Prof. J. Esparza
S. Sickert, J. Krämer

Einführung in die theoretische Informatik

Sommersemester 2017 – Übungsblatt Lösungsskizze 1

Übungsblatt

Wir unterscheiden zwischen Übungs- und Abgabeblättern. Auf diesem *Übungsblatt* finden Sie eine Übersicht über die Kernaspekte, die Sie in Kalenderwoche 18 in den Tutorien diskutieren, üben und vertiefen. Die Aufgaben auf diesem Blatt dienen dem Üben und Verstehen des Vorlesungsstoffes, sowie dem *eigenständigen Erarbeiten* der Kernaspekte. Außerdem sollen Ihnen diese Aufgaben auch helfen, ein Gefühl dafür zu bekommen, was Sie inhaltlich in der Klausur erwartet. Klausuraufgaben können jedoch deutlich von den hier gestellten Aufgaben abweichen. Abschreiben und Auswendiglernen von Lösungen wird Ihnen daher keinen dauerhaften Erfolg in der Vorlesung bringen. Fragen zu den Übungsblättern können Sie montags bis donnerstags von 12 Uhr bis 14 Uhr in der *THEO-Sprechstunde* in Raum 03.11.034 stellen.

Kernaspekte

K1.1 korrektes Wiedergeben der folgenden Definitionen

- Alphabet Σ
- Wort w
- Konkatenation von Wörtern
- Länge eines Wortes
- formale Sprache
- Operationen auf Sprachen (Def. 2.3)
- reflexive transitive Hülle
- reguläre Sprache
- kontextfreie Sprache
- kontextsensitive Sprache
- Sprachen vom Typ-0
- rechtslineare Grammatik
- kontextfreie Grammatik
- kontextsensitive Grammatik
- Grammatiken vom Typ-0

K1.2 Sprachen mithilfe von Operationen verknüpfen
K1.3 Wörter einer Sprache angeben
K1.4 begründen und beweisen, dass eine Grammatik ein Wort (nicht) erzeugt
K1.5 begründen und beweisen, dass eine Grammatik eine Sprache (nicht) erzeugt
K1.6 Aussagen über Verknüpfungen von Sprachen beweisen
K1.7 begründet entscheiden, ob gegebene Beispiele neu eingeführte Definitionen erfüllen
K1.8 Aussagen, mit neu eingeführten Definitionen beweisen oder widerlegen

AUFGABE 1.1.

Stufe B

Gegeben seien folgende drei Grammatiken $\mathsf{G}_1 = (\{\mathsf{S},\mathsf{X}\},\Sigma,\mathsf{P}_1,\mathsf{S})$, $\mathsf{G}_2 = (\{\mathsf{S},\mathsf{X}\},\Sigma,\mathsf{P}_2,\mathsf{S})$ und $\mathsf{G}_3 = (\{\mathsf{S},\mathsf{S}',\mathsf{A},\mathsf{B},\mathsf{C}\},\Sigma,\mathsf{P}_3,\mathsf{S})$ über dem Alphabet $\Sigma = \{\mathsf{a},\mathsf{b},\mathsf{c}\}$ mit folgende Produktionen:

$$\begin{array}{lllllll}
\mathsf{P}_1: & \mathsf{S} \to \mathsf{aX} & \mathsf{X} \to \mathsf{bS} \mid \mathsf{b} & & & & \\
\mathsf{P}_2: & \mathsf{S} \to \mathsf{aSb} \mid \mathsf{SX} \mid \varepsilon & \mathsf{X} \to \mathsf{cX} & & & & \\
\mathsf{P}_3: & \mathsf{S} \to \mathsf{S}' \mid \varepsilon & \mathsf{S}' \to \mathsf{ABC} \mid \mathsf{ABCS}' & & & & \\
 & \mathsf{A} \to \mathsf{a} & \mathsf{B} \to \mathsf{b} & \mathsf{C} \to \mathsf{c} & \mathsf{CB} \to \mathsf{BC} & \mathsf{CA} \to \mathsf{AC} & \mathsf{BA} \to \mathsf{AB}
\end{array}$$

Begründen Sie, warum die folgenden Wörter von den gegebenen Grammatiken erzeugt bzw. nicht erzeugt werden.

