

## 2.6 Еквивалентност между автоматни езици и езици, разпознавани от крайни автомати

Нашата най-близка цел е да докажем, че автоматните езици (или езиците от тип 3) са всъщност езиците, които могат да се разпознаят от крайни автомати. Следващите две теореми реализират тази цел.

**Теорема 2.6.1** *Ако  $A$  е детерминиран краен автомат, то  $L(A)$  е автоматен език.*

**Доказателство:** Нека  $A$  е автомат, за който  $A = \langle K, V, \delta, q_0, F \rangle$  и  $K \cap V = \emptyset$ . Да дефинираме сега граматиката  $\Gamma$ , за която  $L(\Gamma) = L(A) \setminus \{\varepsilon\}$ . Ако  $\varepsilon \in L(A)$ , то  $q_0 \in F$  и ако сме получили  $\Gamma$ , можем да определим  $\Gamma'$  така, че  $L(\Gamma') = L(\Gamma) \cup \{\varepsilon\}$ .

Да положим  $\Gamma = \langle V, W, S, P \rangle$ , където  $W = K$ ,  $S = q_0$  и  $P$  съдържа всички правила от вида  $q_i \rightarrow aq_j$  при  $\delta(q_i, a) = q_j$  и също всички правила от вида  $q_i \rightarrow a$  при  $\delta(q_i, a) \in F$ .

Нека  $\omega = a_1a_2\dots a_s$  е дума, която се разпознава от автомата  $A$ , т.е.  $\omega \in L(A)$ . Ще докажем, че  $\omega \in L(\Gamma)$ , ако  $\omega \neq \varepsilon$ , т.е.  $s \geq 1$ .

Тъй като  $\omega \in L(A)$  следва, че съществуват състояния  $q_{t_1}, q_{t_2}, \dots, q_{t_s} \in K$  така, че  $\delta(q_{t_i}, a_{i+1}) = q_{t_{i+1}}$  за  $i = 1, 2, \dots, s-1$  и  $\delta(q_{t_{s-1}}, a_s) = q_{t_s} \in F$ . Също така  $\delta(q_0, a_1) = q_{t_1}$ . Следователно в  $\Gamma$  има правила  $q_0 \rightarrow a_1q_{t_1}$ ,  $q_{t_i} \rightarrow a_{i+1}q_{t_{i+1}}$  за  $i = 1, 2, \dots, s-1$  и  $q_{t_{s-1}} \rightarrow a_s$ . Като вземем предвид, че  $S = q_0$  се вижда, че тези правила са достатъчни да се направи изводът  $q_0 \stackrel{\Gamma}{\vdash} \omega$ , т.е.  $\omega \in L(\Gamma)$ .

Предоставяме за самостоятелно упражнение да се докаже, че ако  $\omega \in L(\Gamma)$ , то  $\omega \in L(A)$ .  $\square$

**Теорема 2.6.2** *Ако  $L = L(\Gamma)$  е автоматен език, то съществува недетерминиран краен автомат  $A$ , който разпознава езика  $L$ , т.е.  $L = L(A)$ .*

**Доказателство:** Нека  $\Gamma = \langle V, W, S, P \rangle$  е автоматна граматика, която поражда езика  $L$ . Да означим с  $A = \langle K, V, \delta, q_0, F \rangle$  недетерминиран краен автомат, за който  $K = W \cup \{\#\}$ ,  $q_0 = S$ ,

$F = \{\#\}$ , където  $\#$  е символ не принадлежащ на  $W$ . Функцията на прехода  $\delta$  определяме по следния начин: ако правилото  $A \rightarrow a$  е в  $P$ , то  $\delta(A, a)$  съдържа състоянието  $\#$  и ако  $A \rightarrow aB$  е в  $P$ , то  $\delta(A, a)$  съдържа състоянието  $B$ .

Ясно е, че при тази дефиниция на  $\delta$ , на всяко правило от  $P$  съответствува някое подмножество на множеството  $K$ .

Нека  $\omega = a_1a_2\dots a_s$  е дума от езика  $L$ , т.е. съществува изводът

$$S \vdash a_1A_{t_1} \vdash a_1a_2A_{t_2} \vdash \dots \vdash a_1a_2\dots a_{s-1}A_{t_{s-1}} \vdash a_1a_2\dots a_s \quad \text{в } \Gamma.$$

Следователно в  $P$  има правила  $S \rightarrow a_1A_{t_1}$ ,  $A_{t_i} \rightarrow a_{t_{i+1}}A_{t_{i+1}}$ , за всяко  $i = 1, 2, \dots, s-2$  и  $A_{t_{s-1}} \rightarrow a_s$ . Следователно  $A_{t_1} \in \delta(S, a_1)$ ,  $A_{t_{i+1}} \in \delta(A_{t_i}, a_{i+1})$ , за всяко  $i = 1, 2, \dots, s-2$  и  $\# \in \delta(A_{t_{s-1}}, a_s)$ , което означава, че  $\omega$  се разпознава от автомата  $A$ , т.е.  $\omega \in L(A)$ .

