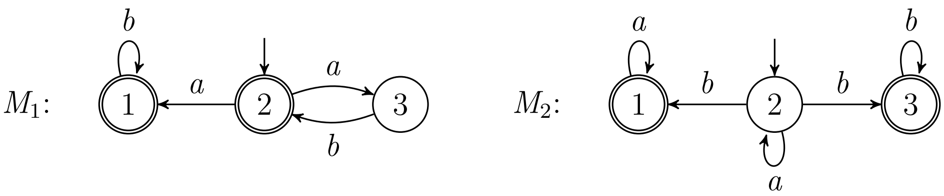
Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

Übungsblatt 9

Aufgabe 1.
Gegeben seien die folgenden NFAs M_1, M_2 (siehe Blatt 5, Aufgabe 2).

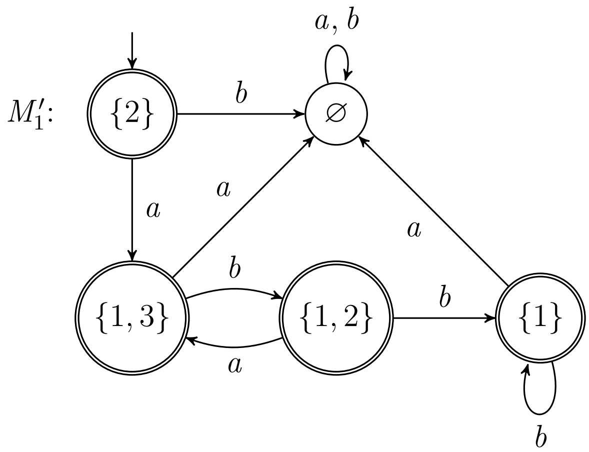


Lösen Sie mit dem Vorgehen aus der Vorlesung das Inklusionsproblem:

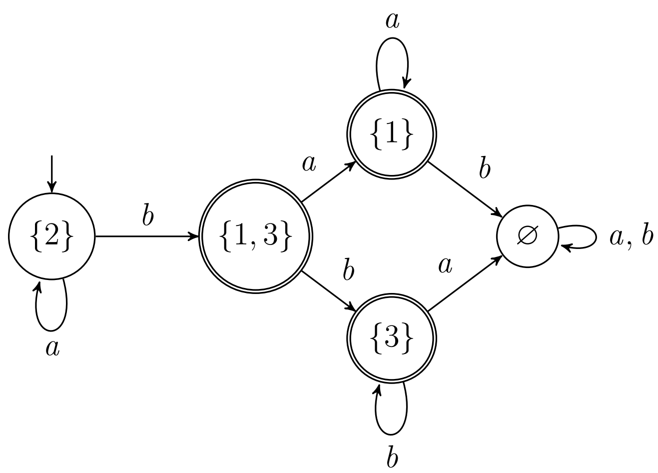
$$T(M_1) \subseteq T(M_2)$$

Lösung zu Aufgabe 1. Da $\varepsilon \in T(M_1)$ aber $\varepsilon \notin T(M_2)$, wissen wir, dass $T(M_1) \subseteq T(M_2)$ nicht gelten kann. $T(M_1) \subseteq T(M_2)$ ist äquivalent zu $T(M_1) \cap \overline{T(M_2)} = \varnothing$. Nun wollen wir mit dem Verfahren aus der Vorlesung zeigen, dass dies nicht gilt.

Zunächst konstruieren wir mittels Potenzmengenkonstruktion die DFAs für M_1 und M_2. Dies haben wir schon auf Blatt 5 erledigt.

