

Automaták és Formális Nyelvek

III. előadás

Kimenettel rendelkező véges automaták

Az ilyen típusú véges automaták a gyakorlati alkalmazásuk szempontjából fontosak:

Egy ilyen determinisztikus automatával ismerkedünk meg a továbbiakban, az ún. Moore automatával.

A Moore automatát formálisan egy $M=(Q, \Sigma, \Delta, q_0, \delta, \lambda)$ "hatosként" foghatjuk fel, ahol

Q -állapotainak a halmaza,

q_0 - a kezdő állapota,

Σ az ún. inputábécé (aminek a szavait beolvassuk)

$\delta: Q \times \Sigma \rightarrow Q$ pedig a (determinisztikus) tranzíciós függvény.

Δ -az ún. kimenő v. outputábécé - ennek betűiből előállított szavak lesznek azok az üzenetek amiket az automata kiír,

$\lambda: Q \rightarrow \Delta$ pedig egy leképezés ami az automata Q állapotainak a függvényeként a kimenő jeleket állítja elő, a következőképpen:

$$\begin{array}{ccccccc} q_0 & \xrightarrow{a_1} & q_1 & \xrightarrow{a_2} & q_2 & \cdots & \xrightarrow{a_n} & q_1 \\ \downarrow & & \downarrow & & \downarrow & & & \downarrow \\ \lambda(q_0) & & \lambda(q_1) & & \lambda(q_2) & \cdots & & \lambda(q_n) \end{array} \quad \delta(q_i, a_{i+1}) = q_{i+1}$$

A $\omega^* = \lambda(q_0)\lambda(q_1)\dots\lambda(q_n) \in \Delta^*$ - egy a Δ ábécé betűiből álló kimenő "üzenet" amit az automata által beolvasott $\omega = a_1a_2\dots a_n \in \Sigma^*$ szónak feleltetünk meg.

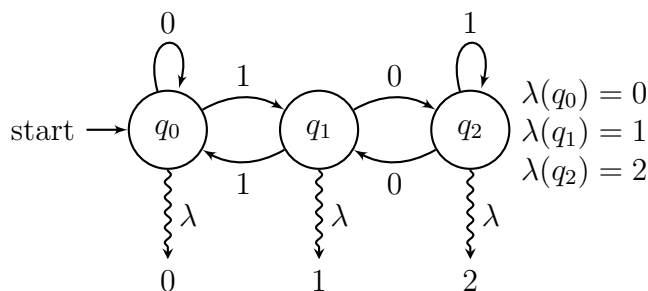
Megjegyezzük, hogy itt végállapotokat nem szükséges feltüntetnünk.

Példa: Legyen $\Sigma = \{0,1\}$, $\Delta = \{0,1,2\}$. Az alábbi Moore automata egy $\omega \in \Sigma^*$ szónak megfelelő 2-es számrendszerbeli szám 3-al való osztási maradékát fogja kiírni (minden szó beolvasása után).

Működési elve - Észrevehetjük, hogy minden 2-számrendszerbeli szám i_0 vagy i_1 alakú és hogy $i_0 = 2i$, $i_1 = 2i + 1$ értékű. Vegyük észre azt is, hogy $[i]_2$ osztási maradékának ismeretében $[i0]_2$ és $[i1]_2$ is kiszámítható a 3-al való osztási maradéka, ahogy azt az alábbi táblázatok mutatják (Itt $[i]_2$ 3-al való osztási maradékát p-vel $[i0]_2$ és $[i1]_2$ osztási maradékát pedig m-el jelöljük.)

$$\begin{array}{l|ll} [i0]_2 = 2[i] & p = 1 & m = 2 \\ & p = 2 & m = 1 \\ & p = 0 & m = 0 \end{array} \quad \begin{array}{l|ll} [i1]_2 = 2[i] + 1 & p = 1 & m = 0 \\ & p = 2 & m = 2 \\ & p = 0 & m = 1 \end{array}$$

