Theoretische Informatik und Logik
Übungsblatt 2 (2019W)
Lösungsvorschlag

Aufgabe 2.1 Geben Sie jeweils eine kontextfreie Grammatik an, welche die folgenden Sprachen erzeugt, sowie eine Linksableitung und einen Ableitungsbaum für ein von Ihnen gewähltes Wort $w \in L$ mit $|w| \geq 6$.
Transformieren Sie Ihre jeweils erhaltene Grammatik schrittweise in Chomsky Normalform.

a) $L = \{\underline{\mathtt{a}}^n \underline{\mathtt{b}}^m \mid m \geq n, m - n \text{ ist gerade}\}$

b) $L = \{\underline{\mathtt{c}}^k \underline{\mathtt{d}}^l \mid k \geq 1, l \geq 2k - 1\}$

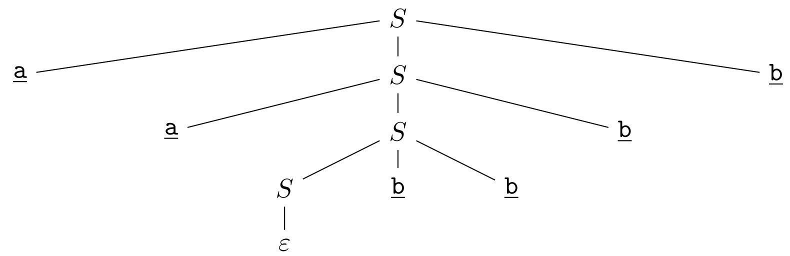
Lösung

a) $G = (\{S\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, \{S \to \underline{\mathtt{a}}S\underline{\mathtt{b}} \mid S\underline{\mathtt{b}}\underline{\mathtt{b}} \mid \varepsilon\}, S)$

Linksableitung für $w = \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}$:

$S \Rightarrow \underline{\mathtt{a}}S\underline{\mathtt{b}} \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}S\underline{\mathtt{b}}\underline{\mathtt{b}} \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}S\underline{\mathtt{b}}^4 \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}^4$

Ableitungsbaum für $w = \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}$:



Transformation von G in Chomsky Normalform:

- **Schritt 1 und 2** *(Elimination nutzloser Produktionen)*: Alle Symbole sind von S aus erreichbar, aus jedem Nonterminal ist ein Terminalwort ableitbar. Es gibt in G also keine nutzlosen Produktionen. Wir erhalten entsprechend $G_1 = G_2 = G$ und es gilt offensichtlich $L(G_1) = L(G_2) = L(G)$.
- **Schritt 3** *(Elimination der ε-Produktionen)*:

  Das Leerwort ist aus dem Startsymbol S ableitbar, also gilt $\varepsilon \in L(G)$!

  Wir entfernen nun die Produktion $S \to \varepsilon$ und ersetzen die rechten Seiten der Produktionen entsprechend (s. Folie 153). Nachdem dadurch keine nutzlosen Produktionen eingeführt werden, erhalten wir

$$\begin{aligned} G_3 &= (\{S\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, P_3, S) \text{ mit} \\ P_3 &= \{S \to \underline{\mathtt{a}}S\underline{\mathtt{b}} \mid \underline{\mathtt{a}}\underline{\mathtt{b}} \mid S\underline{\mathtt{b}}\underline{\mathtt{b}} \mid \underline{\mathtt{b}}\underline{\mathtt{b}}\} \end{aligned}$$

  Und es gilt: $L(G_3) = L(G) - \{\varepsilon\}$.
- **Schritt 4** *(Elimination der Einheits-Produktionen)*: Es gibt keine Einheitsproduktionen in G_3, dementsprechend ist $G_4 = G_3$ mit $L(G_4) = L(G_3) = L(G) - \{\varepsilon\}$:
- **Chomsky Normalform (CNF)**: Um von der reduzierten Grammatik G_4 auf eine äquivalente Grammatik G' in CNF zu kommen, führen wir noch folgende Schritte durch:

* Wir ersetzen die Symbole $\underline{\mathtt{a}}$, $\underline{\mathtt{b}}$ durch entsprechende Nonterminale $X_{\underline{\mathtt{a}}}$, $X_{\underline{\mathtt{b}}}$ und fügen die Produktionen $X_{\underline{\mathtt{a}}} \to \underline{\mathtt{a}}$ bzw. $X_{\underline{\mathtt{b}}} \to \underline{\mathtt{b}}$ hinzu:

$$\begin{aligned} P''' = \{ \quad & S \to X_{\underline{\mathtt{a}}} S X_{\underline{\mathtt{b}}} \mid X_{\underline{\mathtt{a}}} X_{\underline{\mathtt{b}}} \mid S X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}} \mid X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}}, \\ & X_{\underline{\mathtt{a}}} \to \underline{\mathtt{a}}, \quad X_{\underline{\mathtt{b}}} \to \underline{\mathtt{b}} \quad \} \end{aligned}$$

* Nun gibt es noch Produktionen, die auf der rechten Seite mehr als zwei Nonterminale haben. Diese ersetzen wir folgendermaßen:

  $S \to X_{\underline{\mathtt{a}}} S X_{\underline{\mathtt{b}}}$ durch $S \to X_{\underline{\mathtt{a}}} Y_1$, $Y_1 \to S X_{\underline{\mathtt{b}}}$
  $S \to S X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}}$ durch $S \to S Y_2$, $Y_2 \to X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}}$

* Nachdem $\varepsilon \in L(G)$ und S auf der rechten Seite von Produktionen vorkommt, fügen wir ein neues Startsymbol S' sowie die Produktionen $S' \to w$, wobei w die rechte Seite der S-Produktionen ist, und $S' \to \varepsilon$ wieder hinzu, um schließlich die zu G äquivalente Grammatik G' in CNF zu erhalten, mit $L(G') = L(G)$:

  $G' = (\{S', S, X_{\underline{\mathtt{a}}}, X_{\underline{\mathtt{b}}}, , Y_1, Y_2\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, P', S')$ mit

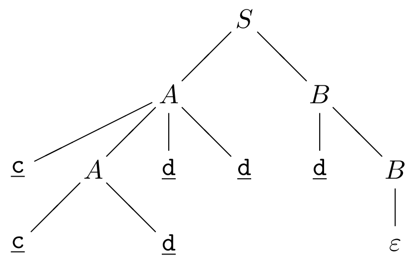
$$\begin{aligned} P' = \{ \quad & S' \to X_{\underline{\mathtt{a}}} Y_1 \mid X_{\underline{\mathtt{a}}} X_{\underline{\mathtt{b}}} \mid S Y_2 \mid X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}} \mid \varepsilon, \\ & S \to X_{\underline{\mathtt{a}}} Y_1 \mid X_{\underline{\mathtt{a}}} X_{\underline{\mathtt{b}}} \mid S Y_2 \mid X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}}, \\ & X_{\underline{\mathtt{a}}} \to \underline{\mathtt{a}}, \quad X_{\underline{\mathtt{b}}} \to \underline{\mathtt{b}}, \\ & Y_1 \to S X_{\underline{\mathtt{b}}}, \quad Y_2 \to X_{\underline{\mathtt{b}}} X_{\underline{\mathtt{b}}} \quad \} \end{aligned}$$

b) $G = (\{S, A, B\}, \{\underline{\mathtt{c}}, \underline{\mathtt{d}}\}, \{S \to AB, A \to \underline{\mathtt{c}}A\underline{\mathtt{d}}\underline{\mathtt{d}} \mid \underline{\mathtt{c}}\underline{\mathtt{d}}, B \to \underline{\mathtt{d}}B \mid \varepsilon\}, S)$

Linksableitung für $w = \underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}$:

$S \Rightarrow AB \Rightarrow \underline{\mathtt{c}}A\underline{\mathtt{d}}\underline{\mathtt{d}}B \Rightarrow \underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}B \Rightarrow \underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}B \Rightarrow \underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}$

Ableitungsbaum für $w = \underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}\underline{\mathtt{d}}$:



Transformation von G in Chomsky Normalform:

- **Schritt 1 und 2** *(Elimination nutzloser Produktionen)*: Alle Symbole sind von S aus erreichbar, aus jedem Nonterminal ist ein Terminalwort ableitbar. Es gibt in G also keine nutzlosen Produktionen. Wir erhalten entsprechend $G_1 = G_2 = G$ und es gilt offensichtlich $L(G_1) = L(G_2) = L(G)$.
- **Schritt 3** *(Elimination der ε-Produktionen)*:

  Das Leerwort ist nur aus dem Nonterminal B ableitbar. (Dementsprechend gilt: $\varepsilon \notin L(G)$)

  Wir entfernen nun die Produktion $B \to \varepsilon$ und ersetzen die rechten Seiten der Produktionen entsprechend (s. Folie 153). Nachdem dadurch keine nutzlosen Produktionen eingeführt werden, erhalten wir

$$\begin{aligned} G_3 &= (\{S, A, B\}, \{\underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P_3, S) \text{ mit} \\ P_3 &= \{S \to AB \mid A, A \to \underline{\mathtt{c}}A\underline{\mathtt{d}}\underline{\mathtt{d}} \mid \underline{\mathtt{c}}\underline{\mathtt{d}}, B \to \underline{\mathtt{d}}B \mid \underline{\mathtt{d}}\} \end{aligned}$$

  Und es gilt: $L(G_3) = L(G)$.

- **Schritt 4** *(Elimination der Einheits-Produktionen)*: Wir entfernen nun die Einheitsproduktion $S \to A$ (wodurch keine nutzlosen Produktionen eingeführt werden) und erhalten G_4 mit $L(G_4) = L(G)$:

  $G_4 = (\{S, A, B\}, \{\underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P_4, S)$ mit

$$
\begin{aligned}
P_4 \quad = \{ \quad & S \to AB \mid \underline{\mathtt{c}}A\underline{\mathtt{d}}\underline{\mathtt{d}} \mid \underline{\mathtt{c}}\underline{\mathtt{d}}, \\
& A \to \underline{\mathtt{c}}A\underline{\mathtt{d}}\underline{\mathtt{d}} \mid \underline{\mathtt{c}}\underline{\mathtt{d}}, \\
& B \to \underline{\mathtt{d}}B \mid \underline{\mathtt{d}}\}
\end{aligned}
$$

- **Chomsky Normalform (CNF)**: Um von der reduzierten Grammatik G_4 auf eine äquivalente Grammatik G' in CNF zu kommen, führen wir noch folgende Schritte durch:
  * Wir ersetzen die Symbole $\underline{\mathtt{c}}$, $\underline{\mathtt{d}}$ überall dort durch entsprechende Nonterminale $X_{\underline{\mathtt{c}}}$, $X_{\underline{\mathtt{d}}}$, wo sie nicht alleine auf der rechten Seite vorkommen und fügen die Produktionen $X_{\underline{\mathtt{c}}} \to \underline{\mathtt{c}}$ und $X_{\underline{\mathtt{d}}} \to \underline{\mathtt{d}}$ hinzu:

$$
\begin{aligned}
P''' \quad = \{ \quad & S \to AB \mid X_{\underline{\mathtt{c}}}AX_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{c}}}X_{\underline{\mathtt{d}}}, \\
& A \to X_{\underline{\mathtt{c}}}AX_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{c}}}X_{\underline{\mathtt{d}}}, \\
& B \to X_{\underline{\mathtt{d}}}B \mid \underline{\mathtt{d}}, \\
& X_{\underline{\mathtt{c}}} \to \underline{\mathtt{c}}, \quad X_{\underline{\mathtt{d}}} \to \underline{\mathtt{d}} \quad \}
\end{aligned}
$$

  * Nun gibt es noch einige Produktionen, die auf der rechten Seite mehr als zwei Nonterminale haben. Diese ersetzen wir folgendermaßen:

$$
\begin{array}{lll}
S \to X_{\underline{\mathtt{c}}}AX_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} & \text{durch} & S \to X_{\underline{\mathtt{c}}}Y_1, \quad Y_1 \to AY_2, \quad Y_2 \to X_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} \\
A \to X_{\underline{\mathtt{c}}}AX_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} & \text{durch} & A \to X_{\underline{\mathtt{c}}}Y_1
\end{array}
$$

  * Nachdem $\varepsilon \notin L(G)$ sind wir bereits fertig, und haben schließlich die zu G äquivalente Grammatik G' in CNF erhalten, mit $L(G') = L(G)$:

    $G' = (\{S, A, B, X_{\underline{\mathtt{c}}}, X_{\underline{\mathtt{d}}}, Y_1, Y_2\}, \{\underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P', S)$ mit

$$
\begin{aligned}
P' \quad = \{ \quad & S \to AB \mid X_{\underline{\mathtt{c}}}Y_1 \mid X_{\underline{\mathtt{c}}}X_{\underline{\mathtt{d}}}, \\
& A \to X_{\underline{\mathtt{c}}}Y_1 \mid X_{\underline{\mathtt{c}}}X_{\underline{\mathtt{d}}}, \\
& B \to X_{\underline{\mathtt{d}}}B \mid \underline{\mathtt{d}}, \\
& X_{\underline{\mathtt{c}}} \to \underline{\mathtt{c}}, \quad X_{\underline{\mathtt{d}}} \to \underline{\mathtt{d}}, \\
& Y_1 \to AY_2, Y_2 \to X_{\underline{\mathtt{d}}}X_{\underline{\mathtt{d}}} \quad \}
\end{aligned}
$$

Aufgabe 2.2 Sind folgende Sprachen L kontextfrei? Falls ja, so beweisen Sie dies mit Hilfe des Satzes von Chomsky-Schützenberger (indem Sie entsprechende Sprachen D_n und R sowie einen entsprechenden Homomorphismus h angeben). Ist hingegen L nicht kontextfrei, so beweisen Sie dies mittels entsprechender Abschlusseigenschaften. (Sie können dabei davon ausgehen, dass eine Sprache der Form $\{\underline{\mathtt{0}}^{kn}\underline{\mathtt{1}}^{ln}\underline{\mathtt{2}}^{mn} \mid n \geq 0\}$ für beliebige (von Ihnen frei wählbare) Konstanten $k, l, m > 0$ als nicht kontextfrei bekannt ist).

a) $L = \{\underline{\mathtt{a}}^i\underline{\mathtt{b}}^k\underline{\mathtt{c}}^j \mid i, j \geq 0, i + j \leq k\}$

b) $L = \{w\underline{\mathtt{c}}^n \mid w \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}^*, |w|_{\underline{\mathtt{a}}} = n \text{ oder } |w|_{\underline{\mathtt{b}}} = n\}$

   (*Hinweis*: $|w|_{\underline{\mathtt{a}}}$ bezeichnet die Anzahl der Symbole $\underline{\mathtt{a}}$ in w)

c) $L = \{\underline{\mathtt{a}}^n\underline{\mathtt{b}}^m\underline{\mathtt{c}}^{n+m}\underline{\mathtt{b}}^m\underline{\mathtt{a}}^n \mid n, m > 0\}$

d) $L = \{\underline{\mathtt{a}}^n\underline{\mathtt{b}}^k\underline{\mathtt{c}}^{4n}\underline{\mathtt{d}}^{3k}\underline{\mathtt{e}}^{3m}\underline{\mathtt{f}}^{3l} \mid n, m, k, l \geq 0\} \cap \{\underline{\mathtt{a}}^m\underline{\mathtt{b}}^k\underline{\mathtt{c}}^{2n}\underline{\mathtt{d}}^{3l}\underline{\mathtt{e}}^{6n}\underline{\mathtt{d}}^{6l} \mid n, m, k, l \geq 0\}$

   (*Hinweis*: Bestimmen Sie zunächst L.)

Lösung

a) **Kontextfrei.** $L = \{\underline{\mathtt{a}}^i\underline{\mathtt{b}}^k\underline{\mathtt{c}}^j \mid i,j \geq 0, i+j \leq k\}$ ist kontextfrei, da $L = h(D_3 \cap R)$, wobei

$$R = \{\underline{[}\}^*\{\underline{]}\}^*\{\underline{(},\underline{)}\}^*\{\underline{\langle}\}^*\{\underline{\rangle}\}^*$$

und

$$h : \{\underline{[},\underline{]},\underline{(},\underline{)},\underline{\langle},\underline{\rangle}\}^* \longrightarrow \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}}\}^* \quad \text{mit} \quad h(\underline{[}) = \underline{\mathtt{a}}, \quad h(\underline{]}) = h(\underline{(}) = h(\underline{\langle}) = \underline{\mathtt{b}}, \quad h(\underline{)}) = \varepsilon, \quad h(\underline{\rangle}) = \underline{\mathtt{c}}$$

b) **Kontextfrei.**
$L = \{w\underline{\mathtt{c}}^n \mid w \in \{\underline{\mathtt{a}},\underline{\mathtt{b}}\}^*, |w|_{\underline{\mathtt{a}}} = n \text{ oder } |w|_{\underline{\mathtt{b}}} = n\}$ ist kontextfrei, da $L = h(D_4 \cap R)$, wobei

$$R = \{\underline{(}_1,\underline{(}_2\}^*\{\underline{)}_1,\underline{)}_2\}^* \cup \{\underline{(}_3,\underline{(}_4\}^*\{\underline{)}_3,\underline{)}_4\}^*$$

und

$$h : \{\underline{(}_1,\underline{)}_1,\underline{(}_2,\underline{)}_2,\underline{(}_3,\underline{)}_3,\underline{(}_4,\underline{)}_4,\}^* \longrightarrow \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}}\}^* \quad \text{mit}$$

$$h(\underline{(}_1) = h(\underline{(}_3) = \underline{\mathtt{a}}, \quad h(\underline{(}_2) = h(\underline{(}_4) = \underline{\mathtt{b}}, \quad h(\underline{)}_1) = h(\underline{)}_4) = \underline{\mathtt{c}}, \quad h(\underline{)}_2) = h(\underline{)}_3) = \varepsilon$$

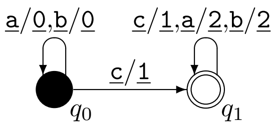
c) **Nicht kontextfrei.**
$L = \{\underline{\mathtt{a}}^n\underline{\mathtt{b}}^m\underline{\mathtt{c}}^{n+m}\underline{\mathtt{b}}^m\underline{\mathtt{a}}^n \mid n,m > 0\}$ ist nicht kontextfrei.

Beweis indirekt. Angenommen, die Sprache L ist kontextfrei. Sei dann

$$M = (\{q_0,q_1\},\{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}}\},\{\underline{\mathtt{0}},\underline{\mathtt{1}},\underline{\mathtt{2}}\},\delta,q_0,\{q_1\})$$

die (deterministische) *gsm* mit

$\delta(q_0,\underline{\mathtt{a}}) = (q_0,\underline{\mathtt{0}})$, $\quad \delta(q_0,\underline{\mathtt{b}}) = (q_0,\underline{\mathtt{0}})$, $\quad \delta(q_0,\underline{\mathtt{c}}) = (q_1,\underline{\mathtt{c}})$,
$\delta(q_1,\underline{\mathtt{a}}) = (q_1,\underline{\mathtt{2}})$, $\quad \delta(q_1,\underline{\mathtt{b}}) = (q_1,\underline{\mathtt{2}})$, $\quad \delta(q_1,\underline{\mathtt{c}}) = (q_1,\underline{\mathtt{1}})$.



Da die Familie der kontextfreien Sprachen gegenüber beliebigen gsm-Abbildungen abgeschlossen ist, müsste auch $M(L) = \{\underline{\mathtt{0}}^n\underline{\mathtt{1}}^n\underline{\mathtt{2}}^n \mid n > 1\}$ kontextfrei sein, was aber nicht der Fall ist. Widerspruch! Somit kann auch L nicht kontextfrei sein.

d) **Nicht kontextfrei.** Wir überlegen zunächst, dass $L = \{\underline{\mathtt{a}}^n\underline{\mathtt{c}}^{4n}\underline{\mathtt{e}}^{12n} \mid n \geq 0\} \cup \{\underline{\mathtt{b}}^{3n}\underline{\mathtt{d}}^{9n} \mid n > 0\}$. L ist also nicht kontextfrei.

Beweis indirekt. Angenommen, L ist kontextfrei. Sei dann

$$h : \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}},\underline{\mathtt{d}},\underline{\mathtt{e}}\}^* \longrightarrow \{\underline{\mathtt{0}},\underline{\mathtt{1}},\underline{\mathtt{2}}\}^*$$

ein Homomorphismus mit

$$h(\underline{\mathtt{a}}) = \underline{\mathtt{0}}, \quad h(\underline{\mathtt{c}}) = \underline{\mathtt{1}}, \quad h(\underline{\mathtt{e}}) = \underline{\mathtt{2}}, \quad h(\underline{\mathtt{b}}) = h(\underline{\mathtt{d}}) = \varepsilon.$$

Da die Familie der kontextfreien Sprachen gegenüber beliebigen Homomorphismen abgeschlossen ist, müsste auch $h(L) = \{\underline{\mathtt{0}}^n\underline{\mathtt{1}}^{4n}\underline{\mathtt{2}}^{12n} \mid n \geq 0\}$ kontextfrei sein, was aber nicht der Fall ist. Widerspruch! Somit kann auch L nicht kontextfrei sein.

Aufgabe 2.3 Sind folgende Aussagen korrekt? Begründen Sie jeweils Ihre Antwort.

a) Es gibt reguläre Ausdrücke R so, dass keine kontextfreie Grammatik G existiert mit $L(G) = L(R)$.

b) Wenn L^* regulär ist, dann ist auch L regulär.

c) Ist L eine nicht-reguläre Sprache, so ist auch L^* nicht-regulär.

d) Wenn A regulär ist, dann ist $A^* - A$ kontextfrei.

e) Sind L_1 und L_2 kontextfrei, so ist auch $L_1 - L_2$ kontextfrei.

f) Die von der Grammatik $G = (\{S, A\}, \{\underline{\mathsf{a}}\}, \{S \to A \mid \underline{\mathsf{a}}, A \to \underline{\mathsf{a}}\}, S)$ erzeugte Sprache ist mehrdeutig.

g) Die von der Grammatik $G = (\{S, A, B\}, \{\underline{\mathsf{a}}\}, \{S \to AB, AB \to S\}, S)$ erzeugte Sprache ist kontextsensitiv.

Lösung

a) **Falsch.** Kann eine Sprache L durch einen regulären Ausdruck beschrieben werden, so ist L regulär. Jede reguläre Sprache ist auch kontextfrei, und kann demnach von einer kontextfreien Grammatik erzeugt werden.

b) **Falsch.** Sei z.B. $L = \{\underline{\mathsf{a}}^p \mid p \text{ ist eine Primzahl}\}$. L ist nicht regulär (siehe z.B. Folie 177). Allerdings gilt: $L^* = \{\varepsilon\} \cup \{\underline{\mathsf{a}}^n \mid n > 1\}$ ist regulär.

c) **Falsch.** siehe b)

d) **Richtig.** Nachdem reguläre Sprachen unter dem Stern-Operator abgeschlossen sind, ist mit A auch A^* regulär. Weiters gilt: $A^* - A = A^* \cap \overline{A}$, was sicher wiederum regulär ist, da reguläre Sprachen auch unter Durchschnitt und Komplement abgeschlossen sind.

e) **Falsch.** Gegenbeispiel: $L_1 = \Sigma^*$. Dann ist $L_1 - L_2$ das Komplement von L_2. Kontextfreie Sprachen sind aber nicht unter Komplement abgeschlossen.

f) **Falsch.** G ist mehrdeutig, da für das Wort $\underline{\mathsf{a}}$ zwei Linksableitungen existieren: $S \Rightarrow A \Rightarrow \underline{\mathsf{a}}$ und $S \Rightarrow \underline{\mathsf{a}}$. Für die von der mehrdeutigen Grammatik G erzeugte Sprache $L(G) = \{\underline{\mathsf{a}}\}$ existiert aber z.B. folgende eindeutige Grammatik: $G = (\{S\}, \{\underline{\mathsf{a}}\}, \{S \to \underline{\mathsf{a}}\}, S)$.

g) **Richtig.** Die Grammatik G ist zwar nicht kontextsensitiv, die von G erzeugte Sprache jedoch schon: $L(G) = \{\}$ ist eine reguläre Sprache, und damit, auf Grund der Chomsky Hierarchie, auch kontextsensitiv.

Aufgabe 2.4 Geben Sie für jede der folgenden Aussagen an, ob diese korrekt ist, oder nicht, und begründen Sie jeweils Ihre Antwort.

a) Aus $A \leq_p SAT$ und $A \in \mathbf{NP}$ folgt, dass auch A **NP**-vollständig ist. (SAT bezeichnet das Satisfiability Problem, siehe Folie 221.)

b) Aus $A \leq_p SAT$ und $SAT \leq_p A$ folgt, dass A **NP**-vollständig ist.

c) Jede Sprache, deren Komplement endlich ist, ist in **NP**.

d) Wenn es eine Sprache $L \in \mathbf{Co} - \mathbf{NP}$ gibt, die **NP**-vollständig ist, dann gilt $\mathbf{Co} - \mathbf{NP} = \mathbf{NP}$.

e) Sei $A \subseteq \Sigma^*$ und $B \subseteq \Gamma^*$. Dann gilt $A \leq_p B$ genau dann wenn $\overline{A} \leq_p \overline{B}$.

f) Jede Sprache, deren Komplement entscheidbar ist, ist in **P**.

g) Ist L in exponentiell beschränkter Zeit von einer deterministischen Turingmaschine entscheidbar, so gilt dies auch für jede Teilmenge von L.

Lösung

a) **Falsch.** Sei $A = \{\}$. Dann ist $A \leq_p SAT$ mittels $w \to x \wedge \neg x$. Aber A ist nicht **NP**-vollständig, da $SAT \not\leq_p A$ ($x \vee \neg x \in SAT$ kann nicht auf ein Wort in A abgebildet werden).

b) **Richtig.** Aus $A \leq_p SAT$ folgt $A \in \mathbf{NP}$ (da $SAT \in \mathbf{NP}$ ist). Da SAT **NP**-hart ist und $SAT \leq_p A$ folgt, dass A **NP**-hart ist. Damit ist A **NP**-vollständig.

c) **Richtig.** Jede Sprache L, deren Komplement endlich ist, ist regulär, und jede reguläre Sprache ist in **P**. Nachdem $\mathbf{P} \subseteq \mathbf{NP}$, ist L damit auch in **NP**.

d) **Richtig.** Sei $L \in \mathbf{Co-NP}$, das heißt, $\overline{L} \in \mathbf{NP}$, und sei L **NP**-vollständig, das heißt, für alle $A \in \mathbf{NP}$ gilt $A \leq_p L$. Dann gilt $\mathbf{NP} \subseteq \mathbf{Co-NP}$, da:
$A \leq_p L$ gilt auch $\overline{A} \leq_p \overline{L}$ (siehe e)). Da $\overline{L} \in \mathbf{NP}$ gilt auch $\overline{A} \in \mathbf{NP}$ und damit ist aber $A \in \mathbf{Co-NP}$.
Umgekehrt gilt aber auch $\mathbf{Co-NP} \subseteq \mathbf{NP}$:
Sei $A \in \mathbf{Co-NP}$, dann ist $\overline{A} \in \mathbf{NP}$ und es gilt $\overline{A} \leq_p L$, da L **NP**-vollständig ist. Damit gilt aber auch $A \leq_p \overline{L}$, und da $\overline{L} \in \mathbf{NP}$ ist auch $A \in \mathbf{NP}$.
Insgesamt ist also unter den angegebenen Voraussetzungen $\mathbf{Co-NP} = \mathbf{NP}$.

(*Anmerkung*: Die Klasse **P** ist abgeschlossen unter Komplementbildung, d.h., es gilt: $\mathbf{P} = \mathbf{Co-P}$. Die Frage hingegen, ob $\mathbf{NP} = \mathbf{Co-NP}$, ist nach wie vor offen.)

e) **Richtig.** Nachdem $A \leq_p B$ gibt es eine Reduktionsfunktion f, welche in polynomieller Zeit berechenbar ist. Daher gilt: $w \in A$ genau dann wenn $f(w) \in B$. Das ist aber äquivalent zu $w \notin A$ genau dann wenn $f(w) \notin B$. Es gilt also auch $\overline{A} \leq_p \overline{B}$.

f) **Falsch.** Jede Sprache, deren Komplement entscheidbar ist, ist selbst entscheidbar, da die rekursiven (entscheidbaren) Sprachen unter Komplementbildung abgeschlossen sind. Allerdings ist nicht jede entscheidbare Sprache in **P**, da **P** eine echte Teilmenge der rekursiven Sprachen ist. ($\mathbf{P} \subseteq \mathcal{L}_{rec}$)

g) **Falsch.** Sei $L = \{\underline{0}, \underline{1}\}^*$. L ist regulär und in **P**, und da $\mathbf{P} \subset \mathbf{EXPTIME}$ also jedenfalls in exponentieller Zeit entscheidbar. Dies gilt allerdings nicht für jede Teilmenge von L. Denn für z.B. das Halteproblem L_u gilt: $L_u \subset L$, aber L_u ist nicht entscheidbar.

