# Kapitola XI. Vlastnosti regulárních jazyků

### Pumping lemma pro RJ

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

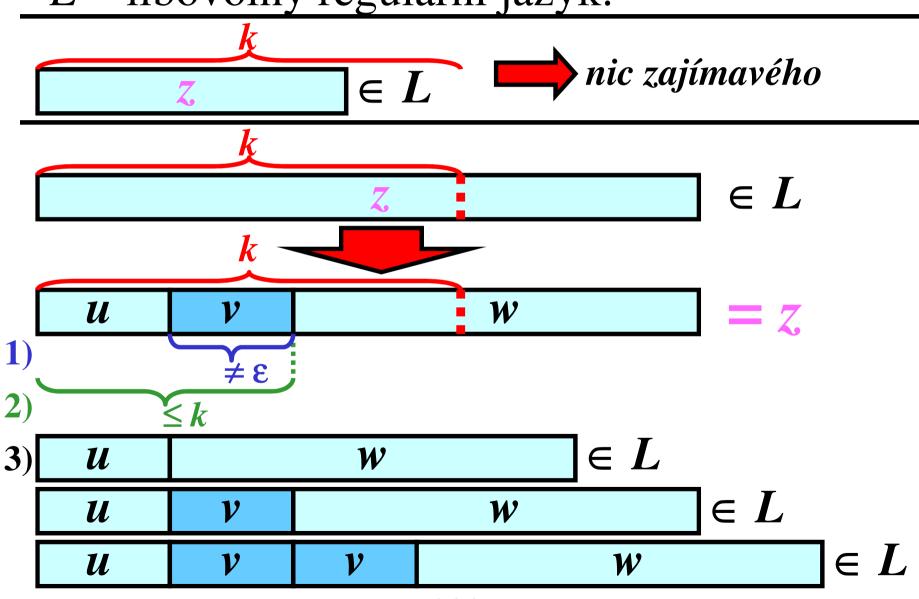
Nechť L je RJ. Pak existuje k≥ 1 takové, že:
pokud z ∈ L a |z| ≥ k, pak existuje u,v,w: z = uvw,
1) v ≠ ε 2) |uv| ≤ k
3) pro každé m ≥ 0, uv<sup>m</sup>w ∈ L

**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 = abc:  $z \in L(r)$  a  $|z| \ge 3$ :  $uv^0w = ab^0c = ac \in L(r)$   $uv^1w = ab^1c = abc \in L(r)$   $uv^2w = ab^2c = abbc \in L(r)$  $v \ne \varepsilon$ ,  $|uv| = 2 \le 3$
- pro z = abbc:  $z \in L(r)$  a  $|z| \ge 3$ :  $uv^0w = ab^0bc = abc \in L(r)$ •  $uv^1w = ab^1bc = abbc \in L(r)$ •  $uv^2w = ab^2bc = abbbc \in L(r)$ 
  - $v \neq \varepsilon$ ,  $|uv| = 2 \le 3$

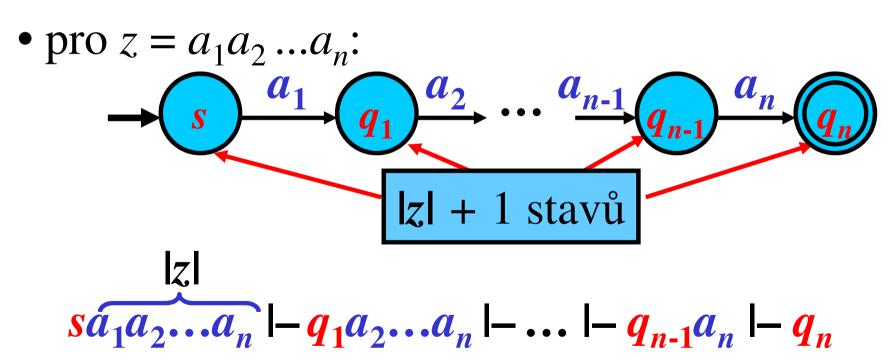
### Pumping lemma: Ilustrace

• L = libovolný regulární jazyk:



