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
Theoretische Informatik für Medieninformatiker

TIMI7-1 *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					

TIMI7-2 ***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

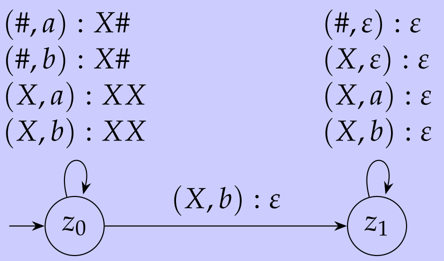
$$P = \{S \to USV \mid Ub,\ U \to a \mid b \mid UU,\ V \to a \mid b\}$$

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.