POLITECNICO DI MILANO





Advanced Operating Systems

Moving to C++

Giuseppe Massari

giuseppe.massari@polimi.it

Outline

C++ Programs

- Memory allocation
- Passing parameters to functions

Classes

- Constructors and assignments
- Object pointers
- Qualifiers

Standard Template Library (STL)

- Templates
- Example: `vector<>`, `list<>`

Memory management

- From raw to smart pointers

Program memory

Memory layout

- Once loaded in the main memory a program is typically divided into segments

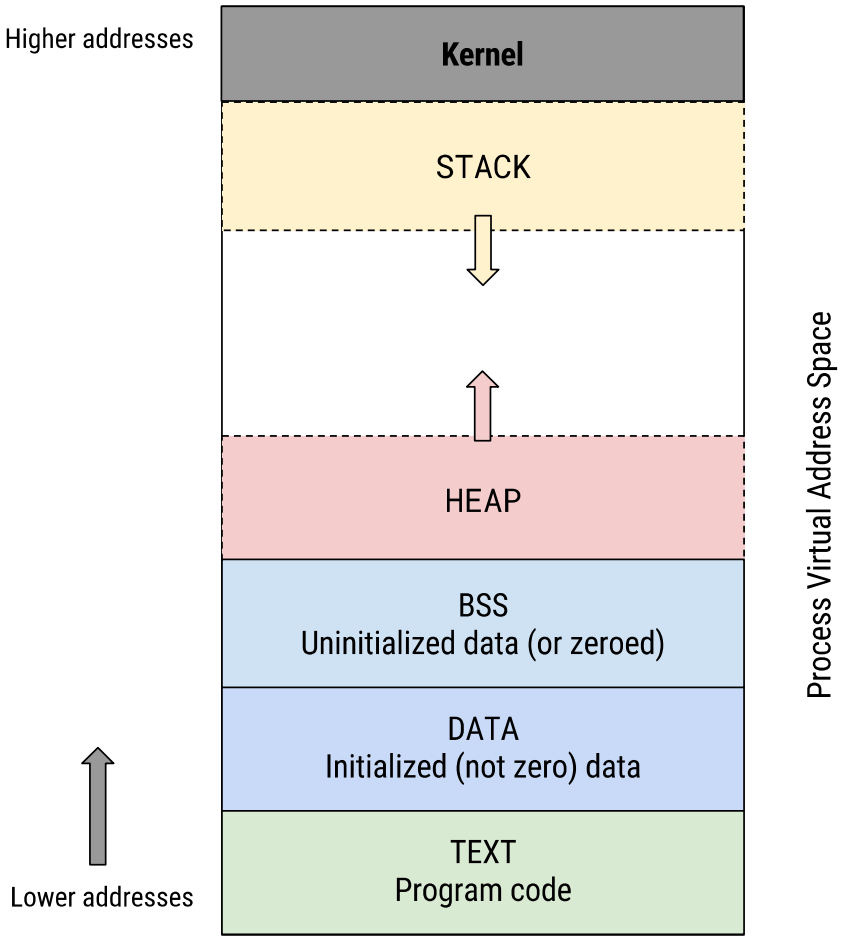
TEXT : Program instructions

DATA : Initialized variables

BSS : Uninitialized or zeroed variables

STACK : Local variables and return address of functions

HEAP : *Dynamically* allocated data, i.e., allocated/de-allocated at run-time



Program memory

Memory allocation

- In C programs dynamic memory allocation and de-allocation is performed through **`malloc()`** and **`free()`** functions

```
char * buff = malloc(buff_size);
```

```
free(buff);
```

- In C++ programs operators **`new`** and **`delete`** are used

  For dynamic array allocation **`new ….[]`** / **`delete[]`**

```
char * buff = new char[5]

Object * obj = new Object();
```

```
delete[] buff;

delete obj;
```

- **`new`**/**`delete`** call class constructor and destructor respectively

  Do not mix **`malloc`**/**`free`** with **`new`**/**`delete`** usage!