Feleljenek most meg a q_0, q_1, q_2 állapotok a 0,1 és 2 maradékoknak. Tegyük fel, hogy a ω szó hatására q_0 -ból a q_i állapotba jutottunk, (azaz a ω -nek megfelelő $[\omega]_2$ bináris szám 3-al való osztási maradéka i), most $q_i \xrightarrow{a} q_j$ akkor és csak akkor teljesül, ha az $a \in \{0,1\}$ jel beolvasása után a $[\omega a]_2$ szám 3-al való osztási maradéka j lesz. Az alábbi tranzíciós diagrammot kapjuk:



Könnyen ellenőrizhető, hogy M most valóban egy 2 számrendszerbeli szám 3-al való osztási maradékát írja ki minden egyes lépés után. Pl. $\omega = 1011$
 $11:3=3$

$[\omega]_2 = 2^3 + 2 + 1 = 11 \xrightarrow{\text{2}} 9$ Vegyük észre, hogy ω hatására valóban q_0 -ból

q_2 -be jutunk.

Mealy automata -jegyzetelni Házi feladat

2. Reguláris kifejezések

A Σ ábécé feletti reguláris kifejezéseket a következőképpen értelmezhetjük:

1) $\emptyset$ reguláris kifejezésnek a $\emptyset$ nyelv felel meg neki, ε reguláris kifejezésnek a $\{\varepsilon\}$ nyelv felel meg.

2) minden $a \in \Sigma$ -ra a egy reguláris kifejezésnek, aminek $\{a\}$ nyelv felel meg.

3) Ha r és s az R illetve S nyelveket jelölő reguláris kifejezések, akkor

$r+s$ is reguláris kifejezés - az $R \cup S$ nyelvet jelöli

$r \cdot s$ is reguláris kifejezés - az $R \cdot S$ nyelvet jelöli

és r^* is reguláris kifejezés - az R^* nyelvet jelöli.

Pl: 00 reguláris kifejezés - a $\{00\}$ nyelvet jelöli

01^* reg. kifejezés - az összes olyan bináris szót jelöli amiben az első karakter 0, utána pedig csak 1-es jöhet

$(0+1)^*$ - reguláris kifejezés az összes bináris szót jelöli

$(0+1)^*00(0+1)^*$ - reguláris kifejezés jelöli az összes olyan bináris szót, ami legalább két egymás melletti 0-t tartalmaz.

3. Reguláris nyelvek

Definíció 1.: Reguláris nyelvek azok az $L \subseteq \Sigma^*$ nyelvek amelyek megadhatók úgy, mint egy reguláris kifejezéshez tartozó szavak összessége.

Pl $L = \{00w \mid w \in (0+1)^*\}$ egy reguláris nyelv mert megadható a $00(0+1)^*$ reguláris kifejezéssel.

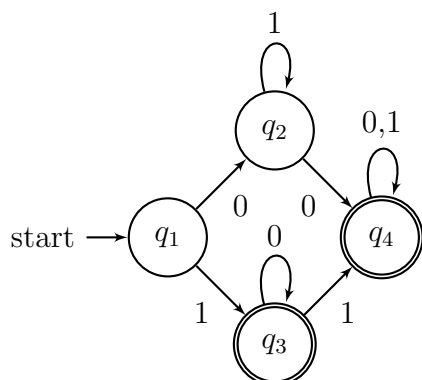
Ekvivalens rekurzív definíció

- 1) Elemi reguláris nyelvek, $\emptyset$, $\{\varepsilon\}$ és minden $a \in \Sigma$ esetén $\{a\}$.
 - 2) Azok és csakis azok a reguláris nyelvek, amelyek az elemi reguláris nyelvekből az $\cup$, $\cdot$ és $*$ művelet véges számú alkalmazásával megkaphatók
- Pl. $L = \{00, 101\}$ nyilván reguláris nyelv hiszen: $L = \{0\} \cdot \{0\} \cup \{1\} \cdot \{0\} \cdot \{1\}$

