

### 3.2 Недетерминирани магазинни автомати

Крайните автомати са удобен инструмент за описание на автоматните граматика. Естествено е да се търсят средства за автоматизирано описание и на безконтекстните граматика, както е естествено да предполагаме, че такива автомати имат по-сложна структура от крайните автомати.

Преди да въведем автомати, с които се описват безконтекстните граматика, да разгледаме една широко използвана структура за описание на данни, известна в програмирането като *стек*. Това е линейна структура с единствена точка на достъп до компонентите ѝ, наречена връх на стека. Основните процедури със стекове са добавяне и извличане на компонента от стека. Тези процедури могат да се осъществят само, чрез върха на стека. От казаното става ясно, че нова компонента може да се добави само във върха на стека и компонента, която подлежи на извличане е само компонентата от върха на стека. Често тази последователност на процедурите се нарича LIFO (last in first out) поради това, че последната компонента, която попада в стека, първа подлежи на извличане. Поради линейността на стека, следва че в него могат да се съхраняват и думи над дадена азбука. *Недетерминиран магазинен автомат (НМА)* може да се смята недетерминиран краен автомат, на който функцията на преход зависи и от съдържанието на някой стек, т.е. възможните състояния на този НКА се определят от текущото състояние, входната дума и от символа, който е във върха на стека.

**Определение 3.2.1 Недетерминиран магазинен автомат (НМА)** се нарича наредена седморка  $M = \langle K, V, W, \delta, q_0, z_0, F \rangle$ , в която:

$K$  е крайно множество от вътрешни състояния на автомата;

$V$  е входна азбука;

$W$  е крайно множество, наречено стекова (магазинна) азбука;

$\delta : K \times (V \cup \{\varepsilon\}) \times W \rightarrow P(K \times W^*)$  е функция на преходите;

$q_0 \in K$  е начално състояние;  
 $z_0 \in W$  е начален магазинен (стеков) символ;  
 $F \subset K$  е множество от крайни (заклучителни) състояния на автомата.

**Определение 3.2.2** Под конфигурация на НМА ще разбираме всяка наредена тройка  $\langle q, \alpha, \gamma \rangle \in K \times V^* \times W^*$ .

Нека  $\alpha = a_1 a_2 \dots a_s$ ,  $\gamma = x_1 x_2 \dots x_m$ . Функцията  $\delta$  определя преход от една конфигурация към следващите конфигурации по следния начин:

- (i) за всяка двойка  $\langle p, \gamma' \rangle \in \delta(q, a_1, x_1)$  конфигурацията  $\langle q, \alpha, \gamma \rangle$  преминава в конфигурацията  $\langle p, \alpha_1, \gamma_1 \rangle$ , където  $\alpha_1 = a_2 a_3 \dots a_s$ ,  $\gamma_1 = \gamma' x_2 \dots x_m$ , което отбелязваме с  $\langle q, \alpha, \gamma \rangle \vdash \langle p, \alpha_1, \gamma_1 \rangle$ .
- (ii) за всяка двойка  $\langle p, \gamma' \rangle \in \delta(q, \varepsilon, x_1)$  конфигурацията  $\langle q, \alpha, \gamma \rangle$  преминава в конфигурацията  $\langle p, \alpha, \gamma' x_2 \dots x_m \rangle$  т.е.

$$\langle q, \alpha, \gamma \rangle \vdash \langle p, \alpha, \gamma' x_2 \dots x_m \rangle.$$

Работата на недетерминирания магазинен автомат върху думата  $\alpha = a_1 a_2 \dots a_s$  се състои в следното:

- от началната конфигурация  $\langle q_0, \alpha, z_0 \rangle$  се получават следващите конфигурации, чрез функцията на преходите  $\delta$ ;
- от новополучените конфигурации се получават всички следващи конфигурации и т.н.

Когато автоматът преминава от една конфигурация в друга, чрез правилото (ii) казваме, че той извършва  $\varepsilon$ -преместване.

