

Einführung in die theoretische Informatik
Sommersemester 2019 – Übungsblatt Lösungsskizze 1

AUFGABE 1.1. (*Wichtige Begriffe*)

Stufe A

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

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

AUFGABE 1.2. (*Automata Tutor*)

Aufgrund der TUMonline Wartung vom 29.04. bis 01.05. wird der Automata Tutor erst am 02.05. verfügbar sein.

Stufe B

AutomataTutor ist ein online Übungstool, mit dem Studierende die Grundlagen der Automatentheorie erlernen können. Es bietet die Möglichkeit, viele der für die Vorlesung relevanten Probleme selbständig zu üben.

In dieser Aufgabe geht es darum, dass Sie das Übungstool AutomataTutor kennenlernen. In den kommenden Wochen werden Sie über das Tool auch Hausaufgabenpunkte sammeln können.

- Rufen Sie die AutomataTutor Version der TUM auf. (<https://automata.model.in.tum.de/>)
- Loggen Sie sich mit Ihrer TUM-kennung (ab12cde) ein.
- Lösen Sie die Aufgaben unter "Contextfree Languages → Find Derivation" und "Contextfree Languages → Words in Grammar".

AUFGABE 1.3.

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

Stufe B

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

Hinweise:

- Mit $|w|_a$ bezeichnen wir die Anzahl der as in w .

Lösungsskizze

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

AUFGABE 1.4.

Gegeben seien folgende drei Grammatiken $G_1 = (\{S, X\}, \Sigma, P_1, S)$, $G_2 = (\{S, X\}, \Sigma, P_2, S)$ und $G_3 = (\{S, S', A, B, C\}, \Sigma, P_3, S)$ über dem Alphabet $\Sigma = \{a, b, c\}$ mit folgende Produktionen:

Stufe B

- $P_1: S \rightarrow aX \quad X \rightarrow bS \mid b$
 $P_2: S \rightarrow aSb \mid SX \mid \varepsilon \quad X \rightarrow cX$
 $P_3: S \rightarrow S' \mid \varepsilon \quad S' \rightarrow ABC \mid ABCS'$
 $A \rightarrow a \quad B \rightarrow b \quad C \rightarrow c \quad CB \rightarrow BC \quad CA \rightarrow AC \quad BA \rightarrow AB$

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

- (a) $w_1 = abab \in L(G_1), w_2 = ba \notin L(G_1), w_3 = \varepsilon \notin L(G_1)$
(b) $u_1 = aabb \in L(G_2), u_2 = aabbcc \notin L(G_2)$
(c) $v_1 = aabbcc \in L(G_3), v_2 = aaabbc \notin L(G_3)$

Lösungsskizze

- (a) $S \rightarrow aX \rightarrow abS \rightarrow abaX \rightarrow w_1, w_2, w_3 \notin L(G_1)$, da der erste Schritt jeder Produktion ein a an der ersten Stelle erzeugt.
- (b) $S \rightarrow aSb \rightarrow aaSbb \rightarrow u_1, u_2$ kann nicht erzeugt werden, da $X \rightarrow Xc$ nie "terminiert".
- (c) $S \rightarrow S' \rightarrow^* ABCABC \rightarrow ABACBC \rightarrow AABCBC \rightarrow AABBC \rightarrow^* aabbcc = v_1$. 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 1.5.

Stufe B

Sei $\Sigma = \{a, b\}$, $A = \{\varepsilon, a, ab\}$ und $B = \{a, ba\}$. 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) $A\emptyset$
- (e) $A \times (\emptyset^*)$
- (f) $A\Delta B := (A \cup B) \setminus (A \cap B)$
- (g) $\overline{B} := \Sigma^* \setminus B$

Lösungsskizze

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

AUFGABE 1.6.

