Prof. Javier Esparza
Kevin Kappelmann, Simon Roßkopf, Lukas Stevens

Technische Universität München
Lehrstuhl für Theoretische Informatik

Einführung in die Theoretische Informatik

Sommersemester 2023 – Hausaufgabenblatt 5

- Bitte beachten Sie, dass in dieser Vorlesung generell Antworten mit Begründung gefordert werden, solange die Aufgabe nicht explizit das Gegenteil sagt.
- Zum Bestehen diese Blattes müssen Sie 50% der Punkte erreichen.
- Update: Es werden die Aufgaben 2(a–c) korrigiert.

AT-Aufgabe H5.1. 0 Punkte

Bearbeiten Sie diese Aufgabe mit Automata Tutor. Konstruieren Sie eine kontextfreie Grammatik für die folgenden Sprachen:

(a) $\{a^i b^j a^j b^i : i, j \geq 0\}$

(b) $\{a^i b^j c^k : j \geq i + k\}$

Lösungsskizze.

(a) $S \to aSb \mid X$
$X \to bXa \mid \epsilon$

(b) $S \to XY$
$X \to aXb \mid Xb \mid \epsilon$
$Y \to bYc \mid \epsilon$

Aufgabe H5.2. (*DER SATZ!*) 3 + 3 + 3 Punkte

Entscheiden Sie, ob folgende Sprachen regulär sind. Bestimmen Sie hierzu die Residualsprachen. Falls die Sprache regulär ist, zeichnen Sie den kanonischen Minimalautomat und beschriften Sie die Zustände mit den entsprechenden Residualsprachen. Falls die Sprache nicht regulär ist, reicht es eine unendliche Menge von Wörtern mit unterschiedlichen Residualsprachen zu bestimmen und zu zeigen, dass die Elemente dieser Menge paarweise verschieden sind.

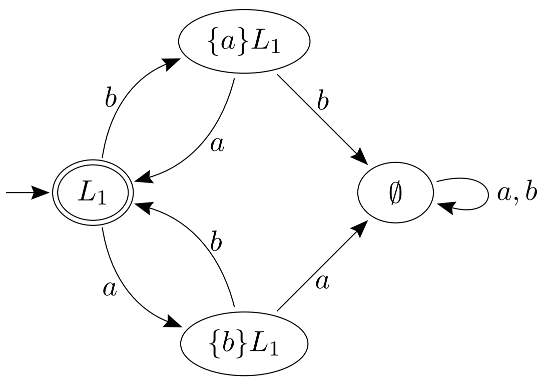
(a) $L_1 := L((\mathsf{ba}|\mathsf{ab})^*)$ mit dem Alphabet $\Sigma := \{a, b\}$

(b) $L_2 := \{a^{2^k} \mid k \in \mathbb{N}\}$ mit dem Alphabet $\Sigma := \{a\}$

(c) $L_3 := \{w \in \Sigma^* \mid |w|_{ab} \neq |w|_{ba}\}$ mit dem Alphabet $\Sigma := \{a, b\}$, wobei $|w|_v$ die Anzahl der Vorkommnisse des Wortes v im Wort w zählt. Beispielsweise gilt $|ababb|_{ab} = 2$ und $|ababb|_{ba} = 1$.

Lösungsskizze.

(a) Kanonischer DFA:

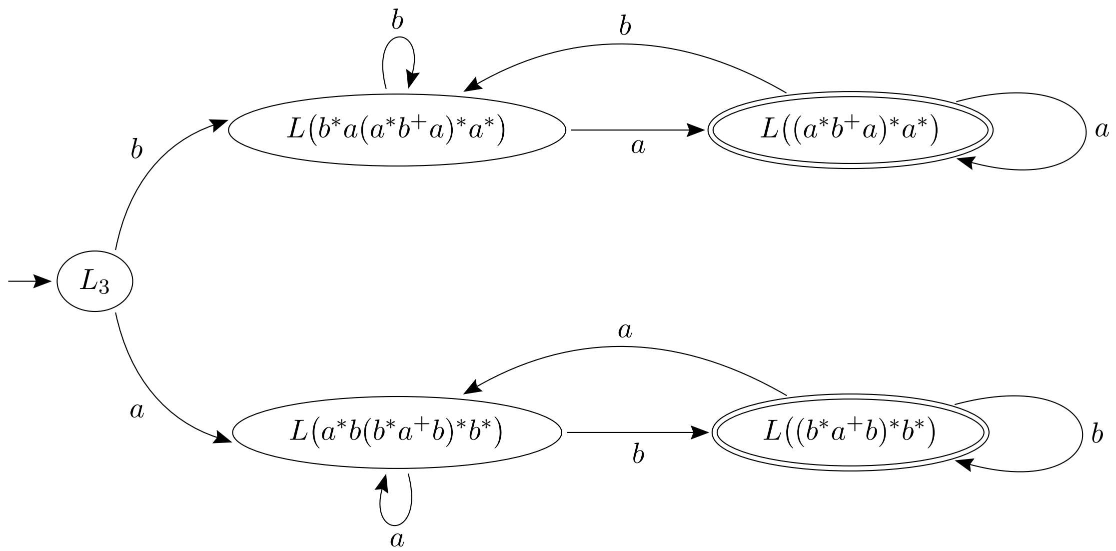


(b) Bestimmen einer unendlichen Menge von Wörtern mit unterschiedlichen Residualsprachen:

$$\{a^{2^i} \mid i \in \mathbb{N}\}$$

Beweis: Sei $i, j \in \mathbb{N}$ verschieden und ohne Einschränkung der Allgemeinheit $i < j$. Dann gilt $a^{2^i}a^{2^i} = a^{2^{i+1}} \in L_2$. Das Wort $w = a^{2^j}a^{2^i}$ hat die Länge $2^i+2^j = 2^i+(2^{j-i}2^i) = 2^i(2^{j-i}+1)$. Damit hat $|w|$ den ungeraden Teiler $2^{j-i}+1$ und ist daher keine Zweierpotenz. Es folgt $w \notin L_2$. Somit $L^{a^{2^i}} \neq L^{a^{2^j}}$ und L_2 ist keine reguläre Sprache.

(c) Kanonischer DFA:



Aufgabe H5.3. (*Eine Runde Konstruktionen für alle*) 2+4+2 Punkte

Sei $\Sigma = \{a, b, c\}$, $L_1 = \{w \mid \exists u, v \in \Sigma^*. w = uabv\}$ die Sprache aller Wörter die das Teilwort ab enthalten, $L_2 = \{c\}\Sigma^*$ die Sprache aller Wörter die mit c beginnen und $L_3 = L_2 \cap \overline{L_1}$.

Geben Sie

(a) DFAs für L_1, L_2

(b) einen DFA für L_3

(c) einen minimalen DFA für L_3

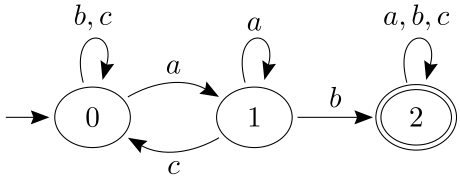
(d) eine rechtslineare Grammatik für L_3

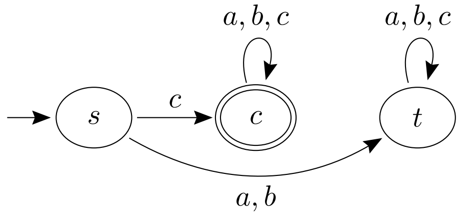
an.

Hinweis: Verwenden Sie jeweils das Ergebnis der vorherigen Teilaufgabe

Lösungsskizze.