**Определение 3.2.3** Казваме, че автоматът  $M$  разпознава думата  $\alpha$ , чрез крайно състояние, ако при работата си върху  $\alpha$  той достигне до конфигурация от вида  $\langle p, \varepsilon, \gamma \rangle$ ,  $p \in F$ .

Да означим с  $T(M)$  езика от думите, които  $M$  разпознава, чрез крайно състояние.

**Определение 3.2.4** Казваме, че думата  $\alpha$  се разпознава от автомата  $M$ , чрез празен магазин, ако след работата си върху  $\alpha$  автомата  $M$ , достигне до конфигурация от вида  $\langle p, \varepsilon, \varepsilon \rangle$ .

Да означим с  $N(M)$  езика, породен от  $M$  по този начин.

**Пример 3.2.1** Нека  $M = \langle \{q_0, q_1, q_2\}, \{a, b\}, \{z_0, z_1\}, \delta, q_0, z_0, \{q_0\} \rangle$  е НМА с функция на преходите

$$\delta(q_0, a, z_0) = \{\langle q_1, z_1 z_0 \rangle\}, \quad \delta(q_1, a, z_1) = \{\langle q_1, z_1 z_1 \rangle\},$$

$$\delta(q_1, b, z_1) = \{\langle q_2, \varepsilon \rangle\}, \quad \delta(q_2, b, z_1) = \{\langle q_2, \varepsilon \rangle\},$$

$$\delta(q_2, \varepsilon, z_0) = \{\langle q_0, \varepsilon \rangle\}.$$

Да разгледаме работата му върху думите  $aabb$  и  $abaab$ .

Последователно получаваме

$$\langle q_0, aabb, z_0 \rangle \vdash \langle q_1, abb, z_1 z_0 \rangle \vdash \langle q_1, bb, z_1 z_1 z_0 \rangle \vdash$$

$$\langle q_2, b, z_1 z_0 \rangle \vdash \langle q_2, \varepsilon, z_0 \rangle \vdash \langle q_0, \varepsilon, \varepsilon \rangle$$

и

$$\langle q_0, abaab, z_0 \rangle \vdash \langle q_1, baab, z_1 z_0 \rangle \vdash \langle q_2, aab, z_0 \rangle.$$

Очевидно първата дума е от  $T(M)$  и  $N(M)$ , докато втората дума не се разпознава от  $M$ .

Ясно е, че върху думи, започващи с  $b$  автоматът не е определен. Когато една дума започва с  $a$ , автоматът преминава в състояние  $q_1$  и за всяко  $a$  добавя в магазина по едно  $z_1$  и остава в състоянието  $q_1$ . След като се появи буквата  $b$  автоматът минава в състояние  $q_2$  и докато чете буквата  $b$ , остава в това състояние, като на всяка стъпка извлича от стека по едно  $z_1$ . Ако с изчерпването на буквите  $b$  се изчерпи и магазина, то автоматът преминава в заключително състояние. Ако обаче отново се появи  $a$ , или  $b$ -тата са повече или по-малко отколкото е буквата  $z_1$  в магазина, то автоматът е неопределен. Следователно

$$T(M) = N(M) = \{a^n b^n \mid n \geq 1\}.$$

Ще покажем, че безконтекстните езици са точно езиците, които могат да бъдат приети от недетерминирани магазинни автомати. Преди да установим този факт да разгледаме отново примера за аритметичните изрази.