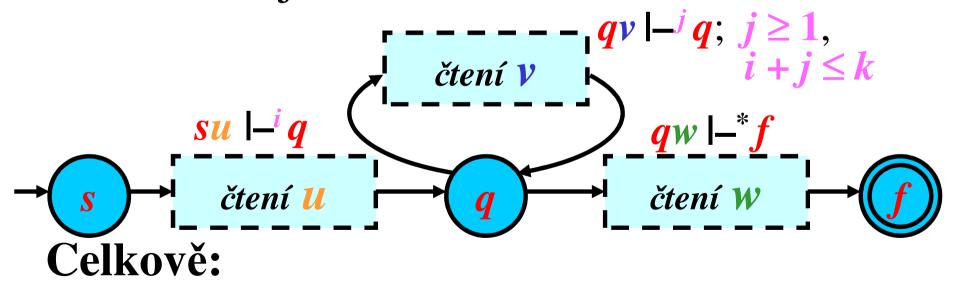
### Důkaz pumping lemmy 1/3

- Nechť 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:



# Důkaz pumping lemmy 2/3

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



$$sz = suvw \mid -iqvw \mid -jqw \mid -*f, f \in F$$

# Důkaz pumping lemmy 3/3

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

$$suw \mid -iqw \mid -*f, f \in F$$

• pro každé m > 0,

#### Celkově:

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

# Pumping lemma: Aplikace I.

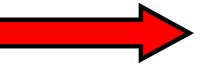
 Pomocí pumping lemmy pro RJ často provádíme důkaz sporem, že daný jazyk <u>není</u> regulární:

Předpokládejme, že L je regulární

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

Pro <u>všechny</u> 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:  $uv^m w \in L$ 

špatný předpoklad

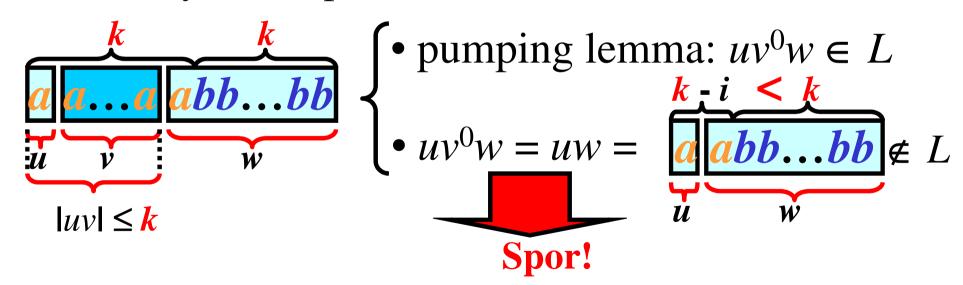


Proto **L není regulární** 

# Pumping Lemma: Příklad

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

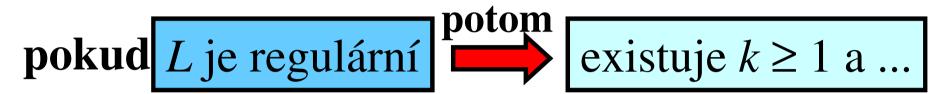
- 1) Předpokládejme, že L je regulární. Nechť  $k \ge 1$  je konstanta z pumping lemmy pro jazyk L.
- 2) Nechť  $z = a^k b^k$ :  $a^k b^k \in L$ ,  $|z| = |a^k b^k| = 2k \ge k$
- 3) Všechny dekompozice z na uvw,  $v \neq \varepsilon$ ,  $|uv| \leq k$ :



4) Proto L není regulární jazyk

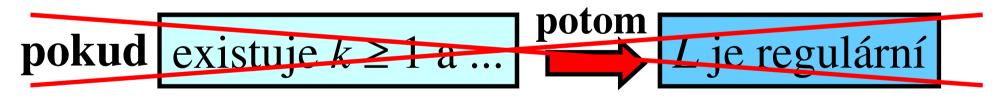
### Poznámka k použití pumping lemmy

• Pumping lemma:



#### Základní aplikace pumping lemmy:

- důkaz sporem, že L není regulární jazyk.
- Ale následující implikace je špatná:



 Nelze použít pumping lemmy k dokázání, že daný jazyk L je regulární!!

# Pumping lemma: Aplikace II. 1/3

• Pumping lemmu je možné použít k dokazování dalších tvrzení.

