

Kapitola V.

Vlastnosti regulárních jazyků

Pumping lemma pro RJ

Myšlenka: Pumping lemma ukazuje nekonečné iterace některých podřetězců v řetězcích v RJ.

- Necht' L je RJ. Pak existuje $k \geq 1$ takové, že: **pokud** $z \in L$ a $|z| \geq k$, **pak** existuje $u, v, w: z = uvw$,
1) $v \neq \varepsilon$ 2) $|uv| \leq k$ 3) pro každé $m \geq 0$, $uv^m w \in L$

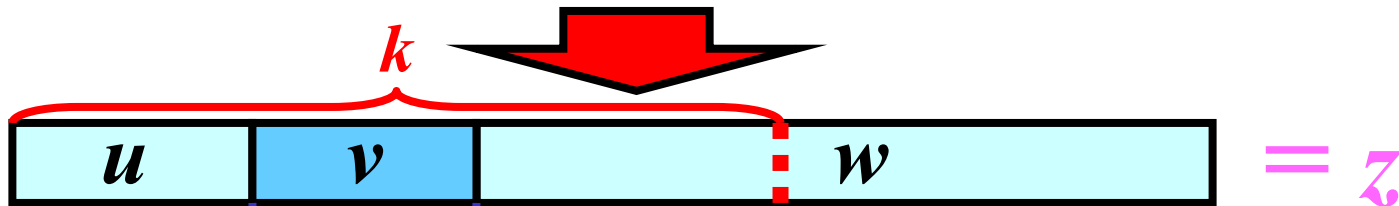
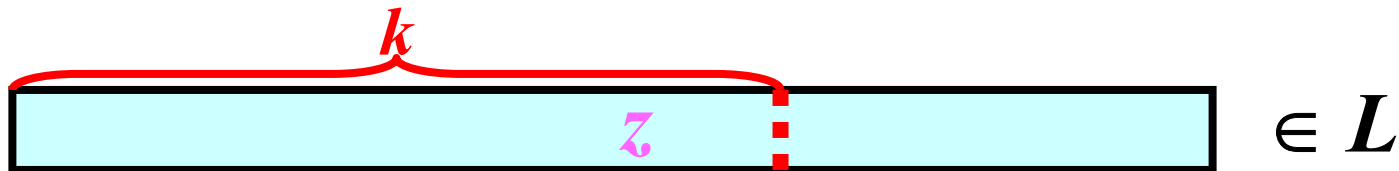
Příklad: pro RV $r = ab^*c$, $L(r)$ je *regulární*.

Pro tento jazyk existuje $k = 3$ takové, že 1), 2) a 3) platí.

- pro $z = \begin{matrix} \text{a} & \text{b} & \text{c} \\ \downarrow & \downarrow & \downarrow \\ u & v & w \end{matrix}$: $z \in L(r)$ a $|z| \geq 3$:
 $uv^0w = ab^0c = ac \in L(r)$
 $uv^1w = ab^1c = abc \in L(r)$
 $uv^2w = ab^2c = abbc \in L(r)$
 \vdots
 $v \neq \varepsilon, |uv| = 2 \leq 3$
- pro $z = \begin{matrix} \text{a} & \text{b} & \text{b} & \text{c} \\ \downarrow & \downarrow & \downarrow & \\ u & v & w & \end{matrix}$: $z \in L(r)$ a $|z| \geq 3$:
 $uv^0w = abb^0c = abc \in L(r)$
 $uv^1w = abb^1c = abbc \in L(r)$
 $uv^2w = abb^2c = abbbc \in L(r)$
 \vdots
 \vdots
 $v \neq \varepsilon, |uv| = 2 \leq 3$

Pumping lemma: Ilustrace

- $L =$ libovolný regulární jazyk:



