Formale Systeme, Automaten, Prozesse, SS 2020
Prof. Dr. P. Rossmanith
J. Dreier, T. Hartmann, H. Lotze, D. Mock

Datum: 8.6.2020

Übungsblatt mit Lösungen 09

Aufgabe T24

Welche der folgenden Sprachen sind kontextfrei, welche nicht? Beweisen Sie jeweils Ihre Behauptung.

1. $L_1 = \{\, a^i b^j c^k \mid i + j = k \,\}$
2. $L_2 = \{\, a^i b^j c^k \mid ij = k \,\}$
3. $L_3 = \{\, pp \mid p \in \{a, b\}^* \,\}$
4. $L_4 = \{\, pqp^R \mid |p| = |q|, p, q \in \{a, b\}^* \,\}$

Lösungsvorschlag

1. Die Sprache L_1 wird durch die folgende Grammatik erkannt:

$$S \to aSc \mid B \mid \epsilon \qquad B \to bBc \mid \epsilon$$

2. Es lässt sich über das Pumping-Lemma zeigen, dass L_2 nicht kontextfrei ist. Für beliebige n, sei das Wort $z = a^n b^n c^{n^2}$. Es liegt offensichtlich in L_2. Es gibt nun viele verschiedene Möglichkeiten, z in $uvwxy$ zu zerlegen. Da aber $|vwx| \leq n$ gelten muss, kann vwx nicht alle drei Terminale (a, b und c) gleichzeitig enthalten. Man sieht leicht, dass ein Aufpumpen zwangsweise dazu führt, dass das entstehende Wort nicht mehr in L_2 liegt.

3. Auch bei L_3 lässt sich anhand des Pumping-Lemmas zeigen, dass diese Sprache nicht kontextfrei ist. Sei n beliebig und $z = a^n b^n a^n b^n \in L_3$. Sei $z = uvwxy$ eine entsprechende Zerlegung. Wir unterscheiden zwischen zwei Fällen:

   1. Fall: vwx ist komplett in der ersten Hälfte oder komplett in der zweiten Hälfte von z enthalten. Durch Pumpen von v und x ändert sich daher eine Hälfte $a^n b^n$ und die andere Hälfte bleibt gleich. Das so entstandene Wort ist somit nicht in L_3 enthalten.

   2. Fall: vwx liegt im Übergang der beiden Wörter, also ist es ein Unterwort von $b^n a^n$. Das heißt, beim pumpen muss sich das erste b^n und das zweite a^n. Da $|vwx| \leq n$, ändert sich jedoch weder das erste a^n noch das zweite b^n. Somit ist das gepumpte Wort nicht mehr in L_3, da es die Form $a^n b^{n+i} a^{n+j} b^n$ mit $i + j \geq 1$ hat.

4. Bei L_4 können wir schon wieder das Pumping-Lemma benutzen, zum Beispiel für das Wort $z = a^n ab^{n-1} a^n \in L_4$ mit $p = p^R = a^n$ und $q = ab^{n-1}$. Für eine Zerlegung $z = uvwxy$ mit $|vwx| \leq n$ gilt $vwx \in L(a^*b^* + b^*a^*)$. Somit können wir beim Pumpen mit $i >= 2$ zwischen drei Fällen unterscheiden:

   1. Fall: Die Länge von q ändert sich, p und p^R bleiben gleich.

   2. Fall: Die Länge von p ändert sich, p^R bleibt gleich.

   3. Fall: Die Länge von p^R ändert sich, p bleibt gleich.

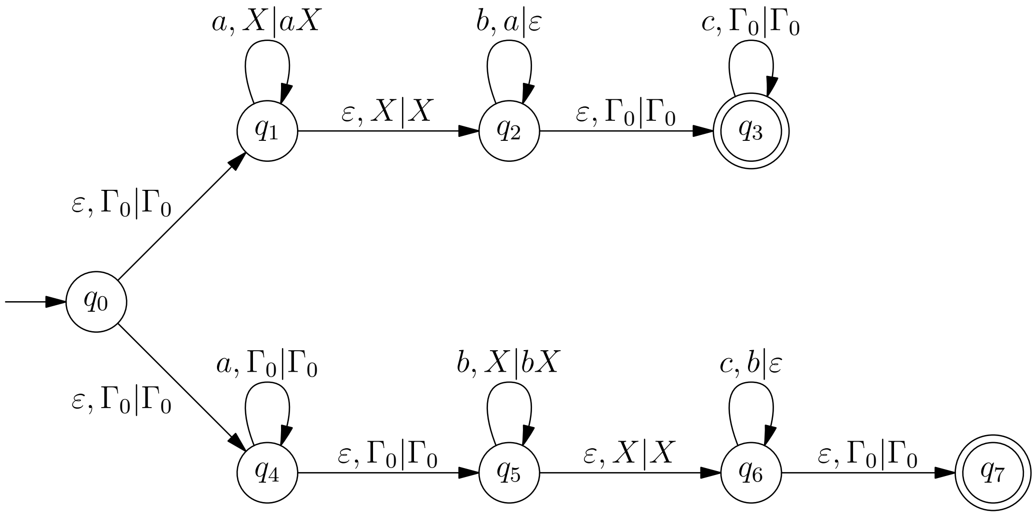
   In jedem der drei Fälle wird das Wort so verändert, dass es nicht mehr in der Sprache ist. Folglich ist L_4 nicht kontextfrei.

