Datum: 15.06.2020

Übungsblatt mit Lösungen 10

Aufgabe T27

Beweisen Sie formal, dass die kontextfreien Sprachen unter Homomorphismen abgeschlossen sind.

Lösungsvorschlag

Seien h ein Homomorphismus, L eine kontextfreie Sprache und $G = (N, T, P, S)$ eine kontextfreie Grammatik, die jene erzeugt. Ohne Einschränkung der Allgemeinheit ist G in Chomsky-Normalform. Außerdem gibt es für jedes Terminal $a \in T$ genau ein Nichtterminal $A \in N$ mit $A \to a$ und umgekehrt leitet jedes Nichtterminal *entweder* zu Nichtterminalen *oder* zu höchstens einem Terminal ab. Jedes Nichtterminal heißt R_a, wenn es nach a ableitet.

Daraus bauen wir eine neue Grammatik $G' = (N, T, P', S)$ für $h(L)$. Jede Regel $A \to BC$ aus P wird in P' übernommen. Kommt $S \to \varepsilon$ in P vor, dann auch in P'. Für jede Regel $R_a \to a$ in P fügen wir $A \to h(a)$ zu P' hinzu. Bemerke, dass G' nicht unbedingt in Chomsky-Normalform ist.

Beweis: Sei $w' \in L(G')$ und sei T' ein Ableitungsbaum von w' bezüglich G'. Wenn wir die Blätter löschen, ist die Front des Baumes $R_{a_1} R_{a_2} \dots R_{a_n}$. Insbesondere ist dieser Baum, nennen wir ihn T^*, auch ein gültiger Ableitungsbaum unter G. Fügen wir nun zu jedem Blatt R_{a_i} dessen eindeutige Produktion $R_{a_i} \to a_i$ hinzu, erhalten wir einen Ableitungsbaum T für ein Wort $w = a_1 \cdot a_2 \cdot \dots \cdot a_n$. Für dieses Wort gilt, dass $h(w) = w'$ nach Konstruktion und weil die Nichtterminale in R genau eine Produktion haben. Also gilt $L(G') \subseteq h(L)$.

Sei $w = a_1 \cdot a_2 \cdot \dots \cdot a_n \in L$ und folgend bauen wir einen Ableitungsbaum für $h(w)$ unter G'. Wieder sei T ein Ableitungsbaum von w unter G und T^* der resultierende Baum, wenn man von T die Blätter löscht. Er ist ein Ableitungsbaum von $R_{a_1} R_{a_2} \dots R_{a_n}$ unter G als auch G'. Nun wenden wir die eindeutigen Produktionen der R_{a_i} unter G' auf die Blätter an und erhalten einen Ableitungsbaum T' für ein Wort w' unter G'. Nach Konstruktion gilt $h(w) = w'$ und somit $h(L) \subseteq L(G')$.

Bemerkung: Hier nutzen wir den Begriff Ableitungsbäume auch wenn die Blätter aus Nichtterminalen bestehen.

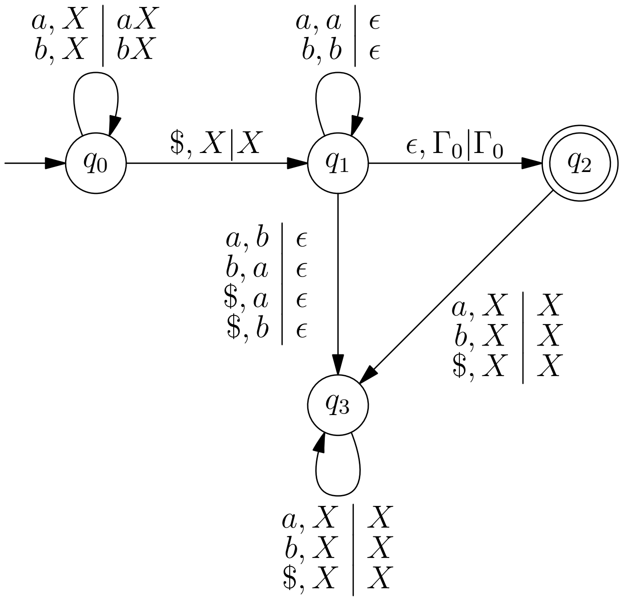
Aufgabe T28

Erstellen Sie einen deterministischen Kellerautomaten für folgende Sprache:

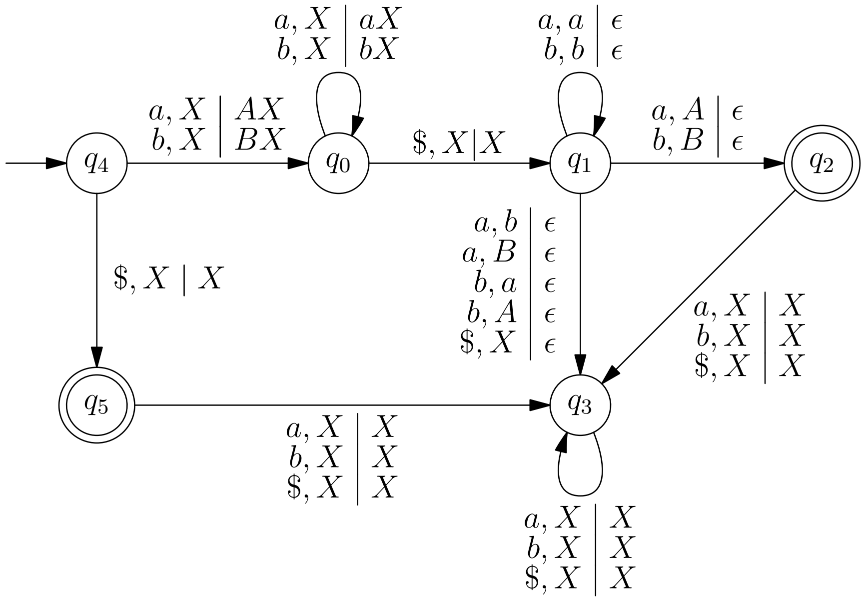
$$L := \{\$, a, b\}^* \setminus \{\, w\$w^R \mid w \in \{a, b\}^* \,\}$$

Lösungsvorschlag

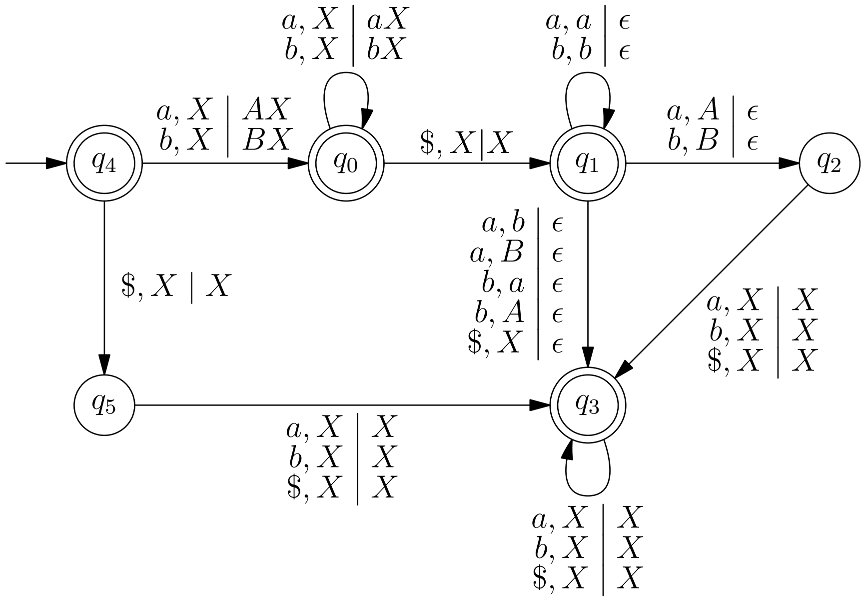
Zunächst erstellen wir einen Automaten für das Komplement, $\{\, w\$w^R \mid w \in \{a, b\}^* \,\}$.



Der Kellerautomat ist zwar deterministisch, enthält aber eine ϵ-Transitionen. Einfaches Tauschen der akzeptierenden Zustände führt daher leider nicht zu einem korrekten Ergebnis. Wir entfernen daher erst die ϵ-Transition, indem wir das erste Symbol auf dem Stack anders notieren, als die kommenden Symbole (und das Wort \$ gesondert behandeln).



Dieser Automat ist deterministisch, enthält keine ϵ-Transitionen und blockiert auch nicht. Letzteres folgt aus der Tatsache, dass alle Zustände außer q_1 vollständig sind und q_1 nur blockieren könnte, falls das Kellerbodensymbol auf dem Stack liegen würde. Dies ist aber in Zustand q_1 nicht möglich, da vorher auf jeden Fall ein A oder ein B auf den Stack geschrieben wurde. Einfaches Austauschen der akzeptierenden Zustände liefert somit einen Automat, der $\{\$, a, b\}^* \setminus \{\, w\$w^R \mid w \in \{a, b\}^* \,\}$ akzeptiert.



