ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

Prof. Dr. Peter Thiemann
Manuel Geffken

18.12.2015
Abgabe bis spätestens Freitag 08.01.2016, 10 Uhr
in die Briefkästen in Gebäude 51

8. Übungsblatt zur Vorlesung
Theoretische Informatik

Aufgabe 1: Eindeutigkeit 2+6 Punkte

Betrachten Sie die Grammatiken $G_1 := (\{S\}, \{a, b\}, P_1, S)$ und $G_2 := (\{S, A, B\}, \{a, b\}, P_2, S)$ mit Produktionen
P_1:

$$
\begin{aligned}
&S \to \varepsilon \\
&S \to aSbS \\
&S \to bSaS
\end{aligned}
$$

P_2:

$$
\begin{aligned}
&S \to \varepsilon, \\
&S \to aB, \;\; S \to bA, \\
&A \to bA'A, \;\; B \to aB'B, \\
&A \to aS, \;\; B \to bS \\
&A' \to bA'A', \;\; B' \to aB'B' \\
&A' \to a, \;\; B' \to b
\end{aligned}
$$

Von Blatt 6 wissen Sie bereits, dass G_1 die Sprache $L = \{w \in \{a, b\}^* \mid \#_a(w) = \#_b(w)\}$ erzeugt.

(a) Zeigen Sie, dass G_1 nicht eindeutig ist.

(b) Welche Sprache L wird von G_2 erzeugt? Geben Sie eine möglichst einfache Beschreibung für L an. (Ein Beweis für $L = \mathcal{L}(G_2)$ ist nicht nötig.)

(c) Zeigen Sie, dass G_2 eindeutig ist, indem Sie beweisen, dass G_2 eine sogenannte LL(1)-Sprache ist. Wenden Sie dazu folgendes Verfahren an:

Seien im Folgenden a, b Terminale; A, B Nichtterminale; X, Y beliebige Grammatiksymbole (Terminale oder Nichtterminale); w eine endliche Folge von Terminalsymbolen (ein Wort); α, β endliche Folgen von Grammatiksymbolen; P die Produktionen einer Grammatik.

(i) Bestimmen Sie FIRST(α) für die rechte Seite jeder Produktion $A \to \alpha$.
Die Funktion FIRST ist wie folgt definiert:

$$\begin{aligned}
\mathrm{FIRST}(a) &:= \{a\} \\
\mathrm{FIRST}(A) &:= \bigcup_{A\to\alpha} \mathrm{FIRST}(\alpha) \\
\mathrm{FIRST}(\varepsilon) &:= \{\varepsilon\} \\
\mathrm{FIRST}(a\beta) &:= \{a\} \\
\mathrm{FIRST}(A\beta) &:= \mathrm{FIRST}^\star(\mathrm{FIRST}(A)\cdot\mathrm{FIRST}(\beta))
\end{aligned}$$

Dabei ist die Funktion FIRST* auf Sprachen wie folgt definiert:

$$\mathrm{FIRST}^\star(L) := \bigcup_{w\in L} \mathrm{FIRST}(w)$$

(ii) Bestimmen Sie FOLLOW(A) für die linke Seite jeder Produktion $A \to \alpha$ für die gilt $\varepsilon \in \mathrm{FIRST}(\alpha)$.
Die Funktion FOLLOW ist wie folgt definiert:

$$\begin{aligned}
\mathrm{FOLLOW}(A) &:= \{\varepsilon \mid A = S\} \\
&\cup \bigcup_{B\to\alpha A\beta} \mathrm{FIRST}^\star(\mathrm{FIRST}(\beta)\cdot\mathrm{FOLLOW}(B))
\end{aligned}$$

In FOLLOW(A) signalisiert ε das Ende der Eingabe. Beachten Sie, dass jeweils $\alpha = \varepsilon$ und $\beta = \varepsilon$ sein kann.
Hinweis: Für die Zwecke dieser Aufgabe folgt aus $\mathrm{FOLLOW}(A) = \mathrm{FOLLOW}(A) \cup U$ dass $\mathrm{FOLLOW}(A) = U$ (für beliebige A, U).

(iii) Die Funktion FIRST$'$ ist wie folgt definiert:

$$\mathrm{FIRST}'(A \to \alpha) := \mathrm{FIRST}^\star(\mathrm{FIRST}(\alpha)\cdot\mathrm{FOLLOW}(A))$$

Zwei Produktionen $A \to \alpha, A \to \beta$ heißen disjunkt gdw.

$$\mathrm{FIRST}'(A \to \alpha) \cap \mathrm{FIRST}'(A \to \beta) = \emptyset$$

Zeigen Sie, dass alle Paare von Produktionen aus P_2, deren linke Seiten gleich sind, disjunkt sind. Damit ist G_2 eine LL(1)-Grammatik.

(iv) Begründen Sie intuitiv, dass LL(1)-Grammatiken stets eindeutig sind.

Aufgabe 2: CYK-Algorithmus 3+3 Punkte

(a) Gegeben sei die folgende kontextfreie Grammatik $G = \langle V, \Sigma, S, P\rangle$ mit $V = \{S, X, Y\}$, $\Sigma = \{a, b, c\}$ und $P = \{S \to aXbYc,\ X \to a \mid b \mid aXXb,\ Y \to bc \mid bYc\}$. Verwenden Sie den CYK-Algorithmus, um zu überprüfen, ob das Wort *ababbccc* in $L(G)$ enthalten ist.

(b) L sei die Sprache der Palindrome der Länge ≥ 1 über dem Alphabet $\{a, b, c\}$. Geben Sie zunächst eine Grammatik G an, so dass $\mathcal{L}(G) = L$ (kein Beweis erforderlich). Verwenden Sie anschließend den CYK-Algorithmus, um zu verifizieren, dass das Wort *abcbcba* ein Palindrom ist.

Aufgabe 3: Abschlusseigenschaften kontextfreier Sprachen I 2 Punkte

Betrachten Sie die Sprache

$$L := \{w \in \{a, b, c\}^* \mid \#_a(w) = \#_b(w) = \#_c(w)\}$$

Zeigen Sie, ohne das Pumping Lemma zu benutzen, dass L nicht kontextfrei ist.

Hinweis: Argumentieren Sie mit den Abschlusseigenschaften von kontextfreien Sprachen und den Sprachen, von denen Sie aus Vorlesung oder Übung bereits wissen, dass diese nicht kontextfrei sind.

Aufgabe 4: Abschlusseigenschaften kontextfreier Sprachen II 2 Punkte

Sind die kontextfreien Sprachen unter der Differenz $\cdot \setminus \cdot$ abgeschlossen? Beweisen Sie Ihre Behauptung.

Aufgabe 5: Weihnachten I 2+2+2+2 Bonuspunkte

Für seinen unendlichen Weihnachtsbaum am Nordpol sucht der Weihnachtsmann noch passende unendliche Lichterketten. Diese müssen allerdings einige Sicherheitsvorschriften erfüllen, und die existierenden, auf DFAs basierenden Zertifizierungsverfahren sind nur für endliche Lichterketten geeignet. Sie sehen darin unmittelbar einen *emerging market* mit Potential für *rapid growth* und brechen Ihr Studium ab, um mit zwei Millionen *venture capital* ein *startup* im Keller des Hauses Ihrer Eltern zu gründen. Bei einer intensiven *brainstorming session* mit Ihren *Vice Presidents* fällt Ihnen spontan Folgendes ein:

Sei Σ ein Alphabet. Die Menge aller unendlichen Wörter über Σ ist

$$\Sigma^\omega := \{w \mid w \colon \mathbb{N} \to \Sigma\}$$

d.h. die Menge der totalen Funktionen, die jeder natürlichen Zahl einen Buchstaben zuordnen. Ein Wort $w \in \Sigma^\omega$ ist also eine Zeichenfolge der Form $a_0a_1a_2\ldots$ über Σ.
Ein Büchi-Automat über Σ ist ein Tupel $A := (\Sigma, Q, \delta, q_0, F)$, wobei jede Komponente genau wie bei DFAs definiert ist.
Ein Lauf über das Wort w für den Büchi-Automaten A ist eine Funktion $\rho \colon \mathbb{N} \to Q$ mit

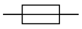
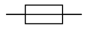
$$\begin{aligned} \rho(0) &= q_0 \\ \rho(n+1) &= \delta(\rho(n), w(n)) \quad \forall n \in \mathbb{N} \end{aligned}$$

Sei $\mathrm{Inf}(\rho)$ die Menge der Zustände, die in ρ unendlich oft besucht werden, d.h.

$$\mathrm{Inf}(\rho) := \big\{q \in Q \;\big|\; |\{n \in \mathbb{N} \mid \rho(n) = q\}| = \infty\big\}$$

Ein Wort w wird von A akzeptiert, falls es einen Lauf ρ über w für A gibt, bei dem mindestens ein akzeptierender Zustand unendlich oft besucht wird, d.h. für den gilt:

$$\mathrm{Inf}(\rho) \cap F \neq \emptyset$$

Mithilfe dieses Formalismus' können Sie nun zur Zertifizierung der Lichterketten schreiten. Eine unendliche Lichterkette ist ein unendliches Wort über dem Alphabet $\Sigma :=$ {—, ,  }[1], wobei — für ein Kabelstück,  für ein Licht und  für eine Sicherung steht.[2]

Formalisieren Sie folgende Anforderungen als Büchi-Automaten:

(a) Die Lichterkette beginnt mit beliebig vielen Kabelstücken.

(b) Die Lichterkette enthält unendlich viele Lichter.

(c) Vor und nach jedem Licht und jeder Sicherung liegt mindestens ein Kabelstück.

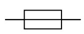
(d) Nach je maximal drei Lichtern muss eine Sicherung vorkommen. Zwischen Lichtern und Sicherungen dürfen dabei beliebig viele Kabelstücke liegen.

Geben Sie jeweils *ein* möglichst einfaches Zustandsdiagramm analog zu dem für DFAs an. Erläutern Sie jeweils kurz, warum Ihr Automat genau die Lichterketten akzeptiert, die die jeweilige Anforderung erfüllen.

Aufgabe 6: Weihnachten II 1 Bonuspunkt

Der Weihnachtsmann hat Ihnen wegen Arbeitsverweigerung fristlos gekündigt, nachdem Sie nicht in der Lage waren, in der veranschlagten Zeit Glühbirnen in alle Fassungen einer unendlichen Lichterkette zu drehen. Im darauf folgenden arbeitsrechtlichen Prozess macht er geltend, Sie hätten doch nur so lange arbeiten müssen, bis Sie keine Glühbirne mehr in eine Fassung drehen können, und das könne ja wohl so schwer nicht sein. Zeigen Sie:

Selbst wenn bereits jede Fassung eine Glühbirne enthält, können Sie noch beliebig viele Glühbirnen in die Fassungen der Lichterkette drehen.

[1] Alternativ können Sie auch die Symbole k statt —, l statt  und s statt  verwenden.

[2] Unendliche Lichterketten müssen mit unendlich starkem Strom betrieben und deshalb separat gesichert werden.