Rozšíření konečných automatů II
Automaty s ε-kroky

Definice 2.46. Nedeterministický konečný automat s ε-kroky je $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$, kde význam všech složek je stejný jako v definici NFA s výjimkou přechodové funkce δ. Ta je definována jako totální zobrazení $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \to 2^Q$.

Rozšířená přechodová funkce

Definujeme funkci $D_\varepsilon : Q \to 2^Q$ následujícím předpisem.
$D_\varepsilon(p)$ je nejmenší množina $X \subseteq Q$ taková, že platí:

- $p \in X$,
- pokud $q \in X$ a $r \in \delta(q, \varepsilon)$, pak také $r \in X$.

Rozšíření funkce D_ε na množiny stavů: pro $Y \subseteq Q$ položíme

$$D_\varepsilon(Y) = \bigcup_{q \in Y} D_\varepsilon(q).$$

Příklad - výpočet D_ε

Definice rozšířené přechodové funkce $\hat{\delta} : Q \times \Sigma^* \to 2^Q$

- $\hat{\delta}(q, \varepsilon) = D_\varepsilon(q)$,
- $\hat{\delta}(q, wa) = \bigcup_{p \in \hat{\delta}(q,w)} D_\varepsilon(\delta(p, a))$.

Lemma 2.47. V přechodovém grafu automatu $\mathcal{M}$ existuje cesta $p_1 \stackrel{\varepsilon}{\to} \ldots \stackrel{\varepsilon}{\to} p_m \stackrel{a}{\to} q_1 \stackrel{\varepsilon}{\to} \ldots \stackrel{\varepsilon}{\to} q_n$, kde $m, n \geq 1$, $a \in \Sigma$, právě když $q_n \in \hat{\delta}(p_1, a)$.

Jazyk přijímaný automatem $\mathcal{M}$ s ε-kroky definujeme

$$L(\mathcal{M}) = \{w \in \Sigma^* \mid \hat{\delta}(q_0, w) \cap F \neq \emptyset\}.$$

Příklad - výpočet rozšířené přechodové funkce pro automat s ε-kroky

Ekvivalence automatů s ε-kroky a NFA

Věta 2.48. Ke každému NFA $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$ s ε-kroky existuje ekvivalentní NFA (bez ε-kroků).

Důkaz. Konstrukce $\overline{\mathcal{M}} = (Q, \Sigma, \gamma, q_0, G)$ bez ε-kroků:

$$\gamma(q, a) = \hat{\delta}(q, a) \text{ pro každé } q \in Q, a \in \Sigma$$

$$G = \begin{cases} F & \text{pokud } D_\varepsilon(q_0) \cap F = \emptyset \\ F \cup \{q_0\} & \text{jinak} \end{cases}$$

Korektnost: Dokážeme, že pro libovolné $p \in Q$, $w \in \Sigma^+$ platí $\hat{\delta}(p, w) = \hat{\gamma}(p, w)$ *(indukcí vzhledem k délce slova w)*.

Algoritmus:

□

Uzávěrové vlastnosti regulárních jazyků

Věta 2.49. a 2.51. Třída regulárních jazyků je uzavřená na **sjednocení**, **průnik**, **rozdíl** a **komplement**.

Důkaz. synchronní paralelní kompozice automatů + komplement. □

Příklad.

$$\begin{aligned} L_1 &= \{a^i b^i c^j \mid 2i \geq j \geq 0\} \\ L_2 &= \{a\}^*.\{b\}^* \\ L_3 &= \{a^n b^n \mid n \geq 0\} \end{aligned}$$

$$L_1 \cap L_2 = L_3$$

Jazyk L_2 je regulární, L_3 není regulární $\implies$ L_1 **není regulární**.

Věta 2.52. Třída regulárních jazyků je uzavřená na **zřetězení**.

Důkaz.

Nechť L_1 a L_2 jsou regulární jazyky, rozpoznávané NFA $\mathcal{M}_1 = (Q_1, \Sigma_1, \delta_1, q_1, F_1)$ a $\mathcal{M}_2 = (Q_2, \Sigma_2, \delta_2, q_2, F_2)$, kde $Q_1 \cap Q_2 = \emptyset$.

Definujeme NFA s ε-kroky $\mathcal{M}_1 \odot \mathcal{M}_2 = (Q_1 \cup Q_2, \Sigma_1 \cup \Sigma_2, \delta, q_1, F_2)$, kde

$$\delta = \delta_1 \cup \delta_2 \cup \{((p, \varepsilon), \{q_2\}) \mid p \in F_1\}.$$

Korektnost: Chceme dokázat $L(\mathcal{M}_1 \odot \mathcal{M}_2) = L(\mathcal{M}_1).L(\mathcal{M}_2)$

$\supseteq$: Nechť $u \in L(\mathcal{M}_1)$, tedy $\exists r \in F_1$ splňující $r \in \widehat{\delta_1}(q_1, u)$.
Nechť $v \in L(\mathcal{M}_2)$, tedy $\widehat{\delta_2}(q_2, v) \cap F_2 \neq \emptyset$. Pak

$$\hat{\delta}(q_1, uv) \;=\; \bigcup_{p \in \hat{\delta}(q_1, u)} \hat{\delta}(p, v) \;\supseteq\; \hat{\delta}(r, v) \;\supseteq\; \hat{\delta}(q_2, v) \;=\; \widehat{\delta_2}(q_2, v).$$

Protože $\widehat{\delta_2}(q_2, v) \cap F_2 \neq \emptyset$, tak $\hat{\delta}(q_1, uv) \cap F_2 \neq \emptyset$.
Tedy $uv \in L(\mathcal{M}_1 \odot \mathcal{M}_2)$.

