Universität Stuttgart

Lösungsblatt 11

Vorbereitungsaufgaben

Vorbereitungsaufgabe 1

Sei $G = (\{S, A, B, C, D, E, F, G\}, \{a, b\}, P, S)$ eine Grammatik mit Produktionen

$$\begin{aligned} S &\to B \mid C \mid GaCB & D &\to aCb \mid ba \mid G \\ A &\to a \mid E & E &\to A \mid F \\ B &\to AD \mid F & F &\to b \mid A \\ C &\to a \mid A & G &\to CE \mid D. \end{aligned}$$

Wandeln Sie G in eine äquivalente Grammatik in Chomsky-Normalform um.

Lösung

Wir führen die Schritte 1 - 5 aus der Vorlesung aus.

Schritt 1

A, E und F bilden eine Ringableitung. Wir ersetzen alle drei Variablen durch H, entfernen die Regel $H \to H$ und erhalten:

$$\begin{aligned} S &\to B \mid C \mid GaCB \\ B &\to HD \mid H \\ C &\to a \mid H \\ D &\to aCb \mid ba \mid G \\ G &\to CH \mid D \\ H &\to a \mid b. \end{aligned}$$

Auch D und G bilden eine Ringableitung. Wir ersetzen beide Variablen durch I, entfernen die Regel $I \to I$ und erhalten:

$$\begin{aligned} S &\to B \mid C \mid IaCB \\ B &\to HI \mid H \\ C &\to a \mid H \\ H &\to a \mid b \\ I &\to aCb \mid ba \mid CH. \end{aligned}$$

Schritt 2

Bei der Anordnung der Variablen muss S vor B und C kommen und sowohl B als auch C müssen vor H kommen. Eine von 10 möglichen Anordnungen ist S, B, C, H, I.

Schritt 3

Durch Beseitigen der Regel $C \to H$ erhalten wir:

$$\begin{aligned}
&S \to B \mid C \mid IaCB\\
&B \to HI \mid H\\
&C \to a \mid b\\
&H \to a \mid b\\
&I \to aCb \mid ba \mid CH.
\end{aligned}$$

Durch Beseitigen der Regel $B \to H$ erhalten wir:

$$\begin{aligned}
&S \to B \mid C \mid IaCB\\
&B \to HI \mid a \mid b\\
&C \to a \mid b\\
&H \to a \mid b\\
&I \to aCb \mid ba \mid CH.
\end{aligned}$$

Durch Beseitigen der Regel $S \to B$ erhalten wir:

$$\begin{aligned}
&S \to a \mid b \mid C \mid HI \mid IaCB\\
&B \to HI \mid a \mid b\\
&C \to a \mid b\\
&H \to a \mid b\\
&I \to aCb \mid ba \mid CH.
\end{aligned}$$

Durch Beseitigen der Regel $S \to C$ erhalten wir:

$$\begin{aligned}
&S \to a \mid b \mid HI \mid IaCB\\
&B \to HI \mid a \mid b\\
&C \to a \mid b\\
&H \to a \mid b\\
&I \to aCb \mid ba \mid CH.
\end{aligned}$$

Schritt 4

Mit den Pseudoterminalen J und K erhalten wir:

$$
\begin{aligned}
&S \to a \mid b \mid HI \mid IJCB \\
&B \to HI \mid a \mid b \\
&C \to a \mid b \\
&H \to a \mid b \\
&I \to JCK \mid KJ \mid CH \\
&J \to a \\
&K \to b.
\end{aligned}
$$

<u>Schritt 5</u>

Wir verkürzen die rechten Seiten der Regeln $S \to IJCB$ und $I \to JCK$ und erhalten die endgültige Regelmenge:

$$
\begin{aligned}
&S \to a \mid b \mid HI \mid IM \\
&B \to HI \mid a \mid b \\
&C \to a \mid b \\
&H \to a \mid b \\
&I \to JN \mid KJ \mid CH \\
&J \to a \\
&K \to b \\
&L \to CB \\
&M \to JL \\
&N \to CK.
\end{aligned}
$$

Vorbereitungsaufgabe 2

Sei $G = (\{A_1, A_2, A_3\}, \{a, b, c\}, P, A_1)$ eine kontextfreie Grammatik mit Produktionen

$$
\begin{aligned}
&A_1 \to A_2a \mid b \\
&A_2 \to A_3A_3 \\
&A_3 \to A_1c,
\end{aligned}
$$

bei der wir die Variablen der Einfachheit halber schon von A_1 bis A_3 durchnummeriert haben.

