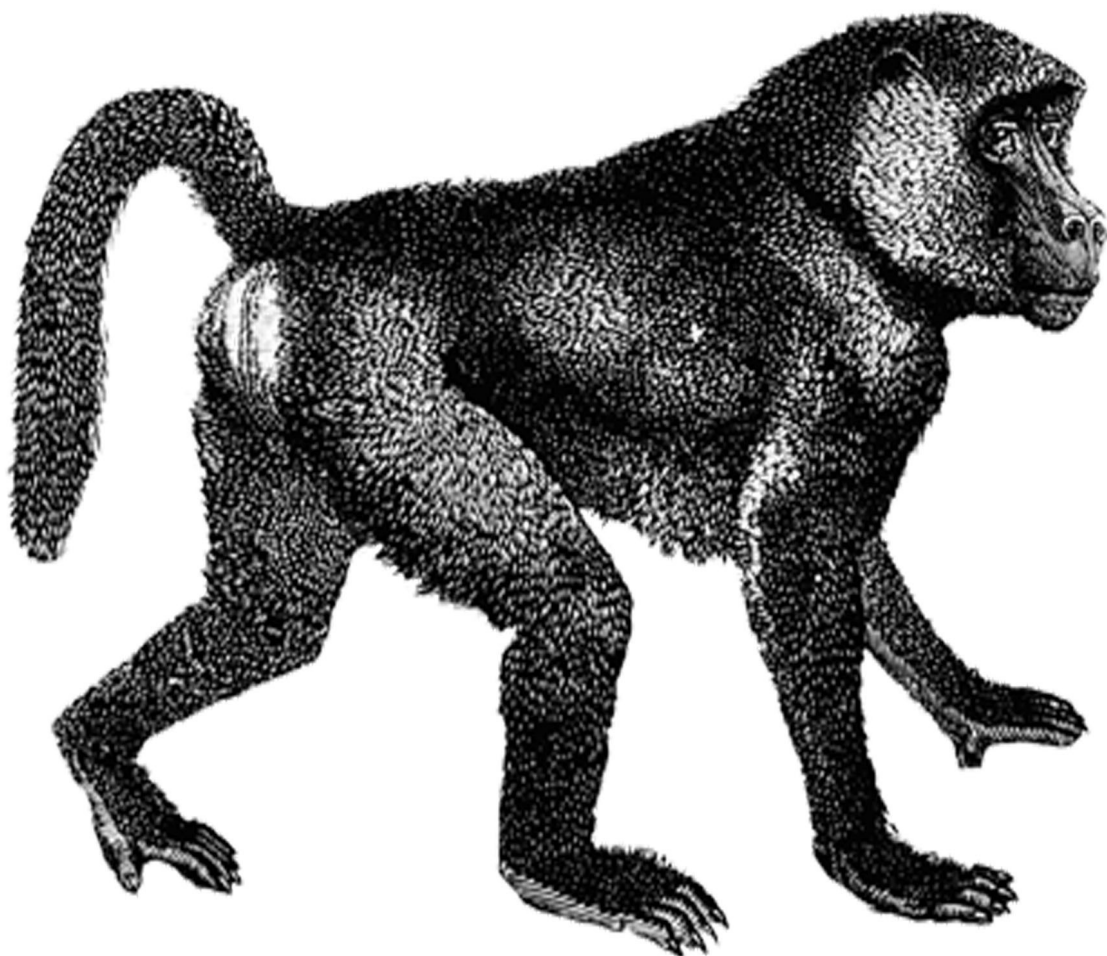


Everything is compiler.

高速上手

C++11/14

THE FASTEST GUIDE



欧长坤 著

cpp.ouchangkun.com

目錄

引言	1.1
第一章 C++11/14 简介	1.2
第二章 语言可用性的强化	1.3
第三章 语言运行期的强化	1.4
第四章 对标准库的扩充: 新增容器	1.5
第五章 对标准库的扩充: 智能指针和引用计数	1.6
第六章 对标准库的扩充: 正则表达式	1.7
第七章 对标准库的扩充: 语言级线程支持	1.8
第八章 其他杂项	1.9
第九章 扩展主题: C++17 简介	1.10

高速上手 C++ 11/14

引言

C++ 算是一个用户群体比较大的语言了，从 C++98 到 C++11 经历了长达十年多之久的积累，C++14 则是作为对 C++11 的重要补充和优化，所有这些新标准中扩充的特性，给 C++ 这门语言注入了新的活力。

那些还在坚持使用传统 C++（本教程把 C++98 及其之前的 C++ 特性均称之为传统 C++）而未接触过 C++11/14 的 C++ 程序员在见到诸如 Lambda 表达式这类全新特性时，甚至会流露出『学的不是同一门语言』的惊叹之情。

