

# C++ Programming Language

## Pointers, References and Dynamic Memory Allocation

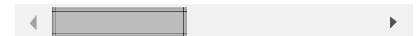
Pointers, References and Dynamic Memory Allocation are the most powerful features in C/C++ language, which allows programmers to directly manipulate memory to *efficiently manage the memory* - the most critical and scarce resource in computer - *for best performance*. However, "pointer" is also the most complex and difficult feature in C/C++ language.

Pointers are extremely powerful because they allows you to access addresses and manipulate their contents. But they are also extremely complex to handle. Using them correctly, they could greatly improve the efficiency and performance. On the other hand, using them incorrectly could lead to many problems, from unreadable and un-maintainable codes, to infamous bugs such as memory leaks and buffer overflow, which may expose your system to hacking. Many new languages (such as Java and C#) remove pointer from their syntax to avoid the pitfalls of pointers, by providing automatic memory management.

Although you can write C/C++ programs without using pointers, however, it is difficult not to mention pointer in teaching C/C++ language. Pointer is probably not meant for novices and dummies.

### TABLE OF CONTENTS (HIDE)

1. Pointer Variables
  - 1.1 Pointer Variables (or Pointers)
  - 1.2 Declaring Pointers
  - 1.3 Initializing Pointers via the Address
  - 1.4 Indirection or Dereferencing C
  - 1.5 Pointer has a Type Too
  - 1.6 Uninitialized Pointers
  - 1.7 Null Pointers
2. Reference Variables
  - 2.1 References (or Aliases) (&)
  - 2.2 How References Work?
  - 2.3 References vs. Pointers
  - 2.4 Pass-By-Reference into Function
  - 2.5 Passing the Function's Return
  - 2.6 Summary
3. Dynamic Memory Allocation
  - 3.1 new and delete Operators
  - 3.2 new[] and delete[] Operators
4. Pointer, Array and Function
  - 4.1 Array is Treated as Pointer
  - 4.2 Pointer Arithmetic
  - 4.3 sizeof Array
  - 4.4 Passing Array In/Out of a Function
  - 4.5 Pass-by-Reference and sizeof
  - 4.6 Operating on a Range of an Array
  - 4.7 C-String and Pointer
5. \*More On Pointers
  - 5.1 Function Pointer
  - 5.2 Generic Pointer or void Pointer
  - 5.3 Constant Pointer vs. Constant



### 1. Pointer Variables

A *computer memory location* has an *address* and holds a *content*. The *address* is a numerical number (often expressed in hexadecimal), which is hard for programmers to use directly. Typically, each address location holds 8-bit (i.e., 1-byte) of data. It is entirely up to the programmer to interpret the meaning of the data, such as integer, real number, characters or strings.

To ease the burden of programming using numerical address and programmer-interpreted data, early programming languages (such as C) introduce the concept of variables. A variable is a *named* location that can store a *value* of a particular *type*. Instead of numerical addresses, names (or identifiers) are attached to certain addresses. Also, types (such as `int`, `double`, `char`) are associated with the contents for ease of interpretation of data.

Each address location typically hold 8-bit (i.e., 1-byte) of data. A 4-byte `int` value occupies 4 memory locations. A 32-bit system typically uses 32-bit addresses. To store a 32-bit address, 4 memory locations are required.

The following diagram illustrate the relationship between computers' memory *address* and *content*; and variable's *name*, *type* and *value* used by the programmers.

Computer		Programmers		
Address	Content	Name	Type	Value
90000000	00	sum	int (4 bytes)	000000FF (255 <sub>10</sub> )
90000001	00			
90000002	00			
90000003	FF			
90000004	FF	age	short (2 bytes)	FFFF (-1 <sub>10</sub> )
90000005	FF			
90000006	1F			
90000007	FF			
90000008	FF	average	double (8 bytes)	1FFFFFFFFFFFFFFFFF (4.45015E-308 <sub>10</sub> )
90000009	FF			
9000000A	FF			
9000000B	FF			
9000000C	FF			
9000000D	FF			
9000000E	90			
9000000F	00			
90000010	00	ptrSum	int* (4 bytes)	90000000
90000011	00			

Note: All numbers in hexadecimal

## 1.1 Pointer Variables (or Pointers)

A *pointer variable* (or *pointer* in short) is basically the same as the other variables, which can store a piece of data. Unlike normal variable which stores a value (such as an int, a double, a char), a *pointer stores a memory address*.

## 1.2 Declaring Pointers

Pointers must be declared before they can be used, just like a normal variable. The syntax of declaring a pointer is to place a \* in front of the name. A pointer is associated with a type (such as int and double) too.

```
type *ptr;    // Declare a pointer variable called ptr as a pointer of type
// or
type* ptr;
// or
type * ptr;  // I shall adopt this convention
```

For example,

```
int * iPtr;    // Declare a pointer variable called iPtr pointing to an int (an int pointer)
               // It contains an address. That address holds an int value.
double * dPtr; // Declare a double pointer
```

Take note that you need to place a \* in front of each pointer variable, in other words, \* applies only to the name that followed. The \* in the declaration statement is not an operator, but indicates that the name followed is a pointer variable. For example,

```
int *p1, *p2, i;    // p1 and p2 are int pointers. i is an int
int* p1, p2, i;     // p1 is a int pointer, p2 and i are int
int * p1, * p2, i;  // p1 and p2 are int pointers, i is an int
```

**Naming Convention of Pointers:** Include a "p" or "ptr" as *prefix* or *suffix*, e.g., iPtr, numberPtr, pNumber, pStudent.