Aufgabe T25

Es sei $L = \{ a^i b^j c^k \mid i = j \text{ oder } j = k \}$. Geben Sie einen nichtdeterministischen Kellerautomaten an, der L mit Endzuständen akzeptiert.

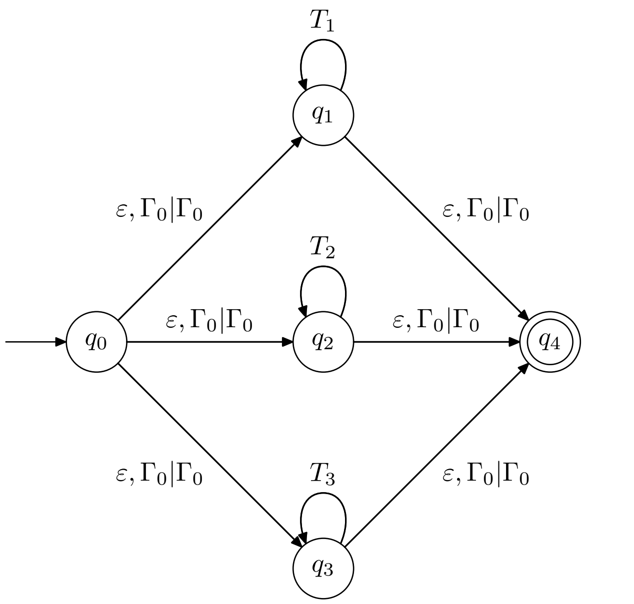
Lösungsvorschlag

Der Kellerautomat rät erst nichtdeterministisch, ob $i = j$ oder $j = k$ gilt. Danach wird der Stack zum Zählen der entsprechenden Symbole genutzt. Hierbei stehe X für ein beliebiges Stacksymbol.



Aufgabe T26

Welche Sprache $L(M)$ erkennt der folgende Kellerautomat? Beweisen Sie Ihre Behauptung. Die Transitionsmengen T_1, T_2, T_3 sind aus der entsprechenden Spalte der Tabelle zu entnehmen.



T_1:	T_2:	T_3:
$a, a\|aa$	$b, b\|bb$	$a, a\|aa$
$b, b\|bb$	$c, c\|cc$	$c, c\|cc$
$a, b\|\varepsilon$	$b, c\|\varepsilon$	$c, a\|\varepsilon$
$b, a\|\varepsilon$	$c, b\|\varepsilon$	$a, c\|\varepsilon$
$a, \Gamma_0\|a\Gamma_0$	$b, \Gamma_0\|b\Gamma_0$	$c, \Gamma_0\|c\Gamma_0$
$b, \Gamma_0\|b\Gamma_0$	$c, \Gamma_0\|c\Gamma_0$	$a, \Gamma_0\|a\Gamma_0$
$c, X\|X$	$a, X\|X$	$b, X\|X$

Lösungsvorschlag

Am Anfang entscheidet sich der Kellerautomat nichtdeterministisch für einen von drei Folgezuständen q_1, q_2, q_3, die sich bis auf Vertauschung von a, b, c zueinander symmetrisch verhalten. Betrachten wir exemplarisch den Zustand q_1: Hier haben offenbar nur die Zeichen a und b einen Effekt. Wann immer ein a gelesen wird, wird entweder ein a auf den Keller gelegt oder ein b vom Keller entfernt. Duales gilt für b. Folglich entspricht die Differenz der Vorkommen von a und b

auf dem Keller der Differenz der Vorkommen von a und b im bisherigen Wort. Andererseits sieht man, dass auf dem Keller niemals gleichzeitig die Zeichen a und b vorkommen können. Folglich ist der Keller genau dann leer, wenn das bisher gelesene Wort gleich viele a wie b enthält.

