TECHNISCHE UNIVERSITÄT DRESDEN

FORMALE SYSTEME

ÜBUNG 3

Eric Kunze

eric.kunze@tu-dresden.de

TU Dresden, 10. November 2021

WER BIN ICH?

- **Eric** Kunze
- `eric.kunze@tu-dresden.de`
- Telegram: `@oakoneric` bzw. `t.me/oakoneric`
- Fragen, Wünsche, Vorschläge, …
- Website mit Material: `https://oakoneric.github.io/fs21`
  keine Garantie für Vollständigkeit/Korrektheit,
  Slides enthalten Lösungs*ansätze* zu ausgewählten Aufgaben

WAS BISHER GESCHAH ...

- Basics Formale Sprachen
- Grammatiken und die Chomsky-Hierarchie
- Reguläre Sprachen

Ein **deterministischer endlicher Automat** (DFA) $\mathcal{M}$ ist ein Tupel $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit den folgenden Bestandteilen:

- Q: endliche Menge von Zuständen
- Σ: Alphabet
- δ: Übergangsfunktion, eine partielle Funktion $Q \times \Sigma \to Q$
- q_0: Startzustand $q_0 \in Q$
- F: Menge von Endzuständen $F \subseteq Q$

Sprache eines DFA: $L(\mathcal{M}) = \{w \in \Sigma^* : \delta^*(q_0, w) \in F\}$

WAS BISHER GESCHAH ...

- **totale Übergangsfunktion**:
  Einführung eines Fangzustandes
- Zusammenhang: reguläre Grammatik $\leftrightarrow$ endlicher Automat:

Satz: Die Klasse der Sprachen, die durch einen DFA erkannt werden können, ist genau die Klasse der regulären Sprachen.

Konstruktion einer regulären Grammatik $G_{\mathcal{M}} = \langle V, \Sigma, P, S \rangle$ aus einem DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$:

- $V := Q$
- $S := q_0$
- P besteht aus den folgenden Produktionsregeln:

$$\begin{aligned} q &\to \mathtt{a}q' && \text{falls } \delta(q, \mathtt{a}) = q' \\ q &\to \mathtt{a} && \text{falls } \delta(q, \mathtt{a}) \in F \\ q_0 &\to \varepsilon && \text{falls } q_0 \in F \end{aligned}$$

WAS BISHER GESCHAH ...

Ein **nichtdeterministischer endlicher Automat** (NFA) $\mathcal{M}$ ist ein Tupel $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F\rangle$ mit den folgenden Bestandteilen:

- Q: endliche Menge von Zuständen
- Σ: Alphabet
- δ: Übergangsfunktion, eine totale Funktion $Q \times \Sigma \to 2^Q$
- Q_0: Menge möglicher Startzustände $Q_0 \subseteq Q$
- F: Menge von Endzuständen $F \subseteq Q$

Alternative Definitionen:

- Übergangsrelation $\Delta \subseteq Q \times \Sigma \times Q$ mit

$$q' \in \delta(q, \sigma) \qquad \text{genau dann wenn} \qquad \langle q, \sigma, q'\rangle \in \Delta$$

- einzelner Startzustand q_0
- einzelner Endzustand q_f

Potenzmengenkonstruktion: NFA $\to$ DFA

Für einen NFA $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F\rangle$ definieren wir den Potenzmengen-DFA $\mathcal{M}_{\text{DFA}} = \langle Q_{\text{DFA}}, \Sigma, \delta_{\text{DFA}}, q_0, F_{\text{DFA}}\rangle$ wie folgt:

- $Q_{\text{DFA}} = 2^Q$
- $\delta_{\text{DFA}}(R, \mathtt{a}) = \bigcup_{q\in R} \delta(q, \mathtt{a})$
- $q_0 = Q_0$
- $F_{\text{DFA}} = \{R \in 2^Q \mid R \cap F \neq \emptyset\}$

Satz (Rabin/Scott): $\mathbf{L}(\mathcal{M}) = \mathbf{L}(\mathcal{M}_{\text{DFA}})$

NFA mit Wortübergängen $\mathcal{M} = \langle Q, \Sigma, \Delta, Q_0, F\rangle$ wobei

- Δ: Übergangsrelation, eine endliche Relation $\Delta \subseteq Q \times \Sigma^* \times Q$

Eliminieren von ε-Übergängen:

Sei $\xrightarrow{\varepsilon}^*$ der reflexive, transitive Abschluss von $\xrightarrow{\varepsilon}$, d.h. die Menge aller Zustandspaare $\langle q, q'\rangle \in Q^2$ für die es Übergänge $q = p_0 \xrightarrow{\varepsilon} p_1 \xrightarrow{\varepsilon} \ldots \xrightarrow{\varepsilon} p_n = q'$ gibt ($n \geq 0$).