Stufe C / D

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

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

Hinweise:

- Wir bezeichnen mit w^R die Spiegelung von w , d.h. $(abb)^R = bba, \varepsilon^R = \varepsilon$.
- 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) $G = (\{S\}, \{0, 1\}, P, S)$ mit $P: S \rightarrow 01S \mid 10S \mid 11S \mid 00S \mid \varepsilon$
- (b) $G = (\{S\}, \{0, 1\}, P, S)$ mit $P: S \rightarrow 0S \mid 1S \mid 0$. Lösung ohne führende Nullen:
 $G = (\{S, X\}, \{0, 1\}, P, S)$ mit $P: S \rightarrow 1X \mid 0 \quad X \rightarrow 0X \mid 1X \mid 0$
- (c) $G = (\{S\}, \{a, b\}, P, S)$ mit $P: S \rightarrow aSa \mid bSb \mid \varepsilon$
- (d) Es gibt hier drei Ansätze zur Lösung:
 - (i) $G = (\{S, X, O, A, B\}, \{a, b\}, P, S)$ mit

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

Wir schreiben alle Buchstaben des 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) $G = (\{S, X, O, A, B, \vec{A}, \vec{B}\}, \{a, b\}, P, S)$ mit

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

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

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

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-Terminals darzustellen. Im Anschluss verschieben wir die Nicht-Terminals in w^R an die richtige Position und ersetzen sie schlussendlich durch Terminals. 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) $G = (\{S, A, \overleftarrow{A}, \overrightarrow{A}, \bullet, \circ, \#, \$, X_\#\}, \{a, b\}, P, S)$ mit

$$\begin{aligned}
 P: \quad & S \rightarrow \bullet AX_\# \$ \quad X_\# \rightarrow \# X_\# \mid \varepsilon \quad A \rightarrow a \\
 & \bullet A \rightarrow A \circ \overleftarrow{A} \quad \overleftarrow{A} A \rightarrow A \overleftarrow{A} \\
 & \circ \overleftarrow{A} \# \rightarrow \bullet AAA \quad A \overleftarrow{A} \# \rightarrow \overleftarrow{A} \# A \\
 & A \overleftarrow{A} \rightarrow \overleftarrow{A} A \\
 & \overleftarrow{A} \$ \rightarrow \varepsilon \quad \circ \rightarrow \varepsilon \quad \circ \overleftarrow{A} \rightarrow \bullet A
 \end{aligned}$$

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

AUFGABE 1.7.

Stufe D

Sei Σ ein Alphabet und $A, B, C \subseteq \Sigma^*$ 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.

- $A^* = A^+$ gdw. $\varepsilon \in A$
- $A(B \cap C) = AB \cap AC$
- Falls $A \subseteq B$, dann $A^n \subseteq B^n$ für jedes $n \in \mathbb{N}_0$.
- Unter der Annahme $A \neq \emptyset$ gilt: $A = AA$ gdw. $A = A^*$.

Lösungsskizze

- Falls $\varepsilon \in A$: $A^+ = A \cup AA^+ = \{\varepsilon\} \cup A \cup AA^+ = \{\varepsilon\} \cup A^+ = A^*$
 Falls $\varepsilon \notin A$: $w \in A^n$ 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 $\varepsilon \notin A^+$ und damit auch $A^* \neq A^+$.
- Wähle $A = \{a\}^*$, $B = \{\varepsilon\}$ und $C = \{a\}$.
 Dann gilt $A(B \cap C) = A\emptyset = \emptyset \neq \{a^i \mid i \geq 1\} = \{a^i \mid i \geq 0\} \cap \{a^i \mid i \geq 1\} = AB \cap AC$.
- Sei $w \in A^n$ und $n \in \mathbb{N}$ beliebig. Dann gilt $w = w^1 \dots w^n$ für $w^i \in A \subseteq B$ nach Def. von A^n bzw. Voraussetzung. Somit auch $w \in B^n$ nach Def. von B^n .
- \Leftarrow : Es gelte $A = A^*$.

Wir zeigen $A^*A^* = A^*$. Sei $w \in A^*A^*$. Dann existieren nach Definition $u, v \in A^*$ mit $uv = w$. Weiterhin existieren $u^1 \dots u^k, v^1 \dots v^l \in A$ mit $k, l \in \mathbb{N}_0$, sodass $u = u^1 \dots u^k$ und $v = v^1 \dots v^l$. Somit $w = u^1 \dots u^k v^1 \dots v^l \in A^*$. Da w beliebig, folgt $A^*A^* \subseteq A^*$. Da $\varepsilon \in A^*$ gilt stets $A^* \subseteq A^*A^*$.
 Dann $A = A^* = A^*A^* = AA$.

