Aufgaben zum Tut am 16.01.2006

Thomas Pajor

23. Januar 2006

Aufgabe 1)

Gegeben sei folgende Grammatik $G := (\{S\}, \{a, b\}, P, S)$ mit

$$
\begin{aligned}
P := \{S &\to aSa, &(1)\\
S &\to bSb, &(2)\\
S &\to aa, &(3)\\
S &\to bb\} &(4)
\end{aligned}
$$

(a) Bestimmen Sie den (maximalen) Chomsky–Typ der Grammatik G.

(b) Wandeln Sie G in Chomsky–Normalform G'.

(c) Bestimmen Sie die durch G erzeugte Sprache $L(G)$ und beweisen Sie ihre Behauptung.

Lösung.

(a) Gleich wegen der ersten Ableitungsregel kann die Grammatik nicht vom Typ CH–3 sein. Da es aber keine Ableitungsregel gibt, die im Widerpsruch zu CH–2 steht, ist G vom Typ CH–2.

(b) Wir führen das Verfahren zur Überführung in CNF durch.

1. Wir führen für jedes $x_i \in \Sigma$ eine neue Variable X_i in V ein, und ersetzen in jeder Produktion x_i durch X_i. Schließlich fügen wir noch $|\Sigma|$ neue Produktion $X_i \to x_i$ für jedes $x_i \in \Sigma$ ein. Dies führt uns zu der Grammatik $G_1 = (\{S, A, B\}, \{a, b\}, P_1, S)$ wobei

$$
\begin{aligned}
P_1 := \{S &\to ASA \mid BSB \mid AA \mid BB,\\
A &\to a,\\
B &\to b\}
\end{aligned}
$$

2. Wir verkürzen alle Produktionen $l \to r$ mit $|r| > 2$. Daraus resultiert die Grammatik $G_2 := (\{S, A, B, X, Y\}, \{a, b\}, P_2, S)$ mit

$$\begin{aligned} P_2 := \{ & S \to AX \mid BY \mid AA \mid BB, \\ & A \to a, \\ & B \to b, \\ & X \to SA, \\ & Y \to SB \} \end{aligned}$$

3. Wir eleminieren alle Kettenregeln (und Kreise). Da es in P_2 keine Kettenregeln gibt, ist hier nichts mehr zu tun.

Es gilt $G' := G_2$ ist die in Chomsky–Normalform äquivalente Grammatik zu G.

(c) **Der Beweis ist etwas anders als im Tutorium, da wir einige Fehler gefunden haben.**

Behauptung: Die durch G erzeugte Sprache ist $L = \{ww^R \mid w \in \Sigma^+\}$, also die Sprache aller Palindrome gerader Länge über Σ. Wir beweisen $L = L(G)$ in zwei Schritten:

1. $L \subseteq L(G)$:
   Wir führen Induktion über die Wortlänge der Wörter in L. Das kürzeste Wort in L hat offensichtlich Länge 2, also

   IA: Sei $w \in L$ mit $|w| = 2$. Dann ist $w = aa$ oder $w = bb$. Diese werden durch G erzeugt in dem man beginnend vom Startsymbol S die dritte oder vierte Ableitungsregel anwendet.

   IV: Die Behauptung gelte für ein beliebiges n (n gerade).

   IS: $n \rightsquigarrow n + 2$:
   Sei $w \in L$ mit $|w| = n + 2$. Dann lässt sich w schreiben als $w = xvx$ mit $v \in L$ und $x \in \Sigma$.
   Da $|v| = n$ ist nach IV $v \in L(G)$, es existiert also eine Folge von Produktionen mit $S \overset{*}{\Rightarrow} v$. Je nach x fügen wir nun die erste bzw. zweite Produktion an den Anfang der Ableitungskette an, und erhalten somit $S \Rightarrow xSx \overset{*}{\Rightarrow} xvx = w$, woraus $w \in L(G)$ folgt.

2. $L(G) \subseteq L$:
   Wir führen Induktion über die Anzahl n der Ableitungsregeln, die Wörter w mit $w \in \Sigma^*$ erzeugen[1].

   IA: Sei $n = 1$, es kommen also nur die Ableitungsfolgen $S \to aa$ bzw. $S \to bb$ in Frage. Sowohl aa alsauch bb sind in L enthalten.

   IV: Die Behauptung gelte für ein beliebiges n.

   IS: $n \rightsquigarrow n + 1$:
   Betrachte die Folge von Ableitungsregeln $S \Rightarrow xSx \overset{*}{\Rightarrow} xvx =: w$. Nach Induktionsvoraussetzung ist $v \in L$, da v nach n Ableitungen produziert wird. Dann ist aber auch $xvx \in L$.

$\Rightarrow L = L(G)$. $\square$

[1] „Unfertige" Ableitungsfolgen, bei denen das Wort noch Variablen enthält, wollen wir außer acht lassen.

