Theoretische Informatik und Logik
Übungsblatt 2 (2017W)
Lösungen

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 die jeweils erhaltene Grammatik schrittweise in Chomsky Normalform.

a) Die Łukasiewicz-Sprache[1] L über dem Alphabet $\{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}$ ist die kleinste Menge, für die gilt:
   - $\underline{\mathtt{b}} \in L$
   - sind $u, w \in L$ so ist auch $\underline{\mathtt{a}}uw \in L$.

b) $L = \{\underline{\mathtt{a}}^m \underline{\mathtt{b}}^n \underline{\mathtt{c}}^p \underline{\mathtt{d}}^q \mid m, n, p, q \geq 0, m + n = p + q\}$

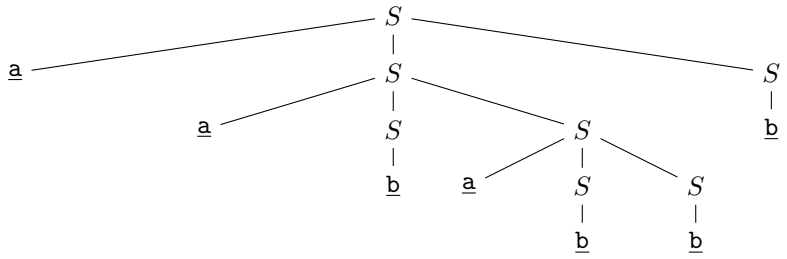
Lösung

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

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

   $S \Rightarrow \underline{\mathtt{a}}SS \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}SSS \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}SS \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{a}}SSS \Rightarrow^3 \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}$

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



   G ist bereits reduziert, wir erhalten nach Ersetzung von Terminalsymbolen und Produktionen mit mehr als drei Nonterminalen auf der rechten Seite folgende Grammatik in CNF:

$$(\{S, X, Y\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, \{S \to XY \mid \underline{\mathtt{b}}, Y \to SS, X \to \underline{\mathtt{a}}\}, S)$$

b) $G = (\{S, T, U, V\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P, S)$, wobei

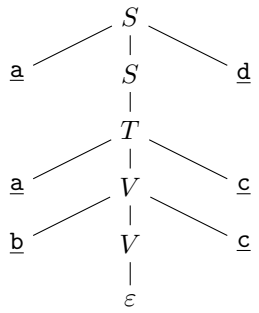
   $P = \{S \to \underline{\mathtt{a}}S\underline{\mathtt{d}} \mid T \mid U, \quad T \to \underline{\mathtt{a}}T\underline{\mathtt{c}} \mid V, \quad U \to \underline{\mathtt{b}}U\underline{\mathtt{d}} \mid V, \quad V \to \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \varepsilon\}$

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

   $S \Rightarrow \underline{\mathtt{a}}S\underline{\mathtt{d}} \Rightarrow \underline{\mathtt{a}}T\underline{\mathtt{d}} \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}V\underline{\mathtt{c}}\underline{\mathtt{d}} \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}V\underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}} \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{c}}\underline{\mathtt{c}}\underline{\mathtt{d}}$

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

[1] Die Łukasiewicz-Sprache beschreibt arithmetische Ausdrücke in der polnischen Notation, wobei das Symbol $\underline{\mathtt{a}}$ für einen zweistelligen Operator und das Symbol $\underline{\mathtt{b}}$ für eine arithmetische Variable bzw. eine Konstante stehen. Dem Wort $\underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{a}}\underline{\mathtt{a}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}\underline{\mathtt{b}}$ entspricht z.B. der Ausdruck $+*2*+3456$ in der polnischen Notation bzw. $(2*((3+4)*5))+6$ in der gewohnten Infix-Notation.



