Automaty a gramatiky (BI-AAG)

2. Deterministické a nedeterministické konečné automaty

Jan Holub

Katedra teoretické informatiky

Fakulta informačních technologií

ČVUT v Praze

Deterministický konečný automat

Definice

Deterministický konečný automat M je pětice $M = (Q, T, \delta, q_0, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná vstupní abeceda,
- δ je zobrazení z $Q \times T$ do Q,
- $q_0 \in Q$ je počáteční stav,
- $F \subseteq Q$ je množina koncových stavů.

$\square$

Konfigurace konečného automatu

Definice

Nechť $M = (Q, T, \delta, q_0, F)$ je konečný automat. Dvojici $(q, w) \in Q \times T^*$ nazveme *konfigurací konečného automatu M*. Konfiguraci (q_0, w) nazveme *počáteční konfigurací* konečného automatu M, konfiguraci (q, ε), kde $q \in F$, nazveme *koncovou konfigurací* konečného automatu M. $\square$

Definice

Nechť $M = (Q, T, \delta, q_0, F)$ je deterministický konečný automat. Relaci $\vdash_M \in (Q \times T^*) \times (Q \times T^*)$ nazveme *přechodem* v automatu M. Jestliže $\delta(q, a) = p$, pak $(q, aw) \vdash_M (p, w)$ pro všechna $w \in T^*$. Symbolem $\vdash_M^k$ označíme k-tou mocninu relace $\vdash_M$. Symboly $\vdash_M^+$ a $\vdash_M^*$ budou označovat tranzitivní a tranzitivně reflexivní uzávěr relace $\vdash_M$. $\square$

Deterministický konečný automat

Definice

Řekneme, že řetězec $w \in T^*$ *je přijat* konečným deterministickým automatem $M = (Q, T, \delta, q_0, F)$, jestliže $(q_0, w) \vdash_M^* (q, \varepsilon)$ pro nějaké $q \in F$.
$L(M) = \{w : w \in T^*, (q_0, w) \vdash^* (q, \varepsilon), q \in F\}$ *je jazyk přijímaný konečným automatem M*. Řetězec $w \in L(M)$, jestliže existuje posloupnost přechodů taková, která z počáteční konfigurace (q_0, w) vede do koncové konfigurace (q, ε). $\square$

Konfigurace det. konečného automatu

Příklad

Mějme deterministický konečný automat
$M = (\{q_0, q_1, q_2, q_3\}, \{0, 1\}, \delta, q_0, \{q_0\})$, kde zobrazení δ je definováno takto:

$$\delta(q_0, 0) = q_2, \quad \delta(q_1, 0) = q_3, \quad \delta(q_2, 0) = q_0,$$
$$\delta(q_3, 0) = q_1, \quad \delta(q_0, 1) = q_1, \quad \delta(q_1, 1) = q_0,$$
$$\delta(q_2, 1) = q_3, \quad \delta(q_3, 1) = q_2.$$

Zobrazení δ můžeme také zapsat ve formě tabulky:

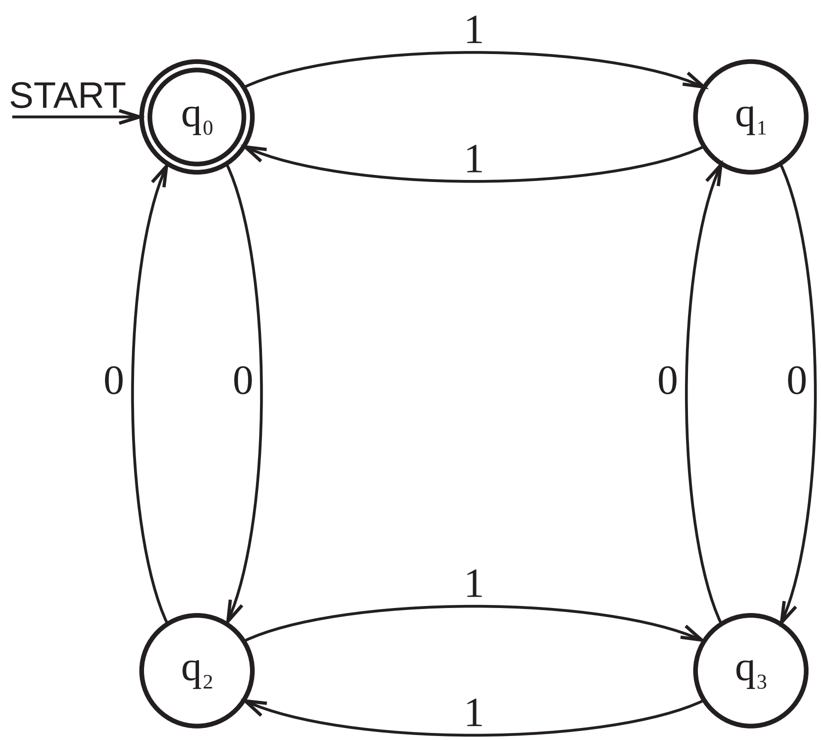
	stav	vstupní symbol	
	δ	0	1
$\rightleftarrows$	q_0	q_2	q_1
	q_1	q_3	q_0
	q_2	q_0	q_3
	q_3	q_1	q_2

Konfigurace konečného automatu

Příklad (pokračování)

$L(M) = \{x : x \in \{0,1\}^*$ a počet nul i jedniček v x je sudý$\}$.
Mějme na vstupu řetězec $x = 110101$. Automat M provede tuto posloupnost přechodů:

$(q_0, 110101) \vdash (q_1, 10101) \vdash (q_0, 0101) \vdash (q_2, 101) \vdash (q_3, 01) \vdash (q_1, 1) \vdash (q_0, \varepsilon)$.



