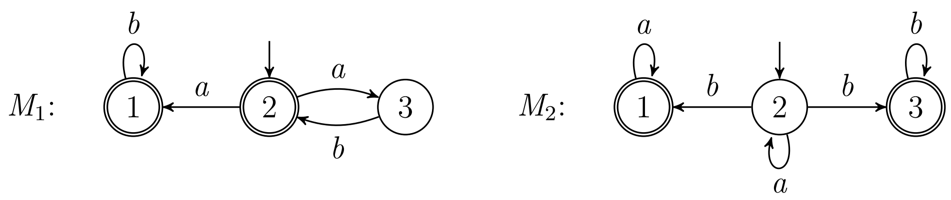
Übungsblatt 7

Aufgabe 1.
Gegeben sind die folgenden NFAs M_1, M_2.

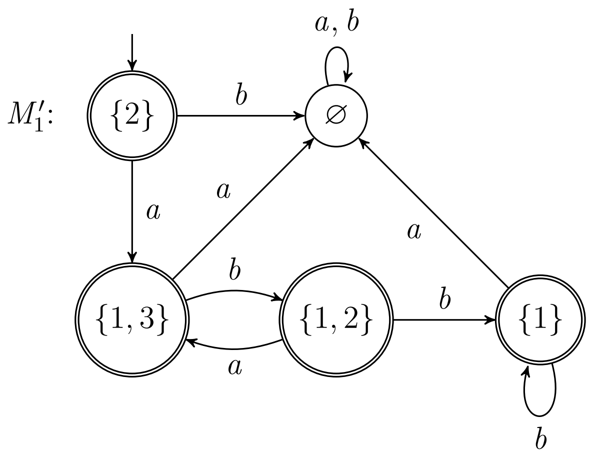


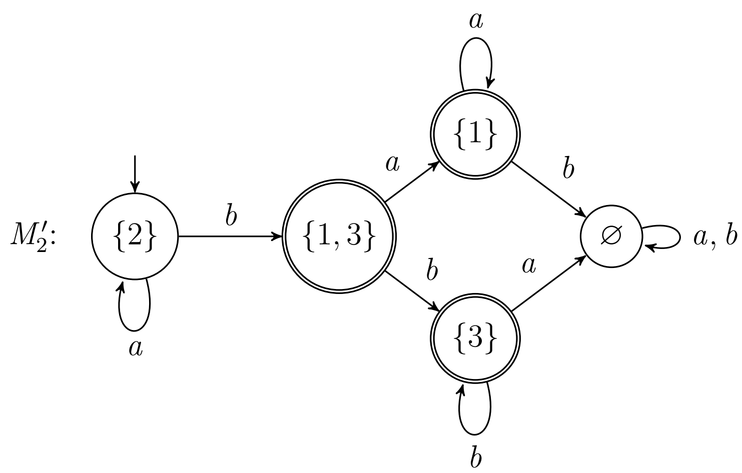
Lösen Sie mit dem Vorgehen aus der Vorlesung das Inklusionsproblem:

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

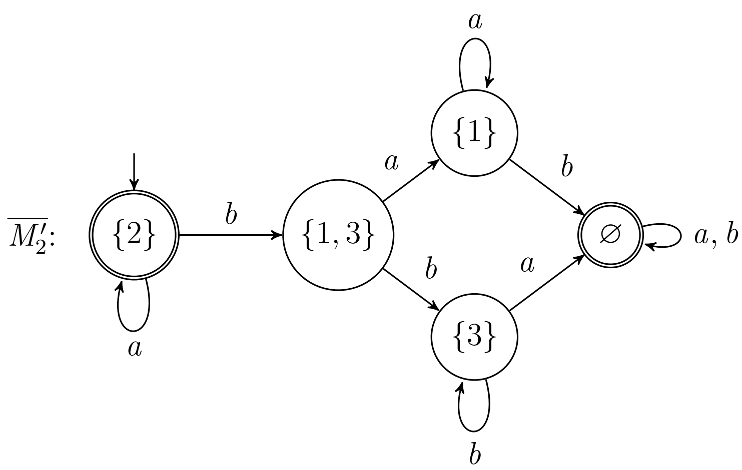
Lösung zu Aufgabe 1. Da $\varepsilon \in T(M_1)$ aber $\varepsilon \notin T(M_2)$, wissen wir, dass $T(M_1) \subseteq T(M_2)$ nicht gelten kann. $T(M_1) \subseteq T(M_2)$ ist äquivalent zu $T(M_1) \cap \overline{T(M_2)} = \varnothing$. Nun wollen wir mit dem Verfahren aus der Vorlesung zeigen, dass dies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

Zunächst konstruieren wir mittels Potenzmengenkonstruktion die DFAs für M_1 und M_2.

