2.2.3 Reguläre Ausdrücke

<u>Def</u>

Reguläre Ausdrücke sind wie folgt def:

- $\emptyset$ ist reg. Ausdruck
- ε — " —
- Für jedes $a \in \Sigma$ ist a ein reg. Ausdr.
- Wenn α und β reg. Ausdr. sind, dann $\alpha\beta$, $(\alpha|\beta)$ und $(\alpha)^*$

<u>Bsp</u> $\Sigma = \{1,0\}$, $(1)^*(0|1)(0)^* \cancel{\overset{?}{=}} (1)^*(0)^*$

Def (Beschriebene Sprache)

Ein regulärer Ausdruck γ beschreibt die Sprache $L(\gamma)$

- Falls $\gamma = \emptyset$, dann $L(\emptyset) = \emptyset$
- Falls $\gamma = \varepsilon$, dann $L(\gamma) = \{\varepsilon\}$
- Falls $\gamma = a$ $(a \in \Sigma)$, dann $L(\gamma) = \{a\}$
- Falls $\gamma = \alpha\beta$, dann $L(\gamma) = L(\alpha)L(\beta)$
  $= \{uv \mid u \in L(\alpha), v \in L(\beta)\}$
- Falls $\gamma = (\alpha|\beta)$, dann $L(\gamma) = L(\alpha) \cup L(\beta)$
- Falls γ $(\alpha)^*$, dann $L(\gamma) = L(\alpha)^*$

Bsp: $(a|b)^*ba$ beschreibt die Spr., die durch unsere Beispielautomaten akzeptiert wird.

<u>Bem</u>: Alle endl. Spr. sind durch reg. beschreibbar.

$L' = \{w_1, w_2, \dots, w_n\}$

$L(w_1 | w_2 | w_3 | \dots | w_n) = L'$

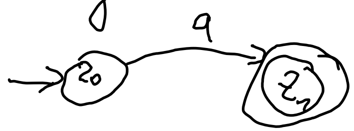
Satz (Kleene)

Die Menge der durch reg. Ausdr. beschreibbaren Sprachen ist genau die Menge der regulären Spr.

<u>Bew</u>:

($\Longrightarrow$) Wir zeigen: Wenn A durch γ beschreibbar ist, dann ex. ein <u>ε-NFA</u> M mit $T(M) = A$.

IA: Für $\gamma = \emptyset$, $\gamma = \varepsilon$ und $\gamma = a$ ($a \in \Sigma$) offensichtlich.



IV: Für reg. Ausdrücke α und β seien

$M_\alpha = (Z_\alpha, \Sigma, \delta_\alpha, S_\alpha, E_\alpha)$ und $M_\beta = (Z_\beta, \Sigma, \delta_\beta, S_\beta, E_\beta)$

$\rightarrow Z_\alpha \cap Z_\beta = \emptyset$,

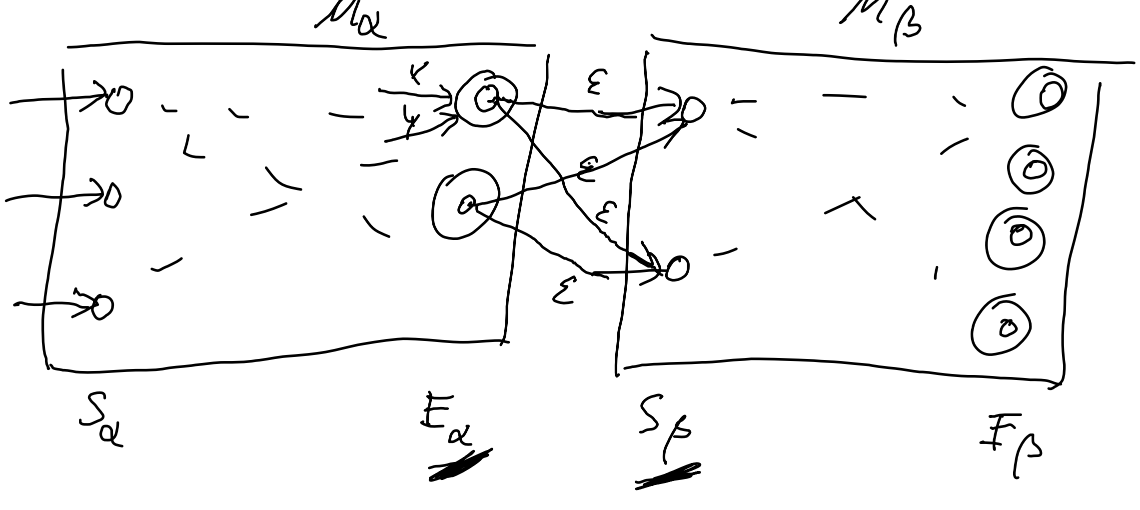
die ε-NFAs mit $L(\alpha) = T(M_\alpha)$ und $L(\beta) = T(M_\beta)$.

