

Introduction into ANTLR v4

Pavel Hřebíček

xhrebi04@stud.fit.vutbr.cz

Faculty of Information Technology, Brno University of Technology

ANTLR (**A**Nother **T**ool for **L**anguage **R**ecognition) is a powerful tool to help you create your own compiler. ANTLR was developed in 1989 and its maintainer is Professor Terence Parr of the University of San Francisco. This tool allows you to create a lexical (scanner), syntactic (parser) and semantic analyzer - the first three parts of the compiler. Input for ANTLR is a context-free grammar in Extended Backus-Naur Form (EBNF). EBNF is a notation to express a context-free grammar. The Output from ANTLR is a source code that can recognize the language. ANTLR parser using a top-down approach. ANTLR v4 support modern programming languages – JAVA, C#, C++, JavaScript, Python2, Python3, Swift and Go.

The first and shorter part of the lecture will focus on introduction ANTLR and its concept. This part will be rather theoretical. On the other hand, the second part will be very practical. I prepared two examples – one simple and the other more complex. The first will familiarize you with the ANTLR technically. What you need to install the tool on your laptop. How ANTLR generates code for the lexical, syntactic and semantic analyzer. How to create, run and test your first grammar. In the second example, I will show you how to use the ANTLR tool in Java and what is the best and easiest way to get an ANTLR into a Java project. Then I will show you walking mechanism (walking over the tree) provided by the ANTLR v4 library.