

編譯程式 程式設計作業

請使用 **C 語言** 寫出一個 **scanner**，能處理類 C 語言程式碼中的下列語法單元類別：

1. 識別字(IDEN): 第一個字必須是英文字母或底線字元，由英文字母、底線及數字組成，長度不限。
2. 保留字(REWD): if, else, while, for, do, switch, case, default, continue, int, float, double, char, break, static, extern, auto, register, sizeof, union, struct, enum, return, goto, const
3. 整數(INTE): 由數字組成
4. 浮點數(FLOT): 由數字及小數點組成，亦可接受科學記號(如.5, 3., 3.5E+5, 3.5e5, 3.5e-5, 3.5e50) $(+|-)(D^*.D^+|D^+.D^+)(\lambda|((E|e)(+|-)\lambda)D^+)$
5. 字元定字(CHAR): 以' '，如 "，'a'
6. 字串定字(STR): 以"包圍的字元"，字串可以跨行
7. 運算子(OPER): +, -, *, /, =, ,, %, >>, <<, ++, --, +=, -=, *=, /=, %=, !, &&, ||, &, [], |, ^, ., ->, >, <, ==, >=, <=, !=, :, (? , :)
8. 特殊符號(SPEC): {, }, (,), ;
9. 單行註解(SC): 以 // 開頭至行末為止
10. 多行註解(MC): 以 /* 開頭至 */ 為止
11. 前端處理(PREP): 所有前端處理指引
12. 處理 escape character
13. octal literal, hex literal

錯誤處理：

遇有 scan error 情形，程式需標示掃描錯誤的位置，並從下一個可以接受的 token 繼續處理

繳交方法

資料夾名稱:學號 EX : u10016042

資料夾內含：

1. scanner.c (主程式)
2. output.txt

將資料夾以壓縮檔(以學號為檔名)上傳至 ee-class

格式不符者扣 30 分計

程式要求

執行 EXE 檔 scan 自己的原始程式碼 自動創建 output.txt

output.txt 內呈現 scanner 執行後的結果

格式為：原始碼行號\token 類別(以編號表示)\token value

範例:

sample.c

```
//This is a sample program file for a C-like simple scanner
main()
{
/* this is a multiline comment
  Hahaha
  still another line comment */
int i, j = 0;
for (i = 0; i = 100; i++)
    printf("%d\n", i);
}
#include <stdio.h>
```

輸出:

result.txt

```
1  SC  //This is a sample program file for a C-like simple scanner
2  IDEN  main
2  OPER  (
2  OPER  )
3  SPEC  {
4-6 MC
7  REWD  int
7  IDEN  i
7  OPER  ,
7  IDEN  j
未完待續...
11 PREP  #include <stdio.h>
```