

**EGE UNIVERSITY**

**FACULTY OF ENGINEERING**

**COMPUTER ENGINEERING DEPARTMENT**

**PROGRAMMING LANGUAGES**

**2020–2021 SPRING SEMESTER**

**PROJECT-1 REPORT**

**(Lexical Analyzer)**

**DELIVERY DATE**

8/6/2021

**PREPARED BY**

05180000072, Ahmet Tartar

05180000076, Yavuz Selim İlik

Contents

[General Explanation 3](#_Toc74067927)

[How Program Works? 3](#_Toc74067928)

[Functions 3](#_Toc74067929)

[Is Keyword 3](#_Toc74067930)

[Is Integer 4](#_Toc74067931)

[Is Variable 4](#_Toc74067932)

[Is Comment 5](#_Toc74067933)

[Is String Constant 6](#_Toc74067934)

[Lexical Analyzer 6](#_Toc74067935)

[Example Inputs and Outputs 11](#_Toc74067936)

[Invalid Inputs: 14](#_Toc74067937)

[Conclusion 15](#_Toc74067938)

# General Explanation

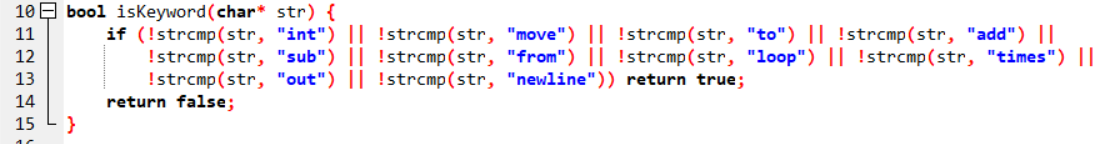
The Project aims to analyze words in BigAdd language. The program analyses words separately. Analyzed words are divided into some types such as identifier, integer, keyword, bracket etc. Every special expression contains distinguishable features. The Lexical Analyzer detects these features.

# How Program Works?

The user should enter the name of the file for example myscript. The program automatically adds .ba and opens the file. After that, the program looks if is there any “{“ character and if there is, erases all of the words inside that brackets. Continuously, the functions are starting work. At this point, there are so many if blocks. After the suitability of the inputs is tested, if there is no error, the type of the input and input itself are printed on the screen.

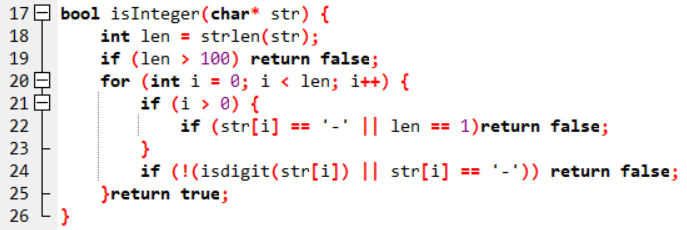
# Functions

## Is Keyword



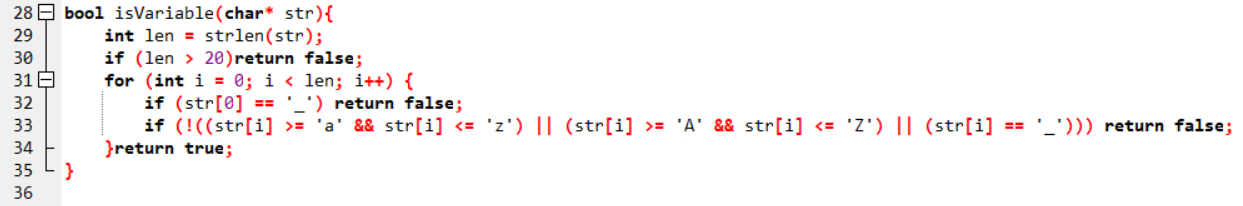
This function detects if the input is a keyword. If it is not, returns false.

## Is Integer



This function looks if the input is appropriate for being an integer. For example, if the input length is just 1, the function returns false.

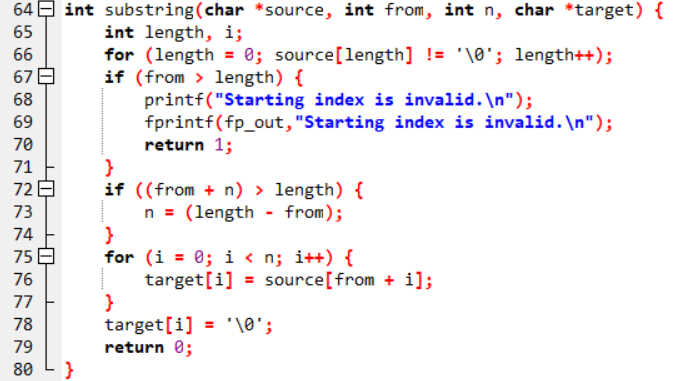
## Is Variable



This function looks if the input is appropriate for being a variable. If the input length bigger than 20 characters, returns false. Moreover, if the input starts with “\_” character also returns false. Finally, if the input word contains any of a character other than alphabetic letters, the function returns false.

