



## HW1 作業說明

在第一個作業，你 (妳) 必須使用 Lex 寫出一個 **Pascal scanner**。你 (妳) 的 scanner 要能處理所有的保留字 (reserved words)、識別字 (Identifiers)、符號 (symbols)、數字(Integer、Real)、註解 (comments)、字串常數(quoted string)。如果處理到不合法的 token，你 (妳) 的 scanner 要產生 error message (error detection)。你 (妳) 的 scanner 要盡可能的處理完所有的輸入 (input)。如果發生錯誤，產生 error message，之後能繼續處理其它的輸入 (error recovery)。

Scanner 的輸出 (output) 會列出每個 token 以及此 token 的型態 (integer, real, ID, reserved word, string, symbol)、此 token 所在的行數、第一個字元所在的位置。

- (1). Pascal 是 case-insensitive。例如：保留字 **program**、**ProGram** 和 **PROGRAM** 是一樣的。

保留字列表：absolute and begin break case const continue do else end  
for function if mod nil not object of or program then to var while  
array integer double write writeln string float read array integer double  
write writeln string float

- (2). 識別字 (Identifiers) 字數最長到 15 個字元。一個識別字的第一個字元必須是**英文字母**或是**底線符號** ( ) 起頭，在第一個字元之後，可以是英文字母、數字和底線符號。識別字裡不可包含空白字元。

例如：

 **peter**、**\_db**、**a1** 這些是合法的識別字。

 **1a**、**^db**、**#db**、**abcdefghijklmnopqrstuvwxyz12345** 這些都不是合法的識別字。


- (3). 符號 (symbols)


符號列表：; : ( ) := > < = == <= >= [ ] + - \* / .




- (4). 實數 (Real) 常數可以有正負，且有小數點 (decimal point) 表示法和科學符號表示法兩種。【參考課本第三章習題五】

例如：

 **1.0** 、 **3.14** 、 **7E-2** 、 **12.25e+6** 、 **-7.5E+3** 這些是合法的實數常數。

 **1.00** 、 **03.0** 、 **12.100** 、 **.1** 、 **1.** 這些都不是合法的實數常數。

- (5). 字串常數 (quoted string) 字數最長到 **30** 個字元，你 (妳) 的 **scanner** 如何處理超長的字串常數？'Pascal' 即是字串常數。你(妳)的 **scanner** 必須能處理下列這些字串：


 空字串，即 ""。


'' 代表一個空白。


 'You'll see' 代表字串 You'll see


- (6). 你 (妳) 的 **scanner** 如何處理註解？註解可以跨行。【參考課本第三章習題六】


例如：

 (\* comment \*) 是一個合法的註解。

 (\* comment  
second line \*) 是一個合法的註解。

 (\*\*\*\*\*) 是一個合法的註解。

 (\* a\*\*b \*) 是一個合法的註解。

 (\*ab\*\*)\*) 不是一個合法的註解。

- (7). 你 (妳) 的 **scanner** 如何在發生錯誤後做 **recover**？當錯誤發生，那些字元要被刪除？請寫在 Readme 中



例子：

#### 輸入

```
program test;  
var  
  i : integer;  
begin  
  read(i);  
end;
```

#### 輸出

Line: 1, 1st char: 1, “**program**” is a “**reserved word**”.  
Line: 1, 1st char: 9, “**test**” is an “**ID**”.  
Line: 1, 1st char: 13, “**;**” is a “**symbol**”.  
Line: 2, 1st char: 1, “**var**” is a “**reserved word**”.  
Line: 3, 1st char: 3, “**i**” is an “**ID**”.  
Line: 3, 1st char: 5, “**:**” is a “**symbol**”.  
Line: 3, 1st char: 7, “**integer**” is a “**reserved word**”.  
Line: 3, 1st char: 14, “**;**” is a “**symbol**”.  
Line: 4, 1st char: 1, “**begin**” is a “**reserved word**”.  
Line: 5, 1st char: 3, “**read**” is a “**reserved word**”.  
Line: 5, 1st char: 7, “**(**” is a “**symbol**”.  
Line: 5, 1st char: 8, “**i**” is an “**ID**”.  
Line: 5, 1st char: 9, “**)**” is a “**symbol**”.  
Line: 5, 1st char: 10, “**;**” is a “**symbol**”.  
Line: 6, 1st char: 1, “**end**” is a “**reserved word**”.  
Line: 6, 1st char: 4, “**;**” is a “**symbol**”.

#### 輸入

```
program test;  
var  
  3i : string;  
begin  
  3i := 'ab';  
end;
```

#### 輸出

Line: 1, 1st char: 1, “**program**” is a “**reserved word**”.  
Line: 1, 1st char: 9, “**test**” is an “**ID**”.  
Line: 1, 1st char: 13, “**;**” is a “**symbol**”.  
Line: 2, 1st char: 1, “**var**” is a “**reserved word**”.  
Line: 3, 1st char: 3, “**3i**” is an invalid “**ID**”.  
Line: 3, 1st char: 6, “**:**” is a “**symbol**”.  
Line: 3, 1st char: 8, “**string**” is a “**reserved word**”.  
Line: 3, 1st char: 14, “**;**” is a “**symbol**”.  
Line: 4, 1st char: 1, “**begin**” is a “**reserved word**”.  
Line: 5, 1st char: 3, “**3i**” is an invalid “**ID**”.  
Line: 5, 1st char: 6, “**:=**” is a “**symbol**”.  
Line: 5, 1st char: 9, “**'ab**” is an invalid “**string**”.  
Line: 5, 1st char: 12, “**;**” is a “**symbol**”.  
Line: 6, 1st char: 1, “**end**” is a “**reserved word**”.  
Line: 6, 1st char: 4, “**;**” is a “**symbol**”.



◆你必須繳交下列檔案：

✧ Scanner，檔名為 – 你/妳的學號.I

✧ makefile 檔

✧ 一個 Readme.pdf 檔，裡面包含

1. Lex,版本
2. 作業平台
3. 執行方式
4. 你/妳如何處理這份規格書上的問題
5. 你/妳寫這個作業所遇到的問題
6. 所有測試檔執行出來的結果，存成圖片或文字檔

請用壓縮軟體將上述這些檔案壓縮成一個檔案，檔名為 – 你/妳的學號\_hw1

◆評分標準(因應老師要求，進行小調整)

1.pass (15%)(全對才給分) 、 2.pass (15%) (全對才給分) 、  
m.pass (20%)(會部份給分) 、 n.pass (20%)( 會部份給分) 、  
助教測資 (20%)(會部份給分) 、 Readme.pdf (8%) 、 makefile(2%)

◆範例檔使用方式

安裝好 lex 後，下載範例檔，在 example 的目錄下輸入 `$ make all`，即可產生 a.out(執行檔)

執行 `a.out < test.in`，可看到輸出結果