4. Крайни автомати. Регулярни езици

1. **Детерминирани крайни автомати**

Деф: Детерминиран краен автомат

Детерминиран краен автомат (ДКА) е петорка $\mathcal{A} = \langle \Sigma, Q, q_{start}, \delta, F \rangle$, където:

- Σ е крайно множество от символи (азбука)
- Q е крайно множество от състояния
- $\delta: Q \times \Sigma \to Q$ е тотална функция, която ще наричаме *функция на преходите*
- $q_{start} \in Q$ е начално състояние
- $F \subseteq Q$ е множество от финални състояния

Деф: Разпознаване

Нека $\alpha \in \Sigma^*$ е дума, където $\alpha = a_0 a_1 \dots a_{n-1}$. Казваме, че α се **разпознава** от автомата $\mathcal{A}$, ако съществува редица от състояние $q_0, q_1, \dots, q_n$, т.ч.

- $q_0 = q_{start}$ - начално състояние на автомата
- $\delta(q_i, a_i) = q_{i+1}$ за всяко $i = 0, \dots, n-1$
- $q_n \in F$

Казваме, че $\mathcal{A}$ **разпознава** езика L, ако $\mathcal{A}$ разпознава точно думите от L, т.е.
$L = \{ \alpha \in \Sigma^* \mid \mathcal{A}$ разпознава $\alpha\}$, където:

- L - формален език над крайна азбука Σ, $L \subseteq \Sigma^*, \emptyset \subseteq \Sigma^*$
- Σ^* - изброимо безкраǐно множество от всички думи w над Σ
- Дума над Σ - крайна редица от символи от Σ

Тогава казваме, че езикът L е **автоматен.**

Деф: Разширена функция на преходите

Разширена функция на преходите δ^*: $Q \times \Sigma^* \to \mathrm{Q}$ дефинирана за всяко $q \in Q$ и $\alpha \in \Sigma^*$ по следния начин:

- Ако $\alpha = \varepsilon$, то $\delta^*(q, \varepsilon) \stackrel{\text{def}}{=} q$
- Ако $\alpha = \beta a$, то $\delta^*(q, \beta a) \stackrel{\text{def}}{=} \delta(\delta^*(q, \beta), a)$

Език на (разпознаван от) автомат $\mathcal{A}$ обикновено означаваме с $\mathcal{L}(\mathcal{A}) \stackrel{\text{def}}{=} \{\alpha \in \Sigma^* \mid \delta^*(q_{start}, \alpha) \in F\}$.
Казваме, че L е **автоматен**, ако $L = \mathcal{L}(\mathcal{A})$

2. **Недетерминирани крайни автомати**

Деф: Недетерминиран краен автомат

Недетерминиран краен автомат (НКА) е петорка $\mathcal{N} = \langle \Sigma, Q, Q_{start}, \Delta, F \rangle$, където:

- Σ е крайно множество от символи (азбука)
- Q - крайно множество от състояния
- $\Delta: Q \times \Sigma \to \mathcal{P}(Q)$ е функция на преходите. Възможно е за двойка $(q, a) \in (Q, \Sigma)$: $\Delta(q, a) = \emptyset$
- $Q_{start} \subseteq Q$ - множество от начални състояния
- $F \subseteq Q$ - множество от финални състояния

Деф: Разширена функция на преходите

Разширена функция на преходите Δ^*: $\mathcal{P}(Q) \times \Sigma^* \to \mathcal{P}(\mathrm{Q})$ дефинираме за произволно множество от състояния $R \subseteq Q$ и дума $\alpha \in \Sigma^*$ по следния начин:

- Ако $\alpha = \varepsilon$, то $\Delta^*(R, \varepsilon) \stackrel{\text{def}}{=} R$
- Ако $\alpha = \beta a$, то $\Delta^*(R, \beta a) \stackrel{\text{def}}{=} \bigcup \{\Delta(p, a) \mid p \in \Delta^*(R, \beta)\}$

Деф: Език, разпознаван от недетерминиран краен автомат $\mathcal{L}(\mathcal{N}) \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid \Delta^*(Q_{start}, \omega) \cap F \neq \emptyset\}$

Деф: Стъпка

Релацията $\vdash_{\mathcal{N}}: Q \times \Sigma^* \to Q \times \Sigma^*$ е т.ч. $(q, a\beta) \vdash_{\mathcal{N}} (p, \beta)$ т.с.т.к. $p \in \Delta(q, a)$;

$q_i \stackrel{a}{\to} q_j$ т.с.т.к. $p \in \Delta(q, a)$

Деф: Релацията $\vdash_{\mathcal{N}}^*$, определяща работата на автомата:

$\vdash_{\mathcal{N}}^*: Q \times \Sigma^* \to Q \times \Sigma^*$ е т.ч. за $q, p \in Q$, $\alpha, \beta \in \Sigma^*$: $(q, \alpha\beta) \vdash_{\mathcal{N}}^* (p, \beta)$ т.с.т.к. $p \in \Delta^*(\{q\}, \alpha)$

$q \stackrel{\alpha}{\Rightarrow} p$ т.с.т.к. $p \in \Delta^*(\{q\}, \alpha)$

3. Представяне на всеки недетерминиран краен автомат с детерминиран

Твърдение 1: За всеки две думи $\alpha, \beta \in \Sigma^*$ и всяко $R \subseteq Q$, $\Delta^*\big(\mathrm{R}, \alpha\beta\big) = \Delta^*\big(\Delta^*(\mathrm{R}, \alpha), \beta\big)$

- Доказва се с индукция по β

Теорема: (Рабин-Скот) За всеки НКА $\mathcal{N}$ съществува еквивалентен на него ДКА $\mathcal{D}$, т.е. $\mathcal{L}(\mathcal{N}) = \mathcal{L}(\mathcal{D})$

Д-во:

Нека $\mathcal{N} = \langle \Sigma, Q, Q_{start}, \Delta, F\rangle$.

