

CSL302: Compiler Design

Lab-3 (Lexical Analysis)

Due Date: 16-Aug-2025 11:59 PM

Instructions:

1. Prepare a zip file and upload your solutions on canvas
2. Include a readme file
3. Submit the solutions on Canvas.

Question:

The C preprocessor is a macro processor that is used automatically by the C compiler to transform your program before actual compilation.

The C pre-processor supports the following features.

1. Inclusion of header files
2. Macro expansion
3. Conditional compilation

In this lab, you are required to implement the lexical analysis for the subset/variant of C pre-processor features. The following gives the list of features.

Keywords: #define, #include, #ifdef, #ifndef, #if, #else, #elif, #endif, #undef, #pragma, #error, #line

String literals: These are strings enclosed within the ""

File name: Any string containing alphabets that end with .h or .c (Eg. newfile.c or newfile.h)

Integers: They can be represented in both decimal and hexa decimal format.

Floating point numbers

Operators: +, -, *, /, ==

Comments:

Single line comments are represented using @@ followed by any comment string.

Multiline comments are enclosed within the quotes "" and ""

Identifiers: Start with \$ followed by at least one capital alphabet

White spaces, tab spaces, line breaks should be ignored

Any other character other than the above should be printed on the console as it is.

For all the recognized words a corresponding token should be printed on the console. The token name can be chosen as per your choice.