## 1.3 Initializing Pointers via the Address-Of Operator (&)

When you declare a pointer variable, its content is not initialized. In other words, it contains an address of "somewhere",

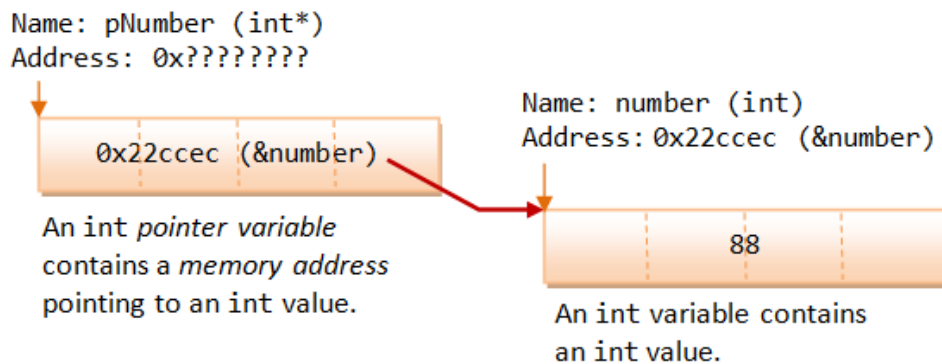
which is of course not a valid location. This is dangerous! You need to initialize a pointer by assigning it a valid address. This is normally done via the *address-of operator* (&).

The *address-of operator* (&) operates on a variable, and returns the address of the variable. For example, if `number` is an `int` variable, `&number` returns the address of the variable `number`.

You can use the address-of operator to get the address of a variable, and assign the address to a pointer variable. For example,

```
int number = 88;           // An int variable with a value
int * pNumber;             // Declare a pointer variable called pNumber pointing to an int (or int pointer)
pNumber = &number;         // Assign the address of the variable number to pointer pNumber

int * pAnother = &number; // Declare another int pointer and init to address of the variable number
```



As illustrated, the `int` variable `number`, starting at address `0x22ccec`, contains an `int` value `88`. The expression `&number` returns the address of the variable `number`, which is `0x22ccec`. This address is then assigned to the pointer variable `pNumber`, as its initial value.

The address-of operator (&) can only be used on the RHS.

## 1.4 Indirection or Dereferencing Operator (\*)

The *indirection operator* (or *dereferencing operator*) (\*) operates on a pointer, and returns the value stored in the address kept in the pointer variable. For example, if `pNumber` is an `int` pointer, `*pNumber` returns the `int` value "pointed to" by `pNumber`.

For example,

```
int number = 88;
int * pNumber = &number; // Declare and assign the address of variable number to pointer pNumber (0x22ccec)
cout << pNumber << endl; // Print the content of the pointer variable, which contain an address (0x22ccec)
cout << *pNumber << endl; // Print the value "pointed to" by the pointer, which is an int (88)
*pNumber = 99;           // Assign a value to where the pointer is pointed to, NOT to the pointer variable
cout << *pNumber << endl; // Print the new value "pointed to" by the pointer (99)
cout << number << endl;  // The value of variable number changes as well (99)
```

Take note that `pNumber` stores a memory address location, whereas `*pNumber` refers to the value stored in the address kept in the pointer variable, or the value pointed to by the pointer.

As illustrated, a variable (such as `number`) *directly* references a value, whereas a pointer *indirectly* references a value through the memory address it stores. Referencing a value indirectly via a pointer is called *indirection* or *dereferencing*.

The indirection operator (\*) can be used in both the RHS (`temp = *pNumber`) and the LHS (`*pNumber = 99`) of an assignment statement.

Take note that the symbol `*` has different meaning in a declaration statement and in an expression. When it is used in a declaration (e.g., `int * pNumber`), it denotes that the name followed is a pointer variable. Whereas when it is used in an expression (e.g., `*pNumber = 99`; `temp << *pNumber`), it refers to the value pointed to by the pointer variable.

## 1.5 Pointer has a Type Too

A pointer is associated with a type (of the value it points to), which is specified during declaration. A pointer can only hold

an address of the declared type; it cannot hold an address of a different type.

```
int i = 88;
double d = 55.66;
int * iPtr = &i;    // int pointer pointing to an int value
double * dPtr = &d; // double pointer pointing to a double value

iPtr = &d;    // ERROR, cannot hold address of different type
dPtr = &i;    // ERROR
iPtr = i;    // ERROR, pointer holds address of an int, NOT int value

int j = 99;
iPtr = &j;    // You can change the address stored in a pointer
```

### Example

```
1  /* Test pointer declaration and initialization (TestPointerInit.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int main() {
6      int number = 88;    // Declare an int variable and assign an initial value
7      int * pNumber;      // Declare a pointer variable pointing to an int (or int pointer)
8      pNumber = &number;  // assign the address of the variable number to pointer pNumber
9
10     cout << pNumber << endl; // Print content of pNumber (0x22ccf0)
11     cout << &number << endl; // Print address of number (0x22ccf0)
12     cout << *pNumber << endl; // Print value pointed to by pNumber (88)
13     cout << number << endl;   // Print value of number (88)
14
15     *pNumber = 99;          // Re-assign value pointed to by pNumber
16     cout << pNumber << endl; // Print content of pNumber (0x22ccf0)
17     cout << &number << endl; // Print address of number (0x22ccf0)
18     cout << *pNumber << endl; // Print value pointed to by pNumber (99)
19     cout << number << endl;   // Print value of number (99)
20                             // The value of number changes via pointer
21
22     cout << &pNumber << endl; // Print the address of pointer variable pNumber (0x22ccec)
23 }
```

Notes: The address values that you get are unlikely to be the same as mine. The OS loads the program in available free memory locations, instead of fixed memory locations.

