Übungsblatt 2 - mit Lösungen

Formale Sprachen und Grammatiken

{Theoretische Informatik}@AIN3
Prof. Dr. Barbara Staehle
Wintersemester 2021/2022
HTWG Konstanz

Aufgabe 2.1 Alphabete und Sprachen

Wir betrachten das Alphabet $\Sigma = \{1, 2, 3, 4, 5, a, b, c\}$, sowie die Worte $\omega_1 = ca5$, $\omega_2 = c$ und $\omega_3 = 321c$.

Teilaufgabe 2.1.1 2 Punkte

a) Geben Sie 3 Wörter an, die Worte über Σ^* (und verschieden zu $\omega_1, \omega_2, \omega_3$) sind, und 2 Wörter, die nicht zu Σ^* gehören.

b) Geben Sie 2 (beliebige) formale Sprachen über Σ^* an.

c) Bestimmen Sie $\omega_1\omega_2$, $\omega_2\omega_1\omega_3$ und ω_1^3.

d) Geben Sie Σ^0, Σ^1 und Σ^2 (andeutungsweise, nicht alle Elemente) an.

e) Bestimmen Sie die Anzahl der Elemente von Σ^5 und geben Sie ein beispielhaftes Wort aus Σ^5 an.

Lösung

a)
- $111, abc, 54a34b2 \in \Sigma^*$
- $x5y, ax, 999 \notin \Sigma^*$

b)
- $L_1 = \{abc, 12345\} \subseteq \Sigma^*$
- $L_2 = \{\varepsilon\} \subseteq \Sigma^*$
- $L_3 = \{5^n \mid n \in \mathbb{N}\} \subseteq \Sigma^*$

c)
- $\omega_1\omega_2 = ca5c$
- $\omega_2\omega_1\omega_3 = cca5321c$
- $\omega_1^3 = ca5ca5ca5$

d)
- $\Sigma^0 = \{\varepsilon\}$
- $\Sigma^1 = \{1, 2, 3, 4, 5, a, b, c\} = \Sigma$
- $\Sigma^2 = \{11, 12, 13, \ldots, 1c, 21, 21, \ldots, cc\}$

e)
- $|\Sigma^5| = 8^5 = 32768$
- $x = 1234a \in \Sigma^5$

Teilaufgabe 2.1.2 3 Punkte

Betrachten Sie zusätzlich $N = \{S, B, Z\}$, sowie die folgenden Grammatiken:

- $G_1 = (N, \Sigma, P_1, S)$ mit $P_1: \; S \rightarrow \varepsilon \mid S1 \mid S2 \mid S3 \mid S4 \mid S5$
- $G_2 = (N, \Sigma, P_2, S)$ mit $P_2: \; S \rightarrow aSa \mid bSb \mid cSc \mid a \mid b \mid c \mid \varepsilon$
- $G_3 = (N, \Sigma, P_3, S)$ mit $P_3: \; \begin{array}{lcl} S & \rightarrow & ZB \\ Z & \rightarrow & 1 \mid 2 \mid 3 \mid 4 \mid 5 \\ B & \rightarrow & a \mid b \mid c \mid aZB \mid bZB \mid cZB \end{array}$

Geben Sie an ob und wenn ja wie (geben Sie also ggf. die Ableitung an) das Wort

a) 12345 aus G_1

b) $12ab$ aus G_1

c) abc aus G_2

d) $aabbcbbaa$ aus G_2

e) $1b2a3c$ aus G_3

f) $2c3bb2$ aus G_3

abgeleitet werden kann.

Lösung

a) 12345 aus G_1: $\begin{array}{lcl} S & \Rightarrow & S5 \\ & \Rightarrow & S45 \\ & \Rightarrow & S345 \\ & \Rightarrow & S2345 \\ & \Rightarrow & 12345 \end{array}$

b) $12ab$ aus G_1: nicht ableitbar, da G_2 keine Buchstaben erzeugt.

c) abc aus G_2: nicht ableitbar, da erster und letzter Buchstabe gleich sein müssen.

