

Token, white space, Comments, variable, Parser

Subject: C language

Date: 21/12/2025

What is Tokens ?

⇒ A token is the smallest individual unit of a program that the Compiler recognizes while converting Source Code into machine Code.

• The text of C/C++ program consists of tokens and white space.

What is Parser ?

⇒ Parser is a program that analyzes the sequence of tokens to check whether they follow the grammatical rules of the language.

OR

Parser checks the syntax of a program using tokens generated by the lexical analyzer.

What is Lexical Analyzer ?

⇒ A lexical analyzer is the first phase of a Compiler that reads the source code character by character and converts it into tokens.

Type of Tokens?

⇒ 1.) keyword

- Reserved words with predefined meaning.
- Cannot be used as identifiers.

Example:- int, float, if, else, while, for, return.

2.) Identifiers

- Names given to variables, functions, array etc.
- Defined by the programmer.

Rules:

- Must start with a letter or underscore (_).
- Cannot be a keyword.

Example:- Sum, total_marks, _Count

3.) Constants

- Fixed values that do not change.

Two types of Constants:-

1) Primary Constants

i.) Integer Constants.

Example:- 10, -5, 100.

ii.) Floating (Real) Constants.

Example:- 3.14, 2.5

iii.) Character Constants

Example — 'A', 'a'

NOTE:- Single digit ya single character ko jab hum single quote me likhte hai, tab use hum Character Constant kahte hain.

2.) Secondary Constants

i.) Array Constants

Example — {1, 2, 3, 4}

ii.) String Constants

Example — "Hello"

iii.) Pointer Constants

Example — Address values like &a

iv.) Structure Constants

v.) Union Constants

vi.) Enumerator Constants

4.) Operators

• Symbols used to perform operations.

Example :- +, -, *, /, %, ==, >=, <=, >, <, &&, ||, etc.

Subject: _____

Date: ____/____/____

5.) Punctuators

- Symbols that separate or structure the program.

Example :- `;`, `()`, `{ }`, `[]`

6.) String Literals

- Sequence of characters inside double quotes.

Example :- "Computer Science"

What is white space?

⇒ Whitespace refers to blank characters that are used to separate tokens and improve readability of the program.

Example :-
Tabs (`\t`)
New lines (`\n`)
Blanks (`' '`)
Form Feed (`\f`)

Is whitespace compulsory for separating tokens in C?

⇒ Whitespace is not always necessary for token separation. The Compiler can separate tokens using operators and punctuators, but whitespace is required where tokens may combine and change meaning.

Comments

- Comments are text ignored by Compiler.
- Comments are useful for documentation of your code and useful for programmers.

There are two types of Comments :-

i) Single line Comment

e.g; // Hello world

ii) Multiple line Comment

e.g; /*

Hello world

Hello world?

*/

What are variables?

⇒ Variable is a name of memory location where we store program's data during execution of program.

- Variable name is any combination of alphabets (a to z or A to Z), digits (0 to 9) and underscore (_).

- No other symbol is allowed.

- Valid variable name cannot start from a digit.

- Variable name cannot be a keyword.