## 1.6 Uninitialized Pointers

The following code fragment has a serious logical error!

```
int * iPtr;
*iPtr = 55;
cout << *iPtr << endl;
```

The pointer `iPtr` was declared without initialization, i.e., it is pointing to "somewhere" which is of course an invalid memory location. The `*iPtr = 55` corrupts the value of "somewhere"! You need to initialize a pointer by assigning it a valid address. Most of the compilers does not signal an error or a warning for uninitialized pointer?!

## 1.7 Null Pointers

You can initialize a pointer to 0 or NULL, i.e., it points to nothing. It is called a *null pointer*. Dereferencing a null pointer (`*p`) causes an `STATUS_ACCESS_VIOLATION` exception.

```
int * iPtr = 0;    // Declare an int pointer, and initialize the pointer to point to nothing
cout << *iPtr << endl; // ERROR! STATUS_ACCESS_VIOLATION exception

int * p = NULL;    // Also declare a NULL pointer points to nothing
```

Initialize a pointer to null during declaration is a good software engineering practice.

C++11 introduces a new keyword called `nullptr` to represent null pointer.

## 2. Reference Variables

C++ added the so-called *reference variables* (or *references* in short). A reference is an *alias*, or an *alternate name* to an existing variable. For example, suppose you make `peter` a reference (alias) to `paul`, you can refer to the person as either `peter` or `paul`.

The main use of references is acting as function formal parameters to support pass-by-reference. In an reference variable is passed into a function, the function works on the original copy (instead of a clone copy in pass-by-value). Changes inside the function are reflected outside the function.

A reference is similar to a pointer. In many cases, a reference can be used as an alternative to pointer, in particular, for the function parameter.

### 2.1 References (or Aliases) (&)

Recall that C/C++ use `&` to denote the *address-of* operator in an expression. C++ assigns an additional meaning to `&` in declaration to declare a reference variable.

The meaning of symbol `&` is different in an expression and in a declaration. When it is used in an expression, `&` denotes the address-of operator, which returns the address of a variable, e.g., if `number` is an `int` variable, `&number` returns the address of the variable `number` (this has been described in the above section).

However, when `&` is used in a *declaration* (including *function formal parameters*), it is part of the type identifier and is used to declare a *reference variable* (or *reference* or *alias* or *alternate name*). It is used to provide *another name*, or *another reference*, or *alias* to an existing variable.

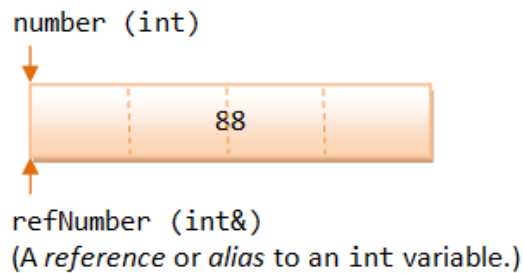
The syntax is as follow:

```
type &newName = existingName;
// or
type& newName = existingName;
// or
type & newName = existingName; // I shall adopt this convention
```

It shall be read as "*newName is a reference to existingName*", or "*newNew is an alias of existingName*". You can now refer to the variable as `newName` or `existingName`.

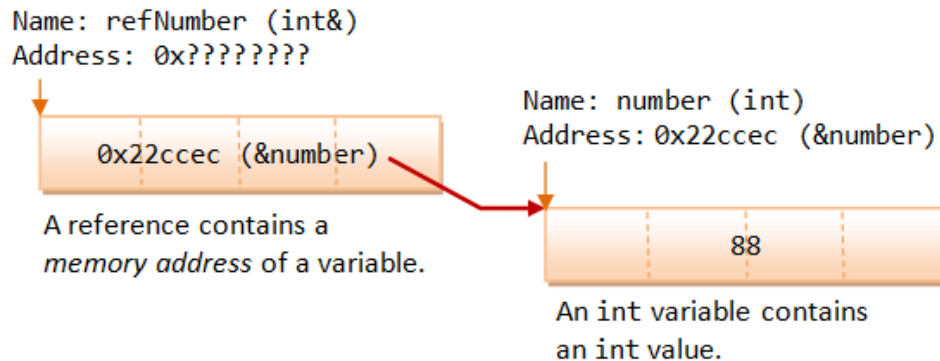
For example,

```
1  /* Test reference declaration and initialization (TestReferenceDeclaration.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int main() {
6      int number = 88;           // Declare an int variable called number
7      int & refNumber = number; // Declare a reference (alias) to the variable number
8                               // Both refNumber and number refer to the same value
9
10     cout << number << endl;    // Print value of variable number (88)
11     cout << refNumber << endl; // Print value of reference (88)
12
13     refNumber = 99;           // Re-assign a new value to refNumber
14     cout << refNumber << endl;
15     cout << number << endl;    // Value of number also changes (99)
16
17     number = 55;              // Re-assign a new value to number
18     cout << number << endl;
19     cout << refNumber << endl; // Value of refNumber also changes (55)
20 }
```



## 2.2 How References Work?

A reference works as a pointer. A reference is declared as an alias of a variable. It stores the address of the variable, as illustrated:



## 2.3 References vs. Pointers

Pointers and references are equivalent, except:

1. A reference is a *name constant for an address*. You need to initialize the reference during declaration.

```
int &iRef;    // Error: 'iRef' declared as reference but not initialized
```

Once a reference is established to a variable, you cannot change the reference to reference another variable.

2. To get the value pointed to by a pointer, you need to use the dereferencing operator `*` (e.g., if `pNumber` is a `int` pointer, `*pNumber` returns the value pointed to by `pNumber`. It is called *dereferencing* or *indirection*). To assign an address of a variable into a pointer, you need to use the address-of operator `&` (e.g., `pNumber = &number`).

