

Rozšíření konečných automatů I

Nedeterministické konečné automaty

Definice 2.42. **Nedeterministický konečný automat (NFA)** je $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$, kde význam všech složek je stejný jako v definici DFA s výjimkou přechodové funkce δ . Ta je definována jako (totální) zobrazení $\delta : Q \times \Sigma \rightarrow 2^Q$.

Rozšířená přechodová funkce $\hat{\delta} : Q \times \Sigma^* \rightarrow 2^Q$:

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

Jazyk přijímaný nedeterministickým konečným automatem \mathcal{M} je

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

Příklad

Ekvivalence DFA a NFA

Věta 2.43. Pro každý NFA $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$ existuje ekvivalentní deterministický konečný automat.

Důkaz. Definujeme DFA $\mathcal{M}' = (Q', \Sigma, \delta', \{q_0\}, F')$

- $Q' = 2^Q$, tj. stavy automatu \mathcal{M}' jsou všechny podmnožiny Q .
- $\delta'(P, a) = \bigcup_{q \in P} \delta(q, a)$.
- Množina koncových stavů F' je tvořena právě těmi podmnožinami Q , které obsahují nějaký prvek množiny F .

Korektnost

- \mathcal{M}' je deterministický konečný automat.

- \mathcal{M} a \mathcal{M}' jsou ekvivalentní:

indukcí k délce $w \in \Sigma^*$ dokážeme: $\hat{\delta}(q_0, w) = \hat{\delta}'(\{q_0\}, w)$

Základní krok $|w| = 0$: Platí $\hat{\delta}(q_0, \varepsilon) = \{q_0\} = \hat{\delta}'(\{q_0\}, \varepsilon)$.

Indukční krok: Nechť $w = va$, pak

$$\begin{aligned}\hat{\delta}(q_0, va) &= \bigcup_{p \in \hat{\delta}(q_0, v)} \delta(p, a) = \delta'(\hat{\delta}(q_0, v), a) \text{ (viz definice } \delta') \\ &= \delta'(\hat{\delta}'(\{q_0\}, v), a) \text{ (indukční předpoklad)} = \hat{\delta}'(\{q_0\}, va).\end{aligned}$$

Pak $L(\mathcal{M}) = L(\mathcal{M}')$, neboť

$$w \in L(\mathcal{M}) \iff \hat{\delta}(q_0, w) \cap F \neq \emptyset \iff$$

$$\hat{\delta}'(\{q_0\}, w) \cap F \neq \emptyset \iff \hat{\delta}'(\{q_0\}, w) \in F' \iff w \in L(\mathcal{M}'). \quad \square$$

Algoritmus transformace NFA na ekvivalentní DFA

Vstup: Nedeterministický konečný automat $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$.

Výstup: Ekvivalentní DFA $\mathcal{M}' = (Q', \Sigma, \delta', \{q_0\}, F')$

bez nedosažitelných stavů a s totální přechodovou funkcí.

```
1  $Q' := \{\{q_0\}\}; \quad \delta' := \emptyset; \quad F' := \emptyset; \quad Done := \emptyset;$   
2 while  $(Q' \setminus Done) \neq \emptyset$  do  
3      $M :=$  libovolný prvek množiny  $Q' \setminus Done$   
4     if  $M \cap F \neq \emptyset$  then  $F' := F' \cup \{M\}$  fi  
5     foreach  $a \in \Sigma$  do  
6          $N := \bigcup_{p \in M} \delta(p, a)$   
7          $Q' := Q' \cup \{N\}$   
8          $\delta' := \delta' \cup \{((M, a), N)\}$  od  
9      $Done := Done \cup \{M\}$   
10 od  
11  $\mathcal{M}' := (Q', \Sigma, \delta', \{q_0\}, F')$ 
```

Důsledky determinizace konečných automatů

Věta 2.44. Pro každé $n \in \mathbb{N}$ existuje NFA o n stavech takový, že ekvivalentní DFA má i po minimalizaci 2^n stavů.

Důkaz: vynechán.

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 $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \rightarrow 2^Q$.