Нека  $\Gamma = \langle V, W, I, P \rangle$  е безконтекстна граматика за която  $V = \{a, b, c, (, ), +, -, *, /\}$ ,  $W = \{I, O, M\}$  и правила

$$I \rightarrow O|I + O|I - O$$

$$O \rightarrow M|O * M|O/M$$

$$M \rightarrow a|b|c|(I).$$

Да построим НМА, който приема  $L(\Gamma)$ . Такъв автомат трябва да "копира" всяка лява схема на извод в  $\Gamma$ , чрез операции над неговия стек. Това може да стане например по следния начин: поставяме началното съдържание на стека да е  $Iz_0$ . На всяко правило  $I \rightarrow O$ ,  $I \rightarrow I + O$ ,  $I \rightarrow I - O$  да съпоставим следните думи над стековата азбука  $Oz_0$ ,  $I + Oz_0$  и  $I - Oz_0$ .

При всяко преминаване в ново състояние на автомата, се извлича по един символ. Ако този символ е нетерминален за  $\Gamma$ , то той се заменя в стека с дясната страна на правилата в  $\Gamma$ , в който символът е лява страна. Ако пък е терминален символ за  $\Gamma$ , той се използва за входен символ и определя следващото състояние на автомата. Да определим следния автомат

$$A = \langle K, V, W', \delta, q_0, z_0, \{q_2\} \rangle,$$

където

$$K = \{q_0, q_1, q_2\},$$

$$W' = V \cup W \cup \{z_0\}.$$

Функцията на преходите се дефинира по следния начин:

$$\delta(q_0, \varepsilon, z_0) = \{\langle q_1, Iz_0 \rangle\}$$

$$\delta(q_1, \varepsilon, I) = \{\langle q_1, O \rangle, \langle q_1, I + O \rangle, \langle q_1, I - O \rangle\}$$

$$\delta(q_1, \varepsilon, O) = \{\langle q_1, M \rangle, \langle q_1, O * M \rangle, \langle q_1, O/M \rangle\}$$

$$\delta(q_1, \varepsilon, M) = \{\langle q_1, a \rangle, \langle q_1, b \rangle, \langle q_1, c \rangle, \langle q_1, (I) \rangle\}$$

$$\delta(q_1, x, x) = \{\langle q_1, \varepsilon \rangle\} \text{ за всяко } x \in V$$

$$\delta(q_1, \varepsilon, z_0) = \{\langle q_2, \varepsilon \rangle\}.$$

Ясно е, че една входна дума е разпозната от  $A$  само, ако чрез нея стекът на  $A$  стане празен и тя привежда  $A$  в състояние  $q_2$ . Да разгледаме например думата  $\omega = a + b * c$ , която е от  $L(\Gamma)$  и да проследим работата на  $A$  върху нея.

Началната конфигурация на  $A$  е  $\langle q_0, \omega, z_0 \rangle$ . Последователно

от нея получаваме конфигурациите

$$\begin{aligned}
&\langle q_1, \omega, z_0 \rangle \vdash \langle q_1, \omega, \text{И}z_0 \rangle \vdash \langle q_1, \omega, \text{И} + \text{О}z_0 \rangle \vdash \langle q_1, \omega, \text{О} + \text{О}z_0 \rangle \vdash \\
&\vdash \langle q_1, \omega, \text{М} + \text{О}z_0 \rangle \vdash \langle q_1, a + b * c, a + \text{О}z_0 \rangle \vdash \langle q_1, +b * c, +\text{О}z_0 \rangle \vdash \\
&\quad \langle q_1, b * c, \text{О}z_0 \rangle \vdash \langle q_1, b * c, \text{О} * \text{М}z_0 \rangle \vdash \langle q_1, b * c, \text{М} * \text{М}z_0 \rangle \vdash \\
&\vdash \langle q_1, b * c, b * \text{М}z_0 \rangle \vdash \langle q_1, *c, * \text{М}z_0 \rangle \vdash \langle q_1, c, \text{М}z_0 \rangle \vdash \langle q_1, c, c z_0 \rangle \vdash \\
&\quad \vdash \langle q_1, \varepsilon, z_0 \rangle \vdash \langle q_2, \varepsilon, \varepsilon \rangle.
\end{aligned}$$

Този пример ни показва, как да построим НМА, които да разпознава даден безконтекстен език.