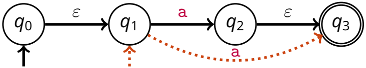
Für einen ε-NFA $\mathcal{M} = \langle Q, \Sigma, \Delta, Q_0, F\rangle$ definieren wir einen NFA $\mathcal{M}' = \langle Q, \Sigma, \delta, Q_0', F\rangle$ wobei

- $\delta(q, \mathtt{a}) = \{q' \mid q \xrightarrow{\mathtt{a}} r \xrightarrow{\varepsilon}^* q' \text{ für ein } r \in Q\}$
- $Q_0' = \{q \mid q_0 \xrightarrow{\varepsilon}^* q \text{ für ein } q_0 \in Q_0\}$

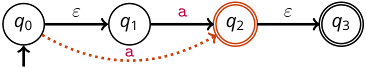
WAS BISHER GESCHAH ...

Eliminierung von ε-Übergängen:

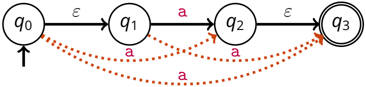
- Verlängerung nach rechts (wie in Definition):
  Achtung: Startzustände



- Verlängerung nach links:
  Achtung: Endzustände



- Verlängerung in beide Richtungen:



Übungsblatt 3

AUFGABE 1

Zeigen Sie konstruktiv, dass

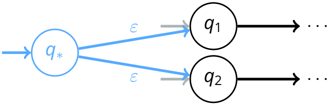
- für jeden NFA $\mathcal{M}$ mit mehreren Startzuständen ein äquivalenter NFA $\mathcal{M}'$ mit nur einem Startzustand existiert bzw.
- für jeden NFA $\mathcal{M}$ mit mehreren Finalzuständen ein äquivalenter NFA $\mathcal{M}'$ mit nur einem Finalzustand existiert. Gilt die letzte Aussage auch für DFAs?

AUFGABE 1 — TEIL (A)

Sei $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F \rangle$ ein NFA mit mehreren Startzuständen (d.h. $|Q_0| \geq 2$).

Wir konstruieren daraus einen NFA $\mathcal{M}'$, bei dem ein neuer, unbenutzter Startzustand vorgeschalten wird und mit ε-Übergängen in die Startzustände von $\mathcal{M}$ übergeht. Es ist also $\mathcal{M}' := \langle Q \cup \{q_*\}, \Sigma, \Delta, \{q_*\}, F \rangle$ mit $q_* \notin Q$ und die Übergangsrelation ist gegeben durch

$$\begin{aligned}\Delta :=& \{\text{originale Übergänge}\} \cup \{\text{neue } \varepsilon\text{-Übergänge}\} \\ :=& \{\langle q, \mathtt{a}, q' \rangle : q' \in \delta(q, \mathtt{a}) \text{ für } q, q' \in Q \text{ und } \mathtt{a} \in \Sigma\} \\ & \cup \{\langle q_*, \varepsilon, q_0 \rangle : q_0 \in Q_0\}\end{aligned}$$



Dieser ε-NFA kann nun in einen äquivalenten NFA umgewandelt werden (durch „Verlängerung nach links").

Es bleibt nun noch zu zeigen, dass der konstruierte Automat $\mathcal{M}'$ korrekt ist, d.h. $\mathbf{L}(\mathcal{M}) = \mathbf{L}(\mathcal{M}')$ gilt.

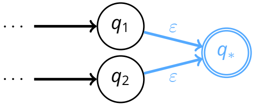
$$
\begin{aligned}
& w \in \mathbf{L}(\mathcal{M}) \\
\Leftrightarrow \quad & w = \mathtt{u_1} \ldots \mathtt{u_n} \text{ mit } \mathtt{u_1}, \ldots, \mathtt{u_n} \in \Sigma \\
\Leftrightarrow \quad & \text{ex. akzeptierender Lauf } q_1, \ldots q_n \text{ in } \mathcal{M} \text{ (d.h. } q_n \in F) \\
\Leftrightarrow \quad & \text{ex. akzeptierender Lauf } q_0, q_1, \ldots, q_n \text{ in } \mathcal{M}' \text{ (d.h. } q_n \in F) \\
\Leftrightarrow \quad & w \in \mathbf{L}(\mathcal{M}')
\end{aligned}
$$

AUFGABE 1 — TEIL (B)

Sei $\mathcal{M} = \langle Q, \Sigma, \delta, Q_0, F\rangle$ ein NFA mit mehreren Finalzuständen (d.h. $|F| \geq 2$).

