https://github.com/914-Modolea-Bogdan/LFTC/tree/main/Lab04

It reads a source code file, tokenizes it, and categorizes the tokens into various groups. The program recognizes operators, separators, reserved words, constants, and identifiers based on predefined patterns and stores this information in a Program Internal Form (PIF) and a symbol table. If any tokens do not match the expected patterns, the program reports lexical errors, indicating the line number and the problematic token.

```java
/**
 * This method reads the operator, separator, and reserved word lists from a file, populating the corresponding lists.
 * @throws FileNotFoundException
 */
public void readTokens() throws FileNotFoundException
```

```java
/**
 * Reads the content of the source code file.
 * @return the content of the source code as a String.
 * @throws FileNotFoundException
 */
private String readProgram() throws FileNotFoundException
```

```java
/**
 * Splits the source code into a list of tokens using separators.
 * Processes the list of tokens and categorizes them.
 * @return a list of pairs (token, line) where each token is associated with the line it appears on.
 */
List<Pair<String, Integer>> getTokens()
```

```java
/**
 * Scans the source code, identifies tokens, and categorizes them as reserved words, operators, separators, constants, or identifiers.
 * Any lexical errors are reported. The results are stored in the PIF and symbol table.
 */
public void scan()
```