ETH zürich
Departement Informatik

Theoretische Informatik

Prof. Dr. Juraj Hromkovič
Dr. Hans-Joachim Böckenhauer
`https://courses.ite.inf.ethz.ch/theo_inf_21`

Lösungsvorschläge – Blatt 4

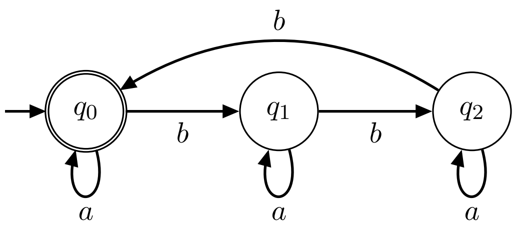
Zürich, 22. Oktober 2021

Lösung zu Aufgabe 10

Die in der Aufgabenstellung gegebene Sprache L lässt sich schreiben als $L = L_1 \cup L_2$ mit

$$\begin{aligned}
L_1 &= \{w \in \{a,b\}^* \mid |w|_a \bmod 3 = |w| \bmod 3\}\\
&= \{w \in \{a,b\}^* \mid |w|_a \bmod 3 = (|w|_a + |w|_b) \bmod 3\}\\
&= \{w \in \{a,b\}^* \mid |w|_b \bmod 3 = 0\}\\
L_2 &= \{w \in \{a,b\}^* \mid w \text{ enthält das Teilwort } ab \text{ und } w \text{ endet mit } b\}\,.
\end{aligned}$$

Für die Sprache L_1 lässt sich der folgende Automat A_1 konstruieren:

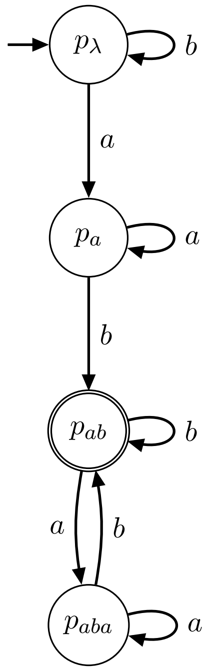


Dieser Automat zählt in seinen Zuständen den Wert von $|w|_b$ modulo 3. Es gilt für $i \in \{0,1,2\}$, dass

$$\mathrm{Kl}[q_i] = \{w \in \{a,b\}^* \mid |w|_b \bmod 3 = i\}\,.$$

(bitte wenden)

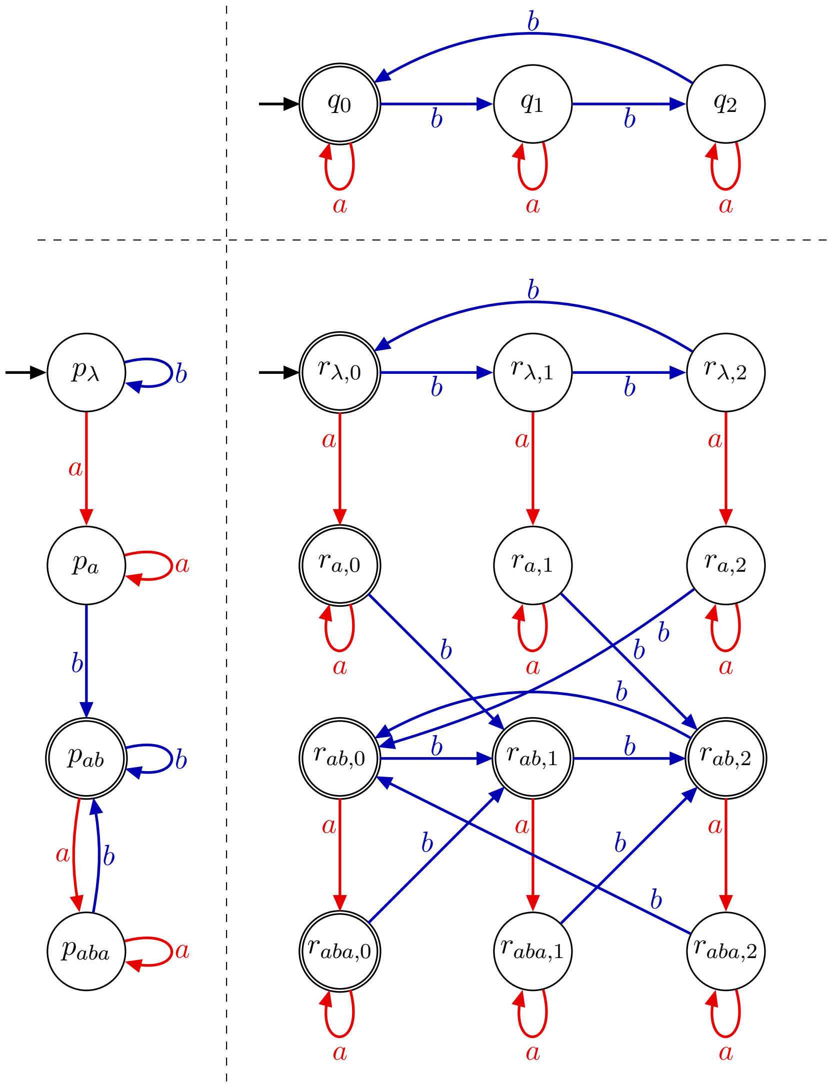
Für die Sprache L_2 lässt sich der folgende Automat A_2 konstruieren:



Die Zustände p_z für $z \in \{\lambda, a, ab\}$ geben an, welches längste Präfix des gesuchten Musters ab aktuell gelesen wurde, p_{ab} ist der akzeptierende Zustand. Der Zustand p_{aba} wird erreicht, wenn das Muster ab bereits gefunden wurde und das aktuell gelesene Präfix des Wortes mit a endet. Damit ergeben sich die folgenden Klassen:

$$\begin{aligned}
\mathrm{Kl}[p_\lambda] &= \{b\}^*\,, \\
\mathrm{Kl}[p_a] &= \{xay \mid x \in \{b\}^* \text{ und } y \in \{a\}^*\}\,, \\
\mathrm{Kl}[p_{ab}] &= L_2\,, \\
\mathrm{Kl}[p_{aba}] &= \{a,b\}^* - \bigcup_{z \in \{\lambda, a, ab\}} \mathrm{Kl}[p_z]\,.
\end{aligned}$$

Mit der Methode des modularen Entwurfs kann man nun aus A_1 und A_2 den folgenden Produktautomaten A konstruieren, der die Sprache $L = L_1 \cup L_2$ akzeptiert. Zur einfacheren Darstellung verwenden wir die Notation $r_{x,y} = \langle p_x, q_y \rangle$. Weil dies ein Produktautomat für die Vereinigung von zwei Sprachen ist, sind alle Zustände akzeptierend, die einen akzeptierenden Zustand aus einem der beiden Teilautomaten enthalten, also die zu q_0 gehörige Spalte und die zu p_{ab} gehörige Zeile.



Lösung zu Aufgabe 11

Um zu zeigen, dass jeder endliche Automat, der die Sprache L akzeptiert, mindestens 5 Zustände benötigt, bestimmen wir 5 Wörter $w_1, \ldots, w_5$ und zeigen, dass diese in einem solchen Automaten in 5 paarweise verschiedene Zustände führen müssen. Falls es einen endlichen Automaten gibt, den das Lesen von zwei verschiedenen Wörtern w_i und w_j in denselben Zustand führt, dann gilt nach Lemma 3.3 aus dem Buch für jedes $z \in \Sigma^*$, dass dann auch $w_i z$ und $w_j z$ den Automaten in denselben Zustand führen. Wir zeigen, dass dies für einen Automaten mit weniger als 5 Zuständen nicht möglich ist, indem wir für je zwei verschiedene Wörter w_i und w_j ein Wort $z_{i,j}$ angeben, so dass

$$w_i z_{i,j} \in L(A) \iff w_j z_{i,j} \notin L(A) \,. \tag{1}$$

Wir wählen die fünf Wörter $w_1 = \lambda$, $w_2 = a$, $w_3 = b$, $w_4 = ab$ und $w_5 = ba$.

(bitte wenden)

Die folgende Tabelle gibt für alle Paare (w_i, w_j) mit $i < j$ jeweils ein Wort $z_{i,j}$ an.

$z_{i,j}$	$w_2 = a$	$w_3 = b$	$w_4 = ab$	$w_5 = ba$
$w_1 = \lambda$	b	a	λ	λ
$w_2 = a$	–	a	λ	λ
$w_3 = b$	–	–	λ	λ
$w_4 = ab$	–	–	–	a

Es ist einfach zu sehen, dass diese Wörter die Bedingung (1) erfüllen, zum Beispiel ist $w_4 z_{4,5} = aba \in L(A)$, aber $w_5 z_{4,5} = baa \notin L(A)$.