Weiterhin ist bei leerem Keller ein ε-Übergang in den akzeptierenden Zustand möglich. Für ein Wort w und ein Zeichen a bezeichne $|w|_a$ dabei die Anzahl von a in w. Der gegebene Kellerautomat erkennt also die Sprache aller Wörter w über $\{a, b, c\}$, die *mindestens eine* der folgenden Bedingungen erfüllen:

- $|w|_a = |w|_b$,
- $|w|_b = |w|_c$,
- $|w|_a = |w|_c$

Aufgabe H30 (10 Punkte)
Welche der folgenden Sprachen sind kontextfrei, welche nicht? Geben Sie hierfür eine kontextfreie Grammatik oder einen Kellerautomaten an oder beweisen Sie, dass die Sprache nicht kontextfrei ist.

1. $L_1 = \{ a^i b^j a^k b^\ell \mid i + j = k + \ell \}$
2. $L_2 = \{a^{2^n} \mid n \geq 0\}$,
3. $L_3 = \{w_1 \dots w_{2m} \mid m \geq 0;\ w_i \in \{a, b, c\}, w_{i+m} \in \{a, b, c\} \setminus \{w_i\} \text{ für alle } i \in \{1, \dots, m\}\}$.

Lösungsvorschlag

1. L_1 ist kontextfrei, da folgende kontextfreie Grammatik L_1 erzeugt:
$$S \to aSb \mid S_{aa} \mid S_{bb}$$
$$S_{aa} \to aS_{aa}a \mid S_{ba}$$
$$S_{bb} \to bS_{ba}b \mid S_{ba}$$
$$S_{ba} \to bS_{ba}a \mid \epsilon$$
2. L_2 ist nicht kontextfrei: Angenommen L_2 ist kontextfrei, dann gilt das Pumping-Lemma für CFL für $n \geq 0$. Betrachte das Wort $a^{2^{n+2}}$, welches $|a^{2^{n+2}}| \geq n$ erfüllt. Nach Pumping-Lemma gibt es Zerlegung $uvwxy = a^{2^{n+2}}$ mit $|vwx| \leq n$ und $|vx| > 0$ so dass für alle $i \geq 0$ auch $uv^iwx^iy \in L_2$. Betrachte das Wort $uv^2wx^2y = a^{2^{n+2}+|vx|}$. Da $0 < |vx| \leq n$, ist $2^{n+2} < 2^{n+2} + |vx| < 2^{n+3}$ und somit $uv^2wx^2y \notin L_2$.
3. L_3 ist nicht kontextfrei: Angenommen L_3 ist kontextfrei, dann gilt das Pumping-Lemma für CFL für $n \geq 0$. Betrachte das Wort $z = a^nb^nb^na^n$, welches $|z| \geq n$ erfüllt. Nach Pumping-Lemma gibt es Zerlegung $uvwxy = a^nb^nb^na^n$ mit $|vwx| \leq n$ und $|vx| > 0$ so dass für alle $i \geq 0$ auch $uv^iwx^iy \in L_3$. Da $|vwx| \leq n$, bildet vwx ein Teilwort der ersten $2n$ Symbole, der $2n$ Symbole gefolgt einschließlich vom ersten b, oder der letzten $2n$ Symbole.

   Falls vwx ein Teilwort der der $2n$ Symbole gefolgt einschließlich vom ersten b, so enthält $z' = uv^0wx^0y$ weniger b's als a's im Widerspruch zu der Definition von L_3.

   Falls vwx ein Teilwort der ersten $2n$ Symbole, so bleibt der Suffix a^n unverändert. Dann muss das Wort $z' = uv^0wx^0y$ die Form $z' = a^jb^na^nba^n$ haben. Dann muss aber gelten $j = n$ sodass $vx = \varepsilon$ im Widerspruch zum Pumping-Lemma,

   Der Fall, dass vwx ein Teilwort der letzten $2n$ Symbole ist, führt analog zu einem Widerspruch.

Aufgabe H31 (10 Punkte)
Für ein Wort w und ein Zeichen a bezeichne $|w|_a$ dabei die Anzahl von a in w.

$$L := \{ w \in \{a, b, c\}^* \mid w = uv \text{ mit } 3|u|_a = |v|_b + |v|_c \vee |u|_c = 2|v|_a \} \subseteq \{a, b, c\}^*$$

a) Geben Sie einen Kellerautomaten für L an.