Deterministický konečný automat

Definice
Deterministický konečný automat $M = (Q, T, \delta, q_0, F)$ nazveme úplný, když zobrazení $\delta(q, a)$ je definováno pro všechny dvojice stavů $q \in Q$ a vstupních symbolů $a \in T$. $\square$

Algoritmus Doplnění konečného automatu na úplný.
Vstup: Konečný automat $M = (Q, T, \delta, q_0, F)$.
Výstup: Úplný konečný automat $M' = (Q', T, \delta', q_0, F)$ takový, že $L(M') = L(M)$.
Metoda:

1. Vytvoříme nový stav $q_\emptyset \notin Q$, který budeme nazývat nulový. $Q' = Q \cup \{q_\emptyset\}$.
2. Jestliže $\delta(q, a)$ není definováno pro nějaké dvojice $q \in Q'$ a $a \in T$, pak definujeme $\delta'(q, a) = q_\emptyset$ pro všechny takové dvojice.
3. Pro ostatní případy $\delta'(q, a) = \delta(q, a)$. $\square$

Deterministický konečný automat

Příklad

Je dán konečný automat
$M = (\{q_0, q_1, q_2\}, \{a, b, c\}, \delta, q_0, \{q_0, q_1, q_2\})$, kde zobrazení δ je zadáno tabulkou:

	δ_M	a	b	c
$\rightleftarrows$	q_0	q_0	q_1	q_2
$\leftarrow$	q_1		q_1	q_2
$\leftarrow$	q_2			q_2

$\Longrightarrow$

	$\delta_{M'}$	a	b	c
$\rightleftarrows$	q_0	q_0	q_1	q_2
$\leftarrow$	q_1	$q_\emptyset$	q_1	q_2
$\leftarrow$	q_2	$q_\emptyset$	$q_\emptyset$	q_2
	$q_\emptyset$	$q_\emptyset$	$q_\emptyset$	$q_\emptyset$

Nedeterministický konečný automat

Definice

Nedeterministický konečný automat M je pětice $M = (Q, T, \delta, q_0, F)$, kde

- Q je konečná množina vnitřních stavů,
- T je konečná vstupní abeceda,
- δ je zobrazení z $Q \times T$ do množiny podmnožin Q (značíme 2^Q),
- $q_0 \in Q$ je počáteční stav,
- $F \subseteq Q$ je množina koncových stavů.

$\square$

Konfigurace nedet. konečného automatu

Definice

Nechť $M = (Q, T, \delta, q_0, F)$ je nedeterministický konečný automat. Relaci $\vdash_M \subseteq (Q \times T^*) \times (Q \times T^*)$ nazveme *přechodem* v automatu M. Jestliže $p \in \delta(q,a)$, pak $(q, aw) \vdash_M (p, w)$, pro libovolné $w \in T^*$. $\square$

Definice

Řekneme, že řetězec $w \in T^*$ *je přijat nedeterministickým konečným automatem* $M = (Q, T, \delta, q_0, F)$, jestliže existuje posloupnost přechodů $(q_0, w) \vdash^* (q, \varepsilon)$ pro nějaké $q \in F$. Potom $L(M) = \{w : w \in T^*, (q_0, w) \vdash^* (q, \varepsilon)$ pro nějaké $q \in F\}$ je *jazyk přijímaný nedeterministickým konečným automatem M.* $\square$