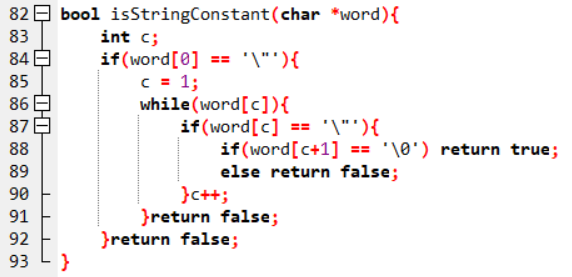
## Is Comment





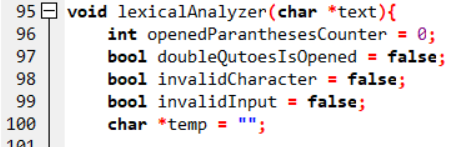
These functions look if the input is a comment line. And also they look if the comment parentheses are closed. After that, they erase those comment lines and then the compiler goes on.

## Is String Constant

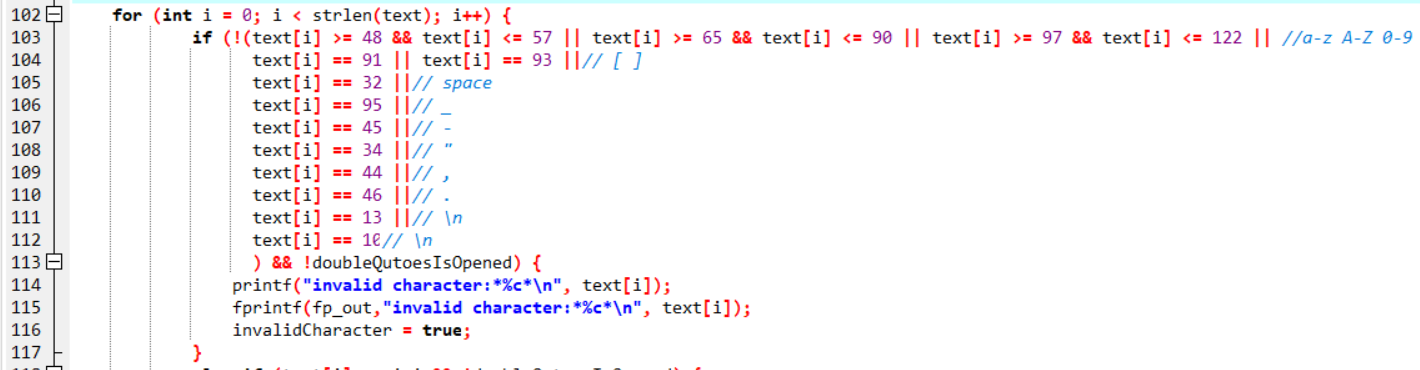


This function looks if the input is a string constant and checks the double quotes in the true form. If the user tries to enter one quote or another quote type (‘’ “ “), the program does not accept that. You can only use “ ” characters.

