# Modern C++ Programming

# 11. Code Conventions

#### Federico Busato

University of Verona, Dept. of Computer Science 2021, v3.04



#### **Table of Context**

# 1 C++ Project Organization

- Project Directories
- Project Files
- src/include directories

# **2** Coding Styles and Conventions

- Coding Styles
- 3 #include

#### **Table of Context**

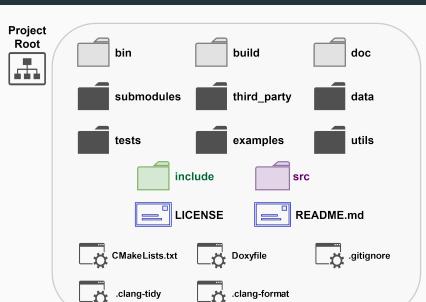
- Macro and Preprocessing
- 5 namespace
- Variables
- Functions
- Structs and Classes

### **Table of Context**

- 9 Control Flow
- **Modern C++ Features**
- Maintainability
- Naming and Formatting
- **E** Code Documentation

# C++ Project **Organization**

# **Project Organization**



#### **Fundamental directories**

include Project public header files
src/source Project source files and private headers
test Source files for testing the project

## **Empty directories**

bin Output executables

build All intermediate files

doc Project documentation

## **Optional directories**

```
submodules Project submodules
```

```
third_party (less often deps/external/extern)
dependencies or external libraries
```

data Files used by the executables or for testing

examples Source files for showing project features

utils (or script) Scripts and utilities related to the
 project

cmake CMake submodules (.cmake)

## **Project Files**

LICENSE Describes how this project can be used and distributed

README.md General information about the project in Markdown format \*

CMakeLists.txt Describes how to compile the project

**Doxyfile** Configuration file used by doxygen to generate the documentation (see next lecture)

others .gitignore, .clang-format, .clang-tidy, etc.

\* Markdown is a language with a syntax corresponding to a subset of HTML tags github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet

## Readme and License

#### **README**

- README template:
  - Embedded Artistry README Template
  - Your Project is Great, So Let's Make Your README Great Too

#### **LICENSE**

- Choose an open source license: choosealicense.com
- License guidelines: Why your academic code needs a software license

#### File extensions

#### Common C++ file extensions:

- header .h .hh .hpp .hxx
- header implementation .i.h, .i.hpp, -inl.h, .inl.hpp
  - separate implementation from interface for inline functions and templates
  - (2) keep implementation "inline" in the header file
- source/implementation .c .cc .cpp .cxx

### **Common conventions:**

- .h .c .cc GOOGLE
- .hh .cc
- .hpp .cpp
- hxx .cxx

#### src/include directories

## Organization:

- Public headers in include
- source files, private headers, header implementations in src/source directory
- The main file (if present) can be placed in src/source and called main.\* or placed in the project root directory with an arbitrary name

#### **Common Rules**

# The file should have the same name of the class/namespace that they implement

my\_class.hpp (MyClass.hpp)
my\_class.i.hpp (MyClass.i.hpp)
my\_class.cpp (MyClass.cpp)

```
mamespace my_np
my_np.hpp (MyNP.hpp)
my_np.i.hpp (MyNP.i.hpp)
my_np.cpp (MyNP.cpp)
```

# **Code Organization Example**

#### include

- my\_interface.hpp

#### src

- my\_class1.cpp
- my\_templ\_class.hpp
- my\_templ\_class.i.hpp (template/inline functions)
- my\_templ\_class.cpp
   (specialization)

#### subdir1

- my\_lib.hpp
- my\_lib.i.hpp
- my\_lib.cpp

- main.cpp (if necessary)
- README.md
- CMakeLists.txt
- Doxyfile
- LICENSE
- build (empty)
- bin (empty)
- doc (empty)
- test
  - test1.cpp
  - test2.cpp

# Coding Styles and Conventions

"one thing people should remember is there is what you can do in a language and what you should do"

Bjarne Stroustrup

# Most important rule: BE CONSISTENT!!

"The best code explains itself"

GOOGLE

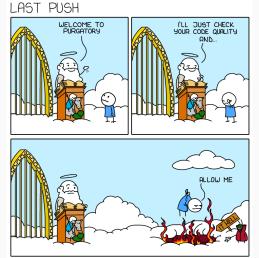
"80% of the lifetime cost of a piece of software goes to maintenance"