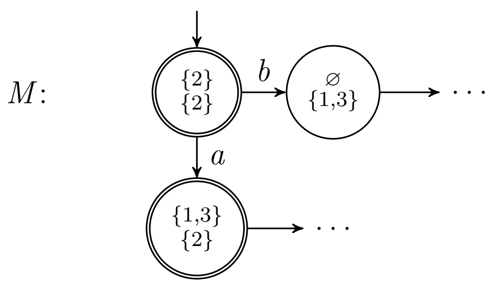




Dann erzeugen wir $\overline{M_2'}$, so dass $T(\overline{M_2'}) = \overline{T(M_2')} = \overline{T(M_2)}$, indem wir Endzustande und Nicht-Endzustände im DFA M_2' vertauschen.
<u>Beachte</u>: NFAs auf diese Art zu „komplementieren“ funktioniert nicht.



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



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.

Aufgabe 2. 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 2. 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.

Beispiel: $w = \sigma_1\sigma_2\sigma_3 a \sigma_5\sigma_6\sigma_7\sigma_8 b \sigma_{10}$ lässt sich unterteilen in $w_1 = \sigma_1\sigma_2\sigma_3 a \sigma_5$ und $w_2 = \sigma_6\sigma_7\sigma_8 b \sigma_{10}$ (wobei jedes σ_i ein beliebiges Symbol aus dem Alphabet $\Sigma = \{a, b\}$ ist).

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 $xa\boldsymbol{y'u'}bv$ (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.

Mit unserem Beispielwort ergibt sich eine Zerlegung in $w_1' = \sigma_1\sigma_2\sigma_3 a \sigma_5\sigma_6\sigma_7$ und $w_2' = \sigma_8 b \sigma_{10}$ bzw. der mittlere Teil $yu = \sigma_5 \;\; \sigma_6\sigma_7\sigma_8$ wird neu aufgeteilt zu $y'u' = \sigma_5\sigma_6\sigma_7 \;\; \sigma_8$.

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.

Für $w \in L$ mit gerader Länge ergibt sich so die Grammatik $G' = (V, \Sigma, P', S)$ mit

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

Wörter ungerader Länge

Zusätzlich benötigen wir noch Wörter ungerader Länge. Für alle $w \in \Sigma^*$ mit ungerader Länge gilt $w \in L$, da ungerade Wörter nie Teil von $\{ww \mid w \in \Sigma^*\}$ sein können.

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

Die komplette Grammatik ist $G = (V, \Sigma, P, S)$ mit

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

Aufgabe 3. Sind die folgenden Aussagen wahr oder falsch?

(a) Wenn L eine Sprache mit $\text{index}(R_L) = \infty$ ist, dann ist L kontextfrei.

(b) Wenn L eine nicht kontextfreie Sprache ist, dann ist $\text{index}(R_L) = \infty$.

(c) Wenn G eine Grammatik in CNF ist, dann ist $L(G)$ kontextfrei und nicht regulär.

(d) Es existieren kontextfreie Sprachen L_1 und L_2 so, dass $L_1 \cap L_2$ auch kontextfrei ist.

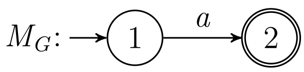
Lösung zu Aufgabe 3.

(a) falsch, weil für alle nicht-regulären Sprachen L auch $\text{index}(R_L) = \infty$ gilt, aber nicht jede nicht-reguläre Sprache ist eine kontextfreie Sprache, zum Beispiel $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$.

(b) wahr, weil jede Sprache, die nicht kontextfrei ist, auch nicht regulär ist und somit $\text{index}(R_L) = \infty$ gilt.

(c) falsch
Sei z.B. $G = (\{S\}, \{a\}, P, S)$ mit $P = \{S \to a\}$.

Diese Grammatik ist in CNF, aber es gilt $L(G) = \{a\}$ und das ist eine reguläre Sprache. Sie wird beispielsweise vom nachfolgenden NFA erzeugt:



(d) wahr
Sei z.B. $L_1 = L_2$, dann ist $L_1 \cap L_2 = L_1 = L_2$ und somit genau dann kontextfrei, wenn L_1 und L_2 kontextfrei sind.

Aufgabe 4. 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 4. 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.

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

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 $B_1 \to SB$.

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