

TRABAJO PRACTICO 1 - TC

GRAMMAR LEXER PARA LENGUAJE C REDUCIDO

Introducción: Problemática abordada

Dado un archivo de entrada en lenguaje C, se debe generar como salida el Árbol Sintáctico (ANTLR) correcto. Para lograr esto se debe construir un parser que tenga como mínimo la implementación de los siguientes puntos:

- Reconocimiento de un bloque de código, que puede estar en cualquier parte del código fuente, controlando balance de llaves.

Y Verificación de:

- _ declaraciones y asignaciones,
- _ operaciones aritmético/lógicas,
- _ declaración/llamada a función.
- _ estructuras de control if, for y while.

Solución propuesta

Para el desarrollo de esta actividad, se definió los TOKENS al principio del código y luego las reglas gramaticales. Se tomo en cuenta definir todos los TOKENS que podían llegar a ser necesarios para las reglas, pero además también TOKENS para símbolos ASCII como los de operaciones matemáticas o lógicas o símbolos de puntuación y exclamación. Además, se tuvieron en cuenta las palabras reservadas propias del lenguaje C y los espacios en blanco y saltos de línea.

El orden de declaración de TOKENS se hizo respetando la regla de generalidad, primero los TOKENS más específicos y luego los más generales. De lo contrario, los más generales “sobrescriben” a los más específicos que se encuentra contenidos en el general.

Además, la estructura con respecto a las reglas gramaticales se realizó de manera encadenada y recursiva para poder tener en cuenta las distintas expresiones posibles. Se definió una expresión general de la cual se desprendieron sub-expresiones.

Conclusión

Para poder realizar esta actividad, se debió investigar sobre la estructura del lenguaje C, teniendo en cuenta qué está permitido y qué no, qué compila y qué da error. Muchas veces se estimó que alguno de los requerimientos seria sencillo y rápido, pero luego de analizarlo se descubrió que cada uno debía seguir varios tipos de caminos distintos. Como, por ejemplo, la declaración de una variable que puede ser solamente la declaración de su tipo y su nombre o puede ser su tipo, su nombre y su valor. Y a su vez su valor puede ser un solo termino o una expresión y así continua. Por ende, la definición de un lenguaje no es tarea sencilla y requiere de un conocimiento extensivo del lenguaje el cual queremos describir.