Wandeln Sie G in eine Grammatik G' in Greibach-Normalform um.

Lösung

Wir gehen vor wie auf den Vorlesungsfolien beschrieben.

Schritt 1 (erster Algorithmus)

Mit dem ersten Algorithmus werden Regeln der Form $A_i \to A_j\alpha$ mit $i \geq j$ entfernt. Für $i = 1$ passiert nichts, da keine Regeln der Form $A_1 \to A_1\alpha$ existieren. Für $i = 2$ passiert ebenfalls nichts, da keine Regeln der Formen $A_2 \to A_1\alpha$ oder $A_2 \to A_2\alpha$ existieren. Für $i = 3$ verändert sich die Regelmenge schrittweise wie folgt.

- Für $j = 1$ erhalten wir:

$$\begin{aligned} A_1 &\to A_2a \mid b \\ A_2 &\to A_3A_3 \\ A_3 &\to A_2ac \mid bc. \end{aligned}$$

- Für $j = 2$ erhalten wir:

$$\begin{aligned} A_1 &\to A_2a \mid b \\ A_2 &\to A_3A_3 \\ A_3 &\to A_3A_3ac \mid bc. \end{aligned}$$

- Durch die Entfernung der Linksrekursion $A_3 \to A_3A_3ac$ erhalten wir schließlich:

$$\begin{aligned} A_1 &\to A_2a \mid b \\ A_2 &\to A_3A_3 \\ A_3 &\to bc \mid bcB \\ B &\to A_3ac \mid A_3acB. \end{aligned}$$

Schritt 2 (zweiter Algorithmus)

Der zweite Algorithmus liefert:

$$\begin{aligned} A_1 &\to bcA_3a \mid bcBA_3a \mid b \\ A_2 &\to bcA_3 \mid bcBA_3 \\ A_3 &\to bc \mid bcB \\ B &\to A_3ac \mid A_3acB. \end{aligned}$$

Schritt 3 (B-Regeln)

Die rechten Seiten der B-Regeln beginnen mit A_3. Durch das Einsetzen von $A_3 \to bc \mid bcB$ in $B \to A_3ac \mid A_3acB$ erhalten wir:

$$\begin{aligned} A_1 &\to bcA_3a \mid bcBA_3a \mid b \\ A_2 &\to bcA_3 \mid bcBA_3 \\ A_3 &\to bc \mid bcB \\ B &\to bcac \mid bcBac \mid bcacB \mid bcBacB. \end{aligned}$$

Schritt 4 (Pseudoterminale)

Durch das Einführen von Pseudoterminalen V_a und V_c erhalten wir die gewünsche Regelmenge P':

$$
\begin{aligned}
A_1 &\to bV_cA_3V_a \mid bV_cBA_3V_a \mid b\\
A_2 &\to bV_cA_3 \mid bV_cBA_3\\
A_3 &\to bV_c \mid bV_cB\\
B &\to bV_cV_aV_c \mid bV_cBV_aV_c \mid bV_cV_aV_cB \mid bV_cBV_aV_cB\\
V_a &\to a\\
V_c &\to c.
\end{aligned}
$$

Ein Pseudoterminal V_b ist in diesem Fall nicht nötig.

Dann ist $G' = (\{A_1, A_2, A_3, B, V_a, V_c\}, \{a, b, c\}, P', A_1)$ eine zu G äquivalente Grammatik in Greibach-Normalform.

Hinweis: Man erkennt, dass beide A_2-Regeln auch entfernt werden könnten ohne die erzeugte Sprache zu verändern.

