

Die Klasse der **regulären Sprachen** über dem Alphabet Σ

(1) ist unter folgenden Operatoren abgeschlossen:

Vereinigung

Konkatenation

Kleeneabschluss

(2) enthält die Sprachen \emptyset , $\{\epsilon\}$ und $\{a\}$ für alle $a \in \Sigma$

Die Klasse der **regulären Sprachen** ist die **kleinste Klasse** von Sprachen mit den Eigenschaften (1) und (2).

1. \emptyset und ε sind reguläre Ausdrücke über Σ .
2. Für $a \in \Sigma$ ist a ein regulärer Ausdruck über Σ .
3. Sind α , β reguläre Ausdrücke über Σ , so sind
 - $(\alpha\beta)$ Konkatenation
 - $(\alpha + \beta)$ Auswahl, Vereinigung
 - (α^*) Kleeneabschlussreguläre Ausdrücke über Σ .
4. Nichts sonst ist ein regulärer Ausdruck über Σ .

saloppe Schreibweise (abstrakte Syntax):

$$\alpha ::= \emptyset \mid \varepsilon \mid a \mid \alpha\beta \mid \alpha + \beta \mid \alpha^*$$

abstrakte Syntax regulärer Ausdrücke über Σ :

$$\alpha ::= \emptyset \mid \varepsilon \mid a \mid \alpha\beta \mid \alpha + \beta \mid \alpha^*$$

induktive Definition von $\mathcal{L}(\alpha)$:

$$\mathcal{L}(\emptyset) = \emptyset$$

$$\mathcal{L}(\varepsilon) = \{\varepsilon\}$$

$$\mathcal{L}(a) = \{a\} \quad \text{für alle } a \in \Sigma$$

$$\mathcal{L}(\alpha\beta) = \mathcal{L}(\alpha)\mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha + \beta) = \mathcal{L}(\alpha) \cup \mathcal{L}(\beta)$$

$$\mathcal{L}(\alpha^*) = \mathcal{L}(\alpha)^*$$

Welche Sprache ...?

1551

welche Sprache beschreibt der reguläre Ausdruck ... ?
über $\Sigma = \{a, b\}$

$$\alpha = (a + b)^* b$$

Antwort: $\mathcal{L}(\alpha) = \{wb : w \in \Sigma^*\}$

$$\beta = ((a + b)^* b)^* = \alpha^*$$

Antwort: $\mathcal{L}(\beta) = \mathcal{L}(\alpha)^* = \{\varepsilon\} \cup \{wb : w \in \Sigma^*\}$

$$\gamma = (a^* b a^* b)^* a^*$$

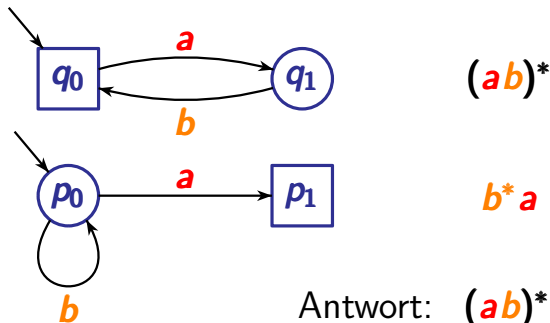
Antwort: $\mathcal{L}(\gamma) = \{w \in \Sigma^* : \#(b, w) \text{ ist gerade}\}$

↑
Anzahl an b 's in w

$$L = \{ w \in \{a, b\}^* : \text{“vor jedem } b \text{ in } w \text{ steht ein } a\text{”} \}$$

Antwort: $(a^*ab)^*a^* = (a^+b)^*a^*$

$L' = \mathcal{L}(\mathcal{M})$ für folgenden NFA \mathcal{M} :



Antwort: $(ab)^* + b^*a$

(semantische) Äquivalenz regulärer Ausdrücke:

$$\alpha \equiv \beta \stackrel{\text{def}}{\iff} \mathcal{L}(\alpha) = \mathcal{L}(\beta)$$

