Prof. T. Nipkow
M. Eberl, J. Rädle, S. Wimmer

Einführung in die theoretische Informatik

Sommersemester 2019 – Übungsblatt 5

AUFGABE 5.1. (*Wichtige Begriffe*) Stufe A

Überprüfen Sie, dass Sie die Folgenden Begriffe korrekt definieren können.

- MyHill-Nerode-Relation
- Äquivalenzklasse
- Kanonischer Minimalautomat

AUFGABE 5.2. (*Myhill-Nerode-Relation und Äquivalenzklassen*) Stufe B

Sei $\mathsf{L} = \mathsf{L}(\mathsf{a}^*\mathsf{b}^*\mathsf{c}^*)$ über dem Alphabet $\Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$.

(a) Entscheiden Sie, welche der folgenden Äquivalenzen wahr sind und begründen Sie Ihre Antwort:

- $\varepsilon \overset{?}{\equiv}_{\mathsf{L}} \mathsf{a}$
- $\mathsf{b} \overset{?}{\equiv}_{\mathsf{L}} \mathsf{c}$
- $\mathsf{abc} \overset{?}{\equiv}_{\mathsf{L}} \mathsf{cba}$

(b) Sei $v = \mathsf{aababc}$. Geben Sie ein Wort $u \neq v$ an, so dass $u \equiv_{\mathsf{L}} v$.

(c) Geben Sie die Mengen $[\mathsf{ab}]_{\mathsf{L}}$, $[\mathsf{bc}]_{\mathsf{L}}$ und $[\mathsf{ca}]_{\mathsf{L}}$ an.

(d) Finden Sie nun L', so dass $\mathsf{c} \equiv_{\mathsf{L}'} \mathsf{ba}$, $\mathsf{c} \not\equiv_{\mathsf{L}'} \mathsf{ab}$ und $\mathsf{aba} \equiv_{\mathsf{L}'} \mathsf{bab}$. Weiterhin soll $\varepsilon, \mathsf{aba} \in \mathsf{L}'$ gelten.

AUFGABE 5.3. (*Anwendungsbeispiel kontextfreie Grammatiken*) Stufe B

Betrachten Sie die folgende Grammatik $\mathsf{G} = (\{J, D, T, N, N', Z, A, S, E, U, B, C, V, U', B'\}$, $\{;, \{, \}, (,), =, a, b, \ldots, y, z, 0, 1, \ldots, 8, 9, +, -, \cdot, /, \%, !, <, >, \&\&, ||\}, \mathsf{P}, J)$ mit den Produktionen $\mathsf{P} :=$

$J \to DS \mid S$
$D \to TN; D \mid TN;$
$T \to int$
$N \to AN'$
$N' \to AN' \mid ZN' \mid A \mid Z$
$A \to a \mid b \mid c \mid \cdots \mid y \mid z$
$Z \to 0 \mid 1 \mid \cdots \mid 8 \mid 9$
$S \to SS \mid ; \mid \{S\} \mid N = E; \mid N = read(); \mid write(E); \mid if(C)\ S\ else\ S \mid while(C)\ S$
$E \to Z \mid N \mid (E) \mid UE \mid EBE$
$U \to -$
$B \to - \mid + \mid \cdot \mid / \mid \%$
$C \to true \mid false \mid (C) \mid EVE \mid U'(C) \mid CB'C$
$V \to == \mid != \mid <= \mid < \mid >= \mid >$
$U' \to !$
$B' \to \&\& \mid ||$

(a) Was für eine Sprache erzeugt diese Grammatik?

(b) Beurteilen Sie die folgende Aussagen: Alle Worte in $\mathsf{L}(\mathsf{G})$ können zu einem funktionierenden Programm compiliert werden.

(c) Geben Sie ein gültiges Wort in der Sprache an, das alle Nichtterminale mindestens einmal verwendet.

(d) Zeichnen Sie den Syntaxbaum für das in Teilaufgabe (c) gefundene Wort.

