Github link: https://github.com/LolutaStefana/UBB-Computer-Science/tree/main/semester5/FLCD/lab3

MyScanner class represents a basic scanner for a simple programming language. The scanner tokenizes a source program, classifies each token, and constructs a Program Internal Form (PIF) and a Symbol Table based on the input program.

Constructor:

- Initializes the symbolTable, the pif, and sets the filePath for reading the program.
  - @param filePath - Represents the filePath from where the program will be read.

readFile():

- Reads the content of the file, replaces tabs with spaces, and returns the file content.
  - @return - The content of the read file.
  - @throws FileNotFoundException if the file doesn't exist.

createListOfProgramElems():

- Prepares the array for tokenization, reads the content, and splits the program into a list of strings.
- Calls the tokenize method to create a list of pairs containing tokens/identifiers/constants and their line numbers.
  - @return - List of pairs with tokens/identifiers/constants and their line numbers.

tokenize(List<String> tokensToBe):

- Tokenizes the program elements based on specific cases like strings, chars, and new lines.
- Builds a list of pairs containing tokens/identifiers/constants and their line numbers.
  - @param tokensToBe - List of program elements (strings) + separators.
  - @return - List of pairs with tokens/identifiers/constants and their line numbers.

scan():

- Scans the program, classifies each token, and populates the PIF and Symbol Table.
- Prints lexical errors if encountered.

getPif():

- Returns the Program Internal Form (PIF).

  - @return - The Program Internal Form.

getSymbolTable():

- Returns the Symbol Table.

  - @return - The Symbol Table.

ProgramInternalForm:

ProgramInternalForm class is designed to store tokens, identifiers, and constants along with their positions in the symbol table and respective categories.

Constructor:

Initializes two lists within the class.

Add Method:

Adds a token/identifier/constant to the list along with its position in the symbol table and the corresponding category.

Parameters:

pair: A pair containing the token/constant/identifier and its position in the symbol table.

type: Category of the token (0-4: constant, identifier, reservedWord, operator, separator).