Prof. Javier Esparza
Philipp Czerner, Kevin Kappelmann, Valentin Krasotin

Technische Universität München
Lehrstuhl für Theoretische Informatik

Einführung in die Theoretische Informatik

Sommersemester 2024 – Übungsblatt 5

- Das Übungsblatt ist in zwei Teile gegliedert: den Vorbereitungsteil, den Sie vor der Übung selbstständig bearbeiten sollen, und den Übungs-/Nachbereitungsteil, der Aufgaben enthält, die in der Übung besprochen werden und von Ihnen anschließend zur Nachbereitung verwendet werden können.

Vorbereitung (→ vor der Übung selbständig zu bearbeiten)

Vorbereitungsaufgabe Ü5.1. (*Wichtige Begriffe*)

Überprüfen Sie, dass Sie die folgenden Begriffe oder Notationen korrekt definieren können.

- Kontextfreie Sprache (*CFL*)
- Kontextfreie Grammatik (*CFG*)
- Syntaxbaum
- Ableitung, Linksableitung, Rechtsableitung
- (inhärent) mehrdeutig

Vorbereitungsaufgabe Ü5.2. (*Automata Tutor: "Contextfree Languages"*)

Lösen Sie die Aufgaben Ü5.2 (a–c) auf Automata Tutor.

Vorbereitungsaufgabe Ü5.3. (*Ableitung und Syntaxbaum*)

Sei $G = (\{S, E, O, A, B, X\}, \{a, b\}, P, S)$ die CFG mit folgenden Produktionen P:

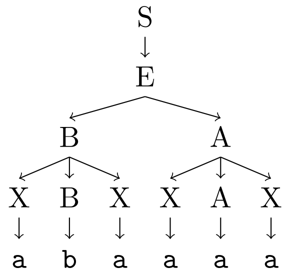
$$\begin{aligned}
S &\to E \mid O \\
E &\to AB \mid BA \\
A &\to XAX \mid a \\
B &\to XBX \mid b \\
O &\to XXO \mid X \\
X &\to a \mid b
\end{aligned}$$

(a) Geben Sie für jedes der folgenden Wörter jeweils eine Linksableitung und eine Rechtsableitung und den entsprechenden Syntaxbaum an:
(i) $abaaaa$ (ii) $babab$ (iii) $aabbaaba$

(b) Geben Sie ein Wort $w \in L(G)$ mit zwei verschiedenen Syntaxbäumen an.

Lösungsskizze.

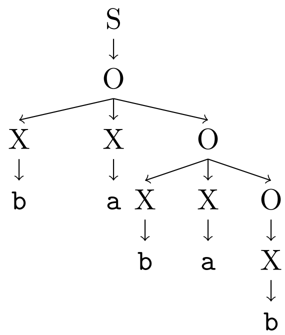
(a) (i)
Linksableitung: $S \to E \to BA \to XBXA \to aBXA \to abXA \to abaA \to abaXAX \to abaaAX \to abaaaX \to abaaaa$
Rechtsableitung: $S \to E \to BA \to BXAX \to BXAa \to BXaa \to Baaa \to XBXaaa \to XBaaaa \to Xbaaaa \to abaaaa$



(ii)

Linksableitung: $S \to O \to XXO \to bXO \to baO \to baXXO \to babXO \to babaO \to babaX \to babab$

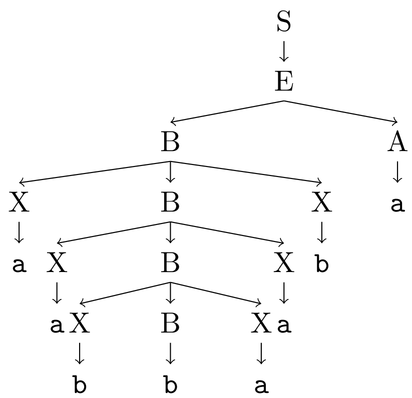
Rechtsableitung: $S \to O \to XXO \to XXXXO \to XXXXX \to XXXXb \to XXXab \to XXbab \to Xabab \to babab$



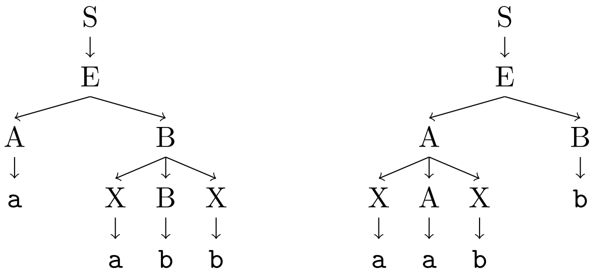
(iii)

