③ Крайни автомати. Регулярни езици. Теорема на Клини

ДКА $A = \cancel{\langle \Sigma, Q, s, F, \delta \rangle}\ \langle Q, \Sigma, s, F, \delta \rangle$

- Σ - кр. мн-во наречено азбука, а неговите ел. - букви
- Q - кр. мн-во от състояния
- $s \in Q$ - стартово състояние
- $F \subseteq Q$ - финални състояния
- $\delta : Q \times \Sigma \to Q$ - частична / ~~тотална~~ функция на преходите
  тотална → тотален краен авт.

Дефинираме $\delta^* : Q \times \Sigma^* \to Q$ — дума - подредица от букви $\alpha = a_1 a_2 \ldots a_n \quad a_i \in \Sigma$

$$\delta^*(q, \alpha) = \begin{cases} q & \alpha = \varepsilon \\ \delta(\delta^*(q, \beta), a) & \alpha = \beta a \end{cases}$$

дума; ? $a \in \Sigma$, $a \in \Sigma^*$?

- Мн-вото от всички думи, които се разпознават от A и нарича език на A

$$L(A) = \{ \alpha \in \Sigma^* \mid \delta^*(s, \alpha) \in F \}$$

НКА

$$N = \cancel{\langle \Sigma, Q, s, F, \delta \rangle} = \langle Q, \Sigma, s, F, \delta \rangle$$

Q - кр. мн. от състояния

Σ - азбука

$s \in Q$ - стартово състояние

$F \subseteq Q$ - мн-во от финални състояния

$\delta : Q \times \Sigma \to 2^Q$ (връща мн-во от състояния)

Разглеждаме разширение на δ ? $\delta^* : Q \times \Sigma^* \to 2^Q$?

$$\delta^*(R, \alpha) = \begin{cases} R, & \text{ако } \alpha = \varepsilon \\ \bigcup \{ \delta(q, a) \mid q \in \delta^*(R, \beta) \}, & \text{ако } \alpha = \beta a \end{cases}$$

Език на НКА:

$$L(N) = \{ w \in \Sigma^* \mid \delta^*(\{s\}, w) \cap F \neq \emptyset \}$$

Регулярни операции   Наименуваме следните ~~операции~~ символи които използвахме по-рано

ε - празната дума - дума с нулева дължина

* - итерация (Звезда на Клини) - $L = \{a\}$ $L^* = \{\varepsilon, a, aa, aaa, \dots\}$ всички възможни конкат.

• - конкатенация - $L_1 = \{a\}, L_2 = \{b\} \Rightarrow L_1 . L_2 = \{a, b\}$ свързване на 2 думи последователно

∨ - обединение - обединение на 2 мн-ва

Нека Σ е кр. азбука и $\Sigma \cap \{\varepsilon, *, ., \vee, (,)\} = \emptyset$

Рег. израз α над Σ дефинираме индуктивно:

а) x за всяко $x \in \Sigma \cup \{\varepsilon\}$ е рег. израз

б) ако α и β са рег. изрази, то $\alpha^*, \alpha.\beta, \alpha \vee \beta$ са рег. изрази

в) няма рег. езици, които за не следват горните 2

Регулярен език $L(\alpha)$ за рег. израз α е функцията L описваща езика, който рег. израз α задава

Дефинираме L индуктивно:

а) $L(x) = \{x\}$ за $\forall x \in \Sigma \cup \{\varepsilon\}$ е рег. език

б) Ако $L(\alpha)$ и $L(\beta)$ са рег. езици за рег. изрази α и β, то $L(\alpha^*) = (L(\alpha))^*$, $L(\alpha.\beta) = L(\alpha).L(\beta)$ и $L(\alpha \vee \beta) = L(\alpha) \cup L(\beta)$ са рег. езици.