#### **Ilustrace:**

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

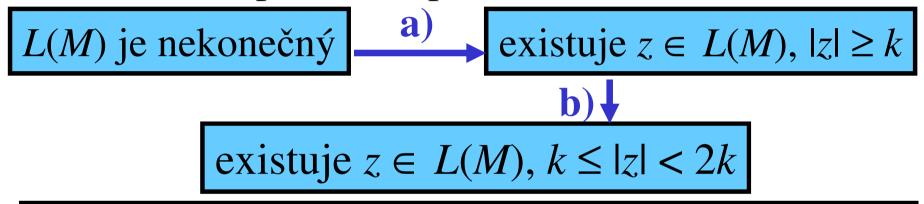
#### Důkaz:

1) existuje  $z \in L(M)$ ,  $k \le |z| < 2k \Rightarrow L(M)$  je nekonečný: pokud  $z \in L(M)$ ,  $k \le |z|$ , potom podle PL: z = uvw,  $v \ne \varepsilon$  a dále pro každé  $m \ge 0$ :  $uv^m w \in L(M)$ 

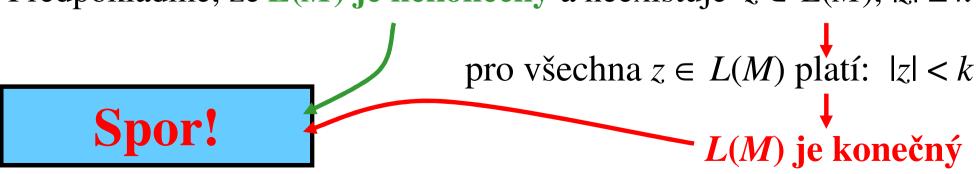
L(M) je nekonečný

# Pumping Lemma: Aplikace II. 2/3

- 2) L(M) je nekonečný  $\Rightarrow$  existuje  $z \in L(M)$ ,  $k \le |z| < 2k$ :
- Dokážeme sporem, že platí:



- a) Dokážeme sporem, že:
- L(M) je nekonečný  $\Rightarrow$  existuje  $z \in L(M)$ ,  $|z| \ge k$ Předpokládme, že L(M) je nekonečný a neexistuje  $z \in L(M)$ ,  $|z| \ge k$



# Pumping Lemma: Aplikace II. 3/3

- b) Dokážeme sporem:
- existuje  $z \in L(M)$ ,  $|z| \ge k \Rightarrow$  existuje  $z \in L(M)$ ,  $k \le |z| < 2k$

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



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

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

Pokud  $z_0 \in L(M)$  a  $|z_0| \ge k$ , PL zaručuje:  $z_0 = uvw$ ,

 $|uv| \le k$  a pro každé  $m \ge 0$ ,  $uv^m w \in L(M)$ 

$$\geq 2k \leq k$$

$$|uw| = |z_0| - |v| \ge k$$

pro 
$$m = 0$$
:  $uv^m w = uw \in L(M)$ 

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

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

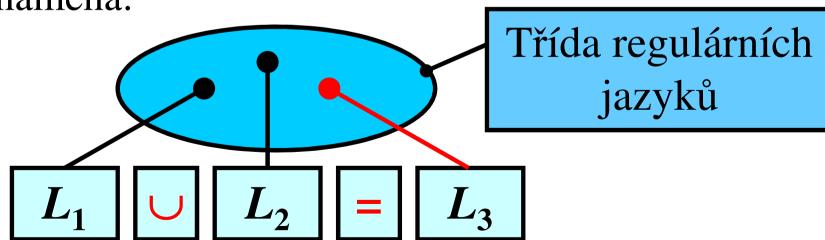
SPOR!

### 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.

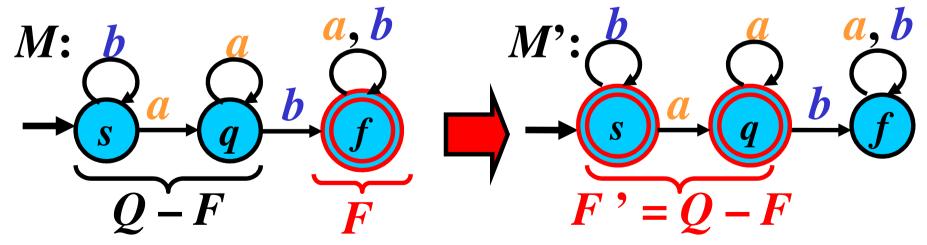
#### Důkaz:

- Nechť  $L_1$ ,  $L_2$  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  $L_1L_2$
  - $r_1 + r_2$  je RV značící  $L_1 \cup L_2$
  - $r_1^*$  je RV značící  $L_1^*$
- Každý RV značí regulární jazyk, tedy
- $L_1L_2$ ,  $L_1 \cup L_2$ ,  $L_1^*$  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:
- $\bullet F' := Q F$

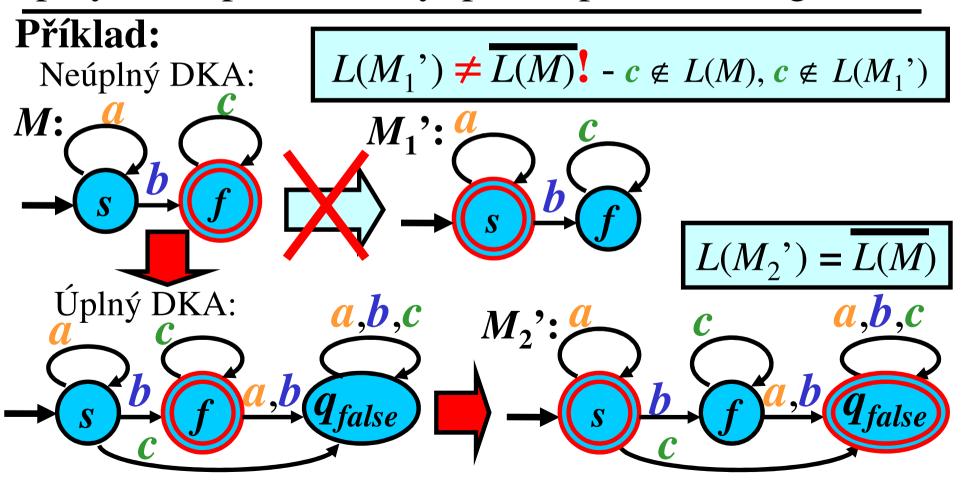
#### Příklad:



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

# KA pro doplněk: Problém

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



### Uzávěrové vlastnosti: Doplněk

Tvrzení: Třída regulárních jazyků je uzavřena vůči doplňku.

#### Důkaz:

- Nechť L je regulární jazyk
- Pak existuje úplný DKA M: L(M) = L
- Můžeme sestrojit úplný DKA M': L(M') = L užitím předchozího algoritmu
- Každý KA definuje regulární jazyk, tedy
   L je regulární jazyk

### Uzávěrové vlastnosti: Průnik

Tvrzení: Třída regulárních jazyků je uzavřena vůči průniku.

#### **Důkaz:**

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

# Boolova algebra jazyků

**Definice:** Nechť je třída jazyků uzavřena vůči sjednocení, průniku a doplňku. Potom tato třída tvoří *Boolovu algebru jazyků*.

**Tvrzení:** Třída regulárních jazyků tvoří Booleovu algebru jazyků.

#### **Důkaz:**

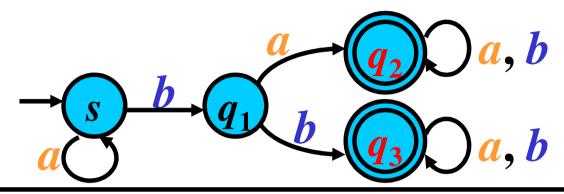
• Třída regulárních jazyků je uzavřená vůči sjednocení, průniku a doplňku.

### Minimalizace: Rozlišitelné stavy

Myšlenka: Řetězec *w rozlišuje* stavy *p* a *q*, pokud se DSKA "dostane" z <u>právě z jedné</u> z konfigurací *pw* a *qw* do koncového stavu.

**Definice:** Necht'  $M = (Q, \Sigma, R, s, F)$  je DSKA a necht'  $p, q \in Q, p \neq q$ . Stavy p a q jsou rozlišiteln'e pokud existuje řetězec  $w \in \Sigma^*$  takový, že:  $pw \vdash p'$  and  $qw \vdash p'$ ,  $q' \in Q$  a  $((p' \in F \text{ a } q' \notin F) \text{ nebo } (p' \notin F \text{ a } q' \in F))$ . Jinak stavy p a q jsou nerozlišiteln'e.