Linksableitung: $S \to E \to BA \to XBXA \to aBXA \to aXBXXA \to aaBXXA \to aaXBXXXA \to aabBXXXA \to aabbXXXA \to aabbaXXA \to aabbaaXA \to aabbaabA \to aabbaaba$

Rechtsableitung: $S \to E \to BA \to Ba \to XBXa \to XBba \to XXBXba \to XXBaba \to XXXBXaba \to XXXBaaba \to XXXbaaba \to XXbbaaba \to Xabbaaba \to aabbaaba$



(b)



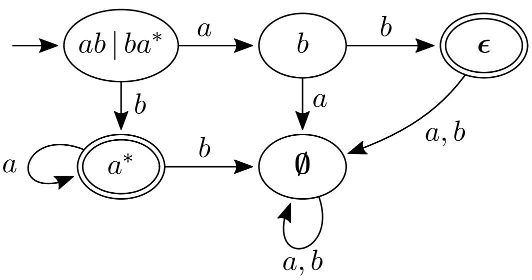
Übung und Nachbereitung

Übungsaufgabe Ü5.4. (*Kanonischer Minimalautomat*)

Konstruieren Sie den kanonischen Minimalautomaten zu dem regulären Ausdruck $r := ab \mid ba^*$ und benennen Sie jeden Zustand mit einem regulären Ausdruck für die entsprechende Residualsprache.

Lösungsskizze.

Wir erhalten folgenden Automaten:



Übungsaufgabe Ü5.5. (*DER Satz*)

Entscheiden Sie, ob folgende Sprachen über dem Alphabet $\Sigma = \{a, b, c\}$ regulär sind. Bestimmen Sie hierzu die Residualsprachen. Falls die Sprache regulär ist, zeichnen Sie den kanonischen Minimalautomaten und beschriften Sie die Zustände mit regulären Ausdrücken für die entsprechenden Residualsprachen. Falls die Sprache nicht regulär ist, reicht es eine unendliche Menge von Wörtern mit unterschiedlichen Residualsprachen zu bestimmen und zu zeigen, dass diese Residualsprachen paarweise verschieden sind.

(a) $L_1 = \{a^{2i} \mid i \in \mathbb{N}\}$

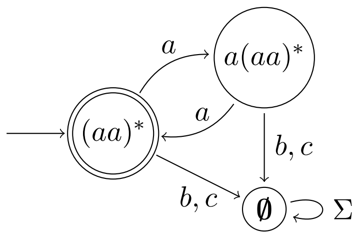
(b) $L_2 = \{a^i b^i c^i \mid i \in \mathbb{N}\}$

(c) $L_3 = \{w \in \Sigma^* \mid |w|_a = 2 \cdot |w|_b\}$

(d) $L_4 = L((a^*(b \mid c))^*)$

Lösungsskizze. Hinweis: Da die Sprachen aus dem Kontext ersichtlich sind, lassen wir den L_i Subscript bei $\equiv$ und $[w]$ weg.

(a) Kanonischer DFA:



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

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

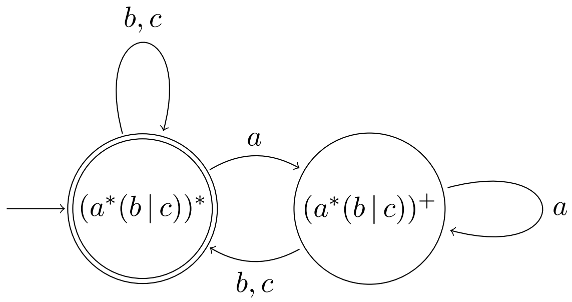
Beweis: Sei $i, j \in \mathbb{N}$ verschieden. Dann $a^i b^i c^i \in L_2$, aber $a^j b^j c^i \notin L_2$. Daher $L_2^{a^i b^i} \neq L_2^{a^j b^j}$. Somit sind alle Residualsprachen unterschiedlich und L_2 keine reguläre Sprache.

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

$$\{b^i \mid i \in \mathbb{N}\}$$

Beweis: Sei $i, j \in \mathbb{N}$ verschieden. Dann $b^i a^{2i} \in L_3$, aber $b^j a^{2i} \notin L_3$. Daher $L_3^{b^i} \neq L_3^{b^j}$. Somit sind alle Residualsprachen unterschiedlich und L_3 keine reguläre Sprache.

(d) Kanonischer DFA:



Übungsaufgabe Ü5.6. (*Sprache einer kontextfreien Grammatik*)

Sei $L = \{w \in \Sigma^* \mid w = w^R\}$ die Sprache der Palindrome über $\Sigma = \{a, b\}$.