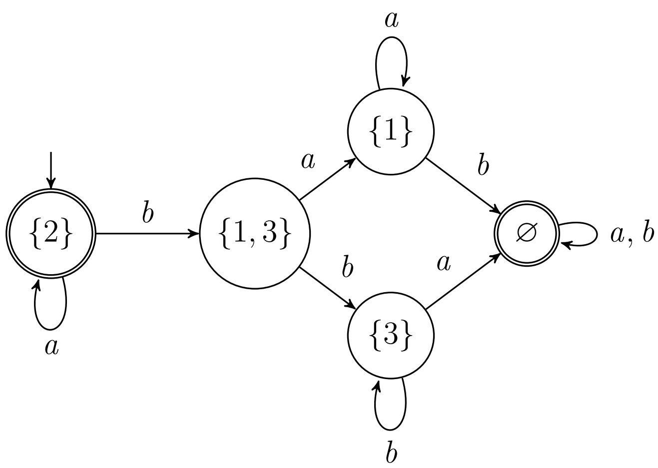


M_2':



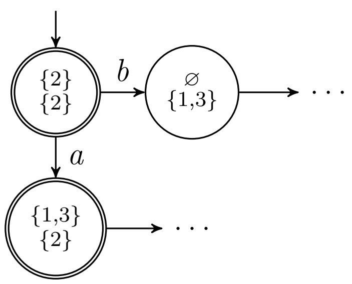
Dann erzeugen wir $\overline{M_2'}$, so dass $T(\overline{M_2'}) = \overline{T(M_2')} = \overline{T(M_2)}$, indem wir Endzustande und Nicht-Endzustände im DFA M_2' vertauschen.
<u>Beachte</u>: NFAs auf diese Art zu „komplementieren“ funktioniert nicht (Übung: Beispiel überlegen).

$\overline{M_2'}$:



Nun bestimmen wir einen DFA M für $T(M_1) \cap \overline{T(M_2)} = T(M_1') \cap T(\overline{M_2'})$ (Kreuzproduktautomat):

M:



Anmerkung: Aus Platzgründen sind die Zustandspaare untereinander geschrieben - oben jeweils der Zustand in M_1', darunter der Zustand in $\overline{M_2'}$.

Wir brauchen nicht den kompletten Automaten zu konstruieren, da wir schon nach wenigen Schritten sehen, dass einen Pfad im Automaten M gibt, der vom Anfangszustand in einen Endzustand führt. Zum Beispiel findet man so, dass $\varepsilon, a \in T(M) = T(M_1) \cap \overline{T(M_2)}$. Somit gilt $T(M_1) \subseteq T(M_2)$ nicht.

Es würde sogar reichen sich zu überlegen, dass die Startzustände in M_1' (bzw. M_1) und $\overline{M_2'}$ beides auch Endzustände sind. Denn dann weiß man bereits, dass ε durch M erkannt wird.

Aufgabe 2. Sei L die Sprache aus Aufgabe 1 von Blatt 8, also

$$L = \{a^m b^n c^n \mid m, n \geq 1\} \cup \{b^m c^n \mid m, n \geq 0\}.$$

Zeigen Sie, dass diese kontextfrei ist, indem Sie eine kontextfreie Grammatik angeben, die die Sprache erzeugt.

Lösung zu Aufgabe 2. Eine Grammatik G mit $L(G) = L$ ist z.B. $G = (V, \Sigma, P, S)$ mit

- $V = \{S, A, B, C, D\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to aA \mid B,\ A \to aA \mid D,\ D \to bDc \mid bc,$
  $B \to bB \mid C,\ C \to cC \mid \varepsilon\}$

Alternativ mit Beachtung der ε-Sonderregelung: $G = (V, \Sigma, P, S)$ mit

- $V = \{S, A, B, C, D\}$
- $\Sigma = \{a, b, c\}$
- $P = \{S \to aA \mid B \mid \varepsilon,\ A \to aA \mid D,\ D \to bDc \mid bc,$
  $B \to bB \mid b \mid C,\ C \to cC \mid c\}$

Aufgabe 3. Sei $\Sigma = \{a, b\}$. Geben Sie eine kontextfreie Grammatik an, die die Sprache

$$L = \Sigma^* \setminus \{ww \mid w \in \Sigma^*\} = \overline{\{ww \mid w \in \Sigma^*\}}$$

erzeugt.
Hinweis: Für jedes $w \in L$ gerader Länge gibt es Wörter $x, y, u, v \in \Sigma^*$ so, dass $w = xayubv$ (oder $w = xbyuav$) ist und $|x| = |u|$ bzw. $|y| = |v|$ gilt.