**Unreal Engine** 

# **Code Quality**

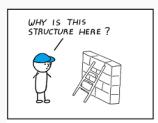
"The worst thing that can happen to a code base is size"

— Steve Yegge



#### **Bad Code**

#### How my code looks like for other people?











**Coding styles** are common guidelines to improve the *readability, maintainability,* prevent *common errors*, and make the code more *uniform* 

- LLVM Coding Standards
   llvm.org/docs/CodingStandards.html
- Google C++ Style Guide google.github.io/styleguide/cppguide.html
- Webkit Coding Style webkit.org/code-style-guidelines
- Mozilla Coding Style firefox-source-docs.mozilla.org

Chromium Coding Style

chromium.googlesource.com

- Unreal Engine Coding Standard
   docs.unrealengine.com/en-us/Programming
- μOS++
  micro-os-plus.github.io/develop/coding-style
  micro-os-plus.github.io/develop/naming-conventions
- High Integrity C++ Coding Standard
   www.perforce.com/resources
- CERT C++ Secure Coding
  Wiki.sei.cmu.edu

#### More comprehensive code guidelines

C++ Guidelines isocpp.github.io/CppCoreGuidelines/CppCoreGuidelines

### Critical system coding standards

- Misra Coding Standard www.misra.org.uk
- Autosar Coding Standard www.misra.org.uk
- Joint Strike Fighter Air Vehicle
   www.perforce.com/blog/qac/jsf-coding-standard-cpp

# Legend

### **※** → Important!

Highlight potential code issues such as bugs, inefficiency, and can compromise readability. Should not be ignored

#### \* $\rightarrow$ Useful

It is not fundamental but it emphasizes good practices. Should be followed if possible

■ → Minor / Obvious
 Style choice or not very common issue



#include

#### Every includes must be self-contained

- include every header you need directly
- the project must compile with any include order
- do not rely on recursive #include

LLVM, GOOGLE, UNREAL, 
$$\mu$$
OS++, CORE

#### \* Include as less as possible, especially in header files

- do not include unneeded headers
- minimize dependencies
- minimize code in headers (e.g. use forward declarations)
- it is not in contrast with the previous rule

```
LLVM, GOOGLE, CHROMIUM, UNREAL, HIC, \muOS++
```

#### Order of #include

#### LLVM, WEBKIT, CORE

- (1) Main Module Header (it is only one)
  - space
- (2) Local project includes (in alphetical order)
  - space
- (3) System includes (in alphetical order)

Note: (2) and (3) can be inverted

GOOGLE

System includes are self-contained, local includes might not

#### Project includes

LLVM, GOOGLE, WEBKIT, HIC, CORE

- \* Use "" syntax
- \* Should be <u>absolute paths</u> from the project include root e.g. #include "directory1/header.hpp"

#### System includes

LLVM, GOOGLE, WEBKIT, HIC

\* Use <> syntax e.g. #include <iostream>

- include guard vs. #pragma once
  - Use  ${\tt include}$  guard if portability is a strong requirement LLVM, GOOGLE, CHROMIUM, CORE
  - #pragma once otherwise

WebKit, Unreal

 #include preprocessor should be placed immediately after the header comment and include guard
 LLVM

#### Forward declarations vs. #includes

- Prefer forward declaration: reduce compile time, less dependency
   CHROMIUM
- Prefer #include: safer

GOOGLE

#### \* Use C++ headers instead of C headers:

```
<cassert> instead of <assert.h>
<cmath> instead of <math.h>, etc.
```

Report at least one function used for each include <iostream> // std::cout, std::cin

#### Example:

# Macro and

Preprocessing

**\*\* Avoid defining macros**, especially in headers GOOGLE

- Do not use macro for enumerators, constants, and functions WebKit, Google

\*\* Use a prefix for all macros related to the project
MYPROJECT\_MACRO
GOOGLE, UNREAL

# #undef macros wherever possible

GOOGLE

- Even in the source files if unity build is used

\* Always use curly brackets for multilines macro

```
#define MACRO \
{
    line1; \
    line2; \
}
```