AUFGABE 5.4. (*Sprache einer kontextfreien Grammatik*) Stufe D

Mit $\mathsf{G} = (\{\mathsf{S}, \mathsf{T}\}, \{\mathsf{a}, \mathsf{b}\}, \mathsf{P}, \mathsf{S})$ sei die CFG mit folgenden Produktionen P bezeichnet:

$$\begin{aligned} \mathsf{S} &\to \mathsf{TaT} \\ \mathsf{T} &\to \mathsf{aTb} \mid \mathsf{bTa} \mid \mathsf{TT} \mid \mathsf{aT} \mid \varepsilon \end{aligned}$$

Die Sprache $\mathsf{L}_{\mathsf{T}}(\mathsf{G})$ die ausgehend von T als Startsymbol erzeugt wird ist $\mathsf{L} = \{w \in \{\mathsf{a}, \mathsf{b}\}^* \mid |w|_{\mathsf{a}} \geq |w|_{\mathsf{b}}\}$.

(a) Wir beweisen zunächst ein Lemma das in der folgenden Teilaufgabe hilfreich sein wird. Zeigen Sie dass für jedes $w \in \mathsf{L}$ das mit b beginnt eine Zerlegung $\mathsf{b}x\mathsf{a}y$ existiert, so dass $x, y \in \mathsf{L}$. Verwenden Sie hierzu eine Höhenfunktion wie in der Vorlesung im Beweis zu Satz 4.16.
Außerdem dürfen Sie im Folgenden ohne Beweis verwenden, dass für alle $w \in \mathsf{L}$ die mit a beginnen und für die $|w|_{\mathsf{a}} = |w|_{\mathsf{b}}$ gilt, eine Zerlegung $\mathsf{a}x\mathsf{b}y$ existiert, so dass $x, y \in \mathsf{L}$.

(b) Zeigen sie nun $\mathsf{L_T(G)} = \mathsf{L}$ unter Verwendung der Lemmas aus der vorigen Aufgabe.
(c) Welche Sprache erzeugt G ausgehend von S?

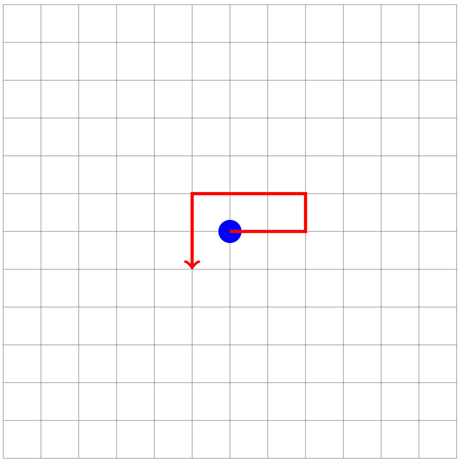
AUFGABE 5.5. (*MyHill-Nerode-Relation*) Stufe D

Entscheiden Sie, ob folgende Sprachen über dem Alphabet $\Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ regulär sind. Bestimmen Sie hierzu die Äquivalenzklassen der dazugehörigen Myhill-Nerode-Relation. Falls die Sprache regulär ist, zählen Sie alle Äquivalenzklassen auf und zeichnen Sie den kanonischen Minimalautomat $M_\mathsf{L} = (\Sigma^*/\equiv_\mathsf{L}, \Sigma, \delta_\mathsf{L}, [\varepsilon]_{\equiv_\mathsf{L}}, F_\mathsf{L})$. Falls die Sprache nicht regulär ist, reicht es eine unendliche Menge von Äquivalenzklassen zu bestimmen.

