Lab 3: Lexical Analysis

I.   Preparation

Install ANTLR as instructed in the guidelines

II.   Practice

- Run the program as instructed and observe the results
- Handle more token types according to the lab instructor's request

III.   Additional Exercises

Handle the following token types

1. *Number:* All numbers L are considered as float number in C/C++, just as real numbers. A number consists of an integer part, a decimal part and an exponent part. The integer part is a sequence of one or more digits with an optional '-' sign if negative. The decimal part is a decimal point followed optionally by some digits. The exponent part starts with the character 'e' or 'E' followed by an optional '+' or '-' sign, and then one or more digits. The decimal part or the exponent part can be omitted.

For example:

0 199 12. 12.3 12.3e3 12.3e-30

2. *Boolean:* A boolean literal is either True or False, formed from ASCII letters.

3. *String:* A string literal includes zero or more characters enclosed by double quotes (").
4. *Array:* An array literal is a comma-separated list of literals enclosed in '[' and ']'. The literal elements are in the same type.

For example, [1, 5, 7, 12] or [[1, 2], [4, 5], [3, 5]].