Theoretische Informatik und Logik
Übungsblatt 2 (2016S)
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 $|w| \geq 6$. Geben Sie weiters jeweils einen Homomorphismus h so an, dass $h(L) = \{\underline{\mathtt{0}}^{2n}\underline{\mathtt{1}}^{4n} \mid n \geq 0\}$.

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

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

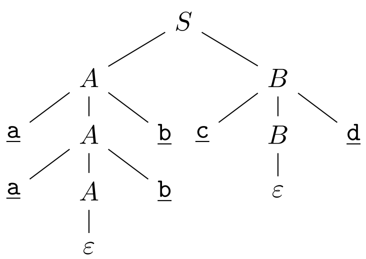
Lösung

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

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

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

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



Z.B. Folgender Homomorphismus $h : \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}},\underline{\mathtt{d}}\}^* \longrightarrow \{\underline{\mathtt{0}},\underline{\mathtt{1}}\}^*$ bildet $L = \{\underline{\mathtt{a}}^i\underline{\mathtt{b}}^i\underline{\mathtt{c}}^j\underline{\mathtt{d}}^j \mid i,j \geq 0\}$ auf $h(L) = \{\underline{\mathtt{0}}^{2n}\underline{\mathtt{1}}^{4n} \mid n \geq 0\}$ ab:

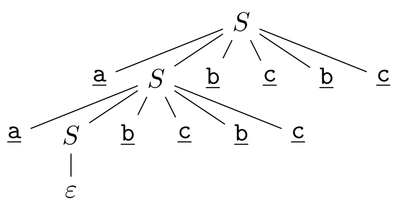
$$h(\underline{\mathtt{a}}) = \underline{\mathtt{0}}^2, \quad h(\underline{\mathtt{b}}) = \underline{\mathtt{1}}^4, \quad h(\underline{\mathtt{c}}) = h(\underline{\mathtt{d}}) = \varepsilon$$

b) $G = (\{S\}, \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}}\}, \{S \to \underline{\mathtt{a}}S(\underline{\mathtt{b}}\underline{\mathtt{c}})^2 \mid \varepsilon\}, S)$

Linksableitung für $w = \underline{\mathtt{a}}\underline{\mathtt{a}}(\underline{\mathtt{b}}\underline{\mathtt{c}})^4$:

$S \Rightarrow \underline{\mathtt{a}}S(\underline{\mathtt{b}}\underline{\mathtt{c}})^2 \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}S(\underline{\mathtt{b}}\underline{\mathtt{c}})^4 \Rightarrow \underline{\mathtt{a}}\underline{\mathtt{a}}(\underline{\mathtt{b}}\underline{\mathtt{c}})^4$

Ableitungsbaum für $w = \underline{\mathtt{a}}\underline{\mathtt{a}}(\underline{\mathtt{b}}\underline{\mathtt{c}})^4$:



Der Homomorphismus $h : \{\underline{\mathtt{a}},\underline{\mathtt{b}},\underline{\mathtt{c}}\}^* \longrightarrow \{\underline{\mathtt{0}},\underline{\mathtt{1}}\}^*$ mit z.B.

$$h(\underline{\mathtt{a}}) = \underline{\mathtt{0}}^2, \quad h(\underline{\mathtt{b}}) = h(\underline{\mathtt{c}}) = \underline{\mathtt{1}}$$

bildet $L = \{\underline{\mathtt{a}}^n(\underline{\mathtt{b}}\underline{\mathtt{c}})^{2n} \mid n \geq 0\}$ auf $h(L) = \{\underline{\mathtt{0}}^{2n}\underline{\mathtt{1}}^{4n} \mid n \geq 0\}$ ab.

Aufgabe 2.2 Sind folgende Sprachen 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), falls nein, so beweisen Sie dies mit Hilfe des Pumping Lemmas für kontextfreie Sprachen.