### Rozlišitelné stavy: Příklad



• s a  $q_1$  jsou rozlišitelné, protože např. pro w = a:

$$sa \vdash s, s \notin F$$
 $q_1a \vdash q_2, q_2 \in F$ 

•  $q_2$  a  $q_3$  jsou **nerozlišitelné**, protože pro každé  $w \in \Sigma^*$ :

$$q_2w \vdash^* q_2, q_2 \in F$$
  
 $q_3w \vdash^* q_3, q_3 \in F$ 

• Ostatní dvojice stavů jsou triviálně **rozlišitelné** pro  $w = \varepsilon$ .

#### Minimální KA

**Definice:** Nechť *M* je **DSKA**. Potom, *M* je *minimální KA*, pokud *M* obsahuje pouze rozlišitelné stavy.

**Tvrzení:** Pro každý DSKA M, existuje ekvivalentní minimální KA  $M_m$ .

Důkaz: Použij následující algoritmus.

### Algoritmus: Minimalizace KA

- Vstup: DSKA  $M = (Q, \Sigma, R, s, F)$
- Výstup: Minimální KA  $M_m = (Q_m, \Sigma, R_m, s_m, F_m)$
- Metoda:
- $Q_m = \{ \{p: p \in F\}, \{q: q \in Q F\} \};$
- repeat

if existuje  $X \in Q_m$ ,  $d \in \Sigma$ ,  $X_1, X_2 \subset X$  takové, že:

$$X = X_1 \cup X_2, X_1 \cap X_2 = \emptyset$$
 and

$$\{q_1: p_1 \in X_1, p_1 d \to q_1 \in R\} \subseteq Q_1, Q_1 \in Q_m,$$

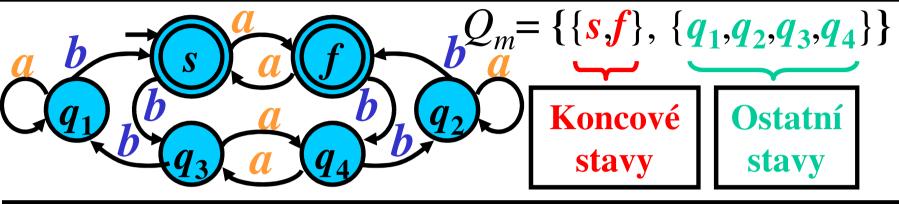
$$\{q_2: p_2 \in X_2, p_2 d \to q_2 \in R\} \cap Q_1 = \emptyset$$

**then** rozštěp X na  $X_1$  a  $X_2$  v  $Q_m$ 

until není možné provést další štěpení;

- $R_m = \{ Xa \rightarrow Y: X, Y \in Q_m, pa \rightarrow q \in R, p \in X, q \in Y, a \in \Sigma \};$
- $s_m = X$ :  $s \in X$ ;  $F_m := \{X : X \in Q_m, X \cap F \neq \emptyset\}$ .

#### Minimalizace: Příklad 1/4



1) 
$$X = \{s, f\}$$
: Z jedné množiny  $d = a$ :  $sa \rightarrow f$   $d = b$ :  $sb \rightarrow q_3$   $q_4$ 

2) 
$$X = \{q_1, q_2, q_3, q_4\}$$
: Z jedné množiny  $d = a$ :  $q_1 a \to q_1$   $d = b$ :  $\{q_1 b \to g_1\}$   $\{q_2 a \to g_2\}$   $\{q_2 b \to g_1\}$   $\{q_3 b \to g_2\}$   $\{q_4 b \to g_2\}$   $\{q_4 b \to g_2\}$  Štěpení:  $\{q_1, q_2, q_3, q_4\}$   $\{q_1, q_2\}$ ,  $\{q_3, q_4\}$   $\{q_1, q_2\}$   $\{q_1, q_2\}$ 

#### Minimalizace: Příklad 2/4

$$Q_m = \{\{s,f\}, \{q_1,q_2\}, \{q_3,q_4\}\}$$

- 1)  $X = \{s, f\}$ : Z jedné množiny Z jedné množiny d = a:  $sa \to f$  d = b:  $sb \to q_3$   $fb \to q_4$
- 2)  $X = \{q_1, q_2\}$ : Z jedné množiny d = a:  $q_1 a \rightarrow q_1$   $q_2 a \rightarrow q_2$  d = b:  $q_1 b \rightarrow s$   $q_2 b \rightarrow f$
- 3)  $X = \{q_3, q_4\}$ : Z jedné množiny d = a:  $q_3 a \rightarrow q_4$  d = b:  $q_3 b \rightarrow q_1$   $q_4 a \rightarrow q_4$  d = b:  $q_4 b \rightarrow q_2$

