Formální jazyky a automaty

Konzervativní rozšíření konečných automatů: Nedeterminismus a ε-kroky

Jan Křetínský

Fakulta informatiky, MU Brno

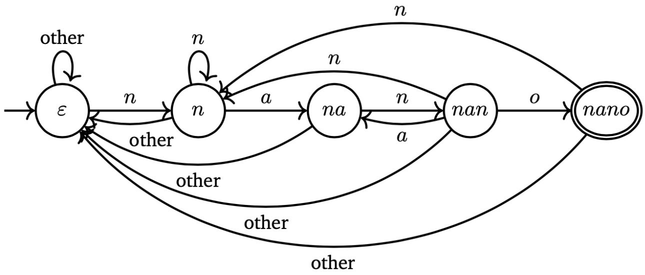
Jaro 2024

Aplikace: Hledání výskytu slova v textu

DFA pro $\Sigma^* nano$:

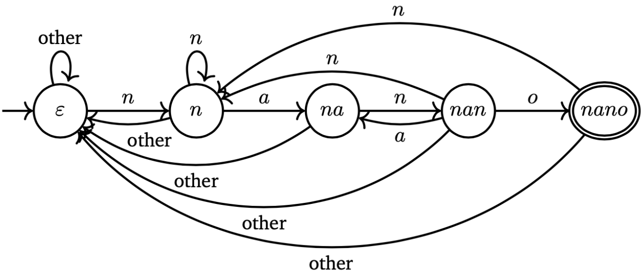
Aplikace: Hledání výskytu slova v textu

DFA pro $\Sigma^* nano$:



Aplikace: Hledání výskytu slova v textu

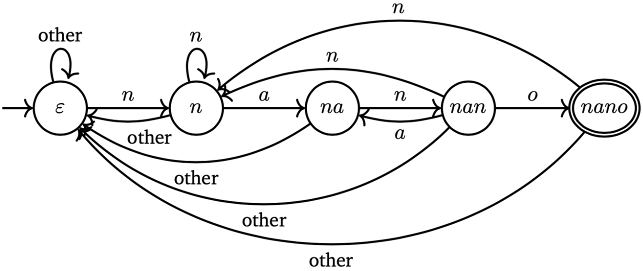
DFA pro $\Sigma^* nano$:



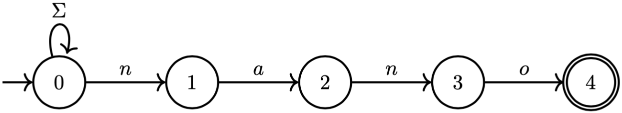
NFA:

Aplikace: Hledání výskytu slova v textu

DFA pro $\Sigma^* nano$:



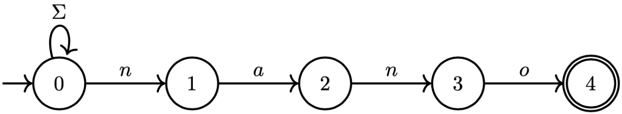
NFA:



Nedeterministický konečný automat (NFA)

Definice 2.42. (NFA)

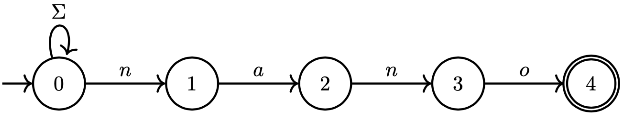
Nedeterministický konečný automat (NFA) je $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$, kde význam všech složek je stejný jako v definici DFA s výjimkou přechodové funkce δ. Ta je definována jako (totální) zobrazení $\delta : Q \times \Sigma \to 2^Q$.



Nedeterministický konečný automat (NFA)

Definice 2.42. (NFA)

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 $\delta : Q \times \Sigma \to 2^Q$.



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

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

Jazyk přijímaný nedeterministickým konečným automatem $\mathcal{M}$:

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

Ekvivalence DFA a NFA: Determinizace

Věta 2.43.