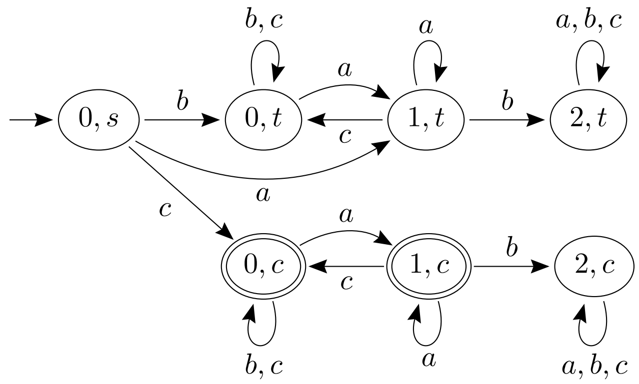
(a) (Minimale) DFAs für L_1, L_2:



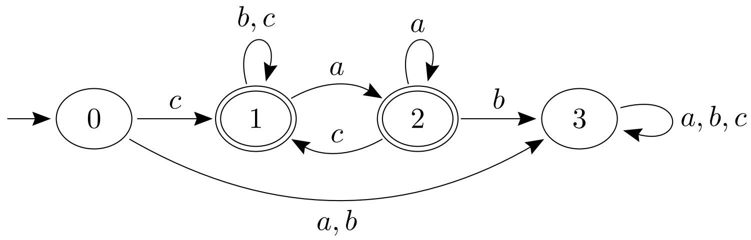


(b) zum Beispiel, nach VL aus (a) erzeugt:

DFA für L_3:



(c) Minimaler DFA für L_3 (aus (b) erzeugt, Zustände umbennant)



(d) zum Beispiel, nach VL aus (c) erzeugt:

Rechtslineare Grammatik für L_3: $([3], \Sigma, P, 0)$, P gegeben durch

- $0 \to a3 \mid b3 \mid c1 \mid c$
- $1 \to a2 \mid b1 \mid c1 \mid a \mid b \mid c$
- $2 \to a2 \mid b3 \mid c1 \mid a \mid c$
- $3 \to a3 \mid b3 \mid c3$

Knobelaufgabe H5.4. (*Irrational irregulär*) nur Respektpunkte, keine echten Punkte

Dr. Evilsparza macht es sich diese Woche gemütlich, um sich von all den Strapazen zu erholen. Damit euch allerdings nicht langweilig wird, hat sich sein Nachwuchsschurke – Mr. Czerner – eine Knobelaufgabe für euch überlegt.

Sei $\Sigma := \{\texttt{0}, ..., \texttt{9}\}$. Für Wörter $u, v \in \Sigma^*$ bezeichnen wir mit $(u.v)_{10} := (u)_{10} + 10^{-|v|} \cdot (v)_{10}$ den Wert eines Dezimalbruches, es gilt also z.B. $(\texttt{1.75})_{10} = \frac{7}{4}$.

Sei $\gamma \in \mathbb{R}$, $0 \leq \gamma < 1$. Zeigen Sie, dass $L(\gamma) := \{w \in \Sigma^* : (\texttt{0.}w)_{10} \leq \gamma\}$ genau dann regulär ist, wenn γ rational ist.

Lösungsskizze. Zunächst stellen wir folgende Beobachtungen über die Residualsprachen fest: Für $c \in \Sigma$ gilt

$$L(x)^c = \{w : (\texttt{0.}cw)_{10} \leq x\} = \{w : (\texttt{0.}w)_{10} \leq 10x - c\} = L(10x - c)$$

Analog folgt $L(x)^w = L(10^{|w|}x - (w)_{10})$ für $w \in \Sigma^*$. „$\Longleftarrow$“: Da $\gamma \in \mathbb{Q}$ gilt, gibt es $k, n \in \mathbb{N}$ mit $\gamma = \frac{k}{n}$. Wir wollen nun zeigen, dass $L(\gamma)$ endlich viele Residualsprachen hat. Insbesondere behaupten wir $\{L(\gamma)^w \mid w \in \Sigma^*\} \subseteq \{L(\frac{i}{n}) \mid i \in \{0,...,n\}\}$. Diese Behauptung zeigen wir mit Induktion über die Länge von w. Die Basis folgt direkt aus $L(\gamma) = L(\frac{k}{n})$ (man bemerke $0 \leq \gamma \leq 1$). Für den Schritt sei $w \in \Sigma^*$ und $c \in \Sigma$ beliebig. Nach Annahme gilt $L(\gamma)^w = L(\frac{i}{n})$ für ein $i \in \mathbb{N}$.

