

Lösungsvorschlag zur Übung 4 zur Vorlesung
Theoretische Informatik für Studierende der Medieninformatik

Wenn Sie Automaten angeben, tun Sie dies immer in Form eines Zustandsgraphen. Andere Formen der Darstellung (z.B. als Liste von Übergängen) werden nicht gewertet, da sie sehr viel aufwändiger zu korrigieren sind. Vergessen Sie nicht, im Zustandsgraph Start- und Endzustände zu markieren.

Reguläre Ausdrücke sind entsprechend Definition 4.7.1 im Vorlesungsskript anzugeben.

TIMI4-1 Pumping-Lemma für reguläre Sprachen (2 Punkte)

Zeigen Sie mit dem Pumping-Lemma für reguläre Sprachen, dass die folgenden Sprachen nicht regulär sind.

- a) $L_1 = \{a^i b^j c^k d^i \mid i, j, k \in \mathbb{N}_{>0}\}$ über dem Alphabet $\Sigma_1 = \{a, b, c, d\}$.

LÖSUNGSVORSCHLAG:

Beweis mit dem Pumping-Lemma.

Sei $n \in \mathbb{N}_{>0}$ beliebig.

Wir wählen $z \in L_1$ als $z = a^n b c d^n$ mit $|z| \geq n$.

Sei $z = uvw$ eine beliebige Zerlegung von z , sodass $|uv| \leq n$, $|v| \geq 1$ und $uv^i w \in L_1$ für jedes $i \in \mathbb{N}$. Da $|uv| \leq n$ ist, ist $v = a^k$ für ein $k \in \mathbb{N}_{>0}$.

Wir wählen $i = 0$. Das Wort $uv^i w$ enthält weniger a 's als d 's und somit ist $uv^i w \notin L_1$. Widerspruch.

- b) $L_2 = L(G_2)$, wobei G_2 eine kontextfreie Grammatik ist mit

$$G_2 = (\{S, A, B\}, \{ (,), [,] \}, P, S)$$

$$P = \{S \rightarrow (S), S \rightarrow [S], S \rightarrow A, S \rightarrow B, A \rightarrow (), A \rightarrow [], B \rightarrow S, B \rightarrow BB\}$$

L_2 ist die Sprache der zueinander passenden eckigen und runden Klammern, d.h. es sind z.B. $([])[] \in L_2$ und $()() \in L_2$, aber $[] \notin L_2$ und $) \notin L_2$.

LÖSUNGSVORSCHLAG:

Beweis mit dem Pumping-Lemma.

Sei $n \in \mathbb{N}_{>0}$ beliebig.

Wir wählen $z \in L_2$ als $z = [^n]^n$ mit $|z| \geq n$.

Sei $z = uvw$ eine beliebige Zerlegung von z , sodass $|uv| \leq n$, $|v| \geq 1$ und $uv^i w \in L_2$ für jedes $i \in \mathbb{N}$. Da $|uv| \leq n$ ist, ist $v = [^k$ für ein $k \in \mathbb{N}_{>0}$. Wir wählen $i = 0$. Das Wort $uv^i w$ hat weniger öffnende als schließende eckige Klammern und somit ist $uv^i w \notin L_2$. Widerspruch.

TIMI4-2 Reguläre Ausdrücke und Abschlusseigenschaften

(0 Punkte)

- a) Betrachten Sie den regulären Ausdruck $\alpha = (a|b)^*(ab|ba)(a|b)^*$.
- i) Geben Sie einen NFA ohne ε -Übergänge an, der $L(\alpha)$ erkennt. Sie können die Algorithmen aus der Vorlesung zur Konstruktion eines NFA aus einem regulären Ausdruck und zur Elimination von ε -Übergängen verwenden, müssen aber nicht.

