C is a middle level language. As a middle level language, C allows the manipulation of bits, bytes and addresses (well suited for system-level programming) – the basic elements with which the computer functions. Despite this fact C code is very portable. Portability means that it is easy to adapt software written for one type of computer or operating system to another. For example, you can easily convert a program written for DOS so that it runs under Windows, that program is portable.

|  |  |
| --- | --- |
| High Level | Ada, Modula-2, Pascal, COBOL, FORTRAN, BASIC |
| Middle Level | Java, C++, C, FORTH |
| Lowest Level | Macro-assembler, Assembler |

All high-level programming languages support the concept of data types.

A ***data type*** defines a set of values that a variable can store along with a set of operations that can be performed on that variable.

Common data types are integer, character, and real.

**Note: C is not a strongly typed language, as are Pascal and Ada. C permits almost all type conversions. For Example, you may free intermix character and integer types in an expression.**

**Unlike a high level language, C performs almost no runtime error checking. For Example, no check is performed to ensure that array boundaries are not overrun. These types of checks are the responsibility of the programmer.**

C is a structured language is not, technically, a block-structured language.

|  |  |
| --- | --- |
| Block structure language | Block structures languages permit procedures or functions to be declared inside other procedures or functions. Since C does not allow the creation of functions within functions, it cannot formally be called block-structured. |
| Structure language | The distinguishing feature of a structure language is compartmentalization of code and data. This is the ability of a language to section off and hide from the rest of the program all information and instructions necessary to perform a specific task.  A structured language allows you a variety of programming possibilities. It directly supports several loop constructs, such while, do-while, and for. In structure language, the use of goto is either prohibited or discouraged and is not the common form of program control. |

The 32 keywords of C language ( 27 were defined by the original version of C. Five were added by the ANSI C committee - enum, const, signed, void and volatile)

**Note: all C and C++ keywords are in lower case.**

|  |  |  |  |
| --- | --- | --- | --- |
| **32 keywords defined by standard C** | | | |
| auto | double | Int | struct |
| break | else | Long | switch |
| case | **enum** | Register | typedef |
| char | extern | Return | union |
| **const** | float | Short | unsigned |
| continue | for | **Signed** | **void** |
| default | goto | Sizeof | **volatile** |
| do | if | Static | while |

**Note:** neither C nor C++ provides any keyword that performs such things as input/output (I/O) operations, high level mathematical computations, or character handling. As a result, most programs include calls to various functions contained in the ***standard library***.

All C++ compilers come with a standard library of functions that perform most commonly needed tasks.

**Expression:** expressions are formed from atomic elements: data and operations. Data may be represented either by variables or by constants.

**There are five atomic data types in C**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Character (**char**) | Integer (**int**) | Floating-point (**float**) | Double floating-point (**double**) | Value less (**void**) |

All other data types in C are based upon one of these types.

**Note:** The size and range of these data types may vary between processor types and compilers. However, in all cases a character is 1 byte.

The size of an integer is usually the same as the word length of the execution environment of the program. For most 16-bit environments, such as DOS or Windows 3.1, an integer is 16 bits. For most 32-bit environments, such as Windows NT, an integer is 32 bits.

Both C and C++ only stipulate the ***minimum range*** of each data type, not its size in bytes.

**Note: To the five basic data types defined by C, C++ add two more: bool and wchar\_t**

**Standard C++ does not specify a minimum size or range for the basic types. Instead it simply states that they must meet the certain requirements. For example, standard C++ states that an int will “have the natural size suggested by the architecture of the execution environment.” In all cases this will meet or exceed the minimum ranges specified by Standard C.**

