Lösungen zum Ergänzungsblatt 8

Hinweis:

Um zu zeigen, dass eine gegebene Sprache in einer bestimmen Sprachklasse liegt, ist es in der Regel ausreichend, eine konkrete Repräsentation (Grammatik, Automat, o. Ä.) anzugeben, wenn ein Korrektheitsbeweis nicht explizit gefordert wird.

Bei kontextfreien Grammatiken reicht es aus, dass jede Produktion eine einzige Variable als linke Seite besitzt, da solche Grammatiken sehr leicht so modifiziert werden können, dass sie die ε-Sonderregel einhalten.

Vorbereitungsaufgaben

Vorbereitungsaufgabe 1

Üben Sie das Schema zur Herstellung einer Grammatik in CNF anhand der Beispiele auf den Vorlesungsfolien 20.2 bis 20.6.

Lösung

Lösungen auf den jeweiligen Folien.

Vorbereitungsaufgabe 2

Seien $\Sigma = \{a, b, c, d\}$ und $\Gamma = \{a, b\}$ zwei Alphabete und $\varphi\colon \Sigma^* \to \Gamma^*$ ein Homomorphismus mit $\varphi(a) = ab$, $\varphi(b) = ba$, $\varphi(c) = aba$ und $\varphi(d) = a$.

1. Bestimmen Sie $\varphi(\{\varepsilon, b, aa, cb, adb\})$.
2. Bestimmen Sie $\varphi^{-1}(\{\varepsilon, bb, aba, ababa\})$.

Erinnerung: Für eine Funktion $f\colon A \to B$ und Mengen $A' \subseteq A$ und $B' \subseteq B$ ist

- $f(A') = \{f(a) \mid a \in A'\}$ das *Bild* von A' unter f und
- $f^{-1}(B') = \{a \mid f(a) \in B'\} = \bigcup_{b \in B'} f^{-1}(b)$ das *Urbild* von B' unter f.

Man beachte, dass $f(A')$ unf $f^{-1}(B')$ beides Mengen sind. Es gilt: $f(A') \subseteq B$ unf $f^{-1}(B') \subseteq A$. Auch $f^{-1}(b)$ ist für jedes $b \in B$ eine Menge: $f^{-1}(b) = \{a \mid f(a) = b\}$.

Lösung

1. Es gilt:

$$\begin{aligned}\varphi(\{\varepsilon, b, aa, cb, adb\}) &= \{\varphi(\varepsilon), \varphi(b), \varphi(aa), \varphi(cb), \varphi(adb)\}\\ &= \{\varepsilon, ba, abab, ababa, ababa\}\\ &= \{\varepsilon, ba, abab, ababa\}.\end{aligned}$$

2. Es gilt:

$$\begin{aligned}\varphi^{-1}(\{\varepsilon, bb, aba, ababa\}) &= \varphi^{-1}(\varepsilon) \cup \varphi^{-1}(bb) \cup \varphi^{-1}(aba) \cup \varphi^{-1}(ababa))\\ &= \{\varepsilon\} \cup \emptyset \cup \{c, ad, db\} \cup \{ac, cb, aad, adb, dbb\}\\ &= \{\varepsilon, c, ad, ac, cb, db, aad, adb, dbb\}.\end{aligned}$$

Präsenzaufgaben

Präsenzaufgabe 1

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

1. $L = \left\{a^k b^l \,\middle|\, k < l\right\}$ über $\Sigma = \{a, b\}$
2. $L = \left\{a^k b^l c^m \,\middle|\, k < l < m\right\}$ über $\Sigma = \{a, b, c\}$
3. $L = \left\{a^k b^l c^m \,\middle|\, k + l = m\right\}$ über $\Sigma = \{a, b, c\}$
4. $L = \left\{a^{kl} \,\middle|\, k, l \geq 2\right\}$ 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. Wir zeigen mit dem Pumping-Lemma, dass L nicht kontextfrei ist. Sei hierfür $n \in \mathbb{N}$ 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 $uv^iwx^iy \notin L$. $\square$

3. 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.

4. 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!

Bemerkung: Die Sprache $\{\varepsilon, a\}$ ist endlich und somit auch regulär (vgl. Ergänzungsblatt 7, Zusatzaufgabe 4).