4. Reguláris nyelvek és véges determinisztikus automaták kapcsolata

Legyen $A = (\{q_1, q_2, \dots, q_n\}, \Sigma, q_1, \delta, F)$ egy véges determinisztikus automata. Minden $i, j, k \in \{1, \dots, n\}$ indexre értelmezzük a következő halmazt:

$R_{i,j}^k = \{w \in \Sigma^* \mid \bar{\delta}(q_i, w) = q_j, \text{ a } w\text{-t előállító tranzíciós útvonal nem halad át } k\text{-nál magasabb indexű állapoton}\}$



$$\begin{aligned} R_{1,4}^2 &= 01^*0 \\ R_{1,4}^3 &= 01^*0 + 10^*1 \\ R_{1,2}^0 &= \{0\}, R_{1,1}^0 = \{\varepsilon\}, R_{1,4}^0 = \emptyset \end{aligned}$$

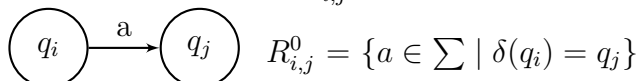
Tétel 3: Egy A véges determinisztikus automata által elfogadott $\mathcal{L}(A)$ nyelv mindig reguláris.

Bizonyítás:

Ha A végállapotai $q_{j_1}, q_{j_2}, \dots, q_{j_k}$ akkor $\mathcal{L}(A) = R_{1,j_1}^n \cup R_{1,j_2}^n \cup \dots \cup R_{1,j_k}^n$ (*)

Be fogjuk bizonyítani, hogy minden $R_{i,j}^n$ ($i, j \in \{1, \dots, n\}, n \in \mathbb{N}$) nyelv reguláris, mégpedig indukcióval bizonyítjuk.

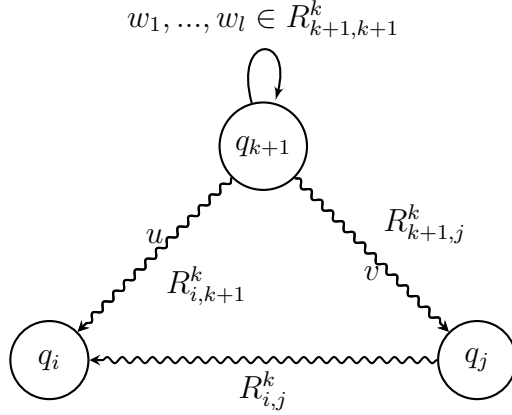
Ha $m = 0$ akkor az $R_{i,j}^0$ nyelv vagy üres vagy egyetlen karaktert tartalmaz:



$i=j$ esetén a lehet ε is. Tehát $R_{i,j}^0$ vagy $\emptyset$, vagy $\{a\}$ vagy $\{\varepsilon\}$ - ezek mind reguláris nyelvek. Tegyük fel most hogy minden $i, j \in \{1, \dots, n\}$ -re

$R_{i,j}^k$ reguláris nyelv és igazoljuk, hogy akkor $R_{i,j}^{k+1}$ is az:

Vegyük észre, hogy q_i -ből q_j -be lényegében egy $u \in R_{i,j}^{k+1}$ szóval kétféleképpen juthatunk el: - vagy hogy érintjük a q_{k+1} állapotot és ez esetleg többször is előfordul, vagy pedig olyan útvonalmentén ami nem érint k -nál magasabb indexű állapotot. A második esetben a szavunk egy $w \in R_{i,j}^k$ szó



Tekintsük most az első lehetőséget, a q_{k+1} állapoton átmenő utat. Ezt így szemléltethetjük:

$$q_i, \dots, q_{k+1}, \dots, q_{k+1}, \dots, q_{k+1}, \dots, q_j,$$

mert q_{k+1} többször is előfordulhat. Az első előfordulásáig a $u \in \Sigma^*$ szóval jutunk el. az egyes előfordulások közötti utaknak vagyis a "hurokban való forgásnak" a $w_1 \cdot w_2 \cdot \dots \cdot w_s$ szavak felelnek meg. Ezekhez tehát olyan utak tartoznak amelyek q_{k+1} -ből q_{k+1} -be mennek, úgy hogy közben nem haladnak át k -nál magasabb indexű állapoton. q_{k+1} utolsó előfordulásából q_j -be a v szóval jutunk el. a szavaink tehát a következő halmazokhoz tartoznak.