Analog zu Teil (a) führen wir einen neuen Finalzustand ein und erreichen diesen über ε-Übergänge von den originalen Finalzuständen. Es ist also $\mathcal{M}' := \langle Q \cup \{q_*\}, \Sigma, \Delta, Q_0, \{q_*\}\rangle$ mit $q_* \notin Q$ und die Übergangsrelation ist gegeben durch

$$\begin{aligned}\Delta :=& \{\text{originale Übergänge}\} \cup \{\text{neue } \varepsilon\text{-Übergänge}\}\\ :=& \{\langle q, \mathtt{a}, q'\rangle : q' \in \delta(q, \mathtt{a}) \text{ für } q, q' \in Q \text{ und } \mathtt{a} \in \Sigma\}\\ &\cup \{\langle q_F, \varepsilon, q_*\rangle : q_F \in F\}\end{aligned}$$



Verlängerung nach rechts eliminiert nun noch die ε-Übergänge und wir erhalten einen NFA.

AUFGABE 2

Sei L eine reguläre Sprache über einem mindestens zweielementigen Alphabet Σ. Zeigen Sie, dass die folgenden Sprachen regulär sind.

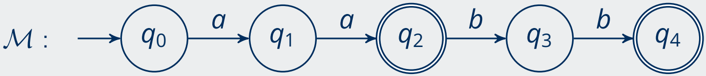
(a) $L_1 = \{x \in L : \text{es gibt kein } y \in \Sigma^+, \text{so dass } xy \in L\}$

(b) $L_2 = \{x \in L : \text{kein echtes Präfix von } x \text{ liegt in } L\}$

AUFGABE 2 — TEIL (A)

Gegeben sei eine reguläre Sprache L — dazu gehört ein DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F\rangle$ mit $L = L(\mathcal{M})$.

Beispiel: $L = \{aa, aabb\}$



Es ist $L_1 = \{aabb\}$, denn für $x = aa$ existiert Verlängerung $y = bb$, sodass $xy = aabb \in L$.

Idee: Unterbinde zu frühe Akzeptanz (hier in q_2)

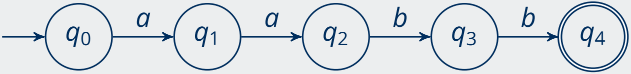
$$F' := \{q \in F : \nexists y \in \Sigma^+ \text{ mit } \delta(q, y) \in F\}$$

F' beschreibt die Finalzustände $q \in F$, die nicht mit einem Wortübergang in einen weiteren Finalzustand verlängert werden können.

Wir definieren den Automaten $\mathcal{M}' := \langle Q, \Sigma, \delta, q_0, F' \rangle$ mit dem bereits angegebenen $F' := \{q \in F : \nexists y \in \Sigma^+ \text{ mit } \delta(q, y) \in F\}$.

Beispiel: $F' = \{q_4\}$

$\mathcal{M}'$:



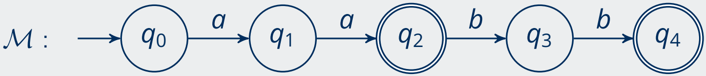
Korrektheit: zu zeigen ist $L(\mathcal{M}') = L_1$

$$
\begin{array}{lll}
 & x \in L(\mathcal{M}') & \\
\Leftrightarrow & \delta(q_0, x) \in F' & \text{(Def. Akzeptanz)} \\
\Leftrightarrow & \delta(q_0, x) \in F \text{ und ex. kein } y \in \Sigma^+ : \delta(q_0, xy) \in F & \text{(Def. } F'\text{)} \\
\Leftrightarrow & x \in L_1 & \text{(Def. von } L_1\text{)}
\end{array}
$$

AUFGABE 2 — TEIL (B)

Gegeben sei eine reguläre Sprache L — dazu gehört ein DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit $L = L(\mathcal{M})$.

Beispiel: $L = \{aa, aabb\}$



Es ist $L_2 = \{aa\}$, denn für *aabb* existiert der Präfix $aa \in L$.