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)*:

  Wir bestimmen zunächst die Menge der Nonterminale, aus denen das Leerwort ableitbar ist:
  * da $V \to \varepsilon$ gilt $N_3^{(1)} = \{V\}$
  * nachdem $T \Rightarrow V \Rightarrow \varepsilon$ und $U \Rightarrow V \Rightarrow \varepsilon$ erhalten wir $N_3^{(2)} = \{U, T, V\}$
  * es gilt aber z.B. auch $S \Rightarrow T \Rightarrow V \Rightarrow \varepsilon$, daher ist $N_3^{(3)} = \{S, U, T, V\} = N$

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

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

  $$\begin{aligned} G_3 &= (\{S,T,U,V\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P_3, S) \text{ mit} \\ P_3 &= \{S \to \underline{\mathtt{a}}S\underline{\mathtt{d}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{d}} \mid T \mid U, \quad T \to \underline{\mathtt{a}}T\underline{\mathtt{c}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{c}} \mid V \quad U \to \underline{\mathtt{b}}U\underline{\mathtt{d}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{d}} \mid V, \quad V \to \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{c}}\} \end{aligned}$$

  Und es gilt: $L(G_3) = L(G) - \{\varepsilon\}$.
- **Schritt 4** *(Elimination der Einheits-Produktionen)*: Wir entfernen nun die Einheitsproduktionen $S \to T$, $S \to U$, $T \to V$ und $U \to V$ (wodurch keine nutzlosen Produktionen eingeführt werden) und erhalten G_4 mit $L(G_4) = L(G) - \{\varepsilon\}$:

  $G_4 = (\{S,T,U,V\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}, P_4, S)$ mit

  $$\begin{aligned} P_4 = \{ \quad & S \to \underline{\mathtt{a}}S\underline{\mathtt{d}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{d}} \mid \underline{\mathtt{a}}T\underline{\mathtt{c}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{c}} \mid \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{c}} \mid \underline{\mathtt{b}}U\underline{\mathtt{d}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{d}}, \\ & T \to \underline{\mathtt{a}}T\underline{\mathtt{c}} \mid \underline{\mathtt{a}}\,\underline{\mathtt{c}} \mid \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{c}}, \\ & U \to \underline{\mathtt{b}}U\underline{\mathtt{d}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{d}} \mid \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{c}}, \\ & V \to \underline{\mathtt{b}}V\underline{\mathtt{c}} \mid \underline{\mathtt{b}}\,\underline{\mathtt{c}} \quad \} \end{aligned}$$

- **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 $\underline{\mathtt{a}}$, $\underline{\mathtt{b}}$, $\underline{\mathtt{c}}$, $\underline{\mathtt{d}}$ durch entsprechende Nonterminale $X_{\underline{\mathtt{a}}}$, $X_{\underline{\mathtt{b}}}$, $X_{\underline{\mathtt{c}}}$, $X_{\underline{\mathtt{d}}}$ und fügen die Produktionen $X_{\underline{\mathtt{a}}} \to \underline{\mathtt{a}}$ etc. hinzu:

    $$\begin{aligned} P''' = \{ \quad & S \to X_{\underline{\mathtt{a}}}SX_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{a}}}X_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{a}}}TX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{a}}}X_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}VX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}UX_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{d}}}, \\ & T \to X_{\underline{\mathtt{a}}}TX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{a}}}X_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}VX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{c}}}, \\ & U \to X_{\underline{\mathtt{b}}}UX_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{d}}} \mid X_{\underline{\mathtt{b}}}VX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{c}}}, \\ & V \to X_{\underline{\mathtt{b}}}VX_{\underline{\mathtt{c}}} \mid X_{\underline{\mathtt{b}}}X_{\underline{\mathtt{c}}}, \\ & X_{\underline{\mathtt{a}}} \to \underline{\mathtt{a}}, \quad X_{\underline{\mathtt{b}}} \to \underline{\mathtt{b}}, \quad X_{\underline{\mathtt{c}}} \to \underline{\mathtt{c}}, \quad X_{\underline{\mathtt{d}}} \to \underline{\mathtt{d}} \quad \} \end{aligned}$$