Lösung zu Aufgabe 3. Wir unterscheiden zwei Fälle, Wörter gerader Länge und Wörter ungerader Länge.

Wörter gerader Länge

Der Hinweis ist so zu verstehen, dass sich jedes Wort $w \in L$ mit Länge $2n$ in $w_1 w_2$ unterteilen lässt, wobei w_1 und w_2 Wörter der Länge n sind, die sich an mindestens einer Stelle unterscheiden.

Da die Wörter x, y, u, v beliebig sein können, solange $|x| = |u|$ und $|y| = |v|$, ist die Menge der Wörter der Form $w = xayubv$ gleich der Menge der Wörter der Form $xay'u'bv$ (analog für den Fall $w = xbyuav$), so dass $w = w_1' w_2'$ und wobei $w_1' = xay'$ und $w_2' = u'bv$ (bzw. $w_1' = xby'$ und $w_2' = u'av$) mit $|x| = |y'|$ und $|u'| = |v|$. Einfach gesagt: In der Mitte steht nach wie vor das gleiche Wort $yu = y'u'$, allerdings teilen wir dieses Wort nun so auf, dass y' die Länge von u und x hat und u' die Länge von y und v hat.

Wörter der Form xay' mit $|x| = |y'|$ können wir mit den Produktionen $\{A \to XAX \mid a,\ X \to a \mid b\}$ aus A erzeugen, Wörter der Form $u'bv$ mit $|u'| = |v|$ durch eine weitere Produktion $B \to XBX \mid b$ aus B.

Die Wörter $w \in L$ mit gerader Länge erhalten wir also mit der Grammatik $G' = (V, \Sigma, P', S)$ mit

- $V = \{S, A, B, X\}$
- $\Sigma = \{a, b\}$
- $P' = \{S \to AB \mid BA,\ A \to XAX \mid a,\ B \to XBX \mid b,\ X \to a \mid b\}$

Wörter ungerader Länge

Für alle $w \in \Sigma^*$ mit ungerader Länge gilt $w \in L$. Um diese Wörter zu erzeugen, können wir die Nicht-Terminale A und B „wiederverwenden", da A alle Wörter ungerader Länge mit einem a in dem Mitte erzeugt, B alle Wörter ungerader Länge mit einem b in der Mitte.

Grammatik

Eine kontextfreie Grammatik, die L erzeugt, ist also $G = (V, \Sigma, P, S)$ mit

- $V = \{S, A, B, X\}$
- $\Sigma = \{a, b\}$
- $P = \{S \to A \mid B \mid AB \mid BA, A \to XAX \mid a, B \to XBX \mid b, X \to a \mid b\}$

Aufgabe 4. Beweisen oder widerlegen Sie:

(a) Sei L kontextfrei und nicht regulär. Dann ist auch L^* kontextfrei und nicht regulär.

(b) Sei L kontextfrei und sei L' nicht regulär. Dann ist $L'' = L \cup L'$ nicht regulär.

Lösung zu Aufgabe 4. (a) Die Aussage ist falsch: Sei L die Sprache aus Aufgabe 3, also

$$L = \overline{\{ww \mid w \in \{a, b\}^*\}}.$$

Dann ist L kontextfrei, denn in Aufgabe 3 haben wir gezeigt, dass wir L mit einer kontextfreien Grammatik erzeugen können. Außerdem ist L nicht regulär: Auf Blatt 7, Aufgabe 3b haben wir gezeigt, dass das Komplement von L, die Sprache $\{ww \mid w \in \{a, b\}^*\}$, nicht regulär ist. Da reguläre Sprachen unter Komplementbildung abgeschlossen sind, kann L also auch nicht regulär sein. Nun gilt aber $a, b \in$ Ł. Daher ist $L^* = \{a, b\}^*$, also ist L^* regulär.

