Theoretische Informatik 2

Sommersemester 2018

Elizaveta Kovalevskaya
Georg Schnitger
Philipp Tertel

GOETHE UNIVERSITÄT
FRANKFURT AM MAIN

Institut für Informatik
AG Theoretische Informatik

Übungsblatt 4

Ausgabe: 08.05.18
Abgabe: 15.05.18

Aufgabe 4.1 *Kontextfrei oder nicht?* (3+4 Punkte)

a) Im Streich-Spiel ist ein Wort $w \in \{a,b\}^*$ gegeben. Wir fragen uns, ob es möglich ist w in einer Folge von Zügen in das leere Wort zu überführen. In einem Zug darf irgendein Teilwort, das nur aus a's oder nur aus b's besteht, gestrichen werden, solange das Teilwort mindestens die Länge zwei hat.

Die Sprache ST besteht aus allen Worten $w \in \{a,b\}^*$, die sich durch irgendeine Folge von Zügen in das leere Wort überführen lassen.

Beispiel: $w := abbaaaba$ gehört zu ST, denn $w \to abbba \to aa \to \varepsilon$. Andererseits gehört $abbaaab$ nicht zu ST.

Zeigen Sie durch eine Angabe einer kontextfreien Grammatik G, dass ST kontextfrei ist. Eine informelle, aber natürlich überzeugende Begründung für $ST = L(G)$ genügt.

b) Sei $\Sigma = \{a,b,c\}$ und sei

$$L_b := \{\, a^i b^i c^j \ : \ i,j \in \mathbb{N},\ i \neq j \,\}.$$

Zeigen Sie, dass L_b nicht kontextfrei ist.

Hinweis: Wenn L_b kontextfrei wäre, dann gibt es mit Ogdens Lemma eine Pumpingkonstante N. Wählen Sie $z := a^N b^N c^{N+?}$. Welche Positionen in z markieren Sie und welchen Wert wählen Sie für das Fragezeichen, so dass Sie, unabhängig von der Zerlegung $z = uvwxy$, nach geeignetem Aufpumpen aus der Sprache geworfen werden?

Aufgabe 4.2 *Eindeutigkeit – Mehrdeutigkeit* (4+4 Punkte)

a) Die Dyck-Sprache D aller wohlgeformten Klammerausdrücke wird sowohl definiert durch die kontextfreie Grammatik $G_1 = (\Sigma, V_1, S_1, P_1)$ mit den Produktionen

$$S_1 \to (S_1) \mid S_1 S_1 \mid \varepsilon$$

wie auch durch die kontextfreie Grammatik $G_2 = (\Sigma, V_2, S_2, P_2)$ mit den Produktionen

$$S_2 \to (S_2) S_2 \mid \varepsilon.$$

Welche der Grammatiken ist eindeutig, welche ist mehrdeutig? Beweisen Sie ihre Antwort. (Sie brauchen nicht zu zeigen, dass die jeweilige Grammatik die Dyck-Sprache erzeugt.)

Bitte wenden!

b) Die Grammatik $G = (\Sigma, V, S, P)$ mit dem Terminalalphabet $\Sigma := \{\mathtt{i}, \mathtt{e}\}$ und den Produktionen

$$S \to \mathtt{i}S\mathtt{e}S \mid \mathtt{i}S \mid \varepsilon$$

modelliert eine Spezifikation der if-then-else-Anweisung. Leider besitzt das Wort

`iie`

zwei unterschiedliche Ableitungsbäume.

Entwerfen Sie eine eindeutige Grammatik G_3 für $L(G)$. Eine *informelle* Begründung, warum Ihre Konstruktion funktioniert, ist völlig ausreichend.

Hinweis: In G_2 sollte einem „`e`“ das nächstliegende linke, offene „`i`“ zugewiesen werden. Und die Konsequenz: Für eine Produktion der Form „$Z \to \mathtt{i}X\mathtt{e}Y$“ ist zu garantieren, dass X nur if-then-else-Ausdrücke mit derselben Anzahl von `i`’s und `e`’s erzeugt.

Aufgabe 4.3 *Kontextfreie Sprachen auf unären Alphabeten* (4+4+1 Punkte)

Sei $\Sigma = \{1\}$ und sei $L \subseteq \Sigma^*$ eine kontextfreie Sprache. Wir möchten mithilfe des Pumping-Lemmas für kontextfreie Sprachen zeigen, dass L regulär ist.

a) Sei N die Pumpingkonstante, wenn das Pumping-Lemma für L angewandt wird. Zeigen Sie: Für jedes $M \geq N$ mit $1^M \in L$ gibt es eine Zahl $1 \leq a_M \leq N$, so dass

$$\{\, 1^{M-a_M} 1^{i \cdot a_M} \ : \ i \in \mathbb{N} \,\} \subseteq L.$$

b) Zeigen Sie, dass $L' := \{1^M \in L \ : \ M \geq N\}$ eine *endliche* Vereinigung von Mengen der Form $\{1^b 1^{i \cdot a} \ : \ i \in \mathbb{N}_{>0}\}$ für Zahlen $a, b \in \mathbb{N}$ mit $a \leq N$ und $b \geq N - a$ ist.

c) Zeigen Sie, dass L regulär ist.

Kommentar: L ist also die Vereinigung von „linearen“ Sprachen.