Prof. Javier Esparza
Philipp Czerner, Martin Helfrich

Technische Universität München
Lehrstuhl für Theoretische Informatik

Einführung in die Theoretische Informatik

Sommersemester 2022 – 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.
- Das ist nicht das Hausaufgabenblatt! Die Hausaufgaben finden Sie auf einem separaten Blatt.

Vorbereitung (→ vor der Übung selbständig zu bearbeiten)

Individualaufgabe Ü5.1. (*Wichtige Begriffe*)

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

- Kontextfreie Sprache (*CFL*)
- Kontextfreie Grammatik (*CFG*)
- rechts-lineare / links-lineare CFG
- (Links-/Rechts-)Ableitung
- Syntaxbaum
- mehrdeutige CFG

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

Lösen Sie die Aufgaben Ü5.2 (a–d) auf Automata Tutor.[1]

Individualaufgabe Ü5.3. (*Anwendungsbeispiel kontextfreie Grammatiken*)

Betrachten Sie die folgende Grammatik $G = (V, \Sigma, P, J)$ mit

$$V = \{J, D, T, N, N', Z, A, S, E, U, B, C, V, U', B'\}$$
$$\Sigma = \{\texttt{;}, \texttt{\{}, \texttt{\}}, \texttt{(}, \texttt{)}, \texttt{=}, \texttt{a}, \texttt{b}, \dots, \texttt{y}, \texttt{z}, \texttt{0}, \texttt{1}, \dots, \texttt{8}, \texttt{9}, \texttt{+}, \texttt{-}, \cdot, \texttt{/}, \texttt{\%}, \texttt{!}, \texttt{<}, \texttt{>}, \texttt{\&\&}, \texttt{||}\}$$

[1] Wenn Sie Automata Tutor noch nicht verwendet haben, folgen Sie erst den Schritten in Ü1.2, um sich richtig zu registrieren.

mit den Produktionen $P :=$

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

(a) Was für eine Sprache erzeugt diese Grammatik?

(b) Beurteilen Sie die folgende Aussagen: Alle Worte in $L(G)$ 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.

Übung und Nachbereitung

Übungsaufgabe Ü5.4. (*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 zzgl. des entsprechenden Syntaxbaums an:
(i) *abaaaa* (ii) *babab* (iii) *aabbaaba*

(b) Entscheiden Sie, ob die Grammatik G mehrdeutig ist oder nicht. Wenn sie mehrdeutig ist, geben Sie ein Wort $w \in L(G)$ mit zwei Syntaxbäumen an. Sonst beweisen Sie, dass G nicht mehrdeutig ist.