d) $aabbcbbaa$ aus G_2 : $\begin{array}{lcl} S & \Rightarrow & aSa \\ & \Rightarrow & aaSaa \\ & \Rightarrow & aabSbaa \\ & \Rightarrow & aabbSbbaa \\ & \Rightarrow & aabbcbbaa \end{array}$

e) $1b2a3c$ aus G_3: $\begin{array}{lcl} S & \Rightarrow & ZB \\ & \Rightarrow & ZbZB \\ & \Rightarrow & ZbZaZB \\ & \Rightarrow & ZbZaZc \\ & \Rightarrow & ZbZa3c \\ & \Rightarrow & Za2a3c \\ & \Rightarrow & Zb2a3c \\ & \Rightarrow & 1b2a3c \end{array}$

f) $2c3bb2$ aus G_3: nicht ableitbar, da keine zwei Buchstaben aufeinander folgend können.

Teilaufgabe 2.1.3 2 Punkte

Geben Sie für jede Grammatik an, welche Sprache diese erzeugt (also $\mathscr{L}(G_1), \mathscr{L}(G_2), \mathscr{L}(G_3)$).

Lösung

a) $\mathscr{L}(G_1) = \{$beliebige Zahl mit mindestens einer Ziffer, die nur aus 1,2,3,4,5 besteht, oder$\varepsilon\} = \{1, 2, 3, 4, 5\}^*$

b) $\mathscr{L}(G_2) = \{$Palindrom aus den Buchstaben a,b,c oder das leere Wort $\} =$
$\{\omega \in \Sigma^* \mid \omega$ von vorne und hinten gelesen ist gleich$\}$

c) $\mathscr{L}(G_3) = \{$beliebig lange Hintereinaderreihung von Bausteinen
aus einer Zahl (1,2,3,4,5) und Buchstabe (a,b,c)$\} = \{1a, 1b, 1c, 2a, \ldots, 1c5a5a4b2c, \ldots\}$

Aufgabe 2.2 Grammatiken, Ableitungen und Syntaxbäume für D_4

Teilaufgabe 2.2.1 Eine Grammatik für die Dyck-Sprache D_4, 1 Punkt

Aus der Vorlesung ist Ihnen die Dyck-Sprache D_4 bekannt, sowie eine Grammatik G_4 mit $\mathscr{L}(G_4) = D_4$.

Geben Sie die Grammatik G_4, welche die Sprache D_4 (alle korrekt geklammerten Ausdrücke mit den Klammerpaaren (), [], { }, <>) erzeugt an.

Lösung

D_4 wird von der Grammatik G_4 erzeugt: $\mathscr{L}(G_4) = D_4$

- $G_4 = \{N, \Sigma, P, S\} = \{\{S\}, \{(,), [,], \{,\}, <, >\}, P, S\}$
- Die Produktionsmenge P besteht aus den Regeln:

$$\begin{array}{ccc} S & \rightarrow & \varepsilon \\ S & \rightarrow & SS \\ S & \rightarrow & [S] \\ S & \rightarrow & (S) \\ S & \rightarrow & \{S\} \\ S & \rightarrow & <S> \end{array}$$

Teilaufgabe 2.2.2 Ableitung des Wortes $[\,] < \{([\,])()\} >$, 3 Punkte

a) Geben Sie eine Linksableitung des Wortes $[\,] < \{([\,])()\} >$ an.

b) Geben Sie eine Rechtsableitung des Wortes $[\,] < \{([\,])()\} >$ an.

LÖSUNG

a) Linksableitung: in jedem Schritt wird das linkeste Nonterminal ersetzt:

$$\begin{aligned} S &\Rightarrow SS \\ &\Rightarrow [S]S \\ &\Rightarrow []S \\ &\Rightarrow [] < S > \\ &\Rightarrow [] < \{S\} > \\ &\Rightarrow [] < \{SS\} > \\ &\Rightarrow [] < \{(S)S\} > \\ &\Rightarrow [] < \{([S])S\} > \\ &\Rightarrow [] < \{([])S\} > \\ &\Rightarrow [] < \{([])(S)\} > \\ &\Rightarrow [] < \{([])()\} > \end{aligned}$$

b) Rechtsableitung: in jedem Schritt wird das rechteste Nonterminal ersetzt:

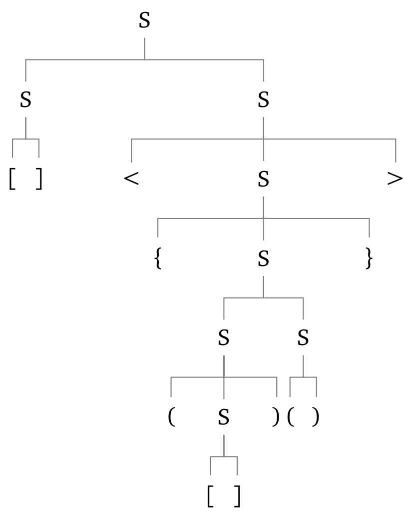
$$\begin{aligned} S &\Rightarrow SS \\ &\Rightarrow S < S > \\ &\Rightarrow S < \{S\} > \\ &\Rightarrow S < \{SS\} > \\ &\Rightarrow S < \{S(S)\} > \\ &\Rightarrow S < \{S()\} > \\ &\Rightarrow S < \{(S)()\} > \\ &\Rightarrow S < \{([S])()\} > \\ &\Rightarrow S < \{([])()\} > \\ &\Rightarrow [S] < \{([])()\} > \\ &\Rightarrow [] < \{([])()\} > \end{aligned}$$

TEILAUFGABE 2.2.3 SYNTAXBAUM ZUR ABLEITUNG DES WORTES $[] < \{([])()\} >$, 2 PUNKTE

a) Geben Sie für Ihre Linksableitung des Wortes $[] < \{([])()\} >$ den dazugehörigen Syntaxbaum an.

b) Geben Sie für Ihre Rechtsableitung des Wortes $[] < \{([])()\} >$ den dazugehörigen Syntaxbaum an.

LÖSUNG



Links- und Rechtsableitung führen zum selben Syntaxbaum. G_4 ist eindeutig:

AUFGABE 2.3 3 PUNKTE

Geben Sie für das Alphabet $\Sigma = \{1, 2, 3, 4, 5, a, b, c\}$ (siehe Aufgabe 2.1) folgende Grammatiken (Chomsky-Typ egal) an:

a) G_1 mit $\mathscr{L}(G_1) = \Sigma^2$; G_1 soll genau die Wörter der Länge 2 über Σ erzeugen

b) G_2 mit $\mathscr{L}(G_2) = \{1a, 1b, \ldots, 4c, 5c\}$; G_2 soll genau die 15 möglichen Kombinationen aus einer Zahl und einem Buchstaben (einstellige korrekte Hausnummer) erzeugen

c) G_3 mit $\mathscr{L}(G_3) = \{a1, a2, \ldots, c4, c5\}$; G_4 soll genau die 15 möglichen Kombinationen aus einem Buchstaben und einer Zahl (einstellige korrekte Gebäudenummer) erzeugen

d) G_4 mit $\mathscr{L}(G_4) = \{$ korrekt formulierte Hausnummern beliebiger Länge über Σ $\}$
Beispiele für korrekt formulierte Hausnummern beliebiger Länge
(die von G_4 erzeugt werden sollen):

$$1a, 5c, 12345b, 54c$$

Beispiele für nicht korrekt formulierte Hausnummern beliebiger Länge
(die von G_4 **nicht** erzeugt werden sollen):

$$1aa, c5c, a12c345b, abc, 11$$

e) G_5 mit $\mathscr{L}(G_5) = \mathscr{L}(G_2) \cup \mathscr{L}(G_3) = \{1a, 1b, \ldots, 4c, 5c, a1, a2, \ldots, c4, c5\}$
$=$ { korrekt formulierte einstellige Haus- oder Gebäudenummer über Σ }