(a) $w_1 = \mathsf{abab} \in \mathsf{L}(\mathsf{G}_1)$, $w_2 = \mathsf{ba} \notin \mathsf{L}(\mathsf{G}_1)$, $w_3 = \varepsilon \notin \mathsf{L}(\mathsf{G}_1)$
(b) $u_1 = \mathsf{aabb} \in \mathsf{L}(\mathsf{G}_2)$, $u_2 = \mathsf{aabbcc} \notin \mathsf{L}(\mathsf{G}_2)$
(c) $v_1 = \mathsf{aabbcc} \in \mathsf{L}(\mathsf{G}_3)$, $v_2 = \mathsf{aaabbc} \notin \mathsf{L}(\mathsf{G}_3)$

Lösungsskizze

(a) $\mathsf{S} \to \mathsf{aX} \to \mathsf{abS} \to \mathsf{abaX} \to w_1$, $w_2, w_3 \notin \mathsf{L}(\mathsf{G}_1)$, da der erste Schritt jeder Produktion ein a an der ersten Stelle erzeugt.
(b) $\mathsf{S} \to \mathsf{aSb} \to \mathsf{aaSbb} \to u_1$, u_2 kann nicht erzeugt werden, da $\mathsf{X} \to \mathsf{Xc}$ nie "terminiert".
(c) $\mathsf{S} \to \mathsf{S}' \to^* \mathsf{ABCABC} \to \mathsf{ABACBC} \to \mathsf{AABCBC} \to \mathsf{AABBCC} \to^* \mathsf{aabbcc} = v_1$. v_2 kann nicht erzeugt werden, da jede Produktionsregel die Invariante $|w|_{\{a,A\}} = |w|_{\{b,B\}} = |w|_{\{c,C\}}$ erhält und v_2 diese nicht erfüllt.

AUFGABE 1.2.

Stufe B

Sei $\Sigma = \{\mathsf{a}, \mathsf{b}\}$ und $\mathsf{W} = \{\mathsf{aa}, \mathsf{aaa}, \mathsf{b}\}$. Geben Sie, wenn möglich, jeweils mindestens drei Wörter an, die innerhalb bzw. außerhalb der folgenden Sprachen liegen.

(a) $A = \{w \mid w \in W^2 \wedge w \in W^3\}$
(b) $B = \{w \in W^* \mid |w| = 3\}$
(c) $C = \{w \in \Sigma^* \mid \exists u \in \Sigma^*. \exists v \in W. w = uv\}$
(d) $D = \{w \in \Sigma^* \mid \exists u \in \Sigma^*. uw = w^2 u\}$
(e) $E = \{(\mathsf{ba}^n\mathsf{b})^n \mid n \in \mathbb{N}_0\}$
(f) $F = \{w \in \Sigma^* \mid \exists n \in \mathbb{N}_0.\ |w|_{\mathsf{a}} = n \cdot |w|_{\mathsf{b}}\}$

Hinweise:

- Mit $|w|_{\mathsf{a}}$ bezeichnen wir die Anzahl der as in w.

Lösungsskizze

(a) $A = \{\mathsf{a}^6\}$, $\mathsf{a}, \mathsf{b}, \mathsf{aaa} \in \Sigma^* \setminus A$
(b) $\mathsf{aaa}, \mathsf{aab}, \mathsf{bbb} \in B$, $\mathsf{a}, \mathsf{bb}, \mathsf{aaab} \in \Sigma^* \setminus B$
(c) $\mathsf{aa}, \mathsf{bbaa}, \mathsf{bab} \in C$, $\mathsf{ba}, \mathsf{aba}, \mathsf{bba} \in \Sigma^* \setminus C$
(d) $D = \{\varepsilon\}$, $\mathsf{a}, \mathsf{b}, \mathsf{aa} \in \Sigma^* \setminus D$
(e) $\varepsilon, \mathsf{bab}, \mathsf{baabbaab} \in E$, $\mathsf{bb}, \mathsf{baab}, \mathsf{abba} \in \Sigma^* \setminus E$
(f) $\varepsilon, b, aaaabb \in F$, $a, aa, aaabb \in \Sigma^* \setminus F$

AUFGABE 1.3.