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

b) $L = \{w\underline{\#}w^r \mid w \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}^*\}$

c) $L = \{ww \mid w \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}^*\}$

Lösung

a) $L = \{\underline{\mathtt{a}}^i\underline{\mathtt{b}}^j\underline{\mathtt{c}}^j\underline{\mathtt{d}}^i \mid i, j \geq 0\}$ ist kontextfrei, da $L = h(D_2 \cap R)$, wobei

$$R = \{\underline{[_1}\}^*\{\underline{[_2}\}^*\{\underline{]_2}\}^*\{\underline{]_1}\}^*$$

und

$$h : \{\underline{[_1}, \underline{]_1}, \underline{[_2}, \underline{]_2}\}^* \longrightarrow \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}^* \quad \text{mit} \quad h(\underline{[_1}) = \underline{\mathtt{a}}, h(\underline{]_1}) = \underline{\mathtt{d}}, h(\underline{[_2}) = \underline{\mathtt{b}}, h(\underline{]_2}) = \underline{\mathtt{c}}$$

b) $L = \{w\underline{\#}w^r \mid w \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}^*\}$ ist kontextfrei, da $L = h(D_5 \cap R)$, wobei

$$R = \{\underline{[_1}, \underline{[_2}, \underline{[_3}, \underline{[_4}\}^*\{\underline{[_5}\}\{\underline{]_5}\}\{\underline{]_1}, \underline{]_2}, \underline{]_3}, \underline{]_4}\}^*$$

und

$$h : \{\underline{[_i}, \underline{]_i} \mid 1 \leq i \leq 5\}^* \longrightarrow \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}, \underline{\#}\}^* \text{ mit}$$

$$h(\underline{[_1}) = h(\underline{]_1}) = \underline{\mathtt{a}}, h(\underline{[_2}) = h(\underline{]_2}) = \underline{\mathtt{b}}, h(\underline{[_3}) = h(\underline{]_3}) = \underline{\mathtt{c}}, h(\underline{[_4}) = h(\underline{]_4}) = \underline{\mathtt{d}}, h(\underline{[_5}) = \underline{\#}, h(\underline{]_5}) = \varepsilon$$

c) $L = \{ww \mid w \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}, \underline{\mathtt{d}}\}^*\}$ ist nicht kontextfrei.

Beweis indirekt. Nehmen wir an, L ist kontextfrei. Sei dann m die Konstante aus dem Pumping Lemma für kontextfreie Sprachen. Wir wählen nun z.B. das Wort

$$s = \underline{\mathtt{a}}^m\underline{\mathtt{b}}^m\underline{\mathtt{a}}^m\underline{\mathtt{b}}^m$$

Für diese Wort s gilt $s \in L$ und $|s| \geq m$, da $|s| = 4m$. Nun zerlegen wir s in $uvxyz$ so, dass $|vxy| \leq m$ und $|vy| \geq 1$. Wir wählen $i = 0$, d.h., $s_0 = uv^0xy^0z = uxz$. Dabei ergeben sich folgende Möglichkeiten:

1. vxy liegt in der ersten Hälfte des Wortes.

   Nachdem aber $|vy| \geq 1$, enthält nun s_0 in der ersten Worthälfte zumindest ein Symbol $\underline{\mathtt{a}}$ oder ein Symbol $\underline{\mathtt{b}}$ weniger als entsprechende Symbole $\underline{\mathtt{a}}$ oder $\underline{\mathtt{b}}$ in der zweiten. Damit hat das Wort aber nicht mehr die Form ww, und kann somit nicht in L sein. WIDERSPRUCH!

2. vxy liegt in der zweiten Hälfte des Wortes. Hier gilt natürlich analog:

   Nachdem $|vy| \geq 1$, enthält nun s_0 in der zweiten Worthälfte zumindest ein Symbol $\underline{\mathtt{a}}$ oder ein Symbol $\underline{\mathtt{b}}$ weniger als entsprechende Symbole $\underline{\mathtt{a}}$ oder $\underline{\mathtt{b}}$ in der ersten. Damit hat das Wort aber nicht mehr die Form ww und kann somit nicht in L sein. WIDERSPRUCH!

3. vxy befindet sich in einer „Mittellage":

   Nachdem $|vy| \geq 1$, fällt nun in s_0 zumindest ein Symbol $\underline{\mathtt{b}}$ in der ersten Worthälfte oder ein Symbol $\underline{\mathtt{a}}$ in der zweiten Worthälfte weg. Auch damit hat das Wort aber nicht mehr die Form ww und kann somit nicht in L sein. WIDERSPRUCH!

Wir haben alle möglichen Zerlegungen von s untersucht, sind aber in jedem Fall auf einen Widerspruch gestoßen, d.h., L kann nicht kontextfrei sein.

Aufgabe 2.3 Geben Sie für jede der folgenden Grammatiken an, ob diese regulär, kontextfrei, monoton, kontextsensitiv und/oder unbeschränkt ist. Geben Sie weiters an, welche Sprache von der jeweils angegebenen Grammatik erzeugt wird, und ob diese regulär, kontextfrei, kontextsensitiv und/oder rekursiv aufzählbar ist.

a) $G_1 = (\{S, B\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}, \underline{\mathtt{c}}\}, \{S \to \underline{\mathtt{c}}\,\underline{\mathtt{b}}\,\underline{\mathtt{a}} \mid \underline{\mathtt{c}}SB\underline{\mathtt{a}}, \quad \underline{\mathtt{a}}B \to B\underline{\mathtt{a}}, \quad \underline{\mathtt{b}}B \to \underline{\mathtt{b}}\,\underline{\mathtt{b}}\}, S)$

b) $G_2 = (\{S\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, \{S \to \underline{\mathtt{a}}S \mid \underline{\mathtt{b}}S \mid S\underline{\mathtt{a}} \mid S\underline{\mathtt{b}}\}, S\})$

c) $G_3 = (\{S\}, \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}, \{S \to Saa \mid a \in \{\underline{\mathtt{a}}, \underline{\mathtt{b}}\}\} \cup \{S \to \varepsilon\}, S)$

Lösung

a) G_1 ist monoton und unbeschränkt, aber wegen z.B. $\underline{\mathtt{a}}B \to B\underline{\mathtt{a}}$ weder regulär, noch kontextfrei, noch kontextsensitiv.

   $\mathcal{L}(G_1) = \{\underline{\mathtt{c}}^n\underline{\mathtt{b}}^n\underline{\mathtt{a}}^n \mid n \geq 1\}$ ist kontextsensitiv und daher auch rekursiv aufzählbar. (aber keinesfalls regulär und kontextfrei, siehe z.B. Folie 174 ff.).

b) G_2 ist kontextfrei, monoton, kontextsensitiv und unbeschränkt, aber nicht regulär (wegen z.B. $S \to S\underline{\mathtt{a}}$).

   $\mathcal{L}(G_2) = \{\}$ ist aber regulär, und damit, aufgrund der Chomsky Hierarchie auch kontextfrei, kontextsensitiv und rekursiv aufzählbar.

c) G_3 ist kontextfrei und unbeschränkt. Wegen z.B. der Produktion $S \to S\underline{\mathtt{a}}\,\underline{\mathtt{a}}$ ist G_3 nicht regulär, und wegen $S \to \varepsilon$ nicht monoton und kontextsensitiv, nachdem S auch auf der rechten Seite von Produktionen vorkommt.

   $\mathcal{L}(G_3) = (\{\underline{\mathtt{a}}\,\underline{\mathtt{a}}\}^* \cup \{\underline{\mathtt{b}}\,\underline{\mathtt{b}}\}^*)^*$ ist aber regulär, und damit auch kontextfrei, kontextsensitiv und rekursiv aufzählbar.

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

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

b) Zu jeder regulären Sprache L gibt es eine unbeschränkte Grammatik, welche L erzeugt.

c) Ist L kontextfrei, so ist auch jede Teilmenge von L kontextfrei.

d) Es gibt kontextfreie Sprachen, für deren Komplement das Wortproblem entscheidbar ist.

e) Wenn $\mathbf{P} = \mathbf{NP}$ ist, kann jedes Problem, das in exponentieller Zeit gelöst werden kann, schon in polynomieller Zeit gelöst werden.

f) Sei A $\mathbf{NP}$-hart und B $\mathbf{NP}$-vollständig. Dann gilt: Wenn B in $\mathbf{P}$ ist, so ist auch A in $\mathbf{P}$.

g) Sei $A \leq_p B$. Dann gilt: Wenn B in $\mathbf{P}$ ist, dann ist auch das Komplement von A in $\mathbf{P}$.

h) Sei $A \leq_p B$. Dann gilt: Wenn A $\mathbf{NP}$-vollständig ist, so ist auch B $\mathbf{NP}$-vollständig.

Lösung

a) Nein. Zwar ist G mehrdeutig, denn z.B. für das Wort $\underline{\mathtt{a}}$ existieren zwei verschiedene Linksableitungen: $S \Rightarrow A \Rightarrow \underline{\mathtt{a}}$ bzw. $S \Rightarrow \underline{\mathtt{a}}$.
   Für die von G erzeugte Sprache $L(G) = \{\underline{\mathtt{a}}\}$ existiert aber z.B. folgende eindeutige Grammatik: $G = (\{S\}, \{\underline{\mathtt{a}}\}, \{S \to \underline{\mathtt{a}}\}, S)$.

b) Das ist korrekt, da jede reguläre Sprache rekursiv aufzählbar ist, und somit auch von einer unbeschränkten Grammatik erzeugt werden kann.

c) Das stimmt sicher nicht. Gegenbeispiel: $\{\underline{\mathsf{a}}^n \mid n \geq 0\}$ ist kontextfrei, enthält aber z.B. die Teilmenge $\{\underline{\mathsf{a}}^{2^n} \mid n \geq 0\}$, welche sicher nicht kontextfrei ist.

d) Das ist korrekt, denn ist L kontextfrei, so ist das Komplement von L entscheidbar. (Kontextfreie Sprachen sind eine echte Teilmenge der rekursiven Sprachen, welche unter Komplement abgeschlossen sind.)

e) Falsch, da $\mathbf{P} \neq \mathbf{EXPTIME}$.

f) Nein. Nachdem A **NP**-hart ist, kann jedes Problem in **NP** in polynomieller Zeit auf A reduziert werden. Es gibt also eine polynomielle Reduktion von B auf A. Das Problem A selbst muss deshalb aber nicht in **NP** bzw. **P** liegen, sondern kann durchaus schwieriger sein (z.B. in **EXPTIME** sein).

g) Ja, denn **P** ist unter Komplement abgeschlossen.

h) Damit diese Aussage zutrifft, muss B in **NP** liegen. B kann aber auch durchaus schwieriger sein, sogar nicht rekursiv aufzählbar.

Aufgabe 2.5

a) Finden und beschreiben Sie den Fehler im folgenden "Beweis" für $\mathbf{P} \neq \mathbf{NP}$:

Wir betrachten folgenden Algorithmus für SAT:

- Durchlaufe für die gegebene Formel ϕ alle möglichen Belegungen der Variablen mit den Wahrheitswerten
- Akzeptiere ϕ, wenn eine der durchlaufenen Belegungen ϕ erfüllt.

Dieser Algorithmus hat eine mit der Anzahl der Variablen exponentiell wachsende Laufzeit. Daher hat das Problem SAT einen exponentiellen Aufwand und kann nicht in **P** liegen. Weil aber SAT in **NP** liegt, muß also $\mathbf{P} \neq \mathbf{NP}$ gelten.

b) Es ist nicht bekannt, ob das Komplement von jeder Sprache in **NP** selbst wiederum in **NP** liegt (d.h., ob **NP** unter Komplement abgeschlossen ist bzw. ob **NP**=**co-NP**).

Finden und beschreiben Sie den Fehler in folgendem vermeintlichen "Beweis", welcher dies aber für alle Probleme in **NP** behauptet:

Satz: Ist $L \in \mathbf{NP}$, so ist auch $\overline{L} \in \mathbf{NP}$.

Beweis: Sei N eine nicht-deterministische Turingmaschine (NTM), welche L in polynomiell beschränkter Zeit akzeptiert. Vertausche nun die Endzustände von N mit Nicht-Endzuständen von N und umgekehrt. Offensichtlich akzeptiert die so konstruierte NTM N' die Sprache $\overline{L}$ in polynomiell beschränkter Zeit. Es gilt also $\overline{L} \in \mathbf{NP}$. □

Lösung

a) Die Vorgabe eines Algorithmus mit exponentiellem Aufwand zur Lösung eines gegebenen Problems bedeutet nicht, dass das Problem eine exponentielle Komplexität besitzt und damit in **NP** liegt. Der Denkfehler besteht darin, dass aus der Tatsache, dass man nicht auf triviale Weise einen effizienten (polynomiellen) Algorithmus findet, geschlossen wird, dass ein solcher effizienter (polynomieller) Algorithmus auch nicht existieren kann.

b) Die Maschine N' akzeptiert genau jene Wörter, für welche es in N mindestens eine Folge möglicher Bewegungen gibt, die von der Startkonfiguration zu einer Endkonfiguration in einem nicht-akzeptierenden Zustand führt. Dies könnten aber durchaus auch Wörter sein, welche von N akzeptiert werden.

N' akzeptiert also nicht $\overline{L}$, sondern eigentlich eine Obermenge von $\overline{L}$.