  Do not use **`malloc`**/**`free`** with C++ objects since they do not call the class constructor/destructor!

Program memory

Memory allocation

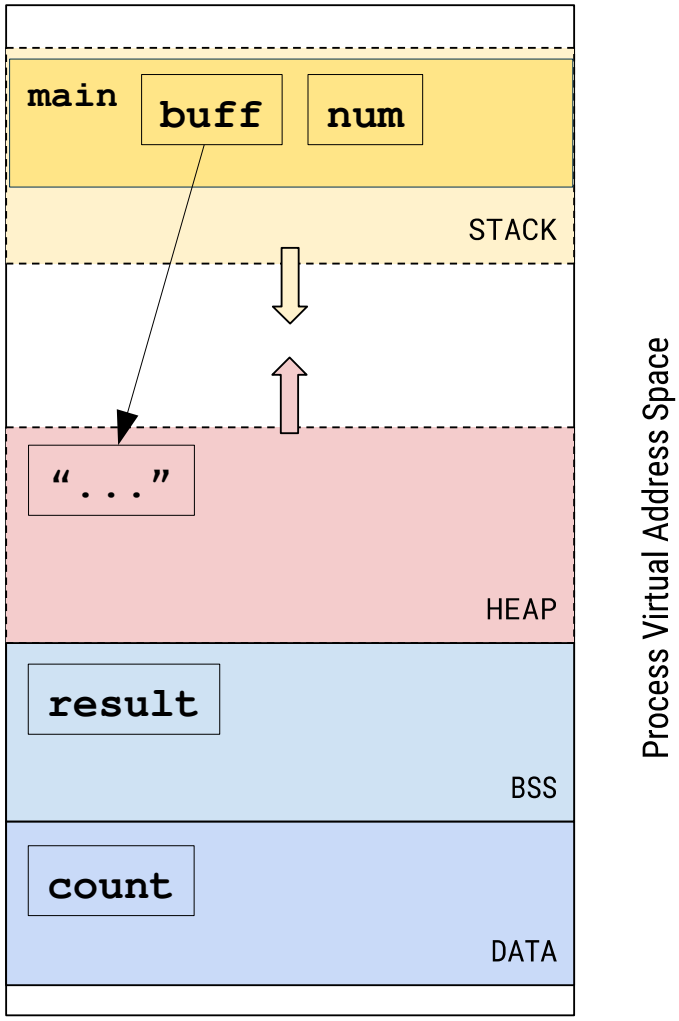
- Example

```
static result;
unsigned int count = 20;

int main() {

    int num;
    char * buff = new char[4];

    …
    delete[] value;
    return 0;
}
```



Memory allocation common errors

Memory leak

- Memory allocated but NOT released
- Application continuously increase memory consumption
  - Reducing the availability for other applications
- Performance slow down
  - Operating system may re-map the virtual memory page to the hard drive
- Application or even the system may stop working
  - Unpredictable behaviour when system memory resource limits are reached

Sources

- We forgotten the related **`delete`** call for some **`new`**
- We changed the value of a pointer variable
  - No other variables point to the previously dynamically allocated data
  - Such data becomes unreachable!

Memory allocation common errors

Memory corruption

- Memory location alteration without an explicit assignment

Attempt of freeing an already freed memory location (dangling pointer)

```
int main() {
    char * buff = new char[4];
    delete[] buff;
    …
    delete[] buff;
}
```

```
$ ./program
*** Error in `./program':
 double free or corruption
: 0x00000000011a5040 ***

Aborted (core dumped)
```

Attempt of accessing a freed (or not allocated yet) memory location

```
int main() {
    char * buff;
    cout << buff << endl;
}
```

```
$ ./program
Segmentation fault
(core dumped)
```