**Теорема 3.2.1** *За всеки безконтекстен език  $L$  съществува недетерминиран магазинен автомат  $M$  така, че  $L = L(M)$ .*

**Доказателство:** Нека  $L$  се поражда от безконтекстната граматика  $\Gamma = \langle V, W, S, P \rangle$ . Да дефинираме автомата  $M$  по следния начин:  $M = \langle K, V, W', \delta, q_0, z_0, \{q_2\} \rangle$ , където

$$\begin{aligned}
K &= \{q_0, q_1, q_2\}, \\
W' &= V \cup W \cup \{z_0\}, \quad z_0 \notin V \cup W
\end{aligned}$$

и функция на преходите  $\delta$  за която:

$$\begin{aligned}
\delta(q_0, \varepsilon, z_0) &= \{\langle q_1, S z_0 \rangle\} \\
\delta(q_1, \varepsilon, A) &= \{\langle q_1, \alpha \rangle \mid \text{съществува правило } A \rightarrow \alpha \text{ в } P\} \\
\delta(q_1, a, a) &= \{\langle q_1, \varepsilon \rangle\} \text{ за всяко } a \in V \\
\delta(q_1, \varepsilon, z_0) &= \{\langle q_2, \varepsilon \rangle\}.
\end{aligned}$$

Тази дефиниция гарантира, че ако  $\omega \in L$ , то съществува лява схема на извод  $S \models \omega$  в  $\Gamma$ , точно тогава когато от конфигурацията  $\langle q_0, \omega, z_0 \rangle$  може да се достигне до конфигурацията  $\langle q_2, \varepsilon, \varepsilon \rangle$ . Следователно

$$\omega \in L \iff S \models \omega \iff \langle q_2, \varepsilon \rangle \in \delta(q_0, \omega, z_0). \quad \square$$

В сила е и обратната теорема, но за нейното доказателство ще ни бъде необходима следната лема.

**Лема 3.2.1** Нека  $M = \langle K, V, W, \delta, q_0, z_0, F \rangle$  е произволен НМА. Съществува недетерминиран магазинен автомат

$$M' = \langle K', V, W', \delta', q'_0, z'_0, F' \rangle,$$

който е еквивалентен на  $M$ , има единствено крайно състояние и след достигането му, стекът на  $M'$  става празен.

**Доказателство:** Нека  $K' = K \cup \{q'_0, q'_1\}$ ,  $\{q'_0, q'_1\} \cap K = \emptyset$ ,  $W' = W \cup \{z'_0\}$ ,  $z'_0 \notin W$  и  $\delta'$  съвпада с  $\delta$  за всички състояния от  $K$  и

$$\begin{aligned} \delta'(q'_0, \varepsilon, z'_0) &= \{\langle q_0, z_0 z'_0 \rangle\} \\ \delta'(q, \varepsilon, A) &= \{\langle q'_1, \varepsilon \rangle\} \text{ за всяко } q \in F, A \in W' \\ \delta'(q'_1, \varepsilon, A) &= \{\langle q'_1, \varepsilon \rangle\} \text{ за всяко } A \in W' \\ F' &= \{q'_1\}. \end{aligned}$$

От първото равенство следва, че в стека на  $M'$  още преди автоматът да започне работата, с която и да е входна дума се поставя символът  $z'_0$ , което гарантира, че ако автоматът достигне до някое състояние от  $F$ , то в стека му ще остане поне един елемент

Второто равенство показва, че в  $q'_1$  автоматът може да достигне и чрез  $\varepsilon$ -преместване, ако преди това е бил в състояние от  $F$ . То ни гарантира, че  $M$  и  $M'$  са еквивалентни автомати т.е. разпознават един и същи език.  $\square$

**Теорема 3.2.2** Ако  $L$  е език, който се разпознава от недетерминиран магазинен автомат, то  $L$  е безконтекстен език.