За самостоятелно упражнение предоставяме да се докаже, че ако  $\omega \in L(A)$ , то  $\omega \in L(\Gamma)$ .  $\square$

Като се възползуваме от Теорема 2.6.2 и Теорема ?? следва, че езикът  $L$  е автоматен точно тогава, когато съществува детерминиран краен автомат  $A$ , за който  $L = L(A)$ .

Да отбележим, че по зададен автоматен език конструктивно (явно) може да се дефинира съответния му ДКА, който го разпознава. Този ДКА може да се използва ефективно за определяне дали дадена дума е от съответния език или не. За целта е достатъчно да се построи ориентираният граф, с който се представя ДКА и по този граф да се определи дали съществува краен път от началното състояние до някое крайно състояние, който съответствува на разглежданата дума. Понеже автоматът е детерминиран, то такъв път е най-много един, а ако този автомат се доопредели така, че за него  $\delta$  да е изображение (виж предходния параграф), то за всяка дума ще съответствува единствен на нея път и проверката се свежда единствено за това, дали този път завършва в крайно състояние на автомата.

Еквивалентността между автоматните езици и езиците, разпознавани от ДКА ни дава възможност да установим някои нови свойства на автоматните езици.

**Теорема 2.6.3** Ако  $L$  е автоматен език, то и езикът  $\overline{L} = V^* \setminus L$  е автоматен.

**Доказателство:** Нека  $\Gamma = \langle V, W, S, P \rangle$  е произволна автоматна граматика, която поражда езика  $L$ , т.е.  $L = L(\Gamma)$  и  $A = \langle K, V, \delta, q_0, F \rangle$  е ДКА, за който  $L = L(A)$ . Без ограничение на общността можем да смятаме, че  $A$  е напълно определен, т.е.  $\delta$  е изображение на  $K \times V$  в  $K$ , тъй като това може да се построи винаги с добавянето в  $K$  на едно допълнително фиктивно състояние (виж предходния параграф).

Очевидно  $\overline{L} = L(\overline{A})$ , където  $\overline{A}$  е ДКА, дефиниран по следния начин

$$\overline{A} = \langle K, V, \delta, q_0, K \setminus F \rangle. \quad \square$$

По-нататък ще покажем, че свойството автоматност на езиците се съхранява, не само при опрациите произведение, обединение и допълнение, а също така и при операцията сечение.

**Теорема 2.6.4** Ако  $L_1$  и  $L_2$  са автоматни езици, то и езикът  $L_1 \cap L_2$  е автоматен.

**Доказателство:** От Теорема 2.6.3 следва, че  $\overline{L_1}$  и  $\overline{L_2}$  са автоматни езици, а от Теорема 2.6.2 следва, че  $\overline{L_1} \cup \overline{L_2}$  е автоматен език. Но от закона на Де Морган имаме  $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$  и следователно  $L_1 \cap L_2$  е автоматен език.  $\square$

### З а д а ч и

1. Да се опише краен автомат, който приема всички думи над българската азбука, започващи с "под" и завършващи на "н".
2. Да се състави алгоритъм и програма на *Pascal*, чрез която се изброяват думите в даден текстов файл, които могат да бъдат разпознати от крайния автомат, определен в зад.1
3. Да се докаже, че ако  $L_1$  и  $L_2$  са автоматни езици, то и  $L_1 \setminus L_2$  е автоматен език.

**У п ъ т в а н е:** Използвайте равенството  $A \setminus B = A \cap \overline{B}$ , където  $A$  и  $B$  са множества.

4. Да се построи ДКА, разпознаващ езика, породен от граматиката:

а)  $\Gamma = \langle \{a, b\}, \{S, A, B, C\}, S, \{S \rightarrow aA|bB, A \rightarrow a|aC, B \rightarrow b|bC, C \rightarrow a|b\} \rangle$ ;

б)  $\Gamma = \langle \{0, 1\}, \{S, A, B\}, S, \{S \rightarrow 0A|0B|0, A \rightarrow 0|0A, B \rightarrow 1|1B\} \rangle$ .