Passing parameters

- By *value*

  Passed variables/objects are COPIED into the called function stack

```
void func(Object var);
```

```
func(var);    // call
```

- By *address* (C approach)

  Caller passes the address of the object/variable to the callee (NO COPY)

  Callee can modify the content of the variable/object

```
void func(Object * var);
```

```
func(&var);    // call
```

- By *reference* (C++ approach)

  Same effect of passing by address (NO COPY)

  Callee can modify the content of the variable/object

```
void func(Object & var);
```

```
func(var);    // call
```

Outline

Syntax

- A C++ class can include a *public* and *private* sections (keywords **`public`** and **`private`**) to set the visibility level of member *data* (attributes) and member *functions* (methods)

- Good practice: member definition and member implementation placed into separated files

  Definition → header file (**`.h`**, **`.hh`**) / *Implementation* → **`.cpp`**, **`.cc`**

- Member function implementation syntax

  `<function type>` ***`ClassName`***`::`***`FunctionName`***`(...) {...}`**

- Accessor methods (**`Get...()`**) are tagged with the keyword **`const`**

- The class definition is closed by a semicolon (`;`)

Syntax

```
#ifndef PEOPLE_H_
#define PEOPLE_H_

#include <string>

class People {
public:
    People();
    int Walk();
    void Talk() const;

private:
    std::string name;
    int step_count;
};

#endif // PEOPLE_H_

// people.h
```

```
#include <iostream>
#include "people.h"

People::People():
    name("guy"),
    step_count(0) {
}

int People::Walk() {
    return ++step_count;
}

void People::Talk() const{
    std::cout << name
        << " made " << step_count
        << " steps" << std::endl;
}

// people.cpp
```

Constructors

- A function having the same class name but not returning any type
- Called every time an instance of a class is created
- If not provided, the compiler will generate the code for one
- A class can have multiple constructors, with different argument lists

```
class People {
public:
    People();
    People(std::string _name);
    …
private:
    std::string name;
    int step_count;
…
};

// people.h
```

```
People::People():
    name("Guy"),
    step_count(0) {
}

People::People(std::string _name):
    name(_name),
    step_count(0) {
}

// people.cpp
```

Constructors

- Called when we instantiate an object, by declaring a variable of the class type, or we dynamically allocate an object of the class type (see later)

```
int main() {
...
    People who1;
    People who2();
    People simon("Simon");


    who1.Talk();
    who2.Talk();
    simon.Talk();
...
}
```

People() version is called

People(std::string) version is called

Output:

```
Guy made 0 steps
Guy made 0 steps
Simon made 0 steps
```

Destructors

- Called when the object is going out of scope or it is explicitly deleted
- Syntax: **`~ClassName()`** (no arguments)

  Very often qualified as "virtual" to prevent derived class from invoking the base class destructor
- Commonly includes the implementation of end of life actions, e.g., clean-up code, memory deallocation,...

```
class People {
public:
    People();
    People(std::string _name);
    virtual ~People();
…
};
// people.h
```

```
…
People::~People() {
    std::cout << name
        << ": i'm dying..."
        << std::endl;
}
…
// people.cpp
```

Copy constructors

- Special constructors having an object of the same class as argument
- If not implemented the compiler will automatically generate one
- The default implementation is a *member-wise copy* of the object

```
People::People( const People & other )
{
    name       = other.name;
    step_count = other.step_count;
}
```

- Member-wise copy can be expensive

  What if member data are further complex objects?

- Depending on the class definition, it may also lead to undesired behaviours

  What if member data are pointers? Pointer object “complex” class types?

Copy constructors

- Example

  Fred is created starting from Simon(do we want Fred to have the same Simon's information?)

  Default copy constructor is called (member-wise copy)

```
int main() {
...
    People simon("Simon");

    for(int i=0; i<3; ++i)
        simon.Walk()
    simon.Talk();

    People fred(simon);
    fred.Talk();
...
}
```

Output:

```
Simon made 3 steps
Simon made 3 steps
```

fred's member data have been initialized to the same value of object **simon**

Copy constructors

- Creating a “People” object starting from another one...

