

Write a lex program to remove the comments, extra white spaces/tabs/newlines etc., and tokenize all valid C program constructs, such as integer constants, floating point constants, character/string constants, identifiers, Keywords, operators like Arithmetic, Relational, Assignment, Conditional, bitwise etc., and install the necessary information of the tokens identifiers and constants, into a symbol table.

The symbol table should maintain at least the following information for the identifiers and constants.

- token name
- lexeme
- value (if any)

For example, for a C statement **int** marks=95;, these values should be,

- token name is id1
- lexeme is marks
- value is 95

Note: There may be many identifiers in your source program. You may identify each of them by giving unique token names (like id1, id2,

The input is a C program whose name is given as command line argument, and output should be a) contents of symbol table b) all other token, displayed on the screen.