(a) Geben Sie eine Grammatik für L an.

(b) Geben Sie eine Grammatik G für $\overline{L}$ an.

(c) Zeigen oder widerlegen Sie: $\overline{L}$ ist regulär.

(d) Zeigen Sie $L(G) = \overline{L}$ formal.

Lösungsskizze.

(a)

$$S \to aSa \mid bSb \mid a \mid b \mid \varepsilon$$

(b)

$$
\begin{aligned}
S &\to aSa \mid bSb \mid aTb \mid bTa \\
T &\to aT \mid bT \mid \varepsilon
\end{aligned}
$$

(c) $\overline{L}$ ist nicht regulär. Angenommen $\overline{L}$ wäre regulär. Da reguläre Sprachen unter Komplement abgeschlossen sind, ist somit auch L regulär. In Ü3.6 (a) haben wir jedoch gezeigt, dass L nicht regulär ist. Widerspruch.

(d) Zunächst zeigen wir $L_G(T) = \Sigma^*$.

(1) Die Richtung $L_G(T) \subseteq \Sigma^*$ ist klar.

(2) Wir zeigen $\Sigma^* \subseteq L_G(T)$ per Induktion über die Wortlänge n von $w \in \Sigma^*$. Im Fall $n = 0$ gilt $w = \epsilon$ und $T \to \varepsilon$.

Im Fall $w \in \Sigma^{n+1}$ erhalten wir als Induktionshypothese, dass $\Sigma^n \subseteq L_G(T)$. Sei nun $x \in \Sigma$ und $v \in \Sigma^n$ mit $w = xv$. Per I.H. gibt es eine Ableitung $T \to^* v$. Falls $x = a$, folgt $T \to aT \to^* av$. Falls $x = b$, folgt $T \to bT \to^* bv$.

Nun zeigen wir $L(G) = \overline{L}$.

(1) Wir zeigen $L(G) \subseteq \overline{L}$ per Induktion über die Erzeugung von Wörtern in $L_G(S)$. Im Fall $S \to aTb$ erhalten wir $S \to aTb \to^* awb$, und es gilt $(awb)^R = bw^Ra \neq awb$. Somit $awb \in \overline{L}$. Der Fall $S \to bwa$ folgt analog.

Im Fall $S \to aSa$ erhalten wir per I.H. $w \in \Sigma^*$ mit $w \neq w^R$ und $S \to aSa \to^* awa$. Es gilt $(awa)^R = aw^Ra \neq awa$, also $awa \in \overline{L}$. Der Fall $S \to bwb$ folgt analog.

(2) Wir zeigen $\overline{L} \subseteq L(G)$ mittels starker Induktion über die Wortlänge n von $w \in \overline{L}$. Da $w \in \overline{L}$, gilt $w \neq w^R$.

In den Fällen $n \in \{0, 1\}$ folgt $w = w^R$. Mit dem Fakt $w \neq w^R$ folgt ein Widerspruch und somit $w \in L(G)$ (ex falso quodlibet).

Im Fall $n \geq 2$ erhalten wir als Induktionshypothese, dass $\{w \in \overline{L} \mid |w| < n\} \subseteq L(G)$. Da $|w| \geq 2$, existiert $x, y \in \Sigma$ und $v \in \Sigma^{n-2}$ mit $w = xvy$. Wir betrachten zwei Fälle:

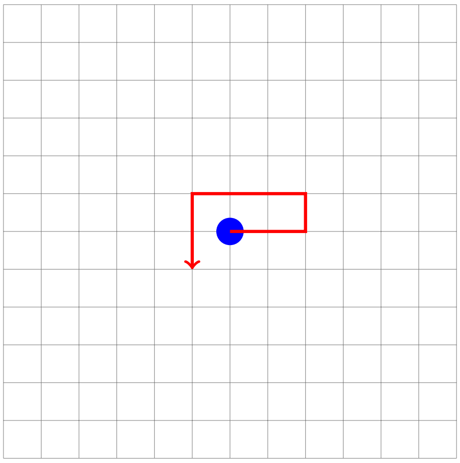
i. Falls $x = y$ folgt wegen $xvx = xvy = w \neq w^R = yv^Rx = xv^Rx$, dass $v \neq v^R$. Somit gilt $v \in \overline{L}$ und $|v| < n$. Per I.H. erhalten wir also eine Ableitung $S \to^* v$. Falls nun $x = a$, folgt $S \to aSa \to^* ava$ und analoges gilt für $x = b$.

ii. Falls $x \neq y$ gibt es eine Ableitung $S \to xTy \to^* xvy$, da wir bereits gezeigt haben, dass $L_G(T) = \Sigma^*$.

