

Harmadik típusú nyelvek és véges automaták

Formális nyelvek, 10. gyakorlat

Célja: Az automata-nyelvekre vonatkozó tételek elmélyítése, gyakorlati alkalmazása

Fogalmak: reguláris kifejezés, Kleene-tétel, általánosított reguláris kifejezés, direkt szorzat automata, maradéknnyelvek és tulajdonságai, MYHILL-NERODE tétel, kis Bar-Hillel lemma

Feladatok jellege: Néhány reguláris és általánosított reguláris kifejezés felírása. Egy egyszerű automata Kleene-nyelveinek elkészítése. Automatakészítés a szimmetrikus differenciához valamely konkrét, nagyon egyszerű VDA-k esetén. Konkrét nyelvek maradéknnyelvei halmazának felírása, ezekből MYHILL-NERODE alapján következtetések levonása. Konkrét nyelvről kis Bar-Hillel lemmával kimutatni, hogy nem reguláris.

2008/09 I. félév

Házi feladatok megoldása

1. feladat

Melyik nyelvet fogadja el a következő automata?

	a	b	c
→ q ₀	q ₁	q ₂	q ₃
q ₁	q ₄	q ₂	q ₄
q ₂	q ₀	q ₄	q ₃
← q ₃	q ₄	q ₃	q ₄
q ₄	q ₄	q ₄	q ₄

Megoldás:

$$(aba \cup ba)^*(abc \cup bc \cup c)b^*$$

Házi feladatok megoldása

2. feladat

Készítsünk KMP automatát a következő mintához!
babbcabc ($T = \{a, b, c\}$)

Megoldás:

	a	b	c
→ q _ε	q _ε	q _b	q _c
q _b	q _{ba}	q _b	q _c
q _{ba}	q _ε	q _{bab}	q _c
q _{bab}	q _{ba}	q _{babb}	q _c
q _{babb}	q _{ba}	q _b	q _{babbc}
q _{babbc}	q _{babbca}	q _b	q _c
q _{babbca}	q _ε	q _{babbcab}	q _c
q _{babbcab}	q _{ba}	q _b	q _F
← q _F	q _F	q _F	q _F

VDA-hoz 3NF nyelvtan készítése

1. Feladat: Adjunk 3NF nyelvtant, mely ugyanezt a nyelvet generálja, amit az automata elfogad!

	a	b
→ q ₀	q ₀	q ₂
q ₁	q ₁	q ₄
q ₂	q ₁	q ₃
← q ₃	q ₀	q ₄
← q ₄	q ₂	q ₃

Megoldás:

$$G = (\{a, b\}, \{q_0, q_1, q_2, q_3, q_4\}, \mathcal{P}, q_0)$$

$$q_0 \rightarrow aq_0 \mid bq_2$$

$$q_1 \rightarrow aq_1 \mid bq_4$$

$$q_2 \rightarrow aq_1 \mid bq_3$$

$$q_3 \rightarrow aq_0 \mid bq_4$$

$$q_4 \rightarrow aq_2 \mid bq_3$$

$$q_3 \rightarrow \varepsilon$$

$$q_4 \rightarrow \varepsilon$$

VDA konstruálása 3.típusú nyelvtanhoz

1. lépés: Normálformára hozzuk a nyelvtant.

2. lépés: Az automata elkészítése

Legyen $\mathcal{G} = \langle T, N, \mathcal{P}, S \rangle$ 3NF nyelvtan.

Készítünk egy \mathcal{A} VDA-t, melyre $L(\mathcal{A}) = L(\mathcal{G})$.

$\mathcal{A} = \langle 2^N, T, \delta, \{S\}, F \rangle$, ahol

$\delta(\{A_1, \dots, A_k\}, t) = \{B \in N \mid \exists 1 \leq i \leq k, A_i \rightarrow tB \in \mathcal{P}\}$,

$F = \{X \in 2^N \mid X \cap \{A \in N \mid A \rightarrow \varepsilon \in \mathcal{P}\} \neq \emptyset\}$.

A gyakorlatban csak azon állapotokra határozzuk meg az átmeneteket, melyek a kezdőállapotból elérhető.