$$L(\gamma)^{wc} = L(\tfrac{i}{n})^c = L(10 \cdot \tfrac{i}{n} - c) = L(\tfrac{10i-nc}{n})$$

Falls $j \coloneqq 10i - nc$ negativ ist, gilt $L(\gamma)^{wc} = \emptyset = L(\frac{0}{n})$, für $j > n$ folgt $L(\gamma)^{wc} = \Sigma^* = L(\frac{n}{n})$, und sonst erhalten wir $L(\gamma)^{wc} = L(\frac{j}{n})$ mit $j \in \{0,...,n\}$. In allen drei Fällen ist die Aussage somit gezeigt.

„$\Longrightarrow$“: Wir zeigen zunächst, dass die Sprache $L(x)$ die Zahl x eindeutig festlegt (solange $0 \leq x < 1$). Seien $x, y \in \mathbb{R}$ mit $0 \leq x < y < 1$. Wir zeigen nun $L(x) \neq L(y)$. Dazu wählen wir ein hinreichend großes n, sodass $10^n(y - x) \geq 1$ und ein $k \in \mathbb{N}$ mit $x < k \cdot 10^{-n} \leq y$. Dann sei $w \in \Sigma^n$ mit $(w)_{10} = k$; es folgt $w \notin L(x)$ und $w \in L(y)$.

Wir wollen nun eine unendliche Folge $\gamma_0, \gamma_1, ... \in \mathbb{R}$ konstruieren, sodass $0 \leq \gamma_i < 1$, $L(\gamma_i)$ eine Residualsprache von $L(\gamma)$ ist, und es ein $k \in \mathbb{Z}$ mit $\gamma_i = 10^i\gamma + k$ gibt, für alle $i \in \mathbb{N}$. Die Folge konstruieren wir induktiv, wir starten mit $\gamma_0 \coloneqq \gamma$.

Sei nun i beliebig. Da $0 \leq \gamma_i < 1$, gibt es genau ein $c \in \Sigma$ mit $0 \leq 10\gamma_i - c < 1$. Mit $\gamma_{i+1} \coloneqq 10\gamma_i - c$ gilt $L(\gamma_i)^c = L(\gamma_{i+1})$, also ist $L(\gamma_{i+1})$ eine Residualsprache von $L(\gamma)$. Schließlich gibt es ein k mit $\gamma_i = 10^i\gamma + k$, und somit $\gamma_{i+1} = 10\gamma_i - c = 10^{i+1}\gamma + 10k - c$.

Da $L(\gamma)$ regulär ist, gibt es nur endlich viele Residualsprachen. Also finden wir $i < j$ mit $L(\gamma_i) = L(\gamma_j)$. Wir wir oben argumentiert haben, folgt daraus $\gamma_i = \gamma_j$. Es gibt außerdem $k, l \in \mathbb{Z}$ mit $\gamma_i = 10^i\gamma + k$ und $\gamma_j = 10^j\gamma + l$. Insgesamt erhalten wir

$$\gamma = 10^{-i}(\gamma_i - k) = 10^{-i}(\gamma_j - k) = 10^{-i}(10^j\gamma + l - k) = 10^{j-i}\gamma + 10^{-i}(l - k)$$

Dies können wir nun nach γ auflösen (da $i \neq j$).

$$\gamma = \frac{10^{-i}(l - k)}{(1 - 10^{j-i})} = \frac{l - k}{10^i(1 - 10^{j-i})} \in \mathbb{Q}$$