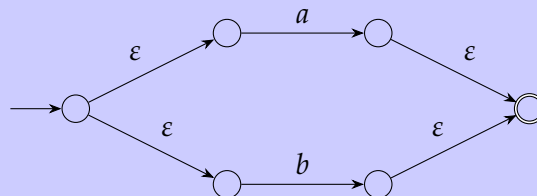
LÖSUNGSVORSCHLAG:

Mit dem Algorithmus aus der Vorlesung ergeben sich folgende NFAs mit ε -Übergängen:

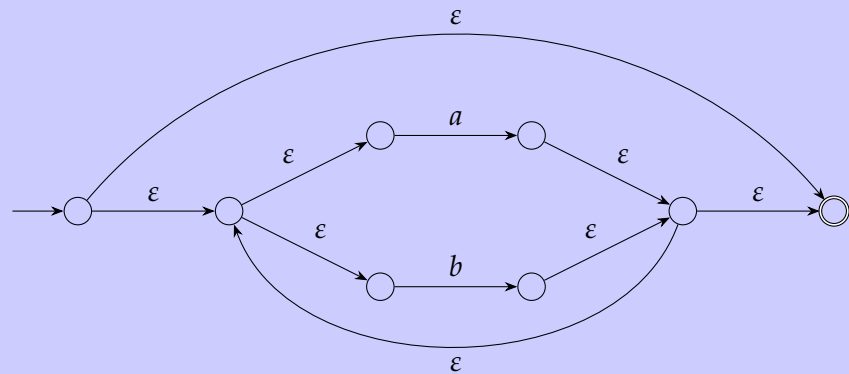
- M_a (und analog M_b):



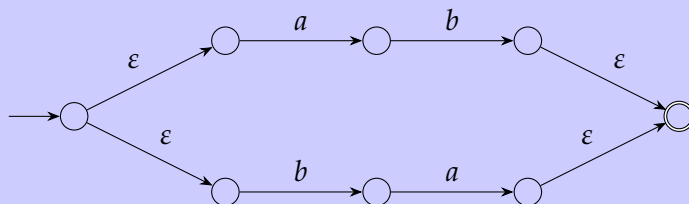
- $M_{a|b}$:



- $M_{(a|b)^*}$:



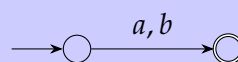
- $M_{ab|ba}$:



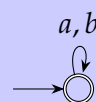
- $M_{(a|b)^*(ab|ba)(a|b)^*}$: $M_{(a|b)^*}$, $M_{ab|ba}$ und $M_{(a|b)^*}$, verbunden mit ϵ -Übergängen.

Sinnvollerweise eliminiert man bereits während der obigen Konstruktion ϵ -Übergänge und entfernt offensichtlich redundante Zustände. So erhält man kompaktere Automaten:

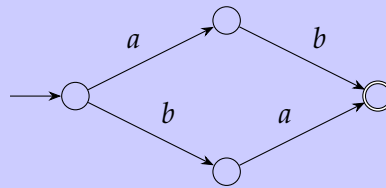
- $M_{a|b}$:



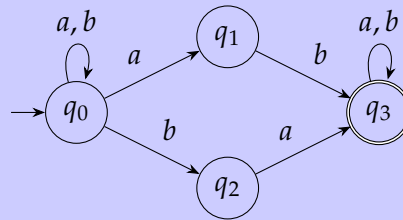
- $M_{(a|b)^*}$:



- $M_{ab|ba}$:



Insgesamt ergibt sich M_α :



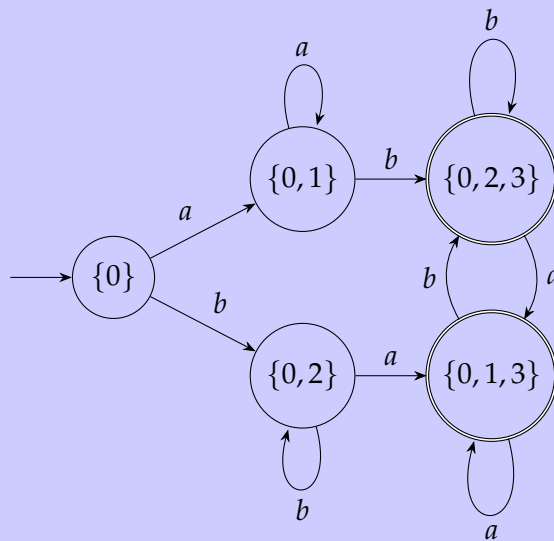
- ii) Geben Sie einen DFA an, der $L(\alpha)$ erkennt. Sie können die Potenzmengenkonstruktion verwenden, müssen aber nicht.