VDA konstruálása 3.típusú nyelvtanhoz

Példa

2. Feladat: Készítsünk VDA-t, $S \rightarrow abA \mid bS$
mely ugyanazt a nyelvet fogadja el, $A \rightarrow bbA \mid B \mid a \mid b$
melyet a nyelvtan generál! $B \rightarrow S \mid abB$

Megoldás:

1.a.) lépés: Láncmentesítés

$S \rightarrow abA \mid bS$
 $A \rightarrow bbA \mid abB \mid abA \mid bS \mid a \mid b$
 $B \rightarrow abA \mid abB \mid bS$

1.b.) lépés: Hosszredukció

$S \rightarrow aK_1 \mid bS$
 $A \rightarrow bK_1 \mid aK_2 \mid aK_1 \mid bS \mid a \mid b$
 $B \rightarrow aK_1 \mid aK_2 \mid bS$
 $K_1 \rightarrow bA$ $K_2 \rightarrow bB$

1.c.) lépés: "A $\rightarrow a$ " alakú szabályok eliminálása

$S \rightarrow aK_1 \mid bS$
 $A \rightarrow bK_1 \mid aK_2 \mid aK_1 \mid bS \mid aF \mid bF$
 $B \rightarrow aK_1 \mid aK_2 \mid bS$
 $K_1 \rightarrow bA$
 $K_2 \rightarrow bB$
 $F \rightarrow \varepsilon$

VDA konstruálása 3.típusú nyelvtanhoz

Példa

$S \rightarrow aK_1 \mid bS$
 $A \rightarrow bK_1 \mid aK_2 \mid aK_1 \mid bS \mid aF \mid bF$
 $B \rightarrow aK_1 \mid aK_2 \mid bS$
 $K_1 \rightarrow bA$
 $K_2 \rightarrow bB$
 $F \rightarrow \varepsilon$

	a	b
$\rightarrow \{S\}$	$\{K_1\}$	$\{S\}$
$\{K_1\}$	$\{\}$	$\{A\}$
$\{\}$	$\{\}$	$\{\}$
$\{A\}$	$\{K_1, K_2, F\}$	$\{K_1, S, F\}$
$\leftarrow \{K_1, K_2, F\}$	$\{\}$	$\{A, B\}$
$\leftarrow \{K_1, S, F\}$	$\{K_1\}$	$\{A, S\}$
$\{A, B\}$	$\{K_1, K_2, F\}$	$\{K_1, S, F\}$
$\{A, S\}$	$\{K_1, K_2, F\}$	$\{K_1, S, F\}$

A VDA-k éppen az \mathcal{L}_3 -beli nyelveket ismerik fel.

Minimális automata készítése

Cél: Adott automatához minél kevesebb állapotú, az eredeti automata által felismert nyelvet elfogadó automatát megadni.

Minimális automata: minimális állapotszámú ilyen automata.

1) Összefüggővé alakítás:

$H_0 = \{q_0\}$.

$H_{i+1} = H_i \cup \{q \in Q \mid \exists q' \in H_i, t \in T, q = \delta(q', t)\}$.

Tartalmazásra nézve monoton növekvő, felülről korlátos halmabsorozat, így stabilizálódik, azaz $\exists i$, hogy $H_i = H_{i+1} =: H$.

Hagyjuk el a $Q \setminus H$ -beli állapotokat és ezekre vonatkozó állapotátmeneteket.

2) Redukció:

Ha a q és q' állapotokból ugyanazon szavak hatására kerül az automata végállapotba, akkor az elfogadás szempontjából mindegy, hogy a működés során a q vagy a q' állapotba kerül az automata.

Célunk az ilyen állapotok összevonása.

Redukció

Ekvivalens állapotok

Az $q \in Q$ állapotra vonatkozó maradék nyelv:

$$L(\mathcal{A}, q) := \{v \mid \delta(q, v) \in F\}.$$

Legyenek $q, q' \in Q$ állapotok. q és q' ekvivalensek, ha

$$L(\mathcal{A}, q) = L(\mathcal{A}, q'). \text{ Jelölése: } q \sim q'.$$

\sim ekvivalencia- és jobbkongruenciareláció és analóg módon kiterjeszhető két automata állapotai közötti relációvá.

$\mathcal{A}, \mathcal{A}'$ automaták (q_0 és q'_0 kezdőállapotokkal) ekvivalensek, (jelölésben $\mathcal{A} \sim \mathcal{A}'$) ha $q_0 \sim q'_0$.