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

$$\begin{array}{lll}
S \to X_{\underline{a}} S X_{\underline{d}} & \text{durch} & S \to X_{\underline{a}} Y_1, \quad Y_1 \to S X_{\underline{d}} \\
S \to X_{\underline{a}} T X_{\underline{c}} & \text{durch} & S \to X_{\underline{a}} Y_2, \quad Y_2 \to T X_{\underline{c}} \\
S \to X_{\underline{b}} V X_{\underline{c}} & \text{durch} & S \to X_{\underline{b}} Y_3, \quad Y_3 \to V X_{\underline{c}} \\
S \to X_{\underline{b}} U X_{\underline{d}} & \text{durch} & S \to X_{\underline{b}} Y_4, \quad Y_4 \to U X_{\underline{d}} \\
T \to X_{\underline{a}} T X_{\underline{c}} & \text{durch} & B \to X_{\underline{b}} Y_2 \\
T \to X_{\underline{b}} V X_{\underline{c}} & \text{durch} & T \to X_{\underline{b}} Y_3 \\
U \to X_{\underline{b}} V X_{\underline{d}} & \text{durch} & B \to X_{\underline{b}} Y_5, \quad Y_5 \to V X_{\underline{d}} \\
U \to X_{\underline{b}} V X_{\underline{c}} & \text{durch} & U \to X_{\underline{b}} Y_3 \\
V \to X_{\underline{b}} V X_{\underline{c}} & \text{durch} & V \to X_{\underline{b}} Y_3
\end{array}$$

* 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 $L(G') = L(G)$:

  $G' = (\{S', S, T, U, V, X_{\underline{a}}, X_{\underline{b}}, X_{\underline{c}}, X_{\underline{d}}, Y_1, Y_2, Y_3, Y_4, Y_5\}, \{\underline{a}, \underline{b}, \underline{c}, \underline{d}\}, P', S')$ mit

$$\begin{aligned}
P' \quad = \{ \quad & S' \to X_{\underline{a}} Y_1 \mid X_{\underline{a}} X_{\underline{d}} \mid X_{\underline{a}} Y_2 \mid X_{\underline{a}} X_{\underline{c}} \mid X_{\underline{b}} Y_3 \mid X_{\underline{b}} X_{\underline{c}} \mid X_{\underline{b}} Y_4 \mid X_{\underline{b}} X_{\underline{d}} \mid \varepsilon, \\
& S \to X_{\underline{a}} Y_1 \mid X_{\underline{a}} X_{\underline{d}} \mid X_{\underline{a}} Y_2 \mid X_{\underline{a}} X_{\underline{c}} \mid X_{\underline{b}} Y_3 \mid X_{\underline{b}} X_{\underline{c}} \mid X_{\underline{b}} Y_4 \mid X_{\underline{b}} X_{\underline{d}}, \\
& T \to X_{\underline{b}} Y_2 \mid X_{\underline{a}} X_{\underline{c}} \mid X_{\underline{b}} Y_3 \mid X_{\underline{b}} X_{\underline{c}}, \\
& U \to X_{\underline{b}} Y_5 \mid X_{\underline{b}} X_{\underline{d}} \mid X_{\underline{b}} Y_3 \mid X_{\underline{b}} X_{\underline{c}}, \\
& V \to X_{\underline{b}} Y_3 \mid X_{\underline{b}} X_{\underline{c}}, \\
& X_{\underline{a}} \to \underline{a}, \quad X_{\underline{b}} \to \underline{b}, \quad X_{\underline{c}} \to \underline{c}, \quad X_{\underline{d}} \to \underline{d}, \\
& Y_1 \to S X_{\underline{d}}, \quad Y_2 \to T X_{\underline{c}}, \quad Y_3 \to V X_{\underline{c}}, \quad Y_4 \to U X_{\underline{d}}, \quad Y_5 \to V X_{\underline{d}} \quad \}
\end{aligned}$$

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). Geben Sie in diesem Fall auch eine kontextfreie Grammatik an, welche L erzeugt. Ist hingegen L nicht kontextfrei, so beweisen Sie dies mittels einer Methode Ihrer Wahl. (Sie können dabei davon ausgehen, dass eine Sprache der Form $\{\underline{a}^{kn}\underline{b}^{ln}\underline{c}^{mn} \mid n \geq 0\}$ für beliebige Konstanten $k, l, m > 0$ als nicht kontextfrei bekannt ist).