LÖSUNG

Mögliche Lösungen:

a)
- $G_1 = (N_1, \Sigma, P_1, S)$
- $N_1 = \{S, T\}$
- $P_1 = \begin{array}{lcl} S & \rightarrow & 1T \mid 2T \mid 3T \mid 4T \mid 5T \mid aT \mid bT \mid cT \\ T & \rightarrow & 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid a \mid b \mid c \end{array}$

b)
- $G_2 = (N_2, \Sigma, P_2, S)$
- $N_2 = \{S, T\}$
- $P_2 = \begin{array}{lcl} S & \rightarrow & 1T \mid 2T \mid 3T \mid 4T \mid 5T \\ T & \rightarrow & a \mid b \mid c \end{array}$

c)
- $G_3 = (N_3, \Sigma, P_3, S)$
- $N_3 = \{S, T\}$
- $P_3 = \begin{array}{lcl} S & \rightarrow & aT \mid bT \mid cT \\ T & \rightarrow & 1 \mid 2 \mid 3 \mid 4 \mid 5 \end{array}$

d)
- $G_4 = (N_4, \Sigma, P_4, S)$
- $N_4 = \{S, T\}$
- $P_4 = \begin{array}{lcl} S & \rightarrow & 1T \mid 2T \mid 3T \mid 4T \mid 5T \\ T & \rightarrow & 1T \mid 2T \mid 3T \mid 4T \mid 5T \mid a \mid b \mid c \end{array}$

e)
- $G_5 = (N_5, \Sigma, P_5, S)$
- $N_5 = \{S, S_1, S_2, T_1, T_2\}$
- $P_5 = \begin{array}{lcl} S & \rightarrow & S_1 \mid S_2 \\ S_1 & \rightarrow & 1T_1 \mid 2T_1 \mid 3T_1 \mid 4T_1 \mid 5T_1 \\ T_1 & \rightarrow & a \mid b \mid c \\ S_2 & \rightarrow & aT_2 \mid bT_2 \mid cT_2 \\ T_2 & \rightarrow & 1 \mid 2 \mid 3 \mid 4 \mid 5 \end{array}$

Aufgabe 2.4 Die Chomsky-Hierarchie, 2 Punkte

Sei $N = \{A, B, C\}$ das Alphabet der Nonterminale, $\Sigma = \{1, 2, 3\}$ das Alphabet der Terminale über welchem verschiedene Grammatiken definiert sind. Im Folgenden ist aus jeder dieser Grammatiken eine Regel angegeben.

Geben Sie für jede der Regeln an, von welchem Chomsky-Typ sie (maximal) ist. Wenn also eine Regel vom Typ 0, 1 und 2 ist, dann ist die Lösung „Typ 2".

Begründen Sie Ihre Entscheidung.

a) $r_1 : B \to 1A$

b) $r_2 : 2CAB \to C3C$

c) $r_3 : C \to A$

d) $r_4 : AB \to 12$

e) $r_5 : C2A \to 23B$

f) $r_6 : 12 \to AB$

g) $r_7 : AB \to 1$

h) $r_8 : 2 \to 1$

i) $r_9 : B \to A1$

Lösung

a) $r_1 : B \to 1A$ Typ 3 (weil der Form $N \to \Sigma N$)

b) $r_2 : 2CAB \to C3C$ Typ 0 (weil der Form $l \to r$ mit $l > r$)

c) $r_3 : C \to A$ Typ 2 (weil der Form $N \to (\Sigma \cup N)^*$)

d) $r_4 : AB \to 12$ Typ 1 (weil der Form $l \to r$ mit $l \leq r$)

e) $r_5 : C2A \to 23B$ Typ 1 (weil der Form $l \to r$ mit $l \leq r$)

f) $r_6 : 12 \to AB$ keine gültige Regel (weil auf der linken Seite nur Terminale stehen)

g) $r_7 : AB \to 1$ Typ 0 (weil der Form $l \to r$ mit $l > r$)

h) $r_8 : 2 \to 1$ keine Regel, weil Terminal auf Terminal abgebildet wird

i) $r_9 : B \to A1$ Typ 2 (weil der Form $N \to (\Sigma \cup N)^*$); alternativ: Typ 3, aber nur wenn nur linkslineare Regeln verwendet werden