Lösung zu Aufgabe 12

(a) Wir zeigen mit Hilfe von Lemma 3.3 aus dem Buch, dass die Sprache

$$L_1 = \{w \in \{a,b,c\}^* \mid w \text{ enthält das Teilwort } ab \text{ gleich oft wie das Teilwort } ba\}$$

nicht regulär ist. Angenommen, L_1 sei regulär. Dann gibt es einen Automaten $A = (Q, \{a,b,c\}, \delta, q_0, F)$ mit $L(A) = L_1$. Sei $m = |Q|$. Wir betrachten die Wörter

$$\lambda, abc, (abc)^2, \ldots, (abc)^m\,.$$

Weil dies $m+1$ Wörter sind, also mehr Wörter als A Zustände hat, gibt es $i, j \in \{0, \ldots, m\}$ mit $i \neq j$, so dass

$$\hat{\delta}(q_0, (abc)^i) = \hat{\delta}(q_0, (abc)^j)\,.$$

Nach Lemma 3.3 gilt nun für alle $z \in \{a,b,c\}^*$, dass

$$(abc)^i z \in L_1 \iff (abc)^j z \in L_1.$$

Die Wahl von $z = (bac)^i$ führt aber zum Widerspruch, weil

$$(abc)^i z = (abc)^i (bac)^i \in L_1$$

und $(abc)^j z = (abc)^j (bac)^i \notin L_1$ gilt. Also ist die Annahme falsch und L_1 ist nicht regulär.

(bitte wenden)

(b) Wir verwenden das Pumping-Lemma, um zu zeigen, dass die Sprache

$$L_2 = \{w \in \{0,1\}^* \mid |w|_0 \neq |w|_1\}$$

nicht regulär ist. Angenommen, L_2 sei regulär. Dann existiert nach dem Pumping-Lemma (Lemma 3.4) eine Konstante $n_0 \in \mathbb{N}$, so dass sich jedes Wort $w \in \{0,1\}^*$ mit $|w| \geq n_0$ in drei Teile y, x und z zerlegen lässt, so dass

1. $|yx| \leq n_0$,
2. $|x| \geq 1$ und
3. entweder $\{yx^kz \mid k \in \mathbb{N}\} \subseteq L_2$ oder $\{yx^kz \mid k \in \mathbb{N}\} \cap L_2 = \emptyset$.

Wir wählen das Wort $w = 0^{n_0}1^{n_0}$. Offenbar gilt $|w| \geq n_0$. Also muss es eine Zerlegung $w = yxz$ von w geben, die die Bedingungen (i), (ii) und (iii) erfüllt. Wegen (i) gilt $|yx| \leq n_0$, also ist $y = 0^l$ und $x = 0^m$ für $l, m \in \mathbb{N}$ mit $l + m \leq n_0$, also insbesondere $m \leq n_0$. Wegen (ii) gilt weiter $m > 0$. Da $w \notin L_2$, gilt nach (iii), dass auch

$$\{yx^kz \mid k \in \mathbb{N}\} = \{0^{n_0+(k-1)m}1^{n_0} \mid k \in \mathbb{N}\} \cap L_3 = \emptyset\,.$$

Dies ist aber ein Widerspruch, da $yx^2z = 0^{n_0+m}1^{n_0} \in L_2$. Also ist die Annahme falsch und L_2 ist nicht regulär.

(c) Wir verwenden die Methode der Kolmogorov-Komplexität, um zu zeigen, dass die Sprache

$$L = L_3 = \{0^{\binom{2n}{n}} \mid n \in \mathbb{N}\}$$

nicht regulär ist. Angenommen, L_3 sei regulär. Für alle $n \geq 1$ definieren wir

$$\Delta_n := \binom{2(n+1)}{n+1} - \binom{2n}{n} = \binom{2n}{n} \cdot \left(\frac{(2n+1)(2n+2)}{(n+1)^2} - 1 \right) = \binom{2n}{n} \cdot \left(3 - \frac{2}{n+1} \right).$$

Für jedes $m \in \mathbb{N}$ ist also $0^{\Delta_m - 1}$ das erste Wort in der Sprache

$$L_{0^{\binom{2m}{m}+1}} = \{y \mid 0^{\binom{2m}{m}+1}y \in L\}\,.$$

Nach Satz 3.1 aus dem Buch existiert eine Konstante c, unabhängig von m, so dass

$$K(0^{\Delta_m - 1}) \leq \lceil \log_2(1+1) \rceil + c = 1 + c\,.$$

Da es nur endlich viele Programme der konstanten Länge höchstens $1 + c$ gibt, aber unendlich viele Wörter der Form $0^{\Delta_m - 1}$, ist dies ein Widerspruch. Also ist die Annahme falsch und L_3 ist nicht regulär.