

Reporte: Fase 1

Para la implementación del analizador lexicográfico (Lexer) del lenguaje propuesto “BasicTran” se utilizó el lenguaje Python3 y la librería PLY, basada en las herramientas de construcción de compiladores lex y yacc. En primer lugar, se definieron los tokens que serán buscados por lexer y que devolverá como una lista con su respectiva posición de fila y columna en el programa realizado en el lenguaje dado.

Los tokens se definieron utilizando las estructuras de lista y diccionario. Una lista para los tokens relacionados con separadores y operadores del lenguaje y un diccionario para agrupar todas las palabras reservadas en un solo token para reducir el número de expresiones regulares que se utilizarán para definir los patrones que deben cumplir los tokens.

Una vez indicados los tokens se definieron las expresiones regulares que harán “match” para identificar los distintos tokens del lenguaje propuesto. Para el caso de los tokens de operadores aritméticos y separadores, las expresiones se establecieron utilizando reglas compatibles con el módulo de expresiones regulares (re) de Python. Para el caso de palabras reservadas se implementaron distintas funciones que utilizan expresiones más complejas y que trabajan con los atributos *t.value* y *t.type* del objeto LexToken. Además se agregaron funciones con expresiones para manejo de errores lexicográficos y obtener el número de columna que corresponde a un token determinado.

Adicionalmente se incluyeron en los siguientes tokens para detectar errores:

- TkError: detecta nombres de variables no válidos.
- TkErrorSol: detecta caracteres no definidos por el lenguaje que no forman parte de ningún otro token.
- TkCharacterError: detecta caracteres de longitud mayor a 1.
- TkCarError: detecta caracteres no válidos dentro de comillas simples.
- TkCarEspecial: detecta un carácter especial seguido de un error.

Finalmente, para ejecutar el lexer en el formato deseado `./Lex <Archivo>` es necesario otorgarle permisos de ejecución con el comando `chmod 777` al archivo `Lex` donde se encuentra un script que permitirá correr el analizador lexicográfico sobre el archivo que contiene el programa en el lenguaje propuesto.