$S \to E \mid O$

$E \to AB \mid BA$

$A \to XAX \mid a$

$B \to XBX \mid b$

$O \to XXO \mid X$

$X \to a \mid b$

$L_G(A) = \{ vaw \mid |v| = |w| \}$

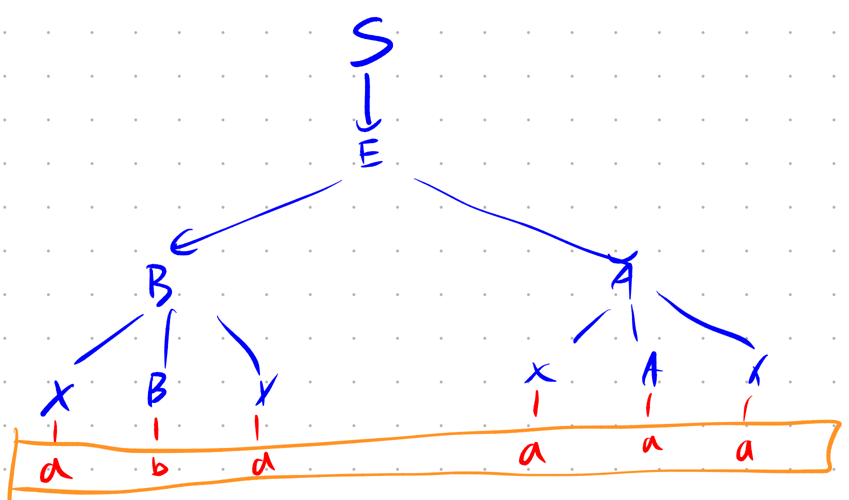
$L_G(B) = \{ vbw \mid |v| = |w| \}$

$L_G(O) = \{ c^{2n+1} \mid c \in \{a, b\} \wedge n \geq 0 \}$

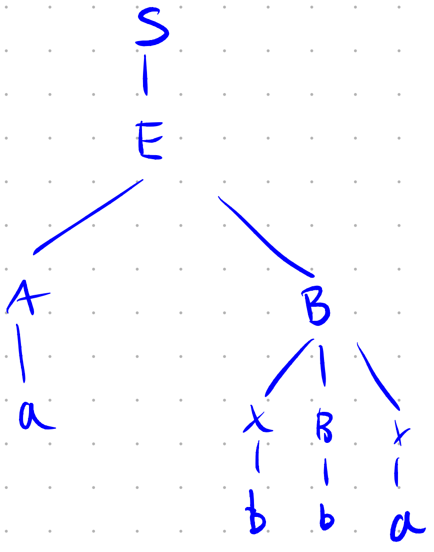
i) $\underbrace{aba}_{B} \mid \underbrace{aaa}_{A}$

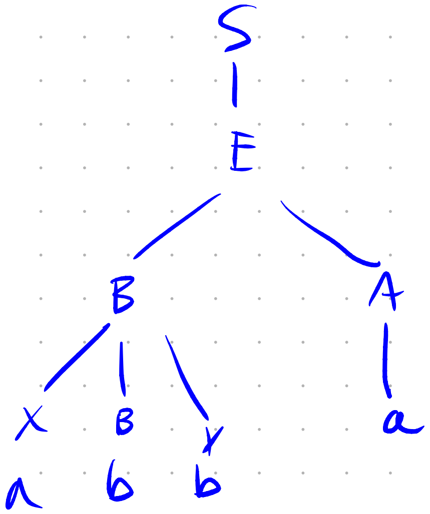
$S \to E \to BA \to XBXA$
$\to aBXA$
$\to abXA$
$\to abaA$
$\to \ldots$

$S \to E \to BA \to BXAX$
$\to BXAa$
$\to BXaa$
$\to Baaa$
$\to \ldots$



$w = abba$





G ist mehrdeutig

$L = \{ w \in A \mid P(a) \}$

Übungsaufgabe Ü5.5. (*Sprache einer kontextfreien Grammatik*)

Gegeben sei die folgende Grammatik G:

$$S \to aSa \mid bSb \mid aTb \mid bTa$$
$$T \to aT \mid bT \mid \epsilon$$

(a) Welche Sprache beschreibt G? Geben Sie eine intensionale Mengendarstellung[2] L für $L(G)$ an.

(b) Zeigen oder widerlegen Sie: L ist regulär.

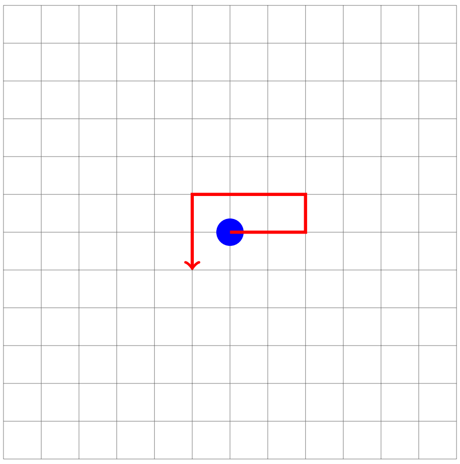
(c) Zeigen Sie $L(G) = L$ formal. [Beweisen Sie dabei auch induktiv, welche Sprache von T erzeugt wird.]

Übungsaufgabe Ü5.6. (*Residualsprachen*)

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.

Übungsaufgabe Ü5.7. (*Pfeilsprachen*)

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.

[2]Das heißt, eine Beschreibung der Form $L := \{w \in A \mid P(a)\}$ für eine geeignete Menge A und Prädikat P.

