

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

輸入多行程式碼，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。
NUM值的範圍： -2147483647~2147483647

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"。
另外OP前後允許有空格，例如(1+2+3)與(1 + 2 + 3)皆可以

表 1.

Terminal	Regular Expression
ID	[A-Za-z_][A-Za-z0-9_]*
SEMICOLON	;
LBR	\(
RBR	\)
OP	[+ - * / =]
NUM	-?[1-9][0-9]*

<p>Sample Input1</p> <p>HELLO87 (1 + 2 + 3)</p> <p>Sample Output1</p> <p>ID HELLO87 NUM 1 LBR (OP + NUM 2 OP + NUM 3 RBR)</p>	<p>Sample Input2</p> <p>HELLO87 (1 + 2 + 3) \$</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>INTEGER = 11111</p> <p>Sample Output4</p> <p>ID INTEGER OP = NUM 11111</p>

在範例2中，雖然「HELLO87 (1 + 2 + 3)」可以被正確分割，但「\$」沒有定義，所以輸出"invalid input"。

在範例3中，「"」沒有定義，所以輸出"invalid input"。