\Rightarrow : Es gelte $A = AA$. Nach Definition gilt $A \subseteq A^*$. Somit bleibt nur noch $A \supseteq A^*$ zu zeigen:

Wir zeigen zuerst $A^0 \stackrel{\text{Def}}{=} \{\varepsilon\} \subseteq A$.

Sei $w \in A = AA$ ein kürzestes Wort. Dann lässt sich w in zwei Wörter $u, v \in A$ mit $|w| = |u| + |v|$ faktorisieren. Da w ein kürzestes Wort ist, folgt $0 = |w| = |u| = |v|$, also $w = \varepsilon$.

Es bleibt zu zeigen, dass $A^k \subseteq A$ für alle $k \geq 1$.

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

AUFGABE 1.8.

Stufe D

Sei $L = \{w \in \Sigma^* \mid |w|_a = 2|w|_b\}$ und seien folgende drei Grammatiken gegeben:

- (a) $S \rightarrow \varepsilon \mid SS \mid aab \mid aba \mid baa$
- (b) $S \rightarrow \varepsilon \mid SS \mid aSaSb \mid aSbSa \mid bSaSa$
- (c) $S \rightarrow \varepsilon \mid SaX \mid XaS \quad X \rightarrow SaSbS \mid SbSaS \mid XX$

Entscheiden Sie, ob für die jeweils angegebenen Grammatiken $L(G) = L$ 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 $aaaabb \notin \{\varepsilon, aab, aba, baa\}^2$
- (b) Die Grammatik $G = (\{S\}, \{a, b\}, P, S)$ mit den folgenden Produktionen P erfüllt die gewünschten Eigenschaften:

Korrektheitsbeweis: Wir zeigen $L = L(G)$ durch $L \subseteq L(G)$ und $L(G) \subseteq L$.

$L \subseteq L(G)$: Mittels Induktion über die Länge von w beweisen wir $w \in L \Rightarrow w \in L(G)$.

IBasis:

- $|w| = 0$: Dann ist $w = \varepsilon$ und kann mit $S \rightarrow \varepsilon$ produziert werden. Nach Definition von $L(G)$ gilt $w \in L(G)$.
- $1 \leq |w| \leq 2$: Es gibt kein $w \in L$ mit Länge 1 oder 2.

ISchritt: Sei $l \in \mathbb{N}_0$ beliebig aber fixiert.

- IAnnahme: Für alle w mit $|w| \leq l$ gilt $w \in L \Rightarrow w \in L(G)$.
- IBehauptung: Für alle w mit $|w| = l + 1$ gilt $w \in L \Rightarrow w \in L(G)$.
- IBeweis: Sei $w \in L$ mit $|w| = l + 1$. Wir verwenden die Funktion h (*Höhe*) um die gewichtete Differenz zwischen der Anzahl der **a**s und **b**s zu beschreiben:

$$h(w) = 2|w|_b - |w|_a.$$

Offensichtlich gilt $h(w) = 0$ gdw. $w \in L$. Im Folgenden sagen wir, dass w eine *Nullstelle* hat, wenn eine nicht-triviale Zerlegung $w = uv$ mit $u \neq \varepsilon, v \neq \varepsilon$ 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 $S \rightarrow SS \rightarrow^* uS \rightarrow^* uv$ und somit $uv \in L(G)$.
- w hat keine Nullstelle. Daher trifft einer der drei folgenden Fälle zu:
 - * $w = bu$ für ein $u \in \Sigma^*$ mit $h(u) = -2$. Da w keine Nullstellen hat und nur immer um genau 1 fallen kann, ist h nicht-negativ für alle Präfixe von w . Weil $h(w) = 0$ gelten muss, kann u in $u = vaa$ mit $v \in \Sigma^*$ und $h(v) = 0$ zerlegt werden. Somit $S \rightarrow bSaSa \rightarrow bSaa \rightarrow^* bvaa = w$ und $w \in L(G)$.
 - * $w = aua$ für ein $u \in \Sigma^*$ mit $h(u) = 2$. Da w keine Nullstellen hat, kann u in $u = v_1bv_2$ mit $v_1, v_2 \in \Sigma^*$ und $h(v_1) = h(v_2) = 0$ zerlegt werden. Somit $S \rightarrow aSbSa \rightarrow^* av_1bv_2a = w$ und $w \in L(G)$.
 - * $w = ub$ für ein $u \in \Sigma^*$ mit $h(u) = -2$. Analog.