Zusätzliche Übungsaufgabe Ü5.7. (*Pfeilsprachen*)

Diese Aufgabe können Sie zusätzlich als Nachbereitung lösen. Es wird kein neuer Inhalt in dieser Aufgabe behandelt.

In dieser Aufgabe betrachten wir Sprachen, deren Worte Linienzüge in einem unendlichen zweidimensionalen Gitter von einem fixen Startpunkt aus beschreiben. Die folgende Grafik zeigt einen Ausschnitt aus dem Gitter:



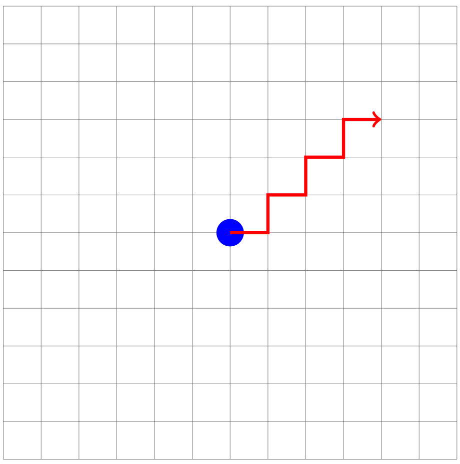
Der Startpunkt ist blau markiert. Linienzüge beschreiben wir im Folgenden als eine Sequenz von Pfeilen, d.h. als Worte über dem Alphabet $\Sigma = \{\rightarrow, \leftarrow, \uparrow, \downarrow\}$. Die Pfeile beschreiben dabei (vom Startpunkt aus gesehen) einen ein Kästchen langen Schritt entlang des Gitters. Wir stellen daher den im Bild rot eingezeichnete Linienzug durch das Wort $w = \rightarrow\rightarrow\uparrow\leftarrow\leftarrow\leftarrow\downarrow\downarrow$ dar.

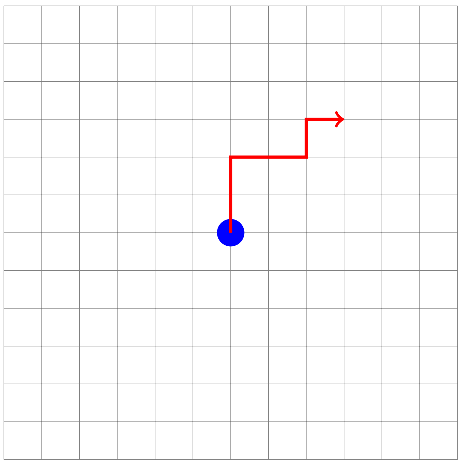
(a) Betrachten Sie die folgenden Beschreibungen in natürlicher Sprache zusammen mit jeweils einem Beispiel, welches in der Sprache liegt (auf der linken Seite), und einem Beispiel, das kein Element der Sprache ist (auf der rechten Seite).

Geben Sie für jede der Sprachen eine intensionale Mengendarstellung der Form $\{w \in \Sigma^* \mid \ldots\}$ an.

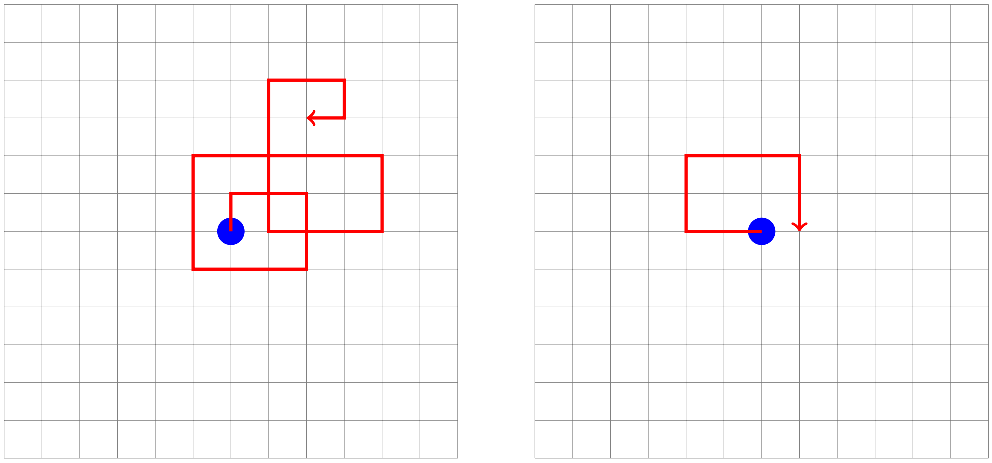
Hinweis: Die Sprachen sind mithilfe der Beispiele nicht eindeutig bestimmt. Ziel der Aufgabe ist es, für die intuitive Beschreibungen (z.B. “Sprache aller Skylines”) zusammen mit den Beispielen eine sinnvolle, formale Definition zu finden.