$$u \in R_{i,k+1}^k, v \in R_{k+1,j}^k, \quad w_1 \cdot \dots \cdot w_s \in R_{k+1,k+1}^k, \text{ így} \\ w_1 \cdot \dots \cdot w_s \in (R_{k+1,k+1}^k)^*$$

$$\text{Tehát } R_{i,j}^{k+1} = R_{i,j}^k \cup R_{i,k+1}^k \cdot (R_{k+1,k+1}^k)^* \cdot R_{k+1,j}^k \quad (**)$$

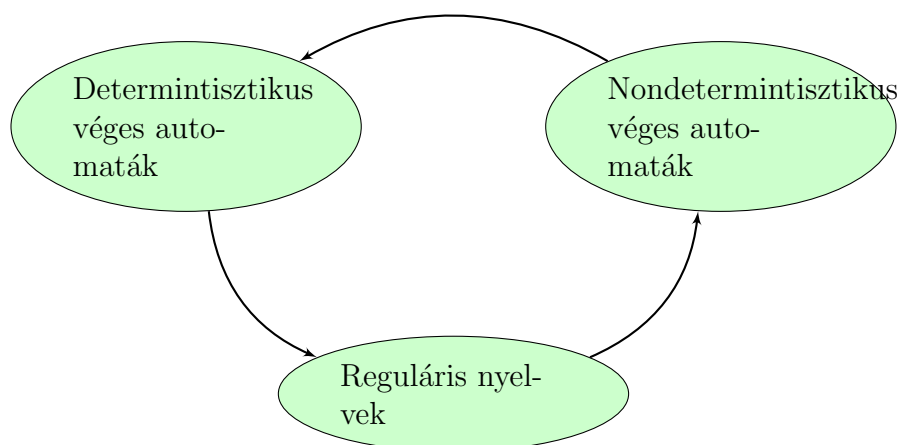
Mivel a szavakban és az $\cup$ -ban szereplő minden szóhalmaz (nyelv) reguláris, ezért $R_{i,j}^{k+1}$ is reguláris $\implies$ az indukció elve alapján következik, hogy minden $R_{i,j}^m$ nyelv reguláris! - De akkor a (*) összefüggésben szereplő k darab nyelv is reguláris, ezért $\mathcal{L}(A)$ is reguláris.

Azonnal látni fogjuk, hogy a Tétel 1 fordítottja is igaz. Ugyanis igazolható az alábbi tétel:

Tétel 2: Minden $L \subseteq \Sigma^*$ reguláris nyelvhez megadható egy olyan N nondeterminisztikus véges automata amire $\mathcal{L}(N) = L$

5. Ekvivalencia tétel

Egy $L \subseteq \Sigma^*$ nyelv akkor és csak akkor reguláris ha előállítható egy véges determinisztikus automatával.



Bizonyításvázlat: Igazoltuk hogy a véges determinisztikus automata nyelvei reguláris nyelvek. Fordítva, legyen most L egy reguláris nyelv. Akkor a Tétel 2 szerint előállítható egy véges nondeterminisztikus N automatával. Előző órán megmutattuk hogy minden véges nondeterminisztikus N automatahoz konstruálható egy olyan A véges determinisztikus automata ami ugyanazt a nyelvet állítja elő mint N , vagyis az L nyelvet. Tehát minden reguláris nyelv előállítható egy véges determinisztikus automatával. Ezeket az átalakításokat a fenti ábra szemlélteti.

Következmény1: Minden $L \subseteq \Sigma^*$ nyelvhez szerkeszthető egy A véges determinisztikus automata ami L -t elfogadja.

- Két automatát A_1 és A_2 -t ekvivalensnek nevezzük ha ugyanazt a nyelvet fogadják el, vagyis $\mathcal{L}(A_1) = \mathcal{L}(A_2)$.

