Einführung in die Computerlinguistik
Hausaufgabe CFG, Abgabe Montag 31.05.2021

Laura Kallmeyer

Sommer 2021, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 *Betrachten Sie folgende CFG:*
$G = \langle \{S, B, C\}, \{b, c\}, \{S \to B \mid SC, B \to b \mid bS, C \to c\}, S \rangle$

1. *Geben Sie alle Linksableitungen und Rechtsableitungen für $w = bbc$ an.*
2. *Geben Sie alle Parsbäume für $w = bbc$ an.*
3. *Ist die Grammatik mehrdeutig (ambig)?*
4. *Welche Sprache wird von G erzeugt?*

Lösung:

1. Linksableitungen: $S \Rightarrow SC \Rightarrow BC \Rightarrow bSC \Rightarrow bBC \Rightarrow bbC \Rightarrow bbc$
   $S \Rightarrow B \Rightarrow bS \Rightarrow bSC \Rightarrow bBC \Rightarrow bbC \Rightarrow bbc$

   Rechtsableitungen: $S \Rightarrow SC \Rightarrow Sc \Rightarrow Bc \Rightarrow bSc \Rightarrow bBc \Rightarrow bbc$
   $S \Rightarrow B \Rightarrow bS \Rightarrow bSC \Rightarrow bSc \to bBc \Rightarrow bbc$

2.

3. Ja, da es Wörter mit mehr als einem Parsbaum (s.2.) gibt.
4. $\{b^n c^m \mid n \geq 1, m \geq 0\}$
   oder (alternativ) die Sprache, die von dem regulären Ausdruck b^+c^* denotiert wird.

Aufgabe 2 *Betrachten Sie nun die folgenden CFGs:*

1. $G = \langle \{S\}, \{a, b, c\}, \{S \to aSbb \mid c\}, S \rangle$
2. $G = \langle \{S\}, \{a, b, c\}, \{S \to aS \mid aSb \mid aSc \mid \varepsilon\}, S \rangle$
3. $G = \langle \{S, A, B\}, \{a, b\}, \{S \to AB, A \to Aa \mid \varepsilon, B \to cBb \mid cb\}, S \rangle$
4. $G = \langle \{S\}, \{a, b, c\}, \{S \to aScc \mid aSb \mid \varepsilon\}, S \rangle$

Welche Sprachen werden jeweils von diesen Grammatiken generiert?

Lösung:

1. $\{a^n c b^{2n} \mid n \geq 0\}$
2. $\{a^n w \mid n \geq 0, |w| \leq n, w \in \{b, c\}^*\}$
3. $\{a^n b^m c^m \mid n \geq 0, m > 0\}$
4. $\{a^n w \mid n \geq 0, w \in \{b, c\}^*, |w|_b + \frac{1}{2} \cdot |w|_c = n\}$

Aufgabe 3 *Geben Sie zu den folgenden Sprachen jeweils eine kontextfreie Grammatik an, die die Sprache generiert.*
(1) $L = \{ab^n c^n d | n \geq 0\}$ *(2)* $L = \{a^n (bc)^n | n > 0\}$ *(3)* $L = \{wa^n b^m w^R | w \in \{c, d\}^+, n \geq m > 0\}$[1]

Lösung:
(Hier nur die Produktionen, eigentlich muss man jedoch auch N, T und das Startsymbol angeben.)

1. $S \to aTd, T \to bTc \mid \varepsilon$
2. $S \to aSbc \mid abc$
3. $S \to cSc \mid dSd \mid cTc \mid dTd, T \to aT \mid aTb \mid ab$

[1] $w^R = w$ in umgekehrter Reihenfolge.