## Lexical Analyzer

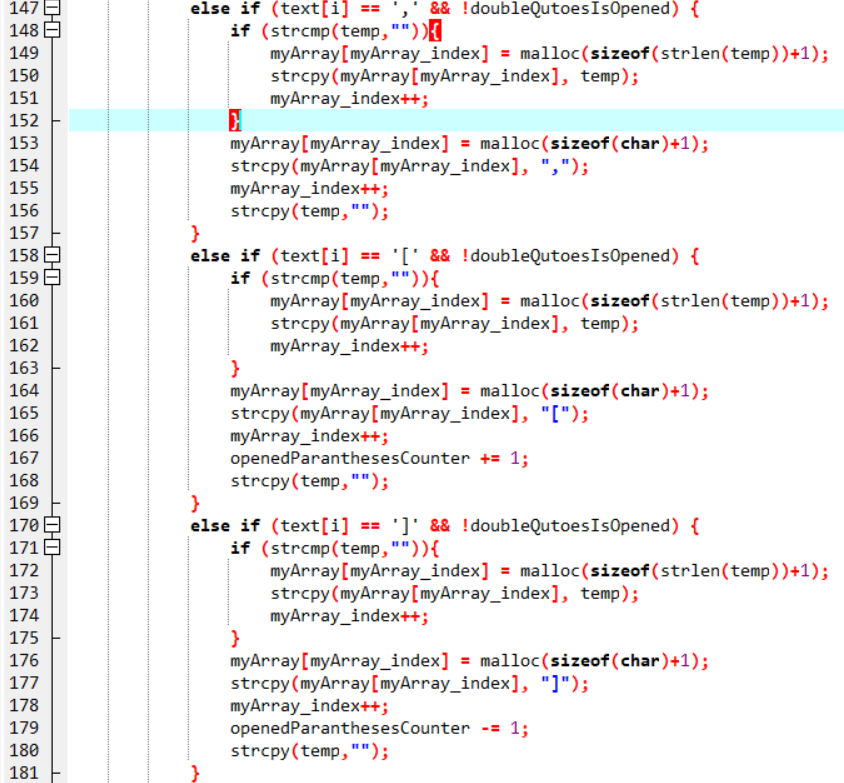


Here, we defined some variables such as parantheses counter and is the double quotes opened.



In this section, the program does not accept the input as a valid input if the input has different characters from what you see above. And then prints “Invalid character”. By the way, there is an exception here, if your input is inside the quotes, then it is a valid character.

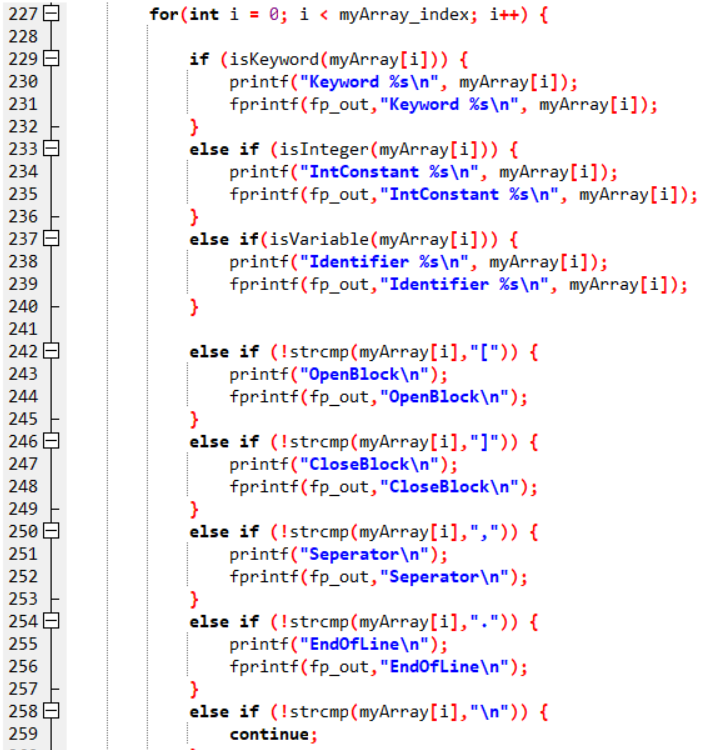




While the characters are adding an array, the program looks is there any space “ ” characters. If any, it is ignored for not to take up space. Additionally, if the text is “[“, the program adds 1 to openedParantesesCounter.



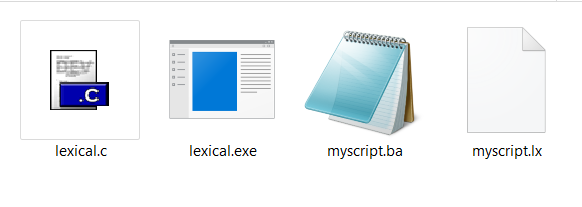
In this section, the program adds all of the inputs to a string constant after the quote marks open until the marks are closed. Even if there are spaces between the characters, they are printed on the screen as a single character, when the quotations are closed.



