

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	;

<p>Sample Input1</p> <p>Str.length("123");</p> <p>Sample Output1</p> <p>ID Str DOT . ID length LBR (STRLIT "123" RBR) SEMICOLON ;</p>	<p>Sample Input2</p> <p>Str.length("123"); 1</p> <p>Sample Output2</p> <p>invalid input</p>
<p>Sample Input3</p> <p>"string;</p> <p>Sample Output3</p> <p>invalid input</p>	<p>Sample Input4</p> <p>"test_string"; Test_ID;</p> <p>Sample Output4</p> <p>STRLIT "test_string" SEMICOLON ; ID Test_ID SEMICOLON ;</p>

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"。