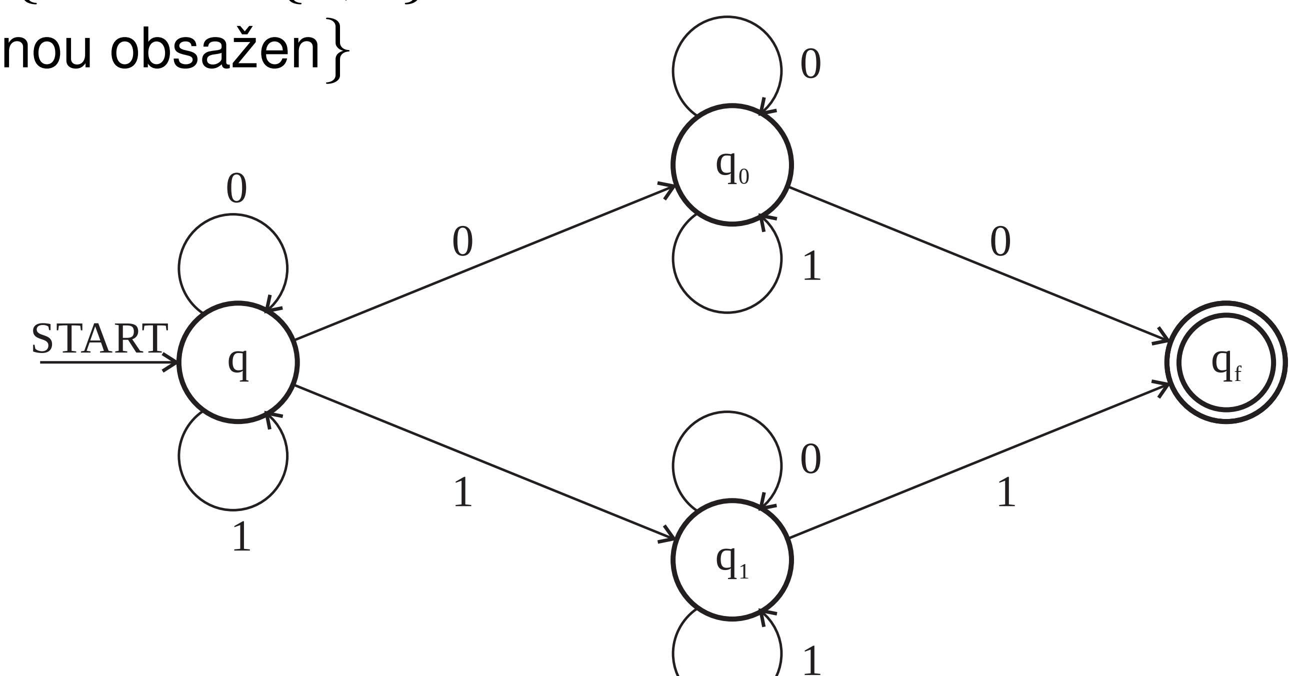
Konfigurace nedet. konečného automatu

Příklad

Mějme nedeterministický konečný automat
$M = (\{q, q_0, q_1, q_f\}, \{0, 1\}, \delta, q, \{q_f\})$, kde δ:

$$\begin{aligned} \delta(q, 0) &= \{q, q_0\}, & \delta(q, 1) &= \{q, q_1\}, \\ \delta(q_0, 0) &= \{q_0, q_f\}, & \delta(q_0, 1) &= \{q_0\}, \\ \delta(q_1, 0) &= \{q_1\}, & \delta(q_1, 1) &= \{q_1, q_f\}. \end{aligned}$$

$L(M) = \{w : w \in \{0, 1\}^*$ a w končí symbolem, který je už ve w aspoň jednou obsažen$\}$



Konfigurace nedet. konečného automatu

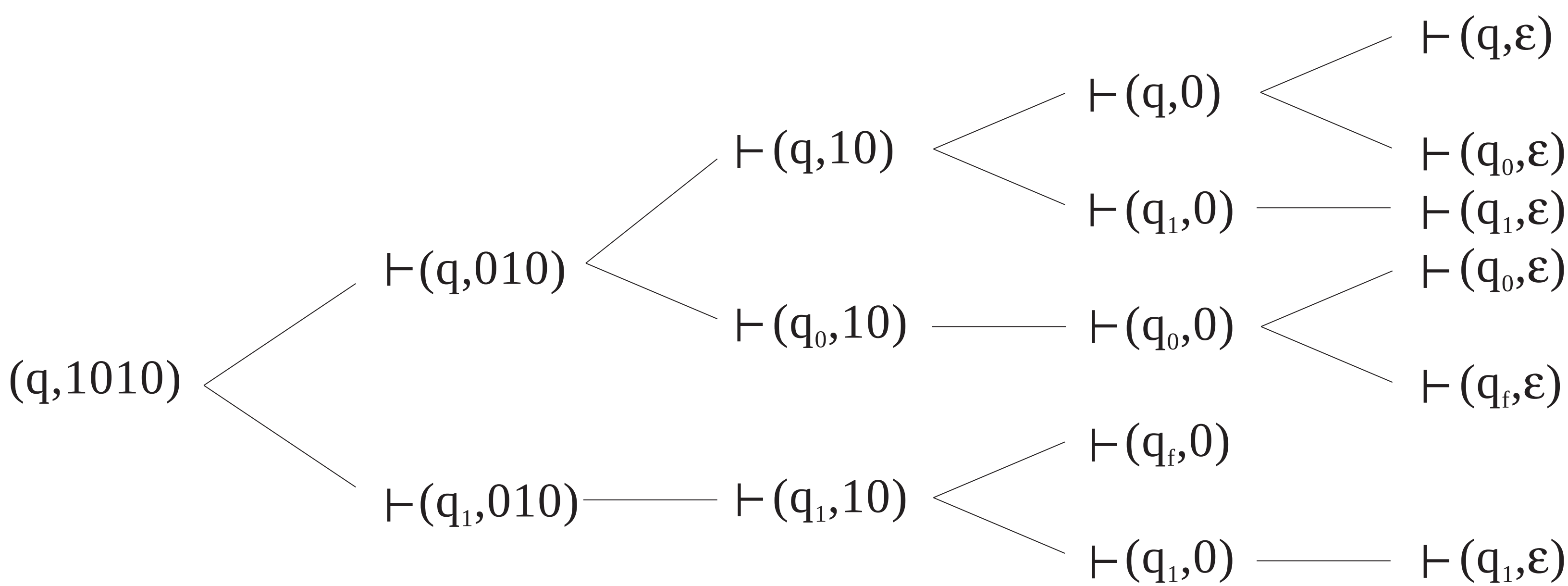
Příklad (pokračování)

Pro řetězec 1010 může automat provést mimo jiné posloupnost přechodů:

$(q, 1010) \vdash (q, 010) \vdash (q_0, 10) \vdash (q_0, 0) \vdash (q_f, \varepsilon)$.

Znázorněme si pro řetězec 1010 všechny možné posloupnosti přechodů.

$\square$



Dosažitelný stav

Definice
Nechť je dán konečný automat $M = (Q, T, \delta, q_0, F)$. Stav $q \in Q$ nazveme *dosažitelný*, pokud existuje řetězec $w \in T^*$ takový, že existuje posloupnost přechodů, která vede z počátečního stavu q_0 do stavu q:

$(q_0, w) \vdash^* (q, \varepsilon)$.

