

Hinweise:

- Die **Hausaufgaben** sollen in Gruppen von je 2 Studierenden aus dem gleichen Tutorium bearbeitet werden.
- Die Lösungen der Hausaufgaben müssen bis Mi., 16.06.2010 im Tutorium abgegeben werden. Alternativ ist es bis 17 Uhr möglich, diese in den Kasten im Flur des LuFG I2 einzuwerfen (Ahornstr. 55, E1, 2. Etage).
- Namen und Matrikelnummern der Studierenden sowie **die Nummer der Übungsgruppe** sind auf jedes Blatt der Abgabe zu schreiben. **Heften bzw. tackern Sie die Blätter!**
- Da am Mi., 9.6.2010 der Studieninformationstag stattfindet, fallen die Tutorien an diesem Tag aus.
- Die **Tutoraufgaben** werden in der kommenden Globalübung am 11.06.2010 vorgerechnet.
- Die **Einsicht** in die Präsenzübungs-Ergebnisse wird am **11.06.2010 zwischen 8:00 und 10:00 in 2356|056** (Ahornstraße 55, Raum 5056) stattfinden. Studenten mit einer Matrikelnummer kleiner als 294000 werden gebeten, zwischen 8:00 und 9:00 zu kommen, alle anderen zwischen 9:00 und 10:00.
- Die **Rückgabe** der korrigierten Präsenzübungen erfolgt in der Einsicht und in den Tutorien am 16.6.2010. Nach der Rückgabe sind **keine Eingaben zur Korrektur möglich!**

Hausaufgabe 1 (Eindeutigkeit):
(1 + 2 + 2 = 5 Punkte)

 Sei $G = (N, T, P, S)$ eine Grammatik mit $N = \{S, A, B\}$, $T = \{a, b\}$ und P wie folgt:

$$S \rightarrow AB \mid SAB$$

$$A \rightarrow Aa \mid \epsilon$$

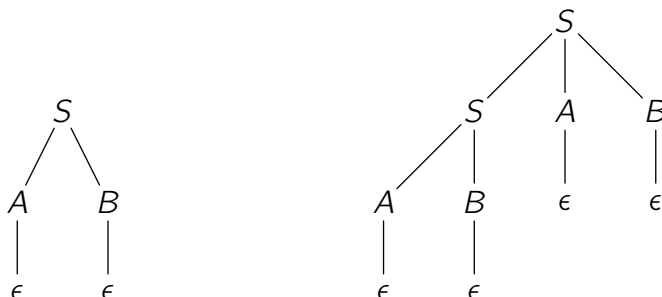
$$B \rightarrow Bb \mid \epsilon$$

- Geben Sie die von G erzeugte Sprache $L(G)$ an (ohne Beweis).
- Beweisen Sie, dass G nicht eindeutig ist.
- Geben Sie eine eindeutige kontextfreie Grammatik G' an mit $L(G') = L(G)$ (ohne Beweis).

Lösung:

- Die von G erzeugte Sprache ist $L(G) = L((a + b)^*)$.

- Bereits das Wort ϵ lässt sich mit verschiedenen Ableitungsbäumen erzeugen:



Damit ist bewiesen, dass G nicht eindeutig ist.

- c) Wir definieren $G' := (\{S\}, T, P', S)$ mit P' wie folgt:

$$S \rightarrow aS \mid bS \mid \epsilon$$

Hausaufgabe 2 (Grammatiken und reguläre Sprachen): (2 + 4 + 2 = 8 Punkte)

Sei $G = (N, T, P, S)$ eine Grammatik. Wenn für G gilt, dass

$$\forall (A \rightarrow \alpha) \in P : (A \in N) \wedge (\alpha \in \{\epsilon\} \cup T \cup T \cdot N)$$

dann nennt man G **linkslinere** Grammatik. In einer solchen Grammatik sind also alle Produktionen so gestaltet, dass sie entweder ein Nonterminalsymbol löschen, es in ein Terminalsymbol oder in ein Terminalsymbol und ein weiteres Nonterminalsymbol umwandeln.

- a) Betrachten Sie die linkslinere Grammatik $G := (N, T, P, S)$ mit $N := \{S, A, B\}$, $T := \{a, b\}$ und P wie folgt:

$$S \rightarrow aA \mid bB$$

$$A \rightarrow aA \mid \epsilon$$

$$B \rightarrow aB \mid bA$$

Geben Sie Ableitungsbäume für die Worte $w_1 = aa$ und $w_2 = baaba$ an.

- b) Nehmen Sie an, dass $G = (N, T, P, S)$ eine beliebige linkslinere Grammatik ist. Zur Vereinfachung fordern wir zusätzlich, dass P keine Regeln der Form $A \rightarrow b$ enthält, sondern alle Regeln die Form $A \rightarrow \epsilon$ oder $A \rightarrow bC$ für beliebige $A, C \in N, b \in T$ haben.¹

Definieren Sie nun einen NFA M_G , für den $L(M_G) = L(G)$ gilt. Sie brauchen diese Gleichheit nicht zu beweisen.