Beispiele: für alle regulären Ausdrücken α , β , γ gilt:

$$\alpha + \beta \equiv \beta + \alpha \longleftarrow \text{Kommutativität von } \cup$$

$$\left. \begin{array}{l} \alpha(\beta + \gamma) \equiv \alpha\beta + \alpha\gamma \\ (\beta + \gamma)\alpha \equiv \beta\alpha + \gamma\alpha \end{array} \right\} \text{Distributiv-} \\ \text{gesetze}$$

⋮

$$\varepsilon + \alpha(\varepsilon + \alpha^+) \equiv \varepsilon + \alpha\varepsilon + \alpha\alpha^+ \equiv \alpha^*$$

$|\alpha|$ = Wortlänge von α als Wort über dem Alphabet $\{\emptyset, \varepsilon, +, *, (,)\} \cup \Sigma$

induktive Definition der Länge regulärer Ausdrücke:

$$|\emptyset| = |\varepsilon| = |a| = 1$$

$$|\alpha\beta| = |\alpha| + |\beta| + 2 \quad \leftarrow \alpha\beta = (\alpha\beta)$$

$$|\alpha + \beta| = |\alpha| + |\beta| + 3 \quad \leftarrow \alpha + \beta = (\alpha + \beta)$$

$$|\alpha^*| = |\alpha| + 3 \quad \leftarrow \alpha^* = (\alpha)^*$$

Beispiele: $|b^*c| = 1 + 1 + 3 + 2 = 7$

$$|bc + c| = 1 + 1 + 1 + 3 + 2 = 8$$

Sei L eine Sprache. Dann gilt:

L ist regulär

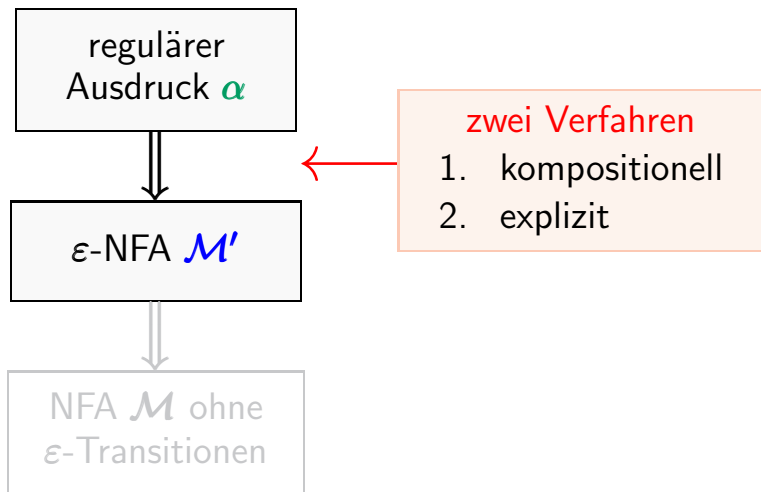
gdw $L = \mathcal{L}(G)$ für eine reguläre Grammatik G

gdw $L = \mathcal{L}(\mathcal{M})$ für einen DFA \mathcal{M}

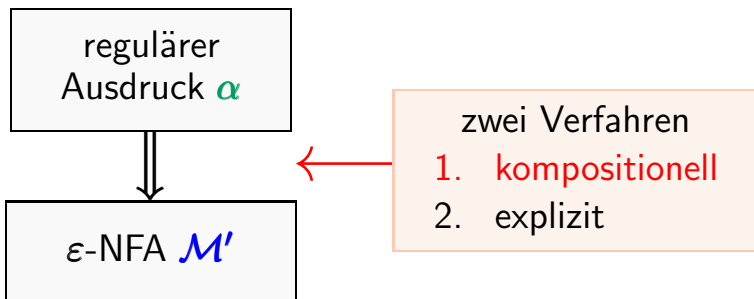
gdw $L = \mathcal{L}(\mathcal{M})$ für einen NFA \mathcal{M}

gdw $L = \mathcal{L}(\alpha)$ für einen regulären Ausdruck α

Zu jedem regulären Ausdruck α gibt es einen NFA \mathcal{M} ,
so dass $\mathcal{L}(\alpha) = \mathcal{L}(\mathcal{M})$.