  Do we want the “new person” Fred to start with the number of steps of the Simon?

- Overload of the copy constructor implementing our custom version

  Set **`step_count`** to zero, since we are constructing a new object that has never walked before!

  Desired behaviour: *Just copy the name of the other person*

```
People::People( const People & other )
{
    name       = other.name;
    step_count = 0;
}
```

Copy constructors

- Example

  We create a second “Simon” object

  Custom copy constructor initializes the object using the same of the first Simon

```
int main() {
...
    People simon1(“Simon”);

    for(int i=0; i<3; ++i)
        simon.Walk()
    simon.Talk();

    People simon2(simon1);
    simon2.Talk();
...
}
```

Output:

```
Simon made 3 steps
Simon made 0 steps
```

simon2 has the same name of **simon1**, but since it has just born, it has never walked before!

Object assignment

- Assignment operation is provided by the compiler
- Also in this case, the default implementation is the member-wise copy of the object

  Conversely from other OO languages (e.g., Java) the object variables content is a *value* and not a reference

```
People & People::operator=( const People & other )
{
    name       = other.name;
    step_count = other.step_count;
}
```

- We can overload the assignment operator (“=”) too!

  Syntax similar to member function implementation, but replacing the function name with **operator**<operator_symbol>

Object assignment

- Conversely from other OO languages (e.g., Java) the object variables content is a *value* and not a reference
- An assignment operation lead to a copy of the entire object

```
int main() {
...
    People simon("Simon");
    People fred;

    fred = simon;
    simon.Walk();
    simon.Talk();

    fred.Talk()
...
}
```

Output:

```
Simon made 1 steps
Simon made 0 steps
```

Member data **name** value has been copied

Walk() has updated only **simon** status

Object pointers

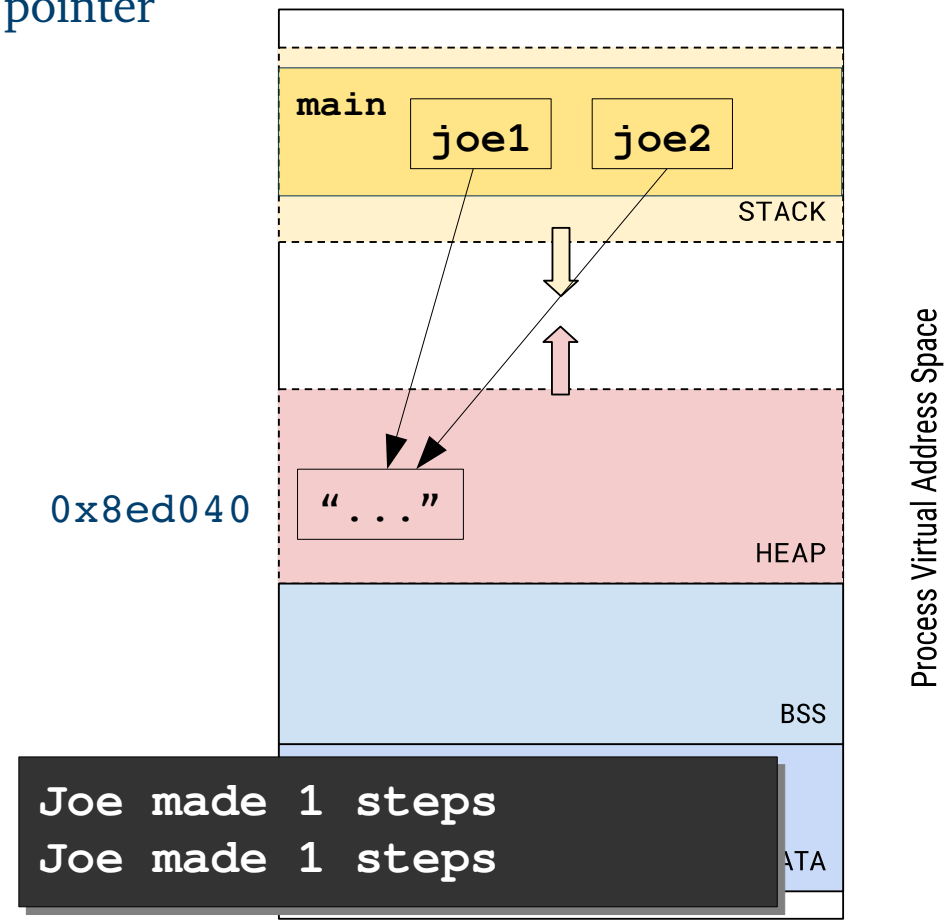
- *Pointers* allows multiple variables to jointly access the same object