\* Always put macros after #include

HIC

Put macros outside namespaces

## Style:

Close #endif with the respective condition of the first #if

```
#if defined(MACRO)
...
#endif // defined(MACRO)
```

 The hash mark that starts a preprocessor directive should always be at the beginning of the line

```
#if defined(MACRO)

# define MACRO2

#endif
```

Place the \ rightmost for multilines macro

Prefer #if defined(MACRO) instead of #ifdef MACRO

# namespace

- \*\* Avoid using namespace -directives at global scope LLVM, GOOGLE, WEBKIT, UNREAL, HIC,  $\mu$ OS++
- \* Limit using namespace -directives at local scope and prefer explicit namespace specification

GOOGLE, WEBKIT, UNREAL

- \*\* Always place code in a namespace to avoid global namespace pollution
  GOOGLE, WEBKIT
- \* Avoid anonymous namespaces in headers GOOGLE, CERT
- anonymous namespace vs. static
  - Prefer anonymous namespaces instead of static variables/functions  ${\small \mbox{Google, Core}}$
  - Use anonymous namespaces only for inline class declaration,
     static otherwise
     LLVM, STATIC<sup>30/68</sup>

#### Style guidelines:

The content of namespaces is not indented

```
LLVM, GOOGLE, WEBKIT
```

Close namespace declarations

```
} // namespace <namespace_identifier> LLVM
} // namespace (for anonymous namespaces) GOOGLE
```

#### Anonymous namespaces and source files:

Items local to a source file (e.g. .cpp) file should be wrapped in an anonymous namespace. While some such items are already file-scope by default in C++, not all are; also, shared objects on Linux builds export all symbols, so anonymous namespaces (which restrict these symbols to the compilation unit) improve function call cost and reduce the size of entry point tables

CHROMIUM, CORE, HIC31/68

## **Variables**

\*\* Place a variables in the *narrowest scope* possible, and *always initialize* variables in the declaration

GOOGLE, ISOCPP, MOZILLA, HIC, *mu*OS, CERT

\* Avoid static (non-const) global variables

LLVM, GOOGLE, CORE, HIC

- Use assignment syntax = when performing "simple" initialization
   CHROMIUM
- Declaration of pointer/reference variables or arguments may be placed with the asterisk/ampersand adjacent to either the type or to the variable name for all in the same way Google
  - char\* c; WebKit, Mozilla, Chromium, Unreal
  - char \*c;
  - char \* c;

- \*\* Use fixed-width integer type (e.g. int64\_t, int8\_t, etc.). Exception: int and unsigned GOOGLE, UNREAL
- \* size\_t vs. int64\_t
  - Use size\_t for object and allocation sizes, object counts, array and pointer offsets, vector indices, and so on. (integer overflow behavior for signed types is undefined)
  - Use int64\_t instead of size\_t for object counts and loop indices
- Use brace initialization to convert (constant) arithmetic types (narrowing) e.g. int64\_t{x}
   GOOGLE
- \* Use  ${\tt true}$ ,  ${\tt false}$  for boolean variables instead numeric values 0, 1 WebKit

- **\* Do not shift**  $\ll$  signed operands HIC, CORE,  $\mu OS$
- **\* Do not directly compare floating point == , < , etc.** HIC
- **\*\* Use signed types for arithmetic** Core

#### Style:

- Use floating-point literals to highlight floating-point data types, e.g. 30.0f
   WEBKIT (opposite)
- $\blacksquare$  Avoid redundant type, e.g. unsigned int , signed int  $$\operatorname{WebKit}$$

# **Functions**

- \* **Split up large functions** into logical sub-functions for improving readability and compile time

Unreal, Google, Core

- Use inline only for small functions (e.g. < 10 lines)  $_{
  m GOOGLE,\ HIC}$
- \*\* Never return pointers for new objects. Use std::unique\_ptr instead CHROMIUM, CORE

```
int* f() { return new int[10]; } // wrong!!
std::unique_ptr<int> f() { return new int[10]; } // correct
```

- f w Prefer pass by-reference instead by-value except for raw arrays and built-in types  $f w_{EBKIT}$
- \* Pass function arguments by const pointer or reference if those arguments are not intended to be modified by the function

  UNREAL
- Do not pass by-const-value for built-in types, especially in the declaration (same signature of by-value)
- \* Prefer returning values rather than output parameters