$L \supseteq L(G)$: Mit Induktion über die Höhe des Ableitungsbaums von w zeigen wir $w \in L(G) \Rightarrow w \in L$.

IBasis: Höhe des Ableitungsbaums ist 1, damit kommt nur $S \rightarrow \varepsilon$ in Frage. Nach Definition gilt $\varepsilon \in L$.

ISchritt: Sei $l \in \mathbb{N}_0$ beliebig aber fixiert.

- IAnnahme: Für alle w die durch Ableitungen mit einer Höhe von maximal l erzeugt werden können gilt $w \in L(G) \Rightarrow w \in L$.
- IBehauptung: Für alle w die durch Ableitungen der Höhe $l + 1$ erzeugt werden gilt $w \in L(G) \Rightarrow w \in L$.
- IBeweis: Sei $w \in L(G)$ durch eine Ableitung der Höhe $l + 1$ erzeugt. Wir zeigen die Behauptung durch Fallunterscheidung über die Wurzel des Ableitungsbaumes. Alle Kinder der Wurzel sind Bäume der Höhe maximal l und damit ist die Induktionsannahme anwendbar (Achtung: Die Höhe der Kinder muss nicht gleich l sein!).
 - $S \rightarrow SS \rightarrow^* w_1w_2$: Nach Induktion $w_1, w_2 \in L$ und somit auch $w_1w_2 \in L$.
 - $S \rightarrow aSaSb \rightarrow^* aw_1aw_2b$: Nach Induktion $w_1, w_2 \in L$ und somit auch $aw_1aw_2b \in L$.
 - Restliche Fälle analog.

Insgesamt also $L = L(G)$, was zu zeigen war.

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

Stufe E

AUFGABE 1.9.

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

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

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

$$S \rightarrow cSc \mid aaXbb \quad X \rightarrow c \mid S$$

- (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. Wir eliminieren zunächst die Kettenproduktionen und entfernen damit X :

$$S \rightarrow cSc \mid aaSbb \mid aacbb$$

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 erhalten wir in unserem Beispiel:

$$S \rightarrow cW \mid aaY \mid aacbb \quad W \rightarrow Sc \quad Y \rightarrow Sbb$$

Haben wir die Grammatik in eine Form überführt, so dass für jede Produktionsregel, die ein Nicht-Terminal enthält, gilt, dass nur noch auf einer Seite des Nicht-Terminals Terminale stehen, so müssen wir Sequenzen von Terminalen ebenfalls in "Hilfs-Nicht-Terminalen" auslagern. Dabei erhalten wir die folgenden Produktionsregeln:

$$\begin{aligned} S &\rightarrow cW \mid aY^1 \mid aZ^1 & W &\rightarrow Sc \\ Y^1 &\rightarrow aY^2 & Y^2 &\rightarrow Y^3b & Y^3 &\rightarrow Sb \\ Z^1 &\rightarrow aZ^2 & Z^2 &\rightarrow cZ^3 & Z^3 &\rightarrow bZ^4 & Z^4 &\rightarrow b \end{aligned}$$

- (b) Start mit beliebiger linearer CFG, Entfernen aller ε -Regeln und aller Kettenproduktionen (Verfahren wird in der Vorlesung im Zusammenhang mit Chomsky-Normalform erklärt werden), schließlich: falls $X \rightarrow awYv$ mit $a \in \Sigma, w \in \Sigma^+, v \in \Sigma^*, Y \in V$, Übergang zu $X \rightarrow aX_{wYv}$ und $X_{wYv} \rightarrow wYv$, wiederholen bis alle Regeln von der Form $X \rightarrow aY$ oder $X \rightarrow Yv$ mit $v \in \Sigma^+$. Entsprechend $X \rightarrow Yv$ noch überführen.