Aufgabe 2.5 Das als **NP**-vollständig bekannte Problem CLIQUE ist wie folgt definiert (Folie 227):

Gegeben:	ungerichteter Graph $G = (V, E)$, $k \in \mathbb{N}$
Gefragt:	Enthält G einen vollständigen Teilgraphen mit k Knoten?

Gegeben ist nun das folgende Problem HALF-CLIQUE:

Gegeben:	ungerichteter Graph $G = (V, E)$
Gefragt:	Enthält G einen vollständigen Teilgraphen mit mindestens $\lvert V\rvert/2$ Knoten?

Beweisen Sie mit Hilfe einer entsprechenden Reduktion CLIQUE $\leq_p$ HALF-CLIQUE, dass auch HALF-CLIQUE **NP**-vollständig ist.

Lösung HALF-CLIQUE ist in **NP**:

HALF-CLIQUE ist in **NP**, da für einen Teilgraphen der Größe $\lvert V\rvert/2$ eines ungerichteten Graphen $G = (V, E)$ in polynomiell beschränkter Zeit geprüft werden kann, ob dieser Teilgraph vollständig und damit eine Clique ist.

HALF-CLIQUE ist **NP**-schwer:

Das Problem CLIQUE ist **NP**-vollständig (siehe Folie 227 ff.). Wenn nun CLIQUE $\leq_p$ HALF-CLIQUE gilt, dann folgt, dass HALF-CLIQUE **NP**-schwer, und damit (nachdem es in **NP** liegt) **NP**-vollständig ist.

Beweis der Reduktion CLIQUE $\leq_p$ HALF-CLIQUE:

Dazu sei eine Instanz (G, k) mit $G = (V, E)$ des Problems CLIQUE gegeben. Die Reduktionsfunktion f unterscheidet drei Fälle:

1. Fall: $k = |V|/2$. Dann ist $f(G, k) = G$.

2. Fall: $k > |V|/2$. Dann bildet $f(G, k)$ auf einen Graphen $G' = (V', E')$ ab, der G beinhaltet und noch $2k - |V|$ zusätzliche Knoten hat, die mit keinem anderen Knoten in V' verbunden sind.

3. Fall: $k < |V|/2$. Dann bildet $f(G, k)$ auf einen Graphen G' ab, der G beinhaltet und noch einen zusätzlichen vollständigen Graph mit $|V| - 2k$ Knoten hat. Jeder dieser Knoten hat eine Kante zu allen Knoten aus G.

Es ist offensichtlich, dass diese Operation in polynomieller Zeit berechnet werden kann, da nur Knoten oder Teilgraphen polynomieller Länge eingefügt werden.

Beweis der Korrektheit: G hat eine k-Clique $\Leftrightarrow$ G' hat eine $|V|/2$-Clique:

$\Rightarrow$*:* $(G, k) \in$ CLIQUE.

G besitzt eine Clique der Größe k, damit hat nach Konstruktion G' eine Clique der Größe $|V(G')|/2$, da im 2. Fall $|V(G')| = 2k - |V| + |V| = 2k$ ist und im 3. Fall $|V(G')| = |V| + |V| - 2k = 2(|V| - k)$ und eine Clique der Größe $k + |V| - 2k = |V| - k$ existiert.

$\Leftarrow$*:* $G' \in$ HALF-CLIQUE.

Im ersten und zweiten Fall ist es offensichtlich, dass G nun eine k-Clique besitzt. Für den dritten Fall beobachtet man Folgendes: wenn G' eine $|V(G')|/2$-Clique besitzt, dann existiert eine solche Clique, in der alle zugefügten Knoten beteiligt sind. Hat man eine Clique, die diese unvollständig benutzt, so kann man sie mit den zugefügten Knoten zu einer größeren Clique ergänzen. In jedem Fall besitzt der ursprüngliche Graph dann mindestens eine k-Clique.