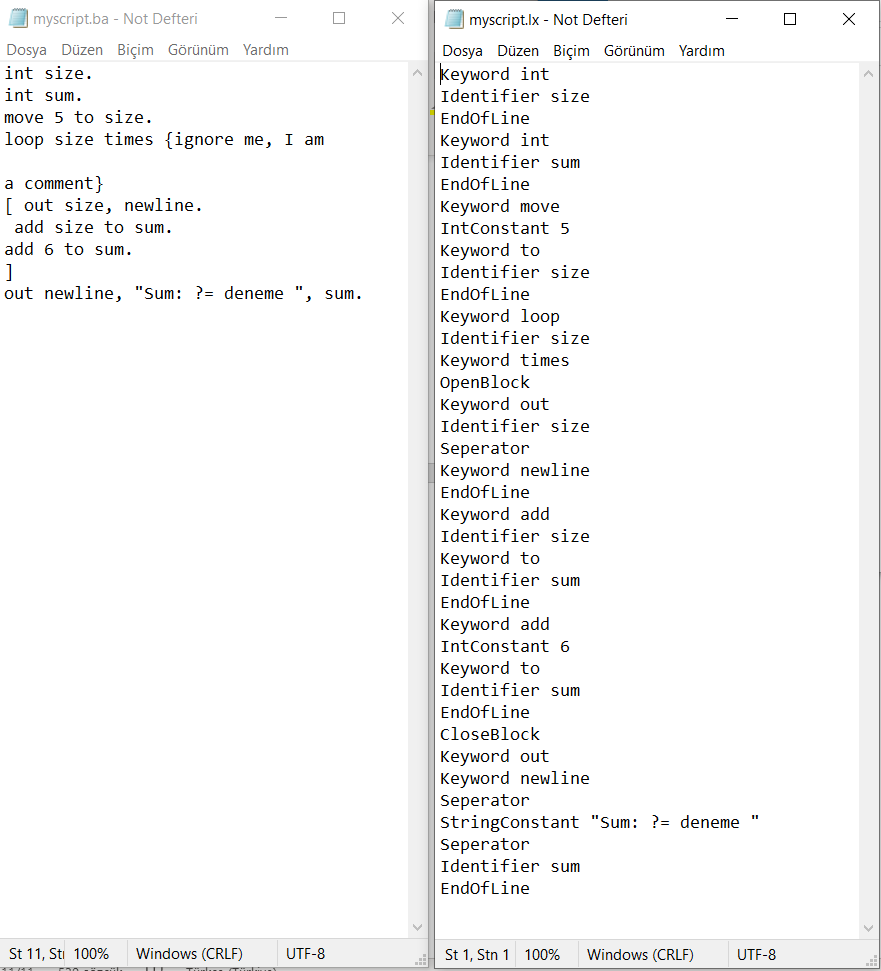


In these for and if blocks, the program prints all of the inputs themselves and their types. If the input does not belong to any type, the program throws error messages as you see above.

