The scanner implementation uses the following approach:

- We initialize the symbol table as a hash table. The has table uses a simple chaining approach where nodes will be used to chain items that land on the same position. The hash function consists of computing the sum of the identifiers' characters ascii values. The insert function adds the identifier based on the hash function and the find function returns the position (a.k.a. the ascii sum) of the searched identifier. To the symbol table we add the tokens which are identifiers or constants.
- We initialize the PIF as a list to which we add every pair of (token, symbol table position).
- The tokens will be read and saved in a list, from a text file using the get_tokens function.
- We read the text file that contains our problem (which we can swap by changing the filename variable at the beginning), splitting every line by every symbol in our language. Reserved words, operators and separators will be added to PIF, but not to the symbol table, while identifiers and constants will be added to both. If any token is not part of any of the above categories, a lexical error will be printed in the console.
- The symbol table and the PIF are both printed in their own files.