Formale Sprachen und Komplexität, SS 18,
Prof. Dr. Volker Heun

Übungsblatt 2

Abgabe: bis Mo. 30.04.2018 8 Uhr

Institut für Informatik **Ludwig-Maximilians-Universität München**
Prof. Dr. Volker Heun

Formale Sprachen und Komplexität, SS 18
Übungsblatt 2

Abgabe: bis Mo. 30.04.2018 8 Uhr

Nach Bearbeitung dieses Übungsblattes sollten Sie:

	Check
Angeben können, zu welchem Typ der Chomsky Hierarchie eine gegebene Grammatik gehört.	
Eine Grammatik zu einer informell beschriebenen Sprache herstellen können.	
Die Sprache einer gegebenen Grammatik informell beschreiben können.	
Systematisch alle Wörter einer formalen Sprache bestimmen können (falls möglich).	
Einfache Veränderungen an den Produktionsregeln einer formalen Sprache vornehmen können, ohne die Sprache zu verändern.	
Produktionen der Form $A \to \epsilon$ ersetzen können, um eine Typ-2 Grammatik zu erhalten.	
Den Unterschied zwischen dem Typ einer Grammatik und dem Typ der zugehörigen Sprache formulieren können.	

Diese Ziele sind wichtige Hinweise für die Klausur!

Hinweis: Eine Grammatik ist bereits vom Typ 0, wenn Sie eine Produktion der Form $A \to \epsilon$ enthält.

Aufgabe 2-1 schriftlich bearbeiten
Grammatiken, Chomsky-Hierarchie

Sei L die Sprache der Literale, die die Programmiersprache Java für `int`-Konstanten im Dezimalsystem erlaubt. Ein solches Literal darf mit höchstens einem Vorzeichen beginnen, muss aber nicht. Danach kommt eine nichtleere Folge von Dezimalziffern, in der keine führenden Nullen erlaubt sind: 0 und +0 und −0 sind erlaubt, aber 00 und +08 und −009 nicht.[1]

a) Geben Sie eine Grammatik $G = (V, \Sigma, P, S)$ mit $L(G) = L$ an.

Lösungsvorschlag:

Die folgende Lösung ist relativ natürlich, aber es gibt viele andere Möglichkeiten.

$G = (V, \Sigma, P, S)$ mit:

$\Sigma = \{+, -, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
$V = \{S, E, F\}$ E für „Eins bis Neun“, F für „Folge“ (nichtleer, ohne führende 0)
$P = \{$

$$S \to 0, \quad S \to +0, \quad S \to -0,$$
$$S \to F, \quad S \to +F, \quad S \to -F,$$
$$F \to E, \quad F \to F0, \quad F \to FE,$$
$$E \to 1, \quad E \to 2, \quad E \to 3, \quad E \to 4, \quad E \to 5,$$
$$E \to 6, \quad E \to 7, \quad E \to 8, \quad E \to 9$$

$\}$

[1]Wer bezweifelt, dass führende Nullen in Java relevant sind, sollte mal folgende Anweisung ausprobieren und das Ergebnis nachrechnen: `System.out.println( 060 < 50 );`
Wer davon noch nicht überzeugt ist, kann anschließend `060` durch `080` ersetzen und neu compilieren.

b) Von welchem Typ der Chomsky-Hierarchie ist Ihre Grammatik?

Lösungsvorschlag:

Typ 2 (und damit auch Typ 1 und Typ 0)

nicht vom Typ 3

c) Geben Sie für jeden Typ der Chomsky-Hierarchie an, ob

- aus Ihren obigen Lösungen folgt, dass die Sprache L von diesem Typ ist;
- aus Ihren obigen Lösungen folgt, dass die Sprache L nicht von diesem Typ ist;
- aus Ihren obigen Lösungen weder das eine noch das andere folgt.

Lösungsvorschlag:

Die obigen Lösungen besagen,
dass G eine Grammatik vom Typ 2 und 1 und 0 ist und dass $L(G) = L$ ist
Daraus folgt dass L eine Sprache vom Typ 2 und 1 und 0 ist.

Aus den obigen Lösungen folgt weder, dass L vom Typ 3 ist,
noch dass L nicht vom Typ 3 ist.

Ob L vom Typ 3 ist, hängt davon ab,
ob es eine Grammatik G' vom Typ 3 gibt mit $L = L(G')$.
(Tatsächlich gibt es solche Grammatiken, aber das folgt nicht aus den obigen Lösungen).

Aufgabe 2-2 schriftlich bearbeiten
Grammatiken, Chomsky-Hierarchie

Sei $G = (V, \Sigma, P, S)$ die Grammatik mit $V = \{S, F, B, Z\}$ und $\Sigma = \{a, b, c, 1, 2\}$ und

$$
\begin{array}{llll}
P = \{ & S \to BF, & & \\
& F \to \varepsilon, & F \to BF, & F \to ZF, \\
& B \to a, & B \to b, & B \to c, \\
& Z \to 1, & Z \to 2 & \quad \}
\end{array}
$$

a) Welche Sprache ist $L(G)$?

Lösungsvorschlag:

Analyse der Grammatik:

Das Alphabet besteht aus Buchstaben und Ziffern.

Die letzten beiden Zeilen der Produktionsmenge bedeuten,
dass man aus B genau die Buchstaben und aus Z genau die Ziffern ableiten kann.

Die Produktionen mit F auf der linken Seite ermöglichen, aus F eine beliebige Folge von Buchstaben und Ziffern abzuleiten.