### Žádné další štěpení !!!

#### Minimalizace: Příklad 3/4

$$Q_m = \{ \{ s, f \}, \{ q_1, q_2 \}, \{ q_3, q_4 \} \}$$

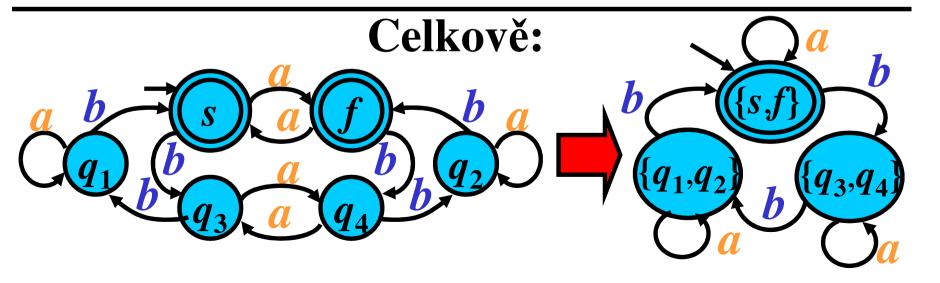
$$\begin{array}{l} sa & \rightarrow f \in R: \\ fa & \rightarrow s \in R: \end{array} \} \Longrightarrow \{s,f\}a & \rightarrow \{s,f\} \in R_m \\ sb & \rightarrow q_3 \in R: \\ fb & \rightarrow q_4 \in R: \end{array} \} \Longrightarrow \{s,f\}b & \rightarrow \{q_3,q_4\} \in R_m \\ q_1a & \rightarrow q_1 \in R: \\ q_2a & \rightarrow q_2 \in R: \end{array} \} \Longrightarrow \{q_1,q_2\}a & \rightarrow \{q_1,q_2\} \in R_m \\ q_1b & \rightarrow s \in R: \\ q_2b & \rightarrow f \in R: \end{cases} \Longrightarrow \{q_1,q_2\}b & \rightarrow \{s,f\} \in R_m \\ q_2b & \rightarrow f \in R: \end{cases} \Longrightarrow \{q_3,q_4\}a & \rightarrow \{q_3,q_4\} \in R_m \\ q_3a & \rightarrow q_3 \in R: \\ q_4a & \rightarrow q_4 \in R: \end{cases} \Longrightarrow \{q_3,q_4\}a & \rightarrow \{q_3,q_4\} \in R_m \\ q_3b & \rightarrow q_1 \in R: \\ q_4b & \rightarrow q_2 \in R: \end{cases} \Longrightarrow \{q_3,q_4\}b & \rightarrow \{q_1,q_2\} \in R_m \end{array}$$

#### Minimalizace: Příklad 4/4

$$\mathbf{s} \in \{\mathbf{s},\mathbf{f}\} \implies s_m := \{\mathbf{s},\mathbf{f}\}$$

$$s \in F$$
:
 $f \in F$ :
 $s \in F$ :
 $s$ 

$$\begin{split} &M_{m} = (Q_{m}, \Sigma, R_{m}, s_{m}, F_{m}), \text{ kde: } \Sigma = \{a, b\}, s_{m} = \{s, f\} \\ &Q_{m} = \{\{s, f\}, \{q_{1}, q_{2}\}, \{q_{3}, q_{4}\}\}, F_{m} = \{\{s, f\}\} \} \\ &R_{m} = \{\{s, f\}a \rightarrow \{s, f\}, \{s, f\}b \rightarrow \{q_{3}, q_{4}\}, \{q_{1}, q_{2}\}a \rightarrow \{q_{1}, q_{2}\}, \{q_{1}, q_{2}\}b \rightarrow \{s, f\}, \{q_{3}, q_{4}\}a \rightarrow \{q_{3}, q_{4}\}, \{q_{3}, q_{4}\}b \rightarrow \{q_{1}, q_{2}\} \} \end{split}$$