- Két automata típust ekvivalensnek nevezünk ha ugyanazt a nyelvosztályt állítják elő.

Következmény2: A DFA-k, az NFA-k és a Moore automata ekvivalens automata típusok.

6. Reguláris nyelvek zártsági tulajdonságai

Tekintsük a Σ (rögzített) ábécé feletti reguláris nyelveket. Tudjuk hogy ha $L_1, L_2 \subseteq \Sigma^*$ reguláris nyelvek, akkor

- $L_1 \cup L_2$ is reguláris nyelv
- $L_1 \cdot L_2$ is reguláris nyelv

- L_1^*, L_2^* is reguláris nyelvek.

Az ilyen tulajdonságokat zártági tulajdonságoknak nevezzük.

Azt is mondhatjuk, hogy a Σ feletti reguláris nyelvek családja, amit $\text{Reg}(\Sigma)$ -val fogunk jelölni zárt az $\cup, \cdot$ és $*$ műveletekre nézve.

Magyarázat: ezek a műveletek nem vezetnek ki a $\text{Reg}(\Sigma)$ nyelvcsaládból! Egy $L \subseteq \Sigma^*$ nyelv $\bar{L}$ komplementere alatt a $\Sigma^* \setminus L$ különbségalmazat értjük. Nyilván $\bar{\bar{L}} = \{w \in \Sigma^* \mid w \notin L\}$

Tétel 3: Ha $L \subseteq \Sigma^*$ egy reguláris nyelv, akkor $\bar{L}$ is reguláris nyelv.

Bizonyítás: Ha L egy reguláris nyelv, akkor az Ekvivalencia tétel szerint létezik egy olyan $A = (Q, \Sigma, \delta, q_0, F)$ véges determinisztikus automata amire $\mathcal{L}(A) = L$. Nyilván itt $F \neq \emptyset$ és $F \subseteq Q$. Tekintsük most az $A' = (Q, \Sigma, \delta, q_0, Q \setminus F)$ automatát. Mivel δ nem változott meg nyilván ez is egy véges determinisztikus automata. Vegyük észre, hogy tetszőleges $w \in \Sigma^*$ szó esetén

$$w \in \mathcal{L}(A') \iff \bar{\delta}(q_0, w) \in Q \setminus F \iff \bar{\delta}(q_0, w) \notin F \iff w \notin \mathcal{L}(A) \iff w \notin L \iff w \in \bar{L}.$$

Ez azt jelenti, hogy $\mathcal{L}(A') = \bar{L}$. Tehát $\bar{L}$ is egy reguláris nyelv.

Következmény1: Ha $L_1, L_2 \in \text{Reg}(\Sigma)$ akkor

- (i) $L_1 \cap L_2 \in \text{Reg}(\Sigma)$
- (ii) $L_1 \setminus L_2 \in \text{Reg}(\Sigma)$

Bizonyítás:

$$(i) L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}. \text{ Ha } L_1, L_2 \in \text{Reg}(\Sigma) \xrightarrow{\text{Tétel 3}} \overline{L_1}, \overline{L_2} \in \text{Reg}(\Sigma) \\ \implies \overline{L_1} \cup \overline{L_2} \in \text{Reg}(\Sigma) \implies \overline{\overline{L_1} \cup \overline{L_2}} \in \text{Reg}(\Sigma).$$

Tehát $L_1 \cap L_2 \in \text{Reg}(\Sigma)$

(ii) $L_1 \setminus L_2 = L_1 \cap \bar{L_2}$. Ha $L_1, L_2 \in \text{Reg}(\Sigma)$ akkor $\bar{L_2} \in \text{Reg}(\Sigma)$ és (i) alapján $L_1 \cap \bar{L_2} \in \text{Reg}(\Sigma)$. Így $L_1 \setminus L_2 \in \text{Reg}(\Sigma)$.

Következmény2: $\text{Reg}(\Sigma)$ zárt a $\cup, \cdot, \cap, \setminus$ halmazműveletekre.

Következmény3: $\text{Reg}(\Sigma)$ egy ún. Boole algebra.