Übungsaufgabe Ü5.5. (*Sprache einer kontextfreien Grammatik*)

Gegeben sei die folgende Grammatik G:

$$\boxed{S \to aSa \mid bSb \mid} aTb \mid bTa$$
$$T \to aT \mid bT \mid \epsilon$$

a) 1. Bemerkung: $L_G(T) = \Sigma^*$

2. Bemerkung: alle $w \in L(G)$ sehen so aus:

$$v \underbrace{a \, \Sigma^* \, b} v^R$$

$$v \underbrace{b \, \Sigma^* \, a} v^R$$

D.h. $L = \{ w \in \Sigma^* \mid w \neq w^R \}$

b) $\overline{L} = \{ w \in \Sigma^* \mid w = w^R \}$
ist nicht regulär. Somit auch L nicht.

c) $\underline{L(G) = L \quad \rightsquigarrow \quad L \subseteq L(G) \;\wedge\; L(G) \subseteq L.}$

c) Wir zeigen zuerst: $L_G(T) = \Sigma^*$.

Es reicht $\Sigma^* \subseteq L_G(T)$ zu zeigen.

Sei $w \in \Sigma^*$. Wir zeigen $w \in L_G(T)$

Beweis: per Induktion auf $n = |w|$.

1. Basis: $n = 0 \Rightarrow w = \varepsilon \Rightarrow T \to \varepsilon = w$. Also $w \in L_G(T)$.

IH: Für alle $v \in \Sigma^n$ gilt $v \in L_G(T)$.

IS: Sei $w \in \Sigma^{n+1}$. Dann hat w die Gestalt cv, wobei $c \in \Sigma$, $v \in \Sigma^n$. Es gilt ja

$$T \to^* v.$$

Somit können wir auch w ableiten, durch

$$T \to cT \to^* cv.$$

(Da $c \in \{a, b\}$, gilt $T \to cT$)

Schlußendlich gilt $w \in L_G(T)$.

Nun zeigen wir $L_G(T) \subseteq \Sigma^*$.

Sei $w \in L_G(T)$. (z.z. $w \in \Sigma^*$)

$\boxed{\text{Fall 1}}$: $T \to aT \xrightarrow{*} av = w$ Die Induktionshypothese lautet: $v \in \Sigma^*$. Also folgt aus $a \in \Sigma$, dass

$$w = av \in \Sigma^*.$$

$\boxed{\text{Fall 2}}$: $T \to bT \xrightarrow{*} bv = w$

IH: $v \in \Sigma^*$.

$\Rightarrow$ $w = bv \in \Sigma^*$ da $b \in \Sigma$.

$\boxed{\text{Fall 3}}$: $T \to \varepsilon = w$. Trivialerweise gilt $w \in \Sigma^*$.

$L = L(G)$ $\qquad$ $L = \{w \in \Sigma^* \mid w \neq w^R\}$

Übungsaufgabe Ü5.5. (*Sprache einer kontextfreien Grammatik*)
Gegeben sei die folgende Grammatik G:

$$S \to aSa \mid bSb \mid aTb \mid bTa$$
$$T \to aT \mid bT \mid \epsilon$$

1 Wir zeigen $L(G) \subseteq L$.
Sei $w \in L(G)$. Dann wird w durch eine der Produktionen $S \to aSa$, $S \to bSb$, ... erzeugt. Zu zeigen: $w \in L$.

<u>Fall 1</u> $S \to aTb \to^* avb = w$ erzeugt. Wir wissen: $v \in \Sigma^*$. D.h. $w = avb \in \Sigma^*$. Es bleibt $w \neq w^R$ zu zeigen. Es gilt:

$$w = avb \neq bv^Ra = w^R$$

Somit $w \in L$.

<u>Fall 2</u> folgt analog.

<u>Fall 3</u> $S \to aSa \to^* ava = w$. Die IH lautet $v \in L$, d.h. $v \neq v^R$. Wir haben also:

$$w = ava \neq av^Ra = (ava)^R = w^R$$

Somit gilt $w \in L$. [$w \in \Sigma^*$ trivialerweise, da $v \in \Sigma^*$ und $a \in \Sigma$]

<u>Fall 4</u> folgt analog

"$L \subseteq L(G)$"

Sei $w \in L$. Es ist zu zeigen $w \in L(G)$.

Beweis per Induktion über $n = |w|$.

