Prof. Dr. Jasmin Blanchette
Jannis Limperg
Lydia Kondylidou

Ludwig-Maximilians-Universität München
Institut für Informatik
Abgabe bis 7. Juni 2023, 10:00 Uhr

Lösungsvorschlag zur Übung 7 zur Vorlesung

Formale Sprachen und Komplexität

FSK7-1 ***CYK-Algorithmus*** (2 Punkte)

Sei G die Grammatik $(\{A_1, A_2, A_3, A_4, A_5\}, \{\$, \#\}, P, A_1)$ mit

$$\begin{aligned} P = \{ & A_1 \to A_3A_4 \mid A_3A_2, \\ & A_2 \to A_1A_4, \\ & A_3 \to \$, \\ & A_4 \to \# \mid A_3A_4, \\ & A_5 \to A_4A_4 \mid \#\} \end{aligned}$$

a) Prüfen Sie mit dem CYK-Algorithmus, ob die folgenden Wörter w_1 und w_2 in $L(G)$ sind.

- $w_1 = \$\$\$\#\#\#$

LÖSUNGSVORSCHLAG:

Wort:	\$	\$	\$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

Da das Startsymbol A_1 in Zeile 6, Spalte 1 enthalten ist, liegt \$\$\$### in $L(G)$.

- $w_2 = \$\$\$\#\$$

LÖSUNGSVORSCHLAG:

Wort:	\$	\$	\$	#	\$
j\ i	1	2	3	4	5
1	A_3	A_3	A_3	A_4, A_5	A_3
2			A_1, A_4		
3		A_1, A_4			
4	A_1, A_4				
5					

Da das Startsymbol A_1 nicht in Zeile 5, Spalte 1 enthalten ist, liegt $$$#$ nicht in $L(G)$.

Erstellen Sie dazu für jedes Wort die entsprechende Tabelle des Algorithmus und erläutern Sie anhand der Tabelle, ob das Wort in $L(G)$ ist.

b) Verwenden Sie Ihre beiden zuvor erstellten Tabellen, um weitere Wörter w zu finden, die in $L(G)$ liegen. Welche $w \in L(G)$ sind aus den Tabellen ablesbar?

LÖSUNGSVORSCHLAG:

Alle Zellen, die das Startsymbol A_1 enthalten, stehen für ein Teilwort, das in $L(G)$ liegt. In der zweiten Tabelle sind es die Wörter $$$#, $$#, $#:

Wort:	$	$	$	#	$
j\ i	1	2	3	4	5
1	A_3	A_3	A_3	A_4, A_5	A_3
2			A_1, A_4		
3		A_1, A_4			
4	A_1, A_4				
5					

Wort:	$	$	$	#	$
j\ i	1	2	3	4	5
1	A_3	A_3	A_3	A_4, A_5	A_3
2			A_1, A_4		
3		A_1, A_4			
4	A_1, A_4				
5					

Wort:	$	$	$	#	$
j\ i	1	2	3	4	5
1	A_3	A_3	A_3	A_4, A_5	A_3
2			A_1, A_4		
3		A_1, A_4			
4	A_1, A_4				
5					

In der ersten Tabelle neben dem getesten Wort auch $$$##, $$$#, $$##, $$#, $#:

Wort:	$	$	$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

Wort:	$	$	$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

Wort:	$	$	$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

Wort:	$	$	$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

Wort:	$	$	$	#	#	#
j\ i	1	2	3	4	5	6
1	A_3	A_3	A_3	A_4, A_5	A_4, A_5	A_4, A_5
2			A_1, A_4	A_5	A_5	
3		A_1, A_4	A_2, A_5			
4	A_1, A_4	A_1, A_2, A_5				
5	A_1, A_2, A_5	A_2				
6	A_1, A_2					

FSK7-2 *Pumping-Lemma für kontextfreie Sprachen* (2 Punkte)

a) Betrachten Sie die Sprache $L_1 = \{ba^{2k}ba^kca^{2k} \mid k \in \mathbb{N}\}$ über dem Alphabet $\{a, b, c\}$. Zeigen Sie mit dem Pumping-Lemma für kontextfreie Sprachen, dass L_1 nicht kontextfrei ist.

LÖSUNGSVORSCHLAG:

Wir zeigen, dass L_1 die Pumpingeigenschaft für kontextfreie Sprachen nicht erfüllt. Zu zeigen ist also:

$$\begin{aligned}\forall n \in \mathbb{N}_{>0}.\ &\exists z \in L_1, |z| \geq n.\ \forall u, v, w, x, y.\\ &z = uvwxy \land |vx| \geq 1 \land |vwx| \leq n \implies \exists i \geq 0 : uv^i wx^i y \notin L_1\end{aligned}$$

Sei $n \geq 1$ beliebig. Wir wählen $z = ba^{2n}ba^n ca^{2n}$. Damit ist $z \in L_1$ und $|z| \geq n$.

Wir betrachten nun jede Zerlegung $z = uvwxy$ mit $|vwx| \leq n$ und $|vx| \geq 1$. Vollständige Fallunterscheidung:

- v oder x enthält ein b. Wähle $i = 0$ und betrachte $z' = uv^0wx^0y$. Es ist $\#_b(z') < 2$ und damit $z' \notin L_1$.
- v oder x enthält ein c. Wieder für $i = 0$ und $z' = uv^0wx^0y$ ist $\#_c(z') < 1$ und $z' \notin L_1$.
- v und x enthalten weder ein b noch ein c, d.h. $v = a^k$ und $x = a^l$. Dann liegt vwx in mindestens einem und höchstens zwei der a-Blöcke von z, da $|vwx| \leq n$. Für $i > 1$ enthalten diese Blöcke in $z' = uv^iwx^iy$ mehr a's als in z, aber der dritte Block enthält gleich viele a's wie in z. Damit ist $z' \notin L_1$.

b) Betrachten Sie die Sprache $L_2 = \{a^p b^p \mid p \text{ ist eine Primzahl}\}$ über dem Alphabet $\{a, b\}$. Zeigen Sie mit dem Pumping-Lemma für kontextfreie Sprachen, dass auch L_2 nicht kontextfrei ist.

LÖSUNGSVORSCHLAG:

Sei $n \geq 1$ beliebig. Wir wählen $z = a^q b^q$, wobei q die kleinste Primzahl mit $q \geq n$ ist. Damit ist $z \in L_2$ und $|z| \geq n$.

Wir betrachten nun jede Zerlegung $z = uvwxy$ mit $|vwx| \leq n$ und $|vx| \geq 1$. Sei $c \in \{a, b\}$ ein Zeichen mit $\#_c(vx) \geq 1$. (Ein solches Zeichen existiert, da $|vx| \geq 1$.) Dann gilt für beliebige i mit $z' = uv^iwx^iy$:

$$\#_c(z') = \#_c(uwy) + \#_c(v^i x^i) = \#_c(uwy) + i \cdot \#_c(vx)$$

Wähle nun $i = \#_c(uwy)$. Dann ist

$$\#_c(z') = \#_c(uwy) + \#_c(uwy) \cdot \#_c(vx) = \#_c(uwy) \cdot (1 + \#_c(vx))$$

und somit ist $\#_c(z')$ nicht prim, d.h. $z' \notin L_2$.

FSK7-3 ***Kontextfreie Sprachen*** (0 Punkte)

Sei L die formale Sprache aller Wörter $w \in \{a,b\}^*$, sodass w in der zweiten Hälfte mindestens ein b enthält:

$$L = \{ubv \mid u, v \in \{a,b\}^*, |u| > |v|\}$$

a) Geben Sie eine kontextfreie Grammatik an, die L erzeugt.

LÖSUNGSVORSCHLAG:

$G = (\{S, U, V\}, \{a, b\}, P, S)$ mit

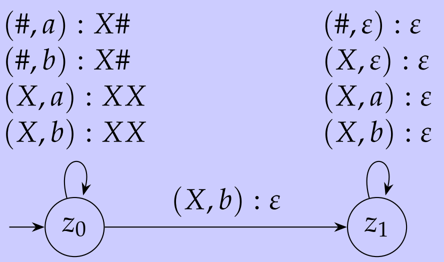
$$P = \{S \to USV \mid Ub,\ U \to a \mid b \mid UU,\ V \to a \mid b\}$$

Erläuterung: S erzeugt eine Satzform der Form $U^i b V^{i-1}$; jedes von V erzeugte Wort hat Länge 1; jedes von U erzeugte Wort hat Länge mindestens 1. Daher erzeugt U^i ein Wort der Länge mindestens i und V^{i-1} ein Wort der Länge genau $i-1$. Damit gilt $|u| > |v|$. Die Nichtterminale U und V können beliebig a's und b's erzeugen.

b) Geben Sei einen Kellerautomaten an, der L akzeptiert. (Entsprechend der Definition aus der Vorlesung soll Ihr Kellerautomat durch leeren Keller akzeptieren.)

Hinweis: Nutzen Sie den Nichtdeterminismus des Automaten aus!

LÖSUNGSVORSCHLAG:



Erläuterung: Im Startzustand z_0 werden beliebige a und b gelesen, wobei jeweils ein X pro Zeichen in den Keller gelegt wird. Der Nichtdeterminismus wird verwendet, um das richtige b zu lesen und in den Zustand z_1 zu wechseln. Mit diesem Wechseln können nur noch soviele a und b gelesen werden, wie zuvor gelesen wurden, da jedes Lesen ein X vom Keller abbaut. Folgen weniger Zeichen, so werden die verbleibenden X durch ε-Übergänge abgebaut. Das Startsymbol im Keller wird durch einen weiteren ε-Übergang entfernt.

FSK7-4 ***Begrenzter Stack*** (0 Punkte)

a) Ein Kellerautomat mit k-begrenztem Stack ist ein Kellerautomat, bei dem maximal k Symbole auf dem Stack liegen können. Wenn der Stack voll ist, können Übergänge, die ein Symbol auf den Stack legen würden, nicht gewählt werden.

Zeigen Sie, dass für alle $k \in \mathbb{N}_{>0}$ die Kellerautomaten mit k-begrenztem Stack genau die regulären Sprachen beschreiben.

LÖSUNGSVORSCHLAG:

Man kann jeden endlichen Automaten als Kellerautomaten mit Endzuständen auffassen, der immer das Symbol # vom Stack nimmt und wieder auf den Stack legt. Dieser akzeptiert die gleiche Sprache. Damit ist jede reguläre Sprache durch einen 1-begrenzten Stack (und damit auch einen k-begrenzten Stack) darstellbar. Aus der Vorlesung wissen wir, dass Akzeptanz durch leeren Keller äquivalent zu Akzeptanz mit Endzuständen ist. Daher gilt die Aussage für Kellerautomaten ohne Endzustände ebenfalls.

Umgekehrt kann ein Kellerautomat mit k-begrenztem Stack nur endlich viele verschiedene Stackkonfigurationen (damit ist nur der Teil $W \in \Gamma^*$ aus der Konfiguration des Kellerautomatens gemeint) haben, da es nur endlich viele verschiedene Stacksymbole gibt und es eine feste obere Grenze für die Stacktiefe gibt.

Sei nun C die Menge dieser möglichen Stackkonfigurationen und Z die Zustandsmenge des Kellerautomaten. Dann können wir einen nichdeterministischen endlichen Automaten (mit ε-Übergängen) konstruieren, dessen Zustandsmenge $C \times Z$ ist. Der Automat hat einen Übergang von Zustand (c, z) zu (c', z') mit Symbol a (oder mit ε) genau dann wenn der Kellerautomat mit a (oder mit ε) in der Stackkonfiguration c von z nach z' mit Stackkonfiguration c' übergehen würde. Die Endzustände des NFA sind Zustände (c, z), bei denen z ein Endzustand des Kellerautomaten ist. (Hier nehmen wir o.b.d.A. an, dass der Kellerautomat mit Endzuständen akzeptiert.)

b) Zeigen Sie, dass für jeden Kellerautomaten K und $i \in \mathbb{N}_0$ gilt:

$$L(K \text{ mit } i\text{-begrenztem Stack}) \subseteq L(K).$$

LÖSUNGSVORSCHLAG:

Wir betrachten o.b.d.A. Kellerautomaten, die mit leerem Keller akzeptieren. Jedes Wort, das von so einem Kellerautomaten akzeptiert werden kann, ohne dass der Stack die Länge k überschreitet, kann auch von dem Kellerautomaten allgemein akzeptiert werden, da die gleiche Konfigurationsfolge auch im Kellerautomaten ohne Stackbegrenzung gültig ist und den Stack leert.

Es gilt also für alle $w \in \Sigma^*$: wenn $w \in L(K \text{ mit } i\text{-begrenztem Stack})$ ist, dann ist auch $w \in L(K)$.