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

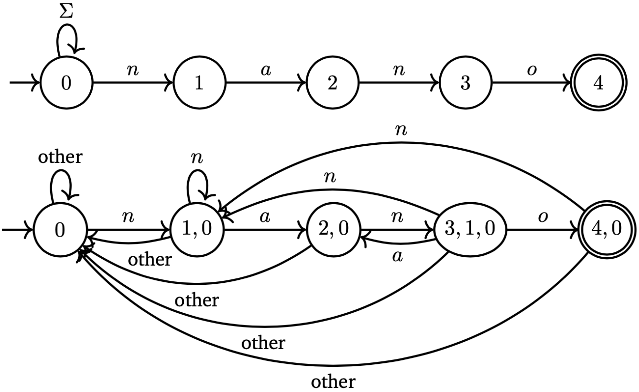
Důkaz Definujeme DFA $\mathcal{M}' = (\mathcal{Q}, \Sigma, \Delta, \{q_0\}, \mathcal{F})$

- $\mathcal{Q} = 2^Q$, tj. stavy automatu $\mathcal{M}'$ jsou všechny podmnožiny Q.
- $\Delta(P, a) = \bigcup_{q \in P} \delta(q, a)$.

- Množina koncových stavů $\mathcal{F}$ je tvořena právě těmi podmnožinami Q, které obsahují nějaký prvek množiny F.

Determinizace: Příklad

- $\mathcal{Q} = 2^Q$
- $\Delta(P, a) = \bigcup_{q \in P} \delta(q, a)$.
- $\mathcal{F} = \{P \mid P \cap F \neq \emptyset\}$



Determinizace: Korektnost

- $\mathcal{M}'$ je deterministický konečný automat.
- $\mathcal{M}$ a $\mathcal{M}'$ jsou ekvivalentní:
  indukcí k délce $w \in \Sigma^*$ dokážeme $\hat{\delta}(q_0, w) = \widehat{\Delta}(\{q_0\}, w)$

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

  Indukční krok: Nechť $w = va$, pak

$$\begin{aligned}
\hat{\delta}(q_0, va) &= \bigcup_{p \in \hat{\delta}(q_0, v)} \delta(p, a) && \text{(viz definici } \hat{}\text{ pro NFA)} \\
&= \Delta(\hat{\delta}(q_0, v), a) && \text{(viz definici } \Delta\text{)} \\
&= \Delta(\widehat{\Delta}(\{q_0\}, v), a) && \text{(indukční předpoklad)} \\
&= \widehat{\Delta}(\{q_0\}, va) && \text{(viz definici } \hat{}\text{ pro DFA)}
\end{aligned}$$

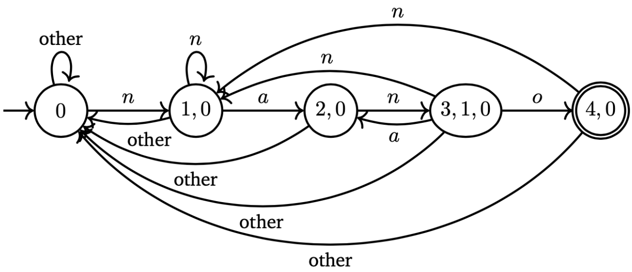
  Pak $L(\mathcal{M}) = L(\mathcal{M}')$, neboť
  $w \in L(\mathcal{M}) \iff \hat{\delta}(q_0, w) \cap F \neq \emptyset \iff \widehat{\Delta}(\{q_0\}, w) \cap F \neq \emptyset$
  $\iff \widehat{\Delta}(\{q_0\}, w) \in \mathcal{F} \iff w \in L(\mathcal{M}')$. $\square$

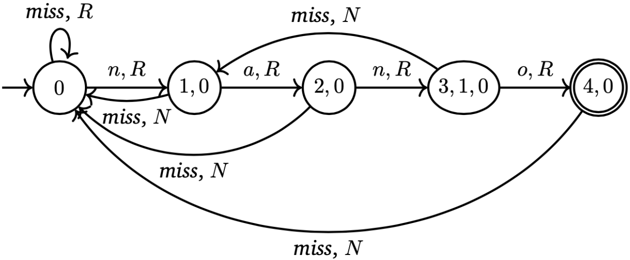
Algoritmus transformace NFA na ekvivalentní DFA

Vstup: Nedeterministický konečný automat $\mathcal{M} = (Q, \Sigma, \delta, q_0, F)$.
Výstup: Ekvivalentní DFA $\mathcal{M}' = (\mathcal{Q}, \Sigma, \Delta, \{q_0\}, \mathcal{F})$
bez nedosažitelných stavů a s totální přechodovou funkcí.