Aufgabe 2)

Sei $G := (V, \Sigma, P, S)$ eine kontextfreie Grammatik und seien induktiv die folgenden Mengen definiert:

$$\begin{aligned} X_0 &:= \Sigma \\ X_{n+1} &:= X_n \cup \{A \mid A \in V, \exists z \in X_n^* : A \to z\} \end{aligned}$$

(a) Zeigen Sie: $X_i \subseteq X_{i+1}$ für alle i und es gibt ein $k \in \mathbb{N}$ mit $X_k = X_{k+1}$ und falls $X_k = X_{k+1}$ so ist auch $X_k = X_{k+r}$ für alle $r \in \mathbb{N}$.

(b) Zu jedem $A \in V$ sei

$$L(A) = \{z \mid z \in \Sigma^*, A \stackrel{*}{\Rightarrow} z\}$$

Zeigen Sie, dass $L(A) \neq \emptyset$ genau dann, wenn es ein $n \in \mathbb{N}$ gibt, mit $A \in X_n$.

(c) Die Definition der X_i lassen sich als Verfahren benutzen um zu prüfen ob die durch eine kontextfreie Grammatik G erzeugte Sprache leer ist. Geben Sie den worst–case Aufwand asymptotisch im $\mathcal{O}$ Kalkül an.

Lösung.

(a) $X_i \subseteq X_{i+1}$ folgt unmittelbar aus der Definition. Da $|V| < \infty$ kann es keine unendliche *echt* aufsteigende Kette geben, somit existiert ein $k \in \mathbb{N}$ mit $X_k = X_{k+1}$.

Dass nun $X_k = X_{k+r}$ für alle $r \in \mathbb{N}$ gilt, beweisen wir durch vollständige Induktion über r.

IA: $r = 1$: Dies ist die Voraussetzung.

IV: Die Behauptung gelte für ein beliebiges aber festes r.

IS: $r \rightsquigarrow r + 1$: Wir rechnen nach:

$$\begin{aligned} X_{k+r+1} &= X_{k+r} \cup \{A \mid A \in V \text{ und es gibt } z \in X_{k+r}^* \text{ mit } A \to z\} \\ &\stackrel{\text{IV}}{=} X_k \cup \{A \mid A \in V \text{ und es gibt } z \in X_k^* \text{ mit } A \to z\} \\ &= X_{k+1} \\ &= X_k \end{aligned}$$

□

(b) Wir zeigen die Behauptung in zwei Schritten.

„$\Leftarrow$“: Wir zeigen $A \in X_n \Rightarrow L(A) \neq \emptyset$ durch Induktion über n:

IA: $n = 1$. Sei $A \in X_1$, dann gibt es nach Definition ein $z \in X_0^* = \Sigma^*$ mit $A \Rightarrow z$, somit ist $z \in L(A)$ und $L(A) \neq \emptyset$.

IV: Die Behauptung gelte für ein festes n.

IS: $n \rightsquigarrow n+1$. Sei $A \in X_{n+1}$, dann ist ist $A \in X_n$ oder es gibt ein $z \in X_n^*$ mit $A \Rightarrow z$.
Ist $A \in X_n$, so folgt $L(A) \neq 0$ aus der IV.
Im anderen Fall betrachten wir jede Variable B die in z enthalten ist. Weil $z \in X_n^*$ ist, ist auch $B \in X_n^*$. Aus der IV folgt nun $L(B) \neq \emptyset$, also insbesondere existiert ein z_B wobei $B \Rightarrow z_B$ mit $z_B \in \Sigma^*$. Also existiert eine Kette von Ableitungen

$$A \Rightarrow z \overset{*}{\Rightarrow} z_A \qquad z_A \in \Sigma^*$$

und somit ist $L(A) \neq \emptyset$.

„$\Rightarrow$“: Sei $L(A) \neq \emptyset$, also gibt es ein Wort $z \in \Sigma^*$ mit $A \overset{*}{\Rightarrow} z$, also eine Ableitungskette der Form

$$A =: z_n \Rightarrow z_{n-1} \Rightarrow \ldots \Rightarrow z_0 := z$$

Wir zeigen jetzt durch Induktion über i dass für jede Variable $B \in V$ gilt: Kommt B in z_i vor, dann gilt $B \in X_i$, also $z_i \in X_i^*$.

IA: Für $i = 0$ ist $z_0 = z \in \Sigma^* = X_0^*$.

IV: Behauptung gelte für festes n.

IS: $n \rightsquigarrow n+1$.
Sei B also in z_{i+1}. Der Ableitungsschritt $z_{i+1} \Rightarrow z_i$ wird durch eine Produktion $Z \Rightarrow R$ erzeugt. Dann gilt $z_{i+1} = uZv$ und $z_i = uRv$. Ist B auch in z_i, so folgt die Behauptung mit der IV.
Sei nun B nicht in z_i, dann wurde B durch die Produktion $Z \Rightarrow R$ gelöscht, also gilt $Z = B$ und nach Konstruktion von X_{i+1} ist $B \in X_{i+1}$.

Also folgt aus $L(A) \neq \emptyset$, dass ein $n \in \mathbb{N}$ existiert mit $A \in X_n$.

$\square$

(c) Sei $|V| =: k$, dann endet nach Aufgabe (a) das Verfahren spätestens nach Konstruktion von X_k. Im schlimmsten Fall muss man in jedem Schritt jede Produktion testen, also ist der Aufwand des Verfahrens in $\mathcal{O}(|P| \cdot |V|)$.