Tutorien-Übungsblatt 3

Aufgabe 1

Gegeben sei die Sprache $\mathcal{L} = \{w \in \{a,b\}^* \mid w \text{ enthält gleich viele } a \text{ wie } b\}$.

1. Wie lautet das Pumping Lemma? Was genau muss man zeigen, falls man die Kontraposition des Pumping Lemmas verwenden will?
2. Zeigen Sie mit Hilfe des Pumping Lemmas, dass L nicht regulär ist!
3. Zeigen Sie mit Hilfe des Pumping Lemmas, dass die Sprache $\mathcal{L}' = \{a^p \mid p \text{ Primzahl}\}$ nicht regulär ist.
4. Betrachten Sie nun die Sprache $\mathcal{L}'' = \{a, aab, aaab\}$! Ist diese regulär? Falls ja, geben Sie einen endlichen Automaten an, der diese Sprache akzeptiert! Kann man mit dem Pumping Lemma zeigen, dass die Sprache regulär ist?

Aufgabe 2

Gegeben sei die folgende Grammatik: $\mathcal{G} = (\mathcal{T}, \mathcal{V}, S, \mathcal{P})$ mit
$\mathcal{T} := \{a,b,c,d\}$, $\mathcal{V} := \{S, A, D, M\}$, $\mathcal{P} := \{S \to AMD \mid M, A \to AA \mid a, D \to DD \mid d, M \to bMc \mid \lambda\}$

1. Geben Sie die erzeugte Sprache an!
2. Wandeln Sie die gegebene kontextfreie Grammatik $\mathcal{G}$ in eine äquivalente kontextfreie Grammatik $\mathcal{G}'$ in Chomsky-Normalform um, indem sie jeden Schritt durch eine neue Grammatik beschreiben!
3. Zeigen oder widerlegen Sie mit Hilfe des CYK-Algorithmus, ob die folgenden Wörter in der Sprache $\mathcal{L}$ liegen, die durch die Grammatik $\mathcal{G}$ erzeugt wird!
   (a) $aabbccdd$
   (b) $abbcc$
   (c) $abcdd$

Aufgabe 3

Gegeben sei folgende Sprache für das Alphabet $\Sigma = \{a,b,c\}$:

$$\mathcal{L} = \{w_1 w_2 \in \Sigma^* \mid w_1 \in \{a,b\}^*, w_2 \in \{b,c\}^*, \#_a w_1 + \#_b w_1 = \#_b w_2 + \#_c w_2\}$$

Hier gibt $\#_x w$ die Häufigkeit des Vorkommens eines Zeichens $x \in \Sigma$ in einem Wort $w \in \Sigma^*$ an.

1. Zeigen Sie, dass $\mathcal{L}$ nicht regulär ist!
2. Geben Sie eine Chomsky-2-Grammatik an, die genau die Sprache $\mathcal{L}$ erzeugt!
3. Geben Sie einen Kellerautomaten $\mathcal{M}$ an, der genau die Sprache $\mathcal{L}$ erkennt! Zeichnen Sie den Zustandsübergangsgraphen für $\mathcal{M}$!

Lösung zu Aufgabe 1

1. Das Pumping Lemma: Sei L eine Sprache über einem Alphabet Σ.
   L regulär $\Rightarrow \exists\, n \in \mathbb{N} : \forall\, \omega \in L, |\omega| \geq n : \exists\, \alpha, \beta, \gamma \in \Sigma^*, \omega = \alpha\beta\gamma :$
   $\beta \neq \lambda \wedge |\alpha\beta| \leq n \wedge \forall\, i \in \mathbb{N}_0 : \alpha\beta^i\gamma \in L$