Präsenzaufgabe 2

Seien Σ ein Alphabet und L eine reguläre Sprache über Σ. Zeigen Sie, dass folgende Sprachen auch regulär sind.

1. $L_1 = \{x \in \Sigma^* \mid \exists y \in L \colon x \text{ ist Infix von } y\}$
2. $L_2 = \{y \in \Sigma^* \mid \exists x \in L \colon x \text{ ist Infix von } y\}$

Lösung

1. Es gilt:

$$L_1 = \{x \in \Sigma^* \mid \exists u, v \in \Sigma^* \colon uxv \in L\} .$$

<u>Mittels Automaten</u>

Da L regulär ist, existiert ein DEA $M = (Q, \Sigma, \delta, s, F)$ mit $T(M) = L$. Definiere den NEA $M' = (Q, \Sigma, \delta', S, F')$ mit $\delta'(q, a) = \{\delta(q, a)\}$, $S = \left\{\hat{\delta}(s, u) \,\middle|\, u \in \Sigma^*\right\}$ und $F' = \left\{q \in Q \,\middle|\, \exists v \in \Sigma^* \colon \hat{\delta}(q, v) \in F\right\}$.

Dann gilt für jedes $x \in \Sigma^*$:

$$\begin{aligned} w \in L_1 &\iff \exists u, v \in \Sigma^* \colon uxv \in L \\ &\iff \exists u, v \in \Sigma^* \colon \hat{\delta}(s, uxv) \in F \\ &\iff \exists u, v \in \Sigma^* \colon \hat{\delta}(\hat{\delta}(\hat{\delta}(s, u), x), v) \in F \\ &\iff \exists u \in \Sigma^* \colon \hat{\delta}(\hat{\delta}(s, u), x) \in F' \\ &\iff \exists p \in S \colon \hat{\delta}(p, x) \in F' \\ &\overset{(*)}{\iff} \hat{\delta}'(S, x) \cap F' \neq \emptyset \\ &\iff x \in T(M'), \end{aligned}$$

d. h. $T(M') = L_1$. $\square$

Bemerkung: An der Stelle $(*)$ wurde die Gleichung

$$\hat{\delta}'(Q', x) = \left\{ \hat{\delta}(q, x) \,\middle|\, q \in Q' \right\}$$

verwendet. Diese kann sehr leicht für alle $Q' \subseteq Q$ und alle $x \in \Sigma^*$ per Induktion nach der Länge von x gezeigt werden.

<u>Mittels Erkennbarkeit durch Monoide</u>

Da L regulär ist, ist L erkennbar, d. h. es existieren ein endliches Monoid $(M, \cdot)$, ein Homomorphismus $\varphi\colon \Sigma^* \to M$ und eine Teilmenge $A \subseteq M$ mit $L = \varphi^{-1}(A)$.

Wähle $A' = \{a \in M \mid \exists u, v \in \Sigma^*\colon \varphi(u) \cdot a \cdot \varphi(v) \in A\}$. Dann gilt $A' \subseteq M$ und

$$\begin{aligned} x \in L_1 &\iff \exists u, v \in \Sigma^*\colon uxv \in L \\ &\iff \exists u, v \in \Sigma^*\colon uxv \in \varphi^{-1}(A) \\ &\iff \exists u, v \in \Sigma^*\colon \varphi(uxv) \in A \\ &\iff \exists u, v \in \Sigma^*\colon \varphi(u) \cdot \varphi(x) \cdot \varphi(v) \in A \\ &\iff \varphi(x) \in A', \end{aligned}$$

für alle $x \in \Sigma^*$. Daraus folgt $L_1 = \varphi^{-1}(A')$, d. h. L_1 wird auch von M durch φ erkannt. Also ist L_1 erkennbar und somit auch regulär. $\square$

<u>Mittels Abschlusseigenschaften</u>

Betrachte das Alphabet $\widehat{\Sigma} = \{\hat{a} \mid a \in \Sigma\}$ und die Homomorphismen $\varphi, \psi\colon (\Sigma \cup \widehat{\Sigma})^* \to \Sigma^*$ mit $\psi(a) = \varphi(a) = \varphi(\hat{a}) = a$ und $\psi(\hat{a}) = \varepsilon$ für alle $a \in \Sigma$.