(a) $\mathsf{L}_1 = \{\mathsf{a}^{2i} \mid i \in \mathbb{N}_0\}$
(b) $\mathsf{L}_2 = \{\mathsf{a}^i\mathsf{b}^i\mathsf{c}^i \mid i \in \mathbb{N}_0\}$
(c) $\mathsf{L}_3 = \{w \in \Sigma^* \mid |w|_\mathsf{a} = 2 \cdot |w|_\mathsf{b}\}$
(d) $\mathsf{L}_4 = \mathsf{L}((\mathsf{a}^*(\mathsf{b} \mid \mathsf{c}))^*)$
(e) $\mathsf{L}_5 = \{w\mathsf{c}w \mid w \in \{\mathsf{a}, \mathsf{b}\}^*\}$

AUFGABE 5.6. (*Pfeilsprachen*) Stufe B – D

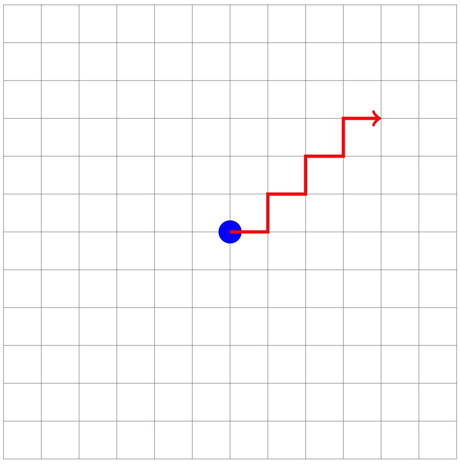
In dieser Aufgabe betrachten wir Sprachen, deren Worte Linienzüge in einem unendlichen zweidimensionalen Gitter von einem fixen Startpunkt aus beschreiben. Die folgende Grafik zeigt einen Ausschnitt aus dem Gitter:

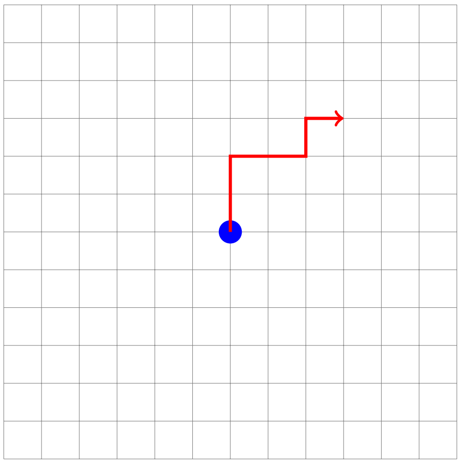


Wir haben Startpunkt blau markiert. Linienzüge beschreiben wir im Folgenden als eine Sequenz von Pfeilen, d.h. als Worte über dem Alphabet $\Sigma = \{\rightarrow, \leftarrow, \uparrow, \downarrow\}$. Die Pfeile beschreiben dabei (vom Startpunkt aus gesehen) einen ein Kästchen langen Schritt entlang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

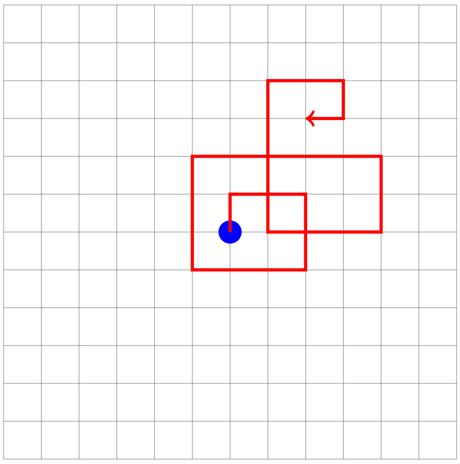
(a) Betrachten Sie die folgenden natürlich sprachlichen Beschreibungen zusammen mit jeweils einem Beispiel, welches in der Sprache liegt (auf der linken Seite), und einem Beispiel, das kein Element der Sprache ist (auf der rechten Seite). *Geben Sie für jede der Sprachen eine formale Definition der Form* $\{w \in \Sigma^* \mid \dots\}$ *an.*[1]