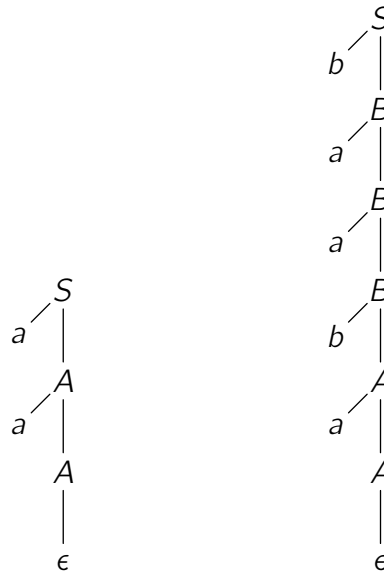
Hinweis: Sie sollten als Zustände für M_G die Nonterminalsymbole N wählen.

- c) Sei $M = (Q, \Sigma, \delta, q_0, F)$ ein NFA. Geben Sie eine linkslinere Grammatik G_M an, so dass $L(G_M) = L(M)$. Sie brauchen diese Gleichheit nicht zu beweisen.

Lösung:

- a) Die Ableitungsbäume können wie folgt dargestellt werden:

¹Dies ist keine wesentliche Einschränkung, denn mit Hilfe eines zusätzlichen Nonterminalsymbols E mit der Produktion $E \rightarrow \epsilon$ können Regeln der Form $A \rightarrow b$ in $A \rightarrow bE$ umgewandelt werden. Damit kann die selbe Sprache wie mit Hilfe der ursprünglichen Grammatik erkannt werden.



b) Wir definieren $M_G := (N, T, \delta, S, F)$ mit

$$\delta(A, b) = \{C \mid (A \rightarrow bC) \in P\} \quad A, C \in N, b \in T$$

und $F = \{A \mid (A \rightarrow \epsilon) \in P\}$.

c) Wir definieren $G_M := (Q, \Sigma, P, q_0)$ mit

$$\begin{aligned}
 P := & \{q \rightarrow aq' \mid q, q' \in Q, a \in \Sigma, q' \in \delta(q, a)\} \\
 & \cup \{q \rightarrow \epsilon \mid q \in F\}
 \end{aligned}$$

Tutoraufgabe 3 (Kontextfreie Grammatiken):

Die Dyck-Sprache ist die Sprache aller korrekt geklammerten Ausdrücke. Z.B. sind "()", "(())" und "()()" korrekt geklammerte Ausdrücke, während ")(" und "((" keine korrekt geklammerten Ausdrücke sind. Zur Vereinfachung der Lesbarkeit folgender Definitionen verwenden wir die Zeichen a und b anstelle von (und). Formal ist die Dyck-Sprache über dem Alphabet $\Sigma = \{a, b\}$ dann folgendermaßen definiert:

$D = \{w \in \Sigma^* \mid \#_a(w) = \#_b(w) \wedge \forall u, v \in \Sigma^* : u \cdot v = w \Rightarrow \#_a(u) \geq \#_b(u)\}$, wobei die Funktionen $\#_a$ und $\#_b$ die Anzahl der vorkommenden Zeichen a bzw. b in einem Wort zählen. Formal:

$\#_a : \Sigma^* \rightarrow \mathbb{N}_0$ mit $\#_a(\epsilon) = 0$, $\#_a(w \cdot a) = \#_a(w) + 1$ und $\#_a(w \cdot b) = \#_a(w)$. Die Funktion $\#_b$ ist analog definiert.

Geben Sie eine kontextfreie Grammatik an, die D erzeugt (Sie brauchen nicht zu beweisen, dass diese Grammatik D erzeugt).

Lösung: _____

Wie in der Vorlesung geben wir nur die Produktionen an. Das Startsymbol ist S .

$$\begin{aligned}
 S & \rightarrow SS \\
 S & \rightarrow aSb \\
 S & \rightarrow \epsilon
 \end{aligned}$$

Alternative Lösung:

$$\begin{aligned} S &\rightarrow aSbS \\ S &\rightarrow \epsilon \end{aligned}$$

Hausaufgabe 4 (Kontextfreie Grammatiken):

(2 + 2 + 3 = 7 Punkte)

Geben Sie jeweils eine kontextfreie Grammatik an, welche die folgenden Sprachen erzeugt (Sie brauchen nicht zu beweisen, dass Ihre Grammatiken die geforderten Sprachen erzeugen).

