19CSE401 - Compiler Design

Lex Programming

Lab sheet-3

Done By:

N Sai Pavan Krishna
AM.EN.U4CSE19347

1. Write a lex program to extract tokens from a given source code.

Source Code:

```
#include<stdio.h>
 main()
{
int fact=1,n;
for(int i=1;i<=n;i++)
{
fact=fact*i;
 }
printf("Factorial Value of N is", fact); getch();
}
```

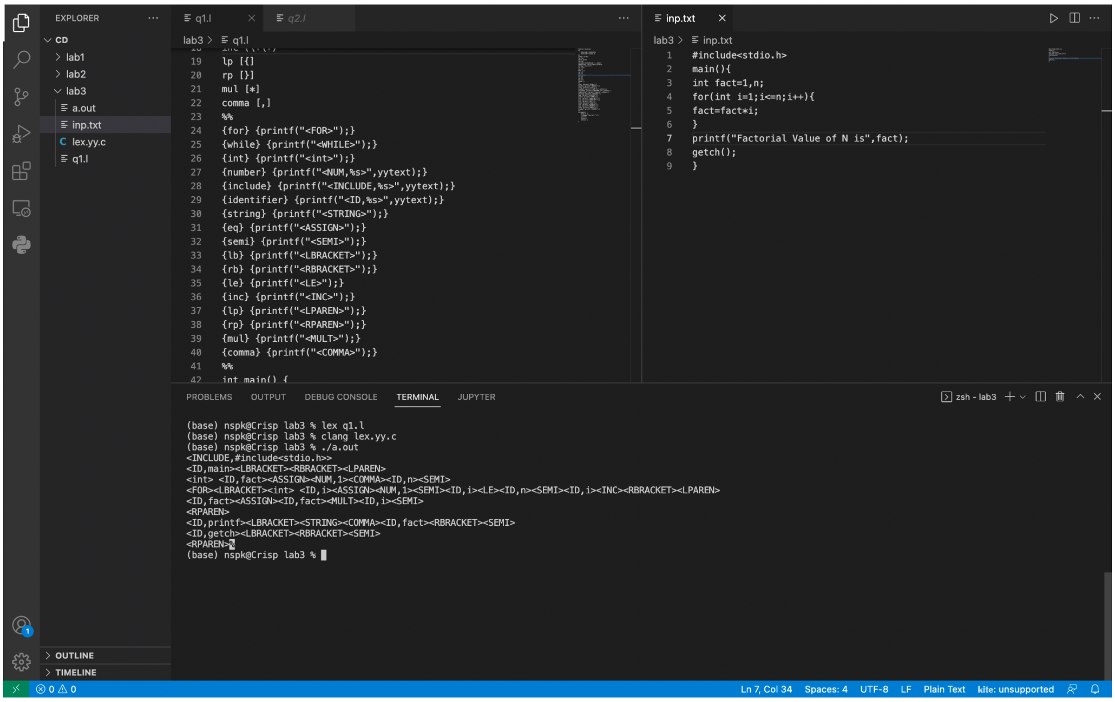
a. For header file the token must be **<INCLUDE, #include<stdio.h>>**

b. For keywords, the token must be **<keyword>**. Example **, <FOR>, <INT>**

c. For variable names and function name the token must be <**ID, name**> **.**
Example <ID, main> , <ID, fact>, <ID, i><ID, n>, <ID, printf>, <ID, getch>, etc

d. For the symbols:

i.	=	-	<ASSIGN>
ii.	;	-	<SEMI>
iii.	(	-	<LBRACKET>
iv.	)	-	<RBRACKET>
v.	<=	-	<LE>
vi.	++	-	<INC>
vii.	{	-	<LPAREN>
viii.	}	-	<RPAREN>
ix.	*	-	< MULT>
x.	"anystring"	-	<STRING>
xi.	,	-	<COMMA>

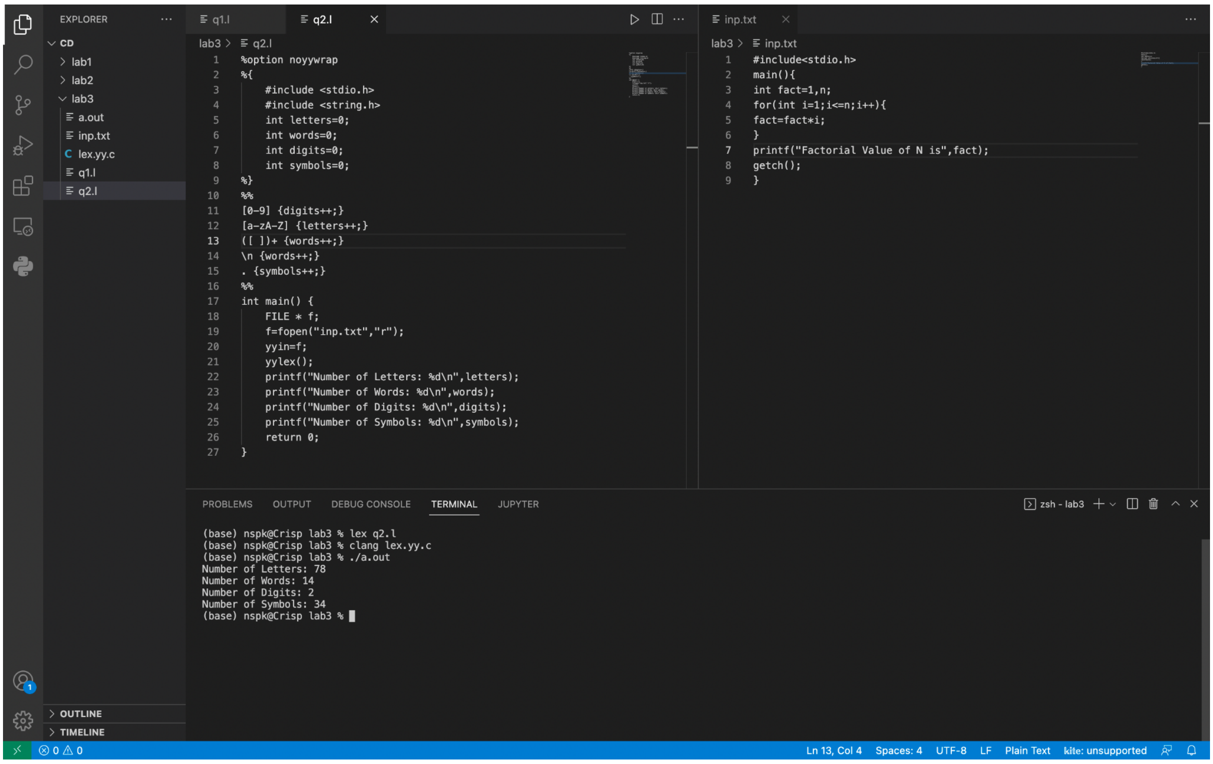


```
lp [{]
rp [}]
mul [*]
comma [,]
%%
{for} {printf("<FOR>");}
{while} {printf("<WHILE>");}
{int} {printf("<int>");}
{number} {printf("<NUM,%s>",yytext);}
{include} {printf("<INCLUDE,%s>",yytext);}
{identifier} {printf("<ID,%s>",yytext);}
{string} {printf("<STRING>");}
{eq} {printf("<ASSIGN>");}
{semi} {printf("<SEMI>");}
{lb} {printf("<LBRACKET>");}
{rb} {printf("<RBRACKET>");}
{le} {printf("<LE>");}
{inc} {printf("<INC>");}
{lp} {printf("<LPAREN>");}
{rp} {printf("<RPAREN>");}
{mul} {printf("<MULT>");}
{comma} {printf("<COMMA>");}
%%
int main() {
```

```
#include<stdio.h>
main(){
int fact=1,n;
for(int i=1;i<=n;i++){
fact=fact*i;
}
printf("Factorial Value of N is",fact);
getch();
}
```

```
(base) nspk@Crisp lab3 % lex q1.l
(base) nspk@Crisp lab3 % clang lex.yy.c
(base) nspk@Crisp lab3 % ./a.out
<INCLUDE,#include<stdio.h>>
<ID,main><LBRACKET><RBRACKET><LPAREN>
<int> <ID,fact><ASSIGN><NUM,1><COMMA><ID,n><SEMI>
<FOR><LBRACKET><int> <ID,i><ASSIGN><NUM,1><SEMI><ID,i><LE><ID,n><SEMI><ID,i><INC><RBRACKET><LPAREN>
<ID,fact><ASSIGN><ID,fact><MULT><ID,i><SEMI>
<RPAREN>
<ID,printf><LBRACKET><STRING><COMMA><ID,fact><RBRACKET><SEMI>
<ID,getch><LBRACKET><RBRACKET><SEMI>
<RPAREN>%
(base) nspk@Crisp lab3 %
```

2. Write a lex program to count the number of letters, words, digits and symbols in a given input file (input.txt).

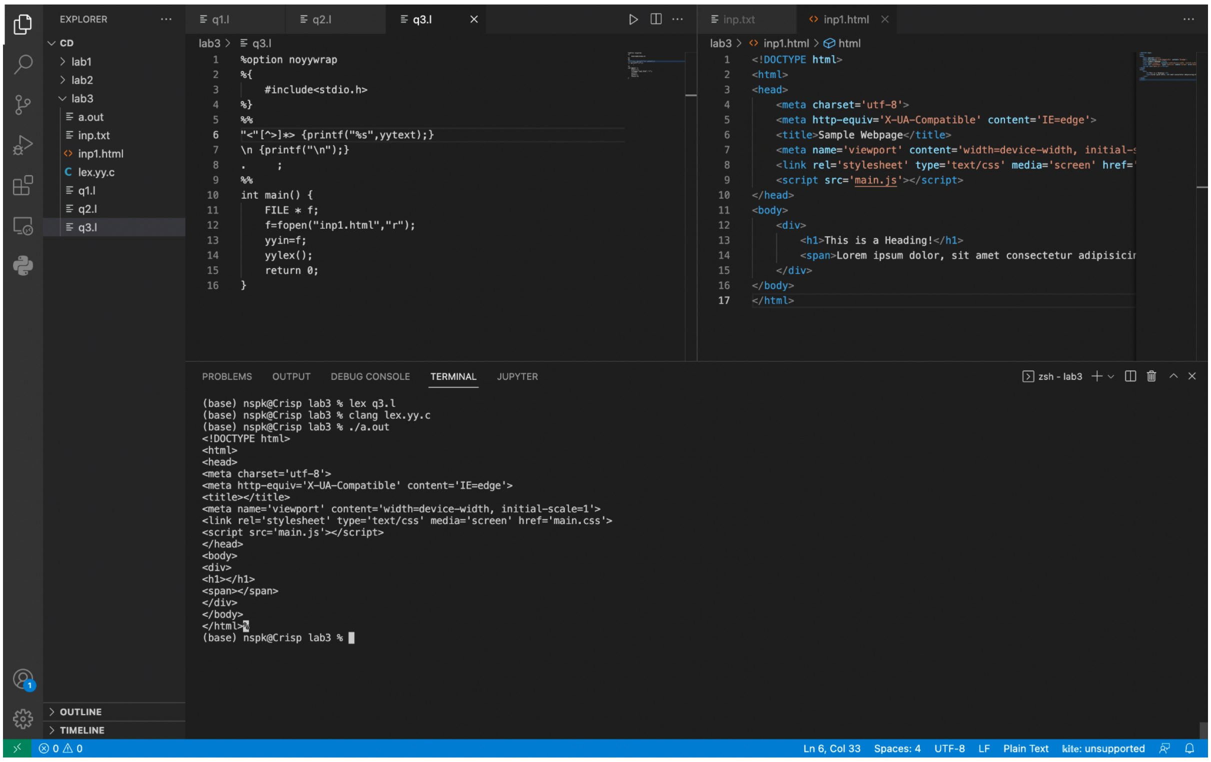


```
%option noyywrap
%{
    #include <stdio.h>
    #include <string.h>
    int letters=0;
    int words=0;
    int digits=0;
    int symbols=0;
%}
%%
[0-9] {digits++;}
[a-zA-Z] {letters++;}
([ ])+ {words++;}
\n {words++;}
. {symbols++;}
%%
int main() {
    FILE * f;
    f=fopen("inp.txt","r");
    yyin=f;
    yylex();
    printf("Number of Letters: %d\n",letters);
    printf("Number of Words: %d\n",words);
    printf("Number of Digits: %d\n",digits);
    printf("Number of Symbols: %d\n",symbols);
    return 0;
}
```

```
#include<stdio.h>
main(){
int fact=1,n;
for(int i=1;i<=n;i++){
fact=fact*i;
}
printf("Factorial Value of N is",fact);
getch();
}
```

```
(base) nspk@Crisp lab3 % lex q2.l
(base) nspk@Crisp lab3 % clang lex.yy.c
(base) nspk@Crisp lab3 % ./a.out
Number of Letters: 78
Number of Words: 14
Number of Digits: 2
Number of Symbols: 34
(base) nspk@Crisp lab3 % 
```

3. Write a LEX program to identify the tags in an HTML file(tag.html)



```
%option noyywrap
%{
    #include<stdio.h>
%}
%%
"<"[^>]*> {printf("%s",yytext);}
\n {printf("\n");}
.     ;
%%
int main() {
    FILE * f;
    f=fopen("inp1.html","r");
    yyin=f;
    yylex();
    return 0;
}
```

```
<!DOCTYPE html>
<html>
<head>
    <meta charset='utf-8'>
    <meta http-equiv='X-UA-Compatible' content='IE=edge'>
    <title>Sample Webpage</title>
    <meta name='viewport' content='width=device-width, initial-s
    <link rel='stylesheet' type='text/css' media='screen' href='
    <script src='main.js'></script>
</head>
<body>
    <div>
        <h1>This is a Heading!</h1>
        <span>Lorem ipsum dolor, sit amet consectetur adipisicir
    </div>
</body>
</html>
```

```
(base) nspk@Crisp lab3 % lex q3.l
(base) nspk@Crisp lab3 % clang lex.yy.c
(base) nspk@Crisp lab3 % ./a.out
<!DOCTYPE html>
<html>
<head>
<meta charset='utf-8'>
<meta http-equiv='X-UA-Compatible' content='IE=edge'>
<title></title>
<meta name='viewport' content='width=device-width, initial-scale=1'>
<link rel='stylesheet' type='text/css' media='screen' href='main.css'>
<script src='main.js'></script>
</head>
<body>
<div>
<h1></h1>
<span></span>
</div>
</body>
</html>%
(base) nspk@Crisp lab3 % 
```