Assignment operation copies the pointer value, i.e., the memory address

```
int main() {
    People * joe1 =
        new People("Joe");
    People * joe2;

    joe2 = joe1;
    joe1->Walk();
    joe1->Talk();
    joe2->Talk()

    delete joe2;
    return 0;
}
```



Constant members

- Data or function members qualified with **`const`** keyword
- Const-qualified variable are used to store data that should not be altered
- Qualifier **`const`** applies to whatever is on its immediate left

  If there is nothing there applies to whatever is its immediate right

```
// k is a variable pointer to a constant integer
const int * k;
// ...alternative syntax which does the same
int const * k;
// k is constant pointer to an integer variable
int * const k;
// k is a constant pointer to a constant integer
int const * const k;
```

Constant members

- The **const** keyword can be used also in functions
- To specify a “read-only” function, i.e., that does not change the object status (member data values)

```
void Talk() const;
```

- To pass (by reference) arguments (typically objects) that we do not want to be modified

```
void TalkWith(const People & other);
```

Function **TalkWith()** cannot access not **const** member functions of **other**

- To prevent alteration of returned variable of pointer types

Alteration attempts are detected at compile-time

```
const People * GetParent() const;
```

Static members

- The **`static`** keyword can be used in a class definition for different purposes, depending on if it is applied to data or functions
- Static member ***data*** are variables shared among all the instances

  Member data initialization cannot be performed in the class

```
class MyClass { // myclass.h
    static int count;
};
```

```
// myclass.cpp
// Initialization
int MyClass::count = 0;
```

- Static member ***functions*** do not require a class instance

  Functions though to be meaningful at *class-level* instead of *instance-level*

  They cannot access *non static* members since not "belonging" to any instance

```
class MyClass { // myclass.h
public:
    static int GetCount();
};
```

```
std::cout
    << MyClass::GetCount()
    << std::endl;
```

Static members

- Example

```
class People {
public:
    People();
…
    static int GetPopulation();
private:
    …
    static int population;
};
// people.h
```

```
int main() {
    People joe;
    std::cout <<
        People::GetPopulation();
    …
}
```

```
//Init
int People::population = 0;

People::People():
    name("guy"),
    step_count(0) {
    population++;
}

People::~People() {
    population--;
}

int People::GetPopulation() {
    return population;
}

// people.cpp
```

Outline

C++ Programs

- Memory allocation
- Passing parameters to functions

Classes

- Constructors and assignments
- Object pointers
- Qualifiers

Standard Template Library (STL)

- Templates
- Example: `vector<>`, `list<>`

Memory management

- From raw to smart pointers

Template

- A way of making functions or classes more abstract
- Focus on the behaviour of the function (or the class) without knowing variables data type
  - A single function / class implementation to cover several data type cases
  - Avoid code replications (simply due to variables data type)
- Data type resolution performed at compile-time
  - Compilers generate additional code
    - Large use of template may strongly increase the executable size
    - Longer compilation times
  - Debugging is complicated by hard to read compiler messages
  - Template code is commonly placed in header files
    - Modify a template class may lead to entire project re-build