Vorbereitungsaufgabe 3

Zeigen Sie mit dem Pumping-Lemma für kontextfreie Sprachen, dass die Sprache

$$L = \left\{a^k b^\ell c^m \,\middle|\, k < \ell < m\right\}$$

über dem Alphabet $\Sigma = \{a, b, c\}$ nicht kontextfrei ist.

Lösung

Sei $n \in \mathbb{N}\backslash\{0\}$ beliebig. Wähle $z = a^n b^{n+1} c^{n+2}$. Dann gilt $z \in L$ und $|z| = 3n+3 \geq n$. Seien $u, v, w, x, y \in \Sigma^*$ beliebig mit $z = uvwxy$, $|vx| \geq 1$ und $|vwx| \leq n$. Wir unterscheiden drei Fälle:

- Fall 1: vx enthält ein a.

  Dann enthält vx kein c. Für $i = 3$ gilt $|uv^iwx^iy|_a \geq |uv^iwx^iy|_c$.

- Fall 2: vx enthält kein a, aber ein b.

  Für $i = 0$ gilt $|uv^iwx^iy|_a \geq |uv^iwx^iy|_b$.

- Fall 3: vx enthält weder ein a noch ein b.

  Dann enthält vx mindestens ein c. Für $i = 0$ gilt $|uv^iwx^iy|_b \geq |uv^iwx^iy|_c$.

In allen Fällen existiert ein $i \in \mathbb{N}$ mit $uv^iwx^iy \notin L$.

Vorbereitungsaufgabe 4

Seien $G = (V, \Sigma, P, S)$ eine Grammatik in Chomsky-Normalform und $w = a_1 \ldots a_n$ ein Wort mit $a_1, \ldots, a_n \in \Sigma$. Wir betrachten die vom CYK-Algorithmus verwendeten Mengen $T_{i,j}$, wenn dieser auf G und w gestartet wird.

Geben Sie für alle $i, j \in \mathbb{N}$ mit $1 \leq i \leq n$ und $1 \leq j \leq n - i + 1$ eine möglichst einfache Definition von $T_{i,j}$ an.

Lösung

$T_{i,j}$ enthält alle Variablen, von denen sich derjenige Infix von w ableiten lässt, der an der i-ten Position in w beginnt und Länge j hat, d. h.:

$$T_{i,j} = \{A \in V \mid A \Rightarrow_G^* a_i \dots a_{i+j-1}\} .$$

Vorbereitungsaufgabe 5

Spielen Sie ein bisschen mit folgender Webseite herum:

`www.xarg.org/tools/cyk-algorithm`

Lösung

Viel Spaß! :-)

Präsenzaufgaben

Präsenzaufgabe 1

Welche der folgenden Sprachen sind kontextfrei und welche nicht? Beweisen Sie Ihre Antworten.

1. $L = \{a^k b^\ell \mid k < \ell\}$ über $\Sigma = \{a, b\}$
2. $L = \{a^k b^\ell c^m \mid k + \ell = m\}$ über $\Sigma = \{a, b, c\}$
3. $L = \{a^k b^\ell c^k d^\ell \mid k, \ell \geq 0\}$ über $\Sigma = \{a, b, c, d\}$
4. $L = \{a^k b^\ell c^m d^n \mid k + m = \ell + n\}$ über $\Sigma = \{a, b, c, d\}$
5. $L = \{a^{k\ell} \mid k, \ell \geq 0\}$ über $\Sigma = \{a\}$
6. $L = \{a^{k\ell} \mid k, \ell \geq 2\}$ über $\Sigma = \{a\}$

Lösung

1. L ist kontextfrei, da L von der kontextfreien Grammatik $G = (\{S\}, \Sigma, P, S)$ mit $S \to aSb \mid Sb \mid b$ erzeugt wird.
2. L ist kontextfrei, da L von der kontextfreien Grammatik $G = (\{S, T\}, \Sigma, P, S)$ mit Produktionen

   $$\begin{aligned} S &\to aSc \mid T \\ T &\to bTc \mid \varepsilon \end{aligned}$$

   erzeugt wird.
3. Wir zeigen mit dem Pumping-Lemma, dass L nicht kontextfrei ist. Sei $n \in \mathbb{N}$ beliebig. Wähle $z = a^n b^n c^n d^n$. Dann gilt $z \in L$ und $|z| = 4n \geq n$. Seien $u, v, w, x, y \in \Sigma^*$ beliebig mit $z = uvwxy$, $|uvw| \leq n$ und $|vx| \geq 1$. Wir unterscheiden vier Fälle:

- Fall 1: vx enthält ein a.