**Доказателство:** Нека  $M = \langle K, V, W, \delta, q_0, z_0, \{q_n\} \rangle$  е автомат, за който  $L = T(M)$  и  $K = \{q_0, q_1, \dots, q_n\}$ . Да построим безконтекстна граматика  $\Gamma$ , за която  $L = L(\Gamma)$ . За всеки символ  $A \in W$  от магазинната азбука да въведем нетерминален символ  $A(i, j)$  на  $\Gamma$ , така че от него с правила да могат да се изведат всички думи  $\omega \in V^*$ , за които  $M$  се премества от състояние  $q_i$  до  $q_j$  и от стека на  $M$  се извлича символът  $A$ . Нека  $\Gamma = \langle V, W', S, P \rangle$ . Да положим

$$W' = \{A(i, j) \mid A \in W, 0 \leq i, j \leq n\}, S = z_0(1, n)$$

Правилата в  $\Gamma$  изглеждат по следния начин:

Ако  $\langle q_j, B_1 B_2 \dots B_m \rangle \in \delta(q_i, a, A)$ , то в  $P$  поставяме правилото

$$A(i, j) \rightarrow a B_1(i, n_1) B_2(n_1, n_2) \dots B_m(n_{m-1}, j)$$

за всяка редица  $n_1, n_2, \dots, n_{m-1}$  за която  $0 \leq n_k \leq n$ . Ако  $\langle q_j, \varepsilon \rangle \in \delta(q_i, a, A)$ , то в  $P$  поставяме правилото  $A(i, j) \rightarrow a$ .

Следователно

$$\begin{aligned} \omega \in T(M) &\iff \langle q_n, \varepsilon \rangle \in \delta(q_0, \omega, z_0) \iff z_0(1, n) \models \omega \iff \\ &\iff \omega \in L(\Gamma). \quad \square \end{aligned}$$

**Теорема 3.2.3** Ако  $L$  е безконтекстен език, а  $R$  е автоматен език, то  $L \cap R$  – безконтекстен език.

**Доказателство:** Нека  $L$  се разпознава от недетерминирания магазинен автомат  $M = \langle K, V, W, \delta, q_0, z_0, F \rangle$  и  $R$  се разпознава от недетерминирания краен автомат  $\tilde{M} = \langle \tilde{K}, V, \tilde{\delta}, \tilde{q}_0, F \rangle$ . Да построим следния недетерминиран магазинен автомат

$$M \times \tilde{M} = \langle K \times \tilde{K}, V, W, \delta', (q_0, \tilde{q}_0), z_0, F \times \tilde{F} \rangle,$$

където функцията  $\delta'$  се определя по следния начин

$$\delta'((q, \tilde{q}), a, A) = \{ \langle (q', \tilde{q}'), x \rangle \mid \langle q', x \rangle \in \delta(q, a, A) \text{ и } \tilde{q}' \in \tilde{\delta}(\tilde{q}, a) \},$$

$$\delta'((q, \tilde{q}), \varepsilon, A) = \{ \langle (q', \tilde{q}'), x \rangle \mid \langle q', x \rangle \in \delta(q, \varepsilon, A) \text{ и } \tilde{q}' = \tilde{q} \}.$$

Индуктивно, по броя на преминаванията от състояние в състояние на автомата  $M \times \tilde{M}$  може да се докаже, че  $\omega \in (M \times \tilde{M})$ , точно тогава, когато  $\omega \in L \cap R$ .  $\square$

Както при автоматните езици, се доказва, че ако  $L$  и  $R$  са два безконтекстни езици, то  $L \cup R$  и  $LR$  са безконтекстни езици.

Следващият пример показва, че  $L \cap R$  не винаги е безконтекстен език, ако  $L$  и  $R$  са безконтекстни.