- a) $\{a^n b^n \mid n \geq 0\}$
- b) $\{ab^n c \mid n > 0\}$
- c) $\{w \in \{a, b, c\}^* \mid \#_a(w) = \#_b(w)\}$

Lösung:

Wie in der Vorlesung geben wir nur die Produktionen an. Das Startsymbol ist jeweils S .

a)

$$\begin{aligned} S &\rightarrow aSb \\ S &\rightarrow \epsilon \end{aligned}$$

b)

$$\begin{aligned} S &\rightarrow aBc \\ B &\rightarrow bB \\ B &\rightarrow b \end{aligned}$$

Alternativ als linkslineare Grammatik:

$$\begin{aligned} S &\rightarrow aA \\ A &\rightarrow bB \\ B &\rightarrow bB \\ B &\rightarrow c \end{aligned}$$

c)

$$\begin{aligned} S &\rightarrow SaSbS \\ S &\rightarrow SbSaS \\ S &\rightarrow cS \\ S &\rightarrow \epsilon \end{aligned}$$

Alternative Lösung mit zusätzlichem Nonterminal C:

$$\begin{aligned} S &\rightarrow SaCbC \\ S &\rightarrow SbCaC \\ S &\rightarrow CaSbC \\ S &\rightarrow CbSaC \\ S &\rightarrow CaCbS \\ S &\rightarrow CbCaS \\ S &\rightarrow C \\ C &\rightarrow cC \\ C &\rightarrow \epsilon \end{aligned}$$

Hausaufgabe 5 (Induktion):

(6 Punkte)

Betrachten Sie folgende kontextfreie Grammatik $G = (N, T, P, S)$ mit $N = \{S\}$, $T = \{a, b\}$ und $P = \{$

$$\begin{aligned} S &\rightarrow abS, \\ S &\rightarrow \epsilon \end{aligned}$$

$\}$
und die Sprache $L = \{(ab)^n \mid n \geq 0\}$.

Beweisen Sie, dass $L(G) = L$ gilt.

Hinweis: Zeigen Sie die Inklusionen $L(G) \subseteq L$ und $L \subseteq L(G)$ separat. Verwenden Sie für die erste Inklusion eine Induktion über die Ableitungslänge der von G erzeugten Wörter und für die zweite Inklusion eine Induktion über die Anzahl von ab Teilworten in $w \in L$.

Lösung:

Wir beweisen zunächst $L(G) \subseteq L$. Sei $w \in L(G)$. Wir zeigen $w \in L$ per Induktion über die Ableitungslänge von \Rightarrow_G .

Im Induktionsanfang ist die Ableitungslänge 1 und wir haben $S \Rightarrow_G w$. Das einzige Terminalwort, das in einem Schritt von S erzeugt werden kann, ist $\epsilon \in L$.

Als Induktionshypothese setzen wir voraus, dass die Aussage für alle Ableitungslängen $m' < m$ gilt.

Im Induktionsschritt gilt nun $S \Rightarrow_G^m w$ mit $m > 1$. Da $m > 1$ gilt, kann der erste Ableitungsschritt nicht die Produktion $S \rightarrow \epsilon$ verwendet haben, da danach kein Nonterminal mehr in der resultierenden Satzform enthalten und die Ableitung beendet wäre. Also lautet der erste Ableitungsschritt $S \Rightarrow_G abS$ und wir haben $abS \Rightarrow_G^{m-1} w$. Da ein Ableitungsschritt einer kontextfreien Grammatik keine Terminale entfernen kann, gilt außerdem $w = abw'$ und $S \Rightarrow_G^{m-1} w'$. Aus der Induktionshypothese folgt nun, dass $w' \in L$ gilt. Damit gilt dann aber auch $abw' = w \in L$.

Nun beweisen wir $L \subseteq L(G)$. Sei $w = (ab)^n \in L$ für ein $n \in \mathbb{N}_0$. Wir zeigen $w \in L(G)$ per Induktion über n .

Im Induktionsanfang ist $n = 0$ und wir haben $w = \epsilon$. Es gilt $S \Rightarrow_G \epsilon$ und damit $\epsilon \in L(G)$.

Als Induktionshypothese nehmen wir an, dass die Aussage für alle $n' < n$ gilt.

Im Induktionsschritt ist nun $w = (ab)^n = ab(ab)^{n-1}$. Es gilt $S \Rightarrow_G abS$. Aus der Induktionshypothese folgt außerdem, dass $(ab)^{n-1} \in L(G)$, d.h. $S \Rightarrow_G^* (ab)^{n-1}$. Zusammen folgt daraus, dass $S \Rightarrow_G abS \Rightarrow_G^* ab(ab)^{n-1} = w$ und damit $w \in L(G)$.

Aus $L(G) \subseteq L$ und $L \subseteq L(G)$ folgt schließlich $L(G) = L$.