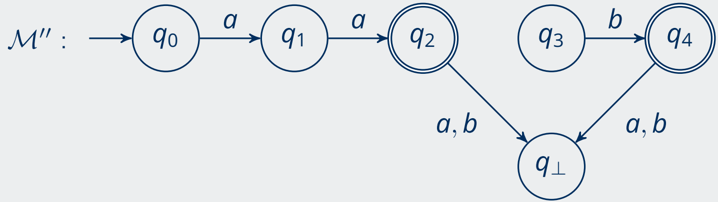
Idee: Verhindere weitere Übergänge von Finalzuständen

Sei $q_\perp \notin Q$ ein neuer, unbenutzter Fangzustand und $a \in \Sigma$ ein beliebiges Symbol.

$$\delta''(q, a) := \begin{cases} q_\perp & \text{falls } q \in F \\ \delta(q, a) & \text{sonst} \end{cases}$$

Wir definieren den Automaten $\mathcal{M}'' := \langle Q \cup \{q_\perp\}, \Sigma, \delta'', q_0, F \rangle$ mit dem bereits angegebenen δ''.

Beispiel: $q_2 \xrightarrow{a,b} q_\perp$, $q_4 \xrightarrow{a,b} q_\perp$



Korrektheit: zu zeigen ist $L(\mathcal{M}'') = L_2$

$$
\begin{aligned}
x \in L(\mathcal{M}'') \quad &\Leftrightarrow \quad \delta''(q_0, x) \in F && \text{(Def. Akzeptanz)} \\
&\Leftrightarrow \quad \delta(q_0, x) \in F \text{ und ex. kein echtes Präfix} \\
&\qquad\quad \overline{x} \text{ von } x \text{ mit } \delta(q_0, \overline{x}) \in F && \text{(Def. } \delta''\text{)} \\
&\Leftrightarrow \quad x \in L_2 && \text{(Def. von } L_2\text{)}
\end{aligned}
$$

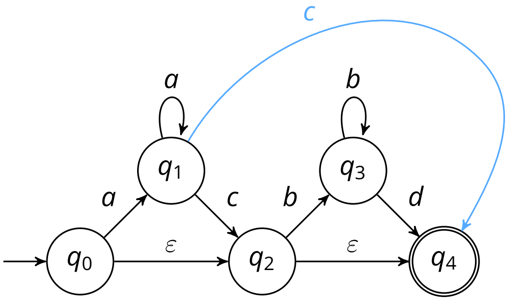
AUFGABE 3

Konstruieren Sie zu dem grafisch angegebenen ε-NFA $\mathcal{M} = (Q, \Sigma, \delta, \{q_0\}, F)$ einen äquivalenten NFA $\mathcal{M}'$. Beschreiben Sie die Komponenten beider Automaten.

Schritt 1: Startzustände anpassen

$$Q_0' = \left\{ q \mid q_0 \xrightarrow{\varepsilon}{}^{*} q \text{ für ein } q_0 \in Q_0 \right\}$$

Schritt 2: Verlängerung nach rechts



REGULÄRE AUSDRÜCKE

Beobachtungen:

- Jede endliche Sprache ist regulär.
- Reguläre Sprachen sind unter $\cup$, $\cdot$, * abgeschlossen.
- Jede reguläre Sprache lässt mittels $\cup$, $\cdot$, * aus endlichen Sprachen konstruieren.

Die Menge der **regulärer Ausdrücke** über einem Alphabet Σ ist induktiv wie folgt definiert:

- $\emptyset$ ist ein regulärer Ausdruck
- ε ist ein regulärer Ausdruck
- $\mathtt{a}$ ist ein regulärer Ausdruck für jedes $\mathtt{a} \in \Sigma$
- Wenn α und β reguläre Ausdrücke sind, dann sind auch $(\alpha\beta)$, $(\alpha \mid \beta)$ und $(\alpha)^*$ reguläre Ausdrücke

Die **Sprache eines regulären Ausdrucks** α wird mit $\mathbf{L}(\alpha)$ bezeichnet und rekursiv definiert:

- $\mathbf{L}(\emptyset) = \emptyset$
- $\mathbf{L}(\varepsilon) = \{\varepsilon\}$
- $\mathbf{L}(\mathtt{a}) = \{\mathtt{a}\}$ für jedes $\mathtt{a} \in \Sigma$
- $\mathbf{L}((\alpha\beta)) = \mathbf{L}(\alpha) \circ \mathbf{L}(\beta)$
- $\mathbf{L}((\alpha \mid \beta)) = \mathbf{L}(\alpha) \cup \mathbf{L}(\beta)$
- $\mathbf{L}((\alpha)^*) = \mathbf{L}(\alpha)^*$

AUFGABE 4

Welche Sprachen $\mathbf{L}(r_i)$ werden durch folgende reguläre Ausdrücke r_i beschrieben?