On the other hand, referencing and dereferencing are done on the references implicitly. For example, if `refNumber` is a reference (alias) to another `int` variable, `refNumber` returns the value of the variable. No explicit dereferencing operator `*` should be used. Furthermore, to assign an address of a variable to a reference variable, no address-of operator `&` is needed.

For example,

```
1  /* References vs. Pointers (TestReferenceVsPointer.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int main() {
6      int number1 = 88, number2 = 22;
7
8      // Create a pointer pointing to number1
9      int * pNumber1 = &number1; // Explicit referencing
10     *pNumber1 = 99;             // Explicit dereferencing
11     cout << *pNumber1 << endl;   // 99
12     cout << &number1 << endl;    // 0x22ff18
13     cout << pNumber1 << endl;    // 0x22ff18 (content of the pointer variable - same as above)
14     cout << &pNumber1 << endl;  // 0x22ff10 (address of the pointer variable)
15     pNumber1 = &number2;        // Pointer can be reassigned to store another address
16
17     // Create a reference (alias) to number1
```

```

18  int & refNumber1 = number1; // Implicit referencing (NOT &number1)
19  refNumber1 = 11;           // Implicit dereferencing (NOT *refNumber1)
20  cout << refNumber1 << endl; // 11
21  cout << &number1 << endl;  // 0x22ff18
22  cout << &refNumber1 << endl; // 0x22ff18
23  //refNumber1 = &number2;    // Error! Reference cannot be re-assigned
24                               // error: invalid conversion from 'int*' to 'int'
25  refNumber1 = number2;       // refNumber1 is still an alias to number1.
26                               // Assign value of number2 (22) to refNumber1 (and number1).
27  number2++;
28  cout << refNumber1 << endl; // 22
29  cout << number1 << endl;   // 22
30  cout << number2 << endl;   // 23
31  }

```

A reference variable provides a new name to an existing variable. It is *dereferenced implicitly* and does not need the dereferencing operator `*` to retrieve the value referenced. On the other hand, a pointer variable stores an address. You can change the address value stored in a pointer. To retrieve the value pointed to by a pointer, you need to use the indirection operator `*`, which is known as *explicit dereferencing*. Reference can be treated as a `const` pointer. It has to be initialized during declaration, and its content cannot be changed.

Reference is closely related to pointer. In many cases, it can be used as an alternative to pointer. A reference allows you to manipulate an object using pointer, but without the pointer syntax of referencing and dereferencing.

The above example illustrates how reference works, but does not show its typical usage, which is used as the function formal parameter for pass-by-reference.

## 2.4 Pass-By-Reference into Functions with Reference Arguments vs. Pointer Arguments

### Pass-by-Value

In C/C++, by default, arguments are passed into functions *by value* (except arrays which is treated as pointers). That is, a clone copy of the argument is made and passed into the function. Changes to the clone copy inside the function has no effect to the original argument in the caller. In other words, the called function has no access to the variables in the caller. For example,

```

1  /* Pass-by-value into function (TestPassByValue.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int square(int);
6
7  int main() {
8      int number = 8;
9      cout << "In main(): " << &number << endl; // 0x22ff1c
10     cout << number << endl; // 8
11     cout << square(number) << endl; // 64
12     cout << number << endl; // 8 - no change
13 }
14
15 int square(int n) { // non-const
16     cout << "In square(): " << &n << endl; // 0x22ff00
17     n *= n; // clone modified inside the function
18     return n;
19 }

```

The output clearly shows that there are two different addresses.

### Pass-by-Reference with Pointer Arguments

In many situations, we may wish to modify the original copy directly (especially in passing huge object or array) to avoid the overhead of cloning. This can be done by passing a pointer of the object into the function, known as *pass-by-reference*. For example,



```

1  /* Pass-by-reference using pointer (TestPassByPointer.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  void square(int *);
6
7  int main() {
8      int number = 8;
9      cout << "In main(): " << &number << endl; // 0x22ff1c
10     cout << number << endl; // 8
11     square(&number); // Explicit referencing to pass an address
12     cout << number << endl; // 64
13 }
14
15 void square(int * pNumber) { // Function takes an int pointer (non-const)
16     cout << "In square(): " << pNumber << endl; // 0x22ff1c
17     *pNumber *= *pNumber; // Explicit de-referencing to get the value pointed-to
18 }

```

The called function operates on the same address, and can thus modify the variable in the caller.

### Pass-by-Reference with Reference Arguments

Instead of passing pointers into function, you could also pass references into function, to avoid the clumsy syntax of referencing and dereferencing. For example,

```

1  /* Pass-by-reference using reference (TestPassByReference.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  void square(int &);
6
7  int main() {
8      int number = 8;
9      cout << "In main(): " << &number << endl; // 0x22ff1c
10     cout << number << endl; // 8
11     square(number); // Implicit referencing (without '&')
12     cout << number << endl; // 64
13 }
14
15 void square(int & rNumber) { // Function takes an int reference (non-const)
16     cout << "In square(): " << &rNumber << endl; // 0x22ff1c
17     rNumber *= rNumber; // Implicit de-referencing (without '*')
18 }

```

Again, the output shows that the called function operates on the same address, and can thus modify the caller's variable.

Take note referencing (in the caller) and dereferencing (in the function) are done implicitly. The only coding difference with pass-by-value is in the function's parameter declaration.

Recall that references are to be initialized during declaration. In the case of function formal parameter, the references are initialized when the function is invoked, to the caller's arguments.

References are primarily used in passing reference in/out of functions to allow the called function accesses variables in the caller directly.

### "const" Function Reference/Pointer Parameters

A const function formal parameter cannot be modified inside the function. Use const whenever possible as it protects you from inadvertently modifying the parameter and protects you against many programming errors.

