Institut für Informatik
M. Mutz

Übungsblatt 5

Theoretische Informatik

Bearbeitungszeit: 13.05.2024 bis 19.05.2024, 16:00 Uhr
Besprechung: TBD (es ist Pfingstmontag), 10:30 Uhr in Hörsaal 5E

Abgabe: als PDF über das ILIAS
Gruppenabgaben möglich und erwünscht

Aufgabe 1 (Pumping-Lemma)10 Punkte

Zeigen Sie mit Hilfe des Pumping-Lemmas, dass die folgende Sprache nicht regulär ist.

$L_1 = \{0^n 1^m 0^k \mid n, m, k \geq 0, m - n < k\}$ über $\Sigma = \{0, 1\}$
Hinweise: Achten Sie dabei auf eine gut ersichtliche Beweisstruktur (Was ist zu zeigen, was wird zu einem Widerspruch geführt, etc.), darauf, dass alle Einzelschritte nachvollziehbar sind (Führen Sie verwendete Regeln auf, begründen Sie, warum Sie mit diesem Wort/dieser Zahl argumentieren dürfen, welche Eigenschaften eine Variable hat, etc.) und definieren Sie alle verwendeten Variablen.

Lösungsvorschlag:
z.z.: $L_1 = \{0^n 1^m 0^k \mid n, m, k \geq 0, m - n < k\} \notin$ REG
Beweis: Wir beweisen, dass die Sprache L_1 nicht regulär ist, mit Hilfe eines Widerpruchbeweis. Angenommen, $L_1 \in$ REG. Sei $s \geq 1$ die Zahl, die nach dem Pumping-Lemma für L_1 existiert. Betrachte das Wort $x = 0^0 1^s 0^{s+1} = 1^s 0^{s+1} \in L_1$ mit der Länge $\mid x \mid = 2s + 1 \geq s$. Da $x \in L_1$, lässt sich x so in $x = uvw = 1^s 0^{s+1}$ zerlegen, dass die Bedingungen (1) - (3) aus dem Pumping-Lemma für reguläre Sprachen gelten. Aus (1) und (2) folgt:

1. $\mid uv \mid \leq s$, d.h. $uv = 1^p$ für $p \leq s$.
2. $\mid v \mid \geq 1$, d.h. $v = 1^q$ mit $q \geq 1$.
3. $(\forall i \geq 0)[uv^i w \in L]$.

Insbesondere gilt für $i = 2$ nach (3):

$$uv^2 w = 1^{p-q} 1^{2q} 1^{s-p} 0^{s+1} = 1^{p-q+2q+s-p} 0^{s+1} = 0^0 1^{s+q} 0^{s+1} \in L$$

Andererseits gilt:

$$(s+q) - 0 \geq s+1.$$

Dies ist aber ein Widerspruch zur Definition von L_1. Also ist die Annahme falsch und somit folgern wir, dass L_1 nicht regulär ist.

Aufgabe 2 (Chomsky-Normalform)10 Punkte

Betrachten Sie die Grammatik $G = (\Sigma, N, S, P)$ mit $\Sigma = \{a, b, c, d, e, f\}$, $N = \{S, A, B, C, E, F\}$ und

$$\begin{aligned} P = \{S &\to A \mid BCd, \\ A &\to BC \mid E, \\ B &\to b \mid Bb, \\ C &\to c \mid \lambda, \\ E &\to e \mid F, \\ F &\to f \mid A\}. \end{aligned}$$

Geben Sie eine zu G äquivalente Grammatik G' in CNF an. Verwenden Sie dazu die Konstruktionsschritte der Vorlesung.