Ще построим детерминиран краен автомат $\mathcal{D} = \langle \Sigma, Q', q_{start}, \delta, F'\rangle$, за който $\mathcal{L}(\mathcal{N}) = \mathcal{L}(\mathcal{D})$ по следния начин:

- $Q' \stackrel{\text{def}}{=} \{R \mid R \subseteq Q\}$
- За произволна буква $a \in \Sigma$ и произволно $R \subseteq Q$:
  $$\delta\Big(\underbrace{R}_{\text{състояние}}, \mathrm{a}\Big) \stackrel{\text{def}}{=} \Delta^*\Big(\underbrace{R}_{\text{множество}}, a\Big)$$
- $q_{start} \stackrel{\text{def}}{=} \mathrm{Q}_{start}$
- $F' \stackrel{\text{def}}{=} \{R \in Q' \mid R \cap F \neq \emptyset\}$

Ще докажем, че за произволна дума α и произволно множество $R \subseteq Q$ е изпълнено следното равенство (*):

$$\Delta^*\Big(\underbrace{R}_{\text{множество}}, \alpha\Big) = \delta^*\Big(\underbrace{R}_{\text{състояние}}, \alpha\Big)$$

Индукция по дължината на думата α:

- Ако $|\alpha| = 0$, т.е. $\alpha = \varepsilon$, то от дефиницията на Δ^* и δ^* имаме, че за всяко $R \subseteq Q$ е изпълнено:
  $\Delta^*(R, \varepsilon) = R = \delta^*(R, \varepsilon)$
- И.П: Да приемем, че (*) е изпълнено за всички думи α с дължина n, т.е.
  $(\forall \alpha \in \Sigma^n)(\forall R \subseteq Q)\big[\Delta^*(R, \alpha) = \delta^*(R, \alpha)\big]$
- Нека α има дължина $n + 1$, т.е. $\alpha = \beta a$, където $|\beta| = n$ и $a \in \Sigma$. Тогава
  $$\delta^*(R, \beta a) \underbrace{=}_{\substack{\text{деф}\\ \text{на } \delta^*}} \delta\left(\delta^*(R, \beta a)\right) \underbrace{=}_{\substack{\text{И.П.}\\ \text{за } \beta}} \delta\big(\Delta^*(R, \beta), a\big) \underbrace{=}_{\substack{\text{деф}\\ \text{на } \delta}} \Delta^*\big(\Delta^*(R, \beta), a\big) \underbrace{=}_{\substack{\text{Тв.1}\\ R\subseteq Q}} \Delta^*(R, \beta a)$$

Доказахме равенството (*). Лесно се съобразява, че:

$$\omega \in \mathcal{L}(\mathcal{D}) \underbrace{\Leftrightarrow}_{\substack{\text{Деф}\\ \mathcal{L}(\mathcal{D})}} \delta^*\big(\mathrm{q}_{start}, \omega\big) \in \mathrm{F}' \underbrace{\Leftrightarrow}_{(*)} \Delta^*\big(\mathrm{Q}_{start}, \omega\big) \cap \mathrm{F} \neq \emptyset \underbrace{\Leftrightarrow}_{\substack{\text{Деф}\\ \mathcal{L}(\mathcal{N})}} \omega \in \mathcal{L}(\mathcal{N})$$

4. Регулярни операции. Регулярни езици

Нека Σ е азбука.

Деф: Конкатенация на думи (индуктивна дефиниция)

$u, v \in \Sigma$

- $u = \varepsilon$, то $\varepsilon \cdot v = v \cdot \varepsilon = v$
- $u = a_1 \dots a_n$ и $v = b_1 \dots b_k$, то $u \cdot v = a_1 \dots a_n b_1 \dots b_k$

Деф: Конкатенация на езици $\cdot$

$L_1, L_2 \subseteq \Sigma^*$, то $L_1 \cdot L_2 = \{u \cdot v \mid u \in L_1 \,\&\, v \in L_2\}$

Деф: Обединение $\cup$

$L_1, L_2 \subseteq \Sigma^*$, $\quad L_1 \cup L_2 = \{u \mid u \in L_1 \vee u \in L_2\}$

Деф: Допълнение $\bar{}$: $L \subseteq \Sigma^*$, $\bar{\mathrm{L}} = \Sigma^* \setminus L$

Деф: *Нека $L \subseteq \Sigma^*$.* Дефинираме $\mathrm{L}^{\mathrm{k}} = \underbrace{\mathrm{L} \cdot \mathrm{L} \cdot \ldots \cdot \mathrm{L}}_{\mathrm{k}}$ индук*тивно:*

- $L^0 \stackrel{\text{def}}{=} \{\varepsilon\}$
- $L^{n+1} \stackrel{\text{def}}{=} \{L^n \cdot L\}$

Деф: Звезда на Клини $*$

Нека $L \subseteq \Sigma^*$. Тогава $L^* = \bigcup_{n \in \mathcal{N}} L^n$; $\;L^+ = L \cdot L^*$

Деф: Регулярен език

Един език $L \subseteq \Sigma^*$ е **регулярен**, ако се получава от основните езици $\emptyset$; $\{\varepsilon\}$; $\{\mathrm{a}\}$ за $a \in \Sigma$

С помощта на регулярните операции $\cup$, $\cdot$, $*$ приложени краен брой пъти.

- (Ако L_1 и L_2 са регулярни езици, то и $L_1 \cup L_2$, $L_1 \cdot L_2$, L_1^* са регулярни езици)

5. Доказателство за затвореност на автоматните езици относно регулярните операции

Теорема: Автоматните езици са затворени относно регулярните операции $(\cdot,\ \cup,\ *,\ ^{-})$

Д-во:

За всяка от операциите ще покажем как може да построим краен автомат, разпознаващ езика, конструиран чрез съответната операция.