A const function parameter can receive both const and non-const argument. On the other hand, a non-const function reference/pointer parameter can only receive non-const argument. For example,

```

1  /* Test Function const and non-const parameter (FuncationConstParameter.cpp) */
2  #include <iostream>
3  using namespace std;

```



```

4
5 int squareConst(const int);
6 int squareNonConst(int);
7 int squareConstRef(const int &);
8 int squareNonConstRef(int &);
9
10 int main() {
11     int number = 8;
12     const int constNumber = 9;
13     cout << squareConst(number) << endl;
14     cout << squareConst(constNumber) << endl;
15     cout << squareNonConst(number) << endl;
16     cout << squareNonConst(constNumber) << endl;
17
18     cout << squareConstRef(number) << endl;
19     cout << squareConstRef(constNumber) << endl;
20     cout << squareNonConstRef(number) << endl;
21     // cout << squareNonConstRef(constNumber) << endl;
22     // error: invalid initialization of reference of
23     // type 'int&' from expression of type 'const int'
24 }
25
26 int squareConst(const int number) {
27     // number *= number; // error: assignment of read-only parameter
28     return number * number;
29 }
30
31 int squareNonConst(int number) { // non-const parameter
32     number *= number;
33     return number;
34 }
35
36 int squareConstRef(const int & number) { // const reference
37     return number * number;
38 }
39
40 int squareNonConstRef(int & number) { // non-const reference
41     return number * number;
42 }

```

## 2.5 Passing the Function's Return Value

### Passing the Return-value as Reference

You can also pass the return-value as reference or pointer. For example,

```

1  /* Passing back return value using reference (TestPassByReferenceReturn.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int & squareRef(int &);
6  int * squarePtr(int *);
7
8  int main() {
9      int number1 = 8;
10     cout << "In main() &number1: " << &number1 << endl; // 0x22ff14
11     int & result = squareRef(number1);
12     cout << "In main() &result: " << &result << endl; // 0x22ff14
13     cout << result << endl; // 64
14     cout << number1 << endl; // 64
15
16     int number2 = 9;
17     cout << "In main() &number2: " << &number2 << endl; // 0x22ff10
18     int * pResult = squarePtr(&number2);
19     cout << "In main() pResult: " << pResult << endl; // 0x22ff10
20     cout << *pResult << endl; // 81

```

```

21     cout << number2 << endl;    // 81
22 }
23
24 int & squareRef(int & rNumber) {
25     cout << "In squareRef(): " << &rNumber << endl; // 0x22ff14
26     rNumber *= rNumber;
27     return rNumber;
28 }
29
30 int * squarePtr(int * pNumber) {
31     cout << "In squarePtr(): " << pNumber << endl; // 0x22ff10
32     *pNumber *= *pNumber;
33     return pNumber;
34 }

```

### You should not pass Function's local variable as return value by reference

```

1  /* Test passing the result (TestPassResultLocal.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int * squarePtr(int);
6  int & squareRef(int);
7
8  int main() {
9      int number = 8;
10     cout << number << endl; // 8
11     cout << *squarePtr(number) << endl; // ??
12     cout << squareRef(number) << endl; // ??
13 }
14
15 int * squarePtr(int number) {
16     int localResult = number * number;
17     return &localResult;
18     // warning: address of local variable 'localResult' returned
19 }
20
21 int & squareRef(int number) {
22     int localResult = number * number;
23     return localResult;
24     // warning: reference of local variable 'localResult' returned
25 }

```

This program has a serious logical error, as local variable of function is passed back as return value by reference. Local variable has local scope within the function, and its value is destroyed after the function exits. The GCC compiler is kind enough to issue a warning (but not error).

It is safe to return a reference that is passed into the function as an argument. See earlier examples.

### Passing Dynamically Allocated Memory as Return Value by Reference

Instead, you need to dynamically allocate a variable for the return value, and return it by reference.

```

1  /* Test passing the result (TestPassResultNew.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int * squarePtr(int);
6  int & squareRef(int);
7
8  int main() {
9      int number = 8;
10     cout << number << endl; // 8
11     cout << *squarePtr(number) << endl; // 64
12     cout << squareRef(number) << endl; // 64
13 }
14

```

```

15  int * squarePtr(int number) {
16      int * dynamicAllocatedResult = new int(number * number);
17      return dynamicAllocatedResult;
18  }
19
20  int & squareRef(int number) {
21      int * dynamicAllocatedResult = new int(number * number);
22      return *dynamicAllocatedResult;
23  }

```

## 2.6 Summary

Pointers and references are highly complex and difficult to master. But they can greatly improve the efficiency of the programs.

For novices, avoid using pointers in your program. Improper usage can lead to serious logical bugs. However, you need to understand the syntaxes of pass-by-reference with pointers and references, because they are used in many library functions.

- In *pass-by-value*, a clone is made and passed into the function. The caller's copy cannot be modified.
- In *pass-by-reference*, a pointer is passed into the function. The caller's copy could be modified inside the function.
- In *pass-by-reference with reference arguments*, you use the variable name as the argument.
- In *pass-by-reference with pointer arguments*, you need to use `&varName` (an address) as the argument.

## 3. Dynamic Memory Allocation

### 3.1 new and delete Operators

Instead of define an `int` variable (`int number`), and assign the address of the variable to the `int` pointer (`int *pNumber = &number`), the storage can be dynamically allocated at runtime, via a `new` operator. In C++, whenever you allocate a piece of memory dynamically via `new`, you need to use `delete` to remove the storage (i.e., to return the storage to the heap).

The `new` operation returns a pointer to the memory allocated. The `delete` operator takes a pointer (pointing to the memory allocated via `new`) as its sole argument.