Aufgabe 2.5 Nutzung einer kontextfreien Grammatik

Gegeben sei die Grammatik $G_1 = (N, \Sigma, P, S) = (\{S\}, \{a, b\}, P, S)$ und

$$P = \quad S \quad \to \quad \varepsilon \mid aSb \mid SS$$

Teilaufgabe 2.5.1 2 Punkte

Nutzen Sie G_1, um aus dem Startsymbol folgende Worte abzuleiten:

a) $\omega_1 = \varepsilon$

b) $\omega_2 = ab$

c) $\omega_3 = abab$

d) $\omega_4 = aabbab$

Lösung

a) $S \Rightarrow \varepsilon$

b) $$\begin{aligned} S &\Rightarrow aSb \\ &\Rightarrow ab \end{aligned}$$

c) $$\begin{aligned} S &\Rightarrow SS \\ &\Rightarrow aSbS \\ &\Rightarrow abS \\ &\Rightarrow abaSb \\ &\Rightarrow abab \end{aligned}$$

d) $$\begin{aligned} S &\Rightarrow SS \\ &\Rightarrow aSbS \\ &\Rightarrow aaSbbS \\ &\Rightarrow aabbS \\ &\Rightarrow aabbaSb \\ &\Rightarrow aabbab \end{aligned}$$

Teilaufgabe 2.5.2 2 Punkte

Geben Sie jeweils den Syntaxbaum für Ihre Ableitung der Worte ω_3 und ω_4 an.

Lösung

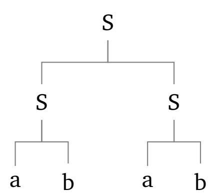


Abbildung 1: ω_3

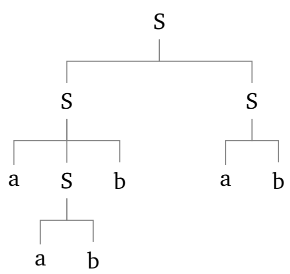


Abbildung 2: ω_4

Teilaufgabe 2.5.3 1 Punkt

Begründen Sie, weshalb man aus G_1 die folgenden Worte NICHT ableiten kann:

a) $\omega_5 = abc$

b) $\omega_6 = ba$

c) $\omega_7 = abbba$

Lösung

a) ω_5 enthält ein c, welches nicht in Σ enthalten ist.

b) ω_6 kann nicht abgeleitet werden, da keine Regel die Anordnung „ba“ erlaubt.

c) ω_7 kann nicht abgeleitet werden, da die Regeln nur eine ineinander verschachtelte Anordnung von as und bs erlauben.

Aufgabe 2.6 Nutzung einer kontextsensitiven Grammatik

Gegeben sei die Grammatik $G_2 = (N, \Sigma, P, S) = (\{S, A, B, C\}, \{a, b, c\}, P, S)$ und

$$P = \begin{array}{lcl} S & \rightarrow & \varepsilon \mid ABCS \\ CA & \rightarrow & AC \\ AC & \rightarrow & CA \\ BA & \rightarrow & AB \\ AB & \rightarrow & BA \\ CB & \rightarrow & BC \\ BC & \rightarrow & CB \\ A & \rightarrow & a \\ B & \rightarrow & b \\ C & \rightarrow & c \end{array}$$

Teilaufgabe 2.6.1 2 Punkte

Nutzen Sie G_2, um aus dem Startsymbol folgende Worte abzuleiten:

a) $\omega_8 = \varepsilon$

b) $\omega_9 = abc$

c) $\omega_{10} = bac$

d) $\omega_{11} = cbaabc$

Lösung

a) $S \Rightarrow \varepsilon$

b) $$\begin{array}{lcl} S & \Rightarrow & ABCS \\ & \Rightarrow & ABC \\ & \Rightarrow & aBC \\ & \Rightarrow & abC \\ & \Rightarrow & abc \end{array}$$

c) $$\begin{array}{lcl} S & \Rightarrow & ABCS \\ & \Rightarrow & ABC \\ & \Rightarrow & BAC \\ & \Rightarrow & bAC \\ & \Rightarrow & baC \\ & \Rightarrow & bac \end{array}$$

d)
$$\begin{aligned}
S &\Rightarrow ABCS\\
&\Rightarrow ABCABCS\\
&\Rightarrow ABCABC\\
&\Rightarrow BACABC\\
&\Rightarrow BCAABC\\
&\Rightarrow CBAABC\\
&\Rightarrow cBAABC\\
&\Rightarrow cbAABC\\
&\Rightarrow cbaABC\\
&\Rightarrow cbaaBC\\
&\Rightarrow cbaabC\\
&\Rightarrow cabaabc
\end{aligned}$$

Teilaufgabe 2.6.2 1 Punkt

Geben Sie die von G_2 erzeugte Sprache $\mathscr{L}(G_2)$ an, bzw. charakterisieren Sie die von G_2 erzeugten Worte so genau wie möglich.