Stufe B

Sei $\Sigma = \{\mathsf{a}, \mathsf{b}\}$, $\mathsf{A} = \{\varepsilon, \mathsf{a}, \mathsf{ab}\}$ und $\mathsf{B} = \{\mathsf{a}, \mathsf{ba}\}$. Wir betrachten nun verschiedene Verknüpfungen. Geben Sie für endliche Mengen alle Elemente und für unendliche drei Beispiele an:

(a) AB
(b) A^2
(c) B^0
(d) $\mathsf{A}\emptyset$
(e) $\mathsf{A} \times (\emptyset^*)$
(f) $\mathsf{A}\Delta\mathsf{B} \coloneqq (\mathsf{A} \cup \mathsf{B}) \setminus (\mathsf{A} \cap \mathsf{B})$
(g) $\overline{B} \coloneqq \Sigma^* \setminus B$

Lösungsskizze

(a) $\mathsf{AB} = \{a, \mathsf{ba}, \mathsf{aa}, \mathsf{aba}, \mathsf{abba}\}$
(b) $\mathsf{A}^2 = \{\varepsilon, \mathsf{a}, \mathsf{ab}, \mathsf{aa}, \mathsf{aab}, \mathsf{aba}, \mathsf{abab}\}$
(c) $\mathsf{B}^0 = \{\varepsilon\}$
(d) $\mathsf{A}\emptyset = \emptyset$
(e) $\mathsf{A} \times (\emptyset^*) = \{(\varepsilon, \varepsilon), (\mathsf{a}, \varepsilon), (\mathsf{ab}, \varepsilon)\}$
(f) $\mathsf{A}\Delta\mathsf{B} = \{\varepsilon, \mathsf{ab}, \mathsf{ba}\}$
(g) $\varepsilon, \mathsf{b}, \mathsf{ab} \in \overline{B}$

AUFGABE 1.4.

Stufe C / D

Bestimmen Sie für jede der folgenden Sprachen eine passende Grammatik G, so dass $\mathsf{L}(\mathsf{G})$ genau die Sprache ist.

(a) $A = \{w \in \{0,1\}^* \mid |w| \mod 2 = 0\}$
(b) $B = \{w \in \{0,1\}^+ \mid w \mod 2 = 0\}$, wobei w eine Binärzahl mit most-significant-bit-first Encoding darstellt
(c) $C = \{ww^R \mid w \in \{\mathsf{a}, \mathsf{b}\}^*\}$
(d) $D = \{ww \mid w \in \{\mathsf{a}, \mathsf{b}\}^*\}$
(e) $E = \{\mathsf{a}^{n^2} \mid n \geq 0\}$

Hinweise:

- Wir bezeichnen mit w^R die Spiegelung von w, d.h. $(\mathsf{abb})^R = \mathsf{bba}, \varepsilon^R = \varepsilon$.
- Eine mögliche Lösung von Aufgabenteil (d) erweitert die Grammatik von Aufgabenteil (c) passend.
- Für Aufgabe (e) ist `https://math.stackexchange.com/questions/136237/` hilfreich.
- *Achtung*: Aufgabenteile (d) und (e) sind deutlich anspruchsvoller. Es macht daher nichts, wenn Sie die Aufgaben erst in ein paar Wochen lösen können!

Lösungsskizze

(a) $\mathsf{G} = (\{\mathsf{S}\}, \{0,1\}, \mathsf{P}, \mathsf{S})$ mit $\mathsf{P} : \mathsf{S} \to 01\mathsf{S} \mid 10\mathsf{S} \mid 11\mathsf{S} \mid 00\mathsf{S} \mid \varepsilon$

(b) $\mathsf{G} = (\{\mathsf{S}\}, \{0,1\}, \mathsf{P}, \mathsf{S})$ mit $\mathsf{P} : \mathsf{S} \to 0\mathsf{S} \mid 1\mathsf{S} \mid 0$

(c) $\mathsf{G} = (\{\mathsf{S}\}, \{\mathsf{a},\mathsf{b}\}, \mathsf{P}, \mathsf{S})$ mit $\mathsf{P} : \mathsf{S} \to \mathsf{aSa} \mid \mathsf{bSb} \mid \varepsilon$