Az \mathcal{A} automata \mathcal{A}/\sim faktorautomatájája ekvivalens az eredetivel, redukált (nincsenek különböző ekvivalens állapotai), továbbá izomorfia erejéig az egyetlen összefüggő, redukált \mathcal{A} -val ekvivalens automata.

Állapotok ekvivalenciájának algoritmikus eldöntése

$$q \sim q' \Leftrightarrow L(\mathcal{A}, q) = L(\mathcal{A}, q') \Leftrightarrow \forall u \in T^* : (\delta(q, u) \in F \Leftrightarrow \delta(q', u) \in F).$$

Azt mondjuk, hogy $q \stackrel{i}{\sim} q'$ (q i -ekvivalens q' -vel), ha minden $u \in T^{\leq i}$ esetén $(\delta(q, u) \in F \Leftrightarrow \delta(q', u) \in F)$ ($i \geq 0$).

- $\stackrel{i}{\sim}$ ekvivalenciareláció,
- $q \stackrel{0}{\sim} q'$, ha $(q \in F \Leftrightarrow q' \in F)$,
- minden $q, q' \in Q$ -ra $q \stackrel{i+1}{\sim} q' \Leftrightarrow q \stackrel{i}{\sim} q' \wedge (\forall t \in T : \delta(q, t) \stackrel{i}{\sim} \delta(q', t))$,
- $\stackrel{0}{\sim} \prec \stackrel{1}{\sim} \prec \stackrel{2}{\sim} \prec \dots \prec \sim$, továbbá $q \sim q' \Leftrightarrow \forall i \geq 0 : q \stackrel{i}{\sim} q'$.
($\varrho_1 \prec \varrho_2$, ha minden $q, q' \in Q$ esetén $q\varrho_2q' \Rightarrow q\varrho_1q'$.)

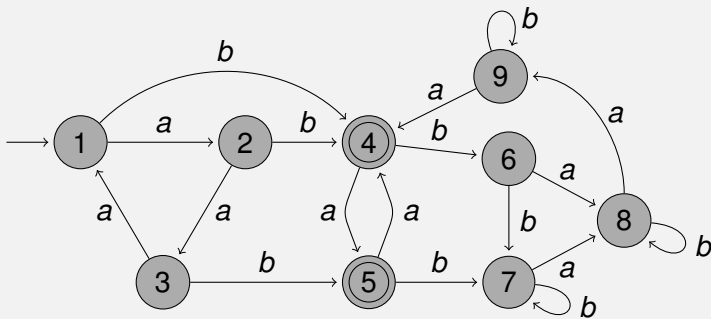
$$i_0 := \min\{j \mid \stackrel{j}{\sim} = \sim\}.$$

Ekkor $i_0 \leq |Q| - 1$, és így $\sim = |Q|^{-1}$.

Redukálás

5. Feladat:

Redukáljuk a következő automatát!



Redukálás

Megoldás:

$$\stackrel{0}{\sim} : \{1, 2, 3, 6, 7, 8, 9\}, \{4, 5\}$$

	a	b
1	{1,2,3,6,7,8,9}	{4,5}
2	{1,2,3,6,7,8,9}	{4,5}
3	{1,2,3,6,7,8,9}	{4,5}
6	{1,2,3,6,7,8,9}	{1,2,3,6,7,8,9}
7	{1,2,3,6,7,8,9}	{1,2,3,6,7,8,9}
8	{1,2,3,6,7,8,9}	{1,2,3,6,7,8,9}
9	{4,5}	{1,2,3,6,7,8,9}

$$\stackrel{1}{\sim} : \{1, 2, 3\}, \{4, 5\}, \{6, 7, 8\}, \{9\}$$

	a	b		a	b
4	{4,5}	{6,7,8}	1	{1,2,3}	{4,5}
5	{4,5}	{6,7,8}	2	{1,2,3}	{4,5}
			3	{1,2,3}	{4,5}
			6	{6,7,8}	{6,7,8}
			7	{6,7,8}	{6,7,8}
			8	{6,7,8}	{9}

$$\stackrel{2}{\sim} : \{1, 2, 3\}, \{4, 5\}, \{6, 7\}, \{8\}, \{9\}$$

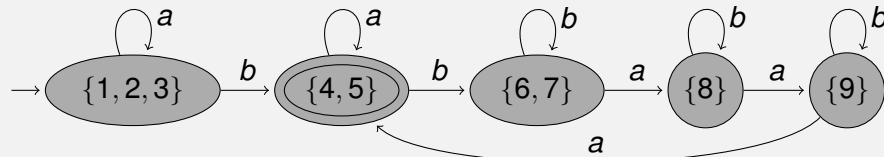
Redukálás

$\tilde{\sim}$: {1,2,3}, {4,5}, {6,7}, {8}, {9}

	a	b	1	a	b	6	a	b
4	{4,5}	{6,7}	2	{1,2,3}	{4,5}	7	{8}	{6,7}
5	{4,5}	{6,7}	3	{1,2,3}	{4,5}	7	{8}	{6,7}