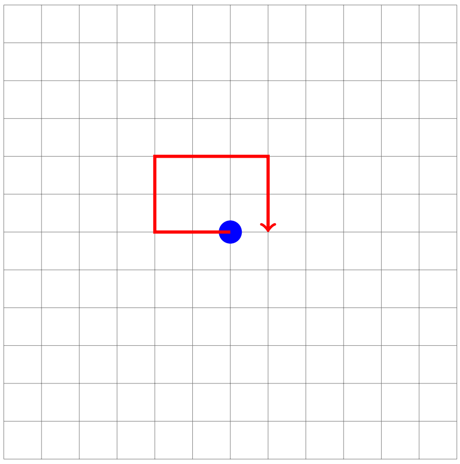
(i) die Sprache aller Treppen über dem Alphabet $\Sigma' = \{\rightarrow, \uparrow\}$



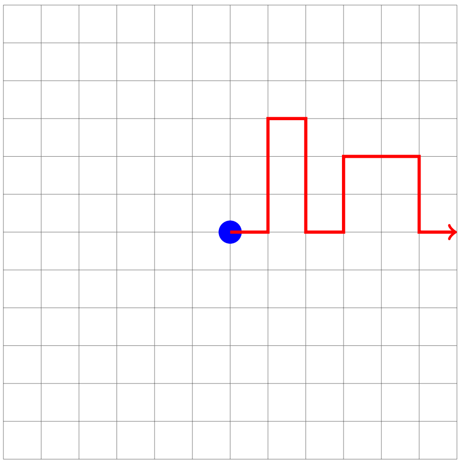


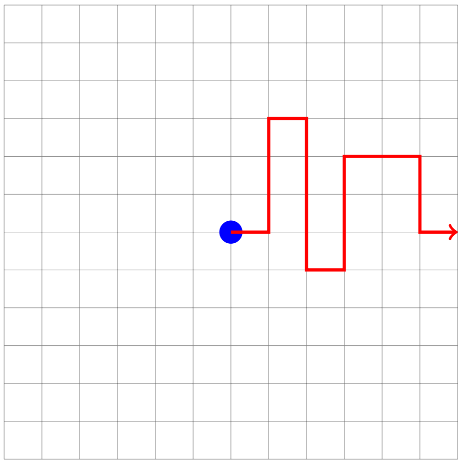
(ii) die Sprache aller im Uhrzeigersinn laufenden Spiralen über dem Alphabet Σ, die vom Startpunkt aus zuerst nach oben laufen





(iii) die Sprache aller "Skylines" über dem Alphabet $\Sigma'' = \{\rightarrow, \uparrow, \downarrow\}$





[1] Das heißt insbesondere, dass Sie in diesem Aufgabenteil keinen Automaten, keinen regulären Ausdruck, keine Grammatik oder ähnliches angeben sollen, die die Sprache beschreiben.

(iv) die Sprache aller Quadrate über dem Alphabet Σ

Hinweis: Die Sprachen sind mithilfe der Beispiele nicht eindeutig bestimmt! Ziel der Aufgabe ist es, die intuitive Beschreibung (z.B. "Sprache aller Quadrate") zusammen mit den Beispielen in eine möglichst allgemeine Sprachdefinition zu bringen.

(b) *Stellen Sie Vermutungen auf, ob die obigen Sprachen regulär, kontextfrei bzw. kontextsensitiv sind.* Begründen Sie Ihre Antwort möglichst anschaulich anhand des Beispiels. Überlegen Sie dazu, was es für Linienzüge heißt, regulär, kontextfrei bzw. kontextsensitiv zu sein.

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

(d) Verallgemeinern Sie Aufgabenteil (a), iv) zu der Sprache aller Rechtecke, indem Sie zunächst Beispiele von Wörtern angeben, die in der Sprache liegen bzw. nicht in der Sprache liegen, und dann die Aufgabenteile (a) bis (c) auch für diese Sprache lösen.