Rozšířená přechodová funkce

Definujeme funkci $D_\varepsilon : Q \rightarrow 2^Q$ 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 $\hat{\delta} : Q \times \Sigma^* \rightarrow 2^Q$

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

Lemma 2.47. V přechodovém grafu automatu \mathcal{M} existuje cesta $p_1 \xrightarrow{\varepsilon} \dots \xrightarrow{\varepsilon} p_m \xrightarrow{a} q_1 \xrightarrow{\varepsilon} \dots \xrightarrow{\varepsilon} q_n$, 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 je

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

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. \square

Příklad.

$$L_1 = \{a^i b^i c^j \mid 2i \geq j \geq 0\}$$

$$L_2 = \{a\}^* \cdot \{b\}^*$$

$$L_3 = \{a^n b^n \mid n \geq 0\}$$

$$L_1 \cap L_2 = L_3$$

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 $r \in \hat{\delta}_1(q_1, u)$.
Nechť $v \in L(\mathcal{M}_2)$, tedy $\hat{\delta}_2(q_2, v) \cap F_2 \neq \emptyset$. Pak

$$\hat{\delta}(q_1, uv) = \bigcup_{p \in \hat{\delta}(q_1, u)} \hat{\delta}(p, v) \supseteq \hat{\delta}(r, v) \supseteq \hat{\delta}(q_2, v) = \hat{\delta}_2(q_2, v).$$

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

\subseteq :

□

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

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

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:

$$L(\epsilon) \stackrel{def}{=} \{\epsilon\}$$

$$L(\emptyset) \stackrel{def}{=} \emptyset$$

$$L(a) \stackrel{def}{=} \{a\} \text{ pro každé } a \in \Sigma$$

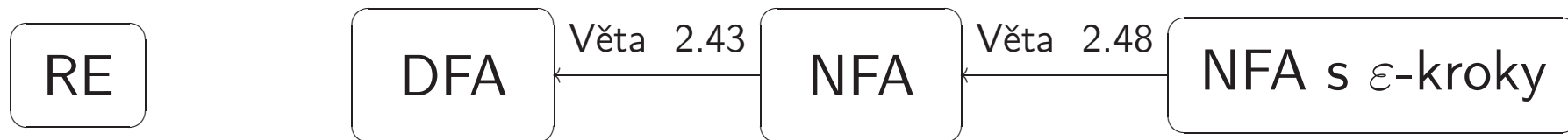
$$L(E.F) \stackrel{def}{=} L(E).L(F)$$

$$L(E + F) \stackrel{def}{=} L(E) \cup L(F)$$

$$L(E^*) \stackrel{def}{=} L(E)^*$$

Příklady

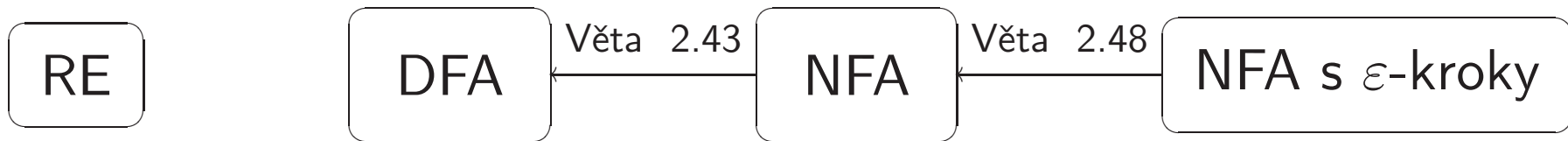
Ekvivalence konečných automatů a reg. výrazů



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. □

Příklad



Věta 2.62. Nechť L je akceptovaný nějakým DFA, pak L je popsateľný nějakým regulárním výrazem.

Důkaz bude podán později pomocí regulárních přechodových grafů.

Kleeneho věta 2.63. Libovolný jazyk je popsateľný regulárním výrazem právě když je rozpoznateľný konečným automatem.

Ekvivalentní formulace: Třída jazyků rozpoznateľných konečnými automaty je nejmenší třída, která obsahuje všechny konečné množiny a je uzavřena na sjednocení, zřetězení a iteraci.

Regulární přechodový graf

Definice 2.64.

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

- Q je neprázdňá konečňá množina stavů.
- Σ je vstupňí abeceda.
- $\delta : Q \times Q \rightarrow RE(\Sigma)$ je parciální přechodová funkce.
- $I \subseteq Q$ je množina počátečňích stavů.
- $F \subseteq Q$ je množina koncových stavů.

Příklad

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

- existuje posloupnost stavů q_0, \dots, 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 $w = v_1 \dots v_n$ 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 $I \cap F \neq \emptyset$.

Převod regulárního přechodového grafu na NFA

Motivace

Věta 2.65. Pro libovolný regulární přechodový graf $\mathcal{M} = (Q, \Sigma, \delta, I, F)$ existuje ekvivalentní NFA \mathcal{M}' s ε -kroky.

