

- ★ 最多 3 人一組，用自己熟悉的程式語言撰寫
- ★ 輸入一般 .txt 文字檔（可於執行時輸入檔名），切出各 token，判斷各 token 類別
- ★ 輸出 .txt 文字檔，若輸入檔名為 test.txt，輸出檔名為 scanner_test.txt
- ★ 計算 token 總數及各類別 token 個數，列於輸出檔最後端
- ★ 找出 undefined token，從下一行開始再次 scan，同時列出 skipped token
- ★ 期末 demo 並繳交書面報告，說明程式架構及 token 判斷方式

★ Token 類別

類別	格式	說明
Reserved word	include main char int float if else elseif for while do return switch case printf scanf	允許各式大小寫組合
Library name	<xxx.h>	前面必須存在 #include，否則視為 undefined token
Identifier	char/int/float 有定義的變數	由英文字母、底線及數字組成，限定英文字母開頭，否則視為 undefined token
Character	'*'	'*' 算 1 個 token
Number	整數/浮點數/負數 負數可以表示成 -3 或 (-3)	算 1 個 token
Pointer	*identifier	*identifier 算 1 個 token
Bracket	() [] { }	
Operator	+ - * / % ^ ++ -- & =	
Comparator	== < > <= >= !=	
Address	&identifier identifier 必須已定義	&identifier 算 1 個 token
Punctuation	, ; : # " '	
Format specifier	%d %f %c *	
Printed token	printf 或 scanf 裡直接印出的 token	以空格區隔
Comment	/* ... */ 或 // ...	不管內容，都算 1 個 token
Undefined token		
Skipped token		不管內容，都算 1 個 token

```

#include<stdio.h>
int i, k, m, *id;
for (i = 1; i < 10; i++) {
    k = k + i;
    m = 100 - m;
}
printf("xxx yyy %d\n", k); // This is a comment
/* This is also a comment. */

```

-
1. **token # belongs to** punctuation
 2. **token include belongs to** reserved word
 3. **token <stdio.h> belongs to** library name
 4. **token int belongs to** reserved word
 5. **token i belongs to** identifier
 6. **token , belongs to** punctuation
 7. **token k belongs to** identifier
 - ...
 11. **token *id belongs to** pointer
 12. **token ; belongs to** punctuation
 13. **token for belongs to** reserved word
 - ...
 23. **token i belongs to** identifier
 24. **token ++ belongs to** operator
 - ...
 40. **token printf belongs to** reserved word
 41. **token (belongs to** bracket
 42. **token " belongs to** punctuation
 43. **token xxx belongs to** printed token
 44. **token yyy belongs to** printed token
 45. **token %d belongs to** format specifier
 46. **token \n belongs to** format specifier
 - ...
 52. **token // This is a comment belongs to** comment
 53. **token /* This is also a comment. */ belongs to** comment

Total: 53 tokens

Reserved word: 4

include

int

for

printf

Library name: 1

<stdio.h>

Identifier: 12

i (x5)

k (x4)

m (x3)

Number: 3

1

10

100

Pointer: 1

*id

Bracket: 6

((x2)

) (x2)

{

}

Operator: 6

= (x3)

++

+

-

Comparator: 1

<

Punctuation: 13

#

, (x4)

; (x6)

" (x2)

Format specifier: 2

%d

\n

Printed token: 2

xxx

yyy

Comment: 2

// This is a comment

/* This is also a comment. */

```

<stdio.h>
int ia, *ie, forif;
char ca;
ia = (*ie) + ib; ia = ia * 2;
ca = 'A';
whilefor (ia<3) forif = forif + 1;
sCAnf("%c", &cb);
ia = *forif + (-2);
ia = 5 * ie;

```

1. **token** < **belongs to** comparator
2. **token** stdio.h **belongs to** undefined token
3. **token** > **belongs to** skipped token
4. **token** int **belongs to** reversed word
5. **token** ia **belongs to** identifier
6. **token** , **belongs to** punctuation
7. **token** *ie **belongs to** pointer
8. **token** , **belongs to** punctuation
9. **token** forif **belongs to** identifier
10. **token** ; **belongs to** punctuation
11. **token** char **belongs to** reversed word
12. **token** ca **belongs to** identifier
13. **token** ; **belongs to** punctuation
14. **token** ia **belongs to** identifier
15. **token** = **belongs to** operator
16. **token** (**belongs to** bracket
17. **token** *ie **belongs to** pointer
18. **token**) **belongs to** bracket
19. **token** + **belongs to** operator
20. **token** ib **belongs to** undefined token
21. **token** ; ia = ia *2; **belongs to** skipped token
22. **token** ca **belongs to** identifier
23. **token** = **belongs to** operator
24. **token** 'A' **belongs to** character
25. **token** ; **belongs to** punctuation
26. **token** whilefor **belongs to** undefined token
27. **token** (ia<3) forif = forif + 1; **belongs to** skipped token
28. **token** sCAnf **belongs to** reversed word

29. **token (belongs to** bracket
30. **token " belongs to** punctuation
31. **token %c belongs to** format specifier
32. **token " belongs to** punctuation
33. **token , belongs to** punctuation
34. **token &cb belongs to** undefined token
35. **token); belongs to** skipped token
36. **token ia belongs to** identifier
37. **token = belongs to** operator
38. **token * belongs to** operator
39. **token forif belongs to** identifier (可將*forif 判斷為 undefined token)
40. **token + belongs to** operator
41. **token (-2) belongs to** number
42. **token ; belongs to** punctuation
43. **token ia belongs to** identifier
44. **token = belongs to** operator
45. **token 5 belongs to** number
46. **token *ie belongs to** pointer (可分別判斷為 operator 和 undefined token)
47. **token ; belongs to** punctuation

Total: 47 tokens

Reserved word: 3

int

char

sCAnf

Identifier: 8

ia (x4)

forif (x2)

ca (x2)

Character: 1

'A'

Number: 2

(-2)

5

Pointer: 3

*ie (x3)

Bracket: 3

((x2)

)

Operator: 7

= (x4)

+ (x2)

*

Comparator: 1

<

Punctuation: 10

, (x3)

; (x5)

" (x2)

Format specifier: 1

%c

Undefined token: 4

stdio.h

ib

whilefor

&cb

Skipped token: 4

>

; ia = ia *2;

(ia<3) forif = forif + 1;

)