For example,

```

// Static allocation
int number = 88;
int * p1 = &number; // Assign a "valid" address into pointer

// Dynamic Allocation
int * p2;           // Not initialize, points to somewhere which is invalid
cout << p2 << endl; // Print address before allocation
p2 = new int;       // Dynamically allocate an int and assign its address to pointer
                  // The pointer gets a valid address with memory allocated

*p2 = 99;
cout << p2 << endl; // Print address after allocation
cout << *p2 << endl; // Print value point-to
delete p2;          // Remove the dynamically allocated storage

```

Observe that `new` and `delete` operators work on *pointer*.

To initialize the allocated memory, you can use an initializer for fundamental types, or invoke a constructor for an object. For example,

```

// use an initializer to initialize a fundamental type (such as int, double)
int * p1 = new int(88);
double * p2 = new double(1.23);

// C++11 brace initialization syntax
int * p1 = new int {88};
double * p2 = new double {1.23};

// invoke a constructor to initialize an object (such as Date, Time)

```

```
Date * date1 = new Date(1999, 1, 1);
Time * time1 = new Time(12, 34, 56);
```

You can dynamically allocate storage for *global* pointers inside a function. Dynamically allocated storage inside the function remains even after the function exits. For example,

```
1  // Dynamically allocate global pointers (TestDynamicAllocation.cpp)
2  #include <iostream>
3  using namespace std;
4
5  int * p1, * p2; // Global int pointers
6
7  // This function allocates storage for the int*
8  // which is available outside the function
9  void allocate() {
10     p1 = new int; // Allocate memory, initial content unknown
11     *p1 = 88; // Assign value into location pointed to by pointer
12     p2 = new int(99); // Allocate and initialize
13 }
14
15 int main() {
16     allocate();
17     cout << *p1 << endl; // 88
18     cout << *p2 << endl; // 99
19     delete p1; // Deallocate
20     delete p2;
21     return 0;
22 }
```

The main differences between static allocation and dynamic allocations are:

1. In static allocation, the compiler allocates and deallocates the storage automatically, and handle memory management. Whereas in dynamic allocation, you, as the programmer, handle the memory allocation and deallocation yourself (via `new` and `delete` operators). You have full control on the pointer addresses and their contents, as well as memory management.
2. Static allocated entities are manipulated through named variables. Dynamic allocated entities are handled through pointers.

### 3.2 new[] and delete[] Operators

Dynamic array is allocated at runtime rather than compile-time, via the `new[]` operator. To remove the storage, you need to use the `delete[]` operator (instead of simply `delete`). For example,

```
1  /* Test dynamic allocation of array (TestDynamicArray.cpp) */
2  #include <iostream>
3  #include <cstdlib>
4  using namespace std;
5
6  int main() {
7     const int SIZE = 5;
8     int * pArray;
9
10     pArray = new int[SIZE]; // Allocate array via new[] operator
11
12     // Assign random numbers between 0 and 99
13     for (int i = 0; i < SIZE; ++i) {
14         *(pArray + i) = rand() % 100;
15     }
16     // Print array
17     for (int i = 0; i < SIZE; ++i) {
18         cout << *(pArray + i) << " ";
19     }
20     cout << endl;
21
22     delete[] pArray; // Deallocate array via delete[] operator
23     return 0;
```

```
24 }
```

C++03 does not allow you to initialize the dynamically-allocated array. C++11 does with the brace initialization, as follows:

```
// C++11
int * p = new int[5] {1, 2, 3, 4, 5};
```

## 4. Pointer, Array and Function

### 4.1 Array is Treated as Pointer

In C/C++, an array's name is a pointer, pointing to the first element (index 0) of the array. For example, suppose that `numbers` is an `int` array, `numbers` is also an `int` pointer, pointing at the first element of the array. That is, `numbers` is the same as `&numbers[0]`. Consequently, `*numbers` is `numbers[0]`; `*(numbers+i)` is `numbers[i]`.

For example,

```
1  /* Pointer and Array (TestPointerArray.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int main() {
6      const int SIZE = 5;
7      int numbers[SIZE] = {11, 22, 44, 21, 41}; // An int array
8
9      // The array name numbers is an int pointer, pointing at the
10     // first item of the array, i.e., numbers = &numbers[0]
11     cout << &numbers[0] << endl; // Print address of first element (0x22fef8)
12     cout << numbers << endl;    // Same as above (0x22fef8)
13     cout << *numbers << endl;    // Same as numbers[0] (11)
14     cout << *(numbers + 1) << endl; // Same as numbers[1] (22)
15     cout << *(numbers + 4) << endl; // Same as numbers[4] (41)
16 }
```

### 4.2 Pointer Arithmetic

As seen from the previous section, if `numbers` is an `int` array, it is treated as an `int` pointer pointing to the first element of the array. `(numbers + 1)` points to the next `int`, instead of having the next sequential address. Take note that an `int` typically has 4 bytes. That is `(numbers + 1)` increases the address by 4, or `sizeof(int)`. For example,

```
int numbers[] = {11, 22, 33};
int * iPtr = numbers;
cout << iPtr << endl;        // 0x22cd30
cout << iPtr + 1 << endl;    // 0x22cd34 (increase by 4 - sizeof int)
cout << *iPtr << endl;      // 11
cout << *(iPtr + 1) << endl; // 22
cout << *iPtr + 1 << endl;  // 12
```

### 4.3 sizeof Array

The operation `sizeof(arrayName)` returns the *total bytes* of the array. You can derive the length (size) of the array by dividing it with the size of an element (e.g. element 0). For example,

```
int numbers[100];
cout << sizeof(numbers) << endl;    // Size of entire array in bytes (400)
cout << sizeof(numbers[0]) << endl; // Size of first element of the array in bytes (4)
cout << "Array size is " << sizeof(numbers) / sizeof(numbers[0]) << endl; // (100)
```