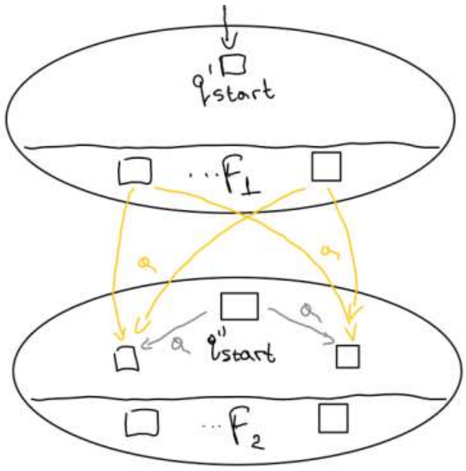
1. *Конкатенация*

Нека L_1 и L_2 са произволни автоматни езици. Ще докажем, че $L_1 \cdot L_2$ също е автоматен.
Нека $\mathcal{A}_1, \mathcal{A}_2$ са ДКА, т.ч.

- $\mathcal{A}_1 = \langle \Sigma, Q_1, \delta_1, q'_{start}, F_1 \rangle$, където $\mathcal{L}(\mathcal{A}_1) = L_1$
- $\mathcal{A}_2 = \langle \Sigma, Q_2, \delta_2, q''_{start}, F_2 \rangle$, където $\mathcal{L}(\mathcal{A}_2) = L_2$

Ще дефинираме автомата $\mathcal{N} = \langle \Sigma, Q, Q_{start}, \Delta, F \rangle$, така че $\mathcal{L}(\mathcal{N}) = \mathrm{L}_1 \cdot \mathrm{L}_2 = \mathcal{L}(\mathcal{A}_1) \cdot \mathcal{L}(\mathcal{A}_2)$:

- $Q \overset{\text{def}}{=} Q_1 \cup Q_2$
- $Q_{start} \overset{\text{def}}{=} \{q'_{start}\}$
- $F \overset{\text{def}}{=} \begin{cases} F_1 \cup F_2, & \text{ако } q''_{start} \in F_2 \\ F_2, & \text{иначе} \end{cases}$
- $\Delta(q,a) \overset{\text{def}}{=} \begin{cases} \{\delta_1(q,a)\}, & \text{ако } q \in Q_1 \setminus F_1 \;\&\; a \in \Sigma \\ \{\delta_1(q,a), \delta_2(q''_{start}, a)\}, & \text{ако } q \in F_1 \& \; a \in \Sigma \\ \{\delta_2(q,a)\}, & \text{ако } q \in Q_2 \;\&\; a \in \Sigma \end{cases}$



Първо ще докажем, че $\mathcal{L}(\mathcal{A}_1) \cdot \mathcal{L}(\mathcal{A}_2) \subseteq \mathcal{L}(\mathcal{N})$.

Нека $\alpha \in \mathcal{L}(\mathcal{A}_1), \beta \in \mathcal{L}(\mathcal{A}_2)$. Следователно

- $(q'_{start}, \alpha) \vdash^*_{\mathcal{A}_1} (q_1, \varepsilon)$ за някое $q_1 \in F_1$
- $(q''_{start}, \alpha) \vdash^*_{\mathcal{A}_2} (q_2, \varepsilon)$ за някое $q_2 \in F_1$

От деф на НКА $\mathcal{N}$ имаме, че $(q'_{start}, \alpha) \vdash^*_{\mathcal{N}} (q_1, \varepsilon)$ за някое $q_1 \in F_1$

Ако:

- $\beta = \varepsilon$, то $q''_{start} \in F_2$, значи $F_1 \subseteq F$
- $\beta = b\gamma$ за някое $\gamma \in \Sigma^*$, то имаме: $(q''_{start}, b\gamma) \vdash_{\mathcal{A}_2} (q, \gamma) \vdash^*_{\mathcal{A}_2} (q_2, \varepsilon)$ за някое $q_2 \in F_2$, където $q = \delta_2(q''_{start}, b)$

От деф. на НКА $\mathcal{N}$ имаме, че $(q, \gamma) \vdash^*_{\mathcal{N}} (q_2, \varepsilon)$ за някое $q_2 \in F_2$
Също, $q \in \Delta(q_1, b)$, защото $q_1 \in F_1$, т.е. $(q_1, b\gamma) \vdash_{\mathcal{N}} (q, \gamma)$

Следователно $(q_1, \beta) \vdash^*_{\mathcal{N}} (q_2, \varepsilon)$ за някое $q_2 \in F_2$.
Получихме: $(\mathrm{q}'_{\mathrm{start}}, \alpha\beta) \vdash^*_{\mathcal{N}} (q_1, \beta) \vdash^*_{\mathcal{N}} (q_2, \varepsilon)$ за някое $q_2 \in F_2$, т.е. $\alpha \cdot \beta \in \mathcal{L}(\mathcal{N})$

Сега ще докажем $\mathcal{L}(\mathcal{N}) \subseteq \mathcal{L}(\mathcal{A}_1) \cdot \mathcal{L}(\mathcal{A}_2)$

Нека $\omega \in \mathcal{L}(\mathcal{N})$, където $|\omega| = n$. Нека $(q_i)_{i=0}^{n}$ е редица от състояния, която описва приемащо изчисление на $\mathcal{N}$ върху ω, следователно:

- $q_0 = q_{start}$
- $q_{i+1} \in \Delta(q_i, \omega[i])$ за $i < n$
- $q_n \in F$

Ако:

- $q_n \in F_1$, от деф. на $\mathcal{N}$, $\varepsilon \in \mathcal{L}(\mathcal{A}_2)$, $q_i \in Q_1$, за $i \in \{0, \dots, n\}$, следователно $\omega \in \mathcal{L}(\mathcal{A}_1)$
- $q_n \in F_2$, от деф на $\mathcal{N}$ не можем да преминем от състояние от Q_2 в състояние от Q_1. Значи можем да разбием $(q_i)_{i=0}^{n}$ на непразни подредици:
  - $(q_i)_{i=0}^{l}$ - състоянията от Q_1, $(\mathrm{q_i})_{\mathrm{i=l+1}}^{\mathrm{n}}$ - състоянията от Q_2

  Нека $\omega_1 = \omega[:l]$ и $\omega_2 = \omega[l:]$.