Stav, který není dosažitelný, nazveme *nedosažitelný* stav. $\square$

Dosažitelný stav

Algoritmus Nalezení a odstranění nedosažitelných stavů.
Vstup: Konečný automat $M = (Q, T, \delta, q_0, F)$.
Výstup: Konečný automat $M'(Q', T, \delta', q_0, F')$, který nemá žádné nedosažitelné stavy takový, že $L(M) = L(M')$.
Metoda:

1. Určíme množinu všech dosažitelných stavů Q_a takto:
   a) $Q_0 = \{q_0\}, i := 1$.
   b) $Q_i = \{q : q \in \delta(p, a), p \in Q_{i-1}, a \in T\} \cup Q_{i-1}$.
   c) Jestliže $Q_i \neq Q_{i-1}$, pak $i := i + 1$ a jdi na krok b), jinak $Q_a = Q_i$.
2. $M' = (Q_a, T, \delta', q_0, F \cap Q_a)$, kde δ':
   $\delta'(q, a) = \delta(q, a)$ pro všechna $q \in Q_a$. $\square$

Dosažitelný stav

Příklad

Je dán konečný automat $M = (\{p, q, r\}, \{a, b\}, \delta, p, \{r\})$, kde δ:

		a	b
$\rightarrow$	p	p, r	r
	q	r	
$\leftarrow$	r	p, r	p

Pomocí algoritmu zjistíme, že $Q_0 = \{p\}, Q_1 = \{p, r\}, Q_2 = \{p, r\}$.
Stav q je nedosažitelný.
$M' = (\{p, r\}, \{a, b\}, \delta', p, \{r\})$, kde δ':

		a	b
$\rightarrow$	p	p, r	r
$\leftarrow$	r	p, r	p

Užitečný/zbytečný stav

Definice

Nechť je dán konečný automat $M = (Q, T, \delta, q_0, F)$. Stav $q \in Q$ nazveme *užitečný*, pokud existuje řetězec $w \in T^*$ takový, že existuje posloupnost přechodů, která vede ze stavu q do nějakého koncového stavu:

$(q, w) \vdash^* (p, \varepsilon),\ p \in F$.

Stav, který není užitečný, nazveme *zbytečný* stav. $\square$

Užitečný/zbytečný stav

Algoritmus Nalezení užitečných stavů a odstranění zbytečných stavů.

Vstup: Konečný automat $M = (Q, T, \delta, q_0, F)$.

Výstup: Konečný automat $M' = (Q', T, \delta', q_0, F)$, který nemá žádné zbytečné stavy takový, že $L(M) = L(M')$.

Metoda: Určíme množinu všech užitečných stavů Q_u takto:

1.
   a) $Q_0 = F, i := 1$.

   b) $Q_i = \{q : p \in \delta(q, a), p \in Q_{i-1}\} \cup Q_{i-1}$.

   c) Jestliže $Q_i \neq Q_{i-1}$, pak $i := i + 1$ a jdi na krok b), jinak $Q_u = Q_i$.

2. $M' = (Q_u, T, \delta', q_0, F)$, kde δ' bude zkonstruováno takto:

   $\delta'(q, a) = \delta(q, a) \cap Q_u$ pro všechna $q \in Q_u$. $\square$

Užitečný/zbytečný stav

Příklad

Je dán konečný automat $M = (\{p, q, r\}, \{a, b\}, \delta, p, \{r\})$, kde zobrazení δ je zadáno tabulkou přechodů:

		a	b
$\rightarrow$	p	q, r	p, r
	q	q	q
$\leftarrow$	r	p	r

Pomocí algoritmu zjistíme, že $Q_0 = \{r\}, Q_1 = \{p, r\}, Q_2 = \{p, r\}$.
Proto $Q_u = \{p, r\}$ a je vidět, že stav q je zbytečný.
$M' = (\{p, r\}, \{a, b\}, \delta', p, \{r\})$, kde δ':

		a	b
$\rightarrow$	p	r	p, r
$\leftarrow$	r	p	r

Konečné automaty s ε-přechody

Definice

Nedeterministický konečný automat s ε-přechody je pětice $M = (Q, T, \delta, q_0, F)$, kde Q, T, q_0, F jsou stejné jako v definici nedeterministického konečného automatu. Zobrazení δ je definováno takto:

δ je zobrazení z $Q \times (T \cup \{\varepsilon\})$ do množiny podmnožin Q. $\square$

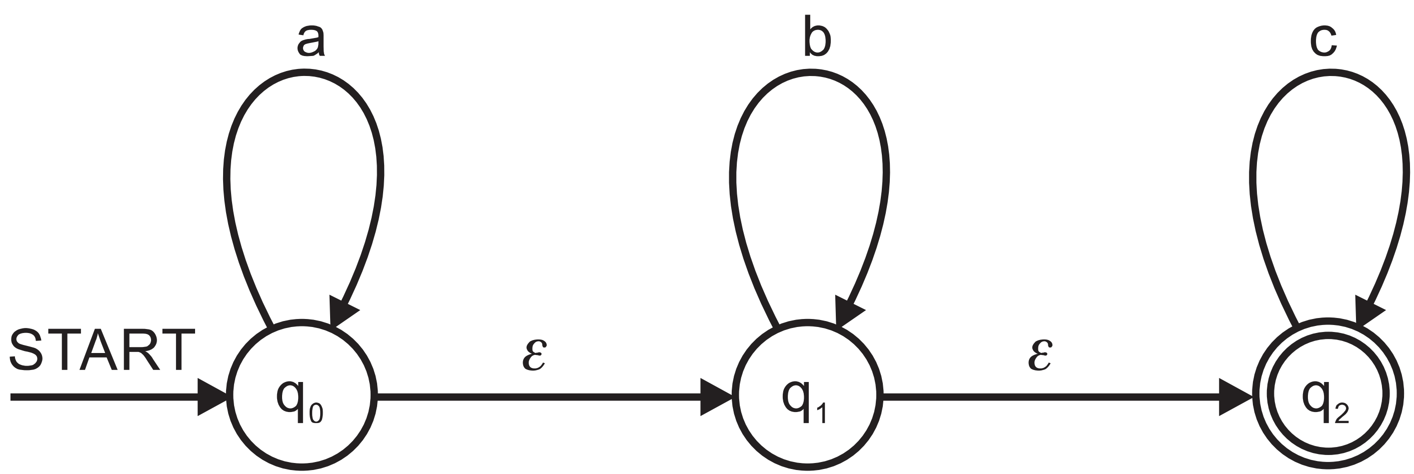
Konečné automaty s ε-přechody

Příklad

$M = (\{q_0, q_1, q_2\}, \{a, b, c\}, \delta, q_0, \{q_2\})$, kde δ:

$$\begin{aligned} &\delta(q_0, a) = \{q_0\}, \quad \delta(q_0, \varepsilon) = \{q_1\}, \\ &\delta(q_1, b) = \{q_1\}, \quad \delta(q_1, \varepsilon) = \{q_2\}, \\ &\delta(q_2, c) = \{q_2\}. \end{aligned}$$

V ostatních případech $\delta(q, x) = \emptyset$.

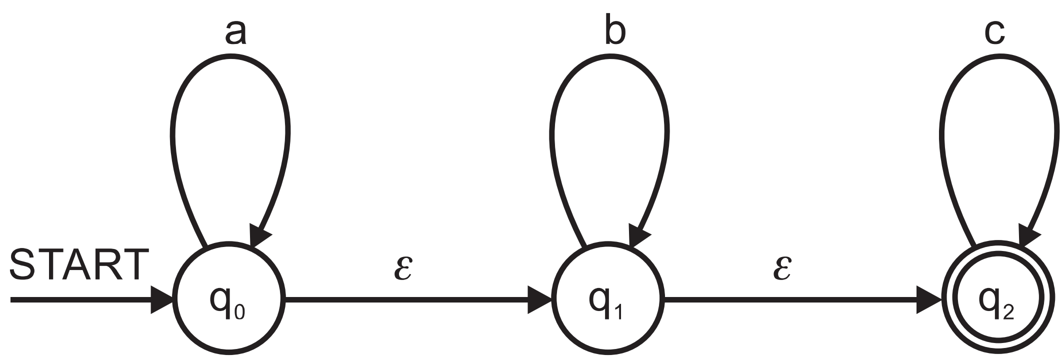


Konečné automaty s ε-přechody

Definice

Nechť $M = (Q, T, \delta, q_0, F)$ je nedeterministický konečný automat s ε–přechody. Relaci $\vdash_M \subseteq (Q \times T^*) \times (Q \times T^*)$ nazveme *přechodem* v automatu M. Jestliže $p \in \delta(q, a),\ a \in T \cup \{\varepsilon\}$, pak $(q, aw) \vdash_M (p, w)$ pro libovolné $w \in T^*$. $\square$

Příklad



Konečný automat provede pro vstupní řetězec abc tuto posloupnost přechodů:

$(q_0, abc) \vdash (q_0, bc) \vdash (q_1, bc) \vdash (q_1, c) \vdash (q_2, c) \vdash (q_2, \varepsilon)$.

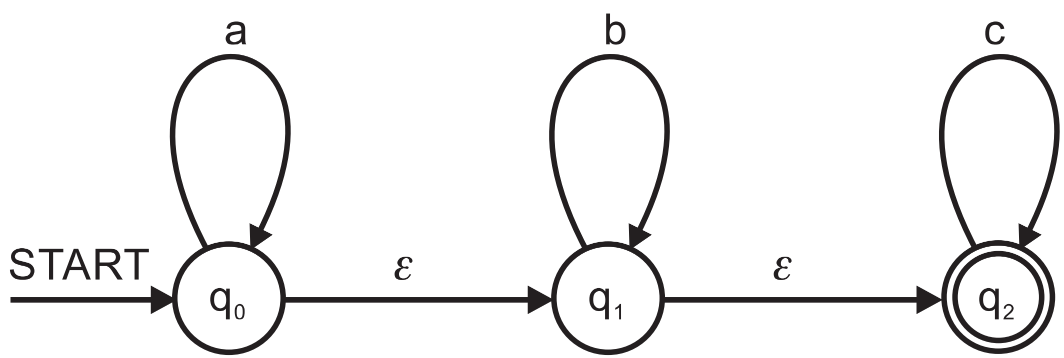
ε-*CLOSURE*

Definice

Funkce ε-*CLOSURE* pro konečný automat $M = (Q, T, \delta, q_0, F)$ je definována takto:

ε-*CLOSURE*$(q) = \{p : (q, \varepsilon) \vdash^* (p, \varepsilon), p \in Q\}$. $\square$

Příklad



$$\begin{aligned}
\varepsilon\text{-}\mathit{CLOSURE}(q_0) &= \{q_0, q_1, q_2\}, \\
\varepsilon\text{-}\mathit{CLOSURE}(q_1) &= \{q_1, q_2\}, \\
\varepsilon\text{-}\mathit{CLOSURE}(q_2) &= \{q_2\}.
\end{aligned}$$

Odstranění ε-přechodů

Algoritmus Převod nedeterministického konečného automatu s ε-přechody na nedeterministický konečný automat bez ε-přechodů.

Vstup: Konečný automat $M = (Q, T, \delta, q_0, F)$ s ε–přechody.

Výstup: Konečný automat $M' = (Q, T, \delta', q_0, F')$ bez ε-přechodů takový, že $L(M) = L(M')$.