### 4.4 Passing Array In/Out of a Function

An array is passed into a function as a *pointer to the first element* of the array. You can use array notation (e.g., `int[]`) or pointer notation (e.g., `int*`) in the function declaration. The compiler always treats it as pointer (e.g., `int*`). For example,

the following declarations are equivalent:

```
int max(int numbers[], int size);
int max(int *numbers, int size);
int max(int number[50], int size);
```

They will be treated as `int*` by the compiler, as follow. The size of the array given in `[]` is ignored.

```
int max(int*, int);
```

Array is passed by reference into the function, because a pointer is passed instead of a clone copy. If the array is modified inside the function, the modifications are applied to the caller's copy. You could declare the array parameter as `const` to prevent the array from being modified inside the function.

The size of the array is not part of the array parameter, and needs to be passed in another `int` parameter. Compiler is not able to deduce the array size from the array pointer, and does not perform array bound check.

**Example:** Using the usual array notation.

```
1  /* Passing array in/out function (TestArrayPassing.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  // Function prototypes
6  int max(const int arr[], int size);
7  void replaceByMax(int arr[], int size);
8  void print(const int arr[], int size);
9
10 int main() {
11     const int SIZE = 4;
12     int numbers[SIZE] = {11, 22, 33, 22};
13     print(numbers, SIZE);
14     cout << max(numbers, SIZE) << endl;
15     replaceByMax(numbers, SIZE);
16     print(numbers, SIZE);
17 }
18
19 // Return the maximum value of the given array.
20 // The array is declared const, and cannot be modified inside the function.
21 int max(const int arr[], int size) {
22     int max = arr[0];
23     for (int i = 1; i < size; ++i) {
24         if (max < arr[i]) max = arr[i];
25     }
26     return max;
27 }
28
29 // Replace all elements of the given array by its maximum value
30 // Array is passed by reference. Modify the caller's copy.
31 void replaceByMax(int arr[], int size) {
32     int maxValue = max(arr, size);
33     for (int i = 0; i < size; ++i) {
34         arr[i] = maxValue;
35     }
36 }
37
38 // Print the array's content
39 void print(const int arr[], int size) {
40     cout << "{";
41     for (int i = 0; i < size; ++i) {
42         cout << arr[i];
43         if (i < size - 1) cout << ",";
44     }
45     cout << "}" << endl;
46 }
```

Take note that you can modify the contents of the caller's array inside the function, as array is passed by reference. To prevent accidental modification, you could apply `const` qualifier to the function's parameter. Recall that `const` inform the

compiler that the value should not be changed. For example, suppose that the function `print()` prints the contents of the given array and does not modify the array, you could apply `const` to both the array name and its size, as they are not expected to be changed inside the function.

```
void print(const int arr[], int size);
```

Compiler flags out an error "assignment of read-only location" if it detected a `const` value would be changed.

**Example:** Using pointer notation.

```
1  /* Passing array in/out function using pointer (TestArrayPassingPointer.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  // Function prototype
6  int max(const int *arr, int size);
7
8  int main() {
9      const int SIZE = 5;
10     int numbers[SIZE] = {10, 20, 90, 76, 22};
11     cout << max(numbers, SIZE) << endl;
12 }
13
14 // Return the maximum value of the given array
15 int max(const int *arr, int size) {
16     int max = *arr;
17     for (int i = 1; i < size; ++i) {
18         if (max < *(arr+i)) max = *(arr+i);
19     }
20     return max;
21 }
```

## 4.5 Pass-by-Reference and sizeof

```
1  /* Test sizeof array (TestSizeofArray.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  // Function prototypes
6  void fun(const int *arr, int size);
7
8  // Test Driver
9  int main() {
10     const int SIZE = 5;
11     int a[SIZE] = {8, 4, 5, 3, 2};
12     cout << "sizeof in main() is " << sizeof(a) << endl;
13     cout << "address in main() is " << a << endl;
14     fun(a, SIZE);
15 }
16
17 // Function definitions
18 void fun(const int *arr, int size) {
19     cout << "sizeof in function is " << sizeof(arr) << endl;
20     cout << "address in function is " << arr << endl;
21 }
```

```
sizeof in main() is 20
address in main() is 0x22fefc
sizeof in function is 4
address in function is 0x22fefc
```

The address of arrays in `main()` and the function are the same, as expected, as array is passed by reference.

In `main()`, the `sizeof` array is 20 (4 bytes per `int`, length of 5). Inside the function, the `sizeof` is 4, which is the `sizeof` `int` pointer (4-byte address). This is why you need to pass the size into the function.



## 4.6 Operating on a Range of an Array

```

1  /* Function to compute the sum of a range of an array (SumArrayRange.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  // Function prototype
6  int sum(const int *begin, const int *end);
7
8  // Test Driver
9  int main() {
10     int a[] = {8, 4, 5, 3, 2, 1, 4, 8};
11     cout << sum(a, a+8) << endl;      // a[0] to a[7]
12     cout << sum(a+2, a+5) << endl;    // a[2] to a[4]
13     cout << sum(&a[2], &a[5]) << endl; // a[2] to a[4]
14 }
15
16 // Function definition
17 // Return the sum of the given array of the range from
18 // begin to end, exclude end.
19 int sum(const int *begin, const int *end) {
20     int sum = 0;
21     for (const int *p = begin; p != end; ++p) {
22         sum += *p;
23     }
24     return sum;
25 }

```

Program Notes:

- To write a function that operates on a range of the given array, you can pass the begin pointer and the end pointer into the function. By convention, the operation shall start at the begin pointer, up to the end pointer, but excluding the end pointer.
- In "const int \*p", \*p (content pointed-to) is constant, but p is not constant.