Template functions

- Example of global function performing addition between two variables of a generic type

```
…
template<class T>
T add(T a1, T a2)
{
    return a1 + a2;
}
…
int main() {
    int i1, i2;
    float f1, f2;
    …
    cout << add<int>(i1, i2)
        << endl;
    cout << add<float>(f1, f2)
        << endl;
    return 0;
}
```

The compiler will generate two versions of function **add()**

```
int add(int a1, int a2)
{
    return a1 + a2;
}
```

```
float add(float a1, float a2)
{
    return a1 + a2;
}
```

Template classes

- A class including generic type members

```
#ifndef MATH_H_
#define MATH_H_
namespace aos {

template <class T>
class Math {
public:
    T add(T a1, T a2);
    T mul(T m1, T m2);
    T div(T m1, T m2);
};

} // namespace aos
#endif // MATH_H_

// math.h
```

```
#include "math.h"
namespace aos {

template <class T>
T Math<T>::add(T a1, T a2) {
    return a1 + a2;
}
template <class T>
T Math<T>::mul(T m1, T m2) {
    return m1 * m2;
}
template <class T>
T Math<T>::div(T d1, T d2) {
    return d1 / d2;
}
template class Math<int>;
template class Math<float>;
} // namespace aos
```

Template classes

- A class including generic type members

  Two instances are used to work on integer and floating points data

```
#include <iostream>
#include "math.h"
using namespace std;

int main() {
    aos::Math<int>   mI;
    aos::Math<float> mF;
    // Integer operations
    cout << "Add: " << mI.add(6, 10) << endl;
    cout << "Mul: " << mI.mul(12, 2) << endl;
    cout << "Div: " << mI.div(7, 3)  << endl;
    // Floating point operations
    cout << "Div: " << mF.div(7, 3)  << endl;
    return 0;
}
```

```
Add: 16
Mul: 24
Div: 2
Div: 2.33333
```

Standard Template Library (STL)

STL

- A powerful set of C++ template classes
- Provides general-purpose template-based classes and functions
- Implement commonly used algorithms and data structure
- *Containers* are the most noticeable example of data structures
  - List, queues, stack, map, set, …
  - Used to manage a collection of object of a certain class
  - The differ for functionality, typology (sequential, associative, unordered), and complexity of accessing and manipulation operations
- *Iterators* are used to step through the elements of the containers
- *Algorithms* includes function manipulating containers
  - Search, copy, sort, transform,...

For a complete list: http://www.cplusplus.com/reference/stl/

vector

- Sequence containers representing dynamic arrays (mutable size)
  - Contiguous of storage of elements with random access

```
int main() {
    People claire("Claire");
    People john("John");
    vector<People> citizens;      // Container of People
    citizens.push_back(john);     // Add to vector
    citizens.push_back(claire);   // Add to vector
    vector<People>::iterator it;
    for(it = citizens.begin(); it != citizens.end(); ++it) {
        cout << "Name: " << it->GetName() << endl;
    }
    cout << "Citizen 1 is" << citizens[1].GetName() << endl;
    return 0;
}
```

```
Name: John
Name: Claire
Citizen 1 is Claire
```

list

- Sequence containers implementing double-linked lists

  Constant time insert and erase operations on the object collection

```
int main() {
    People claire("Claire");
    People john("John");
    list<People> citizens;          // Container of People
    citizens.push_back(john);       // Add to vector
    citizens.push_back(claire);     // Add to vector
    list<People>::iterator it;
    for(it = citizens.begin(); it != citizens.end(); ++it) {
        cout << "Name: " << it->GetName() << endl;
    }
    return 0;
}
```

```
Name: John
Name: Claire
```

Outline

C++ Programs

- Memory allocation
- Passing parameters to functions

Classes

- Constructors and assignments
- Object pointers
- Qualifiers

Standard Template Library (STL)

- Templates
- Example: `vector<>`, `list<>`

Memory management

- From raw to smart pointers

Raw pointers

- Using (*raw*) pointers is a powerful C/C++ feature, but in general difficult to handle and error-prone
- Consider the following example:

```
int main() {
    People * joe = new People("Joe");
    People * ann = new People("Ann");
    People * alex = new People("Alex");
    joe->SetChild(alex);
    ann->SetChild(alex);

    delete joe;
    delete ann;
    delete alex;
    return 0;
}
```

```
class People {
public:
    virtual ~People();
    void SetChild(
        People * _child);
private:
    People * child;
}
```

```
People::~People() {
    delete child;
}
```

This ends in a "Segmentation fault" since "**alex**" is deallocated multiple times!

Smart pointers

- To solve the problem we should start from the concept of *ownership*
- To own a dynamically allocated object means to be the entity in charge of deallocating it

- *Smart pointers* is a feature introduced in C++ since standard C++11, to simplify the memory management in C++ programs
  - Concept of ownership introduced and exploited
  - Limited *garbage collection* facilities
- Base idea: *To wrap raw pointers in stack allocated objects that could control the life-cycle of the dynamically allocated object*
- We are going to focus on just two class of smart pointer
  - Unique pointers (class **`unique_ptr<>`**)
  - Shared pointers (class **`shared_ptr<>`**)

Shared pointers

- Pointers of type *shared* allows us to keep track of the number of shared ownership, i.e. references to the objects
- Reference counting based mechanism
  - Assignments and scope exit alter the reference counters
  - If the count is decreased to zero, the object is automatically deallocated
- This facility does not come for free
  - Reference counting mechanism requires a little additional memory overhead

```
#include <memory>
using namespace std;
int main() {
    shared_ptr<People> jack = make_shared<People>("Jack");
…
}
```

- Shared pointers can rely the developer from the burden of using explicit **`new/delete`** constructs

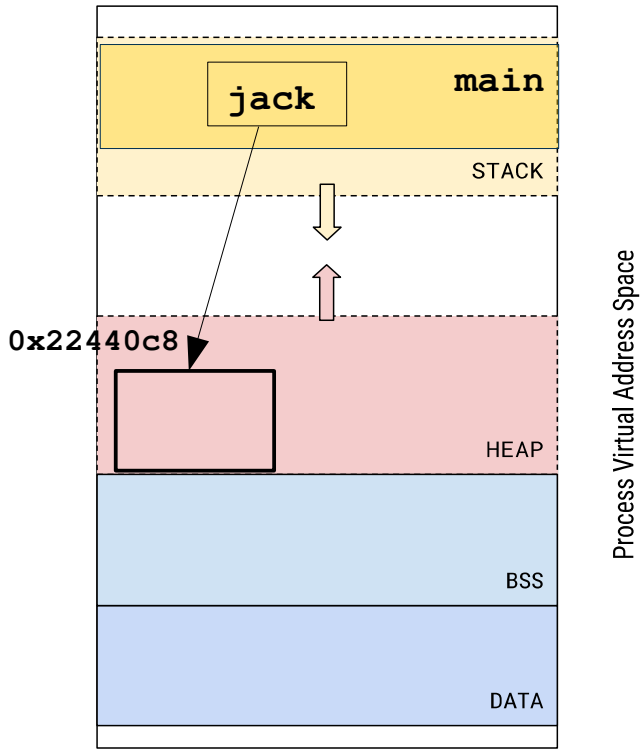
Memory management

Shared pointers

- Similarly to pointer variables...
  - Object “jack” (of class **shared_ptr<>**) is allocated onto the stack
  - Members of the pointed object are accessed by using the “**->**” operator
- **shared_ptr<>** member function **get()** returns the raw pointer value
  - Using “ . ” operator to access **shared_ptr<>** member functions

```
…
  shared_ptr<People> jack
        = make_shared<People>("Jack");
  cout << "Name = "<< jack->GetName() << endl;
  cout << "Raw ptr = " << jack.get() << endl;
…
```

```
Name = Jack
Raw pointer = 0x22440c8
```