# Typy KA: Shrnutí

	KA	KA bez e-přech.	DKA	Úplný KA	DSKA	Minimální KA
Počet všech pravidel tvaru $p \rightarrow q$ , kde $p, q \in Q$	0-n	0	0	0	0	0
Počet pravidel tvaru $pa \rightarrow q$ , pro libovolné $p \in Q$ a libovolné $a \in \Sigma$	0-n	0-n	0-1	1	1	1
Počet všech nedostupných stavů	0-n	0-n	0-n	0-n	0	0
Počet všech neukončujících stavů	0-n	0-n	0-n	0-n	0-1	0-1
Počet všech možných těchto automatů pro jeden regulární jazyk	8	8	8	8	8	1

# Hlavní rozhodnutelné problémy

#### 1. Problém členství:

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

#### 2. Problém prázdnosti:

• 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  $sw \vdash f, f \in F$  then napiš('ANO')
  else napiš('NE')

#### **Celkově:**

Problém členství je pro KA rozhodnutelný

### Algoritmus: Problém prázdnosti

- 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 prázdnosti 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'  $k = \operatorname{card}(Q)$
- if existuje  $z \in L(M)$ ,  $k \le |z| < 2k$  then napiš('NE')
  else napiš('ANO')

**Pozn.:** Tento algoritmus je založen na tvrzení: L(M) je nekonečný  $\Leftrightarrow$  existuje z:  $z \in L(M)$ ,  $k \le |z| < 2k$ 

#### Celkově:

Problém konečnosti je pro KA rozhodnutelný

# Rozhodnutelné problémy: Příklad

$$M: a$$
  $b$   $b$   $a$ 

```
Otázka: ab \in L(M)?

sab \mid -sb \mid -f, f \in F
```

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

Otázka:  $L(M) = \emptyset$ ?

$$Q_0 = \{ f \}$$

1.  $qa' \rightarrow f$ ;  $q \in Q$ ;  $a' \in \Sigma$ :  $sb \rightarrow f$ ,  $fa \rightarrow f$  $Q_1 = \{f\} \cup \{s, f\} = \{f, s\}$  ... s je ukončující

Odpověď: NE, protože s je ukončující

Otázka: Je L(M) konečný? k = Card(Q) = 2Všechny řetězce  $z \in \Sigma^*$ :  $2 \le |z| < 4$ :  $aa, bb, ab \in L(M)$ , ...

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

### Algoritmus: Problém ekvivalence

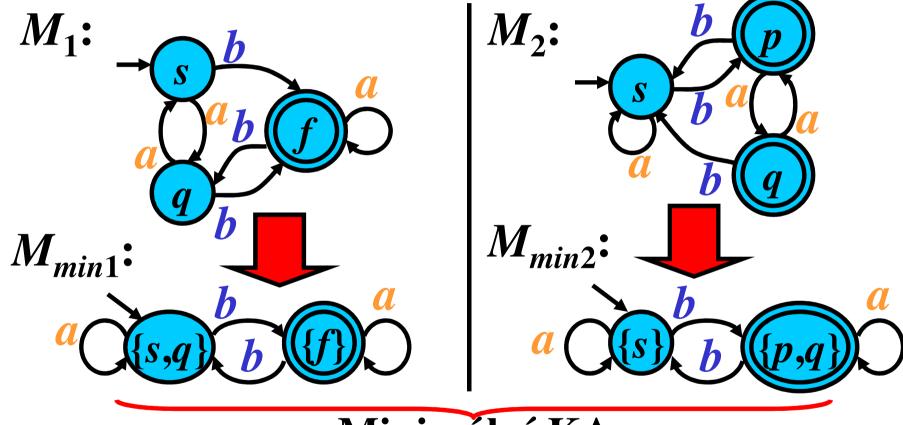
- Vstup: Dva minimální KA,  $M_1$ a  $M_2$
- Výstup: ANO, pokud  $L(M_1) = L(M_2)$ NE, pokud  $L(M_1) \neq L(M_2)$
- Metoda:
- if M<sub>1</sub> má stejnou strukturu jako M<sub>2</sub> až na pojmenování stavů
  then napiš('ANO')
  else napiš('NE')

#### **Celkově:**

Problém ekvivalence je pro KA rozhodnutelný

#### Problém ekvivalence: Příklad

**Otázka:**  $L(M_1) = L(M_2)$ ?



#### Minimální KA

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