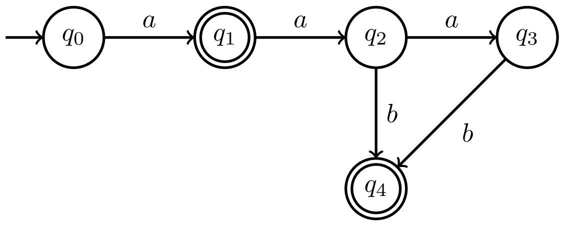
   Kontraposition des Pumping Lemmas: Sei L eine Sprache über einem Alphabet Σ.
   $\forall\, n \in \mathbb{N} : \exists\, \omega \in L, |\omega| \geq n : \forall\, \alpha, \beta, \gamma \in \Sigma^*, \omega = \alpha\beta\gamma :$
   $\beta = \lambda \vee |\alpha\beta| > n \vee \exists\, i \in \mathbb{N}_0 : \alpha\beta^i\gamma \notin L \Rightarrow L$ nicht regulär

   Falls man also die Kontraposition des Pumping Lemmas verwenden will, muss man zu jeder natürlichen Zahl als Mindestwortlänge ein beliebiges, frei wählbares Wort ω finden, dass jede mögliche Zerlegung von ω eine der drei Eigenschaften des Pumping Lemmas verletzt. Dazu betrachtet man der Einfachheit wegen nur alle Zerlegungen von ω, die die ersten beiden Eigenschaften des Pumping Lemmas erfüllen, und zeigt dann, dass diese die dritte Eigenschaft verletzen.

2. Behauptung: $\mathcal{L}$ nicht regulär
   Beweis: Verwendung der Kontraposition des Pumping Lemmas
   Sei $n_0 \in \mathbb{N}$ beliebig.
   Wähle $\omega \in \mathcal{L}$ geeignet mit $|\omega| \geq n_0$, wähle also beispielsweise $\omega = a^{n_0}b^{n_0}$. $\Rightarrow$
   $\forall\, \alpha, \beta, \gamma \in \{a, b\}^*, \omega = \alpha\beta\gamma, \beta \neq \lambda, |\alpha\beta| \leq n_0 : \alpha\beta = a^k$ mit $k \in \mathbb{N}_0, 1 \leq k \leq n_0$
   Seien also $\beta = a^i$ mit $i \in \mathbb{N}_0, 1 \leq i \leq n_0$, $\alpha = a^j$ mit $j \in \mathbb{N}_0, 0 \leq j \leq n_0 - i$ und $\gamma = a^{n_0-i-j}b^{n_0}$. $\Rightarrow$
   Es gelten $\omega = \alpha\beta\gamma, \beta \neq \lambda, |\alpha\beta| \leq n_0$.
   Aber $\alpha\beta^2\gamma = a^{n_0+i}b^{n_0} \notin L$ wegen $i \geq 1$. $\Rightarrow L$ nicht regulär

3. Behauptung: $\mathcal{L}'$ nicht regulär
   Beweis: Verwendung der Kontraposition des Pumping Lemmas
   Sei $n_0 \in \mathbb{N}$ beliebig.
   Wähle $\omega \in \mathcal{L}$ geeignet mit $|\omega| \geq n_0$, wähle also beispielsweise $\omega = a^p$ mit p Primzahl, $p \geq n_0 + 2$. $\Rightarrow$
   $\forall\, \alpha, \beta, \gamma \in \{a\}^*, \omega = \alpha\beta\gamma, \beta \neq \lambda, |\alpha\beta| \leq n_0 : \alpha\beta = a^k$ mit $k \in \mathbb{N}_0, 1 \leq k \leq n_0$
   Seien also $\beta = a^i$ mit $i \in \mathbb{N}_0, 1 \leq i \leq n_0$, $\alpha = a^j$ mit $j \in \mathbb{N}_0, 0 \leq j \leq n_0 - i$ und $\gamma = a^{p-i-j}$. $\Rightarrow$
   Es gelten $\omega = \alpha\beta\gamma, \beta \neq \lambda, |\alpha\beta| \leq n_0$.
   Aber $\alpha\beta^{p-i}\gamma = a^j a^{i\cdot(p-i)} a^{p-i-j} = a^{(i+1)\cdot(p-i)} \notin L$, denn wegen $i \geq 1$ folgt $i + 1 \geq 2$ und wegen $p \geq n_0 + 2$ und $i + j \leq n_0$ folgt $p - i \geq p - i - j = p - (i + j) \geq p - n_0 \geq 2$. $\Rightarrow L$ nicht regulär

4. Ja, die Sprache ist regulär, denn jede endliche Sprache ist regulär. Folgender endliche Automat akzeptiert $\mathcal{L}''$:

   Man kann mit dem Pumping Lemma niemals zeigen, dass eine Sprache regulär ist.