Mit der ersten Produktion gilt damit $L(G) = \{w \in \Sigma^* \mid w \text{ beginnt mit einem Buchstaben}\}$

Die Grammatik erzeugt die Sprache aller Wörter über dem gegebenen Alphabet,
die in Programmiersprachen üblicherweise als Bezeichner zugelassen wären.

Man könnte die Sprache auch so spezifizieren: $L(G) = \{a, b, c\}\, \Sigma^*$

b) Von welchem Typ der Chomsky-Hierarchie ist die gegebene Grammatik?

Lösungsvorschlag:

Typ 0

wegen der ε-Produktion nicht Typ 1 (und damit auch nicht Typ 2 und nicht Typ 3)
alle anderen Produktionen sind vom Typ 2

c) Konstruieren Sie aus G eine Grammatik G' vom Typ 2 (kontextfrei) mit $L(G) = L(G')$.

Lösungsvorschlag:

Siehe: Verfahren im Buch, Seite 10/11.

$$\text{Zerlege } V = \underbrace{\{S, B, Z\}}_{\substack{V_1 \\ \not\Rightarrow^* \varepsilon}} \cup \underbrace{\{F\}}_{\substack{V_2 \\ \Rightarrow^* \varepsilon}}$$

Modifiziere P so, dass $F \Rightarrow^* \varepsilon$ „vorweggenommen" wird

$$\begin{array}{llllll} P' = \{ & \boxed{S \to B}, & S \to BF, & & & \\ & \boxed{F \to B}, & F \to BF, & \boxed{F \to Z}, & F \to ZF, & \\ & B \to a, & B \to b, & B \to c, & & \\ & Z \to 1, & Z \to 2 & & \} & \end{array}$$

Diese Grammatik G' ist vom Typ 2 (kontextfrei), aber nicht vom Typ 3 (regulär).

d) Von welchem Typ der Chomsky-Hierarchie ist die Sprache $L(G)$?

Lösungsvorschlag:

Wegen $L(G) = L(G')$ und G' vom Typ 2
ist $L(G)$ eine Sprache vom Typ 2 und 1 und 0.

Weder G noch G' sind Grammatiken vom Typ 3.

Trotzdem ist $L(G)$ eine Sprache vom Typ 3,
denn sie wird auch von folgender Grammatik erzeugt:

$$\begin{array}{llllllllll} S \to a, & S \to b, & S \to c, & F \to a, & F \to b, & F \to c, & F \to 1, & F \to 2, \\ S \to aF, & S \to bF, & S \to cF, & F \to aF, & F \to bF, & F \to cF, & F \to 1F, & F \to 2F \end{array}$$

Aufgabe 2-3 Grammatiken, Chomsky-Hierarchie

Sei $G = (V, \Sigma, P, S)$ die Grammatik mit $V = \{S, B, W\}$ und $\Sigma = \{b, e, i, n, r, s, t, u, w\}$ und $P = \{\ S \to weissB, \quad S \to weissW, \quad B \to bier, \quad W \to wein, \quad W \to wurst\ \}$.

Geben Sie eine Grammatik G' vom Typ 3 (regulär) an mit dem selben Alphabet Σ und mit $L(G) = L(G')$.

Lösungsvorschlag:

$V' = \{S, S_1, S_2, S_3, S_4, B, B_1, B_2, B_3, W, E, E_1, E_2, U, U_1, U_2, U_3\}$

$$
\begin{array}{llllll}
P' = \{ & S \to wS_1, & S_1 \to eS_2, & S_2 \to iS_3, & S_3 \to sS_4, & \\
& S_4 \to sB, & B \to bB_1, & B_1 \to iB_2, & B_2 \to eB_3, & B_3 \to r, \\
& S_4 \to sW, & W \to wE, & E \to eE_1, & E_1 \to iE_2, & E_2 \to n, \\
& & W \to wU, & U \to uU_1, & U_1 \to rU_2, & U_2 \to sU_3, \quad U_3 \to t \quad \}
\end{array}
$$

Nach diesem Prinzip kann man jede Produktion der Form $A \to xB$ oder $A \to x$ mit $x \in \Sigma^+$ in eine endliche Menge von regulären Produktionen umwandeln.

Eine Folge davon ist: Hat G nur Produktionen der Gestalt $A \to xB$ oder $A \to x$ mit $x \in \Sigma^+$, so ist $L(G)$ regulär (obwohl G nicht regulär ist).

Aufgabe 2-4 Grammatiken, Wortproblem

Sei $G = (V, \Sigma, P, S)$ die Grammatik mit $V = \{S, A\}$ und $\Sigma = \{a, -\}$ und
$P = \{\ S \to A, \qquad S \to A - S, \qquad A \to a\ \}$.

a) Geben Sie den Baum aller Wörter $w \in (V \cup \Sigma)^*$ an mit $S \Rightarrow^* w$ und $|w| \leq 5$.
(Siehe die Bemerkung am Ende von Abschnitt 1.1.1 im Buch.)

Lösungsvorschlag:

Hinweis: S. 8 im Buch

Baum aller w mit $S \Rightarrow^* w$ und $|w| \leq 5$ $\boxed{\text{auf Folie}}$.

b) Begründen Sie, warum $a - a - a \in L(G)$ und $a - - a \notin L(G)$ ist.

Lösungsvorschlag:

$a - a - a \in L(G)$, weil Ableitungen dafür existieren.

$a - - a \notin L(G)$, weil keine Ableitungen dafür existieren.
(Das Wort hat die Länge 4,
kann also in den mit $\boxed{> 5}$ markierten Teilbäumen nicht vorkommen.)