$\subseteq$:

$\square$

Věta 2.53. Třída regulárních jazyků je uzavřená na **iteraci**.

Důkaz.

Nechť L je regulární jazyk akceptovaný NFA $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$.

Definujeme NFA s ε-kroky $\mathcal{M}_* = (Q \cup \{p\}, \Sigma, \delta', p, \{p\})$, kde $p \notin Q$ a

$$\delta' = \delta \cup \{((p, \varepsilon), \{q_0\})\} \cup \{((q, \varepsilon), \{p\}) \mid q \in F\}. \qquad \square$$

Uzávěrové vlastnosti regulárních jazyků - Shrnutí

Regulární výrazy

Definice 2.58. Množina **regulárních výrazů** nad abecedou Σ, označovaná $RE(\Sigma)$, je definována induktivně takto:

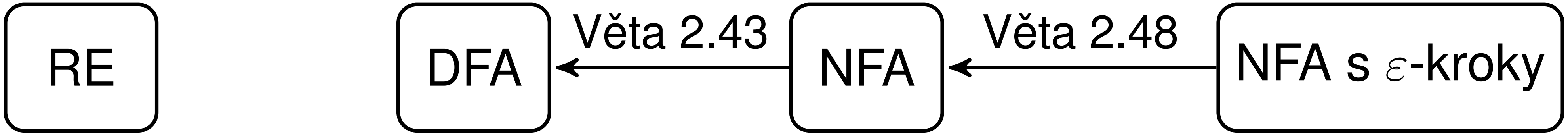
1. ε, $\emptyset$ a a pro každé $a \in \Sigma$ jsou regulární výrazy nad Σ *(tzv. základní regulární výrazy)*.
2. Jsou-li E, F regulární výrazy nad Σ, jsou také $(E.F)$, $(E + F)$ a (E^*) regulární výrazy nad Σ.
3. Každý regulární výraz vznikne po konečném počtu aplikací kroků 1-2.

Každý regulární výraz E nad abecedou Σ **popisuje** (jednoznačně určuje) **jazyk** $L(E)$ nad abecedou Σ podle těchto pravidel:

$$
\begin{aligned}
L(\varepsilon) &\stackrel{def}{=} \{\varepsilon\} \\
L(\emptyset) &\stackrel{def}{=} \emptyset \\
L(a) &\stackrel{def}{=} \{a\} \text{ pro každé } a \in \Sigma \\
L(\, E.F \,) &\stackrel{def}{=} L(E).L(F) \\
L(\, E + F \,) &\stackrel{def}{=} L(E) \cup L(F) \\
L(\, E^* \,) &\stackrel{def}{=} L(E)^*
\end{aligned}
$$

Příklady

Převod regulárních výrazů na konečné automaty



Věta 2.60. Nechť E je regulární výraz. Pak existuje konečný automat rozpoznávající $L(E)$.

Důkaz. Pro libovolnou abecedu Σ lze zkonstruovat konečné automaty, které rozpoznávají jazyky popsané základními regulárními výrazy, tj. $\emptyset$, $\{\varepsilon\}$ a $\{a\}$ pro každé $a \in \Sigma$. Tvrzení věty pak plyne z uzavřenosti jazyků rozpoznatelných konečnými automaty vůči operacím sjednocení, zřetězení a iteraci. $\square$

Příklad

Regulární přechodový graf

Definice 2.64. Regulární přechodový graf $\mathcal{M}$ je pětice $(Q, \Sigma, \delta, I, F)$, kde

- Q je neprázdná konečná množina stavů,
- Σ je vstupní abeceda,
- $\delta : Q \times Q \to RE(\Sigma)$ je parciální přechodová funkce,
- $I \subseteq Q$ je množina počátečních stavů,
- $F \subseteq Q$ je množina koncových stavů.

Příklad

Slovo $w \in \Sigma^*$ je **akceptováno** grafem $\mathcal{M}$, právě když

- existuje posloupnost stavů $q_0, \ldots, q_n$, kde $n \geq 0$, $q_0 \in I$, $q_n \in F$
- a $\delta(q_{i-1}, q_i)$ je definováno pro každé $0 < i \leq n$

taková, že

- w lze rozdělit na n částí $w = v_1 \ldots v_n$ tak, že
- $v_i \in L(\delta(q_{i-1}, q_i))$ pro každé $0 < i \leq n$.

Slovo ε je akceptováno také tehdy, je-li $I \cap F \neq \emptyset$.