GOOGLE

\* Do not declare functions with an excessive number of parameters. Use a wrapper structure instead HIC, CORE

 All parameters should be aligned if they do not fit in a single line (especially in the declaration)

```
void f(int a,
     const int* b);
```

- Parameter names should be the same for declaration and definition
   CLANG-TIDY
- Do not use inline when declaring a function (only in the definition)
- Do not separate declaration and definition for template and inline functions
   GOOGLE

# \_\_\_\_\_

**Structs and Classes** 

- \* Use a struct only for passive objects that carry data; everything else is a class Google
- $\boldsymbol{\texttt{*}}$  Objects are fully initialized by constructor call  ${\tt Google, WebKit, Core}$
- \* Prefer in-class initializers to member initializers Core
- Use delegating constructors to represent common actions for all constructors of a class
- Initialize member variables in the order of member declaration

Core, Hic

\* Do not define implicit conversions. Use the explicit keyword for conversion operators and constructors

GOOGLE, CORE

Prefer = default constructors over user-defined / implicit
 default constructors
 MOZILLA, CHROMIUM, CORE, HIC

• Use = delete to mark deleted functions Core, Hic

Mark destructors noexcept

- Use braced initializer lists for aggregate types  $A\{1, 2\}$ ; LLVM, GOOGLE
- Do not use braced initializer lists {} for constructors. It can be confused with std::initializer\_list object LLVM

\* Avoid virtual method calls in constructors

GOOGLE, CORE, CERT

- \* Default arguments are allowed only on non-virtual functions
  GOOGLE, CORE, HIC
- \* Multiple inheritance and virtual inheritance are discouraged GOOGLE, CHROMIUM
- \* Prefer *composition* over *inheritance*

GOOGLE

\* A polymorphic class should suppress copying

Core

\* A class with a *virtual function* should have a *virtual or protected destructor* (e.g. interfaces and abstract classes)

CORE40/68

- \* Declare class data members in special way\*. Examples:
  - Trailing underscore (e.g. member\_var\_)

Google,  $\mu$ OS, Chromium

Leading underscore (e.g. \_member\_var )

.NET

- Public members (e.g. m\_member\_var )

WebKit

Class inheritance declarations order:
 public, protected, private

Google,  $\mu$ OS

- First data members, then function members
- If possible, avoid this-> keyword
- \*
- It helps to keep track of class variables and local function variables
- The first character is helpful in filtering through the list of available variables  $^{41/6}$

```
int x;
   float y;
};
class B {
public:
   B();
   void public_function();
protected:
                            // in general, it is not public in
   int _a;
                            // derived classes
   void _protected_function(); // "protected function()" is not wrong
                            // it may be public in derived classes
private:
   int x;
   float _y;
   void _private_function();
};
```

 In the constructor, each member should be indented on a separate line, e.g.
 WEBKIT, MOZILLA

```
A::A(int x1, int y1, int z1):
    x(x1),
    y(y1),
    z(z1) {
```

# Control Flow

- \* Avoid redundant control flow (see next slide)
  - Do not use else after a return / break  $LLVM,\;Mozilla,\;Chromium,\;WebKit$
  - Avoid return true/return false pattern
  - Merge multiple conditional statements
- \* Prefer switch to multiple if -statement

Core

Avoid do-while loop

CORE

Avoid goto

 $\mu$ OS, Core

 Do not use default labels in fully covered switches over enumerations

LLVM

```
if (condition) { // wrong!!
  < code1 >
  return;
else // <-- redundant
   < code2 >
if (condition) { // Corret
 < code1 >
  return;
< code2 >
if (condition) // wrong!!
  return true;
else
  return false;
return condition; // Corret
```

Use early exits (continue, break, return) to simplify the code

#### LLVM

Turn predicate loops into predicate functions

#### LLVM

\*\* Tests for null/non-null , and zero/non-zero should all be done with equality comparisons Core, Webkit (opposite) Mozilla

```
** Prefer (ptr == nullptr) and x > 0 over (nullptr == ptr) and 0 < x CHROMIUM
```

Do not compare to true/false, e.g. if (x == true)