IS:

1) $\gamma = \alpha\beta$. konstr. $M_\gamma = (Z_\alpha \cup Z_\beta, \Sigma, \delta_\gamma, S_\alpha, E_\beta)$

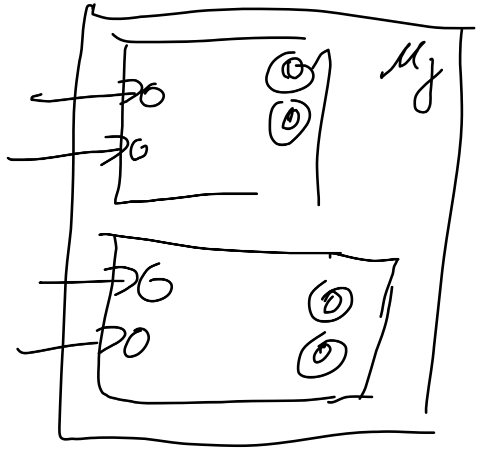
$$\delta_\gamma = \delta_\alpha \cup \delta_\beta \cup \delta'$$

$$\delta'(z, \varepsilon) = S_\beta \qquad z \in E_\alpha$$



2) $\gamma = (\alpha \mid \beta)$. konstr. M_γ!

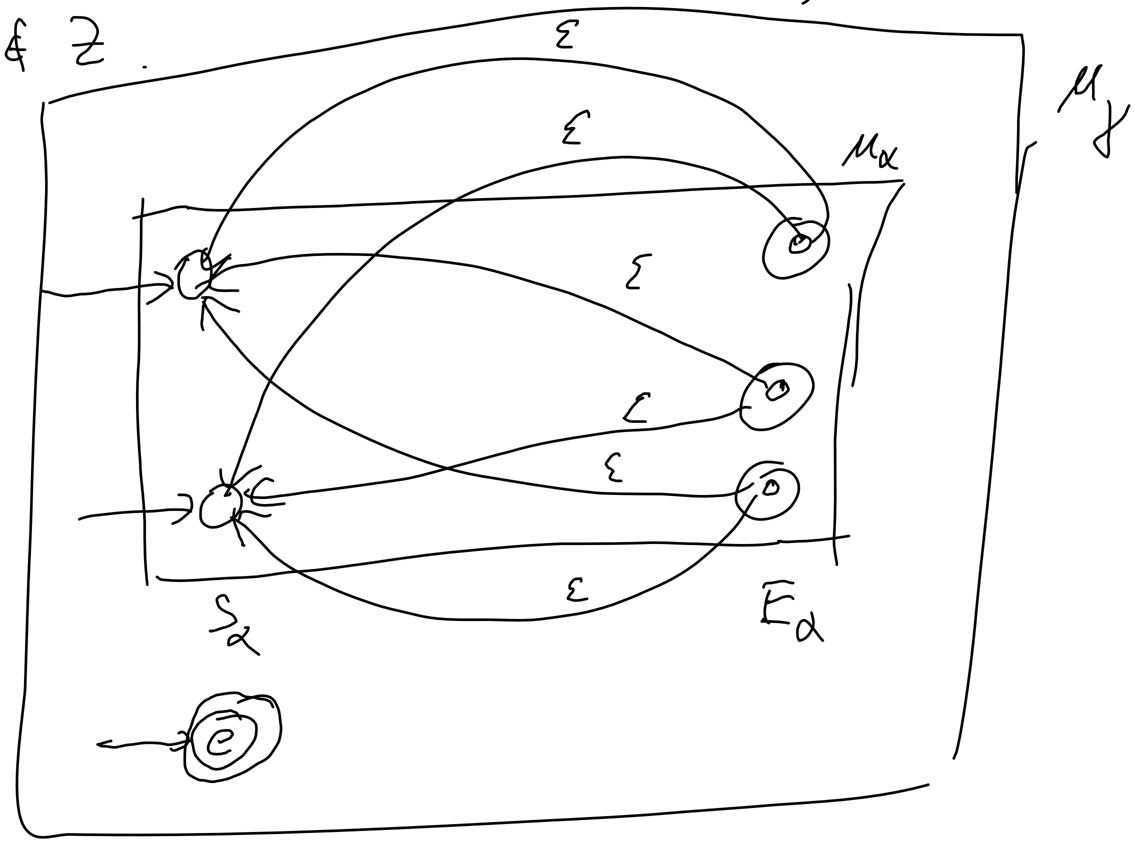
$$M_\gamma = (Z_\alpha \cup Z_\beta,\ \Sigma,\ \delta_\alpha \cup \delta_\beta,\ S_\alpha \cup S_\beta,\ F_\alpha \cup F_\beta)$$



