Institut für Theoretische Informatik
Lehrstuhl Prof. Dr. D. Wagner

Übungsblatt 6

Vorlesung Theoretische Grundlagen der Informatik im WS 21/22

Ausgabe: 14. Januar 2022
Abgabe: 28. Januar 2022 (digital im ILIAS)

Bitte bearbeiten Sie die Aufgaben **handschriftlich** und laden Sie eine eingescannte PDF-Version im Übungsmodul Ihrer ILIAS-Tutoriumsgruppe hoch! Beschriften Sie Ihren handschriftlichen Aufschrieb gut sichtbar mit Name und Matrikelnummer. Nicht handschriftliche oder unbeschriftete Abgaben werden nicht akzeptiert!

Aufgabe 1 (2 + 1 + 1 + 2 = 6 Punkte)

Gegeben sei die kontextfreie Grammatik $G = (\Sigma, V, S, R)$ in Chomsky-Normalform mit Terminalen $\Sigma = \{\mathsf{a}, \mathsf{b}\}$, Nichtterminalen $V = \{S, A, B, C, D, E\}$, Startsymbol S und Produktionen

$$\begin{aligned} R = \{&S \to AB \mid SS \mid \mathsf{a}, \\ &A \to BS \mid CD \mid \mathsf{b}, \\ &B \to DD \mid \mathsf{b}, \\ &C \to DE \mid \mathsf{a} \mid \mathsf{b}, \\ &D \to \mathsf{a}, \\ &E \to SS\} \end{aligned}$$

(a) Überprüfen Sie mit dem CYK-Algorithmus, ob das Wort abaab in der Sprache $L(G)$ enthalten ist.

a	b	a	a	b

(b) Geben Sie einen Syntaxbaum für das Wort abaab an.

Betrachten Sie für eine Sprache L die Sprache $Pre(L) \subseteq L$, die ein Wort $w \in L$ genau dann enthält, wenn auch alle Präfixe von w in L enthalten sind, d.h.

$$Pre(L) = \{ w \mid w[1 \dots i] \in L \text{ für alle } i \in \{1, \dots |w|\} \} ,$$

wobei $w[1 \dots i]$ das Präfix von w der Länge i ist.

(c) Sei $L = \{w \in \{\mathsf{a}, \mathsf{b}\}^* \mid w \text{ ist ein Palindrom}\}$. Geben Sie $Pre(L)$ an.

(d) Entwerfen Sie einen möglichst effizienten Algorithmus, der prüft, ob $w \in Pre(L(G))$ für ein gegebenes Wort $w \in \Sigma^*$ und eine kontextfreie Grammatik G in Chomsky-Normalform über einem endlichen Alphabet Σ. Welche Laufzeit hat Ihr Algorithmus?

Aufgabe 2 (5 Punkte)

Sei eine Grammatik $G = (\Sigma, V, S, R)$ durch $V = \{S, A, B, C, E, F\}, \Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ und folgende Regelmenge R gegeben:

$$\begin{aligned}
S &\to B \mid C\mathsf{a}C \\
A &\to \mathsf{a} \mid D \\
B &\to CC \mid E \mid \mathsf{a}C\mathsf{b} \\
C &\to \mathsf{c} \mid A \mid BC \mid \varepsilon \\
D &\to EC \\
E &\to \mathsf{b} \mid A
\end{aligned}$$

Verwenden Sie das Verfahren aus der Vorlesung, um G in eine äquivalente Grammatik in erweiterter Chomsky-Normalform zu überführen. Es genügt dabei, wenn sie nach jedem Schritt bzw. jeder Phase das jeweilige Ergebnis angeben.

Aufgabe 3 (2 + 2 = 4 Punkte)

Sei $G = (\Sigma, S, V, R)$ eine kontextfreie Grammatik in Chomsky-Normalform.

(a) Sei $w \in L(G)$. Wie viele Ableitungsschritte sind nötig, um w aus S abzuleiten? Beweisen Sie.

(b) Sei $L(G)$ nun eine endliche Sprache. Wie lang kann ein Wort $w \in L(G)$ höchstens sein? Beweisen Sie eine möglichst scharfe obere Schranke in Abhängigkeit von G.

Aufgabe 4 (2 + 3 = 5 Punkte)

Geben Sie für die folgenden Sprachen jeweils eine Grammatik mit maximalem Chomsky-Typ an. Erläutern Sie die Idee der von Ihnen angegebenen Regeln.

