https://github.com/andreicupes/FLCD/tree/main/Lab2/src

Lab 4 Documentation

Besides the Symbol Table I have created for the 2nd Lab, I have created 3 more entities, respectively called:

- ProgramInternalForm
- LanguageSpecification
- MyScanner

The Class ProgramInternalForm is basically a list of pairs, while the 2nd element from the pair being a pair himself. It is used to correlate what we do have in our language specification (token table) with its respective pair in the Symbol Table. The class has just an add method in order to add a new element, and the toString method has been overridden in order to display nicely.

The Class LanguageSpecification is what I have used to represent my Token Table. We have as attributes hardcoded separators, operators and reserved words. We then add those in a hash map we will use later in our PIF. We needed some methods in order to check if a given token is part of the different attributes of this class, and I have used REGEX in order to do some matching to see if a given token is a valid constant, either a numerical one, a char or a string.

The Class MyScanner has as attributes the languageSpecification, the PIF, the Symbol Table and its capacity. We also have 2 different attributes in order to manage the files we are working with, the input and the output. We declare a method called scan that will run through the input file and build our tables. It will call the method tokenize, which is used in order to classify our tokens. For each case, we will call different methods that will help us check with the language specification and return the proper token as a string. Moreover, we use a function in order to create our PIF, and in this method we will take the tokens we have after scanning and put them in our table accordingly. We use a flag and whenever we hit an invalid token we signal the lexical error for the user. The last function in this class is the one used to print in an output file both the pif and symbol table.