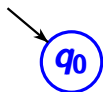
Zu jedem regulären Ausdruck α gibt es einen NFA \mathcal{M} ,
so dass $\mathcal{L}(\alpha) = \mathcal{L}(\mathcal{M})$.



Idee für das kompositionelle Verfahren:

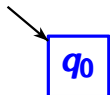
konstruiere $\mathcal{M}' = \mathcal{M}_\alpha$ durch strukturelle Induktion
über den syntaktischen Aufbau von α

$$\alpha = \emptyset$$



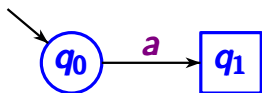
NFA ohne Endzustände
akzeptiert \emptyset

$$\alpha = \varepsilon$$



NFA für $\{\varepsilon\}$

$$\alpha = a \in \Sigma$$



NFA für $\{a\}$

$$\alpha = \beta + \gamma$$

konstruiere \mathcal{M}_β und \mathcal{M}_γ und
setze $\mathcal{M}_\alpha = \mathcal{M}_\beta \uplus \mathcal{M}_\gamma$

$$\alpha = \beta \circ \gamma$$

konstruiere \mathcal{M}_β und \mathcal{M}_γ und
setze $\mathcal{M}_\alpha = \mathcal{M}_\beta \circ \mathcal{M}_\gamma$

$$\alpha = \beta^*$$

konstruiere \mathcal{M}_β und setze $\mathcal{M}_\alpha = \mathcal{M}_\beta^*$

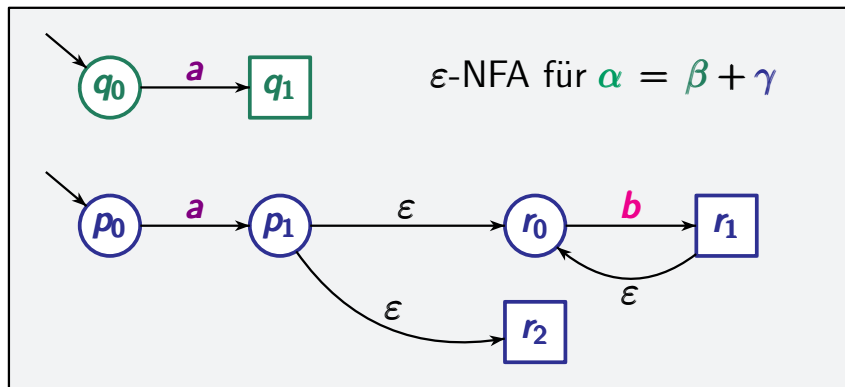
Beispiel: kompositionelles Verfahren

1679A

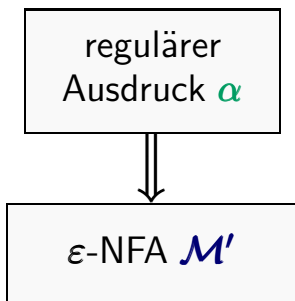
regulärer Ausdruck $\alpha = a + ab^*$

\uparrow \uparrow

β γ



Zu jedem regulären Ausdruck α gibt es einen NFA \mathcal{M} ,
so dass $\mathcal{L}(\alpha) = \mathcal{L}(\mathcal{M})$ und $|\mathcal{M}| = \mathcal{O}(|\alpha|)$.



zwei Verfahren

1. kompositionell
2. **explizit**



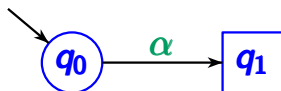
Idee für das explizite Verfahren:

\mathcal{M}' entsteht durch sukzessives Einfügen
von Zuständen und Transitionen

Explizite Konstruktion eines ε -NFA

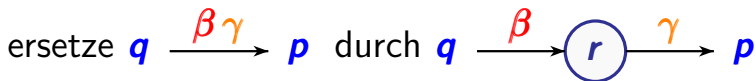
1722

Initialisierung:

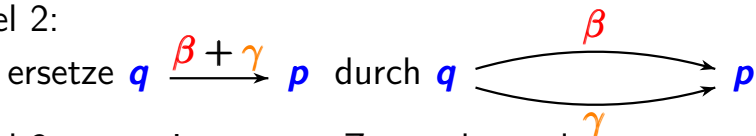


Iteration: solange möglich, wende eine der drei Regeln an

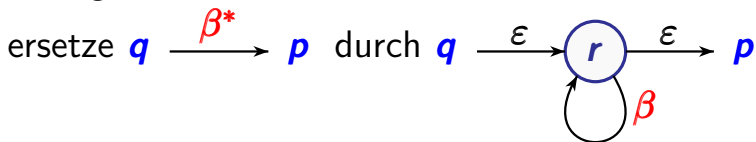
Regel 1: generiere neuen Zustand r und



Regel 2:



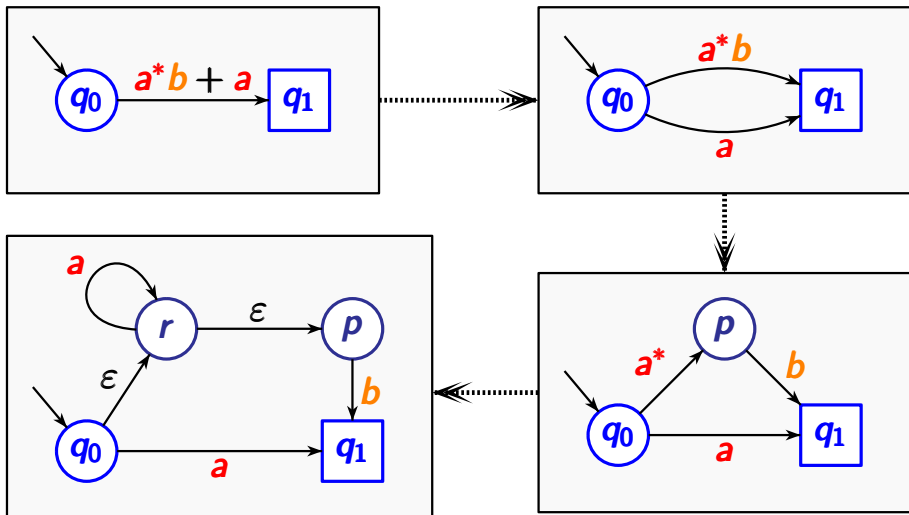
Regel 3: generiere neuen Zustand r und



Beispiel: explizite Konstruktion

1726

Konstruktion eines ε -NFA für $a^*b + a$:



gegeben: NFA $\mathcal{M} = (Q, \Sigma, \delta, Q_0, F)$

gesucht: regulärer Ausdruck α mit $\mathcal{L}(\alpha) = \mathcal{L}(\mathcal{M})$

Ersetzungsmethode:

für jeden Zustand $q \in Q$ berechne einen regulären Ausdruck α_q für die Sprache

$$L_q = \{ w \in \Sigma^* : \text{es gibt ein } p \in F \text{ mit } q \xrightarrow{w} p \}$$

durch Lösen eines Gleichungssystems

$$\mathcal{L}(\mathcal{M}) = \bigcup_{q \in Q_0} L_q = \mathcal{L}(\alpha), \quad \text{wobei } \alpha = \sum_{q \in Q_0} \alpha_q$$