Definition (Monoid)

Ein Monoid $\langle M, \circ, 1\rangle$ besteht aus einer (Träger-)Menge M, einer assoziativen Abbildung $\circ\colon M \times M \to M$ und einem bzgl. $\circ$ neutralen Element $1 \in M$ (d.h. $\forall m \in M.\ m \circ 1 = m = 1 \circ m$). Ist $\circ$ kommutativ, dann wird $\langle M, \circ, 1\rangle$ als kommutatives Monoid bezeichnet. Gilt $\forall m \in M.\ m \circ m = m$, so wird $\langle M, \circ, 1\rangle$ als **idempotentes** Monoid bezeichnet.

Definition (Kleene Algebra)

Eine *Kleene Algebra* $\langle K, +, \cdot, ^*, 0, 1\rangle$ besteht aus einer Trägermenge K, den binären Operationen $+\colon K \times K \to K$ (Addition), $\cdot\colon K \times K \to K$ (Multiplikation), der unären Operation $^*\colon K \to K$ (Stern) und zwei Konstanten $0, 1 \in K$. Mittels der Addition definiert man die binäre Relation $\sqsubseteq$ auf K durch

$$a \sqsubseteq b \overset{\text{Def}}{\longleftrightarrow} a + b = b$$

Wie üblich schreibt man kurz ab für $a \cdot b$. Um Klammern zu sparen, gilt "Stern vor Punkt vor Strich".
Eine Kleene Algebra erfüllt folgende Eigenschaften für alle $a, b, c, x \in K$:

Ax1: $\langle K, +, 0\rangle$ ist ein kommutatives und idempotentes Monoid.

Ax2: $\langle K, \cdot, 1\rangle$ ist ein Monoid.

Ax3: $a(b + c) = ab + ac$ und $(a + b)c = ac + bc$.

Ax4: $a0 = 0 = 0a$.

Ax5: $1 + aa^* \sqsubseteq a^*$ und $1 + a^*a \sqsubseteq a^*$.

Ax6: $b + ax \sqsubseteq x \to a^*b \sqsubseteq x$ und $b + xa \sqsubseteq x \to ba^* \sqsubseteq x$.

AUFGABE 5.7. (*Abstraktion von Problemen*) Stufe E

In dieser Aufgabe abstrahieren wir von der konkreten Interpretation von regulären Ausdrücken als Konstruktionsbeschreibungen regulärer Sprachen. Ziel ist es den Zusammenhang mit Pfadproblemen im Bereich der Informatik und dem Lösen linearer Gleichungssysteme zu verdeutlichen.

(a) Sei Σ ein Alphabet. Beweisen Sie, dass $\langle 2^{\Sigma^*}, \cup, \circ, {}^*, \emptyset, \{\varepsilon\}\rangle$ eine Kleene Algebra ist für

$$LL' := L \circ L' := \{ww' \mid w \in L, w' \in L'\} \qquad \text{und} \qquad L^* := \bigcup_{k \in \mathbb{N}_0} L^k$$

(b) Die Addition und das Minimum auf $\mathbb{R}$ seien auf $[-\infty, \infty] = \mathbb{R} \cup \{\pm\infty\}$ wie folgt erweitert:

$$\infty + a = \infty \quad -\infty + a = -\infty \quad -\infty + \infty = \infty$$

$$\min(a, \infty) = a \quad \min(a, -\infty) = -\infty \quad \min(-\infty, \infty) = -\infty$$

Weiterhin gelte:

$$a^* := \begin{cases} -\infty & \text{falls } a \in [-\infty, 0) \\ 0 & \text{falls } a \in [0, \infty] \end{cases}$$

Beweisen Sie, dass $\langle \mathbb{R} \cup \{\pm\infty\}, \min, +, {}^*, \infty, 0\rangle$ eine Kleene Algebra ist.