Lösungsvorschlag:
Wir entfernen die λ-Regel:

$$\begin{aligned} P_{\lambda-frei} = \{S &\to A \mid BCd \mid \mathbf{Bd} \\ A &\to BC \mid \mathbf{B} \mid E, \\ B &\to b \mid Bb, \\ C &\to c \cancel{\mid \lambda}, \\ E &\to e \mid F, \\ F &\to f \mid A\} \end{aligned}$$

Identifiziere Zyklus $A \to E \to F \to A$:

$$
\begin{aligned}
P_{\circ-frei} = \{S &\to \cancel{A}\ \mathbf{X} \mid BCd \mid Bd,\\
\cancel{A}\ \mathbf{X} &\to BC \mid B \cancel{\mid E},\\
B &\to b \mid Bb,\\
C &\to c,\\
\cancel{E}\ \mathbf{X} &\to e \cancel{\mid F},\\
\cancel{F}\ \mathbf{X} &\to f \cancel{\mid A}\}
\end{aligned}
$$

Einfache Regeln entfernen (hier zwei Schritte, erst $X \to B$, dann $S \to X$):

$$
\begin{aligned}
P_{schwierig} = \{S &\to\ X \mid BCd \mid Bd\ ,\\
X &\to BC \mid \cancel{B \mid}\ \mathbf{b} \mid \mathbf{Bb} \mid e \mid f,\\
B &\to b \mid Bb,\\
C &\to c\}
\end{aligned}
$$

$$
\begin{aligned}
P_{schwierig} = \{S &\to\ \cancel{X \mid}\ BCd \mid Bd \mid \mathbf{BC} \mid \mathbf{Bb} \mid \mathbf{b} \mid \mathbf{e} \mid \mathbf{f},\\
X &\to BC \mid b \mid Bb \mid e \mid f,\\
B &\to b \mid Bb,\\
C &\to c\}
\end{aligned}
$$

optional: Wegschmeißen von Regeln, die nicht erreichbar sind:

$$
\begin{aligned}
P_{erreichbar} = \{S &\to BCd \mid Bd \mid BC \mid Bb \mid b \mid e \mid f,\\
B &\to b \mid Bb,\\
C &\to c\}
\end{aligned}
$$

letzter Schritt: Nichtterminale Einfügen, Regeln in Form $A \to BC$ oder $A \to d$ bringen.

$$
\begin{aligned}
P_{CNF} = \{ & S \to \mathbf{BX_0} \mid \mathbf{BD} \mid BC \mid \mathbf{BB_1} \mid b \mid e \mid f, \\
& \mathbf{X_0} \to \mathbf{CD}, \\
& B \to b \mid \mathbf{BB_1}, \\
& \mathbf{B_1} \to \mathbf{b}, \\
& C \to c, \\
& \mathbf{D} \to \mathbf{d} \}
\end{aligned}
$$

Also ist die äquivalente Grammatik in CNF $G' = (\{a, b, c, d, e, f\}, N', S, P_{CNF})$ mit $N' = \{S, X_0, B, B_1, C, D\}$

Aufgabe 3 (Pumping Lemma für kontextfreie Sprachen)20 Punkte

Zeigen Sie anhand des Pumping-Lemmas, dass die folgende Sprache nicht kontextfrei ist:

(i) $L_1 = \{a^i b^j \mid i \geq 0, j = i^2\}$ über $\Sigma = \{a, b\}$

(ii) $L_2 = \{a^i b^i c a^i b^i \mid i \geq 0\}$ über $\Sigma = \{a, b\}$

Lösungsvorschlag:

(i)

Gegeben: Die Sprache $L_1 = \{a^i b^j \mid i \geq 0, j = i^2\}$
z.z.: L_1 ist keine kontextfreie Sprache.
Beweis: Wir zeigen $L_1 \notin$ CF durch Widespruchsbeweis. Wir nehmen an, dass L_1 eine kontextfreie Sprache ist. Sei n die Zahl, die nach dem Pumping-Lemma für kontextfreie Sprachen für L_1 existiert. Wir betrachten dann das Wort $z = a^n b^{n^2} \in L_1$ mit $|z| = n^2 + n \geq n$. Nach dem Pumping-Lemma lässt sich das Wort z so zerlegen ($z = uvwxy = a^n b^{n^2}$), dass die folgenden drei Bedingungen gelten:

1. $|vwx| \leq n$,
2. $|vx| \geq 1$ und
3. $(\forall i \geq 0)[uv^i wx^i y \in L_1]$.