HIC

- \* Do not mix signed and unsigned types
- Prefer signed integer (better 64-bit) for loop indices
- Prefer enum to bool on function parameters
- Prefer empty() method over size() to check if a container has no items
  MOZILLA
- Ensure that all statements are reachable

HIC

\* Avoid RTTI (dynamic\_cast) or exceptions if possible LLVM, GOOGLE, MOZILLA

\* The if and else keywords belong on separate lines

```
if (c1) <statement1>; else <statement2> // wrong!!

GOOGLE. WEBKIT
```

- Multi-lines statements and complex conditions require curly braces

  GOOGLE
- Boolean expression longer than the standard line length requires to be consistent in how you break up the lines

# Modern C++

# **Features**

#### Use modern C++ features wherever possible

- \* static\_cast reinterpret\_cast instead of old style cast (type) GOOGLE,  $\mu$ OS, HIC
- \* Do not define implicit conversions. Use the explicit keyword for conversion operators and constructors  ${\rm Google},\ \mu{\rm OS}$

- \* Use constexpr instead of macro Google, WebKit
- \* Use using instead typedef
- \* Prefer enum class instead of plain enum UNREAL,  $\mu OS$
- \* static\_assert compile-time assertion UNREAL, HIC
- \* lambda expression UNREAL
- **\* move semantic** UNREAL
- \*\* nullptr instead of 0 or NULL LLVM, GOOGLE, UNREAL, WEBKIT, MOZILLA, HIC,  $\mu$ OS 5

\* Use range-for loops whatever possible

LLVM, WebKit, Unreal, Core

\* Use auto to avoid type names that are noisy, obvious, or unimportant

- \* Use [[deprecated]] / [[noreturn]] /
  [[nodiscard]] to indicate deprecated functions / that do
  not return / result should not be discarded
- Avoid throw() expression. Use noexpect instead

- \*\* Use always override/final function member keyword
  WebKit, Mozilla, Unreal, Chromium, Hic
- \* Use braced direct-list-initialization or copy-initialization for setting default data member value. Avoid initialization in constructors if possible

  UNREAL

```
struct A {
   int x = 3;  // copy-initialization
   int x { 3 };  // direct-list-initialization (best option)
};
```

- Use = default constructors
- Use = delete to mark deleted functions
- Prefer uniform initialization when it cannot be confused with
   std::initializer\_list
   CHROMIUM<sup>53/68</sup>

Maintainability

#### Readability

- \* Write all code in English
- **\* Avoid complicated template programming** GOOGLE

- $\mbox{\ensuremath{\mathtt{\#}}}$  Use the assert to document preconditions and assumptions  $$\operatorname{LLVM}$$
- \* Use symbolic names instead of literal values in code HIC

\* Do not overload operators with special semantics && HIC

#### Readability

Prefer consecutive alignment

```
int     var1 = ...
long long int var2 = ...
```

- \* Write self-documenting code, e.g.  $(x + y 1) / y \rightarrow ceil\_div(x, y)$  UNREAL
- Minimize the number of empty rows
- Do not use more than one empty line

GOOGLE

Do not write excessive long file



#### **Spacing**

- Use always the same style for braces
  - Same line
  - Its own line

WEBKIT (func. only), MOZILLA
UNREAL, WEBKIT (function)
MOZILLA (Class)

### **Spacing**

- \* Use always the same indentation style
  - tab  $\rightarrow$  2 spaces

Google, Mozilla, Hic,  $\mu$ OS

- tab  $\rightarrow$  4 spaces

LLVM, Webkit, Hic,  $\mu$ OS

- tab = 4 spaces

UNREAL

 $\boldsymbol{\mathsf{X}}$  Separate commands, operators, etc., by a space  $\mathsf{LLVM},\ \mathsf{Google},\ \mathsf{WebKit}$ 

```
if(a*b<10&&c) // wrong!!
if (a * c < 10 && c) // correct
```

**\*\* Limit line length (width)** to be at most 80 characters long (or  $120) \rightarrow$  help code view on a terminal

LLVM, GOOGLE, MOZILLA,  $\mu$ OS

### Maintainability

- Address compiler warnings. Compiler warning messages mean something is wrong
   UNREAL
- Ensure ISO C++ compliant code and avoid non-standard extension, deprecated features, or asm declarations, e.g.
   register, \_\_attribute\_\_