(1) Die Sprache aller Treppen über dem Alphabet $\Sigma' := \{\rightarrow, \uparrow\}$.

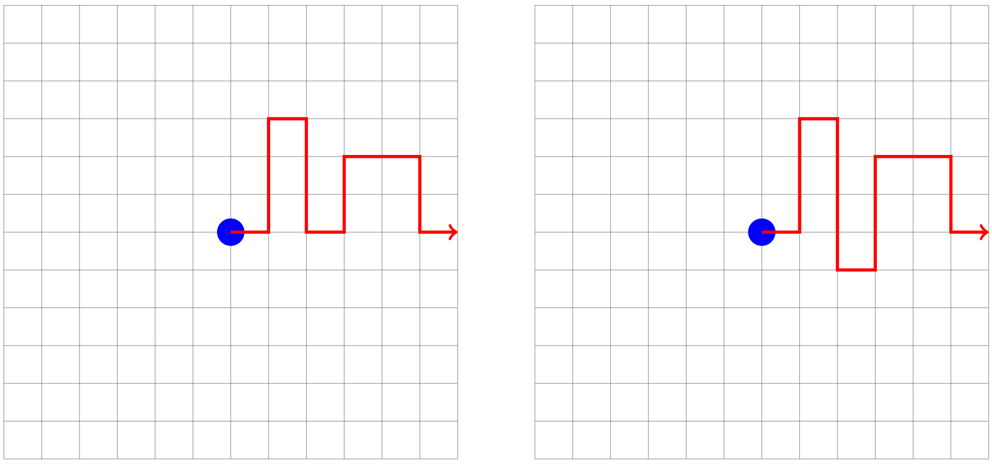




(2) Die Sprache aller im Uhrzeigersinn laufenden Spiralen über dem Alphabet Σ, die vom Startpunkt aus zuerst nach oben laufen.



(3) Die Sprache aller “Skylines” über dem Alphabet $\Sigma'' \coloneqq \{\rightarrow, \uparrow, \downarrow\}$.



(b) Stellen Sie Vermutungen auf, ob die obigen Sprachen regulär oder kontextfrei sind. Begründen Sie Ihre Antwort möglichst anschaulich anhand des Beispiels.

(c) Geben Sie zu jeder der Sprachen L aus Aufgabenteil (a) eine Grammatik G an.

Lösungsskizze.

(a) (1) $L = \{w \in \Sigma^* \mid \exists u, v.\ w = uv \wedge u \in L((\rightarrow\uparrow)^*) \wedge v \in \{\varepsilon, \rightarrow\}\}$

(2) $L = \{u \in \Sigma^* \mid \exists v \in \Sigma^*.\ uv \in L((\uparrow^+\rightarrow^+\downarrow^+\leftarrow^+)^*)\}$

(3) $L = \{w \in \Sigma^* \mid w \in \{\uparrow^n\rightarrow^m\downarrow^n \mid n \in \mathbb{N}, m \in \mathbb{N}_+\}^*\}$ (wir erlauben, dass zwei “Häuser” an derselben x-Koordinate beendet, d.h. Pfeil runter, und begonnen, d.h. Pfeil rauf, werden).

(b) (1) regulär, da nur eine Alternierung konstanter Pfade verlangt wird.

(2) regulär, da nur die Laufrichtung wichtig ist, aber nicht die Länge.

(3) nicht regulär, da zwei Längen verglichen werden müssen; aber kontextfrei.

(c) (1) $S \mapsto \rightarrow\uparrow S \mid \rightarrow \mid \varepsilon$

(2) $S \mapsto \uparrow S \mid \uparrow T \mid \varepsilon \qquad T \mapsto \rightarrow T \mid \rightarrow U \mid \varepsilon \qquad U \mapsto \downarrow U \mid \downarrow V \mid \varepsilon \qquad V \mapsto \leftarrow V \mid \leftarrow S \mid \varepsilon$

(3) $S \mapsto \varepsilon \mid TS \quad T \mapsto \uparrow T \downarrow \mid \to R \qquad R \mapsto \to R \mid \varepsilon$