**ScannerTest**

### **Description**

在編譯器中， Token 是組成程式碼的最小單位，需要由 Scanner 先把 Input text 轉換成 Token，Parser 才會進行後續的工作。請利用表 1.，撰寫一個 Scanner 來取得 Token 並輸出。

### **Input Format**

輸入一行程式碼，每一行程式碼換行皆以 ‘\n’ 表示。

部分輸入的程式碼將會有不符合 Regular Expression 的情況發生。

### **Output Format**

若測試資料符合規定，請輸出該行程式碼所有 tokens 的型別並以換行隔開。

若不符合，即使只出現一個錯誤，也只印出"invalid input"。

表 1.

|  |  |
| --- | --- |
| **Terminal** | **Regular Expression** |
| ID | [A-Za-z\_][A-Za-z0-9\_]\* |
| STRLIT | “[^”]\*” |
| LBR | \( |
| RBR | \) |
| DOT | \. |
| SEMICOLON | ; |

|  |  |
| --- | --- |
| **Sample Input1**  Str.length("123");  **Sample Output1**  ID Str  DOT .  ID length  LBR (  STRLIT "123"  RBR )  SEMICOLON ; | **Sample Input2**  Str.length("123");  1  **Sample Output2**  invalid input |
| **Sample Input3**  "string;  **Sample Output3**  invalid input | **Sample Input4**  "test\_string";  Test\_ID;  **Sample Output4**  STRLIT "test\_string"  SEMICOLON ;  ID Test\_ID  SEMICOLON ; |

在範例2中，雖然「Str.length("123");」可以被正確分割，但「1」沒有定義，所以輸出"invalid input"。

在範例3中，「Strlint」Token應該要有兩個「"」，一個「"」沒有定義，所以輸出"invalid input"。