Lösung zu Aufgabe 2

1. Die gesuchte Sprache ist $\mathcal{L}(\mathcal{G}) = \{a^m b^n c^n d^k \mid m, k \in \mathbb{N}, n \in \mathbb{N}_0\} \cup \{b^n c^n \mid n \in \mathbb{N}_0\}$.

2. Wir gehen in 3 Schritten vor und konstruieren zu $\mathcal{G}$ äquivalente Grammatiken $\mathcal{G}_i = (\mathcal{T}, \mathcal{V}_i, S, \mathcal{P}_i)$ für $i \in \{1, 2, 3\}$, die der Chomsky-Normalform immer näher kommen.

   Für jedes Terminalzeichen $a \in \mathcal{T}$ und für jede Produktion, die dieses Terminalzeichen a auf der rechten Seite enthält, ausgenommen natürlich Produktionen der Form $V \to a$ mit $V \in \mathcal{V}$, wird jedes Vorkommen von a in der Produktion durch ein neues Nichtterminalzeichen A, welches in die Variablenmenge aufgenommen wird, ersetzt und zusätzlich wird noch die Produktion $A \to a$ hinzugefügt.
   Dies ergibt: $\mathcal{V}_1 := \{S, A, B, C, D, M\}$ und
   $\mathcal{P}_1 := \{S \to AMD \mid M, A \to AA \mid a, D \to DD \mid d, M \to BMC \mid \lambda, B \to b, C \to c\}$

   In diesem Schritt werden alle Produktionen der Form $V \to \lambda$ für $V \in \mathcal{V}, V \neq S$ entfernt. Dazu müssen diese Produktion aber vorher rekursiv durch ihre "Vorwegnahme" mit den anderen Produktionen "verschmolzen" werden, es wird also für jede Produktion mit einem der obigen V auf der rechten Seite eine neue Produktion ohne dieses V hinzugefügt.
   Dies ergibt: $\mathcal{V}_2 := \{S, A, B, C, D, M\}$ und
   $\mathcal{P}_2 := \{S \to AMD \mid AD \mid M \mid \lambda, A \to AA \mid a, D \to DD \mid d, M \to BMC \mid BC, B \to b, C \to c\}$

Im letzten Schritt werden die Produktionen, deren rechte Seite nur aus einer einzigen oder mehr als zwei Variablen bestehen, durch einen Trick ersetzt. Bei den einzelnen Variablen werden diese Produktionen wieder durch ihre “Vorwegnahme” mit den anderen Produktionen “verschmolzen”. Für die Produktionen mit mehr als zwei Variablen auf der rechten Seite werden neue Nichterminale eingeführt und dazu passende Produktionen hinzugefügt.
Dies ergibt: $\mathcal{V}_3 := \{S, A, B, C, D, M, X, Y\}$ und
$\mathcal{P}_3 := \{S \to AX \mid AD \mid BY \mid BC \mid \lambda, A \to AA \mid a, D \to DD \mid d, M \to BY \mid BC, X \to MD, Y \to MC,$
$B \to b, C \to c\}$

3. (a) $aabbccdd \in \mathcal{L}$?

S							
S	S						
	S	X					
	$*$	X					
		S, M					
			Y				
A			S, M			D	
A	A	B	B	C	C	D	D
a	a	b	b	c	c	d	d

Markierung $*$ für nächste Teilaufgabe!