(d) Es gibt hier drei Ansätze zur Lösung:

(i) $\mathsf{G} = (\{S, X, O, A, B\}, \{\mathsf{a},\mathsf{b}\}, P, S)$ mit

$$\begin{array}{lllllll} P: & S \to XO & X \to XaA \mid XbB \mid \varepsilon & O \to \varepsilon & & & \\ & Aa \to aA & Ab \to bA & Ba \to aB & Bb \to bB & AO \to Oa & BO \to Ob \end{array}$$

Wir schreiben alle Buchstaben das Wortes doppelt und verschieben A und B bis zum Ende (markiert mit O) und wandeln dort das Zeichen wieder um. Vielen Dank an David Schneller für diese Lösung.

(ii) $\mathsf{G} = (\{S, X, O, A, B, \overrightarrow{A}, \overrightarrow{B}\}, \{\mathsf{a},\mathsf{b}\}, P, S)$ mit

$$\begin{array}{llll} P: & S \to XO & X \to aXA \mid bXB \mid O & O \to \varepsilon \\ & OA \to O\overrightarrow{A} & \overrightarrow{A}A \to A\overrightarrow{A} & \overrightarrow{A}B \to B\overrightarrow{A} \\ & OB \to O\overrightarrow{B} & \overrightarrow{B}A \to A\overrightarrow{B} & \overrightarrow{B}B \to B\overrightarrow{B} \\ & \overrightarrow{A}O \to a & \overrightarrow{A}a \to aa & \overrightarrow{A}b \to ab \\ & \overrightarrow{B}O \to b & \overrightarrow{B}a \to ba & \overrightarrow{B}b \to bb \end{array}$$

Wir erzeugen erst wOw^RO und verschieben dann alle Buchstaben von der Mitte beginnend an das Ende (O). Vielen Dank an Jan Schuchardt für diese Lösung.

(iii) $\mathsf{G} = (\{S, S', A, \overleftarrow{A}, B, \overleftarrow{B}, \overrightarrow{X}, \bullet, \circ\}, \{\mathsf{a},\mathsf{b}\}, \mathsf{P}, \mathsf{S})$ mit

$$\begin{array}{lllll} P: & S \to S'\bullet & S' \to AS'A \mid BS'B \mid \circ & A \to a & B \to b \\ & A\bullet \to \overleftarrow{A}\circ & B\bullet \to \overleftarrow{B}\circ & & \\ & A\overleftarrow{A} \to \overleftarrow{A}A & A\overleftarrow{B} \to \overleftarrow{B}A & B\overleftarrow{A} \to \overleftarrow{A}B & B\overleftarrow{B} \to \overleftarrow{B}B \\ & \circ\overleftarrow{A} \to A\bullet & \circ\overleftarrow{B} \to B\bullet & & \\ & \bullet \to \circ\overrightarrow{X} & & & \\ & \overrightarrow{X}A \to A\overrightarrow{X} & \overrightarrow{X}B \to B\overrightarrow{X} & & \\ & \overrightarrow{X}\circ \to \bullet & & & \\ & \circ\bullet \to \varepsilon & \bullet\circ \to \varepsilon & & \end{array}$$

Ein beliebiges Wort w zu erzeugen ist leicht. Im Aufgabenteil (c) sieht man, wie man ww^R bilden kann. Die Idee hinter dieser Aufgabe ist, zuerst ww^R zu bilden, aber alle Elemente von w^R voerst als Nicht-Terminale darzustellen. Im Anschluss verschieben wir die Nicht-Terminale in w^R an die richtige Position und ersetzen Sie schlussendlich durch Terminale. Die Intuition der Lösung ist, dass $\bullet, \overrightarrow{X}, \overleftarrow{X}$ den aktiven Bereich der Berechnung darstellen, wobei $X \in \{A, B\}$. $\circ$ ist der deaktivierte Begrenzer, der markiert, bis zu welcher Position wir noch Buchstaben verschieben müssen. $\circ$ trennt zunächst die w und w^R voneinander. Mit jedem Buchstaben, den wir vom Ende von w^R nach vorne verschieben, wandert $\circ$ weiter nach rechts.