в) няма др., които да не следват горните две

Пример:

Рег. израз	Рег. език
$a \vee \beta$	$\{a, b\}$
$a.(a \vee \beta)$	$\{aa\}, \{a, b\}$
a^*	$\{\varepsilon, a, aa, a^3 \dots\}$
$a^*(a \vee b)$	$\{a, b, aa, ab, a^2a, \dots$

Представяне на всеки НКА с ДКА

Th: За всеки НКА N, съществува еквивалентен на него ДКА A т.ч. $L(N)=L(A)$

Доказателство: (скица)

Нека $N=\langle Q,\Sigma,s,F,\delta\rangle$ е Н.К.А.

Ще построим еквивалентен ДКА

$$A=\langle 2^Q,\Sigma,\{s\},F',\delta'\rangle$$

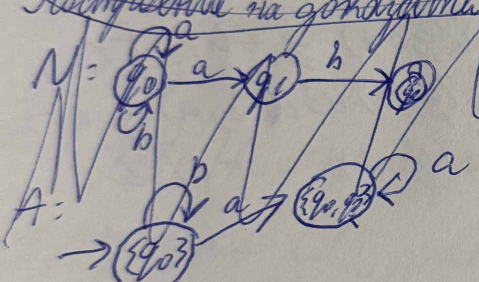
$$F'=\{M\subseteq Q \mid M\cap F\neq\emptyset\}$$

$$\delta'(M,a)=\bigcup_{q\in M}\delta(q,a)$$

твърдя, че $L(A)=L(N)$. Наистина

$$L(N)=\{w\in\Sigma^* \mid \delta^*(\{s\},w)\cap F\neq\emptyset\}=\{w\in\Sigma^* \mid \delta^*(\{s\},w)\in F\}=L(A)$$

~~Построение на доказателството (едва ли трябва)~~



Затвореност относно рег. опер.

- обединение

$A_1=\langle Q_1,\Sigma,s_1,F_1,\delta_1\rangle$, $A_2=\langle Q_2,\Sigma,s_2,F_2,\delta_2\rangle$

Б.О.О. $Q_1\cap Q_2=\emptyset$

$A=\langle Q_1\cup Q_2\cup\{s\},\Sigma,s,F',\delta'$

$s\notin Q_1\cup Q_2$

$$F'=\begin{cases} F_1\cup F_2\cup\{s\} & s_1\in F_1 \lor s_2\in F_2 \\ F_1\cup F_2 & \text{иначе}\end{cases}$$

$$\delta'(q,a)=\begin{cases}\delta_1(q,a) & q\in Q_1\\ \delta_2(q,a) & q\in Q_2\\ \delta_1(s_1,a)\cup\delta_2(s_2,a) & q=s\end{cases}$$

Твърдя, че $L(A) = L(A_1) \cup L(A_2)$

1сл. $w = \varepsilon$ $\quad w \in L(A) \iff \delta^*(\{s\}, \varepsilon) \subseteq F' \iff \{s\} \subseteq F' \iff s \in F' \iff s_1 \in F_1 \vee s_2 \in F_2$

$\iff w \in L(A_1) \vee w \in L(A_2)$

2сл. $w = au$, $a \in \Sigma$, $u \in \Sigma^*$

$au \in L(A) \iff \delta^*(\delta'(\{s\}, a), u) \subseteq F' \iff$

$(\exists q \in Q_1 \cup Q_2)(\delta'(s, a) \ni q \wedge \delta'^*(q, u) \subseteq F') \iff$

$[(\exists q \in Q_1)(\delta'(s,a) \ni q \wedge \delta'^*(q, u \subseteq F_1) \vee (\exists q \in Q_2)(\delta'(s,a) \ni q \wedge \delta'^*(q,u) \subseteq F_2)$

$\iff w \in L(A_1) \vee w \in L(A_2)$

- **Конкатенация**

$A_1 = \langle Q_1, \Sigma, s_1, F_1, \delta_1 \rangle$, $A_2 = \langle Q_2, \Sigma, s_2, F_2, \delta_2 \rangle$

Б.О.О. $Q_1 \cap Q_2 = \emptyset$

$A = \langle Q_1 \cup Q_2, \Sigma, s_1, F', \delta' \rangle$

$$F' = \begin{cases} F_1 \cup F_2 & s_2 \in F_2 \\ F_2 & \text{иначе} \end{cases} \qquad \delta'(q,a) = \begin{cases} \delta_1(q, a) & q \in Q_1 \setminus F_1 \\ \delta_1(q,a) \cup \delta_2(s_2, a) & q \in F_1 \\ \delta_2(q, a) & \text{else} \end{cases}$$

$L(A) = L(A_1) \cdot L(A_2)$

- **Звезда на Клини**

$A_1 = \langle Q_1, \Sigma, s_1, F_1, \delta_1 \rangle$

$A = \langle Q_1 \cup \{s\}, \Sigma, s, F_1 \cup \{s\}, \delta' \rangle$

$s \notin Q_1$

$$\delta'(q,a) = \begin{cases} \delta_1(q,a) & q \in Q_1 \setminus F_1 \\ \delta_1(q,a) \cup \delta_1(s_1,a) & q \in F \\ \delta_1(s_1, a) & q = s \end{cases}$$

Th Клини

v_1: Един език е регулярен т.т.к. се разпознава от краен автомат

v_2. Всеки автоматен език се описва с рег. израз т.е. е регулярен

Доказателство:

Нека $L = L(A)$ за A - ДКА.

Да фиксираме изброяване на състоянията $Q = \{q_0, q_1 \dots q_{n-1}\}$ $\underset{\{s\}}{q_{start}} = q_0$

$L(i,j,k)$ - мн-во от тези думи, които могат да се разпознаят от автомата по път, започващ от q_i и завършващ в q_j и междинните състояния има индекс $< k$

$$L(A) = \cup \{ L(0,j,|Q|) \mid q_j \in F = \bigcup_{q_j \in F} \{L(0,j,|Q|)$$

С индукция по k ще покажем, че $L(i,j,k)$ се описва от рег. израз.

База $k = 0$

- $i \neq j$ $L(i,j,0) = \{ a \in \Sigma \mid \delta(q_i, a) = q_j \}$
- $i = j$ $L(i,j,0) = \{\varepsilon\} \cup \{ a \in \Sigma \mid \delta(q_i, a) = q_j \}$

И.П. $\forall_i \forall_j$ $L(i,j,k)$ е регулярен за някое k.

Ще разгледаме $L(i,j,k+1)$ за произволни i и j

- нека $w \in L(i,j,k+1)$
  - q_k не е мр вътрешните състояния, участващи в изчислението на w
    $\Rightarrow w \in L(i,j,k)$
  - q_k е - - - - - - - -
    $\Rightarrow w \in L(i,k,k) . L(k,k,k)^* . L(k,j,k)$

$\Rightarrow L(i,j,k+1) = L(i,j,k) \cup L(i,k,k) . (L(k,k,k))^* . L(k,j,k)$

от и.п. са рег. $\Rightarrow$ се описват от рег.

$\Rightarrow L(i,j,k+1)$ е рег. (описва се от рег израз

$\Rightarrow L(A)$ — // —

Лема за покачването

Ако L е рег. език, то:

$\exists p \in \mathbb{N}^+$, такава че за $\forall$ дума $w \in L$, ако $|w| \geq p$, то

$\exists x, y, z \in \Sigma^* (w = xyz \wedge |y| \geq 1 \wedge |xy| \leq p \wedge (\forall i \in \mathbb{N} (xy^iz \in L))$

Д-во: Щом L е рег., то L е автоматен.

Нека $A = \langle Q, \Sigma, s, F, \delta \rangle$ е КДА за L. Полагаме $p = |Q|$.

Разглеждаме дума от L с дължина k ($k \geq p = |Q|$). Първите k стъпки от изчислението на думата:

$q_{start} \xrightarrow{a_1} q_1 \xrightarrow{a_2} \dots \xrightarrow{a_p} q_p$

участват $p+1$ състояния $\Rightarrow$ по Дирихле $(\exists i, j)(0 \leq i \leq j \leq n)(q_i = q_j)$

Разделяме w на 3 части

$\underbrace{a_1 \dots a_i}_{x} \underbrace{a_{i+1} \dots a_j}_{y} \underbrace{a_{j+1} \dots a_n}_{z}$

- $|y| \geq 1$ защото $i < j$
- $|xy| = j \leq p$

$\Rightarrow$ наистина щом $xyz \in L$, то $xz, xyyz \dots \in L$ $(\forall i \in \mathbb{N})(xy^iz \in L)$

Пример за нерегулярен език:

$L_1 = \{a^n b^n \mid n \in \mathbb{N}\}$ $\qquad L_2 = \{w w^{rev} \mid w \in \Sigma^*\}$

Пример за рег. език:

$L_1 = \{w \in \{a\}^* \mid w$ има поне 1 $a\}$

Минимизация на състоянията

Деф.: КДА A_M, разпознаващ автоматния език L, наричаме минимален за език L, ако за всеки друг автомат A, разпознаващ L, е изпълнено $|Q_M| \leq |Q|$, където Q_M и Q са мн-вата от състояния на A_M и A.

КДА A_1 и A_2 се нар. еквивалентни ако $L(A_1) = L(A_2)$

Минимизация на състоянията – задачата за намиране на минималния автомат за автоматния език L по зададен КДА A.

Релация на Майхил-Нероуд за L

$$R_L = \{ \langle x, y \rangle \in \Sigma^* \times \Sigma^* \mid \forall z \in \Sigma^* : xz \in L \Leftrightarrow yz \in L \}$$

Автоматна релация за $A = \langle Q, \Sigma, s, F, \delta \rangle$, $L(A) = L$

$$R_A = \{ \langle x, y \rangle \in \Sigma^* \times \Sigma^* \mid \delta^*(s, x) = \delta^*(s, y) \}$$

R_L и R_A са рел на еквивалентност.

Тв.1: R $\forall w \in \Sigma^* \; [w]_{R_A} \subseteq [w]_{R_L}$ (R_A прецизира R_L)

Автомат на Нероуд

$$M = \langle \Sigma, \{[w_1] \dots [w_k]\}, [\varepsilon], F', \delta' \rangle$$

$$F' = \{ [w] \mid w \in F \}$$

$$\delta'([w], a) = [wa]$$

Th Нероуд

L е регулярен $\Leftrightarrow |R_L| < \infty$

Док: ($\Rightarrow$) Разглеждаме обратното

$|R_L| = \infty \Rightarrow L$ не е рег.

Да допуснем, че L е рег.

$\Rightarrow \exists$ КДА A $A = \langle Q, \Sigma, s, F, \delta \rangle$ $L(A) = L$

R_A прецизира R_L

$\Rightarrow |Q| \geq |R_A| \geq |R_L| = \infty$

Но Q трябва да е крайно $\Rightarrow$ противоречие

L е рег.

($\Leftarrow$) Нека $|R_L| < \infty$

$\Rightarrow$ автоматът на Нероуд разпознава L

$\Rightarrow L$ е автоматен $\Leftrightarrow L$ е регулярен

Алг. за констр. на мин. авт по даден ДКА.

Нека $A = \langle Q, \Sigma, s, F, \delta \rangle$ е КДА

$L(q) = \{ w \in \Sigma^* \mid \delta^*(q, w) \in F \}$

Тогава $L(s) = L(A)$.

$\equiv_A \subseteq Q \times Q \qquad p \equiv_A q \iff L(p) = L(q)$

Търсим класовете на екв. на $\equiv_A$

$p \equiv_A^n q \iff L_A^n(p) = L_A^n(q)$

$L_A^n(q) = \{ w \mid |w| \leq n$ и $\delta^*(q, w) \in F \}$

$\equiv_A^n$ е апроксимация на $\equiv_A$

$\equiv_A^{n+1}$ е по фина от $\equiv_A^n$

понеже

$p \equiv_A^{n+1} q \iff p \equiv_A^n q \vee (\forall a \in \Sigma)(\delta(p, a) \equiv_A^n \delta(q, a))$

Класовете на екв. на $\equiv_A$ са състоянията на мин. авт.