$\tilde{\sim}^3 = \tilde{\sim} = \sim$

A redukált automata:



Az átmenetek és elfogadó állapotok meghatározásához tetszőleges reprezentánst tekinthetünk. Az eredeti kezdőállapotot tartalmazó ekvivalenciaosztály lesz az új kezdőállapot.

\mathcal{L}_3 -beli-e egy nyelv?

Egy L nyelv $p \in T^*$ -ra vonatkozó maradéknyelve $L_p := \{v \mid pv \in L\}$.

Myhill-Nerode tétel

$L \in \mathcal{L}_3$ akkor és csak akkor, ha $|\{L_p\}_{p \in T^*}| < \infty$, ahol $T = T(L)$ az L nyelv ábécéje.

Kis Bar-Hillel lemma

Minden $L \in \mathcal{L}_3$ nyelvhez van olyan $n = n(L) \in \mathbb{N}$ nyelvfüggő konstans, hogy minden $u \in L$, $\ell(u) \geq n$ szó esetén van u -nak olyan $u = xyz$ felbontása ($x, y, z \in T(L)^*$), melyre

- $\ell(xy) \leq n$,
- $\ell(y) > 0$,
- minden $i \geq 0$ esetén $xy^i z \in L$.

Myhill-Nerode tétel

Maradéknyelvek

Határozzuk meg a maradéknyelveit az alábbi nyelveknek!

1. $L = \{a, abb, bb, b\}$
2. $L = \{0, 1\}^* 00 \cup \{0\}$
3. HE (helyes zárójelezések nyelve)

Megoldások:

1.feladat

$L_\varepsilon = L$, $L_a = \{\varepsilon, bb\}$, $L_b = \{\varepsilon, b\}$, $L_{ab} = \{b\}$, $L_{bb} = \{\varepsilon\}$, $L_{abb} = \{\varepsilon\}$.
 $L_u = \emptyset \quad \forall u \notin \text{Pre}(L)$.

A maradéknyelvek halmaza tehát:

$\{\emptyset, \{\varepsilon\}, \{b\}, \{\varepsilon, b\}, \{\varepsilon, bb\}, \{a, abb, bb, b\}\}$

Myhill-Nerode tétel

Maradéknyelvek

2.feladat

$$L_u = \begin{cases} L \setminus \{0\} & 1 \in \text{Suf}(u) \\ L & u = \varepsilon \vee 0 \in \text{Suf}(u) \wedge 00 \notin \text{Suf}(u) \wedge u \neq 0 \\ L \cup \{\varepsilon\} & u = 0 \vee 00 \in \text{Suf}(u) \end{cases}$$

3.feladat Legyen $k \in \mathbb{N}$ -re:

$$\text{HE}_k^P := \{u \in \{(,)\}^* \mid \ell_l(u) - \ell_r(u) = k \wedge \ell_l(v) \geq \ell_r(v), \forall v \in \text{Pre}(u)\}$$

$$\text{HE}_k^S := \{u \in \{(,)\}^* \mid \ell_r(u) - \ell_l(u) = k \wedge \ell_r(v) \geq \ell_l(v), \forall v \in \text{Suf}(u)\}$$

$$\text{Ekkor } L_u = \begin{cases} \text{HE}_k^S & u \in \text{HE}_k^P \\ \emptyset & u \notin \bigcup_{k=0}^{\infty} \text{HE}_k^P \end{cases} \quad (k \in \mathbb{N})$$

Myhill-Nerode tétele alapján mivel az első két nyelvnek véges sok (6 illetve 3) maradéknyelve van, míg a harmadiknak végtelen, ezért az első két nyelv \mathcal{L}_3 -beli, a helyes zárójelezések nyelve viszont nem.