  Dann enthält vx kein c. Für $i = 0$ gilt $|uv^iwx^iy|_a < |uv^iwx^iy|_c$.

- Fall 2: vx enthält kein a, aber ein b.

  Dann enthält vx kein d. Für $i = 0$ gilt $|uv^iwx^iy|_b < |uv^iwx^iy|_d$.

- Fall 3: vx enthält weder ein a noch ein b, aber ein c.

  Für $i = 0$ gilt $|uv^iwx^iy|_c < |uv^iwx^iy|_a$.

- Fall 4: vx enthält nur ds.

  Für $i = 0$ gilt $|uv^iwx^iy|_d < |uv^iwx^iy|_b$.

In allen Fällen existiert ein $i \in \mathbb{N}$ mit $uv^iwx^iy \notin L$.

4. L ist kontextfrei, da L von der kontextfreien Grammatik $G = (\{S, T, U, V\}, \Sigma, P, S)$ mit Produktionen

$$\begin{aligned} S &\to aSd \mid TUV \\ T &\to aTb \mid \varepsilon \\ U &\to bUc \mid \varepsilon \\ V &\to cVd \mid \varepsilon \end{aligned}$$

   erzeugt wird.

5. Da k bzw. ℓ auch den Wert 1 haben können, gilt $L = \{a\}^*$. Also ist L regulär und somit auch kontextfrei.

6. Wäre L kontextfrei, dann wäre L auch regulär, da L eine unäre Sprache ist. Da die Klasse der regulären Sprachen unter Vereinigung und Komplement abgeschlossen ist, wäre dann auch

$$\overline{L \cup \{\varepsilon, a\}} = \{a^p \mid p \text{ prim}\}$$

   regulär, ein Widerspruch zur Vorlesung!

   Erinnerung: Jede endliche Sprache ist regulär, also auch $\{\varepsilon, a\}$.

Präsenzaufgabe 2

Welche der folgenden Aussagen sind für beliebige Sprachen A und B richtig und welche falsch? Beweisen Sie Ihre Antworten.

1. Wenn $A \cap B$ kontextfrei ist, dann sind auch A und B kontextfrei.
2. Wenn A^* kontextfrei ist, dann ist auch A kontextfrei.
3. Wenn $A \subseteq B$ gilt und B kontextfrei ist, dann ist auch A kontextfrei.
4. Wenn $A \cup B$ kontextfrei ist, dann sind A und B kontextfrei.
5. Wenn $A \subseteq B$ gilt und A kontextfrei ist, dann ist auch B kontextfrei.
6. Wenn AB kontextfrei ist, dann sind auch A und B kontextfrei.

Lösung

Alle Aussagen sind falsch. Für die Gegenbeispiele verwenden wir eine beliebige nicht kontexfreie Sprache L, z. B. $L = \{a^p \mid p \text{ prim}\}$.

1. Für $A = \emptyset$ und $B = \{a^p \mid p \text{ prim}\}$ ist $A \cap B = \emptyset$ kontextfrei, aber B nicht.
2. Für $A = \{a^p \mid p \text{ prim}\}$ ist $A^* = \{a\}^*$ kontextfrei, aber A nicht.
3. Für $A = \{a^p \mid p \text{ prim}\}$ und $B = \{a\}^*$ gilt $A \subseteq B$ und B ist kontextfrei, aber A nicht.
4. Für $A = \{a^p \mid p \text{ prim}\}$ und $B = \{a\}^*$ ist $A \cup B = \{a\}^*$ kontextfrei, aber A nicht.
5. Für $A = \emptyset$ und $B = \{a^p \mid p \text{ prim}\}$ gilt $A \subseteq B$, aber B ist nicht kontextfrei.
6. Für $A = \emptyset$ und $B = \{a^p \mid p \text{ prim}\}$ ist $AB = \emptyset$ kontextfrei, aber B nicht.

