Übung zur Vorlesung

Theoretische Informatik I

Prof. Dr. Christoph Kreitz
Universität Potsdam, Theoretische Informatik, WS 2008/09

Blatt 12 (Version 1) — Abgabetermin: 02.02.2009, 12:30 Uhr

Vorbereitung auf die nächste Vorlesung: Arbeiten Sie sich in das Thema "Turingmaschinen und Programmiertechniken" ein. Verwenden Sie hierzu z.B. die Vorlesungsfolien der Einheit 4.1, die Kapitel 8.2 und 8.3 des Buches von Hopcroft, Motwani und Ullman, eines der empfohlenen Bücher oder das Internet.

Aufgabe 12.1 (Substitutionen auf Sprachen)

Sei σ eine Abbildung von Wörtern in Sprachen mit $\sigma(0) = \{a^{m^2} \mid m \in \mathbb{N}\}$ und $\sigma(1) = \{bb\}$.

1. Geben Sie die Abbilder der Wörter 00, 01 und 11 unter σ an.
2. Beschreiben Sie die Sprache $\sigma(\{0^n 1^n \mid n \in \mathbb{N}\})$.
3. Nutzen Sie eine geeignete Abbildung σ', um zu beweisen, dass $L_k = \{a^n b^{2n} \mid n \in \mathbb{N}\}$ kontextfrei ist.

Aufgabe 12.2 (Abschlusseigenschaften)

Untersuchen Sie ausschließlich mit Hilfe der Abschlusseigenschaften kontextfreier Sprachen, ob die folgenden Sprachen kontextfrei sind. Verwenden Sie dabei, daß die Sprache $L_{01} = \{0^n 1^n \mid n \in \mathbb{N}\}$ kontextfrei und daß $L_{012} = \{0^n 1^n 2^n \mid n \in \mathbb{N}\}$ nicht kontextfrei ist.

1. $L_1 = \{0^n 1^{n+m} 2^m \mid n, m \in \mathbb{N}\}$
2. $L_2 = \{0^n 1^{n+m} 0^m \mid n, m \in \mathbb{N}\}$
3. $L_3 = \{w \in \{0,1,2\}^* \mid \#_0(w) = \#_1(w) = \#_2(w)\}$
4. $L_4 = \{0^{n_1} 2 1^{n_1} 0^{n_2} 2 1^{n_2} \ldots 0^{n_m} 2 1^{n_m} \mid m, n_1, n_2, .., n_m \in \mathbb{N}\}$

Aufgabe 12.3 (Chomsky-Normalform)

Sei $G = (\{S, A, B, C\}, \{0, 1, 2\}, P, S)$ mit $P = \{S \to ASA \mid 1B,\ A \to S \mid B,\ B \to \varepsilon \mid 0 \mid C,\ C \to 2C\}$. Transformieren Sie G mit dem in der Vorlesung beschriebenen Verfahren in eine Chomsky-Normalform, indem Sie die folgenden Schritte abarbeiten.

1. Eliminieren Sie alle ε-Produktionen.
2. Eliminieren Sie alle Einheitsproduktionen.
3. Eliminieren Sie unnütze Symbole.
4. Separieren Sie die Terminalsymbole und Variablen in den Produktionen.
5. Spalten Sie alle Produktionen $A \to \alpha$ mit $|\alpha| > 2$ auf.

Aufgabe 12.4 (Syntaxanalyse mit dem Cocke-Younger-Kasami Algorithmus)

Gegeben sei die kontextfreie Grammatik $G = (\{S, A, B, C\}, \{a, b\}, P, S)$ in Chomsky-Normalform mit $P = \{S \to a \mid AB \mid BC,\ A \to a \mid AA \mid AB \mid BB,\ B \to b \mid BC,\ C \to BA\}$.

Prüfen Sie mit Hilfe des Cocke-Younger-Kasami Algorithmus, ob die Wörter $ababa$, $aabba$ und $bbaaa$ zur Sprache von G gehören.

Aufgabe 12.5 (Etwas Theorie: Zwei-Stack PDAs)

Die Sprache $L_{012} = \{0^n 1^n 2^n \mid n \in \mathbb{N}\}$ kann von einem PDA akzeptiert werden, wenn dieser einen zweiten zusätzlichen Stack besitzt.

Erklären Sie, wie dieser PDA arbeiten würde und begründen Sie intuitiv, warum ein einziger Stack nicht ausreicht, um L_{012} zu akzeptieren.

Hausaufgabe 12.6 (Abschlusseigenschaften)

Prüfen Sie ausschließlich mit Hilfe der Abschlusseigenschaften regulärer und kontextfreier Sprachen, ob die Sprachen

1. $L_1 = \{a^n b^{2n} c^m \mid n, m \in \mathbb{N}\}$ und $L_2 = \{a^n b^m c^{2m} \mid n, m \in \mathbb{N}\}$.
2. $L_3 = \{a^i b^j c^k \mid i, j, k \in \mathbb{N} \wedge (j{=}2i \vee k{=}2j)\}$
3. $L_4 = L_1 \cap L_2$

kontextfrei sind. Sie dürfen die in der Vorlesung oder den Übungsaufgaben als regulär oder kontextfrei nachgewiesenen Sprachen benutzten, wenn Sie die entsprechenden Folien oder Nummern der Übungsaufgaben angeben.

Hausaufgabe 12.7 (Chomsky-Normalform)

Transformieren Sie die Grammatik $G = (\{S, A, B, C, D, E\}, \{a, b\}, P, S)$ mit
$P = \{S \to \varepsilon \mid aAa \mid bBb,\ A \to S \mid C,\ B \to aB \mid Bb \mid BD,\ C \to \varepsilon \mid S,\ D \to a \mid Eb,\ E \to b \mid Bb \mid BC\}$
in eine Chomsky-Normalform. Gehen Sie dabei nach dem in der Vorlesung beschriebenen Verfahren vor und geben Sie alle Zwischenschritte an.

Hausaufgabe 12.8 (Syntaxanalyse mit dem Cocke-Younger-Kasami Algorithmus)

Gegeben sei die Grammatik $G = (\{S\}, \{a, b\}, P, S)$ mit den Regeln $P = \{S \to aa \mid abS \mid bSaa\}$.

Prüfen Sie mit dem Algorithmus von Cocke, Younger und Kasami, ob die Wörter $aabba$, $abbaaaa$ und $ababaa$ in $L(G)$ enthalten ist.

Bonusaufgabe (3 Sonderpunkte) Für einen Informatiker ist es eigentlich wenig sinnvoll, Computer zu spielen und ein bekanntes Verfahren von Hand auszuführen. Es ist besser, den Algorithmus zu implementieren und ihn dann die Lösung erzeugen zu lassen.

Sie erhalten 2 Sonderpunkte, wenn Sie das CYK Verfahren so implementieren, daß es bei Eingabe eines Wortes (als einfacher ASCII Text) die Tabelle ausdruckt und zum Schluß entscheidet, ob das Wort zur Sprache der Grammatik gehört. Sie enthalten einen dritten Sonderpunkt, wenn es auch die Grammatik als Eingabe verwalten kann. Das Programm muß eine kurze Bedienungs-/Installationsanleitung enthalten und sollte auf einer CD beigelegt oder per e-mail an `tim@cs.uni-potsdam.de` und an `kreitz@cs.uni-potsdam.de` geschickt werden. Es sollte ohne größeren Aufwand zum Laufen gebracht werden können.