$\mathcal{Q} := \emptyset;\quad \Delta := \emptyset;\quad \mathcal{F} := \emptyset;$
$\mathit{Todo} := \{\{q_0\}\};$
while $\mathit{Todo} \neq \emptyset$ **do**
    odstraň libovolný P z Todo
    přidej P do $\mathcal{Q}$
    if $P \cap F \neq \emptyset$ **then** přidej P do $\mathcal{F}$ **end if**
    for all $a \in \Sigma$ **do**
        $P_a := \bigcup_{p \in P} \delta(p, a)$
        přidej $((P, a), P_a)$ do Δ
        if $P_a \notin \mathcal{Q}$ **then** přidej P_a do Todo
return $\mathcal{M}' := (\mathcal{Q}, \Sigma, \Delta, \{q_0\}, \mathcal{F})$

Aplikace: Knuth-Morris-Pratt



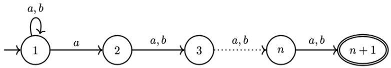


Velikosti NFA a DFA

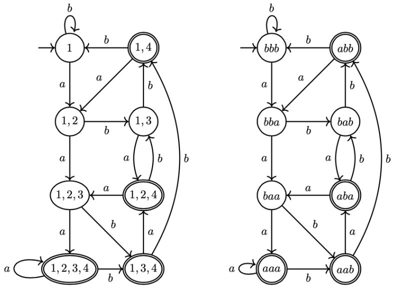
Věta 2.44.

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

Důkaz zde s jiným automatem: n-té písmeno od konce je a



(a) NFA for L_n.



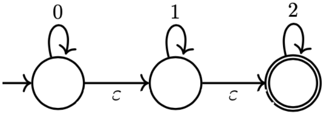
(b) DFA for L_3 and its interpretation.

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

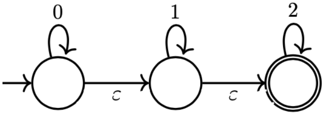
Příklad pro $0^*1^*2^*$:



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

Příklad pro $0^*1^*2^*$: 

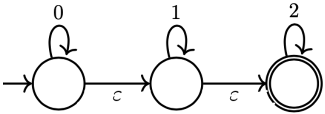
Rozšířená přechodová funkce

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

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



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 $p_1 \xrightarrow{\varepsilon} \ldots \xrightarrow{\varepsilon} p_m \xrightarrow{a} q_1 \xrightarrow{\varepsilon} \ldots \xrightarrow{\varepsilon} q_n$, 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 definujeme

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

Ekvivalence automatů s ε-kroky a NFA

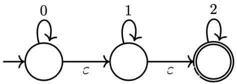
Věta 2.48. (Odstranění ε-kroků)

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

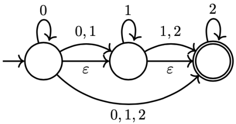
Ekvivalence automatů s ε-kroky a NFA

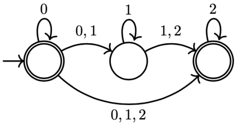
Věta 2.48. (Odstranění ε-kroků)

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



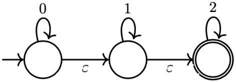
(a) NFA-ε accepting $\mathcal{L}(0^*1^*2^*)$.



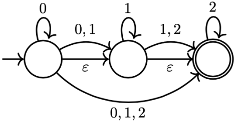


(b) After saturation.

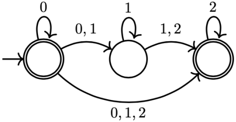
(c) After marking the initial state final and removing all ε-transitions.



(a) NFA-ε accepting $\mathcal{L}(0^*1^*2^*)$.



(b) After saturation.



(c) After marking the initial state final and removing all ε-transitions.

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).* $\square$