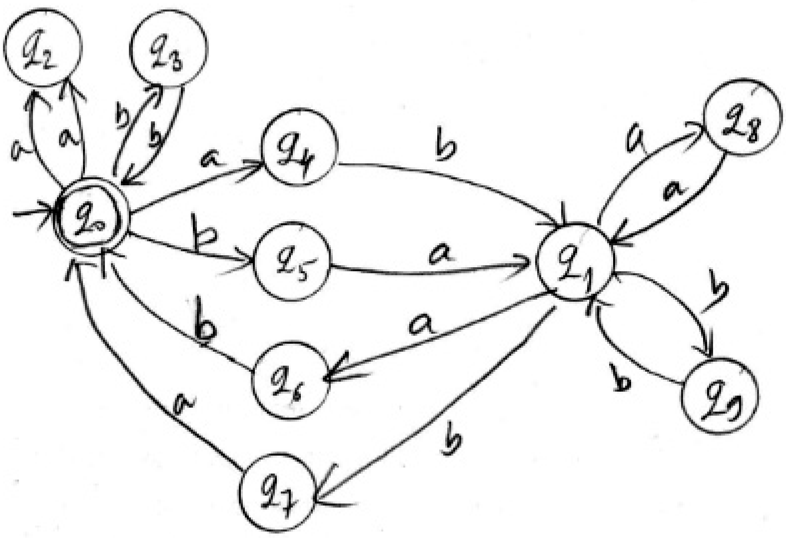
Oldjuk meg a fordított feladatot, azaz rendeljünk ehhez a reguláris kifejezéshez egy véges automatát, alkalmazva a megfelelő algoritmust.


ε
ε
aa+bb+(ab+ba)(aa+bb)*(ab+ba)


aa+bb
q0
(ab+ba)(aa+bb)*(ab+ba)


aa
bb
aa+bb
ab+ba
q0
q1
ab+ba





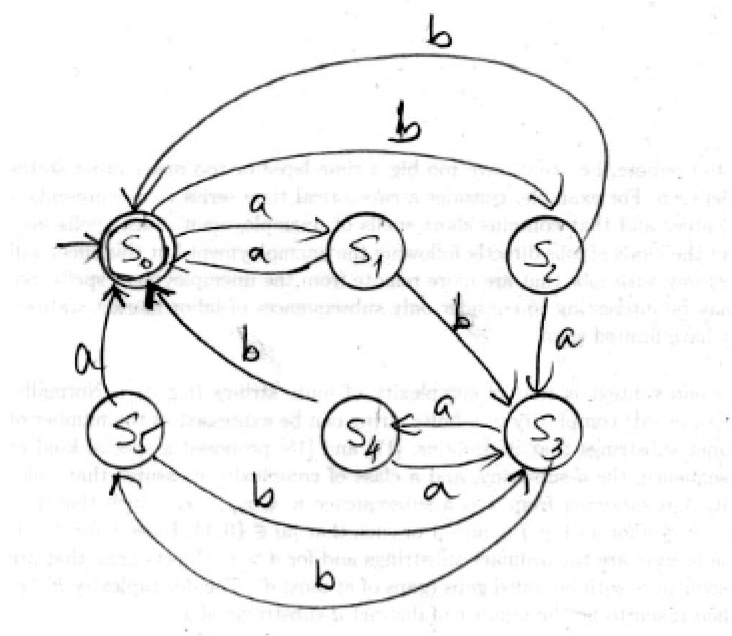
Rendeljünk ehhez az automatához egy vele ekvivalens determinisztikusat. A nemdeterminisztikus véges automata átmenettáblázata:

δ	a	b
q_0	$\{q_2, q_4\}$	$\{q_3, q_5\}$
q_1	$\{q_6, q_8\}$	$\{q_7, q_9\}$
q_2	$\{q_0\}$	$\emptyset$
q_3	$\emptyset$	$\{q_0\}$
q_4	$\emptyset$	$\{q_1\}$
q_5	$\{q_1\}$	$\emptyset$
q_6	$\emptyset$	$\{q_0\}$
q_7	$\{q_0\}$	$\emptyset$
q_8	$\{q_1\}$	$\emptyset$
q_6	$\emptyset$	$\{q_1\}$

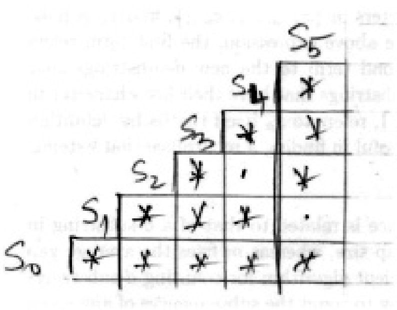
Innen kapjuk az ekvivalens determinisztikus véges automata átmenettáblázatát:

δ'	a	b
$S_0 = \{q_0\}$	$\{S_1\}$	$\{S_2\}$
$S_1 = \{q_2, q_4\}$	$\{S_0\}$	$\{S_3\}$
$S_2 = \{q_3, q_5\}$	$\{S_3\}$	$\{S_0\}$
$S_3 = \{q_1\}$	$\{S_4\}$	$\{S_5\}$
$S_4 = \{q_6, q_8\}$	$\{S_3\}$	$\{S_0\}$
$S_5 = \{q_7, q_9\}$	$\{S_0\}$	$\{S_3\}$

és az átmenetgráfot:



Ezt az automatát minimalizáljuk.



Először megcsillagozunk minden olyan állapotpárt, amelynek egyik tagja az S_0. Majd sorra vizsgáljuk a következő párokat:

$\{S_1, S_2\}$ a-**ra:** $\{S_0, S_3\}$, **amely meg van csillagozva, tehát** $\{S_1, S_2\}$-**t is megcsillagozzuk.**

$\{S_1, S_3\}$ a-**ra:** $\{S_0, S_4\}$ $*$

$\{S_1, S_4\}$ a-**ra:** $\{S_0, S_3\}$ $*$

$\{S_1, S_5\}$ a-**ra:** $\{S_0, S_0\}$
b-**re:** $\{S_3, S_3\}$ **egy ponttal jelöljük, hogy megvizsgáltuk, de nem tudtuk megcsillagozni**

$\{S_2, S_3\}$ a-**ra:** $\{S_3, S_4\}$
b-**re:** $\{S_0, S_5\}$ $*$

$\{S_2, S_4\}$ a-**ra:** $\{S_3, S_3\}$
b-**re:** $\{S_0, S_0\}$ **egy ponttal jelöljük, hogy megvizsgáltuk**

$\{S_2, S_5\}$ a-**ra:** $\{S_3, S_0\}$ $*$

$\{S_3, S_4\}$ a-**ra:** $\{S_4, S_4\}$
b-**re:** $\{S_5, S_0\}$ $*$

$\{S_3, S_5\}$ a-**ra:** $\{S_4, S_0\}$ $*$

$\{S_4, S_5\}$ a-**ra:** $\{S_3, S_0\}$ $*$

Az eredmény a bal oldali ábra, mellette jobb oldalon az eredeti automata: