University of Asia Pacific

Department of Computer Science & Engineering

CSE 430: Compiler Design

Lab #1

Write a program for lexical analysis i.e. take a line from file or keyboard and specify each word or character into the following tokens. **USE PYTHON TO DO IT.**

- Any word either combination of characters and digits or combination of characters: **Identifier**.
- Any number : **Constant**
- Single character token :
    - **Parenthesis** : ( ), { }, [ ]
    - **Punctuation** : ;(semicolon) , : (colon) , , (coma)
    - **Arithmetic Operator** : + , -, *, /, =
- **Logical Operator** : >, >=, <, <=, ==, !=
- **Keyword** (if, else, …)

N.B: Remove the comments as well

Sample Input (Console Input / File Input) 1:

```
/*Multi line comment

2nd line

*/



void main()

{

int a, b, c;

//comment

int a = b*c + 10;

if(a!=2)

   a = 0;

}
```

Sample output 1:

Keyword (3): if, int, void
Identifier (4): a, b, c, main
Constant (3): 0, 10, 2
Arithmetic Operator (3): *, +, =
Logical Operator (1): !=
Punctuation (2): , , ;
Parenthesis (4): (,), {, }

Sample Input (Console Input / File Input) 2 (BONUS):

```
void main()

{

   int a, b, c;

   float d = 3.14;

   char ch = 'a';

   //comment

   int a = b*c + 10;

}
```

Sample output 2:

Keyword (4): char, float, int, void
Identifier (6): a, b, c, ch, d, main
Constant (3): 'a', 10, 3.14
Arithmetic Operator (3): *, +, =
Punctuation (2): ,, ;
Parenthesis (4): (,), {, }