**Пример 3.2.2** Ако  $L = \{a^n b^n c^m | m, n \geq 1\}$  и  $R = \{a^m b^n c^n | m, n \geq 1\}$ , то тези два езика са безконтекстни, защото се пораждат от грамматики със следните правила за:

$$L \quad - \quad \{S \rightarrow AB, A \rightarrow aAb|ab, B \rightarrow cB|c\}$$

$$R \quad - \quad \{S \rightarrow AB, B \rightarrow bBc|bc, A \rightarrow ab|a\}.$$

Сечението  $L \cap R$  е езикът  $L \cap R = \{a^n b^n c^n | n \geq 1\}$ . По-нататък ще покажем, че той не е безконтекстен.

С подходящи примери може да се покаже, че ако  $L$  е безконтекстен език, то  $\overline{L} = V^* \setminus L$  не винаги е безконтекстен език.

Когато функцията на преходите  $\delta$  на един НМА е дефинирана от  $K \times (V \cup \{\varepsilon\}) \times W^*$  в множеството  $K \times W^*$ , а не в  $P(K \times W^*)$  и от  $\delta(q, \varepsilon, z) \neq \emptyset$  да следва  $\delta(q, a, z) = \emptyset$  при  $a \in V$ ,

$$\delta(q, a, A) = \emptyset,$$

то автоматът се нарича *детерминиран магазинен автомат (ДМА)*. Език, който се разпознава от някой ДМА се нарича *детерминиран*.

Очевидно езикът от Пример 3.2.1 е породен от ДМА, въпреки че не е автоматен език. От друга страна всеки автоматен език може да се разпознае от ДМА. Наистина всеки краен автомат е еквивалентен на ДМА, който не използва своя магазин.

Оказва, се че не всеки безконтекстен език, може да се разпознае от някои ДМА т.е. не всеки безконтекстен език е детерминиран.

За детерминирани езици са в сила следните твърдения.

(i) Ако  $L$  е детерминиран, то и  $\overline{L} = V^* \setminus L$  е детерминиран език.

(ii) Ако  $L$  и  $R$  са детерминирани езици, то не винаги  $L \cup R, L \cap R, L^*$  и  $LR$  са детерминирани.

### З а д а ч и

1. Да се построи НМА, който разпознава, чрез крайни състояния, следните езици:



- а)  $\{a^n b^{2n} \mid n \geq 1\}$ ;
  - б)  $\{\omega \in \{a, b\}^* \mid \text{в } \omega \text{ има равен брой } a\text{-та и } b\text{-та}\}$ ;
  - в)  $\{a^n b^n a^n \mid n \geq 1\}$ .
2. Да се построи НМА, който разпознава, чрез празен магазин, следните езици:
- а)  $\{\omega \in \{a, b\}^* \mid \text{за всяко } \omega_1 \in \{a, b\}^*, \omega \neq \omega_1 \omega_1\}$ ;
  - б)  $\{\omega \in \{a, b\}^* \mid \text{съществува } \omega_1 \in \{a, b\}^*, \text{ така че } \omega = \omega_1 \omega'_1, \text{ където } \omega'_1 \text{ е думата } \omega_1 \text{ написана в обратен ред}\}$ .
3. Да се докаже, че ако  $L$  е детерминиран език, а  $R$  – автоматен, то  $L \cap R$  е детерминиран език.
4. Да се докаже, че езикът  $L = \{\omega c \omega' \mid \omega \in \{a, b\}^*\}$  е детерминиран език над азбуката  $\{a, b, c\}$ .
5. Да се построи НМА, които разпознава езика, породен от граматиката с правила
- $$S \rightarrow aA|aBB, \quad A \rightarrow Ba|Sb, \quad B \rightarrow bAS|\varepsilon.$$
6. Да се построи недетерминиран магазинен автомат, който разпознава:
- а) всички аритметични изрази на три променливи(букви);
  - б) езика  $L = \{a^k b^l a^m b^n \mid l \text{ и } n \text{ четни, } k \text{ и } m \text{ цели неотрицателни}\}$ .