  $(q_0, \omega) \vdash^*_{\mathcal{N}} (q_l, \omega[l:]) \vdash_{\mathcal{N}} (q_{l+1}, \omega[l+1:]) \vdash^*_{\mathcal{N}} (q_n, \varepsilon)$

От деф на $\mathcal{N}$:

- $(q_i)_{i=0}^{l}$ описва приемащо изчисление на $\mathcal{A}_1$ върху ω_1
- От $q_{l+1} \in \Delta(q_l, a_l)$, то $q_l \in F_1$ и $\delta_2\left(q''_{start}, a_l\right) = q_{l+1}$

$(q_0, \omega_1) \vdash^*_{\mathcal{A}_1} (q_l, \varepsilon)$. От $q_0 = q'_{start}$ и $q_t \in F_1$, то $\omega_1 \in \mathcal{L}(\mathcal{A}_1)$
$\left(q''_{start}, \omega_2\right) \vdash^*_{\mathcal{A}_2} (q_n, \varepsilon)$. От $q_n \in F_2$, то $\omega_2 \in \mathcal{L}(\mathcal{A}_2) \Rightarrow \omega \in \mathcal{L}(\mathcal{A}_1) \cdot \mathcal{L}(\mathcal{A}_2)$

1. *Обединение*

Нека L_1 и L_2 са произволни автоматни езици. Ще докажем, че $L_1 \cup L_2$ също е автоматен.
Нека $\mathcal{A}_1, \mathcal{A}_2$ са ДКА, т.ч.

- $\mathcal{A}_1 = \langle \Sigma, Q_1, \delta_1, S_1, F_1 \rangle$, където $\mathcal{L}(\mathcal{A}_1) = L_1$
- $\mathcal{A}_2 = \langle \Sigma, Q_2, \delta_2, S_2, F_2 \rangle$, където $\mathcal{L}(\mathcal{A}_2) = L_2$

Ще дефинираме автомата $\mathcal{N} = \langle \Sigma, Q, Q_{start}, \Delta, F \rangle$,
така че $\mathcal{L}(\mathcal{N}) = L_1 \cup L_2 = \mathcal{L}(\mathcal{A}_1) \cup \mathcal{L}(\mathcal{A}_2)$:

- $Q_{start} = \{S_1, S_2\}$
- $Q \overset{\text{def}}{=} Q_1 \cup Q_2$
- $F \overset{\text{def}}{=} F_1 \cup F_2$
- $\Delta(q, a) \overset{\text{def}}{=} \begin{cases} \{\delta_1(q, a)\}, & \text{ако } q \in Q_1 \text{ и } a \in \Sigma \\ \{\delta_2(q, a)\}, & \text{ако } q \in Q_2 \text{ и } a \in \Sigma \end{cases}$

Д-во:
Нека $\omega \in L_1 \cup L_2 \Rightarrow \omega \in L_1$ и/или $\omega \in L_2$. Индукция по ω:

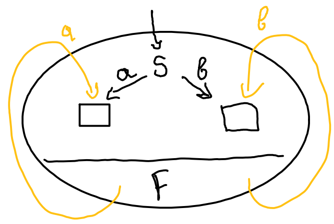
- $\omega = \varepsilon \Rightarrow \begin{cases} S_1 \in F_1 \Rightarrow \omega \in L_1 \\ S_2 \in F_2 \Rightarrow \omega \in L_2 \end{cases}$
- $\omega = au,\ a \in \Sigma,\ u \in \Sigma^*$
  - $\omega \in L_1$, след. $(S_1, au) \vdash_{\mathcal{A}_1} (q, u) \vdash^*_{\mathcal{A}_1} (f, \varepsilon),\ f \in F_1$
    Но същите състояния ги има и в $\mathcal{N}$,т.е. $(S_1, au) \vdash_{\mathcal{N}} (q, u) \vdash^*_{\mathcal{N}} (f, \varepsilon),\ f \in F_1 \subseteq F$
    Така $\omega = au \in \mathcal{L}(\mathcal{N})$
  - $\omega \in L_2$, аналогично

Нека $\omega \in \mathcal{L}(\mathcal{N})$. Индукция по ω:

- $\omega = \varepsilon$, значи $Q_{start} \cap F \neq \emptyset$. Нека $S \in Q_{start} \cap F,\ F = F_1 \cup F_2$
  $\Rightarrow S \in F_1 \vee S \in F_2 \Rightarrow \omega \in L_1 \vee \omega \in L_2 \Rightarrow \omega \in L_1 \cup L_2$
- $\omega = au, a \in \Sigma, u \in \Sigma^*$
  $\omega \in \mathcal{L}(\mathcal{N}) \Rightarrow (S, au) \vdash_{\mathcal{N}} (q, u) \vdash^*_{\mathcal{N}} (f, \varepsilon)$, за някое $S \in Q_{start}, f \in F$
  - Ако $q \in Q_1$, то $(S_1, au) \vdash_{\mathcal{A}_1} (q, u) \vdash^*_{\mathcal{A}_1} (f, \varepsilon),\ f \in F_1,\ \omega \in L_1$
  - Ако $q \in Q_2$ - аналогично

