apl. Prof. Dr. Thomas Noll
Philipp Berger, Matthias Volk

Compiler Construction 2018/19

— Exercise Sheet 3 —

Hand in until November 5th before the exercise class.

General Remarks

- There is *no* practical exercise this week.

Exercise 1 (3 Points)

Consider the context-free grammar G given by the following rules:

$$S \rightarrow a \mid (S) \mid S \cdot S \mid S + S \mid -S$$

(a) Provide a leftmost analysis of the string $a \cdot (-a + (a))$.

(b) Provide a rightmost analysis of the string $a \cdot (-a + (a))$.

(c) Prove or disprove: G is unambiguous.

Exercise 2 (4 Points)

Complete the correctness proof of Theorem 6.1 by showing the direction omitted in the lecture. More precisely, let $G = \langle N, \Sigma, P, S \rangle \in CFG_\Sigma$ and $NTA(G)$ as in the lecture (lecture 6, slide 5). Show that for each $w \in \Sigma^*$ and $z \in [p]^*$ it holds that

$$(w, S, \varepsilon) \vdash^* (\varepsilon, \varepsilon, z) \quad \text{implies} \quad z \text{ is a leftmost analysis of } w.$$

Exercise 3 (5 Points)

Consider the following grammar G which generates binary strings:

$$num \rightarrow 11 \mid 1001 \mid num\ 0 \mid num\ num$$

(a) Show that all binary strings generated by G have values divisible by 3.

(b) Can all multiples of 3 be generated by this grammar? Justify your answer.

Exercise 4 (3 Points)

Consider the following grammar G:

$$\begin{aligned} S &\rightarrow Sc \mid Aa \mid Bb \mid aD \\ A &\rightarrow aB \mid D \\ B &\rightarrow bA \mid B \\ C &\rightarrow aCa \mid c \\ D &\rightarrow Bc \mid a \end{aligned}$$

(a) Compute the first (fi) sets for all nonterminal symbols in G.

(b) Compute the follow (fo) sets for all nonterminal symbols in G.

(c) Does $G \in LL(1)$ hold? Justify your answer.

Exercise 5 (5 Points)

Two characterizations of $LL(1)$ have been given in the lecture.

First, by Lemma 6.5, a context free grammar $G = \langle N, \Sigma, P, S\rangle$ is in $LL(1)$ if and only if for all leftmost derivations of the form

$$S \ \Rightarrow_l^* \ wA\alpha \begin{cases} \Rightarrow_l w\beta\alpha \\ \Rightarrow_l w\gamma\alpha \end{cases}$$

such that $\beta \neq \gamma$, it follows that $\mathsf{fi}_1(\beta\alpha) \cap \mathsf{fi}_1(\gamma\alpha) = \emptyset$.

Second, by Theorem 6.10, G is in $LL(1)$ if and only if for all pairs of rules $A \to \beta \mid \gamma \in P$ (where $\beta \neq \gamma$) we have

$$\mathsf{la}(A \to \beta) \cap \mathsf{la}(A \to \gamma) \ = \ \emptyset.$$

We now lift the latter statement to $LL(k)$ for arbitrary $k \in \mathbb{N}_{>0}$:

Definition: $G \in LL(k)$ iff for all pairs of rules $A \to \beta \mid \gamma \in P$ (where $\beta \neq \gamma$):

$$\mathsf{la}_k(A \to \beta) \cap \mathsf{la}_k(A \to \gamma) = \emptyset$$

where $\mathsf{la}_k(A \to \beta) = \mathsf{fi}_k(\beta \cdot \mathsf{fo}_k(A))$.

Show that the two characterizations are not equivalent for $k > 1$ by giving a grammar that is in $LL(2)$ according to the first characterization, but not according to the lifted version of the second one.