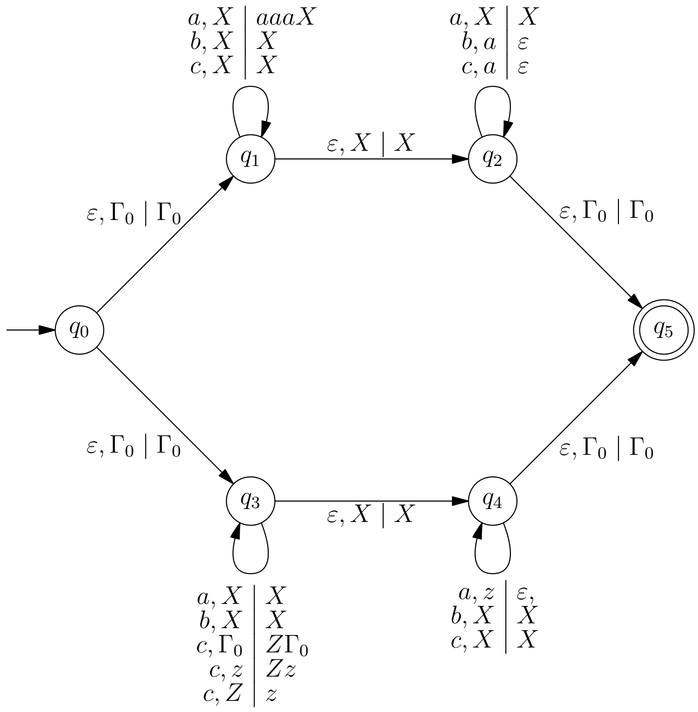
b) Beweisen Sie die Korrektheit und Vollständigkeit des Automaten.

Lösungsvorschlag

a)

Anfangs entscheidet sich der Automat nichtdeterministisch für die Bedingung $3|u|_a = |v|_b + |v|_c$ oder $|u|_c = 2|v|_a$. Im folgenden beschreiben wir, wie der Automat die erste Bedingung richtig erkennt Wir legen zuerst für jedes gelesene a zwei a auf den Keller und ignorieren jedes gelesene b oder c. Dann wechseln wir nichtdeterministisch in eine zweite Phase (Zustand q_2), in der wir für jedes gelesene b oder c ein a vom Keller nehmen und gelesene a ignorieren. Wenn nur noch das Kellerbodensymbol auf dem Keller liegt, können wir in den Endzustand wechseln.

Um die zweite Bedingung zu überprüfen, legen wir für jedes zweite gelesene c in w das Symbol z auf den Keller ab. Dies bewerkstelligen wir, indem wir für das erste gelesene c ein Z ablegen und für das zweite ersetzen wir das Z auf dem Keller durch z. Gelesene a und b werden ignoriert. Dann wechseln wir nichtdeterministisch in eine zweite Phase (Zustand q_4) und gehen sicher, dass wir eine gerade Anzahl von c gelesen haben. In dieser Phase nehmen wir für jedes gelesene a ein z vom Keller und ignorieren gelesene b und c. Wenn nur noch das Kellerbodensymbol auf dem Keller liegt, können wir in den Endzustand wechseln.



b)

Behauptung: Für ein $w \in \{a, b, c\}^*$ gilt $w \in L$ gdw. $(q_0, w, \Gamma_0) \vdash^* (q_5, \varepsilon, \beta)$.

Sei $w \in L$. Dann kann das Wort zerlegt werden in $w = uv$ mit den oben beschriebenen Bedingungen. Im Falle der ersten Bedingung erhalten wir folgende Konfigurationsfolge

$$(q_0, w, \Gamma_0) \vdash^* (q_1, v, a^{3|u|_a}\Gamma_0) \vdash (q_2, v, a^{3|u|_a}\Gamma_0) \vdash^* (q_2, \varepsilon, \Gamma_0) \vdash (q_5, \varepsilon, \Gamma_0),$$

d.h. der Automat akzeptiert w. Im Falle der zweiten Bedingung erhalten wir analog die Konfigurationsfolge

$$(q_0, w, \Gamma_0) \vdash^* (q_3, v, z^{|u|_c/2}\Gamma_0) \vdash (q_4, v, z^{|u|_a/2}\Gamma_0) \vdash^* (q_4, \varepsilon, \Gamma_0) \vdash (q_5, \varepsilon, \Gamma_0),$$

d.h. der Automat akzeptiert auch hier w.