(e) $\mathsf{G} = (\{S, A, \overleftarrow{A}, \overrightarrow{A}, \bullet, \circ, \#, \$, X_\#\}, \{a, b\}, \mathsf{P}, \mathsf{S})$ mit

$$\begin{array}{llll} P: & S \to \bullet AX_\#\$ & X_\# \to \#X_\# \mid \varepsilon & A \to a \\ & \bullet A \to A \circ \overrightarrow{A} & \overrightarrow{A}A \to A\overrightarrow{A} & \\ & \circ\overrightarrow{A}\# \to \bullet AAA & A\overrightarrow{A}\# \to \overleftarrow{A}\#A & \\ & A\overleftarrow{A} \to \overleftarrow{A}A & & \\ & \overrightarrow{A}\$ \to \varepsilon & \circ \to \varepsilon & \circ\overleftarrow{A} \to \bullet A \end{array}$$

Die Grammatik rät die natürliche Zahl n und erzeugt alle ungerade Zahlen $1, 3, \ldots 2n+1$ (1 wird mit einem Nicht-Terminal dargestellt, welches am Ende durch ein Terminal ersetzt wird, 3 mit drei solchen Nicht-Terminalen, etc.). Die Summe all dieser ungeraden Zahlen ergibt gerade n^2, damit erhalten wir gerade eine Zeichensequenz mit n^2 Nicht-Terminalen, die im letzten Schritt dann mit Terminalen ersetzt werden.

AUFGABE 1.5.

Stufe D

Sei Σ ein Alphabet und $\mathsf{A}, \mathsf{B}, \mathsf{C} \subseteq \Sigma^*$ beliebig. Entscheiden Sie, ob die folgenden Aussagen wahr oder falsch sind. Geben Sie für korrekte Aussagen einen Beweis an oder widerlegen Sie falsche mithilfe eines geeigneten Gegenbeispiels.

(a) $\mathsf{A}^* = \mathsf{A}^+$ gdw. $\varepsilon \in \mathsf{A}$

(b) $\mathsf{A}(\mathsf{B} \cap \mathsf{C}) = \mathsf{AB} \cap \mathsf{AC}$

(c) Falls $\mathsf{A} \subseteq \mathsf{B}$, dann $\mathsf{A}^n \subseteq \mathsf{B}^n$ für jedes $n \in \mathbb{N}_0$.

(d) Unter der Annahme $\mathsf{A} \neq \emptyset$ gilt: $\mathsf{A} = \mathsf{AA}$ gdw. $\mathsf{A} = \mathsf{A}^*$.

Lösungsskizze

(a) Falls $\varepsilon \in \mathsf{A}$: $\mathsf{A}^+ = \mathsf{A} \cup \mathsf{A}\mathsf{A}^+ = \{\varepsilon\} \cup \mathsf{A} \cup \mathsf{A}\mathsf{A}^+ = \{\varepsilon\} \cup \mathsf{A}^+ = \mathsf{A}^*$
Falls $\varepsilon \notin \mathsf{A}$: $w \in \mathsf{A}^n$ impliziert $|w| \geq n$, da w sich in n nicht leere Wörter aus A faktorisieren lässt. Damit hat jedes Wort aus A^+ positive Länge, somit $\varepsilon \notin \mathsf{A}^+$ und damit auch $\mathsf{A}^* \neq \mathsf{A}^+$.

(b) Wähle $\mathsf{A} = \{\mathsf{a}\}^*$, $\mathsf{B} = \{\varepsilon\}$ und $\mathsf{C} = \{\mathsf{a}\}$.
Dann gilt $\mathsf{A}(\mathsf{B} \cap \mathsf{C}) = \mathsf{A}\emptyset = \emptyset \neq \{\mathsf{a}^i \mid i \geq 1\} = \{\mathsf{a}^i \mid i \geq 0\} \cap \{a[i] \mid i \geq 1\} = \mathsf{AB} \cap \mathsf{AC}$.

(c) Sei $w \in \mathsf{A}^n$ und $n \in \mathbb{N}$ beliebig. Dann gilt $w = w^1 \dots w^n$ für $w^i \in \mathsf{A} \subseteq \mathsf{B}$ nach Def. von A^n bzw. Voraussetzung. Somit auch $w \in \mathsf{B}^n$ nach Def. von B^n.

