

TP 2 - Técnicas de Compilación-

CARRERA:

- Ingeniería Informática

MATERIA:

- Técnicas de Compilación

DOCENTE:

- Maximiliano A. Eschoyez

INTEGRANTES:

- Bobadilla Barceló, Daniel Agustin
- Genaro, Kevin Luis

INFORME

CONSIGNA

Dado un archivo de entrada en lenguaje C, se debe generar como salida el Árbol Sintáctico (ANTLR) correcto y mostrar por consola o archivo el contenido de la Tabla de Símbolos de cada contexto. Esto es válido únicamente cuando el archivo de entrada es correcto.

Para esta versión se deberá realizar un control de errores básicos y generar un reporte de lo encontrado para poder analizar archivos con posibles errores de codificación. El reporte de errores podrá realizarse por consola o archivo de texto.

Los errores a detectar deben ser los siguientes:

- Errores sintácticos comunes:
 - Falta de un punto y coma,
 - Falta de apertura de paréntesis,
 - Formato incorrecto en lista de declaración de variables
- Errores semánticos comunes:
 - Doble declaración del mismo identificador,
 - Uso de un identificador no declarado,
 - Uso de un identificador sin inicializar,
 - Identificador declarado pero no usado,
 - Tipos de datos incompatibles.

Cada error reportado debe indicarse si es sintáctico o semántico.

OBJETIVO

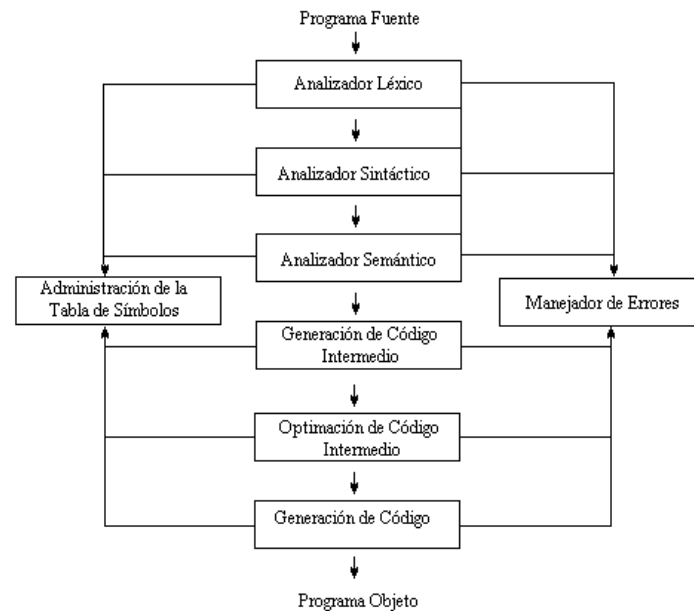
El objetivo de este Trabajo Práctico es agregar al Trabajo Práctico 1 la tabla de símbolos y el reporte de errores. El programa a desarrollar para esta etapa tiene como objetivo tomar un archivo de código fuente en una versión reducida del lenguaje C, posiblemente con

errores léxicos, sintácticos y semánticos, y generar como salida el árbol Sintáctico (ANTLR) y la Tabla de Símbolos y reporte de errores.

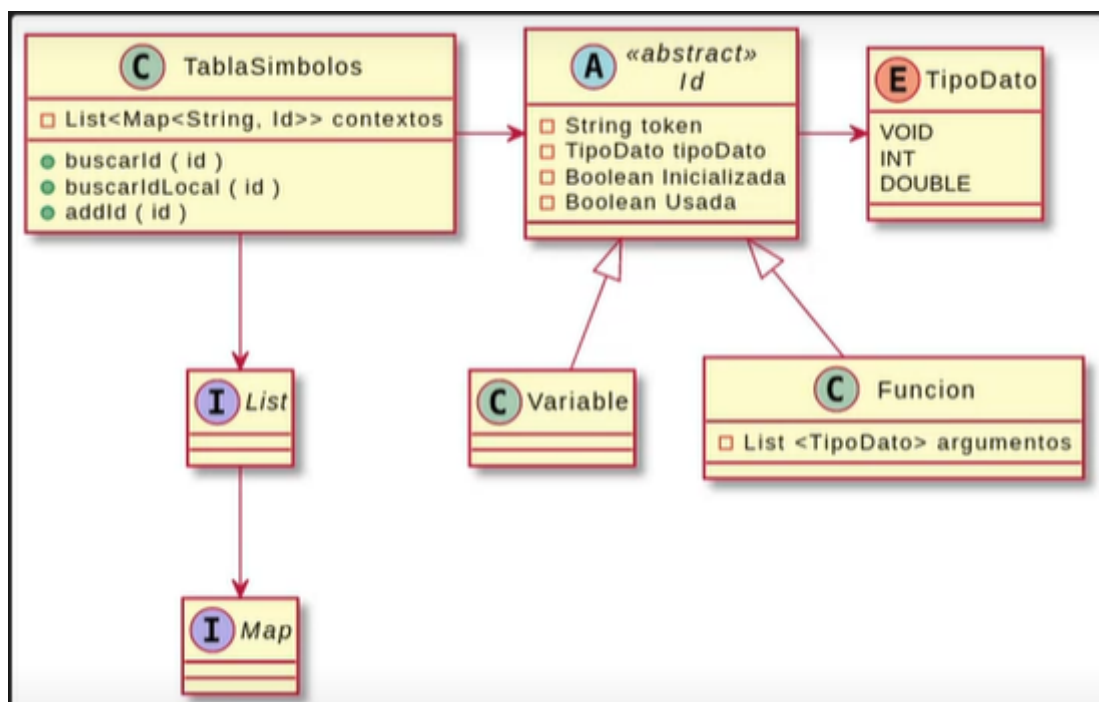
DESARROLLO

Para el TP1 tuvimos en cuenta el **Análisis Léxico** gracias a las Expresiones Regulares que revisaban la ortografía determinando si eran palabras propias del lenguaje y a esta secuencia de caracteres las convertía en Tokens (**Secuencia de Tokens**). Luego esta secuencia ingresaba al **Análisis Sintáctico** que se encargaba de analizar si los tokens conformaban una estructura correcta (**Oraciones bien escritas**).

Para este TP2 avanzamos en el proceso y llegamos al **Análisis Semántico**, en donde verificamos que todo lo que está escrito tenga algún sentido. Para poder avanzar con este análisis debemos incorporar una **Tabla de símbolos** en donde vamos a ir guardando los Tokens (nombre de funciones y variables). Esta tabla nos va a permitir saber si la variable a analizar está declarada o no, inicializada, usada, también si existe o no en un contexto esperado o está en uno de mayor jerarquía, etc. En el caso de las funciones va a pasar lo mismo, vamos a poder guardar la lista de argumentos de cada una con sus respectivos tipos, el tipo de dato que retornan, cantidad de argumentos que reciben, etc. Cabe aclarar que los identificadores se van agregar en las etapas tempranas pero la Tabla de Símbolos se usa siempre, inclusive luego de compilar y ejecutar el archivo ejecutable. Junto con el análisis semántico incorporamos el **Manejo de errores**, algo que no veníamos utilizando hasta ahora.



Partimos de la base de lo explicado en clases antes de ponernos a desarrollar sabiendo que:



A tener en cuenta:

- La Tabla de Símbolos para nosotros es una **Lista de Contextos** en donde cada nodo representa un contexto diferente con un **diccionario** de identificadores asociado.
- Un identificador puede ser una **variable** o **función**.

Luego de comprender toda esta estructura comenzamos a desarrollar teniendo en cuenta también todo lo aprendido en clases.

CONCLUSIÓN

Este trabajo práctico nos pareció mucho más desafiante que la primera parte ya que debimos refactorizar nuestras reglas gramaticales para facilitar la creación de la Tabla de Símbolos así como agregar y eliminar contextos de la misma.

También nos pareció interesante cómo funcionan los Listener y el manejo de errores personalizados que no tuvimos en cuenta en el TP1.

Ver las clases en video nos ayudó mucho para reforzar lo visto en clases y luego de realizar esta parte de la materia creemos que estamos capacitados para comenzar con el código de 3 direcciones.