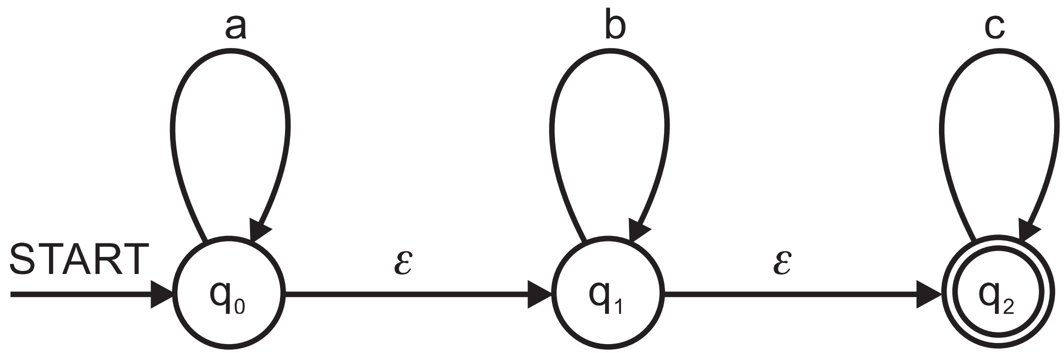
Metoda:

1.
$$\delta'(q, a) = \bigcup_{p \in \varepsilon\text{-}\mathit{CLOSURE}(q)} \delta(p, a).$$

2. $F' = \{q : \varepsilon\text{-}\mathit{CLOSURE}(q) \cap F \neq \emptyset, q \in Q\}$. $\square$

Odstranění ε-přechodů

Příklad



$M = (Q, T, \delta', q_0, F')$, kde:

$$\delta'(q_0, a) = \delta(q_0, a) \cup \delta(q_1, a) \cup \delta(q_2, a) = \{q_0\} \cup \emptyset \cup \emptyset = \{q_0\},$$
$$\delta'(q_0, b) = \delta(q_0, b) \cup \delta(q_1, b) \cup \delta(q_2, b) = \emptyset \cup \{q_1\} \cup \emptyset = \{q_1\},$$
$$\delta'(q_0, c) = \delta(q_0, c) \cup \delta(q_1, c) \cup \delta(q_2, c) = \emptyset \cup \emptyset \cup \{q_2\} = \{q_2\},$$
$$\delta'(q_1, a) = \delta(q_1, a) \cup \delta(q_2, a) = \emptyset \cup \emptyset = \emptyset,$$
$$\delta'(q_1, b) = \delta(q_1, b) \cup \delta(q_2, b) = \{q_1\} \cup \emptyset = \{q_1\},$$
$$\delta'(q_1, c) = \delta(q_1, c) \cup \delta(q_2, c) = \emptyset \cup \{q_2\} = \{q_2\},$$
$$\delta'(q_2, a) = \emptyset,\ \delta'(q_2, b) = \emptyset,\ \delta'(q_2, c) = \{q_2\}.$$

Odstranění ε-přechodů

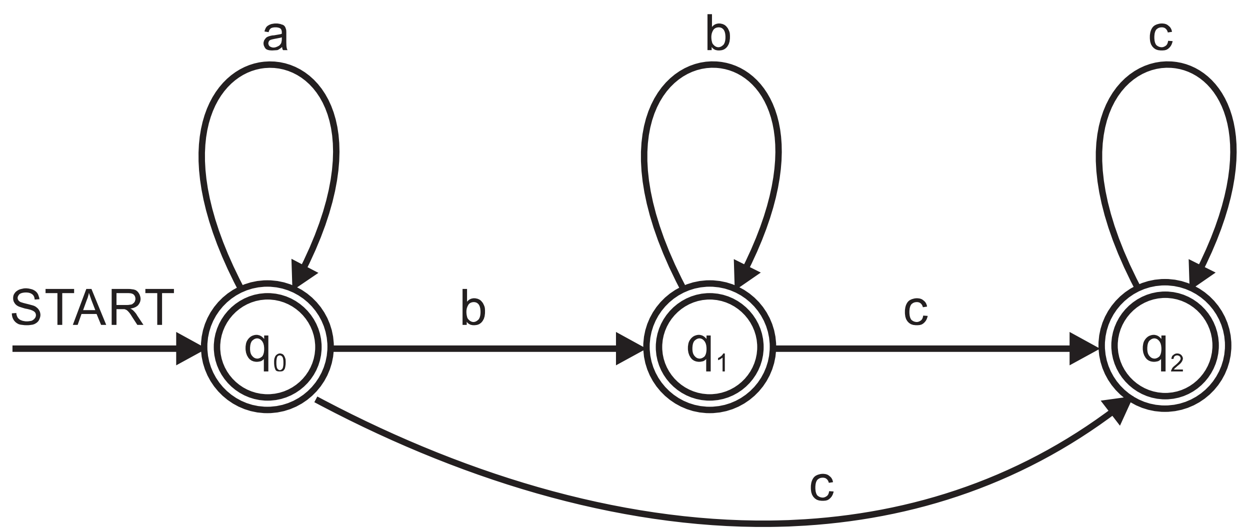
Příklad (pokračování)

$F' = \{q_0, q_1, q_2\}$, protože

ε-*CLOSURE*$(q_0) \cap F = \{q_2\}$,

ε-*CLOSURE*$(q_1) \cap F = \{q_2\}$,

ε-*CLOSURE*$(q_2) \cap F = \{q_2\}$.



Automat je deterministický.

