Parsing

Homework 1 (CFG), due 27 April 2020

Laura Kallmeyer

SS 2020, Heinrich-Heine-Universität Düsseldorf

Question 1 (Grammars)

Consider the following three languages:

- $L_1 = \{a^k b^n (cd)^m \mid k > 0, n, m \geq 0\}$
- $L_2 = \{(ab)^n c d^m e^{n-1} \mid n, m > 0\}$
- $L_3 = \{a^n b (cd)^{n-1} e^{n+1} \mid n \geq 2\}$

One of the languages is regular, one context-free and not regular and one not context-free. Which are the regular and the non-regular context-free languages? Justify your answer by giving the corresponding grammars.

Solution:

L_1 is regular: $S \to aS \mid aT, T \to bT \mid U, U \to cdU \mid \varepsilon$.

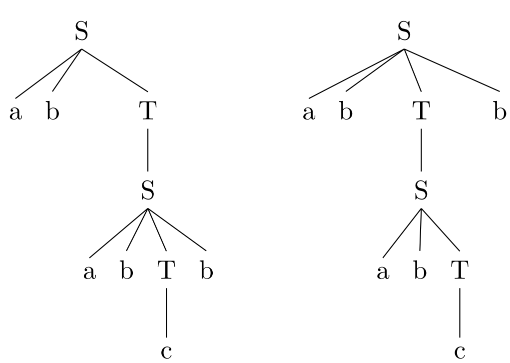
L_2 is context-free: $S \to abSe \mid abT, T \to cD, D \to d \mid cD$.

Question 2 (CFG)

Consider the CFG G with non-terminals $N = \{S, T\}$, terminals $T = \{a, b, c\}$, start symbol S and productions $S \to abTb \mid abT, T \to S \mid c$.

1. *Is G in Chomsky Normal Form? Explain your answer.*
2. *Is G left-recursive? Explain your answer.*
3. *Give the two parse trees for ababcb that one obtains with G.*
4. *What is the language $L(G)$ generated by this grammar?*

Solution:

1. No because a grammar in CNF must be such that each righthand side of a production either contains a single terminal or exactly two non-terminals. $S \to aTb$ for instance does not satisfy this requirement.
2. No, because neither $S \overset{+}{\Rightarrow} S\alpha$ nor $T \overset{+}{\Rightarrow} T\alpha$ for any $\alpha \in (N \cup T)^*$ is possible.
3. 

4. $\{(ab)^n c b^m \mid n \geq 1, n \geq m \geq 0\}$