Für $n \in \{0, 1\}$ gibt es keine $w \in L$ mit $|w| = n$. Also ist $n = 2$ die IB. D.h. $w \in \{ab, ba\}$. In beiden Fällen gibt es eine Ableitung $S \rightarrow aTb \rightarrow ab$ bzw. $S \rightarrow bTa \rightarrow ba$. Also gilt $w \in L(G)$.

I.H: Angenommen, für alle $v \in L$ mit $|v| < n$ gilt $v \in L(G)$.

IS:

Fall 1: $w = ava$, $a \in \Sigma$, $v \in \Sigma^{n-2}$.

Da $w \in L$, gilt $w = ava \neq av^R a = w^R$. Also muss $v \neq v^R$ gelten, d.h. $v \in L$. Aber v hat Länge $n-2 < n$. Laut IH gilt dann

$$S \xrightarrow{*} v,$$

Schlußendlich:

$$S \rightarrow aSa \xrightarrow{*} ava = w.$$

Also $w \in L(G)$.

Fall 2: $w = bvb$, $b \in \Sigma$, $v \in \Sigma^{n-2}$ analog

Fall 3: $w = avb$, $b \in \Sigma$, $v \in \Sigma^{n-2}$, $\Rightarrow$

gilt offensichtlich $w \in L$, da $w = avb \neq bv^R a = w^R$.

Außerdem gilt $v \in \Sigma^* = L_G(T)$, d.h.

$$T \xrightarrow{*} v.$$

Folglich:

$$S \rightarrow aTb \xrightarrow{*} avb = w.$$

Also $w \in L(G)$.

$$S \xrightarrow{k+1} w$$

$$S \rightarrow aSa \xrightarrow{k} ava$$

$$S \xrightarrow{k} v$$

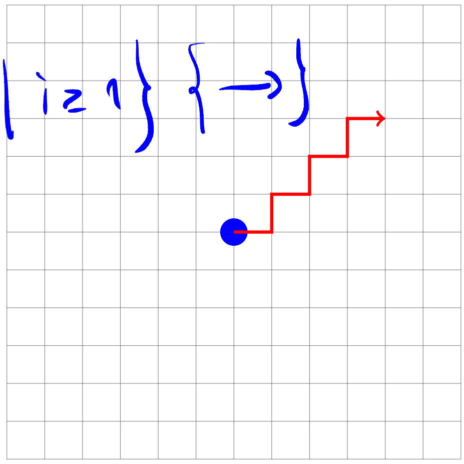
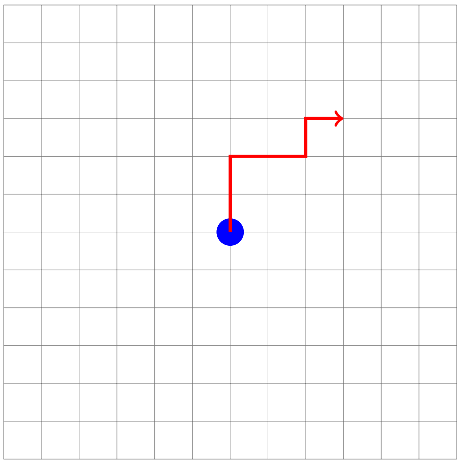
$v \in L$

somit $w \in L$

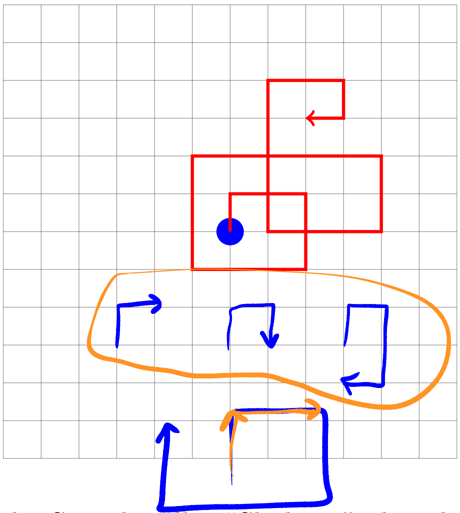
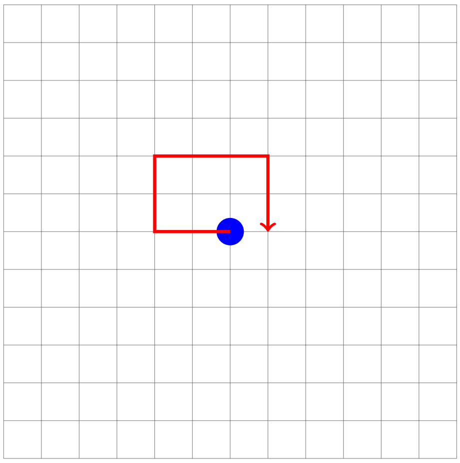
(a) Betrachten Sie die folgenden natürlich sprachlichen Beschreibungen zusammen mit jeweils einem Beispiel, welches in der Sprache liegt (auf der linken Seite), und einem Beispiel, das kein Element der Sprache ist (auf der rechten Seite). *Geben Sie für jede der Sprachen eine formale Definition der Form* $\{w \in \Sigma^* \mid ...\}$ *an.*[3]

