Einführung in die Computerlinguistik
Hausaufgabe (CFG 2), Abgabe 18.06.2018

Laura Kallmeyer

Sommer 2018, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1

1. *Betrachten Sie folgende CFG:*
   $G = \langle \{S, A, B, C, D, E\}, \{a, b, c, e\}, P, S\rangle$ *mit*
   $P = \{S \to AB \,|\, CD, A \to aC \,|\, a, B \to b, C \to cC \,|\, CE, D \to \varepsilon, E \to eD\}$
   Geben Sie eine äquivalente Grammatik ohne nutzlose Symbole an.
2. $G = \langle \{S, T, U\}, \{a, b, c\}, \{S \to aSb \,|\, aTb \,|\, aU, T \to c \,|\, Sb \,|\, UU, U \to ab \,|\, \varepsilon\}, S\rangle$
   Geben Sie eine äquivalente Grammatik ohne ε-Produktionen an.

Lösung:

1. Symbole, aus denen sich terminale Ketten ableiten lassen: $\{A, B, D, S, E\}$
   neue Produktionsmenge ohne C: $S \to AB, A \to a, B \to b, D \to \varepsilon, E \to eD$
   Symbole, die vom Startsymbol erreichbar sind: $\{A, B, a, b\}$
   neue Produktionsmenge $S \to AB, A \to a, B \to b$
   Grammatik: $\langle \{A, B, S\}, \{a, b\}, \{S \to AB, A \to a, B \to b\}, S\rangle$
2. $N_\varepsilon = \{U, T\}$.
   Neue Produktionen: $S \to aSb \,|\, aTb \,|\, aU \,|\, ab \,|\, a, T \to c \,|\, Sb \,|\, UU \,|\, U, U \to ab$

Aufgabe 2 *Geben Sie für die Sprache $\{ww^R \,|\, w \in \{a, b\}^+\}$ (w^R ist w in umgekehrter Reihenfolge)*

1. *eine CFG in Chomsky Normalform und*
2. *eine CFG in Greibach Normalform an.*

Lösung: (hier nur Produktionen, es sollte aber eigentlich das ganze Tupel hingeschrieben werden)

1. $S \to AT \,|\, BU \,|\, AA \,|\, BB, \quad T \to SA, \quad U \to SB, \quad A \to a, B \to b$
2. $S \to aSA \,|\, bSB \,|\, aA \,|\, bB, \quad A \to a, \quad B \to b$

Aufgabe 3 *Betrachten Sie folgende CFG:*
$G = \langle \{S, T\}, \{a, b\}, \{S \to aTbT \,|\, ab, T \to aSa \,|\, bSb\}, S\rangle$
Geben Sie eine äquivalente Grammatik in Chomsky Normalform an.

Lösung:

Einführen von Präterminalen:

$S \to C_aTC_bT \,|\, C_aC_b, T \to C_aSC_a \,|\, C_bSC_b, C_a \to a, C_b \to b$

Binarisierung:

$S \to C_aX_1 \,|\, C_aC_b, X_1 \to TX_2, X_2 \to C_bT,$

$T \to C_aX_3 \,|\, C_bX_4, X_3 \to SC_a, X_4 \to SC_b,$

$C_a \to a, C_b \to b$

Aufgabe 4 *Zeigen Sie, dass die Sprache* $\{a^n e^m b^n d^m c^n \mid n \geq 1, m \geq 0\}$ *nicht kontextfrei ist.*

Tipp: Sie dürfen als bewiesen voraussetzen, dass $\{a^n b^n c^n \mid n \geq 1\}$ *nicht kontextfrei ist. Nutzen Sie die Abgeschlossenheit von kontextfreien Sprachen unter Schnittbildung mit regulären Sprachen.*

Lösung:

Wir nehmen an, $L = \{a^n e^m b^n d^m c^n \mid n \geq 1, m \geq 0\}$ ist kontextfrei. Dann muss der Schnitt von L mit der regulären Sprache $L(a^* b^* c^*)$ auch kontextfrei sein. Dieser Schnitt ist aber $\{a^n b^n c^n \mid n \geq 1\}$, wovon wir schon wissen, dass es nicht kontextfrei ist.

Widerspruch, also ist unsere Annahme falsch und L ist nicht kontextfrei.