**Each C++ compiler specified the size and range of the basic types in the header <climits>**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Type** | **Typical size in bits** | **Minimal range** | **Comments** |  |
| char | 8 | -127 to 127 | Value of type char are generally used to hold values defined by the ASCII character set |  |
| unsigned char | 8 | 0 to 255 |  |  |
| signed char | 8 | -127 to 127 |  |  |
| int | 16 or 32 | -32767 to 32767 | Integers will generally correspond to the natural size of a word on the host computer |  |
| unsigned int | 16 or 32 | 0 to 65535 |  |  |
| signed int | 16 or 32 | Same as int |  |  |
| short int | 16 | -32767 to -32767 |  |  |
| unsigned short int | 16 | 0 to 65535 |  |  |
| signed short int | 16 | Same as short int |  |  |
| long int | 32 | -2147483647 to -2147483647 |  |  |
| signed long int | 32 | Same as long int |  |  |
| unsigned long int | 32 | 0 to 4294967295 |  |  |
| float | 32 | Six digits of precision | Exact format of floating-point values will depend upon how they are implemented. | The range of float and double will depend upon the method used to represent the floating point numbers |
| double | 64 | Ten digits of precision |  |  |
| long double | 80 | Ten digits of precision |  |  |
| void |  |  | Void either explicitly declares a function returning no value or creates generic pointers. |  |

Note: values outside that range may be handled differently by different compilers.

**Modifying the basic Types**

Except for type void, the basic data types may have various modifiers preceding them. Modifiers can be used to alter the meaning of the base type to fit various situations more precisely.

|  |
| --- |
| **List of modifiers** |
| signed |
| unsigned |
| long |
| short |
|  |

Note: The use of signed on integers is allowed, but redundant because the default integer declaration assumes a signed number. The most important use of signed is to modify char in implementations in which char is unsigned by default.

The difference between signed and unsigned integers is in the way that the high order bit if the integer is interpreted. If you specify a signed integer, the compiler generates the code that assumes that the high-order bit of an integer is to be used as a ***sign flag***.

If the sign flag is 0, the number is positive; if it is 1, the number is negative.

In general, negative numbers are represented using the two’s complement approach, which reverses all bits in the number (except the sign flag), adds 1 to this number, and set the sign flag 1.

Example:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Type | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 | 8192 | 16384 | Sign flag | No |
|  | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |  |
| Signed int | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  | -1 |
| Unsigned int | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 65535 |

***Identifiers:*** *in C/C++, the names of variables, functions, labels, and various other user-defined objects are called identifiers.*

*Note: The first characters must be a letter or underscore and subsequent characters must be either letters, digits, or underscores.*

***Correct incorrect***

*Count 1count*

*test23 hi!there*

*high\_balance high…balance*

In C, identifiers may be of any length. If the identifier will be involved in an external link process, then at least the first six characters will be significant. These identifiers, called ***external names***, including function names and global variables that are shared between files.

If the identifier is not used in an external link process, then at least the first 31 characters will be significant. This type of identifier is called an ***internal name*** and includes the name of local variables.

In C++, there is no limit to the length of an identifier, and at least the first 1024 characters are significant.

In an identifier, upper and lowercase are treated as distinct. Hence count, Count and COUNT are three separates identifiers.

An identifier cannot be the same as a C or C++ keyword, and should not have the same name as functions that are in the C or C++ library.

**Variables:** local variables, formal parameters and global variables

**Local variable:** variables that are declared inside a function are called local variables. These variables are referred to as ***automatic*** variables.

The C language contains the keyword **auto**, which can be used to declare local variables. However, since all non-global variables are, by default, assumed to be **auto**. This keyword virtually never used. It has been said that **auto** was included in C to provide for source-level compatibility with its predecessor B. Further, **auto** is supported in C++ to provide compatibility with C.

There is an important difference between C and C++ as to where you can declare local variables. In C, you must declare local variables at the start of the block in which they are defined, prior to any “action” statement.

Local variables are stored on the stack.

**Formal Parameters**: If a function is to use arguments, it must declare variables that will accept the values of the arguments. These variables are called the formal parameters of the function. As local variables, they are also dynamic and are destroyed upon exit from the function.

**Global Variables:** global variables are known throughout the program and may be used by any peace of code. If a global variable and a local variable have the same name, all references to that variable name inside the code block in which the local variable is declared will refer to that local variable and have no effect on the global variable.

Storage for global variables is in a fixed region of memory set aside for this purpose by the compiler.

**Access Modifiers**: There are two modifiers that control how variables may be accessed or modified. These qualifies are **const** and **volatile**.

**const:** The compiler is free to place variables of this type into read-only memory(ROM)