10. Aufgabenblatt zur Vorlesung

Grundlagen der theoretischen Informatik SoSe 2020

Wolfgang Mulzer

Abgabe bis zum 06. Juli 2020, 10 Uhr, im Whiteboard
Dies ist das vorletzte Aufgabenblatt.

Bitte erläutern und begründen Sie alle Ihre Antworten.

Aufgabe 1 Grammatiken und der Kleene-Stern *4+6 Punkte*

(a) Die folgende Konstruktion soll zeigen, dass kontextfreie Sprachen unter dem Kleene-Stern abgeschlossen sind: Sei $G = (\Sigma, V, S, P)$ eine kontextfreie Grammatik. Füge zu G die Ableitung $S \to SS$ hinzu und nenne die resultierende Grammatik G'. Behauptung: $L(G') = L(G)^*$.

Zeigen Sie durch ein Gegenbeispiel, dass die obige Konstruktion im Allgemeinen nicht korrekt ist.

(b) Geben Sie eine korrekte Konstruktion für das Problem aus (a) an und begründen Sie die Korrektheit.

Aufgabe 2 Chomsky-Normalform *10 Punkte*

Wandeln Sie die folgende Grammatik über dem Terminalalphabet $\Sigma = \{0, 1\}$ in Chomsky-Normalform um.

$$\begin{array}{ll} S \to 0A0 \mid 1B1 \mid BB & \quad B \to S \mid A \\ A \to C & \quad C \to S \mid \varepsilon \end{array}$$

Aufgabe 3 Der CYK-Algorithmus *10 Punkte*

Gegeben sei die folgende Grammatik über dem Alphabet $\Sigma = \{a, b\}$.

$$S \to AB \mid BC, A \to BA \mid a, B \to CC \mid b, C \to AB \mid a.$$

Wenden Sie den CYK-Algorithmus an auf die Wörter $w_1 = aaaaa$, $w_2 = aaaaaa$ und $w_3 = baaba$. Zeichnen Sie gegebenenfalls auch einen zugehörigen Syntaxbaum.

Aufgabe 1 Grammatiken und der Kleene-Stern *4+6 Punkte*

Von David Sey und Thore Brehmer

(a) Die folgende Konstruktion soll zeigen, dass kontextfreie Sprachen unter dem Kleene-Stern abgeschlossen sind: Sei $G = (\Sigma, V, S, P)$ eine kontextfreie Grammatik. Füge zu G die Ableitung $S \to SS$ hinzu und nenne die resultierende Grammatik G'. Behauptung: $L(G') = L(G)^*$.

Zeigen Sie durch ein Gegenbeispiel, dass die obige Konstruktion im Allgemeinen nicht korrekt ist.

(b) Geben Sie eine korrekte Konstruktion für das Problem aus (a) an und begründen Sie die Korrektheit.

a) Geg. Bsp.

Sei $G = (\Sigma, V, S, P)$ mit $\Sigma = \{a\}$, $V = \{S\}$ $P: S \to a$

Sei $G' =$ " ——————— " $P: S \to a \mid SS \mid$

$L(G) = \{a\}$

$L(G') = \{a, aa, \ldots\}$

$L(G)^* = \{\varepsilon, a, aa, \ldots\}$

$L(G') \neq L(G)^*$

$S \to SS \mid a \mid Ab \mid BB$

$S \to S_1 S$

$S_1 \to a \mid Ab \mid BB$

b)

Das einzige existierende mit dieser Konstruktion ist in a zu sehen

Daher muss nurnoch die Produktion $S \to \varepsilon$ hinzugefügt werden.

Die Konstruktion ist Korrekt, da Betrachte die Grammatik G

Sei $w_1, w_2 \in L(G)$. Für den Kleenestern muss nun gelten $\varepsilon \in L(G) \wedge w_1 \circ w_2 \in L(G)$.

$\varepsilon \in L(G)$ kann durch $S \to \varepsilon$ erreicht werden

$w_1 \circ w_2 \in L(G)$ kann durch $S \to SS$ erreicht werden, da durch den Start der Variabel S aus G gefolgt von Produktionen immer genau ein Wort ausgegeben wird.

