



## Lex Scanner 評分標準與 DEMO

本次作業滿分為：100 分（佔總成績 15%）

- 公開測資：test1.java, test2.java, test3.java (各 15%)
- 隱藏測資：test4.java, test5.java (各 15%)
- Symbol Table function 實作：10%
- 口頭問答：10%
- Readme.pdf：5%

你（妳）的 Java Scanner 必須能處理以下的情形（同作業規格書）：

### 1. 將來會被傳給 Parser 的 tokens：

- I. **Symbols**：『,』、『:』、『;』、『(』、『[』、『{』
- II. **Operators**：『++』、『--』、『\*』、『/』、『%』、『=』、『<=』、『>=』、『!=』、『&&』、『||』、『!』
- III. **Reserved word**：boolean, break, byte, case, char, catch, class, const, continue, default, do, double, else, extends, false, final, finally, float, for, if, implements, int, long, main, new, print, private, protected, public, return, short, static, string, switch, this, true, try, void, while
- IV. **Identifier(ID)**：Java 定義識別字可用任何 Unicode 符號，但習慣上會以英文二十六大小寫字母為主、另加上數字、底線符號及 \$。  
例如：  
✚ peter 、 \_db 、 a1 這些是合法的 ID。  
✚ 1a 、 ^db 、 #db 這些是不合法的 ID。
- V. **Integer**：整數，有正負之分。
- VI. **Float**：浮點數，有正負之分、另有小數點表示與科學符號表示法。  
例如：  
✚ 1.0 、 3.14 、 7E-2 、 12.25e+6 、 -7.5E+3 、 -7.5E+3f 這些是合法的。  
的。

VII. **String**：字串字元，必須由兩個雙引號構成 string constant。

例如：

✚ 空字串(“”)、空白(“ ”)、"You\’ll see" 代表字串 You’ll see、“\’” 這些是**合法**的字串。

✚ 單引號(‘aaa’)、沒有跳脫字元(“”)、“\’、只有一個雙引號(“ab) 這些是**不合法**的字串。

2. 將來不會傳給 Parser 的 tokens：

I. **Whitespace**：空白、Tabs、換行。

II. **Comments**：註解分可跨行之『/\*與\*/』與不可跨行之『//』。

3. Recovery：

I. 你的 Scanner 可處理完所有 Input，不能遇到錯誤停下來。

4. Symbol Table：

I. 你的 Symbol Tables 可存所有 ID 並在 Scan 結果 Output 所有 ID，而相同名稱的 ID 不能重複印出。

II. 包含 4 個 function：create(), lookup(), insert(), dump)。