(b) Die Aussage ist ebenfalls falsch. Wir wählen $L = \overline{\{ww \mid w \in \{a, b\}^*\}}$ und $L' = \Sigma^* \setminus L = \{ww \mid w \in \{a, b\}^*\}$. Mit den Überlegungen aus Aufgabenteil (a) ist L kontextfrei und L' nicht regulär. Es gilt aber $L'' = \{a, b\}^*$, also ist L'' regulär.

Dies lässt sich verallgemeinern: Für jede kontextfreie, nicht reguläre Sprache kann ihr Komplement nicht regulär sein (wegen Abschluss regulärer Sprachen unter Komplementbildung). Die Vereinigung beider Sprachen ist dann stets Σ^*.

Aufgabe 5. Sei $G = (\{S, A, B\}, \{a, b\}, P, S)$ eine kontextfreie Grammatik, wobei P gegeben ist durch

$$\begin{aligned} S &\to ASB \mid ab \\ A &\to aAS \mid a \\ B &\to SbS \mid A \mid bb. \end{aligned}$$

Geben Sie eine Grammatik G' in CNF so an, dass $L(G') = L(G)$.

Lösung zu Aufgabe 5. Wir verwenden das Verfahren zur Umwandlung einer kontextfreien Grammatik in CNF wie ab Folie 196 im Skript beschrieben.

Schritt 1

G enthält keine Produktion der Form $S \to \varepsilon$.

Schritt 2

Einführen neuer Variablen A_a mit Produktionen $A_a \to a$ für jedes Terminalsymbol $a \in \Sigma$ und Ersetzen aller Vorkommen von a in einer rechten Seite, die nicht bereits nur aus einem Alphabetsymbol auf der rechten Seite bestehen. Änderungen an den Produktionen sind **fettgedruckt** markiert.

- $S \to ASB \mid \boldsymbol{A_a A_b}$
- $A \to \boldsymbol{A_a} AS \mid a$
- $B \to S\boldsymbol{A_b}S \mid A \mid \boldsymbol{A_b A_b}$
- $\boldsymbol{A_a \to a}$
- $\boldsymbol{A_b \to b}$

Schritt 3

Elimination von Kettenregeln der Form $X \to Y$ durch Ersetzen von Y durch die rechten Seiten der Produktion von Y.

- $S \to ASB \mid A_a A_b$
- $A \to A_a AS \mid a$
- $B \to SA_bS \mid \boldsymbol{A_a AS} \mid \boldsymbol{a} \mid A_b A_b$
- $A_a \to a$
- $A_b \to b$

Bemerkung: Falls es Zyklen gibt (wie $A \to B$ und $B \to A$), wird man irgendwann bei $A \to A$ angelangen. Diese Produktion muss man einfach weglassen.

Schritt 4

Elimination von Produktionen der Form $X \to X_1 \ldots X_n$ mit $n \geq 3$ durch Einführen neuer Nicht-Terminale. Beispielsweise schreiben wir $S \to ASB$ um in $S \to AB_1$ und fügen ein neues Nicht-Terminal B_1 hinzu mit einer Produktion $B_1 \to SB$.

- $S \to A\boldsymbol{B_1} \mid A_a A_b$
- $A \to A_a \boldsymbol{B_2} \mid a$
- $B \to S\boldsymbol{B_3} \mid A_a \boldsymbol{B_2} \mid a \mid A_b A_b$
- $A_a \to a$
- $A_b \to b$
- $\boldsymbol{B_1 \to SB}$
- $\boldsymbol{B_2 \to AS}$
- $\boldsymbol{B_3 \to A_b S}$