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

b) $L = \{(\underline{0}^n \underline{1}^n)^n \mid n \geq 2\}$

c) $L = \{ww^r \mid w \in \{\underline{a}, \underline{b}, \underline{c}\}^*\} \cap \{\underline{a}^n \underline{b}^k \underline{a}^n \mid k, n \geq 0\}$
(*Hinweis*: Bestimmen Sie zunächst L.)

d) $L = \{\underline{a}^{4n} \underline{b}^{4n} \underline{c}^{2m} \underline{b}^k \underline{d}^m \mid n, m, k \geq 0\} \cap \{\underline{a}^{2n} \underline{b}^{2m} \underline{c}^{4m} \underline{b}^{2017} \underline{d}^{2n} \mid n, m \geq 0\}$
(*Hinweis*: Bestimmen Sie zunächst L.)

Lösung

a) $L = \{\underline{a}^n \underline{b}^m \mid m \geq n \text{ und } m - n \text{ ist gerade}\}$ ist kontextfrei, da $L = h(D_2 \cap R)$, wobei

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

und

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

z.B. folgende kontextfreie Grammatik erzeugt L:

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

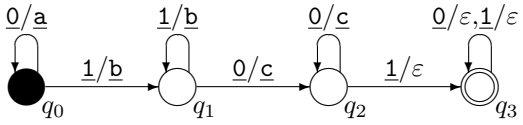
b) $L = \{(\underline{0}^n\underline{1}^n)^n \mid n \geq 2\}$ ist nicht kontextfrei.

Beweis indirekt. Angenommen, die Sprache $L = \{(\underline{0}^n\underline{1}^n)^n \mid n \geq 2\}$ ist kontextfrei. Sei dann

$$M = (\{q_0, q_1, q_2, q_3\}, \{\underline{0}, \underline{1}\}, \{\underline{a}, \underline{b}, \underline{c}\}, \delta, q_0, \{q_3\})$$

die (deterministische) *gsm* mit

$\delta(q_0, \underline{0}) = (q_0, \underline{a})$, $\delta(q_0, \underline{1}) = (q_1, \underline{b})$, $\delta(q_1, \underline{b}) = (q_1, \underline{b})$, $\delta(q_1, \underline{0}) = (q_2, \underline{c})$,
$\delta(q_2, \underline{0}) = (q_2, \underline{c})$, $\delta(q_2, \underline{1}) = (q_3, \varepsilon)$, $\delta(q_3, \underline{0}) = \delta(q_3, \underline{1}) = (q_3, \varepsilon)$.



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

(*Anmerkung*: Man beachte, dass in diesem Falle die Verwendung eines Homomorphismus nicht ausreicht, da der erste Block von Symbolen $\underline{0}$ auf Symbole $\underline{a}$ abgebildet werden muss, während der zweite Block auf eine entsprechende Anzahl von Symbolen $\underline{c}$ abgebildet werden muss.)

c) Wir überlegen zunächst, dass $L = \{a^n b^{2m} a^n \mid n, m \geq 0\}$, also eine kontextfreie Sprache ist, da sie sich als $L = h(D_2 \cap R)$ darstellen lässt, wobei:

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

und

$$h : \{\underline{[}, \underline{]}, \underline{(}, \underline{)}\}^* \longrightarrow \{\underline{a}, \underline{b}\}^* \text{ mit}$$

$$h(\underline{(}) = h(\underline{)}) = \underline{a}, \quad h(\underline{[}) = h(\underline{]}) = \underline{b}$$

Eine kontextfreie Grammatik für L ist z.B.:

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

d) $L = \{\underline{a}^{4n}\underline{b}^{4n}\underline{c}^{8n}\underline{b}^{2017}\underline{d}^{4n} \mid n \geq 0\}$ ist nicht kontextfrei.

Beweis indirekt. Angenommen, $L = \{\underline{a}^{4n}\underline{b}^{4n}\underline{c}^{8n}\underline{b}^{2017}\underline{d}^{4n} \mid n \geq 0\}$ ist kontextfrei. Sei dann $h : \{\underline{a}, \underline{b}, \underline{c}, \underline{d}\}^* \longrightarrow \{\underline{a}, \underline{b}, \underline{c}\}^*$ ein Homomorphismus mit

$h(\underline{a}) = \underline{a}$, $h(\underline{b}) = \varepsilon$, $h(\underline{c}) = \underline{b}$, $h(\underline{d}) = \underline{c}$.

Da die Familie der kontextfreien Sprachen gegenüber beliebigen Homomorphismen abgeschlossen ist, müsste auch $h(L) = \{\underline{a}^{4n}\underline{b}^{8n}\underline{c}^{4n} \mid n \geq 1\}$ 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) Sei L kontextfrei und R regulär. Dann ist $R - L$ kontextfrei.

b) Sei L kontextfrei und R regulär. Dann ist $L - R$ kontextfrei.

c) Eine kontextfreie Grammatik in Chomsky Normalform ist immer eindeutig.

d) Seien $L_1, ...L_k$ Sprachen über einem Alphabet Σ so, dass:
   - für alle $i \neq j$ gilt: $L_i \cap L_j = \{\}$

- $L_1 \cup L_2 \cup ... \cup L_k = \Sigma^*$
- jede dieser Sprachen $L_i, 1 \leq i \leq k$, rekursiv aufzählbar ist.

Dann ist jede dieser Sprachen L_i rekursiv (entscheidbar).

e) Sei $\Sigma = \{\underline{1}\}$. Dann gibt es mindestens eine unentscheidbare Sprache über Σ.

Lösung

a) Falsch. Sei $R = \Sigma^*$, dann ist $R - L$ genau das Komplement von L. Kontextfreie Sprachen sind aber unter Komplement nicht abgeschlossen. Ein konkretes Gegenbeispiel: Sei $R = \{a\}^*\{b\}^*\{c\}^*$ und $L = \{a^i b^j c^k \mid i \neq j \text{ oder } j \neq k\}$. Dann ist $R - L = \{a^n b^n c^n \mid n \geq 0\}$, welches keine kontextfreie Sprache ist.

b) Richtig. Denn es gilt: $L - R = L \cap \overline{R}$. Kontextfreie Sprachen sind abgeschlossen unter Durchschnitt mit regulären Mengen, und reguläre Sprachen sind abgeschlossen unter Komplement.

c) Falsch. Jede kontextfreie Grammatik kann in Chomsky Normalform konvertiert werden. Es gibt aber inhärent mehrdeutige kontextfreie Sprachen, also solche, für die es keine eindeutige Grammatik gibt.

d) Richtig. Für alle i gilt: $\overline{L_i} = L_1 \cup L_2 \cup ... \cup L_{i-1} \cup L_{i+1} \cup ... \cup L_k$. Jede der Sprachen $L_j, 1 \leq j \leq k$, ist rekursiv aufzählbar, also ist auch die Vereinigung dieser Sprachen rekursiv aufzählbar. ($\mathcal{L}_0$ ist abgeschlossen unter Vereinigung.) Nachdem also die Sprache L_i wie auch ihr Komplement rekursiv aufzählbar sind, ist L_i jedenfalls rekursiv (entscheidbar).

e) Richtig. Σ^* ist abzählbar (unendlich), die Menge aller Sprachen $L \subseteq \Sigma^*$ ist überabzählbar. Von diesen ist jedoch nur eine abzählbare Menge entscheidbar. (Es gibt ja nur abzählbar (unendlich) viele Turingmaschinen). Dementsprechend gibt es unentscheidbare Sprachen über Σ.

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

a) Sei $A \leq_p B$. Dann gilt: Wenn B **NP**-vollständig ist, so ist auch A **NP**-vollständig.

b) Sei $A, B \subseteq \Sigma^*$ und $A \leq_p B$. Dann gilt auch $\overline{A} \leq_p \overline{B}$.

c) Angenommen, A ist **NP**-vollständig und $\overline{A} \in \mathbf{NP}$. Dann gilt $\overline{L} \in \mathbf{NP}$ für jede Sprache $L \in \mathbf{NP}$.

d) Das Problem A sei **NP**-hart. Dann gilt: wenn $\mathbf{P} = \mathbf{NP}$ ist, dann kann A in polynomieller Zeit gelöst werden.

e) Wenn $A \leq_p B$ und B und C in **NP** sind, dann ist auch $A \cup C \in \mathbf{NP}$.