Präsenzaufgabe 3

Seien $G = (\{S, A, B, C\}, \{a, b\}, P, S)$ eine Grammatik mit den Produktionen

$$\begin{aligned} S &\to BA \mid CA \mid b, \\ A &\to BA \mid a, \\ B &\to CC \mid b, \\ C &\to AB \mid a. \end{aligned}$$

und $w = abba$ ein Wort.

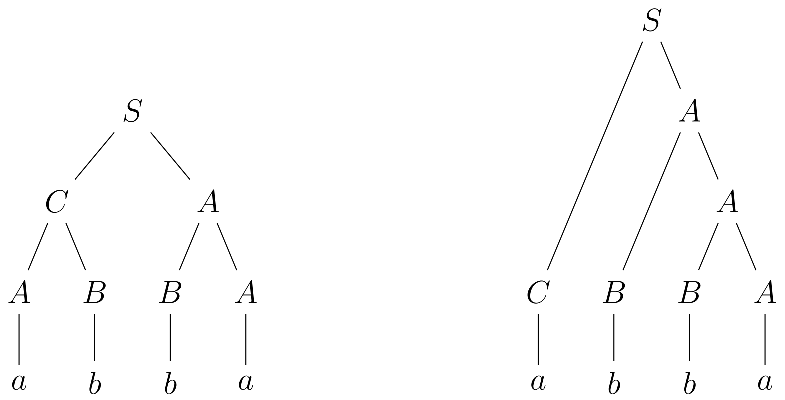
1. Führen Sie den CYK-Algorithmus auf G und w aus.
2. Warum gilt $w \in L(G)$?
3. Geben Sie einen Syntaxbaum für w in G an.
4. Welche Infixe von w sind in L enthalten?

Lösung

1.

$\{S\}$			
$\emptyset$	$\{S, A\}$		
$\{C\}$	$\emptyset$	$\{S, A\}$	
$\{A, C\}$	$\{S, B\}$	$\{S, B\}$	$\{A, C\}$
a	b	b	a

2. Wegen $S \in T_{1,4}$.
3. w besitzt zwei mögliche Syntaxbäume in G:



4. b, ba, bba und $abba$.

Knobelaufgaben

Knobelaufgabe 1

Sei $\Sigma = \{a, b\}$ ein Alphabet. Für Wörter $v, x \in \Sigma^*$ sei $|x|_v$ die Anzahl der Vorkomnisse von v in x. Formal:

$$|x|_v := |\,\{(u, w) \in \Sigma^* \times \Sigma^* \mid x = uvw\}\,|.$$

Beispielsweise gilt $|abababa|_{aba} = 3$ und $|bbbbb|_{bb} = 4$.

Aus Präsenzaufgabe 4 von Ergänzungsblatt 10 wissen wir, dass die Sprache

$$L = \{w \in \Sigma^* \mid |w|_a = |w|_b\}$$

nicht regulär ist.

Was kann über die Regularität der folgenden zwei Sprachen gesagt werden?

1. $A = \{w \in \Sigma^* \mid |w|_{ab} = |w|_{ba}\}$
2. $B = \{w \in \Sigma^* \mid |w|_{aba} = |w|_{bab}\}$

Knobelaufgabe 2

Welche der folgenden Sprachen sind kontextfrei und welche nicht? Beweisen Sie Ihre Antworten.

1. $L_1 = \left\{ a^{k^2+100} \,\middle|\, k \in \mathbb{N} \right\}$ über $\Sigma = \{a\}$
2. $L_2 = \left\{ a^{k^2+\ell} \,\middle|\, k, \ell \in \mathbb{N} \right\}$ über $\Sigma = \{a\}$