Theoretische Informatik 1 – Übungsblatt 4 (SS2006) Lösungen

In den Aufgaben 4.1 und 4.2 bezeichnet G die Grammatik $\langle\{S, A, B\}, \{\underline{\mathsf{a}}, \underline{\mathsf{b}}\}, P, S\rangle$, wobei

$$P \quad = \quad \{S \Rightarrow \varepsilon \mid A\,S\,A,\ A \Rightarrow \underline{\mathsf{a}}\,B \mid B\,\underline{\mathsf{a}},\ B \Rightarrow \varepsilon \mid \underline{\mathsf{b}}\}\ .$$

Aufgabe 4.1 Geben Sie eine Links-, eine Rechts- und eine Parallelableitung für das Wort $\underline{\mathsf{a}}\underline{\mathsf{a}}\underline{\mathsf{b}}\underline{\mathsf{a}}\underline{\mathsf{a}}$ an. Zeigen Sie, dass G mehrdeutig ist.

Lösung Eine Links-, Rechts- bzw. Parallelableitung für $w = \underline{\mathsf{a}}\underline{\mathsf{a}}\underline{\mathsf{b}}\underline{\mathsf{a}}\underline{\mathsf{a}}$ ist:

$$\begin{aligned}
S &\vdash_L A\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,B\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,A\,S\,A\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\,S\,A\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,S\,A\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,A\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\,\underline{\mathsf{a}}\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}
\end{aligned}
\qquad
\begin{aligned}
S &\vdash_R A\,S\,A\\
&\vdash_R A\,S\,\underline{\mathsf{a}}\,B\\
&\vdash_R A\,S\,\underline{\mathsf{a}}\\
&\vdash_R A\,A\,S\,A\,\underline{\mathsf{a}}\\
&\vdash_R A\,A\,S\,B\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R A\,A\,S\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R A\,A\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R A\,\underline{\mathsf{a}}\,B\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R A\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R \underline{\mathsf{a}}\,B\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_R \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}
\end{aligned}
\qquad
\begin{aligned}
S &\vdash_P A\,S\,A\\
&\vdash_P \underline{\mathsf{a}}\,B\,A\,S\,A\,\underline{\mathsf{a}}\,B\\
&\vdash_P \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\,B\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}\\
&\vdash_P \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,\underline{\mathsf{b}}\,\underline{\mathsf{a}}\,\underline{\mathsf{a}}
\end{aligned}$$

Um zu zeigen, dass G mehrdeutig ist, müssen wir ein Wort mit zwei verschiedenen Linksableitungen finden. Wir verwenden das Wort $\underline{\mathsf{a}}\underline{\mathsf{a}}$:

$$\begin{aligned}
S &\vdash_L A\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,B\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}
\end{aligned}
\qquad
\begin{aligned}
S &\vdash_L A\,S\,A\\
&\vdash_L B\,\underline{\mathsf{a}}\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,S\,A\\
&\vdash_L \underline{\mathsf{a}}\,A\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}\,B\\
&\vdash_L \underline{\mathsf{a}}\,\underline{\mathsf{a}}
\end{aligned}$$

Im zweiten Ableitungsschritt wenden wir unterschiedliche Produktionen an, daher handelt es sich um zwei verschiedene Linksableitungen.

Aufgabe 4.2 Ist G bzw. $\mathcal{L}(G)$ kontextfrei? Ist G bzw. $\mathcal{L}(G)$ regulär? Begründen Sie Ihre Antwort.

Lösung Die Grammatik G ist kontextfrei, da alle ihre Produktionen die Form $A \Rightarrow \beta$ besitzen, wobei A ein einzelnes Nonterminal ist. Daher ist auch die dadurch beschriebene Sprache $\mathcal{L}(G)$ kontextfrei.
G ist aber keine reguläre Grammatik, da die Produktionen nicht den entsprechenden Einschränkungen genügen. Die rechte Seite der Produktionen dürfte entweder nur aus ε oder aus einem Terminalsymbol gefolgt von einem Nonterminal bestehen. Das trifft offenbar auf $S \Rightarrow A\,S\,A$ (und andere Produktionen) nicht zu.
$\mathcal{L}(G)$ ist aber regulär: Wir charakterisieren die Sprache, die durch G beschrieben wird, mittels eines regulären Ausdrucks. Um die Diskussion zu erleichtern, erweitern wir den Operator $\mathcal{L}$ auf beliebige Ketten aus Terminal- und Nonterminalsymbolen: $\mathcal{L}(v) = \{w \in T^* \mid v \overset{*}{\vDash} w\}$ für $v \in (V \cup T)^*$. $\mathcal{L}(v)$ ist also die Menge aller Wörter über dem Alphabet T, die aus v mit Hilfe der Produktionen abgeleitet werden können. Insbesondere gilt $\mathcal{L}(G) = \mathcal{L}(S)$, da S die Startvariable der Grammatik G ist.
Wir eliminieren zunächst die Variable B durch Einsetzung:

$$A \quad \Rightarrow \quad \underline{\mathsf{a}}\,(\varepsilon|\underline{\mathsf{b}}) \mid (\varepsilon|\underline{\mathsf{b}})\,\underline{\mathsf{a}}$$

bzw. nach Ausdistribuieren von Verkettung über Vereinigung:

$$A \quad \Rightarrow \quad \underline{\mathtt{a}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{b}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{a}}$$

Wir erhalten $\mathcal{L}(A) = \{\underline{\mathtt{a}}, \underline{\mathtt{a}}\,\underline{\mathtt{b}}, \underline{\mathtt{b}}\,\underline{\mathtt{a}}\}$. Da B von der Startvariablen S aus nicht mehr erreichbar ist, können wir die zugehörigen Produktionen streichen. Bezüglich der S-Produktionen stellen wir fest, dass aus S offenbar eine gerade Anzahl von A's (und sonst nichts) ableitbar ist:

$$S \vdash A\,S\,A \vdash A\,A\,S\,A\,A \vdash \cdots \vdash A^n\,S\,A^n \vdash A^n A^n = A^{2n}$$

Für die Sprache der Grammatik erhalten wir also

$$\begin{aligned}
\mathcal{L}(G) &= \mathcal{L}(S) \\
&= \bigcup_{n=0}^{\infty} \mathcal{L}(A^{2n}) \\
&= \bigcup_{n=0}^{\infty} \mathcal{L}(A)^{2n} \\
&= (\mathcal{L}(A)^2)^* \\
&= (\{\underline{\mathtt{a}}, \underline{\mathtt{a}}\,\underline{\mathtt{b}}, \underline{\mathtt{b}}\,\underline{\mathtt{a}}\}^2)^* \\
&= \{\underline{\mathtt{a}}\,\underline{\mathtt{a}}, \underline{\mathtt{a}}\,\underline{\mathtt{a}}\,\underline{\mathtt{b}}, \underline{\mathtt{a}}\,\underline{\mathtt{b}}\,\underline{\mathtt{a}}, \underline{\mathtt{a}}\,\underline{\mathtt{b}}\,\underline{\mathtt{a}}\,\underline{\mathtt{b}}, \underline{\mathtt{a}}\,\underline{\mathtt{b}}\,\underline{\mathtt{b}}\,\underline{\mathtt{a}}, \underline{\mathtt{b}}\,\underline{\mathtt{a}}\,\underline{\mathtt{a}}, \underline{\mathtt{b}}\,\underline{\mathtt{a}}\,\underline{\mathtt{a}}\,\underline{\mathtt{b}}, \underline{\mathtt{b}}\,\underline{\mathtt{a}}\,\underline{\mathtt{b}}\,\underline{\mathtt{a}}\}^*
\end{aligned}$$

D.h., $\mathcal{L}(G)$ ist regulär.

Aufgabe 4.3 Geben Sie eine kontextfreie Grammatik für die Sprache

$$\{(\underline{\mathtt{a}}\,\underline{\mathtt{b}})^{n_1}(\underline{\mathtt{a}}\,\underline{\mathtt{c}})^{n_2} \mid |n_1 - n_2| \leq 5, n_1, n_2 \geq 0\}$$

an. ($|n|$ bezeichnet den Absolutbetrag der Zahl n.)

Lösung Die Sprache wird durch die kontextfreie Grammatik $\langle \{S, B, C, X\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}\}, P, S \rangle$ generiert, wobei P die Menge der folgenden Produktionen ist:

$$\begin{aligned}
S &\Rightarrow X \mid B\,S\,C \\
X &\Rightarrow B^5 \mid B^4 \mid \cdots \mid B \mid \varepsilon \mid C \mid \cdots \mid C^4 \mid C^5 \\
B &\Rightarrow \underline{\mathtt{a}}\,\underline{\mathtt{b}} \\
C &\Rightarrow \underline{\mathtt{a}}\,\underline{\mathtt{c}}
\end{aligned}$$

Aufgabe 4.4 Eine *Tabelle* im Textsatzystem LaTeX beginnt mit `\begin{tabular}`, gefolgt von einer optionalen Positionsangabe, gefolgt von einer verpflichtenden Spaltenangabe, gefolgt von mehreren Tabellenzeilen (mindestens aber einer), die voneinander durch `\\` getrennt sind, und endet mit `\end{tabular}`. Eine *Positionsangabe* besteht aus dem Zeichen $\underline{\mathtt{b}}$ oder $\underline{\mathtt{t}}$, eingeschlossen in eckigen Klammern. Eine *Spaltenangabe* ist eine nicht-leere Folge der Buchstaben $\underline{\mathtt{c}}$, $\underline{\mathtt{l}}$ und $\underline{\mathtt{r}}$, die in geschwungenen Klammern eingeschlossen ist. Eine *Tabellenzeile* besteht aus einer Folge von Tabelleneinträgen, die voneinander durch das Zeichen $\underline{\mathtt{\&}}$ getrennt sind. Ein *Tabelleneintrag* besteht aus einer möglicherweise leeren Abfolge von Texten und Tabellen. Ein *Text* ist eine nicht-leere Folge von Buchstaben, Ziffern, Leerzeichen, Beistrichen und Punkten.

Beispiel (ohne Unterstreichungen):

```
\begin{tabular}{rrl}
1.Zeile, 1.Spalte rechtsbuendig & 1.Z, 2.Sp rechtsb. & 1.Z, 3.Sp linksb.\\
2.Zeile, 1.Spalte & 2.Z, 2.Sp mit einer \begin{tabular}[t]{c}
                                        Tabelle in der Tabelle
                                        \end{tabular}    & 2.Z, 3.Sp
\end{tabular}
```

Geben Sie für die Sprache der Tabellen eine kontextfreie Grammatik in EBNF an. Verwenden Sie so weit als möglich EBNF-Notationen, um die Grammatik übersichtlich zu halten.

Lösung Eine mögliche Lösung ist $G = \langle V, T, P, \mathit{Tabelle} \rangle$ mit

$$
\begin{array}{rcl}
V & = & \{\, \mathit{Tabelle}, \mathit{Pos}, \mathit{Spalten}, \mathit{Spalte}, \mathit{Zeile}, \mathit{Eintrag}, \mathit{Text}, \mathit{Zeichen} \,\} \ , \\
T & = & \{\underline{\mathtt{a}}, \ldots, \underline{\mathtt{z}}, \underline{\mathtt{A}}, \ldots, \underline{\mathtt{Z}}, \underline{\mathtt{0}}, \ldots, \underline{\mathtt{9}}, \underline{\sqcup}, \underline{\mathtt{,}}, \underline{\mathtt{.}}, \underline{\mathtt{[}}, \underline{\mathtt{]}}, \underline{\mathtt{\{}}, \underline{\mathtt{\}}}, \underline{\mathtt{\backslash}}, \underline{\mathtt{\&}}\} \ , \\
P & = & \{\, \begin{array}[t]{lcl}
\mathit{Tabelle} & \Rightarrow & \underline{\mathtt{\backslash begin\{tabular\}}}\ [\ \mathit{Pos}\ ]\ \mathit{Spalten} \\
 & & \mathit{Zeile}\ \{\ \underline{\mathtt{\backslash\backslash}}\ \mathit{Zeile}\ \} \\
 & & \underline{\mathtt{\backslash end\{tabular\}}} \\
\mathit{Pos} & \Rightarrow & \underline{\mathtt{[}}\ (\ \underline{\mathtt{b}} \mid \underline{\mathtt{t}}\ )\ \underline{\mathtt{]}} \\
\mathit{Spalten} & \Rightarrow & \underline{\mathtt{\{}}\ \mathit{Spalte}\ \{\ \mathit{Spalte}\ \}\ \underline{\mathtt{\}}} \\
\mathit{Spalte} & \Rightarrow & \underline{\mathtt{l}} \mid \underline{\mathtt{c}} \mid \underline{\mathtt{r}} \\
\mathit{Zeile} & \Rightarrow & \mathit{Eintrag}\ \{\ \underline{\mathtt{\&}}\ \mathit{Eintrag}\ \} \\
\mathit{Eintrag} & \Rightarrow & \{\ \mathit{Text} \mid \mathit{Tabelle}\ \} \\
\mathit{Text} & \Rightarrow & \mathit{Zeichen}\ \{\ \mathit{Zeichen}\ \} \\
\mathit{Zeichen} & \Rightarrow & \underline{\mathtt{a}} \mid \cdots \mid \underline{\mathtt{z}} \mid \underline{\mathtt{A}} \mid \cdots \mid \underline{\mathtt{Z}} \mid \underline{\mathtt{0}} \mid \cdots \mid \underline{\mathtt{9}} \mid \underline{\sqcup} \mid \underline{\mathtt{,}} \mid \underline{\mathtt{.}}\ \} \ .
\end{array}
\end{array}
$$

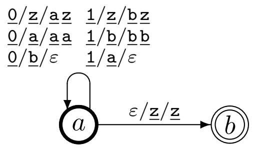
Aufgabe 4.5 Sei $\mathcal{K}$ der Kellerautomat $\mathcal{K} = \langle \{a, b\}, \{\underline{\mathtt{0}}, \underline{\mathtt{1}}\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{z}}\}, \delta, a, \underline{\mathtt{z}}, \{b\} \rangle$, wobei die Übergangsfunktion δ definiert ist durch

$$
\begin{array}{lll}
\delta(a, \underline{\mathtt{0}}, \underline{\mathtt{z}}) = \{(a, \underline{\mathtt{a}}\underline{\mathtt{z}})\} & \delta(a, \underline{\mathtt{1}}, \underline{\mathtt{z}}) = \{(a, \underline{\mathtt{b}}\underline{\mathtt{z}})\} & \delta(a, \varepsilon, \underline{\mathtt{z}}) = \{(b, \underline{\mathtt{z}})\} \\
\delta(a, \underline{\mathtt{0}}, \underline{\mathtt{a}}) = \{(a, \underline{\mathtt{a}}\underline{\mathtt{a}})\} & \delta(a, \underline{\mathtt{1}}, \underline{\mathtt{b}}) = \{(a, \underline{\mathtt{b}}\underline{\mathtt{b}})\} & \\
\delta(a, \underline{\mathtt{0}}, \underline{\mathtt{b}}) = \{(a, \varepsilon)\} & \delta(a, \underline{\mathtt{1}}, \underline{\mathtt{a}}) = \{(a, \varepsilon)\} &
\end{array}
$$

a) Stellen Sie $\mathcal{K}$ graphisch dar.

b) Zeigen Sie, dass $\underline{\mathtt{0}}\underline{\mathtt{0}}\underline{\mathtt{1}}\underline{\mathtt{1}}\underline{\mathtt{1}}\underline{\mathtt{0}}$ in der Sprache $\mathcal{L}(\mathcal{K})$ liegt.

Lösung

a) 

b) Es gilt $(a, \underline{\mathtt{0}}\underline{\mathtt{0}}\underline{\mathtt{1}}\underline{\mathtt{1}}\underline{\mathtt{1}}\underline{\mathtt{0}}, \underline{\mathtt{z}}) \overset{*}{\vdash} (b, \varepsilon, \underline{\mathtt{z}})$:

$$
\begin{array}{lllllll}
(a, \underline{\mathtt{001110}}, \underline{\mathtt{z}}) & \vdash & (a, \underline{\mathtt{01110}}, \underline{\mathtt{az}}) & \vdash & (a, \underline{\mathtt{1110}}, \underline{\mathtt{aaz}}) & \vdash & (a, \underline{\mathtt{110}}, \underline{\mathtt{az}}) \\
 & \vdash & (a, \underline{\mathtt{10}}, \underline{\mathtt{z}}) & \vdash & (a, \underline{\mathtt{0}}, \underline{\mathtt{bz}}) & \vdash & (b, \varepsilon, \underline{\mathtt{z}})
\end{array}
$$

Damit haben wir gezeigt, dass die Anfangskonfiguration $(a, \underline{\mathtt{001110}}, \underline{\mathtt{z}})$ mit Hilfe der Rechenschrittrelation in die Endkonfiguration $(b, \varepsilon, \underline{\mathtt{z}})$ überführt werden kann. Das Wort $\underline{\mathtt{001110}}$ liegt also in $\mathcal{L}(\mathcal{K})$.