# Tokens Types: Lexical Convention of C-

1. Keyword

* while, if, else, int, void, return

1. Special symbols

* + - \* /
* < <= > >= == !=
* =
* ;
* ,
* ( ) [ ] { }
* /\* \*/

1. Identifier (ID = letter letter\*)
2. Number (NUM = digit digit\*)

# Things that should be ignored

1. Whitespace

* blank newline tab

1. Comments

* surrounded by /\* … \*/
* more than one line
* not nested

How to process whitespaces

1. Ignored

* x + y
* xyz = 3456

1. Not ignored

* ID ID
* keyword keyword
* ID NUM
* NUM NUM
* > = (GT ASSIGN)

|  |  |  |
| --- | --- | --- |
| Type | Appearance | Token |
| Keyword | while | WHILE |
| if | IF |
| else | ELSE |
| int | INT |
| void | VOID |
| return | RETURN |
| Special symbol | + | ADD |
| - | SUB |
| \* | MUL |
| / | DIV |
| < | LT |
| <= | LTE |
| > | GT |
| >= | GTE |
| == | EQ |
| != | NEQ |
| = | ASSIGN |
| ; | ENDS |
| , | COMMA |
| ( | LP |
| ) | RP |
| [ | LSB |
| ] | RSB |
| { | LCB |
| } | RCB |
| “/\*” *any*\* “\*/” | COMMENT |
| ID | *letter letter*\* | ID |
| NUM | *digit digit*\* | NUM |