Konečné automaty s více poč. stavy

Definice

Nedeterministický konečný automat s množinou počátečních stavů I je pětice $M = (Q, T, \delta, I, F)$ kde

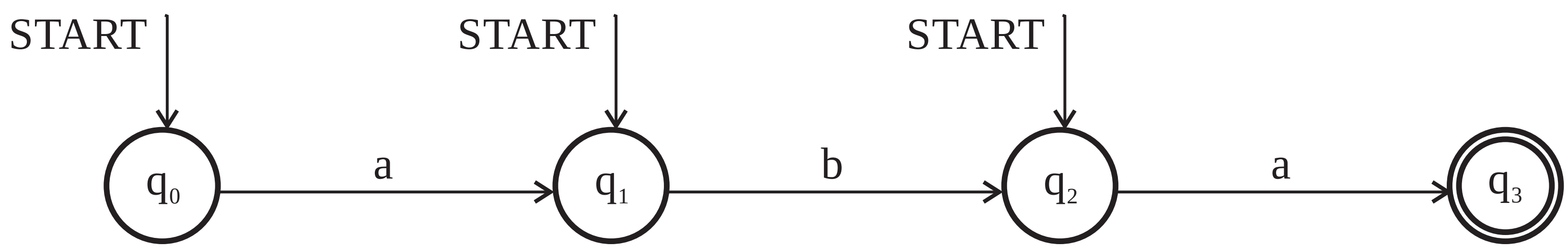
- Q, T, δ, F jsou stejné jako v definici NKA
- I je neprázdná podmnožina množiny stavů, $I \subseteq Q$.

Posloupnost přechodů tohoto konečného automatu může začít v libovolném stavu $q \in I$. $\square$

Konečné automaty s více poč. stavy

Příklad

Je dán konečný automat
$M = (\{q_0, q_1, q_2, q_3\}, \{a, b\}, \delta, \{q_0, q_1, q_2\}, \{q_3\})$, kde δ:
$\delta(q_0, a) = \{q_1\}, \qquad \delta(q_2, a) = \{q_3\},$
$\delta(q_1, b) = \{q_2\}$.
V ostatních případech je $\delta(q, x) = \emptyset$, kde $x \in \{a, b\}$.

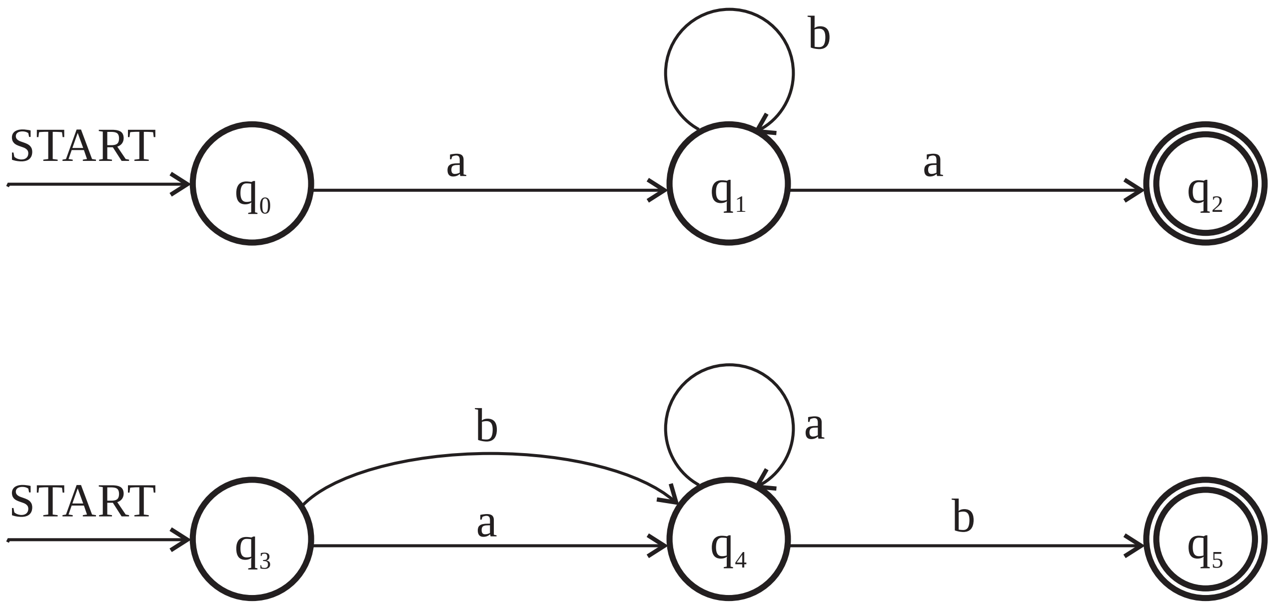


$$M: \begin{array}{r} (q_0, aba) \vdash (q_1, ba) \vdash (q_2, a) \vdash (q_3, \varepsilon), \\ (q_1, ba) \vdash (q_2, a) \vdash (q_3, \varepsilon), \\ (q_2, a) \vdash (q_3, \varepsilon). \end{array}$$

Konečné automaty s více poč. stavy

Příklad

$M = (\{q_0, q_1, q_2, q_3, q_4, q_5\}, \{a, b\}, \delta, \{q_0, q_3\}, \{q_2, q_5\})$, kde δ:



$$M: \begin{array}{l} (q_0, aba) \vdash (q_1, ba) \vdash (q_1, a) \vdash (q_2, \varepsilon), \\ (q_3, bab) \vdash (q_4, ab) \vdash (q_4, b) \vdash (q_5, \varepsilon). \end{array}$$

Převod na NKA s jedním poč. stavem

Algoritmus Převod konečného automatu s více počátečními stavy na automat s jedním počátečním stavem.

Vstup: Konečný automat $M = (Q, T, \delta, I, F),\ |I| > 1$.