Shared pointers

- Example: an assignment operation and a scope exit

```
#include <memory>
using namespace std;
int main() {
    …
    shared_ptr<People> jack = make_shared<People>("Jack");
    cout << "Jack ref.count = " << jack.use_count() << endl;
    {
        shared_ptr<People> jack2;
        jack2 = jack;
        cout << "Jack ref.count = " << jack.use_count() << endl;
    }
    cout << "Jack ref.count = " << jack.use_count() << endl;
    return 0;
}
```

```
Jack ref.count = 1
Jack ref.count = 2  // jack2 = jack
Jack ref.count = 1  // jack2 out of scope => jack ref.count decrease
Jack: i'm dying...  // People object “Jack” destruction
```

Shared pointers

- The previous “parents” example can be rewritten as follows, using shared pointers (changing also **`SetChild()`** signature)

```
#include <memory>
using namespace std;
int main() {
    shared_ptr<People> jack = make_shared<People>("Jack");
    shared_ptr<People> ann  = make_shared<People>("Ann");
    shared_ptr<People> alex = make_shared<People>("Alex");
    jack->SetChild(alex);
    ann->SetChild(alex);
    return 0;
}
```

- No need of explicit delete (objects deallocated at main function exit)
- No multiple deallocation thanks to reference counting

Unique pointers

- Do not share the ownership of a pointed object
  - Unique pointer means unique ownership
  - Only one object responsible of memory deallocation
- Copy assignments not allowed, only move assignments are

```
#include <memory>
using namespace std;
int main() {
    unique_ptr<People> jason(new People("Jason"));
    unique_ptr<People> laura;
    cout << "Jason = " << jason.get() << endl;
    laura = std::move(jason);
    cout << "Jason = " << jason.get() << endl;
    cout << "Laura = " << laura.get() << endl;
}
```

```
Jason = 0x15fd040
Jason = 0
Laura = 0x15fd040
```

Objects and containers

- Containers can be used to store objects (stack allocated)
  - An object copy is performed when a new object is added to the collection
- Containers can be used to store pointers to objects (heap allocated)

```
int main() {
    vector<People *> citizens;  // Using raw pointers
    People * p1 = new People("p1");
    People * p2 = new People("p2");
    citizens.push_back(p1);
    citizens.push_back(p2);
    citizens.clear();           // Clear the container
    return 0;
}
```

- For dynamically allocated objects keep in mind the potential pitfalls about memory management
  - `clear()` calls objects destructors... but what about raw pointers?

Objects and containers

- We may used smart (e.g., shared) pointers also in containers

```
int main() {
    vector<shared_ptr<People>> citizens; // Using shared pointers
    shared_ptr<People> p1 = make_shared<People>("p1");
    shared_ptr<People> p2 = make_shared<People>("p2");
    citizens.push_back(p1);
    citizens.push_back(p2);
    cout << "p1 use count = " << p1.use_count() << endl;
    return 0;
}
```

```
p1 use count = 2
p1: i'm dying...
p2: i'm dying...
p3: i'm dying...
```

- Insert/remove operations affect shared pointers reference counting
- Destroying the vector the reference counters are decreased

References

Web links

- http://www.cplusplus.com
- http://en.cppreference.com/w/
- http://www.horstmann.com/ccj2/ccjapp3.html
- http://www.cprogramming.com/tutorial/constructor_destructor_ordering.html
- http://www.yolinux.com/TUTORIALS/C++MemoryCorruptionAndMemoryLeaks.html
- http://duramecho.com/ComputerInformation/WhyHowCppConst.html
- http://www.tutorialspoint.com/cplusplus/cpp_stl_tutorial.htm