LÖSUNGSVORSCHLAG:

Potenzmengenkonstruktion:

Start	→	Ziel
$\{0\}$	a	$\{0, 1\}$
$\{0\}$	b	$\{0, 2\}$
$\{0, 1\}$	a	$\{0, 1\}$
$\{0, 1\}$	b	$\{0, 2, 3\}$
$\{0, 2\}$	a	$\{0, 1, 3\}$
$\{0, 2\}$	b	$\{0, 2\}$
$\{0, 1, 3\}$	a	$\{0, 1, 3\}$
$\{0, 1, 3\}$	b	$\{0, 2, 3\}$
$\{0, 2, 3\}$	a	$\{0, 1, 3\}$
$\{0, 2, 3\}$	b	$\{0, 2, 3\}$

Der resultierende DFA M'_α :



Die zwei Endzustände könnte man zusammenfassen, aber es war kein Minimalautomat gefragt.

b) Geben Sie reguläre Ausdrücke an, die die folgenden Sprachen erkennen.

- i) Die Sprache L_3 der Wörter über dem Alphabet $\Sigma_1 = \{a, b, c\}$, die mit a oder b anfangen und mindestens ein c enthalten.

LÖSUNGSVORSCHLAG:

$(a|b)(a|b|c)^*c(a|b|c)^*$ oder, äquivalent, $(a|b)(a|b)^*c(a|b|c)^*$.

- ii) Die Sprache L_4 der Wörter über dem Alphabet $\Sigma_2 = \{a, b\}$, die keine zwei a 's hintereinander enthalten.

LÖSUNGSVORSCHLAG:

$(b|ab)^*(a|\varepsilon)$

- c) Zeigen Sie mithilfe der Abschlusseigenschaften regulärer Sprachen, dass die Sprache $L_5 = \{a^i w d^{i+1} \mid i \in \mathbb{N}, w \in \{b, c\}^*\}$ über dem Alphabet $\Sigma_3 = \{a, b, c, d\}$ nicht regulär ist. Sie dürfen annehmen, dass die Sprache L_1 aus Aufgabe TIMI4-1a) nicht regulär ist.

LÖSUNGSVORSCHLAG:

Widerspruchsbeweis: Nimm an, dass L_5 regulär ist.

Wir definieren die Sprache

$$\begin{aligned} L'_5 &= \{a\} \cdot L_5 = \{a^{i+1}wd^{i+1} \mid i \in \mathbb{N}, w \in \{b,c\}^*\} \\ &= \{a^iwd^i \mid i \in \mathbb{N}_{>0}, w \in \{b,c\}^*\} \end{aligned}$$

L'_5 ist regulär, denn $\{a\}$ ist eine reguläre Sprache und die regulären Sprachen sind unter Produkt abgeschlossen.

Nun definieren wir

$$\begin{aligned} L''_5 &= L'_5 \cap L(aa^*bb^*cc^*dd^*) \\ &= L'_5 \cap \{a^ib^jc^kd^l \mid i,j,k,l \in \mathbb{N}_{>0}\} \\ &= \{a^ib^jc^kd^i \mid i,j,k \in \mathbb{N}_{>0}\} \end{aligned}$$

L''_5 ist regulär, denn $L(aa^*bb^*cc^*dd^*)$ ist die Sprache eines regulären Ausdrucks (und damit regulär) und die regulären Sprachen sind unter Schnitt abgeschlossen.

Allerdings wissen wir aus Aufgabe TIMI4-1a), dass $L''_5 = L_1$ nicht regulär ist. Die Annahme, dass L_5 regulär sei, führt somit zu einem Widerspruch.