2. *Звезда на Клини*

Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$ е НКА, разпознаващ езика L.
Ще построим НКА $\mathcal{N} = \langle \Sigma, Q, \Delta, S, F \rangle$, т.ч. $\mathcal{L}(\mathcal{N}) = \left(\mathcal{L}(\mathcal{A})\right)^+$:



- $\Delta(q, a) \overset{\text{def}}{=} \begin{cases} \{\delta(q, a)\}, & \text{ако } q \notin F \\ \{\delta(q, a), \delta(S, a)\}, & \text{ако } q \in F \end{cases}$

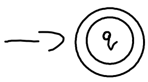
Д-во: Нека $\omega \in \left(\mathcal{L}(\mathcal{A})\right)^+$. Индукция по ω

- $\omega = \varepsilon \Rightarrow S \in F \Rightarrow \omega \in \mathcal{L}(\mathcal{N})$
- $\omega \neq \varepsilon, \omega = u_1 \dots u_n, u_i \in \mathcal{L}(\mathcal{A}), i = 1, \dots, n,\ \ u_i = a_i v_i,\ \ i = 1, \dots, n$
  $(S, a_1 v_1 u_2 \dots u_n) \vdash_{\mathcal{N}} (q_0, v_1 u_2 \dots u_n) \vdash^*_{\mathcal{N}} (f_0, u_2 \dots u_n) \vdash_{\mathcal{N}} \dots \vdash^*_{\mathcal{N}} (f_n, \varepsilon)$
  т.е. $\omega = a_1 v_1 a_2 v_2\ \dots a_n v_n \in \mathcal{L}(\mathcal{N})$

Нека $\omega \in \mathcal{L}(\mathcal{N})$

- $\omega = \varepsilon, \quad S \in F \Rightarrow \omega \in \mathcal{L}(\mathcal{A}) \subseteq \left(\mathcal{L}(\mathcal{A})\right)^+$
- $\omega = a_0 u_0 a_1 u_1 \dots a_n u_n$,
  $(S, a_0 u_0 a_1 u_1 \dots a_n u_n) \vdash_{\mathcal{N}} (q_0, u_0 a_1 u_1 \dots a_n u_n) \vdash^*_{\mathcal{N}} (f_0, a_1 u_1 \dots a_n u_n) \vdash_{\mathcal{N}} \dots \vdash^*_{\mathcal{N}} (f_n, \varepsilon)$
  $q_i \in \Delta(f_{i-1}, a_i),\ i \geq 1$ - нови преходи за $\mathcal{N}$
    Имаме, че $(S, a_i u_i) \vdash_{\mathcal{A}} (S, u_i) \vdash^*_{\mathcal{A}} f_i,\ f_i \in F \Rightarrow a_i u_i \in L,$ за $i \in \{0, \dots, n\}$
    Значи $\omega \in L^+ = \left(\mathcal{L}(\mathcal{A})\right)^+$

Построихме НКА, който разпознава $\left(\mathcal{L}(\mathcal{A})\right)^+$.
Следният автомат разпознава езика $\{\varepsilon\}$:



От вече доказаната затвореност относно обединението, имаме че $\{\varepsilon\} \cup L^+ = L^*$ също е автоматен

4. *Допълнение*
   Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$ е ДКА (**тотален**), разпознаващ езика L.

   Крайният автомат $\mathcal{N} = \langle \Sigma, Q, \Delta, S, Q \setminus F \rangle$ е т.ч. $\mathcal{L}(\mathcal{N}) = \overline{\mathcal{L}(\mathcal{A})} = \bar{L} = \Sigma^* \setminus L$

   Нека $\omega \in \mathcal{L}(\mathcal{A})$, значи $(S, \omega) \vdash^*_{\mathcal{A}} (f, \varepsilon)$ и $f \in F$, т.е. $(S, \omega) \vdash^*_{\mathcal{N}} (f, \varepsilon)$ и $f \notin Q \setminus F \Rightarrow \omega \notin \mathcal{L}(\mathcal{N})$

   Нека $\omega \notin \mathcal{L}(\mathcal{A})$, т.е. $(S, \omega) \vdash^*_{\mathcal{A}} (q, \varepsilon)$, $q \notin F$, значи $q \in Q \setminus F \Rightarrow (S, \omega) \vdash^*_{\mathcal{N}} (q, \varepsilon) \Rightarrow \omega \in \mathcal{L}(\mathcal{N})$
   $\Rightarrow \omega \notin \mathcal{L}(\mathcal{A}) \Leftrightarrow \omega \in \mathcal{L}(\mathcal{N})$

Останалите операции като сечение, разлика могат да се изразят, използвайки вече доказаните и следователно имаме затвореност и относно тях.

4. **Теорема на Клини**

Теорема: (Клини) L е регулярен $\Leftrightarrow$ L е автоматен
Д-во:

$\Rightarrow$ ДКА разпознават $\emptyset, \{\varepsilon\}$ и всеки език, съставен от една буква. Доказахме, че са затворени относно регулярните операции. Следователно всеки регулярен език се разпознава от краен автомат.

$\Leftarrow$ Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$ е ДКА. Ще конструираме регулярен език L, т.ч. $L = \mathcal{L}(\mathcal{A})$
Нека $Q = \{q_0, \dots, q_{n-1}\}$ е изброяване на състоянията и $S = q_0$.

Нека $L(i, j, k)$ е множеството от думи, които могат да се разпознаят от $\mathcal{A}$ по път, започващ от q_i и завършващ в q_j, като междинните състояния имат индекси $< k$

Тогава за $n = |Q|$, имаме $L(i, j, n) = \{\alpha \in \Sigma^* \mid \delta^*(q_i, \alpha) = q_j\}$. Така

$$\mathcal{L}(\mathcal{A}) = \bigcup \left\{ L(0, j, n) \mid q_j \in F \right\} = \bigcup_{q_j \in F} L(0, j, n)$$

Нека $P(k) = (\forall i < n)(\forall j < n)\left[L(i, j, k) \text{ е регулярен}\right]$
С индукция по k ще докажем, че $(\forall k \leq n)P(k)$:

- Нека $k = 0$, $q_i, q_j \in Q$ - произволни.
  - $i = j$, то $L(i, j, 0) = \{\varepsilon\} \cup \left\{a \in \Sigma \mid \delta(q_i, a) = q_j\right\}$
  - $i \neq j$, то $L(i, j, 0) = \left\{a \in \Sigma \mid \delta(q_i, a) = q_j\right\}$

    $L(i, j, 0)$ - краен език, следователно е регулярен
- ИП: Нека е в сила: $(\forall i < n)(\forall j < n)\left[L(i, j, k) \text{ е регулярен}\right]$
- Стъпка: Ще докажем $P(k+1)$:

  Нека $\alpha \in L(i, j, k+1)$, $\alpha = a_1 a_2 \dots a_s$, $a_i \in \Sigma$, $\mathrm{i} = 1, \dots, \mathrm{s}$
  Нека $q_{l_1}, q_{l_2}, \dots, q_{l_{s-1}} \in Q$ са т.ч. $\mathrm{q_i} \xrightarrow{a_1} q_{l_1} \xrightarrow{a_2} q_{l_2} \xrightarrow{a_3} \dots \xrightarrow{a_{s-1}} q_{l_{s-1}} \xrightarrow{a_s} q_j$
  - $q_k \notin q_{l_1}, q_{l_2}, \dots, q_{l_{s-1}} \Rightarrow \alpha \in L(i, j, k)$
  - $q_k \in q_{l_1}, q_{l_2}, \dots, q_{l_{s-1}}$. Нека q_k се среща m пъти и нека $\alpha = \alpha_1 \alpha_2 \dots \alpha_{m+1}$ т.ч.
    $q_i \overset{\alpha_1}{\Rightarrow} q_k \overset{\alpha_2}{\Rightarrow} q_k \to \cdots \to q_k \to \cdots \to q_k \overset{\alpha_{m+1}}{\Longrightarrow} q_j$
    За всяко $l = 1, \dots, m+1$, q_k не е вътрешно състояние за изчислението на α_l, т.е. индексите на всички вътрешни състояния са $< k$.
    Значи $\alpha_1 \in L(i, k, k)$, $\alpha_l \in L(k, k, k)$ *за* $l = 2, \dots, m$ *и* $\alpha_{m+1} \in L(k, j, k)$

    $$\Rightarrow \alpha \in L(i, k, k) \cdot L(k, k, k)^{m-1} \cdot L(k, j, k)$$

    Това разсъждение е за произволно m, значи ако q_k се среща между вътрешните състояния на изчислението на α, то

    $$\alpha \in L(i, k, k) \cdot L(k, k, k)^* \cdot L(k, j, k)$$

    $$\Rightarrow \mathrm{L}(i, j, k+1) \subseteq L(i, j, k) \cup L(i, k, k) \cdot L(k, k, k)^* \cdot L(k, j, k)$$

Имаме и обратното включване от определението на $L(i,j,k)$.
От ИП всички компоненти отдясно са регулярни езици, композирани чрез регулярни операции, следователно $L(i,j,k+1)$ също е регулярен език.

$\Rightarrow (\forall \mathrm{i}, \mathrm{j}, < |Q|)(\forall k \leq |Q|)$

5. **Формулировка и доказателство на лемата разрастване за регулярни езици (uvw-лема)**

Лема: Лема за покачването

Нека L е безкраен регулярен език. $\exists p \in \mathbb{N}, p \geq 1$, т.ч. $\forall \alpha \in L$, т.ч. $|\alpha| \geq p$:
$\exists u, v, w \in \Sigma^*: \alpha = uvw$ и
1. $v \neq \varepsilon$ $(|v| \geq 1)$
2. $|uv| \leq p$
3. $(\forall n \in \mathbb{N})[uv^n w \in L]$

Д-во:

Нека L е регулярен език, следователно е и автоматен. Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$ е т.ч. $L = \mathcal{L}(\mathcal{A})$ и нека $p = |Q|$. Нека $\alpha \in L$, $\alpha = a_1 a_2 \ldots a_k$ и $k \geq p$.

$$S \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \ldots \xrightarrow{a_p} q_p \ldots \xrightarrow{a_k} q_k$$

Да разгледаме първите р стъпки. Нека бележим $q_0 \coloneqq S$. Имаме, че $|Q| = p$, а в

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \ldots \xrightarrow{a_p} q_p$$

Участват $p+1$ състояния $q_0, q_1, \ldots, q_p$, то по принципа на Дирихле съществуват числа i, j за които $0 \leq i < j \leq p$ и $q_i = q_j$. Нека разделим α на три части:

$$\underbrace{a_1 \ldots a_i}_{u}\ \underbrace{a_{i+1} \ldots a_j}_{v}\ \underbrace{a_{j+1} \ldots a_k}_{w}$$

Имаме, че $|v| \geq 1$ и $|uv| = j \leq p$

$$(q_0, uvw) \vdash_{\mathcal{A}}^* (q_i, vw) \vdash_{\mathcal{A}}^* (q_j, w) \vdash_{\mathcal{A}}^* (q_k, \varepsilon)$$

За $n = 0$: Думата $uv^0 w = uw \in L$, от $q_i = q_j$ имаме: $q_0 \underbrace{\xrightarrow{a_1} q_1 \ldots \xrightarrow{a_i}}_{u} q_i \underbrace{\xrightarrow{a_{j+1}} q_{j+1} \ldots \xrightarrow{a_k}}_{w} q_k \in F$

За $n = 2$: Думата $uv^2 w = uvvw \in L$, защото:

$$q_0 \underbrace{\xrightarrow{a_1} q_1 \ldots \xrightarrow{a_i}}_{u} q_i \underbrace{\xrightarrow{a_{i+1}} q_{i+1} \ldots \xrightarrow{a_j}}_{v} \underbrace{q_j}_{q_i} \underbrace{\xrightarrow{a_{i+1}} q_{i+1} \ldots \xrightarrow{a_j}}_{v} \underbrace{q_j}_{q_i} \underbrace{\xrightarrow{a_{j+1}} q_{j+1} \ldots \xrightarrow{a_k}}_{w} q_k \in F$$

Аналогично за произволно $n \geq 2$ имаме:

$$q_0 \underbrace{\xrightarrow{a_1} q_1 \ldots \xrightarrow{a_i}}_{u} q_i \underbrace{\underbrace{\xrightarrow{a_{i+1}} q_{i+1} \ldots \xrightarrow{a_j}}_{v} q_i \underbrace{\xrightarrow{a_{i+1}} q_{i+1} \ldots \xrightarrow{a_j}}_{v} q_i \ldots q_i}_{n \text{ пъти}} \underbrace{\xrightarrow{a_{j+1}} q_{j+1} \ldots \xrightarrow{a_k}}_{w} q_k$$

Така за всяко естествено число n е изпълнено $uv^n w \in L$.

6. **Примери за нерегулярни езици**

1. $L = \{a^n b^n \mid n \geq 0\}$
   Допускаме, че е регулярен, по лемата $\Rightarrow$
   $\exists \mathrm{p} \in \mathbb{N}: \forall \alpha \in \mathrm{L}: |\alpha| \geq p$, то $\exists \mathrm{u}, \mathrm{v}, \mathrm{w} \in \Sigma^*: \alpha = \mathrm{uvw}, |v| \geq 1$, $|uv| \leq p$ и $uv^k w \in L, k \in \mathbb{N}$
   Нека $\alpha = a^p b^p$, от $|uv| \leq p$, то $uv = a^i, i \in \{1, \ldots, p\}$. Тогава $uw \in L$ по лемата, но $a^{p-i} b^p \notin L$.
   Противоречие

2. $L = \{a^n b^k \mid n < k\}$
   Чупи се с $\alpha = a^p b^{p+1}$ и $n = 2$.

7. **Формулировка и доказателство на теоремата на Майхил-Нероуд**

Деф: Еквивалентни думи в език и автомат (релация на Майхил-Хероуд)

Нека $L \subseteq \Sigma^*$ е език и $u, v \in \Sigma^*$.

Казваме, че u и v са еквивалентни спрямо L $(u \approx_L \mathrm{v})$ т.с.т.к. $\forall x \in \Sigma^* [\, ux \in L \Leftrightarrow vx \in L]$.

Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$, т.ч. ДКА.

Казваме, че u и v са еквивалентни спрямо $\mathcal{A}$ $(u \sim_{\mathcal{A}} \mathrm{v})$, ако $\delta^*(S, u) = \delta^*(S, v)$

$\approx_L$ и $\sim_{\mathcal{A}}$ са релация на еквивалентност. Бележим с $[x]_{\approx_L}$ класовете на еквив. породени от $x \in \Sigma^*$

Деф: Автомат на Майхил-Нероуд

За даден език L, дефинираме автоматът на Майхил-Нероуд $M_L = \left\langle \Sigma, Q_{M_L}, \delta_{M_L}, S_{M_L}, F_{M_L} \right\rangle$ по следния начин:

- $Q_{M_L} \stackrel{\text{def}}{=} \{[\alpha]_L \mid \alpha \in \Sigma^*\}$
- $\delta_{M_L}([\alpha]_L, \mathrm{b}) \stackrel{\text{def}}{=} [\alpha b]_L$ за всяка $\alpha \in \Sigma^*$ и $b \in \Sigma$
- $S_{M_L} \stackrel{\text{def}}{=} [\varepsilon]_L$
- $F_{M_L} \stackrel{\text{def}}{=} \{[\alpha]_L \mid [\alpha]_L \subseteq L\}$

Твърдение: Нека L е език. За всеки $\alpha, \beta \in \Sigma^*, x \in \Sigma$ е изпълнено:

$\alpha \approx_L \beta \Rightarrow \alpha x \approx_L \beta x$

Д-во:

$\alpha \approx_L \beta \Rightarrow (\forall z \in \Sigma^*)(\alpha z \in L \Leftrightarrow \beta z \in L) \Rightarrow (\forall xz \in \Sigma^*)(\alpha(xz) \in L \Leftrightarrow \beta(xz) \in L)$

$\Rightarrow (\forall x \in \Sigma^*)(\forall z \in \Sigma^*)\left((\alpha x)z \in L \Leftrightarrow (\beta x)z \in L\right) \Rightarrow (\forall x \in \Sigma)(\alpha x \approx_L \beta x)$

Твърдение: δ_{M_L} е добре дефинирано: $\delta^*_{M_L}([\alpha]_L, \beta) = [\alpha\beta]_L, \alpha, \beta \in \Sigma^*$

Д-во:

Индукция по β:

- $\beta = \varepsilon$: $\delta^*_{M_L}([\alpha]_L, \varepsilon) = [\alpha]_L = [\alpha\varepsilon]_L$
- $\beta = \mathrm{b}\beta_1$: $\delta^*_{M_L}([\alpha]_L, \mathrm{b}\beta_1) = \delta^*_{M_L}\left(\delta_{M_L}([\alpha]_L, \mathrm{b}), \beta_1\right) = \delta^*_{M_L}([\alpha b]_L, \beta) \stackrel{\text{ИП}}{=} [\alpha b\beta_1] = [\alpha\beta]$

Твърдение: $\mathcal{L}(M_L) = L$

Д-во:

$\omega \in \mathcal{L}(M_L) \Leftrightarrow \delta^*_{M_L}([\varepsilon]_L, \omega) \in \{[\omega]_L \mid \omega \in L\} \Leftrightarrow [\omega]_L \in \{[\omega]_L \mid \omega \in L\} \Leftrightarrow \omega \in L$

Твърдение: Ако L е регулярен език и $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$, т.ч. $\mathcal{L}(\mathcal{A}) = \mathrm{L}$ и $u, v \in \Sigma^*$ то:

$u \sim_{\mathcal{A}} v \Rightarrow u \approx_{L=\mathcal{L}(\mathcal{A})} v \quad \left(\forall \omega \in \Sigma^* \quad [\omega]_{\sim_{\mathcal{A}}} \subseteq [\omega]_L\right)$

Д-во:

Нека $u \sim_{\mathcal{A}} v$, значи $\delta^*(S, u) = \delta^*(S, v)$

$\Rightarrow \forall z \in \Sigma^*$: $\delta^*(S, uz) = \delta^*(\delta^*(S, u), z) = \delta^*(\delta^*(S, v), z) = \delta^*(S, vz) = \mathrm{q}$

$q \in F \Rightarrow uz \in L$ и $vz \in L$, също $q \notin F \Rightarrow uz \notin L$ и $vz \notin L$

$(uz \in L \Leftrightarrow vz \in L) \Rightarrow u \approx_L v$

Значи $\sim_{\mathcal{A}}$ прецизира $\approx_L$. Следствие е, че $|\sim_{\mathcal{A}}| \geq |\approx_L|$

Теорема: (Майхил-Нероуд)

Един език L е регулярен т.т.к. M_L има краен брой състояния. Освен това, M_L е минимален ДКА за L

Д-во:

Да допуснем, че L е регулярен и $\left|\{[\alpha]_L \mid \alpha \in \Sigma^*\}\right| = \infty$

От L - регулярен $\Rightarrow \exists$ ДКА $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$, т.ч. $\mathcal{L}(\mathcal{A}) = \mathrm{L}$

Имаме, че $\sim_{\mathcal{A}}$ прецизира $\approx_L \Rightarrow |\mathrm{Q}| \geq |\sim_{\mathcal{A}}| \geq |\approx_L| = \infty$. Противоречие

Но горната закономерност важи за всички ДКА, разпознаващи L. Следователно M_L е ДКА с най-малък брой състояния.

Обратно, ако $|M_L| < \infty$, то той разпознава език $L \Rightarrow \mathrm{L}$ е регулярен.

8. Алгоритъм за конструиране на минимален краен детерминиран тотален автомат, еквивалентен на даден ДКА

Нека $\mathcal{A} = \langle \Sigma, Q, \delta, S, F \rangle$ е ДКА.

Деф:

$\mathcal{L}_{\mathcal{A}}(q) = \{\omega \in \Sigma^* \mid \delta^*(q, \omega) \in F\}$

$\mathcal{L}^n_{\mathcal{A}}(q) = \{\omega \in \Sigma^* \mid |\omega| \leq n \ \&\ \delta^*(q, \omega) \in F\}$

$p \equiv_{\mathcal{A}} q \stackrel{def}{\Longleftrightarrow} \mathcal{L}_{\mathcal{A}}(p) = \mathcal{L}_{\mathcal{A}}(q)$

$p \equiv^n_{\mathcal{A}} q \stackrel{def}{\Longleftrightarrow} \mathcal{L}^n_{\mathcal{A}}(p) = \mathcal{L}^n_{\mathcal{A}}(q)$

Тогава $\mathcal{L}_{\mathcal{A}}(S) = \mathcal{L}(\mathcal{A})$.
Релациите $\equiv_{\mathcal{A}}^{n}$ са апроксимация на $\equiv_{\mathcal{A}}$.

За всяко n, $\equiv_{\mathcal{A}}^{n}$ е по-груба от $\equiv_{\mathcal{A}}^{n+1}$. Алгоритъмът ни строи $\equiv_{\mathcal{A}}^{n}$, докато не срещнем n, т.ч. $\equiv_{\mathcal{A}}^{n} = \equiv_{\mathcal{A}}^{n+1}$.
Тогава ще имаме, че $\equiv_{\mathcal{A}}^{n} = \equiv_{\mathcal{A}}$

$$p \equiv_{\mathcal{A}}^{n+1} q \iff p \equiv_{\mathcal{A}}^{n} q \ \& \ (\forall a \in \Sigma)\left(\delta(p,a) \equiv_{\mathcal{A}}^{n} \delta(q,a)\right)$$

1. За $\equiv_{\mathcal{A}}^{0}$, класовете на еквивалентност са F и $Q \setminus F$.
2. Нека предположим, че сме открили класовете за еквивалентност за $\equiv_{\mathcal{A}}^{n}$.
   Тогава $\forall p, q \in Q$: $p \equiv_{\mathcal{A}}^{n+1} q \Leftrightarrow p \equiv_{\mathcal{A}}^{n} q \ \& \ \forall a \in \Sigma \left(\delta(p,a) \equiv_{\mathcal{A}}^{n} \delta(q,a)\right)$. По този начин разбираме дали две състояния отиват в един и същ клас или не и така конструираме $\equiv_{\mathcal{A}}^{n+1}$
3. Продължаваме така, докато не получим $\equiv_{\mathcal{A}}^{n} = \equiv_{\mathcal{A}}^{n+1}$.

Алгоритъмът е краен, защото на всяка стъпка броят на класовете на еквивалентност се увеличава поне с 1, а не може да има повече класове на еквивалентност от състояния на D.