Automata készítése maradéknnyelvek segítségével

A maradéknnyelvek segítségével készíthető 3. típusú nyelvekhez VDA:

$$\mathcal{A} = \langle \{L_p\}_{p \in T^*}, T, \delta, L_\varepsilon, \{L_p \mid \varepsilon \in L_p\} \rangle, \text{ ahol } \delta(L_p, t) = L_{pt}.$$

1.feladat

	<i>a</i>	<i>b</i>
$\rightarrow \{a, abb, bb, b\}$	$\{\varepsilon, bb\}$	$\{\varepsilon, b\}$
$\leftarrow \{\varepsilon, bb\}$	\emptyset	$\{b\}$
$\leftarrow \{\varepsilon, b\}$	\emptyset	$\{\varepsilon\}$
\emptyset	\emptyset	\emptyset
$\{b\}$	\emptyset	$\{\varepsilon\}$
$\leftarrow \{\varepsilon\}$	\emptyset	\emptyset

2.feladat

	0	1
$\rightarrow L$	$LU\{\varepsilon\}$	$L\{0\}$
$\leftarrow LU\{\varepsilon\}$	$LU\{\varepsilon\}$	$L\{0\}$
$L\{0\}$	L	$L\{0\}$

Megjegyzés: az így kapott VDA minimális automata is lesz.

Kis Bar-Hillel Lemma

Legyen L_{Kif} a csak az *a* változót tartalmazó helyes kifejezések nyelve. (Kifejezések: az 1. gyakorlaton tanultak szerint.)

Feladat: $L_{\text{Kif}} \stackrel{?}{\in} \mathcal{L}_3$

Megoldás:

A Kis Bar-Hillel lemma segítségével bebizonyítjuk, hogy $L_{\text{Kif}} \notin \mathcal{L}_3$.

Elegendő belátni, hogy L_{Kif} nem rendelkezik a Kis Bar-Hillel lemmában leírt, minden \mathcal{L}_3 -beli nyelvre igaz tulajdonsággal, azaz:

Minden $n \in \mathbb{N}$ -re van olyan $u_n \in L_{\text{Kif}}$, $\ell(u_n) \geq n$ szó, hogy minden $u_n = xyz$, $\ell(xy) \leq n$, $\ell(y) > 0$ felbontás esetén létezik olyan $i \in \mathbb{N}$, melyre $xy^i z \notin L_{\text{Kif}}$.

Kis Bar-Hillel Lemma

Tehát minden n -re válasszunk egy u_n kifejezést. Legyen mondjuk $u_n = ({}^n a)^n$. Ekkor nyilván $u_n \in L_{\text{Kif}}$.

Azt kell belátni, hogy u_n n hosszúságú prefixének semelyik nemüres részszava sem iterálható be. De egy ilyen részszó mindenképp $y = ({}^d$ alakú, ahol $d > 0$. Ha y beiterálható lenne akkor a Kis Bar-Hillel lemma szerint $({}^{n-d-k} \{d\}^i ({}^k a)^n = ({}^{n+(i-1)d} a)^n$ (a képletben " $\{$ " és " $\}$ " metazárójelek!) is L_{Kif} eleme lenne minden i természetes számra. De ez $i \neq 1$ esetén nem igaz, mert nem ugyanannyi (" $\{$ " és " $\}$ ") szerepel benne.

Házi feladat

- Készítsünk VDA-t a következő nyelvtanhoz! \rightarrow

	<i>a</i>	<i>b</i>
1	7	6
2	1	2
3	9	1
4	1	4
5	6	4
6	6	2
7	9	8
8	8	2
9	1	8

$S \rightarrow acA \mid bB \mid \varepsilon$
 $A \rightarrow B \mid b \mid C$
 $B \rightarrow S \mid abB \mid a$
 $C \rightarrow acbC \mid B$
- Készítsünk a következő automatával ekvivalens minimális állapotszámú véges determinisztikus automatát! \leftarrow
- Határozzuk meg a palindromák nyelvének $(L = \{u \in T^* \mid u = u^{-1}\})$ maradéknnyelveit! (T tetszőleges.)
- Bizonyítsuk be, hogy a palindromák nyelve nem \mathcal{L}_3 -beli a Myhill-Nerode tétel illetve a Kis Bar-Hillel lemma segítségével! ($|T| \geq 2$) \leftarrow