  HIC
- Prefer sizeof(variable/value) instead of sizeof(type)

GOOGLE

Enforce const-correctness

Unreal

# Naming and

Formatting

#### **Naming Conventions**

#### General rule:

- imes Use full words, except in the rare case where an abbreviation would be more canonical and easier to understand Webkit
- Avoid short and very long names

## Style Conventions

Camel style Uppercase first word letter (sometimes called *Pascal style* or *Capital case*) (less readable, shorter names)

CamelStyle

**Snake style** Lower case words separated by single underscore (good readability, longer names)

snake\_style

**Macro style** Upper case words separated by single underscore (sometimes called *Screaming style*) (good readability, longer names)

MACRO\_STYLE

#### Variable Variable names should be nouns

- Camel style e.g. MyVar
- Snake style e.g. my\_var

LLVM. UNREAL Google,  $\mu$ OS

**Constant** • Camel style + k prefix, e.g. kConstantVar

- Google. Mozilla
- Macro style e.g. CONSTANT\_VAR WEBKIT, OPENSTACK
- **Enum** Camel style + k

e.g. enum MyEnum { kEnumVar1, kEnumVar2 }

GOOGLE

 Camel style e.g. enum MyEnum { EnumVar1, EnumVar2 }

LLVM, WebKit

Namespace • Snake style, e.g. my\_namespace

Google, LLVM

Camel style, e.g. MyNamespace

WebKit

#### **Typename** Should be nouns

- Camel style (including classes, structs, enums, typedefs, etc.)
  - e.g. HelloWorldClass

LLVM, GOOGLE, WEBKIT

Snake style

 $\mu$ OS (class)

#### **Functions**

\* Should be descriptive verb (as they represent actions)

WebKit

- \* Functions that return boolean values should start with boolean verbs, like is, has, should, does  $\mu OS$
- ullet Use set prefix for modifier methods Webkit
- $\hbox{ \begin{tabular}{ll} \blacksquare \end{tabular} Do not use $\gcd$ for observer (const) methods without \\ parameters $WEBKIT$ \\ \end{tabular}$
- Style:
  - Lowercase Camel style, e.g. myFunc()

LLVM

Uppercase Camel style for standard functions
 e.g. MyFunc()
 GOOGLE, MOZILLA, UNREAL

 Snake style for cheap functions e.g. my\_func()

GOOGLE, STD

#### **Macro and Files**

Macro Macro style e.g. MY\_MACRO

GOOGLE

File • Snake style (my\_file)

GOOGLE

Camel style (MyFile)

LLVM

#### **Other Naming Issues**

**\*** Do not use reserved names

Cert

- double underscore followed by any character \_\_var
- single underscore followed by uppercase \_VAR
- Use common loop variable names
  - i, j, k, l used in order
  - it for iterators

 $\blacksquare$  Never put trailing white space or tabs at the end of a line  $\frac{\text{Google, Mozilla}}{\text{Google, Mozilla}}$ 

Declare each identifier on a separate line in a separate declaration

HIC

ullet Only one space between statement and comment Webkit

**\*\* Use the same line ending** (e.g. ' ' ') for all files MOZILLA, CHROMIUM

\* Do not use UTF characters for portability, prefer ASCII

\* If UTF is needed, prefer UTF-8 encoding for portability Chromium

Close files with a blank line

Mozilla, Unreal

## Code

Documentation

\* Any file start with a license

LLVM, UNREAL

#### \* Each file should include

- Qauthor name, surname, affiliation, email
- @version
- Qdate e.g. year and month
- **@file** the purpose of the file

in both header and source files

- Document methods/classes/namespaces only in header files
- Include @param[in], @param[out], @param[in,out],
  @return tags
- Document ranges, impossible values, status/return values
   meaning
   UNREAL $^{68/68}$

- Use always the same style of comment
- Be aware of the comment style, e.g.

```
- Multiple lines
/**

* comment1

* comment2

*/
```

- single line
  /// comment
- Prefer // comment instead of /\* \*/  $\rightarrow$  allow string-search tools like grep to identify valid code lines HIC,  $\mu OS$
- Use anchors for indicating special issues: TODO , FIXME , BUG , etc.
   WEBKIT, CHROMIUM69/68