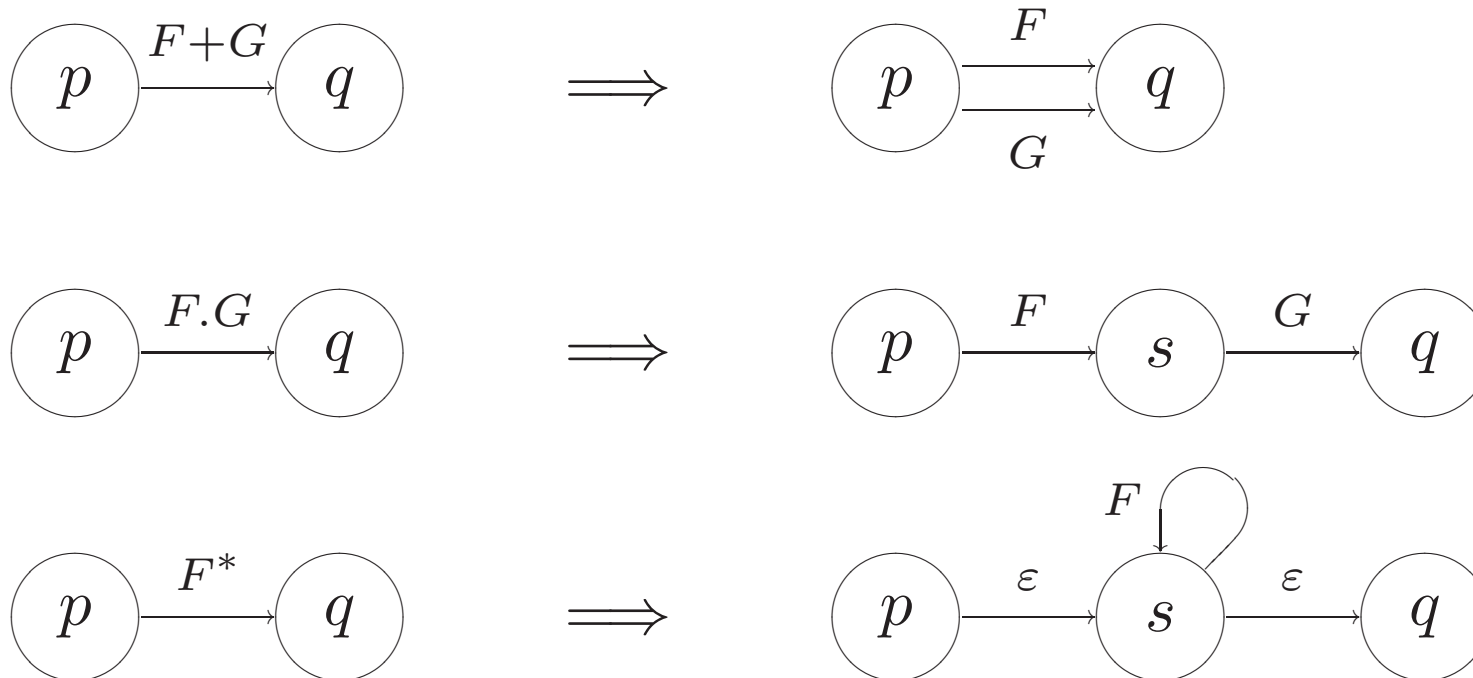
Důkaz. Algoritmus konstrukce NFA \mathcal{M}' s ε -kroky.

Krok 1 Ke grafu \mathcal{M} přidáme nový stav q_0 a hranu $q_0 \xrightarrow{\varepsilon} q$ pro každé $q \in I$. Stav q_0 bude (jediným) počátečním stavem automatu \mathcal{M}' , prvky F jeho koncovými stavy.

Krok 2 Opakovaně realizuj kroky **(a)** a **(b)** dokud přechodový graf obsahuje alespoň jednu hranu ohodnocenou symbolem, který nepatří do $\Sigma \cup \{\varepsilon\}$, tedy je tvaru $F + G$, $F.G$, F^* nebo \emptyset .

(a) Odstraň všechny hrany, které jsou ohodnoceny symbolem \emptyset .

(b) Vyber libovolnou hranu $p \xrightarrow{E} q$, kde $E \notin \Sigma \cup \{\varepsilon\}$, odstraň ji a proveď následující:



Konečnost algoritmu \mathcal{M} obsahuje pouze konečně mnoho hran, tj. konečně mnoho regulárních výrazů. Každý regulární výraz obsahuje jen konečně mnoho výskytů $+$, $.$ a $*$. V každém kroku 2(b) jeden výskyt odstraníme.

Korektnost algoritmu

- Výsledný graf je přechodovým grafem nedeterministického konečného automatu \mathcal{M}' s ε -kroky.
- Přímo z definice regulárního přechodového grafu se snadno ověří, že kroky 1 a 2 popsaného transformačního algoritmu zachovávají ekvivalenci automatů, proto $L(\mathcal{M}) = L(\mathcal{M}')$.

