Parsing Beyond CFG
Homework 1: CFG

Laura Kallmeyer

Winter 2021/22

Question 1 *Give a CFG G_1 for the following language L_1:*

$$L_1 = \{w \in \{a, b, c\}^* \mid w = a^{2n} c^m (ab)^n \text{ where } n \geq 1, m \geq 0\}$$

Solution:

$G = \langle\{S, A, B, C\}, \{a, b, c\}, \{S \to ASB \mid ACB, A \to aa, B \to ab, C \to cC \mid \varepsilon\}, S\rangle$

Question 2 *Betrachten Sie folgende CFGs:*

1. $G_1 = \langle\{S, A, B\}, \{a, b, c\}, \{S \to AB, A \to aAb \mid \epsilon, B \to cB \mid \epsilon\}, S\rangle$
2. $G_2 = \langle\{S, A, B, D\}, \{a, b, d\}, \{S \to ASB \mid D, A \to a \mid aa, B \to b, D \to dD \mid d\}, S\rangle$

Welche Sprachen werden von den beiden Grammatiken erzeugt?

Solution:

1. $\{a^m b^m c^n | m, n \geq 0\}$
2. $\{a^n d^k b^m \mid 0 \leq m \leq n \leq 2m, k \geq 1\}$

Question 3 *Show that the following language is not context-free:*

$$L = \{a^n b^m a^n b^m \mid n, m \geq 0\}$$

Hint: Assume that L is context-free. Then it must satisfy the pumping lemma. Show that this is not the case. Consequently, the initial assumption does not hold.

Solution:

We assume that L is context-free. Then it satisfies the pumping lemma with a certain constant $k \geq 1$.

We consider the word $a^k b^k a^k b^k$. According to the pumping lemma this word must have a form xv_1yv_2z with $|v_1yv_2| \leq k$ such that v_1 and v_2 can be iterated. Because of $|v_1yv_2| \leq k$, v_1v_2 cannot contain elements from both groups of as or from both groups of bs. With this and with $v_1v_2 \neq \varepsilon$, we have that after the first pumping, we obtain a word $xv_1^2yv_2^2z$ for which either the numbers of as in the two groups differ or the numbers of bs in the two b groups. This word is therefore not in L. Contradiction to the pumping lemma. Consequently, L does not satisfy the pumping lemma and is therefore not a CFL.

Question 4 *Nehmen Sie an, dass wir schon wissen, dass $L_3 = \{a^n b^n c^n \mid n \geq 1\}$ keine kontextfreie Sprache ist.*

Zeigen Sie, dass $L = \{d^m b^n c^n d^k a^n \mid n \geq 0, m \geq 0, k \geq 0\}$ dann auch keine kontextfreie Sprache ist.

Tipp: Nutzen Sie sowohl die Abgeschlossenheit kontextfreier Sprachen unter Homomorphismen als auch die Abgeschlossenheit unter Schnittbildung mit regulären Sprachen aus.

Lösung:

Wir nehmen an, dass L kontextfrei ist.

Dann muss auch $L' = h(L)$ mit h eine Homomorphismus, $h(a) = c, h(b) = a, h(c) = b$, $h(d) = \varepsilon$ kontextfrei sein. $L' = \{a^n b^n c^n \mid n \geq 0\}$.

Wenn L' kontextfrei ist, dann muss auch $L'' = L' \cap L(a^+b^+c^+)$ kontextfrei sein.

Es gilt aber $L'' = L_3$, und von L_3 wissen wir, dass es nicht kontextfrei ist. Damit haben wir unsere Annahme zu einem Widerspruch geführt, und somit ist sie falsch. Es gilt also, dass L nicht kontextfrei ist.