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, maradéknyelvek és tulajdonságaik, 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 maradéknyelvei halmazának felírása, ezekből MYHILL-NERODE alapján következtetések levonása. Konkrét nyelvről kis Bar-Hillel lemmával kimutatni, hogy nem reguláris.

2005/06 II. félév

Házi feladatok megoldása

1. feladat

Készítsünk VDA-t, mely épp az $L = a(b \cup c(ba)^*ca)^*b \cup b$ nyelv szavait fogadja el! $(T = \{a, b, c\})$

Megoldás:

	a	b	c
$\to q_0$	q_1	q_2	q_{zs}
q_1	q_{zs}	q_3	q_4
$\leftarrow q_2$	q_{zs}	q_{zs}	q_{zs}
$\leftarrow q_3$	q_{zs}	q_3	q_4
q_4	q_{zs}	q_5	q_6
q_5	q_4	q_{zs}	q_{zs}
q_6	q_1	q_{zs}	q_{zs}
q_{zs}	q_{zs}	q_{zs}	q_{zs}

Házi feladatok megoldása

2. feladat

Melyik nyelvet fogadja el a következő automata?

	a	b	c
$\to q_0$	q_1	q_2	q_3
q_1	q_4	q_2	q_4
q_2	q_0	q_4	q_3
$\leftarrow q_3$	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

3. feladat

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

Megoldás:

	a	b	c
$\to q_\varepsilon$	q_ε	q_b	q_ε
q_b	q_{ba}	q_b	q_ε
q_{ba}	q_ε	q_{bab}	q_ε
q_{bab}	q_{ba}	q_{babb}	q_ε
q_{babb}	q_{ba}	q_b	q_{babbc}
q_{babbc}	q_{babbca}	q_b	q_ε
q_{babbca}	q_ε	$q_{babbcab}$	q_ε
$q_{babbcab}$	q_{ba}	q_b	q_F
$\leftarrow q_F$	q_F	q_F	q_F

VDA-hoz 3NF

Adjunk 3NF nyelvtant, mely ugyanezt a nyelvet generálja, amit az automata elfogad!

	a	b
$\rightarrow q_0$	q_0	q_2
q_1	q_1	q_4
q_2	q_1	q_3
$\leftarrow q_3$	q_0	q_4
$\leftarrow q_4$	q_2	q_3

Megoldás:

$G = \langle \{a, b\}, \{q_0, q_1, q_2, q_3, q_4\}, \mathcal{P}, q_0 \rangle$
$q_0 \rightarrow aq_0 \mid bq_2$
$q_1 \rightarrow aq_1 \mid bq_4$
$q_2 \rightarrow aq_1 \mid bq_3$
$q_3 \rightarrow aq_0 \mid bq_4$
$q_4 \rightarrow aq_2 \mid bq_3$
$q_3 \rightarrow \varepsilon$
$q_4 \rightarrow \varepsilon$

VDA konstruálása 3.típusú nyelvtanhoz

3NF-re alakítás

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

Megoldás:

1. lépés: Láncmentesítés

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

VDA konstruálása 3.típusú nyelvtanhoz

3NF-re alakítás

2. lépés: Hosszredukció

$S \rightarrow aK_1 \mid aB \mid bS$
$A \rightarrow cK_2 \mid aK_3 \mid b \mid aK_1 \mid aB \mid bS \mid a$
$B \rightarrow aK_3 \mid aB \mid bS \mid aK_1 \mid b$

$K_1 \rightarrow bA$
$K_2 \rightarrow cA$
$K_3 \rightarrow bB$

3. lépés: “$A \rightarrow a$” alakú szabályok eliminálása

$S \rightarrow aK_1 \mid aB \mid bS$
$A \rightarrow cK_2 \mid aK_3 \mid bF \mid aK_1 \mid aB \mid bS \mid aF$
$B \rightarrow aK_3 \mid aB \mid bS \mid aK_1 \mid bF$

$K_1 \rightarrow bA$
$K_2 \rightarrow cA$
$K_3 \rightarrow bB$
$F \rightarrow \varepsilon$

VDA konstruálása 3.típusú nyelvtanhoz

1. lépés: NDA

$S \rightarrow aK_1 \mid aB \mid bS$
$A \rightarrow cK_2 \mid aK_3 \mid bF \mid aK_1 \mid aB \mid bS \mid aF$
$B \rightarrow aK_3 \mid aB \mid bS \mid aK_1 \mid bF$

$K_1 \rightarrow bA$
$K_2 \rightarrow cA$
$K_3 \rightarrow bB$
$F \rightarrow \varepsilon$

	a	b	c
$\rightarrow S$	K_1, B	S	
A	B, K_1, K_3, F	S, F	K_2
B	K_1, K_3, B	S, F	
K_1		A	
K_2			A
K_3		B	
$\leftarrow F$			

VDA konstruálása 3.típusú nyelvtanhoz
2. lépés: NDA-ból VDA

$\mathcal{A} = \langle A, T, \delta, a_0, F\rangle$ NDA-hoz $\mathcal{A}' = \langle A', T, \delta', a_0', F'\rangle$ VDA

Ötlet: $\mathcal{A}'$ kövesse végig $\mathcal{A}$ összes lehetséges működését!