(b) $abbcc \in \mathcal{L}$?
Es gilt $abbcc \notin \mathcal{L}$, da dann das mit $*$ markierte Feld in der vorherigen Teilaufgabe die Startvariable S enthalten müsste! Man kann also die Lösung des CYK-Algorithmus zu einem bestimmten Wort für Teilworte weiterverwenden!

(c) $abcdd \in \mathcal{L}$?

S				
S	X			
	X			
	S, M		D	
A	B	C	D	D
a	b	c	d	d

Lösung zu Aufgabe 3

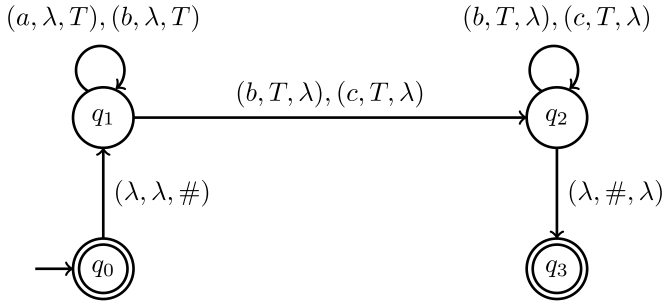
1. Wir verwenden das Pumping-Lemma. Man wählt das Wort $w = a^n c^n \in \mathcal{L}$, wobei $n \in \mathbb{N}$ die Konstante aus dem Pumping-Lemma ist. Dann gilt mit $w = \alpha\beta\gamma$, wobei $|\alpha\beta| \leq n$ und $\beta \neq \lambda$, dass α, β und γ von der Form $\alpha = a^{n-r_1-r_2}$, $\beta = a^{r_1}$ $(r_1 > 0)$, $\gamma = a^{r_2} c^n$ $(r_2 \geq 0)$ sind. Damit ist $w' = \alpha\beta^2\gamma = a^{n+r_1} c^n \notin \mathcal{L}$, da die Zerlegung von $w' = w_1 w_2$ eindeutig in $w_1 = a^{n+r_1}$ und $w_2 = c^n$ vorgegeben ist und $\#_a w_1 \neq \#_c w_2$ gilt. Damit ist $\mathcal{L}$ nicht regulär.

2. Die Grammatik $\mathcal{G} = (\mathcal{T}, \mathcal{V}, S, \mathcal{P})$ mit den Terminalen $\mathcal{T} = \{a, b, c\}$, den Nichtterminalen $\mathcal{V} = \{S, X\}$ und den Produktionen $\mathcal{P}$

$$S \to aXb \mid aXc \mid bXb \mid bXc \mid ab \mid ac \mid bb \mid bc \mid \lambda$$
$$X \to aXb \mid aXc \mid bXb \mid bXc \mid ab \mid ac \mid bb \mid bc$$

leistet das Gewünschte. Der Induktionsbeweis hierzu besteht darin zu zeigen, dass für alle Teilableitungen $w = w_1 X w_2$ gilt $w_1 \in \{a, b\}^*$, $w_2 \in \{b, c\}^*$ und $\#_a w_1 + \#_b w_1 = \#_b w_2 + \#_c w_2$. Man sieht recht einfach, dass der erste Ableitungsschritt hierfür die Induktionsverankerung und der zweite den Induktionsschritt liefert.

3. Ein Kellerautomat $\mathcal{M} = (\mathcal{Q}, \Sigma, \Gamma, \delta, q_0, \mathcal{F})$ mit $\Gamma = \{\#, T\}$ und der Akzeptionsbedingung “leerer Keller”, der die Sprache $\mathcal{L}$ akzeptiert, kann beispielsweise so aussehen:


$(a, \lambda, T), (b, \lambda, T)$
$(b, T, \lambda), (c, T, \lambda)$
$(b, T, \lambda), (c, T, \lambda)$
q_1
q_2
$(\lambda, \lambda, \#)$
$(\lambda, \#, \lambda)$
q_0
q_3