Insbesondere gilt für $i = 2$:

$$|uv^2wx^2y| = |uvwxy| + |vx| \overset{(1)}{\leq} n^2 + n + n < n^2 + 2n + n + 2 = (n+1)^2 + (n+1)$$

Da $n^2 + n < |uv^2wx^2y| < (n+1)^2 + (n+1)$ gilt, ist die Länge des Wortes uv^2wx^2y eine Zahl, die nicht gleich $i^2 + i$ für ein $i \geq 0$ ist. Daraus können wir folgern, dass

$uv^2wx^2y \notin L_1$, und hiermit folgt ein Widerspruch zur Annahme, da Bedingung (3) verletzt ist. Also ist L_1 keine kontextfreie Sprache.

(ii)

Gegeben: Die Sprache $L_2 = \{a^ib^ica^ib^i \mid i \geq 0\} \subseteq \{a, b, c\}^*$
z.z.: L_2 ist keine kontextfreie Sprache.
Beweis: Wir zeigen $L_2 \notin$ CF durch Widespruchsbeweis. Wir nehmen an, dass L_2 eine kontextfreie Sprache ist. Sei n die Zahl, die nach dem Pumping-Lemma für kontextfreie Sprachen für L_2 existiert. Wir betrachten dann das Wort $z = a^nb^nca^nb^n \in L_2$ mit $|z| = 4n + 1 \geq n$. Nach dem Pumping-Lemma lässt sich das Wort z so zerlegen ($z = uvwxy = a^nb^nca^nb^n$), dass die folgenden drei folgenden Bedingungen gelten:

1. $|vwx| \leq n$,
2. $|vx| \geq 1$ und
3. $(\forall i \geq 0)[uv^iwx^iy \in L_2]$.

Insbesondere gilt für $i = 0$:

$$uv^0wx^0y = uwy \in L_2$$

Um die Zerlegungen und das aus 3. neu erstelle Wort besser untersuchen zu können, schauen wir uns die folgenden Fälle an:

(a) $c \sqsubseteq v \lor c \sqsubseteq x$
$\implies c \not\sqsubseteq u \land c \not\sqsubseteq w \land c \not\sqsubseteq y$
$\implies c \not\sqsubseteq uwy$
$\implies uwy \notin L_2$

(b) $c \not\sqsubseteq v \lor c \not\sqsubseteq x \Leftrightarrow c \sqsubseteq uwy$

   (a) $c \sqsubseteq u$
   $\implies u = a^nb^nc \implies vwxy = a^nb^n$
   $\implies v^0wx^0y = a^lb^k | l + k < 2n$ (da $|vx| \geq 1$)
   $\implies uv^0wx^0y = a^nb^nca^lb^k$
   da aber $a^nb^n \neq a^lb^k \implies uwy \notin L_2$

   (b) $c \sqsubseteq y$
   $\implies y = ca^nb^n \implies uvwx = a^nb^n$

$\implies uv^0wx^0 = a^lb^k | l+k < 2n$ (da $|vx| \geq 1$)
$\implies uv^0wx^0y = a^lb^kca^nb^n$
da aber $a^nb^n \neq a^lb^k \implies uwy \notin L_2$

(c) $c \sqsubseteq w$
zusammen mit 2. ($|vwx| \leq n$) ergibt sich
$v = b^l, x = a^s$ mit $1 \leq l + s$
$\implies uv^0wx^0 = a^nb^{n-l}ca^{n-s}b^n$
da entweder $b^{n-l} \neq b^n$ oder $a^{n-s} \neq a^n$ gilt auch hier: $uwy \notin L_2$

Da wir nun für jede mögliche Zerlegung gezeigt haben, dass die nach Pumping Lemma ebenfalls enthaltenen Wörter offenbar nicht in der Sprache liegen, können wir folgern, dass unsere Annahme falsch war und L_2 keine kontextfreie Sprache ist.