(d) $\Leftarrow$: Es gelte $\mathsf{A} = \mathsf{A}^*$.
Wir zeigen $\mathsf{A}^*\mathsf{A}^* = \mathsf{A}^*$. Sei $w \in \mathsf{A}^*\mathsf{A}^*$. Dann existieren nach Definition $u, v \in \mathsf{A}^*$ mit $uv = w$. Weiterhin existieren $u^1 \dots u^k, v^1 \dots v^l \in \mathsf{A}$ mit $k, l \in \mathbb{N}_0$, sodass $u = u^1 \dots u^k$ und $v = v^1 \dots v^l$. Somit $w = u^1 \dots u^k v^1 \dots v^l \in \mathsf{A}^*$. Da w beliebig, folgt $\mathsf{A}^*\mathsf{A}^* \subseteq \mathsf{A}^*$. Da $\varepsilon \in \mathsf{A}^*$ gilt stets $\mathsf{A}^* \subseteq \mathsf{A}^*\mathsf{A}^*$.
Dann $\mathsf{A} = \mathsf{A}^* = \mathsf{A}^*\mathsf{A}^* = \mathsf{AA}$.
$\Rightarrow$: Es gelte $\mathsf{A} = \mathsf{AA}$. Nach Definition gilt $\mathsf{A} \subseteq \mathsf{A}^*$. Somit bleibt nur noch $\mathsf{A} \supseteq \mathsf{A}^*$ zu zeigen:
Wir zeigen zuerst $\mathsf{A}^0 \overset{\mathsf{Def}}{=} \{\varepsilon\} \subseteq \mathsf{A}$.
Sei $w \in \mathsf{A} = \mathsf{AA}$ ein kürzestes Wort. Dann lässt sich w in zwei Wörter $u, v \in \mathsf{A}$ mit $|w| = |u| + |v|$ faktorisieren. Da w ein kürzestes Wort ist, folgt $0 = |w| = |u| = |v|$, also $w = \varepsilon$.
Es bleibt zu zeigen, dass $\mathsf{A}^k \subseteq \mathsf{A}$ für alle $k \geq 1$.

Nach Annahme gilt die Induktionsbasis. Für $k > 1$ beliebig, aber fest folgt induktiv: $\mathsf{A}^{k+1} = \mathsf{A}^k\mathsf{A} \overset{\mathsf{IH}}{\subseteq} \mathsf{AA} = \mathsf{A}$.
Somit $\mathsf{A}^* = \bigcup_{k \geq 0} \mathsf{A}^k \subseteq \mathsf{A}$.

AUFGABE 1.6. Stufe D

Sei $\mathsf{L} = \{w \in \Sigma^* \mid |w|_{\mathsf{a}} = 2\,|w|_{\mathsf{b}}\}$ und seien folgende drei Grammatiken gegeben:

(a) $\mathsf{S} \to \varepsilon \mid \mathsf{SS} \mid \mathsf{aab} \mid \mathsf{aba} \mid \mathsf{baa}$
(b) $\mathsf{S} \to \varepsilon \mid \mathsf{SS} \mid \mathsf{aSaSb} \mid \mathsf{aSbSa} \mid \mathsf{bSaSa}$
(c) $\mathsf{S} \to \varepsilon \mid \mathsf{SaX} \mid \mathsf{XaS} \qquad \mathsf{X} \to \mathsf{SaSbS} \mid \mathsf{SbSaS} \mid \mathsf{XX}$

Entscheiden Sie, ob für die jeweils angegebenen Grammatiken $\mathsf{L(G)} = \mathsf{L}$ gilt. Begründen Sie Ihre Antwort, indem Sie entweder ein Gegenbeispiel angeben oder die Äquivalenz mit vollständiger Induktion beweisen.