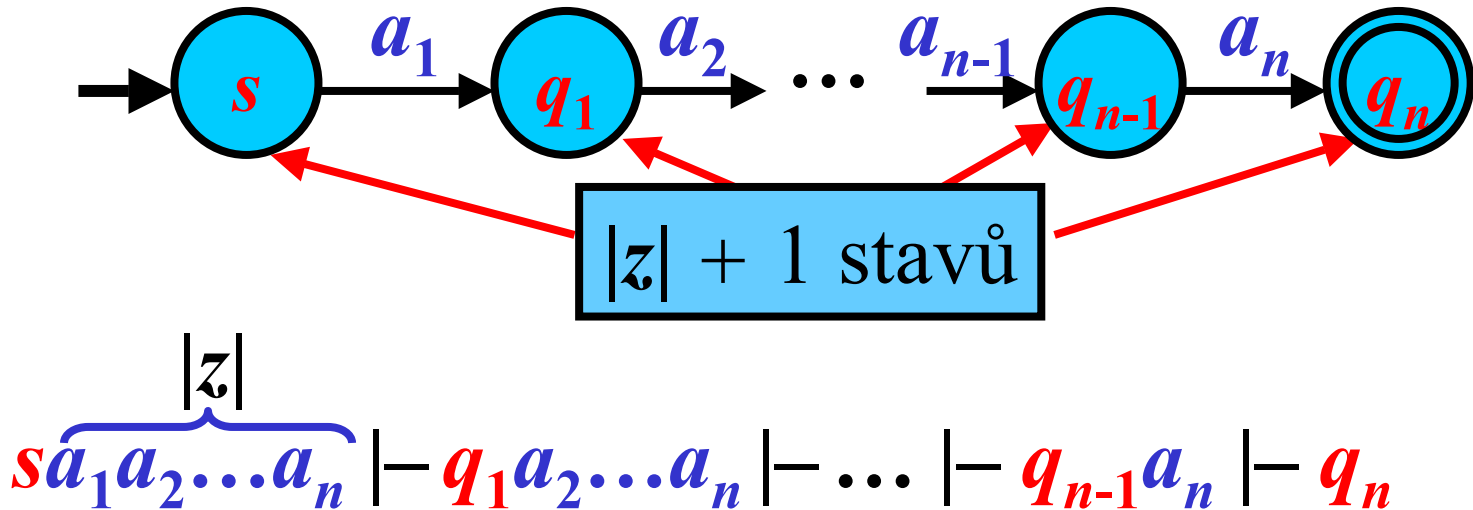
- 2) $|v| \leq k$



...

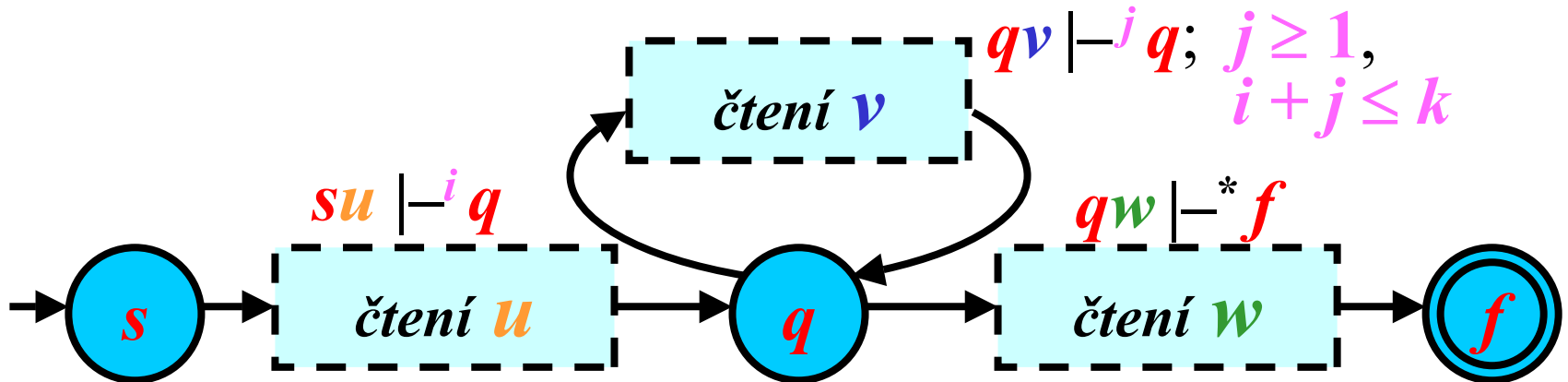
Důkaz pumping lemma 1/3

- Necht' L je libovolný regulární jazyk. Potom existuje **DKA** $M = (Q, \Sigma, R, s, F)$ a $L = L(M)$.
- Pro $z \in L(M)$, M provede $|z|$ přechodů a M navštíví $|z| + 1$ stavů:
- pro $z = a_1 a_2 \dots a_n$:



Důkaz pumping lemma 2/3

- Necht' $k = \text{card}(Q)$ (celkový počet stavů v M).
Pro každé $z \in L$ a $|z| \geq k$, M navštíví nejméně $k + 1$ stavů. Protože $k + 1 > \text{card}(Q)$, musí existovat stav q , který M navštíví nejméně dvakrát.
- Pro z existuje u, v, w takové, že: $z = uvw$:



Celkově:

$$sz = suvw \mid -^i qvw \mid -^j qw \mid -^* f, f \in F$$

Důkaz pumping lemma 3/3

- Obecně tedy M může provést přechody:
 - ① $su \xrightarrow{i} q$; ② $qv \xrightarrow{j} q$; ③ $qw \xrightarrow{*} f, f \in F$, tedy:
- pro $m = 0$, $uv^m w = uv^0 w = uw$,

$$suw \xrightarrow{\text{①}} qw \xrightarrow{\text{③}} f, f \in F$$

- pro každé $m > 0$,

$$suv^m w \xrightarrow{\text{①}} qv^m w \xrightarrow{\text{②}} qv^{m-1} w \xrightarrow{\text{②}} \dots \xrightarrow{\text{②}} qw \xrightarrow{\text{③}} f, f \in F$$

Celkově:

- 1) $qv \xrightarrow{j} q, j \geq 1$; proto $|v| \geq 1$, tedy $v \neq \varepsilon$
- 2) $su \xrightarrow{i} qv \xrightarrow{j} q, i + j \leq k$; proto $|uv| \leq k$
- 3) Pro každé $m \geq 0$: $suv^m w \xrightarrow{*} f, f \in F$, proto $uv^m w \in L$

CBD

Pumping lemma: Aplikace I.

- Pomocí pumping lemma pro RJ často provádíme důkaz sporem, že daný jazyk **není** regulární:

Předpokládejme, že L je regulární

Uvažujme PL konstantu k a vyberme $z \in L$, jehož délka je závislá na k tak, že $|z| \geq k$ je vždy pravdivé

Pro všechny dekompozice z na uvw , $v \neq \varepsilon$, $|uv| \leq k$ ukážeme:
 existuje $m \geq 0$, pro které $uv^m w \notin L$;
 ale podle PL platí vztah: $uv^m w \in L$ } **SPOR**

špatný předpoklad

Proto
 L není regulární

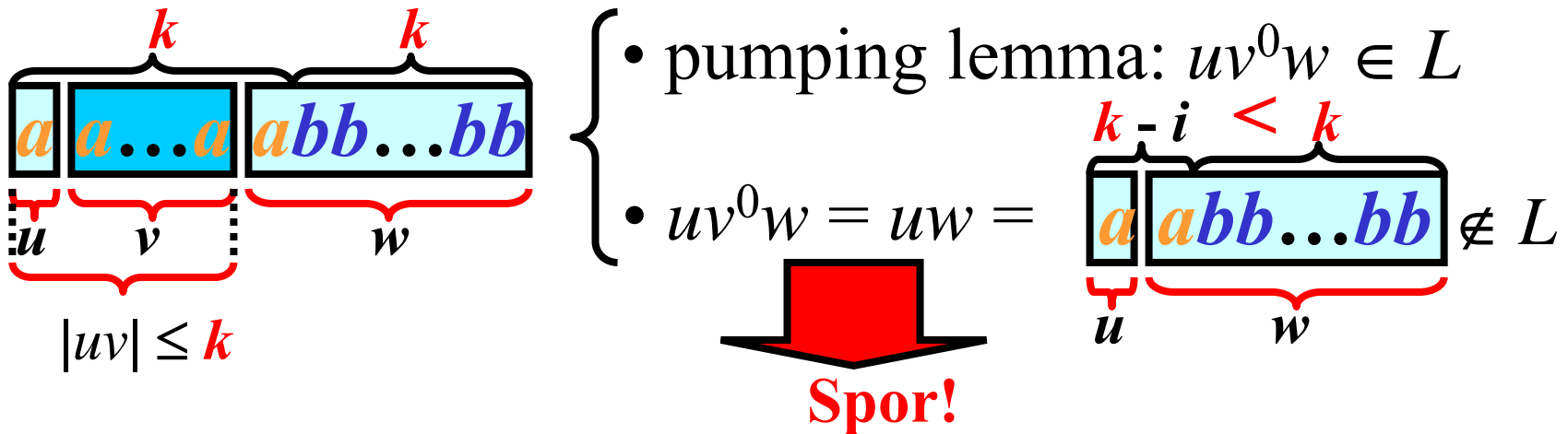
Pumping Lemma: Příklad

Dokažme, že $L = \{a^n b^n : n \geq 0\}$ není regulární:

1) Předpokládejme, že L je regulární. Necht' $k \geq 1$ je konstanta z pumping lemma pro jazyk L .

2) Necht' $z = a^k b^k : a^k b^k \in L, |z| = |a^k b^k| = 2k \geq k$

3) Všechny dekompozice z na uvw , $v \neq \varepsilon, |uv| \leq k$:



4) Proto L není regulární jazyk

Poznámka k použití pumping lemma

• Pumping lemma:

pokud L je regulární $\xrightarrow{\text{potom}}$ existuje $k \geq 0$ a ...

Základní aplikace pumping lemma:

- důkaz sporem, že L není regulární jazyk.

