

Lab Session 4

Context-free Grammars and LL(1) Parsing

Interpretation and Compilation of Languages

Nova School of Science and Technology
Mário Pereira mjp.pereira@fct.unl.pt

Version of March 31, 2025

1 Context-Free Grammars

Exercise 1. Choose the correct answer: “The output of the lexical analyzer is

- A** a set of string characters”. **C** a set of regular expressions”.
B an abstract syntax tree”. **D** a set of tokens”.

□

Exercise 2. Consider the following grammar

$$S \rightarrow \text{int} \mid \text{bool} \mid \text{if } S \text{ then } S \mid \text{if } S \text{ then } S \text{ else } S$$

where **if**, **then**, **else**, **int**, and **bool** are terminal symbols and S is the axiom of the grammar. Show that this grammar is ambiguous (so called “dangling else” problem) by giving two distinct parse trees for the word “**if bool then if bool then int else int**”.

□

Exercise 3. Consider three simple grammars below that all recognize the same language of balanced parentheses, *i.e.* they recognize words such as ϵ , $()$, $(())$, $(((())))$, and don’t recognize words such as $)$, $) ($, $((($.

$$P \rightarrow \epsilon \mid (P) \mid PP$$

Grammar G_1

$$P \rightarrow \epsilon \mid (P)P$$

Grammar G_2

$$P \rightarrow \epsilon \mid P(P)$$

Grammar G_3

Only one grammar is LL(1). For the other two, explain why these are not LL(1). In particular, for an ambiguous grammar show the ambiguity by providing two distinct derivation trees for the same word $(())$.

□

Exercise 4. For the grammar that is LL(1) assume that we add a non-terminal symbol A as a new axiom and we add a rule (where $\#$ is the end-of-file terminal symbol)

$$A \rightarrow P\#$$

Give NULL, FIRST, FOLLOW, and the expansion tables for that grammar.

□

2 LL(1) Parsing for Roman Numbers

Consider the following grammar for roman numbers from 0 to 9:

$$\begin{aligned} S &\rightarrow N \text{ EOF} \\ N &\rightarrow iA \mid \mathcal{I}_3 \mid v\mathcal{I}_3 \mid \epsilon \\ A &\rightarrow v \mid x \\ \mathcal{I}_3 &\rightarrow \mathcal{I}_2 i \mid \epsilon \\ \mathcal{I}_2 &\rightarrow \mathcal{I}_1 i \mid \epsilon \\ \mathcal{I}_1 &\rightarrow i \mid \epsilon \end{aligned}$$

Exercise 5. Write the derivation of the number viii.

□

Exercise 6. Is this grammar ambiguous? If no, explain why. If yes, give two derivation trees for the same word recognized by this grammar.

□

Exercise 7. Is this grammar LL(1)? Motivate your answer by either explaining why it is not LL(1) or build an LL(1) parsing table.

□

Exercise 8. Consider another variant of the grammar:

$$\begin{aligned} S &\rightarrow N \text{ EOF} \\ N &\rightarrow iA \mid v\mathcal{I}_3 \mid \epsilon \\ A &\rightarrow \mathcal{I}_2 \mid v \mid x \\ \mathcal{I}_3 &\rightarrow i\mathcal{I}_2 \mid \epsilon \\ \mathcal{I}_2 &\rightarrow i\mathcal{I}_1 \mid \epsilon \\ \mathcal{I}_1 &\rightarrow i \mid \epsilon \end{aligned}$$

Give NULL, FIRST, FOLLOW, and the expansion table for that grammar.

□