Aufgabe T29

Geben Sie für folgende Sprache eine kontextsensitive Grammatik an:
$L = \{ a^n b^n c^n \mid n \in \mathbf{N} \}$

Lösungsvorschlag

$$\begin{aligned}
S &\to aSBC \mid \varepsilon \\
CB &\to CD \\
CD &\to ED \\
ED &\to EC \\
EC &\to BC \\
aB &\to ab \\
bB &\to bb \\
bC &\to bc \\
cC &\to cc
\end{aligned}$$

Aufgabe H33 (10 Punkte)

Beweisen Sie formal, dass die kontextfreien Sprachen unter *Spiegelung* abgeschlossen sind, d.h. dass falls L kontextfrei ist, dann ist auch L^R kontextfrei.

Lösungsvorschlag

Seien L eine kontextfreie Sprache und $G = (N, T, P, S)$ eine kontextfreie Grammatik, die jene erzeugt. Ohne Einschränkung der Allgemeinheit ist G in Chomsky-Normalform.

Wir fügen jede Regel $A \to a \in P$ zu P' hinzu. Falls $S \to \varepsilon \in P$, dann fügen wir auch diese zu P' hinzu. Für jede Regel $A \to BC$ fügen wir $A \to CB$ hinzu.

Wir zeigen mittels einer Induktion über n, dass $S \Rightarrow_G^n \alpha$ genau dann wenn $G \Rightarrow_{G'}^n \alpha^R$. Für $i = 1$ ist die Aussage klar.

Sei $n+1 \in \mathbf{N}$ und es gelte die Aussage bereits für n. Sei $S \Rightarrow_G^{n+1} \alpha$. Dann gibt es Satzformen β, γ und eine Produktion $A \to BC$ bzw. $A \to a$ aus P, sodass $\alpha = \beta BC\gamma$ oder $\alpha = \beta a\gamma$ und $S \Rightarrow_G^n \beta A\gamma$. Nach Induktionsvorraussetzung gilt, dass $S \Rightarrow_{G'}^n \gamma^R A\beta^R$. Daraus kann mit $A \in CB \in P'$ (bzw. $A \to a$) $\gamma^R BC\beta^R$ (bzw $\gamma^R a\beta^R$) herleiten. Die umgekehrte Richtung kann man analog begründen (die Konstruktion ist selbst-dual). Damit ist die Aussage für $n+1$ gezeigt und nach dem Induktionsprinzip gilt die Aussage für alle n, also insbesondere beliebig lange Herleitungen, gezeigt.

Aufgabe H34 (10 Punkte)
Wir definieren die Menge der *linearen Sprachen* als Menge der Grammatiken, in der nur folgende Ableitungsregeln zugelassen sind: $A \to a$, $A \to aB, A \to Ba$ für $a \in T \cup \{\varepsilon\}, A, B \in N$.

Sie bestehen also aus allen gültigen Ableitungsregeln für rechts- und linkslineare Grammatiken.

Ordnen Sie die linearen Sprachen in die Chomsky-Hierachie ein. Zeigen Sie dafür gültige Teilmengenrelationen zwischen den linearen Sprachen und den regulären sowie kontextfreien Sprachen.

Es reicht, wenn Sie nur eins der beiden Sprachverhältnisse zeigen, also entweder linear zu regulär oder linear zu kontextfrei.

Lösungsvorschlag
Offensichtlich sind die linearen Sprachen eine Teilmenge der kontextfreien Sprachen, da diese über eingeschränkte kontextfreie Grammatiken erzeugt werden.

Sie sind jedoch eine echte Obermenge der regulären Sprachen: Sei $G = (\{A, B\}, \{a, b\}, P, A)$ eine lineare Grammatik mit folgenden Produktionsregeln P: $A \to aB \mid \varepsilon\ B \to Ab$ Diese Grammatik erzeugt die Sprache $a^n b^n$, welche, wie in der Vorlesung bereits gezeigt, nicht regulär ist.

Aufgabe H35 (10 Punkte)
Geben Sie für folgende Sprache eine kontextsensitive Grammatik an:
$L = \{\, w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c \,\}$

Lösungsvorschlag

$$
\begin{aligned}
&S \to SABC \mid \varepsilon \\
&A \to a \\
&B \to b \\
&C \to c
\end{aligned}
$$

und für alle $D, E \in \{A, B, C\}$ mit $D \neq E$ mit jeweils neuen Symbolen F, G:

$$
\begin{aligned}
DE &\to DG \\
DG &\to FG \\
FG &\to EG \\
EG &\to ED
\end{aligned}
$$