3) $\gamma = (\alpha)^*$

Konstr $M_\gamma = (Z_\alpha \cup \{e\}, \Sigma, \delta_\alpha \cup \delta', S_\alpha \cup \{e\}, E_\alpha \cup \{e\})$

wobei $e \notin Z$.



$\delta'(z, \varepsilon) = S$ für alle $z \in E_\alpha$.

Offensichtlich: $L(\gamma) = T(M_\gamma)$

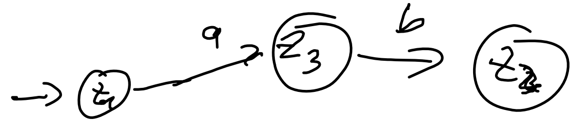
(⇐) Wir konstruieren induktiv für einen gegebenen DFA M einen reg. Ausdruck γ mit $T(M) = L(\gamma)$.

Alle Zustände in M sind von 1 bis $n = |Z|$ durchnummeriert. Sei $R^k_{i,j}$ für $0 \le i,j,k \le n$ die folgende Menge:

$$R^k_{i,j} = \{ w \in \Sigma^* \mid w \text{ überführt } M \text{ gestartet in } z_i \text{ in } z_j \ (\hat{\delta}(z_i, w) = z_j), \text{ ohne dass ein Zustand außer } z_i \text{ und } z_j \text{ mit einem Index größer als } k \text{ besucht wird} \}$$

$R^0_{i,j} = \{ a \in \Sigma \mid \delta(z_i, a) = z_j \}$ für den Fall $i \neq j$

$R^0_{i,j} = \{ a \in \Sigma \mid \delta(z_i, a) = z_j \} \cup \{ \varepsilon \}$ für den Fall $i = j$



$R^0_{1,3} = \{ a \}$

$R^0_{1,2} = \emptyset$

Für $k=0$ ist $R^k_{i,j}$ endlich, d.h. es ex. ein reg Ausdruck, der $R^k_{i,j}$ beschreibt.

Induktion über k

$(k \to k+1)$:

$$R^{k+1}_{i,j} = R^k_{i,j} \cup R^k_{i,k+1} \left(R^k_{k+1,k+1}\right)^* R^k_{k+1,j}$$

$$\begin{tikzcd} z_i \arrow[r, dashed] & z_{k+1} \arrow[loop above, dashed] \arrow[r, dashed] & z_j \end{tikzcd}$$

Falls $\alpha^k_{i,j}$ ein reg Ausdruck für $R^k_{i,j}$ ist, so gilt

$$\alpha^{k+1}_{i,j} = \left(\alpha^k_{i,j} \mid \alpha^k_{i,k+1} \left(\alpha^k_{k+1,k+1}\right)^* \alpha^k_{k+1,j}\right)$$

Da $T(M) = \bigcup_{z_i \in E} R^n_{1,i}$ beschreibt $(\alpha^n_{1,i_1} \mid \dots \mid \alpha^n_{1,i_m})$ falls $E = \{z_{i_1}, \dots, z_{i_m}\}$ ist $\square$

2.2.4 Pumping-Lemma

Hilfsmittel um zu zeigen, dass eine Spr. nicht regulär ist.

Satz (Pumping-Lemma, Bar-Hillel-Lemma, uvw-Theorem)

Sei L ein reg. Spr. Dann gibt es eine Zahl $n \in \mathbb{N}$, so dass sich für alle Wörter $x \in L$ mit $|x| \geq n$ eine Zerlegung $x = uvw$ finden lässt, so dass folgende Bedingungen gelten:

(1) $|v| \geq 1$

(2) $|uv| \leq n$

(3) Für alle $i \geq 0$ gilt: $uv^iw \in L$

$\forall L$ reg $\Rightarrow (\exists n \in \mathbb{N} : (\forall x \in L : |x| \geq n \Rightarrow (\exists u, v, w : x = uvw \wedge |v| \geq 1 \wedge |uv| \leq n \wedge (\forall i \in \mathbb{N} : uv^iw \in L))))$