Lösungsskizze

(a) Die Grammatik erfüllt nicht die gewünschte Eigenschaft. Das Wort aaaabb ist in der Sprache L enthalten, da es vier Mal a und zwei Mal b enthält. Es kann aber nicht von der Grammatik erzeugt werden, da $\mathsf{aaaabb} \notin \{\varepsilon, \mathsf{aab}, \mathsf{aba}, \mathsf{baa}\}^2$

(b) Die Grammatik $\mathsf{G} = (\{\mathsf{S}\}, \{\mathsf{a}, \mathsf{b}\}, \mathsf{P}, \mathsf{S})$ mit den folgenden Produktionen P erfüllt die gewünschten Eigenschaften:
Korrektheitsbeweis: $\mathsf{L} \subseteq \mathsf{L(G)}$. Induktion über die Länge von $w \in \mathsf{L}$.

- $|w| = 0$: Dann ist $w = \varepsilon$ und kann mit $S \to \varepsilon$ produziert werden. Deshalb $w \in \mathsf{L(G)}$.
- $1 \leq |w| \leq 2$: Es gibt kein $w \in \mathsf{L}$ mit Länge 1 oder 2.
- $|w| \geq 3$: Sei $w \in \mathsf{L}$. Wir verwenden die Funktion h (*Höhe*) um die gewichtete Differenz zwischen der Anzahl der as und bs zu beschreiben:

$$h(w) = 2\,|w|_b - |w|_a$$

Offensichtlich gilt $h(w) = 0$ gdw. $w \in \mathsf{L}$. Im Folgenden sagen wir, dass w eine Nullstelle hat, wenn eine nicht-triviale Zerlegung $w = uv$ mit $u \neq \varepsilon \neq v$ und $h(u) = h(v) = 0$ existiert. Wir betrachten folgende Fälle:
  - w hat eine Nullstelle. Da $|u| < |w| > |v|$ gilt, können wir die Induktionshypothese anwenden und erhalten $S \to \mathsf{SS} \to^* u\mathsf{S} \to^* uv$ und somit $uv \in\ L(G)$.
  - w hat keine Nullstelle. Daher trifft einer der drei folgenden Fälle zu:
    * $w = \mathsf{b}u$ für ein $u \in \Sigma^*$ mit $h(u) = -2$. Da w keine Nullstellen hat und nur 1-Schritten fallen kann, ist h nicht-negativ für alle Präfixe von w. Weil $h(w) = 0$ gelten muss, kann u in $u = v\mathsf{aa}$ mit $v \in \Sigma^*$ und $h(v) = 0$ zerlegt werden. Somit $\mathsf{S} \to \mathsf{bSaSa} \to \mathsf{bSaa} \to^* \mathsf{b}v\mathsf{aa} = w$ und $w \in \mathsf{L(G)}$.
    * $w = \mathsf{a}u\mathsf{a}$ für ein $u \in \Sigma^*$ mit $h(u) = 2$. Da w keine Nullstellen hat, kann u in $u = v_1\mathsf{b}v_2$ mit $v_1, v_2 \in \Sigma^*$ und $h(v_1) = h(v_2) = 0$ zerlegt werden. Somit $\mathsf{S} \to \mathsf{aSbSa} \to \mathsf{bSaSa} \to^* \mathsf{a}v_1\mathsf{b}v_2\mathsf{a} = w$ und $w \in \mathsf{L(G)}$.
    * $w = u\mathsf{b}$ für ein $u \in \Sigma^*$ mit $h(u) = -2$. Analog.

$\mathsf{L} \supseteq \mathsf{L(G)}$. Induktion über die Länge der Ableitung von $w \in \mathsf{L}(G)$.
- $\mathsf{S} \to \varepsilon$: $\varepsilon \in L$
- $\mathsf{S} \to \mathsf{SS} \to^* w_1w_2$: Nach Induktion $w_1, w_2 \in \mathsf{L}$ und somit auch $w_1w_2 \in \mathsf{L}$.
- $\mathsf{S} \to \mathsf{aSaSb} \to^* \mathsf{a}w_1\mathsf{a}w_2\mathsf{b}$: Nach Induktion $w_1, w_2 \in \mathsf{L}$ und somit auch $\mathsf{a}w_1\mathsf{a}w_2\mathsf{b} \in \mathsf{L}$.

- Restliche Fälle analog.

