TP1 – Lexical Analyzer

Exercice 1 :

1. What is the alphabet of the language ?

The alphabet of the language is the ASCII table.

1. What are the lexemes (as regular expressions) of the language ?

The lexemes as regular expressions of the language are:

1. Write the table that is matching the lexemes, the tokens, and the attributes of the tokens.

|  |  |  |  |
| --- | --- | --- | --- |
| Lexeme | Regex | Token | |
| Name | Attributes |
| PRINT | PRINT | PRINT | - |
| IF | IF | IF | - |
| THEN | THEN | THEN | - |
| = | = | RELOP | EQ |
| > | > | RELOP | GT |
| < | < | RELOP | LT |
| <> | (><)|(<>) | RELOP | NE |
| <= | <= | RELOP | LE |
| >= | >= | RELOP | GE |
| GOTO | GOTO | GOTO | - |
| INPUT | INPUT | INPUT | - |
| , | , | COMA | - |
| ( | ( | OPEN PARENTESIS | - |
| ) | ) | CLOSE PARENTESIS | - |
| LET | LET | LET | - |
| GOSUB | GOSUB | GOSUB | - |
| RETURN | RETURN | RETURN | - |
| END | END | END | - |
| REM | ^REM[space \t \f].\*$ | - | - |
| a\_name | [a-z A-Z][a-z A-Z 0-9]\* | ID | Reference to the symbol table of a\_name |
| 10.9 | [0-9]+(.+[0-9])? | NUM | 10.9 |
| “Hello” | “[^”]” | STRING | HELLO |
| Skipping | [space \t \f]+ | - | - |

Exercice n°2 :