A determinisztikussá tett automata állapotainak halmaza a nemdeterminisztikus automata hatványhalmaza, azaz $A' := 2^A$

A $\{b_1, \ldots b_s\}$ állapothoz és a t betûhöz a VDA állapotátmenet-függvénye a nemdeterminisztikus automata állapotátmenet-függvényének a b_i állapotokhoz és t betûhöz tartozó képeinek (azaz: állapotok halmazainak) unióját rendelje, azaz:

$$\delta'(\{b_1, \ldots b_s\}, t) := \bigcup_{i=1}^{s} \delta(b_i, t).$$

A kezdőállapot $a_0' := \{a_0\}$, az elfogadó állapotok F' halmaza, pedig azon állapotokból álljon, melyek tartalmaznak F-beli állapotot.

VDA konstruálása 3.típusú nyelvtanhoz
Nemdeterminisztikus automatából determinisztikus

		a	b	c
$\rightarrow$	$\{S\}$	$\{B, K_1\}$	$\{S\}$	$\{\}$
	$\{B, K_1\}$	$\{B, K_1, K_3\}$	$\{S, A, F\}$	$\{\}$
	$\{\}$	$\{\}$	$\{\}$	$\{\}$
	$\{B, K_1, K_3\}$	$\{B, K_1, K_3\}$	$\{S, A, B, F\}$	$\{\}$
$\leftarrow$	$\{S, A, F\}$	$\{B, K_1, K_3, F\}$	$\{S, F\}$	$\{K_2\}$
$\leftarrow$	$\{S, A, B, F\}$	$\{B, K_1, K_3, F\}$	$\{S, F\}$	$\{K_2\}$
$\leftarrow$	$\{B, K_1, K_3, F\}$	$\{B, K_1, K_3\}$	$\{S, A, B, F\}$	$\{\}$
$\leftarrow$	$\{S, F\}$	$\{B, K_1\}$	$\{S\}$	$\{\}$
	$\{K_2\}$	$\{\}$	$\{\}$	$\{A\}$
	$\{A\}$	$\{B, K_1, K_3, F\}$	$\{S, F\}$	$\{K_2\}$

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

Kis Bar-Hillel lemma

Minden $L \in \mathcal{L}_3$ nyelvhez van olyan $n = n(L) \in \mathbb{N}$ nyelvfüggő konstans, hogy minden $u \in L$ szó esetén ha tekintjünk egy tetszőleges $u = \alpha_1 u' \alpha_2$ olyan felbontását, ahol $l(u') \geq n$, akkor van u'-nek olyan v részszava ($u' = \beta_1 v \beta_2$), hogy $0 < l(v) \leq n$, és minden $i \geq 0$ esetén $\alpha_1 \beta_1 v^i \beta_2 \alpha_2 \in L$.

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 $\left|\{L_p\}_{p \in T^*}\right| < \infty$, ahol $T = T(L)$ az L nyelv ábécéje.

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

Megoldások:

1.feladat

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

A maradéknyelvek halmaza tehát:
$\{\emptyset, \{\varepsilon\}, \{b\}, \{\varepsilon, b\}, \{\varepsilon, bb\}, L\}$

Myhill-Nerode tétel

Maradéknyelvek

2.feladat

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

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

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

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

Ekkor $L_u = \begin{cases} \mathrm{HE}_k^S & u \in \mathrm{HE}_k^P \\ \emptyset & u \notin \mathrm{HE}_k^P \end{cases}$

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.

Kis Bar-Hillel Lemma

Feladat:

Kifejezések $\overset{?}{\in} \mathcal{L}_3$

Megoldás:

Nem. Indirekt módon. Tegyük fel, hogy Kifejezések $\in \mathcal{L}_3$. A Kis Bar-Hillel lemma alapján ekkor létezik $n = n(L)$. Vegyük a következő kifejezést: $u := (^n x)^n$, és legyen $u' = (^n$. Mivel $l(u') \geq n$, így alkalmazhatjuk a Kis Bar-Hillel lemmát.
Ezek szerint u'-ben létezik v nemüres, beiterálható részszó. Legyen $v := (^d$, ahol $d > 0$. A lemma szerint $(^{n-d-k}\{(^d\}^i(^k x)^n = (^{n+(i-1)d}x)^n$ eleme a Kifejezések nyelvének. (Az előző képletben "{" és "}" metazárójelek!) Ez viszont $i \neq 1$ esetén a nyelv definíciója miatt nem teljesül, tehát ellentmondásra jutottunk.

Házi feladat

1. Készítsünk VDA-t a következő nyelvtanhoz!
   $S \to acA \mid bB \mid \varepsilon$
   $A \to B \mid b \mid C$
   $B \to S \mid abB \mid a$
   $C \to acbC \mid B$
2. Határozzuk meg a palindromák nyelvének ($L = \{u \in T^* \mid u = u^{-1}\}$) maradéknyelveit! ($T$ tetszőleges.)
3. Bizonyítsuk be, hogy a palindromák nyelve nem $\mathcal{L}_3$-beli a Myhill-Nerode tétel illetve a Kis Bar-Hillel lemma segítségével! ($|T| \geq 2$)