(c) Zeigen Sie, dass in jeder Kleene Algebra $\langle K, +, \cdot, {}^*, 0, 1\rangle$ für beliebige $a, b, c, d, e, f \in K$ gilt:

(i) $\sqsubseteq$ ist eine partielle Ordnung auf K, die monoton bzgl. Addition, Multiplikation und Stern ist, d.h.:

$$a \sqsubseteq b \to (ca \sqsubseteq cb \wedge ac \sqsubseteq bc \wedge a + c \sqsubseteq b + c \wedge a^* \sqsubseteq b^*)$$

(ii) a^*b ist die bzgl. $\sqsubseteq$ kleinste Lösung der linearen Ungleichung $b + aX \sqsubseteq X$ in K (X Variable), genauer:

$$b + a(a^*b) \sqsubseteq a^*b \qquad \wedge \qquad \forall x \in K\colon b + ax \sqsubseteq x \to a^*b \sqsubseteq x$$

Entsprechend ist ba^* die kleinste Lösung in K von $b + Xa \sqsubseteq X$.

Man kann zeigen, dass jedes lineare Ungleichungssystem

$$\begin{aligned} a_{1,1}X_1 + a_{1,2}X_2 + \ldots a_{1,n}X_n + b_1 \quad &\sqsubseteq \quad X_1 \\ &\vdots \\ a_{n,1}X_1 + a_{n,2}X_2 + \ldots a_{n,n}X_n + b_n \quad &\sqsubseteq \quad X_n \end{aligned}$$

mit $a_{i,j}, b_i \in K$ und $X_1, \ldots, X_n$ Variablen stets eine eindeutige $\sqsubseteq$-kleinste Lösung in K hat, d.h. dass es konkrete Elemente $x_1, \ldots, x_n \in K$ gibt, so dass für $X_1 = x_1, \ldots, X_n = x_n$ alle obigen Ungleichungen erfüllt sind, und für jede weitere Lösung $X_1 = y_1, \ldots, X_n = y_n \in K$ stets $x_i \sqsubseteq y_i$ gilt. Insbesondere gilt

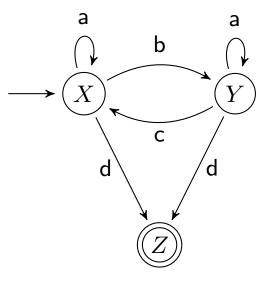
$$x_1 = a_{1,1}^*(a_{1,2}x_2 + \ldots + a_{1,n}x_n + b_1)$$

Somit kann das Gauß-Verfahren zur Bestimmung von $x_1, \ldots, x_n$ verwendet werden.

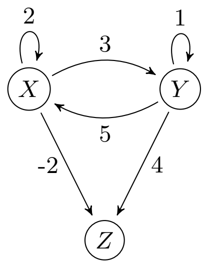
(d) Bestimmen Sie die kleinste Lösung x_1, x_2, x_3, x_4 für folgendes System:

$$\begin{aligned} aX + bY + eZ \quad &\sqsubseteq \quad X \\ cX + dY + fZ \quad &\sqsubseteq \quad Y \\ 1 \quad &\sqsubseteq \quad Z \end{aligned}$$

(e) Bestimmen Sie den regulären Ausdruck für den Zustand X. Durch welche Werte muss man a, b, c, d, e, f konkret ersetzen, damit sich die in (d) berechneten Terme in $\langle 2^{\Sigma^*}, \cup, \circ, {}^*, \emptyset, \{\varepsilon\}\rangle$ zu dem gewünschten Ausdruck auswertet?



(f) Bestimmen Sie die Länge eines kürzesten Pfades von jedem Knoten zum Knoten Z in folgendem gewichteten Graphen. Durch welche Werte muss man a, b, c, d, e, f konkret ersetzen, damit sich die in (d) berechneten Terme in $\langle \mathbb{R} \cup \{\pm\infty\}, \min, +, \infty, 0\rangle$ zu den gesuchten Pfadlängen auswerten?


2
3
1
X
Y
5
-2
4
Z