(a) $r_1 = bb^* \mid (bb)^*a$

$$\mathbf{L}(r_1) = \{b^n : n \geq 1\} \cup \left\{b^{2m}a : m \geq 0\right\}$$

(b) $r_2 = a^*b(aa^*b)^*b(a \mid b)^*$

$$\begin{aligned} L_0 &:= \{a^n b : n \geq 0\} \\ L_1 &:= \{a^n b : n \geq 1\} \qquad \Rightarrow \quad \mathbf{L}(r_2) = L_0 \cdot L_1^* \cdot \{b\} \cdot L_n \\ L_n &:= \{a, b\}^* \end{aligned}$$

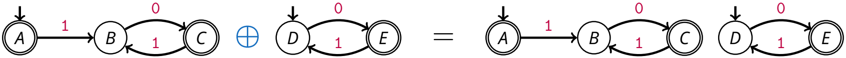
(c) $r_3 = a^* \mid a^*(b \mid bb)(aa^*(b \mid bb))^*a^*$

$$\begin{aligned} L_0 &:= \{a^n b^m : n \geq 0, m \in \{1, 2\}\} \\ L_1 &:= \{a^n b : n \geq 1, m \in \{1, 2\}\} \end{aligned}$$

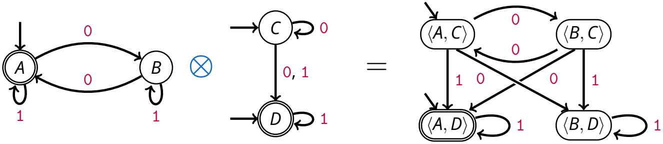
$$\Rightarrow \quad \mathbf{L}(r_3) = \{a\}^* \cup (L_0 \cdot L_1^* \cdot \{a\}^*)$$

OPERATIONEN AUF AUTOMATEN

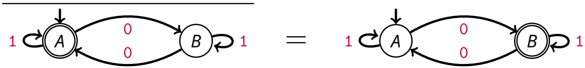
$\cup$



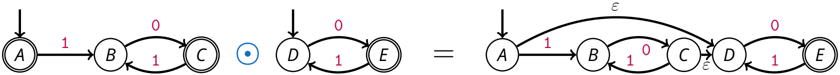
$\cap$



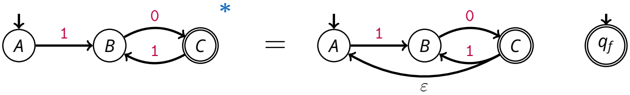
$-$



$\circ$

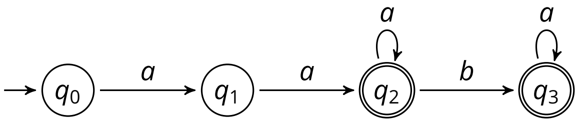


$*$

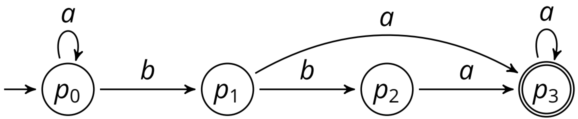


AUFGABE 5 — TEIL (A)

δ_1:



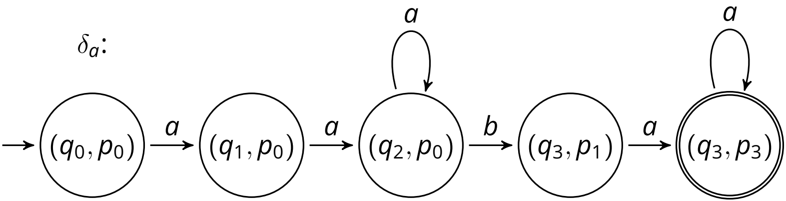
δ_2:



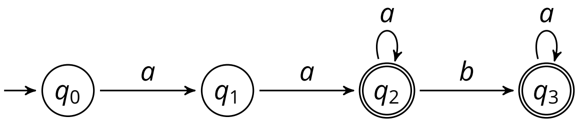
(a) Konstruieren Sie einen ε-freien NFA $\mathcal{M}_a$ mit $L(\mathcal{M}_a) = L(\mathcal{M}_1) \cap L(\mathcal{M}_2)$. Dabei dürfen Sie sich auf die vom Startzustand erreichbaren Zustände beschränken.

$$\begin{aligned} \mathcal{M}_a &:= \langle Q, \Sigma, \delta_a, \{(q_0, p_0)\}, F \rangle \\ \text{mit} \quad Q &:= \{(q_0, p_0), (q_1, p_0), (q_2, p_0), (q_3, p_1), (q_3, p_3)\} \\ \Sigma &:= \{a, b\} \\ F &:= \{(q_3, p_3)\} \end{aligned}$$

δ_a:



δ_1:



(b) Konstruieren Sie einen ε-freien NFA $\mathcal{M}_b$ mit $L(\mathcal{M}_b) = L(\mathcal{M}_1)^*$.

$$\mathcal{M}_b := (\mathcal{M}_1)^* = \langle Q_b, \Sigma, \delta_b, Q_0, F_b \rangle$$

$$\text{mit} \quad Q_b := \{q_0, q_1, q_2, q_3, q_\varepsilon\}$$

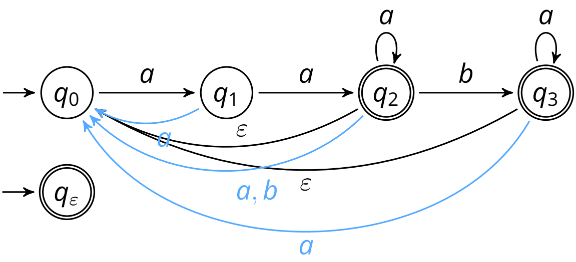
$$\Sigma := \{a, b\}$$

$$Q_0 := \{q_0, q_\varepsilon\}$$

$$F_b := \{q_3, q_\varepsilon\}$$

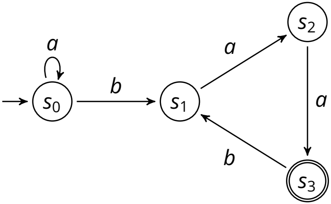
Eliminierung der ε-Übergänge durch Verlängerung nach rechts

${\delta_b}'$:

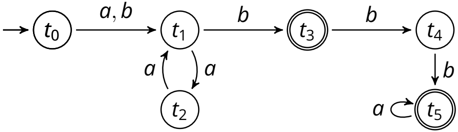


AUFGABE 5 — TEIL (C)

δ_3:



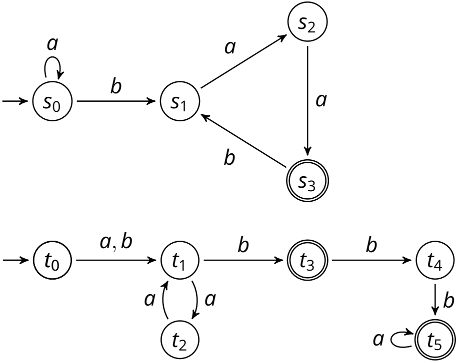
δ_4:



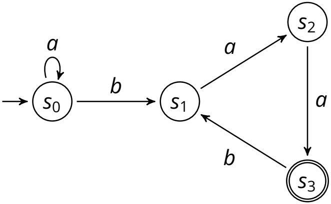
(c) Konstruieren Sie einen ε-freien NFA $\mathcal{M}_c$ mit $L(\mathcal{M}_c) = L(\mathcal{M}_3) \cup L(\mathcal{M}_4)$.

$\mathcal{M}_c = \mathcal{M}_3 \oplus \mathcal{M}_4$, d.h. beide Automaten „nebeneinander schreiben“

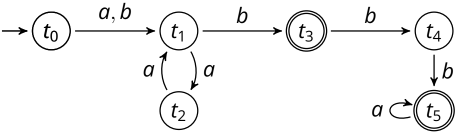
δ_c:



δ_3:



δ_4:



(d) Konstruieren Sie einen ε-freien NFA $\mathcal{M}_d$ mit $L(\mathcal{M}_d) = L(\mathcal{M}_3) \circ L(\mathcal{M}_4)$.

$$\mathcal{M}_d := \mathcal{M}_3 \odot \mathcal{M}_4$$

Eliminierung der ε-Übergänge durch Verlängerung nach rechts

$\delta_d{}'$: