Practical 1

Compiler Construction

2CS701

Mistry Unnat

20BCE515

Department of Computer Science and Engineering

Institute of Technology

Nirma University

Ahmedabad

Aim :

To implement lexical analyse to recognize all distinct token classes:

(Use flex/lex tool to recognize all distinct token classes (Data type, Identifier, constant (Integer, Float, Char, String), Operator (Arithmetic, Relational, Assign, Unary +/-, Increment), Single line/Multi-line comments, Special symbol(;,{}()) .

lex.l :

```
%{
#include <stdio.h>
%}
DATATYPES       (int|float|double|char)
ID              [A-Za-z][a-zA-Z0-9_]*
CONSTANT        ([0-9]+|[0-9]+.[0-9]+|'[a-zA-Z]+')
SIN_LINE_CMT    "//".*
MUL_LINE_CMT    "/*".*"*/"
OPERATOR        ("+"|"-"|"/"|"*"|">>"|"<<"|"++"|"--"|"&&"|"||"|"="|"==")
SPECIAL_SYM     [{}();.?]
%%
\n      {}
" "     {}
{DATATYPES}     { printf("\tDatatype : %s\n", yytext); }
{ID}            { printf("\tID : %s\n", yytext);}
{CONSTANT}      { printf("\tConstant : %s\n", yytext); }
{OPERATOR}      { printf("\tOperator : %s\n", yytext); }
{SIN_LINE_CMT}  { printf("\tSingle Line Commnent : %s\n", yytext); }
{MUL_LINE_CMT}  { printf("\tMulti Line Comment : %s\n", yytext); }
{SPECIAL_SYM}   { printf("\tSpecial Symbol : %s\n", yytext); }

%%
int main(){
    yylex();
    return 0;
}

int yywrap(){
    return 0;
}
```

Output :

```
unnat@unnat:~/Desktop$ ./a.out
char ch = 'a';
        Datatype : char
        ID : ch
        Operator : =
        Constant : 'a'
        Special Symbol : ;
// Single line comment
        Single Line Commnent : // Single line comment
/* Multi line comment */
        Multi Line Comment : /* Multi line comment */
int num_1 = 3653;
        Datatype : int
        ID : num_1
        Operator : =
        Constant : 3653
        Special Symbol : ;
float num_2 = 142.25;
        Datatype : float
        ID : num_2
        Operator : =
        Constant : 142.25
        Special Symbol : ;
int sum = 35 + 36;
        Datatype : int
        ID : sum
        Operator : =
        Constant : 35
        Operator : +
        Constant : 36
        Special Symbol : ;
num1++;
        ID : num1
        Operator : ++
        Special Symbol : ;
^C
unnat@unnat:~/Desktop$
```

Conclusion :

From this practical, we learnt what is lexical analysis and wrote a lex code to identify various token classes.