Angenommen der Automat akzeptiert w. Dann gibt es entweder eine Konfigurationsfolge der Form

$$(q_0, w, \Gamma_0) \vdash^* (q_1, v, a^{3m}\Gamma_0) \vdash (q_2, v, a^{3m}\Gamma_0) \vdash^* (q_2, \varepsilon, \Gamma_0) \vdash (q_5, \varepsilon, \Gamma_0),$$

oder

$$(q_0, w, \Gamma_0) \vdash^* (q_3, v, z^m\Gamma_0) \vdash (q_4, v, z^m\Gamma_0) \vdash^* (q_4, \varepsilon, \Gamma_0) \vdash (q_5, \varepsilon, \Gamma_0),$$

für ein $m \in \mathbf{N}$. Dann kann w zerlegt werden in $w = uv$. Nach Konstruktion des Automaten muss entweder $|v|_b + |v|_c = 3m$ gelten oder $|v|_a = m$. Im ersten Fall folgt aus der Konstruktion des Automaten $|u|_a = m$, im zweiten $|u|_c = m/2$. In jedem Fall gilt $w \in L$.

Aufgabe H32 (10 Punkte)

In der Vorlesung haben wir PDAs wie folgt definiert:
Ein *Kellerautomat* (PDA) $M = (Q, \Sigma, \Gamma, \delta, q_0, \Gamma_0, F)$ ist ein 7-Tupel, wobei

- Q die endliche Menge der *Zustände*,
- Σ das *Eingabealphabet*,
- Γ das *Kelleralphabet*
- $\delta\colon Q \times (\Sigma \cup \{\epsilon\}) \times \Gamma \to 2^{Q\times\Gamma^*}$, wobei jedes Bild eine *endliche* Menge von Paaren ist, die *Übergangsfunktion*
- $q_0 \in Q$ der *Startzustand*
- $\Gamma_0 \in \Gamma$ das *Kellerbodensymbol*
- $F \subseteq Q$ die Menge der Endzustände.

Wenn wir die Zustandsübergangsfunktion wie untenstehend verändern, ändert sich dann die Menge der Sprachen, die von solchen Automaten akzeptiert werden, wenn das Akzeptanzkriterium ein leerer Keller ist? Falls ja, geben Sie eine möglichst präzise Charakterisierung dieser Menge an Sprachen an.

> $\delta\colon Q \times \Sigma \times \Gamma \to 2^{Q\times\Gamma^*}$, wobei jedes Bild eine *endliche* Menge von Paaren ist, die *Übergangsfunktion*

Lösungsvorschlag

Das Akzeptanzverhalten ändert sich genau dann wenn $\varepsilon \in N(M)$.

Jeder Kellerautomat M' mit einer Übergangsfunktion $\delta\colon Q\times\Sigma\times\Gamma \to 2^{Q\times\Gamma^*}$ kann ε nicht akzeptieren. Denn dazu müsste er mindestens einen Schritt gehen um das Kellerboden-Symbol Γ_0 vom Keller zu nehmen. Da die Eingabe jedoch ε kann M' mit der modifizierten Übergangsfunktion jedoch kein Schritt tätigen.

Es bleibt also der Fall zu betrachten, dass $\varepsilon \notin N(M)$. Betrachten wir die äquivalente CFG G mit $L(G) = N(M)$. Da $\varepsilon \notin L(G)$, gibt es eine äuivalente Grammatik G' mit $L(G') = L(G)$ in Greibach-Normalform, daher G' hat Regeln der Form $N \to TN^*$, insbesondere ohne ε-Produktionen. Nun können wir diese analog zur Vorlesung wieder in einen Kellerautomat M' umwandeln, wobei unsere Übergangsfunktion δ' immer ein Zeichen liest. Dazu passen wir die Übergangsfunktion wie folgt an:

$$\delta(q, a, A) = \{(q, \alpha) \mid A \to a\alpha \in P\}, \quad \text{für alle } a \in T, A \in N.$$

Für alle Wörter $w \in \Sigma^*$ gilt $(q, w, S) \vdash^* (q, \varepsilon, \varepsilon)$ genau dann wenn $S \Rightarrow^* w$, analog wie in der Vorlesung. Damit haben wir einen Kellerautomat M' konstruiert mit angepasster Übergangsfunktion und mit $L(M) = L(M')$.