## 4.7 C-String and Pointer

C-string (of the C language) is a character array, terminated with a null character '\0'. For example,

```

1  /* Testing C-string (TestCString.cpp) */
2  #include <iostream>
3  #include <cstring>
4  using namespace std;
5
6  int main() {
7     char msg1[] = "Hello";
8     char *msg2 = "Hello";
9     // warning: deprecated conversion from string constant to 'char*'
10
11     cout << strlen(msg1) << endl;    // 5
12     cout << strlen(msg2) << endl;
13     cout << strlen("Hello") << endl;
14
15     int size = sizeof(msg1)/sizeof(char);
16     cout << size << endl; // 6 - including the terminating '\0'
17     for (int i = 0; msg1[i] != '\0'; ++i) {
18         cout << msg1[i];
19     }
20     cout << endl;
21
22     for (char *p = msg1; *p != '\0'; ++p) {
23         // *p != '\0' is the same as *p != 0, is the same as *p
24         cout << *p;
25     }
26     cout << endl;

```

```
27 }
```

Take note that for C-String function such as `strlen()` (in header `cstring`, ported over from C's `string.h`), there is no need to pass the array length into the function. This is because C-Strings are terminated by `'\0'`. The function can iterate thru the characters in the array until `'\0'`. For example,

```
1  /* Function to count the occurrence of a char in a string (CountChar.cpp) */
2  #include <iostream>
3  #include <cstring>
4  using namespace std;
5
6  int count(const char *str, const char c); // No need to pass the array size
7
8  int main() {
9      char msg1[] = "Hello, world";
10     char *msg2 = "Hello, world";
11
12     cout << count(msg1, 'l') << endl;
13     cout << count(msg2, 'l') << endl;
14     cout << count("Hello, world", 'l') << endl;
15 }
16
17 // Count the occurrence of c in str
18 // No need to pass the size of char[] as C-string is terminated with '\0'
19 int count(const char *str, const char c) {
20     int count = 0;
21     while (*str) { // same as (*str != '\0')
22         if (*str == c) ++count;
23         ++str;
24     }
25     return count;
26 }
```

## 5. \*More On Pointers

### 5.1 Function Pointer

In C/C++, functions, like all data items, have an address. The name of a function is the starting address where the function resides in the memory, and therefore, can be treated as a pointer. We can pass a function pointer into function as well. The syntax for declaring a function pointer is:

```
// Function-pointer declaration
return-type (* function-ptr-name) (parameter-list)

// Examples
double (*fp)(int, int) // fp points to a function that takes two ints and returns a double (function-pointer)
double *dp;           // dp points to a double (double-pointer)
double *fun(int, int) // fun is a function that takes two ints and returns a double-pointer

double f(int, int);    // f is a function that takes two ints and returns a double
fp = f;               // Assign function f to fp function-pointer
```

#### Example

```
1  /* Test Function Pointers (TestFunctionPointer.cpp) */
2  #include <iostream>
3  using namespace std;
4
5  int arithmetic(int, int, int (*)(int, int));
6      // Take 3 arguments, 2 int's and a function pointer
7      // int (*)(int, int), which takes two int's and return an int
8  int add(int, int);
9  int sub(int, int);
10
11  int add(int n1, int n2) { return n1 + n2; }
```

```

12 int sub(int n1, int n2) { return n1 - n2; }
13
14 int arithmetic(int n1, int n2, int (*operation) (int, int)) {
15     return (*operation)(n1, n2);
16 }
17
18 int main() {
19     int number1 = 5, number2 = 6;
20
21     // add
22     cout << arithmetic(number1, number2, add) << endl;
23     // subtract
24     cout << arithmetic(number1, number2, sub) << endl;
25 }

```

## 5.2 Generic Pointer or void Pointer (void \*)

A void pointer can hold address of any data type (except function pointer). We cannot operate on the object pointed to by void pointer, as the type is unknown. We can use a void pointer to compare with another address.

[TODO] Example

## 5.3 Constant Pointer vs. Constant Pointed-to Data

1. Non-constant pointer to constant data: Data pointed to CANNOT be changed; but pointer CAN be changed to point to another data. For example,

```

int i1 = 8, i2 = 9;
const int * iptr = &i1; // non-constant pointer pointing to constant data
// *iptr = 9; // error: assignment of read-only location
iptr = &i2; // okay

```

2. Constant pointer to non-constant data: Data pointed to CAN be changed; but pointer CANNOT be changed to point to another data. For example,

```

int i1 = 8, i2 = 9;
int * const iptr = &i1; // constant pointer pointing to non-constant data
                        // constant pointer must be initialized during declaration
*iptr = 9; // okay
// iptr = &i2; // error: assignment of read-only variable

```

3. Constant pointer to constant data: Data pointed to CANNOT be changed; and pointer CANNOT be changed to point to another data. For example,

```

int i1 = 8, i2 = 9;
const int * const iptr = &i1; // constant pointer pointing to constant data
// *iptr = 9; // error: assignment of read-only variable
// iptr = &i2; // error: assignment of read-only variable

```

4. Non-constant pointer to non-constant data: Data pointed to CAN be changed; and pointer CAN be changed to point to another data. For example,

```

int i1 = 8, i2 = 9;
int * iptr = &i1; // non-constant pointer pointing to non-constant data
*iptr = 9; // okay
iptr = &i2; // okay

```

[Link to "C++ References & Resources"](#)

Feedback, comments, corrections, and errata can be sent to Chua Hock-Chuan (ehchua@ntu.edu.sg) | [HOME](#)