Lösung

a) Das muss nicht sein. Auf Grund der Reduktion ergibt sich nur, dass A jedenfalls in **NP** liegen muss. Es muss deshalb aber nicht notwendigerweise **NP**-vollständig sein.

b) 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 $\overline{A} \leq_p \overline{B}$.

c) Nachdem A **NP**-vollständig ist, muss es auch **NP**-hart sein. Das bedeutet für alle Sprachen $L \in \mathbf{NP}$ gibt es eine Reduktion $L \leq_p A$. Das würde dann aber auch bedeuten, dass $\overline{L} \leq_p \overline{A}$ (siehe b)). Nachdem laut Annahme $\overline{A} \in \mathbf{NP}$, wäre nun auch für alle Sprachen $L \in \mathbf{NP}$ auch ihr Komplement in **NP**.
(*Anmerkung*: Es ist nach wie vor ein offenes Problem, ob **NP=co-NP**, also ob **NP** unter Komplement abgeschlossen ist.)

d) Das muss nicht sein. Wir wissen ja nur, dass A **NP**-hart ist, es könnte also auch außerhalb von **NP** liegen.

e) Korrekt. Nachdem $B \in \mathbf{NP}$ und $A \leq_p B$ muss auch A in **NP** sein. Die Klasse **NP** ist abgeschlossen unter Vereinigung: Seien L_1 und L_2 Sprachen in **NP**. Dann gibt es nichtdeterministische Turingmaschinen M_1 und M_2, die L_1 und L_2 in polynomieller Zeit entscheiden. Daraus kann leicht eine NTM M konstruiert werden, welche $L_1 \cup L_2$ in polynomieller Zeit entscheidet: Auf dem Input w simuliert M zunächst M_1 und akzeptiert, wenn M_1 akzeptiert. Ansonsten simuliert M die Maschine M_2. Akzeptiert diese die Eingabe w, so akzeptiert auch M. In jedem anderen Fall verwirft M die Eingabe w.

   Nachdem also A und C in **NP** sind gilt auch $A \cup C \in \mathbf{NP}$.

Aufgabe 2.5 Sie wollen zur Festzeit Ihr Haus mit bunten Lampen dekorieren. Dazu haben Sie Ihren guten Freund, einen Elektrotechnikstudenten, um Hilfe gebeten und nun ist die Beleuchtung derart kompliziert, dass Sie Schwierigkeiten haben, sie überhaupt anzuschalten.
Ein *Boole'sches Schaltnetz* besteht aus ein einem gerichteten, azyklischen Graphen mit Eingangsknoten $x_1, x_2, ..., x_n$ sowie einem Ausgangsknoten x_{out}. Die Eingangsknoten haben dabei nur eine ausgehende Kante und der Ausgangsknoten nur eine eingehende Kante. Alle weiteren Knoten sind markiert als ODER-, UND- oder NICHT-Gatter, darüberhinaus sind auch Verzweigungen möglich.

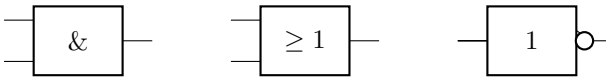
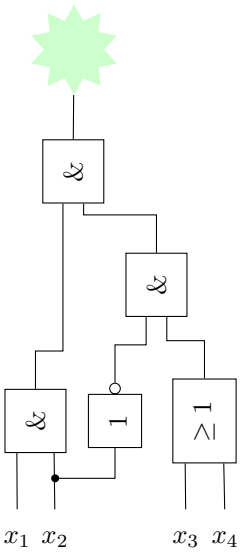