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

$L = \{(\rightarrow\uparrow)^i \mid i \geq 1\}\{\rightarrow\}$

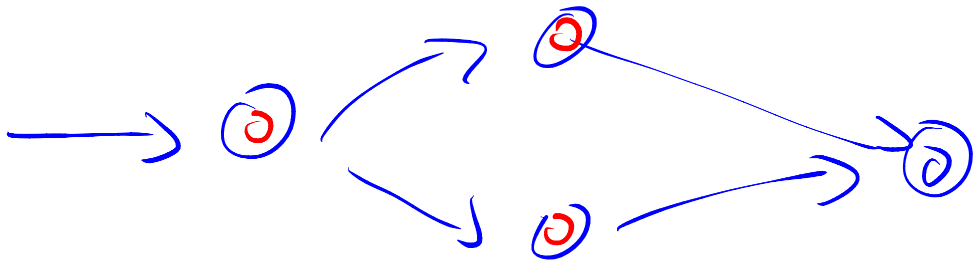
(2) die Sprache aller im Uhrzeigersinn laufenden Spiralen über dem Alphabet Σ, die vom Startpunkt aus zuerst nach oben laufen

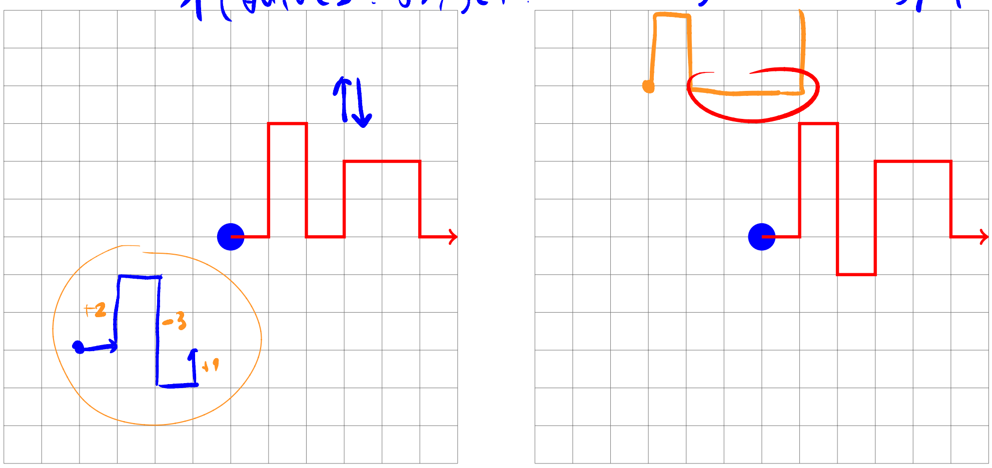
(3) die Sprache aller "Skylines" über dem Alphabet $\Sigma'' = \{\rightarrow, \uparrow, \downarrow\}$

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

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



$$L = \{ w \in \Sigma^* \mid (\forall u, v \in \Sigma^*.\ uv = w \rightarrow |u|_\uparrow \geq |u|_\downarrow) \land (\forall u, v \in \Sigma^*.\ \forall x, y \in \Sigma^*.\ w = uxyv \rightarrow xy \notin \{\uparrow\downarrow, \downarrow\uparrow\}) \}$$



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

(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.

((   ))

( )) (