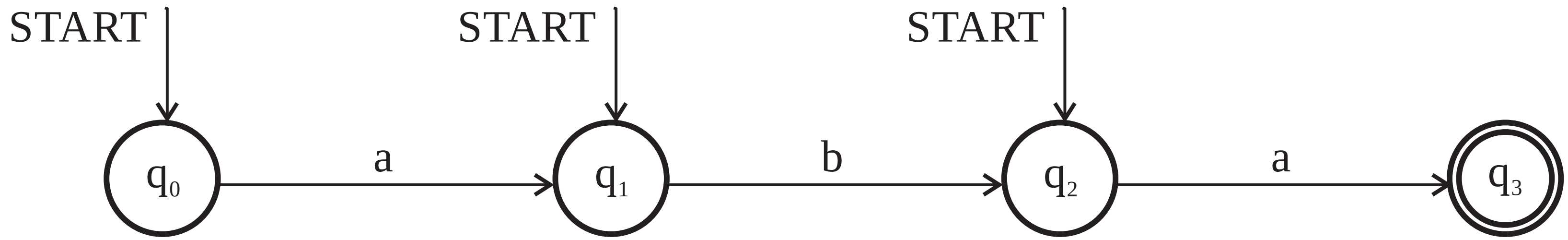
Výstup: Konečný automat $M' = (Q', T, \delta', q_0, F')$ takový, že $L(M) = L(M')$.

Metoda: M':

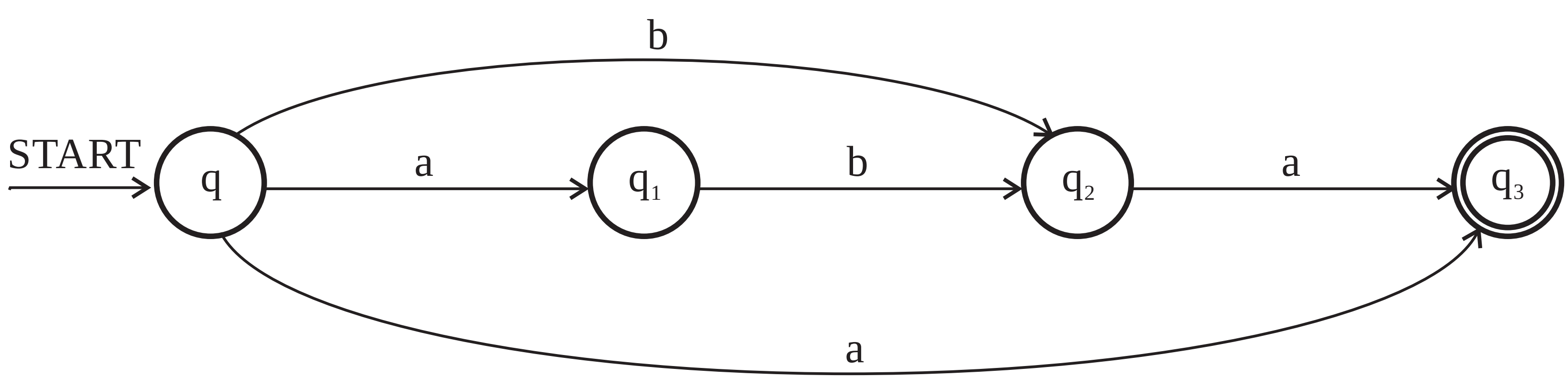
1. $q_0 = I$,
2. $\delta'(q_0, a) = \bigcup_{q \in I} \delta(q, a)$ pro všechna $a \in T$,
3. $\delta'(q, a) = \delta(q, a), \forall a \in T, \forall q \in Q$,
4. $Q' = Q \cup \{q_0\}$,
5. $F' = F$, jestliže $F \cap I = \emptyset$,
6. $F' = F \cup \{q_0\}$, jestliže $F \cap I \neq \emptyset$. $\square$

Převod na NKA s jedním poč. stavem

Příklad

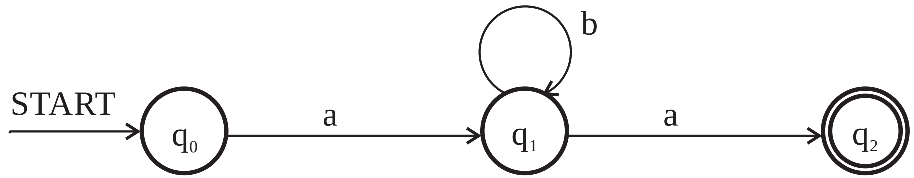


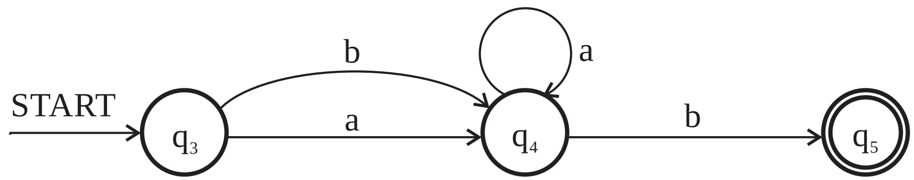
Pro konečný automat z příkladu sestrojíme ekvivalentní automat s jedním počátečním stavem:
$M' = (Q', T, \delta', q, F')$, kde $q = \{q_0, q_1, q_2\}$,
$\delta'(q, a) = \{q_1, q_3\}$, $\delta'(q, b) = \{q_2\}$.
V ostatních případech pro $p \in Q,\ x \in T$ bude $\delta'(p, x) = \delta(p, x)$.
$Q' = Q \cup \{q\},\ F' = F = \{q_3\}$.



Převod na NKA s jedním poč. stavem

Příklad





$M' = (Q', T, \delta', q, F)$, kde $q = \{q_0, q_3\}$, $\delta'(q, a) = \{q_1, q_4\}$, $\delta'(q, b) = \{q_4\}$.

V ostatních případech pro $p \in Q'$, $x \in T$ bude $\delta'(p, x) = \delta(p, x)$.

$Q' = \{q, q_1, q_2, q_4, q_5\}$.