(a) $L_1 = \{w \mid w \in \{\mathsf{a}, \mathsf{b}\}^*,\ w \neq w^R\}$
Dabei bezeichnet w^R das Spiegelwort von w.

(b) $L_2 = \{w{:}w \mid w \in \{\mathsf{a}, \mathsf{b}\}^*\} \subseteq \{\mathsf{a}, \mathsf{b}, :\}^*$

Aufgabe 5 (2 + 2 = 4 Punkte)

Ein Roboter R bewegt sich auf dem unendlichen zweidimensionalen Gitter $\mathbb{Z} \times \mathbb{Z}$. Dieser kann vom Ursprung $(0, 0)$ aus nach oben (O), unten (U), links (L) oder rechts (R) gesteuert werden. Wir können also einen Pfad des Roboters als Wort über dem Alphabet $\Sigma = \{\mathsf{O}, \mathsf{U}, \mathsf{L}, \mathsf{R}\}$ auffassen.

Geben Sie jeweils für die folgenden Sprachen eine kontextfreie Grammatik (mit kurzer Erklärung) an oder zeigen Sie, dass die Sprache nicht kontextfrei ist.

(a) $L_1 = \{w \in \Sigma^* \mid \text{Pfad } w \text{ endet im Ursprung}\}$

(b) $L_2 = \{w \in \Sigma^* \mid \text{Pfad } w \text{ befindet sich zu keinem Zeitpunkt unterhalb der } x\text{-Achse}\}$

Aufgabe 6 (2 + 1 + 2 + 2 = 7 Punkte)

Seien Σ und Γ zwei Alphabete. Eine Abbildung $h\colon \Sigma^* \to \Gamma^*$ ist ein Homomorphismus, wenn für alle Worte $u, v \in \Sigma^*$ gilt: $h(uv) = h(u)h(v)$. Beachten Sie, dass ein Homomorphismus vollständig durch die Bilder $h(a)$ für alle $a \in \Sigma$ festgelegt ist und dass $h(\varepsilon) = \varepsilon$ gilt.

(a) Zeigen Sie, dass die Klasse der kontextfreien Sprachen unter Homomorphismen abgeschlossen ist, d.h. für jede kontextfreie Sprache L und jeden Homomorphismus h gilt: $h(L) = \{h(w) \mid w \in L\}$ ist auch kontextfrei.

Beachten Sie, dass die Alphabete Σ und Γ nicht disjunkt sein müssen.

(b) Gilt auch die Umkehrung? Also ist L kontextfrei, wenn $h(L)$ kontextfrei und h ein Homomorphismus ist? Begründen Sie.

Im Folgenden sollen Sie zeigen, dass die Klasse der kontextsensitive Sprachen nicht unter Homorphismen abgeschlossen ist. Betrachten Sie dazu eine beliebige Grammatik $G = (\Sigma, S, V, R)$ und die Sprache

$$L_G = \{w_1\#w_2\#\ldots\#w_k \mid w_1 = S,\ w_k \in \Sigma^* \text{ und } w_i \to_G w_{i+1} \text{ für } i = 1, \ldots, k-1\}$$

Die Sprache L_G beschreibt genau die Menge aller in G erlaubten Folgen von Wortableitungen.

(c) Zeigen Sie, dass L_G eine kontextsensitive Sprache ist, indem Sie eine linear beschränkte TM angeben, die L_G entscheidet.

Nun erweitern wir das Alphabet um neue Symbole $\widetilde{\Sigma} = \{\widetilde{a} \mid a \in \Sigma\}$ und definieren für ein Wort $w = a_1 \ldots a_n \in \Sigma^*$ das Wort $\widetilde{w} = \widetilde{a_1} \ldots \widetilde{a_n} \in \widetilde{\Sigma}^*$. Damit modifizieren wir die Sprache L_G zu:

$$\widetilde{L_G} = \{w_1\#w_2\#\ldots\#\widetilde{w_k} \mid w_1\#w_2\#\ldots\#w_k \in L_G\}$$

Offensichtlich ist $\widetilde{L_G}$ wie auch L_G kontextsensitiv.

(d) Folgern Sie, dass es eine kontextsensitive Sprache L' und einen Homomorphismus h gibt, sodass $h(L')$ nicht kontextsensitiv ist.