Mit $S \to SS$ können also aneinander gereihte Wörter wie $w_1 \circ w_2$ ausgegeben werden.

Aufgabe 2 Chomsky-Normalform *10 Punkte*

Wandeln Sie die folgende Grammatik über dem Terminalalphabet $\Sigma = \{0,1\}$ in Chomsky-Normalform um.

$$S \to 0A0 \mid 1B1 \mid BB \qquad B \to S \mid A$$
$$A \to C \qquad C \to S \mid \varepsilon$$

Gegeben:

$S \to 0A0 \mid 1B1 \mid BB$
$A \to C$
$B \to S \mid A$
$C \to S \mid \varepsilon$

Schritt 1: Trennen von Terminalsymbolen

$S \to V_0AV_0 \mid V_1BV_1 \mid BB$ $\qquad V_0 \to 0$
$A \to C$ $\qquad V_1 \to 1$
$B \to S \mid A$
$C \to S \mid \varepsilon$

Schritt 2: Beseitigen langer Produktionen

$S \to V_0D \mid V_1E \mid BB$ $\qquad V_0 \to 0$
$A \to C$ $\qquad V_1 \to 1$
$B \to S \mid A$ $\qquad D \to AV_0$
$C \to S \mid \varepsilon$ $\qquad E \to BV_1$

Schritt 3: Beseitigen von ε-Produktionen

- 3a: Makiere alle Variabeln mit $X \xrightarrow{*} \varepsilon$

$S \to V_0D \mid V_1E \mid BB$ $\qquad V_0 \to 0$
$A \to C$ $\qquad V_1 \to 1$
$B \to S \mid A$ $\qquad D \to AV_0$
$C \to S \mid \varepsilon$ $\qquad E \to BV_1$

- 3b: Erzeuge Produktionen wie folgt: $X \to YZ$ erzeuge $X \to Y$, $X \to YZ$ erzeuge $X \to Z$, $X \to YZ$ erzeuge $X \to Y \wedge X \to Z$

$S \to V_0D \mid V_1E \mid BB \mid B \mid \cancel{B}$ $\qquad V_0 \to 0$
$A \to C$ $\qquad V_1 \to 1$
$B \to S \mid A$ $\qquad D \to AV_0 \mid V_0$
$C \to S \mid \varepsilon$ $\qquad E \to BV_1 \mid V_1$

- **3c:** Entferne Produktionen der Form $X \to X$ und $X \to \varepsilon$

$$
\begin{aligned}
S &\to V_1D \mid V_1E \mid BB \mid B & V_0 &\to 0\\
A &\to C & V_1 &\to 1\\
B &\to S \mid A & D &\to AV_0 \mid V_0\\
C &\to S \mid \cancel{\varepsilon} & E &\to BV_1 \mid V_1
\end{aligned}
$$

- **3d:** Falls S makiert war, erstelle eine neue Startvariabel \$

$$
\begin{aligned}
S &\to V_1D \mid V_1E \mid BB \mid B & V_0 &\to 0\\
A &\to C & V_1 &\to 1\\
B &\to S \mid A & D &\to AV_0 \mid V_0\\
C &\to S & E &\to BV_1 \mid V_1\\
\$ &\to V_1D \mid V_1E \mid BB \mid B \mid \varepsilon
\end{aligned}
$$

Schritt 4: Beseitegen von Kettenregeln

- **4a:** Beseitige Kreise, indem alle V. im Kreis durch eine neue V. ersetzt werden.

  - Wähle neue Variabel F

$$
\begin{tikzcd}
S \arrow[r, leftrightarrow] & B \arrow[d] & \$ \arrow[l] \\
C \arrow[u] & A \arrow[l] &
\end{tikzcd}
\qquad
\begin{tikzcd}
V_0 & V_1 \\
D \arrow[u] & E \arrow[u]
\end{tikzcd}
$$

$$
\begin{aligned}
F &\to V_1D \mid V_1E \mid FF \mid F & V_0 &\to 0\\
F &\to F & V_1 &\to 1\\
F &\to F \mid F & D &\to FV_0 \mid V_0\\
F &\to F & E &\to FV_1 \mid V_1\\
\$ &\to V_1D \mid V_1E \mid FF \mid F \mid \varepsilon
\end{aligned}
$$

  - Zusammenfassen und $X \to X$ löschen

$$
\begin{tikzcd}
F & \$ \arrow[l]
\end{tikzcd}
\qquad
\begin{tikzcd}
V_0 & V_1 \\
D \arrow[u] & E \arrow[u]
\end{tikzcd}
$$

$$
\begin{aligned}
F &\to V_1D \mid V_1E \mid FF \mid \cancel{F} & V_0 &\to 0\\
&\cancel{F \to F} & V_1 &\to 1\\
&\cancel{F \to F} \mid F & D &\to FV_0 \mid V_0\\
&\cancel{F \to F} & E &\to FV_1 \mid V_1\\
\$ &\to V_1D \mid V_1E \mid FF \mid F \mid \varepsilon
\end{aligned}
$$

- **4b:** Sotiere den Graphen topologisch und führe die Kettenregel von rechts nach links aus.
Kettenregel: Falls auf der rechten Seite einer P. eine V. alleine steht, ersetze die V. in dieser P. durch alle P. aus der V.

Topo: $\{D, V_0, E, V_1, \$, F\}$

$$
\begin{aligned}
F &\to V_1D \mid V_1E \mid FF & V_0 &\to 0\\
& & V_1 &\to 1\\
& & D &\to FV_0 \mid 0\\
\$ &\to V_1D \mid V_1E \mid FF \mid \cancel{F} \mid \varepsilon \mid \overbrace{\cancel{V_1D \mid V_1E \mid FF}}^{\text{Doppelt}} & E &\to FV_1 \mid 1
\end{aligned}
$$

$\Rightarrow$

$$F \to V_0 D \mid V_1 E \mid FF$$
$$S \to V_0 D \mid V_1 E \mid FF \mid \varepsilon$$

$$V_0 \to 0$$
$$V_1 \to 1$$
$$D \to F V_0 \mid 0$$
$$E \to F V_1 \mid 1$$

Aufgabe 3 Der CYK-Algorithmus *10 Punkte*

Gegeben sei die folgende Grammatik über dem Alphabet $\Sigma = \{a, b\}$.

$$S \to AB \mid BC, A \to BA \mid a, B \to CC \mid b, C \to AB \mid a.$$

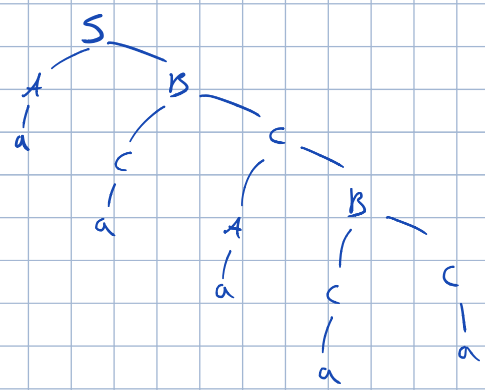
Wenden Sie den CYK-Algorithmus an auf die Wörter $w_1 = aaaaa$, $w_2 = aaaaaa$ und $w_3 = baaba$. Zeichnen Sie gegebenenfalls auch einen zugehörigen Syntaxbaum.

CYK-Alg. wurde nach der Vorlesung durchgeführt

$\Sigma = \{a, b\}$

Produktionen: $S \to AB \mid BC$
$A \to BA \mid a$
$B \to CC \mid b$
$C \to AB \mid a$

	S,A,C				
	B	B			
	S,A,C	S,A,C	S,A,C		
	B	B	B	B	
	A,C	A,C	A,C	A,C	A,C
w_1 =	a	a	a	a	a



	B					
	S,A,C	S,A,C				
	B	B	B			
	S,A,C	S,A,C	S,A,C	S,A,C		
	B	B	B	B	B	
	A,C	A,C	A,C	A,C	A,C	A,C
w_2 =	a	a	a	a	a	a

$\Rightarrow$ Das Wort w_2 kann nicht durch die Grammatik dargestellt werden

	S,A,C				
	–	C,A,S			
	–	B	B		
	S,A	B	C,S	S,A	
	B	A,C	A,C	B	A,C
w_3 =	b	a	a	b	a