Abbildung 1: UND-, ODER- und NICHT-Gatter

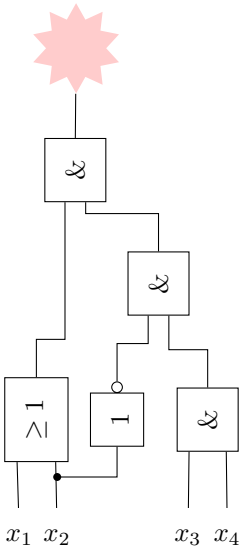
Für das Entscheidungsproblem *BOOLEAN CIRCUIT* ist ein solches Schaltnetz gegeben und die zugehörige Frage ist, ob es dafür eine erfüllende Belegung gibt.

a) Die folgenden Schaltungen hat Ihr Freund in die Beleuchtung verbaut, mit der Behauptung, dass die jeweilige Lichterkette angeht, wenn man die Schalter $x_1, x_2, ..., x_n$ in die richtige Stellung bringt. Ist dies tatsächlich für beide Schaltungen möglich? Wenn ja, geben Sie eine Schalterstellung an, die Ihr Haus festlich beleuchtet.

   Schaltung 1:



   Schaltung 2:



b) Reduzieren Sie das oben beschriebene Problem *BOOLEAN CIRCUIT* auf das Problem *SAT* (s. Folie 221), d.h., zeigen Sie $BOOLEANCIRCUIT \leq_p SAT$.

(Denken Sie dabei an Tannengeruch, Sternspritzer, oder einen duftenden Truthahn. ☺)

Lösung

a) Schaltung 1 ist unerfüllbar, Schaltung 2 ist erfüllbar durch $x_1 = x_3 = x_4 = 1, x_2 = 0$.

b) Um $BOOLEANCIRCUIT \leq_p SAT$ zu zeigen müssen wir in polynomieller Zeit aus einem Boole'schen Schaltnetz C eine Formel F so bestimmen, dass gilt:

$$C \text{ erfüllbar } \iff F \text{ erfüllbar}$$

Wir stellen nun jedes AND-, OR- und NOT-Gatter als eine Aussagenlogische Formel dar, um daraus dann eine Formel für die gesamte Schaltung zu erstellen.

Dazu ordnen wir zunächst jedem Ausgang eines Gatters eine neue Variable $y_1, y_2, ...$ zu.

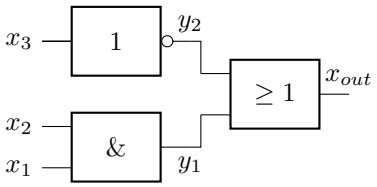
Für jedes Gatter bilden wir eine Formel und betrachten anschließend deren Konjunktion (zusammen mit der Formel x_{out}):

- ist y_i die Beschriftung des Ausgangs eines NOT-Gatters und u jene des Eingangs, so betrachten wir die Formel $y_i \leftrightarrow \neg u$.
- ist y_i die Beschriftung des Ausgangs eines AND- bzw. OR-Gatters und sind u, v die Beschriftungen der Eingänge, so betrachten wir die Formel $y_i \leftrightarrow u \wedge v$ bzw. $y_i \leftrightarrow u \vee v$

Beachte: die Konstruktion der Formel F ist in polynomieller Zeit möglich.

(Zur Veranschaulichung ein kleines Beispiel:

Aus folgender Schaltung C



erhalten wir nach obiger Konstruktion die Formel

$$F = (y_1 \leftrightarrow x_1 \wedge x_2) \wedge (y_2 \leftrightarrow \neg x_3) \wedge (x_{out} \leftrightarrow y_1 \vee y_2) \wedge x_{out}$$

)

Darüber hinaus gilt: Die so konstruierte Formel F ist genau dann erfüllbar, wenn C erfüllbar ist.

(Denn: Hat C eine erfüllende Belegung, dann ist jede Verbindung wohldefiniert, und der Output der Schaltung ist 1. Damit ist aber auch jedes Konjunkt in F gleich 1, und F wird insgesamt zu 1 evaluieren.

Auf der anderen Seite führt aber auch eine erfüllende Belegung von F zu einer Schaltung C mit Output 1.) □