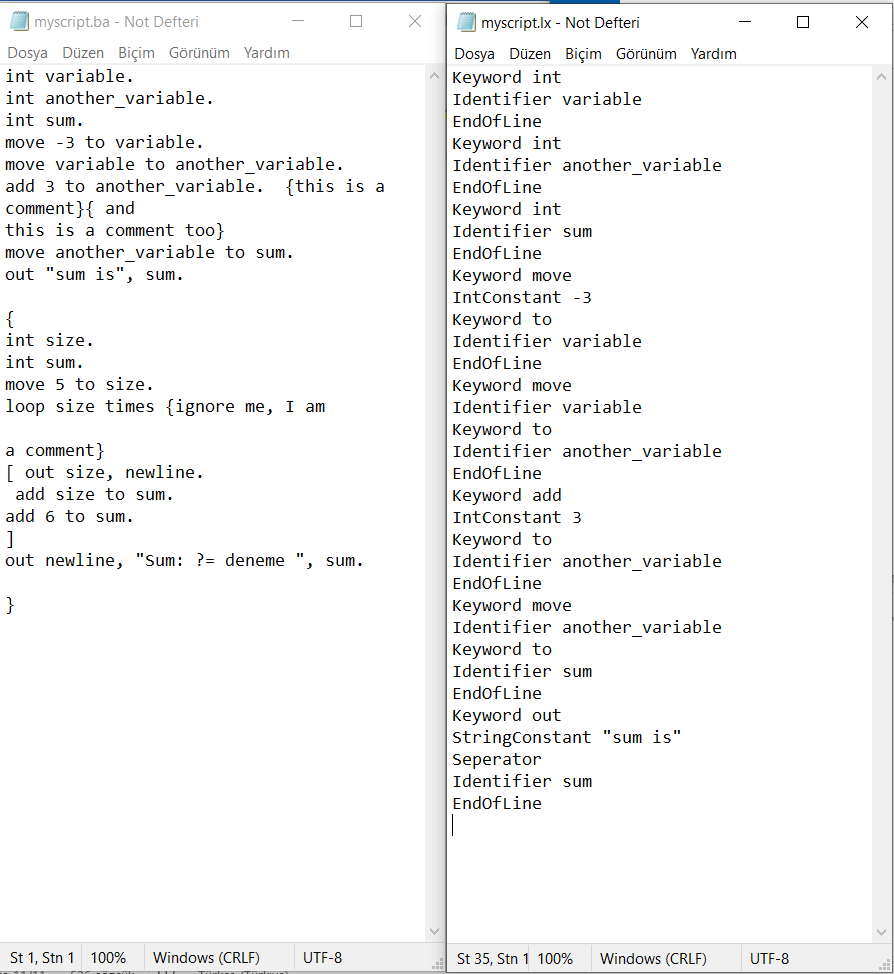
# Example Inputs and Outputs



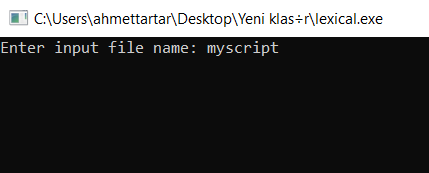
The project folder seems like the above. The user should create a script.ba file. The program read .ba file and creates .lx file. You can read these files with open as a txt file.

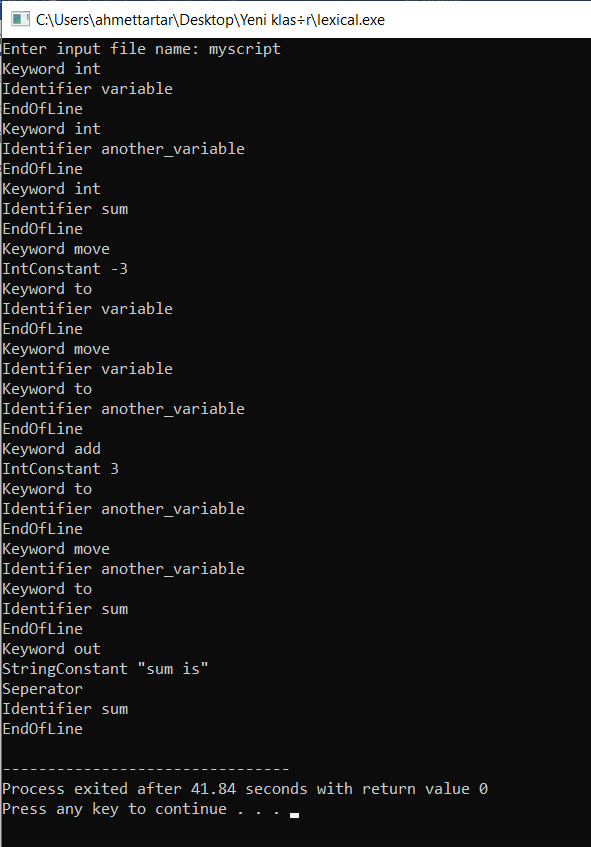


An example for a normal run.

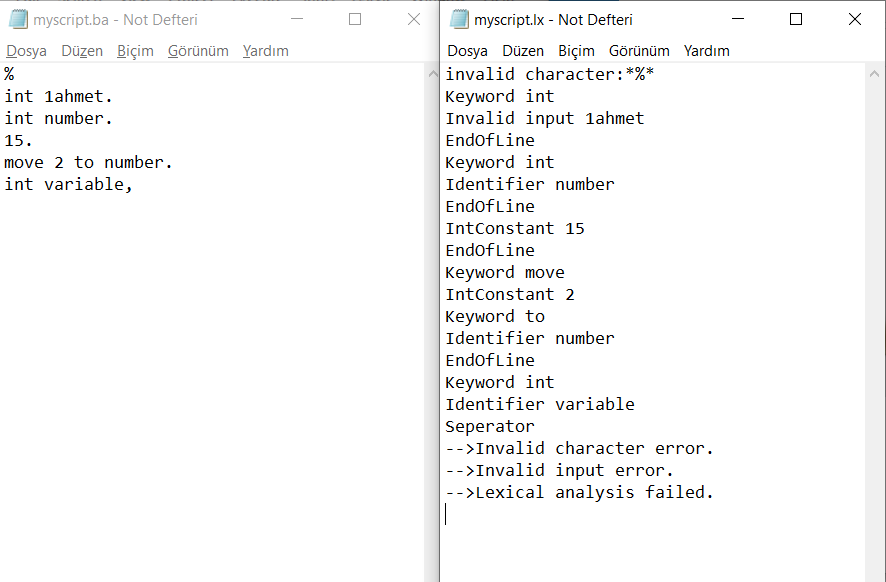


Another example for a normal run. By the way, you can see the outputs on the terminal :





# Invalid Inputs:



If you enter invalid characters or you try to define inappropriate variables, the program prints “invalid character” or “invalid input”.

# Conclusion

In this project, we designed software for making lexical analysis for a ''BigAdd" programming language.