Dann gilt:

$$L_1 = \psi(\varphi^{-1}(L) \cap \widehat{\Sigma}^* \Sigma^* \widehat{\Sigma}^*).$$

Da die Klasse der regulären Sprachen abgeschlossen ist unter Konkatenation, Schnitt, Homomorphismen und inversen Homomorphismen, ist auch L_1 regulär. $\square$

Bemerkung: Intuitiv entfernt φ alle Markierungen auf Buchstaben und ψ entfernt alle markierten Buchstaben, z. B. $\varphi(b\hat{c}a) = bca$ und $\psi(b\hat{c}a) = ba$.

<u>Mittels syntaktischer Kongruenz</u>

Wir zeigen, dass $\equiv_L$ eine Verfeinerung von $\equiv_{L_1}$ ist, indem wir für alle $w_1, w_2 \in \Sigma^*$ die Implikation

$$w_1 \equiv_L w_2 \implies w_1 \equiv_{L_1} w_2$$

zeigen.

Seien $w_1, w_2 \in \Sigma^*$ beliebig mit $w_1 \equiv_L w_2$. Dann gilt für alle $x, y \in \Sigma^*$:

$$\begin{aligned} xw_1y \in L_1 &\iff \exists u, v \in \Sigma^*\colon uxw_1yv \in L \\ &\iff \exists u, v \in \Sigma^*\colon uxw_2yv \in L \\ &\iff xw_2y \in L_1. \end{aligned}$$

Wegen $|\Sigma^*/{\equiv_{L_1}}| \leq |\Sigma^*/{\equiv_L}| < \infty$ ist $\mathrm{Synt}(L_1) = \Sigma^*/{\equiv_{L_1}}$ endlich und somit L_1 regulär. $\square$

2. Es gilt:

$$L_2 = \{y \in \Sigma^* \mid \exists x \in L, u, v \in \Sigma^* \colon y = uxv\} = \{uxv \in \Sigma^* \mid x \in L \wedge u, v \in \Sigma^*\}$$

<u>Mit regulären Audrücken</u>

Seien $a_1, \ldots, a_k$ die Buchstaben in Σ und γ ein regulärer Ausdruck für L. Dann wird L_2 von dem regulären Ausdruck $\gamma' = (a_1| \ldots |a_k)^* \gamma (a_1| \ldots |a_k)^*$ beschrieben.

<u>Mit Abschlusseigenschaften</u>

Da Σ^* regulär ist und die Klasse der regulären Sprachen unter Konkatenation abgeschlossen ist, ist $L_2 = \Sigma^* L \Sigma^*$ auch regulär.

Zusatzaufgaben

Zusatzaufgabe 1

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

1. $L = \left\{a^k b^l c^k d^l \,\middle|\, k, l \in \mathbb{N}\right\}$ über $\Sigma = \{a, b, c, d\}$
2. $L = \left\{a^k b^l c^m d^n \,\middle|\, k + m = l + n\right\}$ über $\Sigma = \{a, b, c, d\}$
3. $L = \left\{a^{kl} \,\middle|\, k, l \in \mathbb{N}\right\}$ über $\Sigma = \{a\}$

Lösung

1. 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 $|uv^i wx^i y|_a < |uv^i wx^i y|_c$.

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

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

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

     Für $i = 0$ gilt $|uv^i wx^i y|_c < |uv^i wx^i y|_a$.

   - Fall 4: vx enthält nur ds.

     Für $i = 0$ gilt $|uv^i wx^i y|_d < |uv^i wx^i y|_b$.

   In allen Fällen existiert ein $i \in \mathbb{N}$ mit $uv^i wx^i y \notin L$. □

2. 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.

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

Zusatzaufgabe 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 fahren fort wie auf den Vorlesungsfolien 21.1 bis 21.4 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:

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

<u>Schritt 2</u> (zweiter Algorithmus)

Der zweite Algorithmus liefert:

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

<u>Schritt 3</u> (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:

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

<u>Schritt 4</u> (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.