• Ale následující implikace je **špatná**:

~~pokud existuje $k \geq 0$ a ... $\xrightarrow{\text{potom}}$ L je regulární~~

- **Nelze** použít pumping lemma k dokázání, že daný jazyk L je regulární!!

Pumping lemma: Aplikace II. 1/3

- Pumping lemma je možné použít k dokazování dalších tvrzení.

Ilustrace:

- Necht' M je DKA a k konstanta z pumping lemma (k je počet stavů v M). Potom platí:
 $L(M)$ je nekonečný \Leftrightarrow existuje $z \in L(M)$, $k \leq |z| < 2k$

Důkaz:

1) existuje $z \in L(M)$, $k \leq |z| < 2k \Rightarrow L(M)$ je nekonečný:

pokud $z \in L(M)$, $k \leq |z|$, potom podle PL:

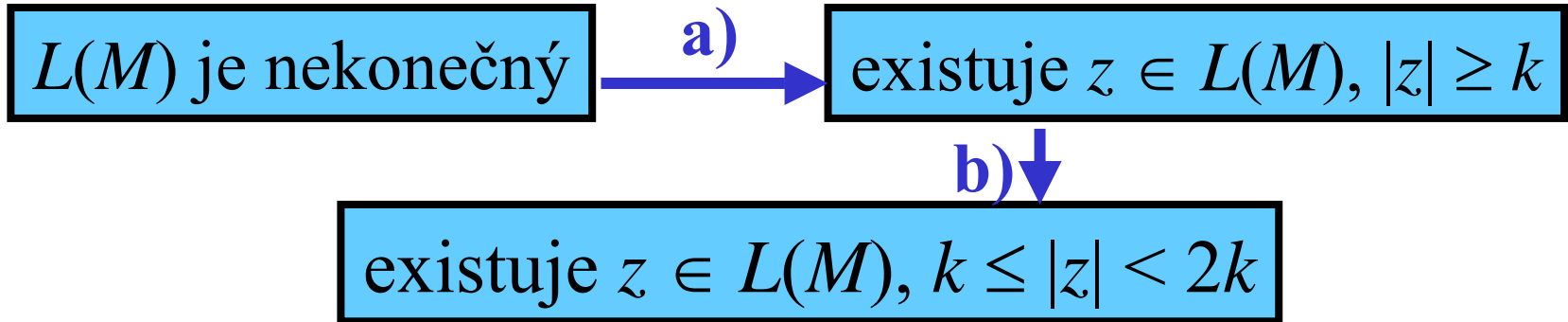
$z = uvw$, $v \neq \varepsilon$ a dále pro každé $m \geq 0$: $uv^m w \in L(M)$

$L(M)$ je nekonečný

Pumping Lemma: Aplikace II. 2/3

2) $L(M)$ je nekonečný \Rightarrow existuje $z \in L(M)$, $k \leq |z| < 2k$:

- Dokážeme sporem, že platí:



- a) Dokážeme sporem, že:

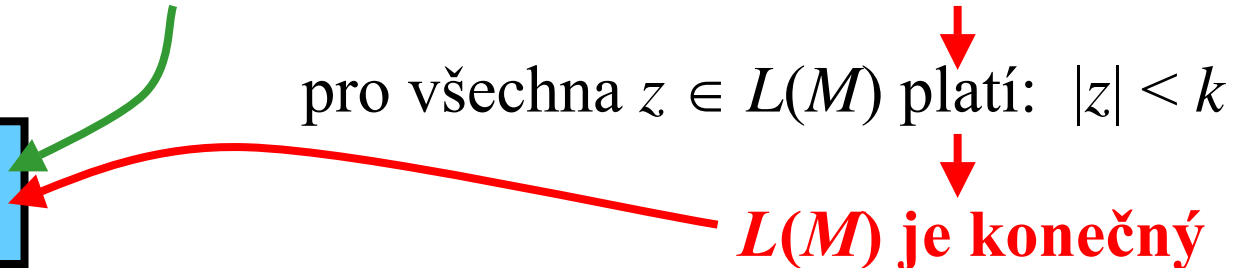
- $L(M)$ je nekonečný \Rightarrow existuje $z \in L(M)$, $|z| \geq k$

Předpokládáme, že $L(M)$ je nekonečný a neexistuje $z \in L(M)$, $|z| \geq k$

pro všechna $z \in L(M)$ platí: $|z| < k$

Spor!

$L(M)$ je konečný

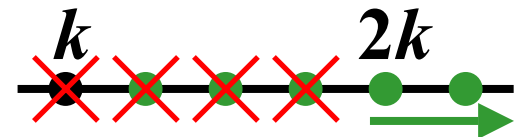


Pumping Lemma: Aplikace II. 3/3

b) Dokážeme sporem:

- existuje $z \in L(M)$, $|z| \geq k \Rightarrow$
existuje $z \in L(M)$, $k \leq |z| < 2k$

Předpokl., že existuje $z \in L(M)$, $|z| \geq k$
a neexistuje $z \in L(M)$, $k \leq |z| < 2k$



Nechť z_0 je **nejkratší řetězec** splňující $z_0 \in L(M)$, $|z_0| \geq k$

Protože neexistuje $z \in L(M)$, $k \leq |z| < 2k$, musí: $|z_0| \geq 2k$

Pokud $z_0 \in L(M)$ a $|z_0| \geq k$, PL zaručuje: $z_0 = uvw$,
 $|uv| \leq k$ a pro každé $m \geq 0$, $uv^m w \in L(M)$

$|uw| = \underbrace{|z_0|}_{\geq 2k} - \underbrace{|v|}_{\leq k} \geq k$ pro $m = 0$: $uv^m w = uw \in L(M)$

Celkově: $uw \in L(M)$, $|uw| \geq k$ a $|uw| < |z_0|$!

z_0 **není nejkratší řetězec** splňující $z_0 \in L(M)$, $|z_0| \geq k$

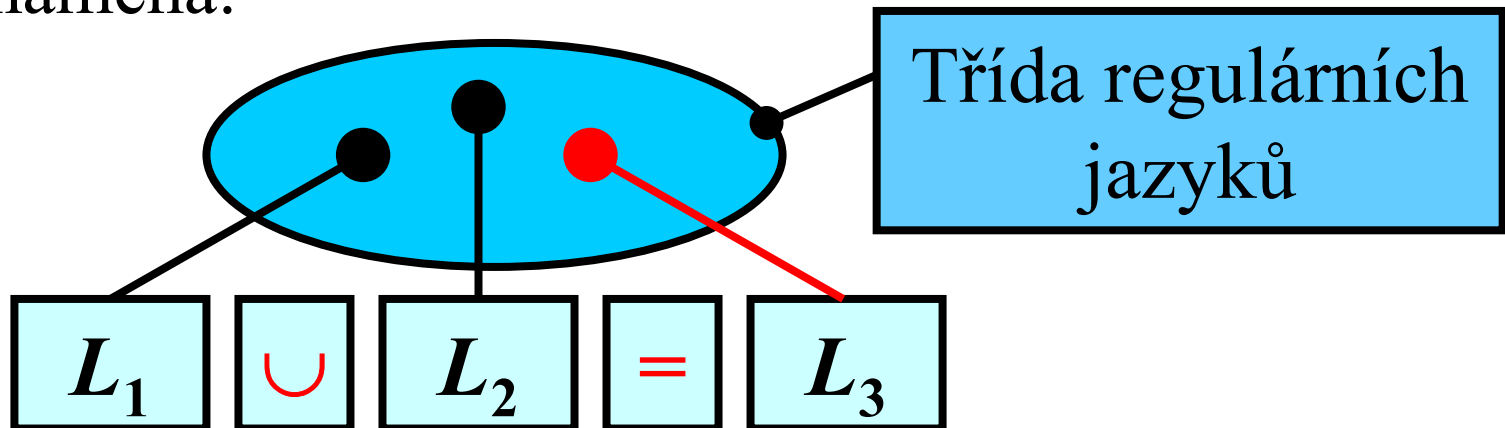
SPOR!

Uzávěrové vlastnosti 1/2

Definice: Třída regulárních jazyků je uzavřená vůči operaci o , pokud výsledek operace o na libovolné regulární jazyky je opět regulární jazyk.

Ilustrace:

- Třída regulárních jazyků je uzavřená vůči *sjednocení*.
To znamená:



Uzávěrové vlastnosti 2/2

Tvrzení: Třída regulárních jazyků je uzavřena vůči: **sjednocení, konkatenaci, iteraci.**

Důkaz:

- Necht' L_1, L_2 jsou dva **regulární jazyky**
- Potom, existují dva RV r_1, r_2 : $L(r_1) = L_1, L(r_2) = L_2$;
- Podle definice regulárních výrazů:
 - $r_1.r_2$ je RV značící $L_1 L_2$
 - $r_1 + r_2$ je RV značící $L_1 \cup L_2$
 - r_1^* je RV značící L_1^*
- Každý RV značí regulární jazyk, tedy $L_1 L_2, L_1 \cup L_2, L_1^*$ jsou **regulární jazyky**

Algoritmus: KA pro doplněk

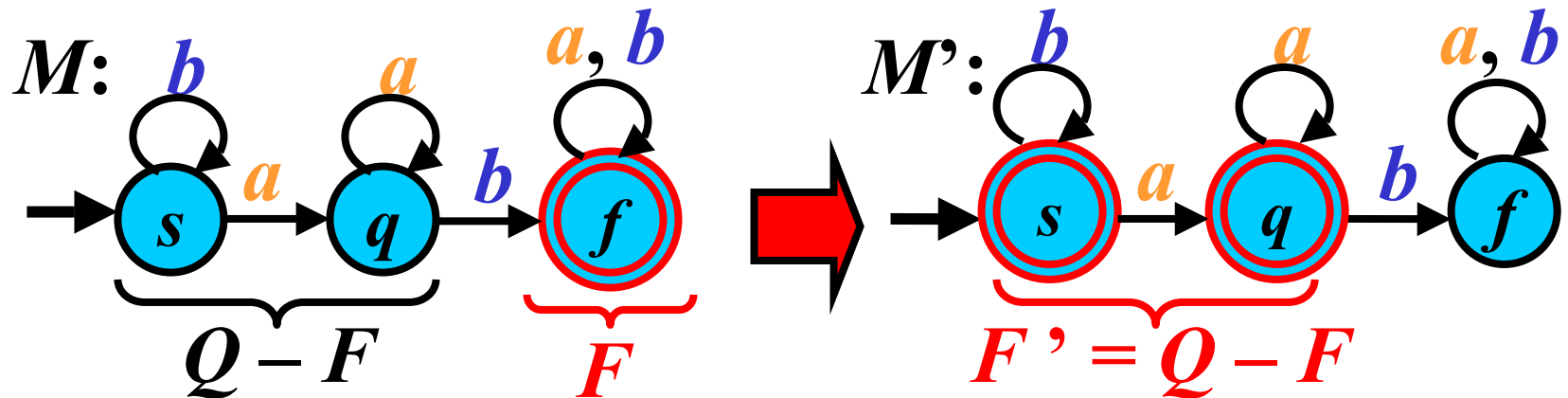
- **Vstup:** Úplný KA: $M = (Q, \Sigma, R, s, F)$
- **Výstup:** Úplný KA: $M' = (Q, \Sigma, R, s, F')$,

$$L(M') = \overline{L(M)}$$

• Metoda:

- $F' := Q - F$

Příklad:



$L(M) = \{x: ab \text{ je podřetězec } x\}$; $L(M') = \{x: ab \text{ není podřetězec } x\}$

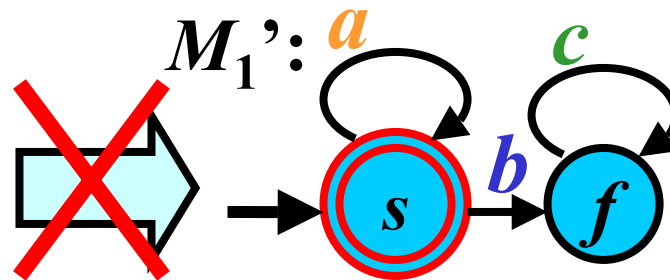
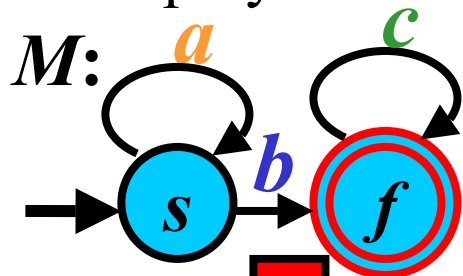
KA pro doplněk: Problém

- Předchozí algoritmus vyžaduje **úplný** KA
- Pokud M není úplný KA, potom M musí být převeden na úplný KA a pak může být použit předchozí algoritmus

Příklad:

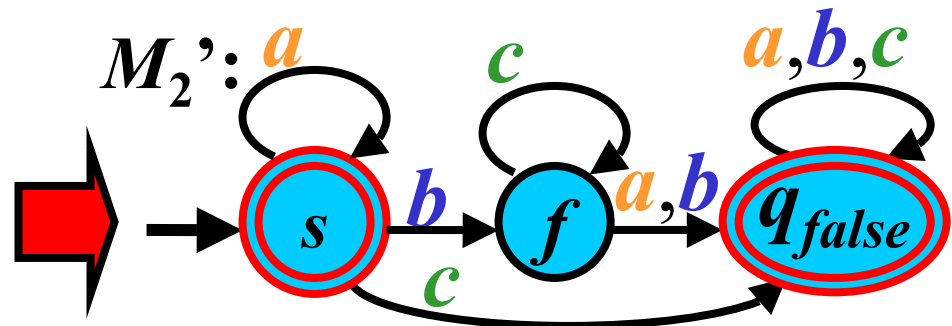
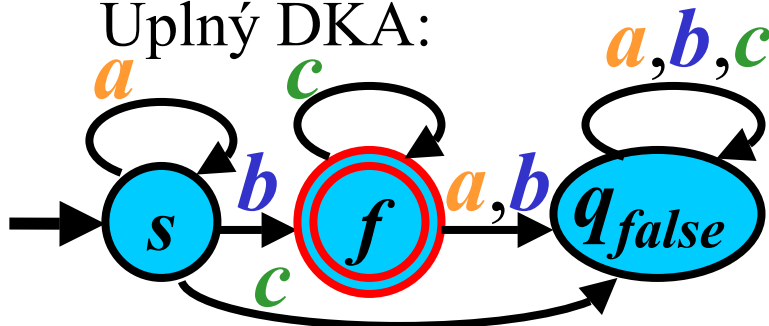
Neúplný DKA:

$$L(M_1') \neq \overline{L(M)}! - c \notin L(M), c \notin L(M_1')$$



$$L(M_2') = \overline{L(M)}$$

Úplný DKA:



Uzávěrové vlastnosti: Doplněk

Tvrzení: Třída regulárních jazyků je uzavřena vůči **doplňku**.

Důkaz:

- Necht' L je **regulární jazyk**
- Pak existuje úplný DKA $M: L(M) = L$
- Můžeme sestrojít úplný DKA $M': L(M') = \overline{L}$ užitím předchozího algoritmu
- Každý KA definuje regulární jazyk, tedy \overline{L} je **regulární jazyk**

Uzávěrové vlastnosti: Průnik

Tvrzení: Třída regulárních jazyků je uzavřena vůči **průniku**.

Důkaz:

- Necht' L_1, L_2 jsou dva **regulární jazyky**
 - $\overline{L_1}, \overline{L_2}$ jsou **regulární jazyky**
(třída regulárních jazyků je uzavřena vůči doplňku)
 - $\overline{L_1} \cup \overline{L_2}$ je **regulární jazyk**
(třída regulárních jazyků je uzavřena vůči sjednocení)
 - $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$ je **regulární jazyk**
(třída regulárních jazyků je uzavřena vůči doplňku)
- (De-Morganovy zákony)

Boolova algebra jazyků

Definice: Necht' je třída jazyků uzavřena vůči sjednocení, průniku a doplňku. Potom tato třída tvoří *Booleovu algebru jazyků*.

Tvrzení: Třída regulárních jazyků tvoří Booleovu algebru jazyků.

Důkaz:

- Třída regulárních jazyků je uzavřená vůči sjednocení, průniku a doplňku.

Hlavní rozhodnutelné problémy

1. Problém členství:

- **Instance:** FA M , $w \in \Sigma^*$; **Otázka:** $w \in L(M)$?

2. Problém prázdnoty:

- **Instance:** FA M ; **Otázka:** $L(M) = \emptyset$?

3. Problém konečnosti:

- **Instance:** FA M ; **Otázka:** Je $L(M)$ konečný?

4. Problém ekvivalence:

- **Instance:** FA M_1, M_2 ; **Otázka:** $L(M_1) = L(M_2)$?

Algoritmus: Problém členství

- **Vstup:** DKA $M = (Q, \Sigma, R, s, F)$; $w \in \Sigma^*$
 - **Výstup:** **ANO**, pokud $w \in L(M)$
NE, pokud $w \notin L(M)$
-

- **Metoda:**
 - **if** $sw \stackrel{*}{\vdash} f$, $f \in F$ **then** napiš('ANO')
else napiš('NE')
-

Celkově:

Problém členství je pro KA rozhodnutelný

Algoritmus: Problém prázdnoti

- **Vstup:** KA $M = (Q, \Sigma, R, s, F)$;
 - **Výstup:** **ANO**, pokud $L(M) = \emptyset$
NE, pokud $L(M) \neq \emptyset$
-

- **Metoda:**
 - **if** s je neukončující **then** napiš('ANO')
else napiš('NE')
-

Celkově:

Problém prázdnoti je pro KA rozhodnutelný

Algoritmus: Problém konečnosti

- **Vstup:** DKA $M = (Q, \Sigma, R, s, F)$;
 - **Výstup:** **ANO**, pokud $L(M)$ je konečný
NE, pokud $L(M)$ je nekonečný
-
- **Metoda:**
 - Necht' $k = \text{card}(Q)$
 - if existuje $z \in L(M)$, $k \leq |z| < 2k$ then napiš('NE')
else napiš('ANO')

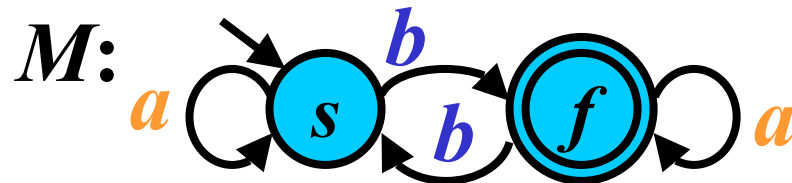
Pozn.: Tento algoritmus je založen na tvrzení:

$L(M)$ je nekonečný \Leftrightarrow existuje $z: z \in L(M)$, $k \leq |z| < 2k$

Celkově:

Problém konečnosti je pro KA rozhodnutelný

Rozhodnutelné problémy: Příklad



Otázka: $ab \in L(M)$?

$sab \vdash sb \vdash f, f \in F$

Odpověď: **ANO**, protože $sab \vdash^* f, f \in F$

Otázka: $L(M) = \emptyset$?

$Q_0 = \{f\}$

1. $qa' \rightarrow f; q \in Q; a' \in \Sigma: sb \rightarrow f, fa \rightarrow f$

$Q_1 = \{f\} \cup \{s, f\} = \{f, s\} \dots s$ je ukončující

Odpověď: **NE**, protože s je ukončující

Otázka: Je $L(M)$ konečný? $k = \text{Card}(Q) = 2$

Všechny řetězce $z \in \Sigma^*: 2 \leq |z| < 4: aa, bb, ab \in L(M), \dots$

Odpověď: **NE**, protože existuje $z \in L(M), k \leq |z| < 2k$

Algoritmus: Problém ekvivalence

- **Vstup:** Dva minimální KA, M_1 a M_2
 - **Výstup:** **ANO**, pokud $L(M_1) = L(M_2)$
NE, pokud $L(M_1) \neq L(M_2)$
-

- **Metoda:**

- **if** M_1 má stejnou strukturu jako M_2 až na pojmenování stavů

then napiš('ANO')

else napiš('NE')

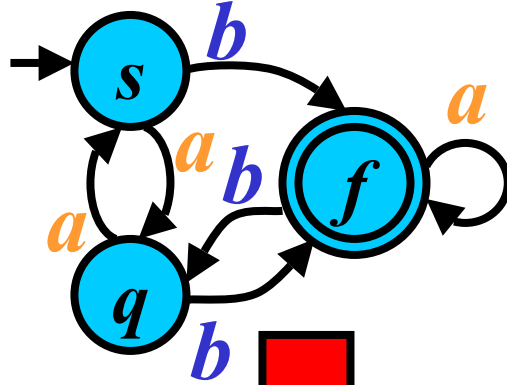
Celkově:

Problém ekvivalence je pro KA rozhodnutelný

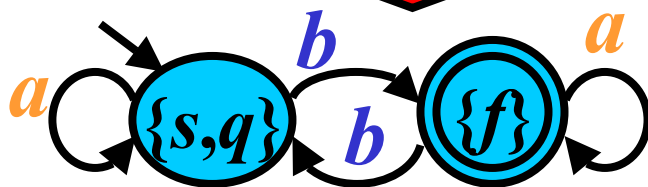
Problém ekvivalence: Příklad

Otázka: $L(M_1) = L(M_2)$?

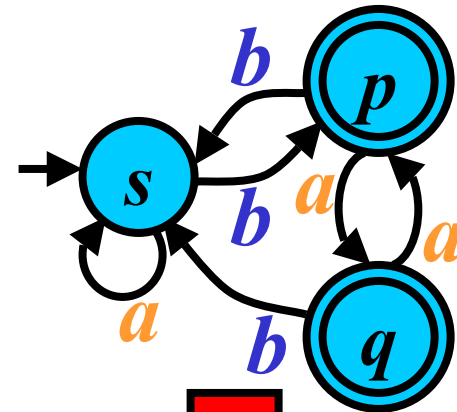
M_1 :



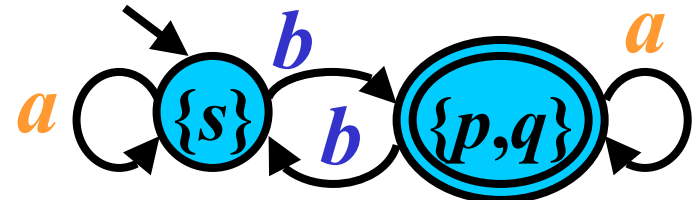
M_{min1} :



M_2 :



M_{min2} :



Minimální KA

Odpověď: **ANO**, protože M_{min1} má stejnou strukturu jako M_{min2}