Lösung

$L(G_2) = \{w \in \{a, b, c\}^* \mid w \text{ enthält gleichviele } a, b \text{ und } c\}$

Aufgabe 2.7 Die Sprache der ganzen Zahlen, 3 Punkte

$L_Z \subseteq \{-, 0, 1, \ldots, 9\}^*$ mit $L_Z = \{\ldots, -78562, -11, -10, \ldots -1, 0, 1, \ldots 9, 10, \ldots, 5906, \ldots, \}$ sei die Sprache der ganzen Zahlen.

a) Geben Sie eine Grammatik an, welche L_Z erzeugt.

b) Welchen Chomsky-Typ hat Ihre Grammatik?

c) Können Sie Ihre Grammatik so umformen, dass sie regulär ist?

d) Können Sie einen regulären Ausdruck angeben, welcher L_Z erzeugt?

Lösung

a) L_Z wird von der Grammatik G_Z erzeugt: $\mathscr{L}(G_Z) = L_Z$

- $G_Z = \{N, \Sigma, P, S\} = \{\{S, T, U\}, \{-, 0, 1, 2, \ldots, 9\}, P, S\}$
- $P = \begin{array}{lcl} S & \rightarrow & -T \mid 1 \mid 2 \mid \ldots \mid 9 \mid 1U \mid 2U \mid \ldots \mid 9U \\ T & \rightarrow & 1 \mid 2 \mid \ldots \mid 9 \mid 1U \mid 2U \mid \ldots \mid 9U \\ U & \rightarrow & 0 \mid 1 \mid \ldots \mid 9 \mid 0U \mid 1U \mid \ldots \mid 9U \end{array}$

b) G_N ist vom Typ 3 (regulär), da alle Regeln der Form $N \rightarrow \Sigma \mid \Sigma N$ sind.

c) Ist schon regulär.

d) $r_N = 0|(-?[1-9][0-9]^*)$

Aufgabe 2.8 Die OTTO-Zahlen, 3 Punkte

$L_O \subseteq L_N \subseteq \{0, 1, \ldots, 9\}^*$ mit $L_O = \{0, 1, \ldots 9, 11, 22 \ldots, 99, 101, 111, 121, \ldots, 573375, \ldots\}$, sei die Sprache der OTTO-Zahlen, also der natürlichen Zahlen, die von vorne und hinten gelesen gleich sind.

a) Geben Sie eine Grammatik an, welche L_O erzeugt.

b) Welchen Chomsky-Typ hat Ihre Grammatik?

c) Können Sie Ihre Grammatik so umformen, dass sie regulär ist?

Lösung

a) L_O wird von der Grammatik G_O erzeugt: $\mathscr{L}(G_O) = L_O$

- $G_O = \{N, \Sigma, P, S\} = \{\{S, S_2\}, \{0, 1, 2, \ldots, 9\}, P, S\}$
- Die Produktionsmenge P besteht aus (zwei) Regeln:

$$\begin{array}{rcl} S & \rightarrow & 0 \mid 1 \mid \ldots \mid 9 \mid 1S_21 \mid 2S_22 \mid \ldots \mid 9S_29 \\ S_2 & \rightarrow & 0 \mid 1 \mid \ldots \mid 9 \mid 0S_20 \mid 1S_21 \mid \ldots \mid 9S_29 \mid \varepsilon \end{array}$$

b) G_O ist vom Typ 2 (kontextfrei), da alle Regeln der Form $l \rightarrow r$, mit $l \leq r$ und $l \in N$ sind. Erlaubte Ausnahme ist die Regel $S_2 \rightarrow \varepsilon$.

c) Nein, da G_O vom Typ 2 ist. Mittels Pumping-Lemma lässt sich auch nachweisen, dass L_O nicht regulär ist, somit kann G_O nicht zu einer regulären Grammatik umgeformt werden.
Alternativ kann man sich auch überlegen, dass man für die Otto-Zahlen keinen regulären Ausdruck finden kann.