(c) Die Grammatik erfüllt nicht die gewünschte Eigenschaft, da sie das Wort aabba erzeugt ($\mathsf{S} \to \mathsf{SaX} \to \mathsf{SaXX} \to^* \mathsf{SaSaSbSSbSaS} \to^* \mathsf{aabba}$). Da das Wort aber drei Mal a und zwei Mal b enthält, ist es nicht in der Sprache L enthalten.

AUFGABE 1.7. Stufe E

Eine Sprache $\mathsf{L} \subseteq \Sigma^*$ heißt *linear*, falls es eine kontextfreie Grammatik G gibt, so dass auf der rechten Seite einer jeden Produktion höchstens ein Nichtterminal steht (d.h. $\mathsf{P} \subseteq \mathsf{V} \times (\Sigma^*\mathsf{V}\Sigma^* \cup \Sigma^*)$). Die folgende Aussage gilt:

Ist L linear mit $\varepsilon \notin \mathsf{L}$, dann gibt es eine CFG G, so dass $\mathsf{L} = \mathsf{L}(\mathsf{G})$ und für jede Produktion $\mathsf{X} \to \gamma$ gilt: $\gamma \in \Sigma\mathsf{V} \cup \mathsf{V}\Sigma \cup \Sigma$.

(a) Vergewissern Sie sich zunächst, dass die Aussage korrekt ist, indem Sie die Grammatik G in eine Form bringen, so dass für jede Produktion $\mathsf{X} \to \gamma$ die Bedingung $\gamma \in \Sigma\mathsf{V} \cup \mathsf{V}\Sigma \cup \Sigma$ gilt.

$$\mathsf{S} \to \mathsf{cSc} \mid \mathsf{aaXbb} \qquad \mathsf{X} \to \mathsf{c} \mid \mathsf{S}$$

(b) Beweisen Sie nun die obige Aussage.

Lösungsskizze

(a) Die Grammatik ist linear, da auf jeder rechten Seite einer Produktion maximal ein Nicht-Terminal steht. Die Produktionen, die das Nicht-Terminal X auf der linken Seite stehen haben, erfüllen bereits die Bedingung. Wann immer auf beiden Seiten eines Nicht-Terminals Terminale vorkommen, müssen wir eine Seite stehen lassen und die andere in eine Art "Hilfs-Nicht-Terminal" verschieben. So wird aus $\mathsf{S} \to \mathsf{cSc} \mid \mathsf{aaXbb}$ in unserem Beispiel $\mathsf{S} \to \mathsf{cY} \mid \mathsf{aaZ} \quad \mathsf{Y} \to \mathsf{Sc} \quad \mathsf{Z} \to \mathsf{Xbb}$. Haben wir die Grammatik in eine Form überführt, so dass für jede Produktionsregel, die ein Nicht-Terminal enhält, gilt, dass nur noch auf einer Seite des Nicht-Terminals Terminale stehen, so müssen wir Sequenzen von Terminalen ebenfalls in "Hilfs-Nicht-Terminale" auslagern. Dabei erhalten wir die Produktionsregeln: $\mathsf{S} \to \mathsf{cY} \mid \mathsf{aZ}^1 \quad \mathsf{Z}^1 \to \mathsf{aZ} \quad \mathsf{Y} \to \mathsf{Sc} \quad \mathsf{Z} \to \mathsf{S}^1\mathsf{b} \quad \mathsf{S}^1 \to \mathsf{Xb}$.

(b) Start mit beliebiger linearer CFG, Entfernen aller ε-Regeln und aller Kettenproduktionen, schließlich: falls $\mathsf{X} \to \mathsf{a}w\mathsf{Y}v$ mit $a \in \Sigma, w \in \Sigma^+, v \in \Sigma^*, \mathsf{Y} \in \mathsf{V}$, Übergang zu $\mathsf{X} \to \mathsf{aX}_{w\mathsf{Y}v}$ und $\mathsf{X}_{w\mathsf{Y}v} \to w\mathsf{Y}v$, wiederholen bis alle Regeln von der Form $\mathsf{X} \to \mathsf{aY}$ oder $\mathsf{X} \to \mathsf{Y}v$ mit $v \in \Sigma^+$. Entsprechend $\mathsf{X} \to \mathsf{Y}v$ noch überführen.