

如何 (正確的) "抓下一個 token"

當 scanner 要去「真正的 input」"抓"「下一個 token」時(不管是因為 parser 在向 scanner 要求下一個 token 還是 whatever)，scanner 應該如此做：

從下一個(尚未讀進來的(或是「已讀進來但尚未處理」的))字元開始

// 反正就是在概念上尚未處理到它的那個渾蛋字元

```
// 隨時要注意 END-OF-FILE ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! !
```

一、跳過所有 white-spaces、直到碰到「非 white-space」(或 END-OF-FILE)為止

二、從那個「非 white-space」開始 "繼續讀入"、直到碰到「separator-字元」為止

(如果一開始的那個「非 white-space」已經是個 separator ,
then you are already done(翻譯：別再讀了！)
)

// 小心哪！後來碰到的那個「separator-字元」就是「已讀進來但尚未處理」！

三、(此時你已讀入一個 **token** 了，問題只是「它是誰」而已！
所以現在(剩下)要做的是「判定」的工作)

判定此 token 是啥！

// 注意：「什麼都不是的」就是 SYMBOL。

工作完畢。

註： separator df= white-space OR '(' OR ')' OR single-quote ('\')

OR double-quote ('\"')

OR ';' OR ','

加註一： '!'不是 separator

加註二：'!'可以是 SYMBOL 的一部分，'!'甚至可以是一個 SYMBOL 的開頭，

但 SYMBOL 不能光是一個'!'。

(如果光是一個'!'，那就是所謂的 DOT、而不是 symbol)