C++1x (本教程中指 C++11/14, 甚至 C++17) 为传统 C++ 注入的大量特性使得整个 C++ 变得更加像一门现代化的语言。C++1x 不仅仅增强了 C++ 语言自身的可用性，`auto` 关键字语义的修改使得我们更加有信心来操控极度复杂的模板类型。同时还对语言运行期进行了大量的强化，Lambda 表达式的出现让 C++ 具有了『匿名函数』的『闭包』特性，而这一特性几乎在现代的编程语言（诸如 Python/Swift/...）中已经司空见惯，右值引用的出现解决了 C++ 长期以来被人诟病的临时对象效率问题等等。

C++1x 为自身的标准库增加了非常多的工具和方法，诸如在语言层面上提供了 `std::thread` 支持了并发编程，在不同平台上不再依赖于系统底层的 API，实现了语言层面的跨平台支持；`std::regex` 提供了完整的正则表达式支持等等。C++98 已经被实践证明了一种非常成功的『范型』，而 C++1x 的出现，则进一步推动这种范型，让 C++ 成为系统程序设计和库开发更好的语言。

目标读者

1. 本教程假定读者已经熟悉了传统 C++，至少在阅读传统 C++ 代码上不具备任何困难。换句话说，那些长期使用传统 C++ 进行编码的人、渴望在短时间内迅速了解 `Modern C++` 特性的人非常适合阅读本书；
2. 本教程一定程度上介绍了一些 C++1x 的黑魔法，但这些魔法毕竟有限，不适合希望进阶学习 C++1x 的读者，本教程的定位系 **C++1x** 的快速上手。当然，希望进阶学习的读者可以使用本书来回顾并检验自己对 `Modern C++` 的熟悉度。

本书目的

本教程虽号称高速上手教程，但实际上对 C++11/14 的相关特性做了一个较为全面的介绍，读者可以自行根据下面的目录选取感兴趣的内容进行学习，快速熟悉需要了解的内容，这从某种意义上来说，也算是高速上手了。

这些特性并不需要全部掌握，只需针对特定的应用场景，学习、查阅最适合自己的新特性即可。

值得一提的是，本教程在介绍这些特性的过程中，尽可能简单明了的介绍了这些特性产生的历史背景和技术需求，这为理解这些特性、运用这些特性提供了很大的帮助。

内容一览

- 第一章 C++11/14 简介
 - 概述
 - 被弃用的特性
 - 与 C 的兼容性
- 第二章 语言可用性的强化
 - `nullptr` 与 `constexpr`
 - 类型推导
 - `auto`
 - `decltype`
 - 尾返回类型、`auto` 与 `decltype` 配合
 - 区间迭代
 - 基于范围的 for 循环
 - 初始化列表
 - `std::initializer_list`
 - 统一初始化语法
 - 模板增强
 - 外部模板
 - 尖括号 `>`
 - 类型别名模板

- 变长参数模板
- 面向对象增强
 - 委托构造
 - 继承构造
 - 显式虚函数重载
 - `override`
 - `final`
 - 显式禁用默认函数
- 强类型枚举
- 第三章 语言运行期的强化
 - lambda 表达式
 - lambda 表达式基础
 - 值捕获
 - 引用捕获
 - 隐式捕获
 - 表达式捕获
 - 泛型 lambda
 - 函数对象包装器
 - `std::function`
 - `std::bind/std::placeholder`
 - 右值引用
 - 左值、右值的纯右值、将亡值、右值
 - 右值引用和左值引用
 - 移动语义
 - 完美转发
- 第四章 对标准库的扩充: 新增容器
 - `std::array`
 - `std::forward_list`
 - `std::unordered_set`
 - `std::unordered_map`
 - `std::tuple`
 - 基本操作
 - 运行期索引
 - 合并与迭代
- 第五章 对标准库的扩充: 智能指针和引用计数
 - RAII 与引用计数

- `std::shared_ptr`
- `std::weak_ptr`
- `std::unique_ptr`
- 第六章 对标准库的扩充: 正则表达式库
 - 正则表达式简介
 - 普通字符
 - 特殊字符
 - 限定符
 - `std::regex` 及其相关
 - `std::regex`
 - `std::regex_match`
 - `std::match_results`
- 第七章 对标准库的扩充: 语言级线程支持
 - `std::thread`
 - `std::mutex`
 - `std::unique_lock`
 - `std::future`
 - `std::packaged_task`
 - `std::condition_variable`
- 第八章 其他杂项
 - 新类型
 - `long long int`
 - `noexcept` 的修饰和操作
 - 字面量
 - 原始字符串字面量
 - 自定义字面量
- 第九章 扩展主题: C++17 简介
 - 主要入选特性
 - 非类型模板参数的 `auto`
 - `std::variant<>`
 - 结构化绑定(Structured bindings)
 - 变量声明的强化
 - 未入选特性
 - Concepts

赞助

如果你认为本教程对你起到了帮助并希望赞助作者，可以通过下面的二维码给予支持。



交流

1. 本教程在每节的最下方提供了评论，如果读者发现教程中内容的错误，可以使用评论或者通过发 `issue` 来指出；
2. 本书依然有很多特性没有参与介绍，例如 `alignas` 内存对齐、无限制联合等，主要是考虑到这些特性的使用频次实在是太低，故没有多做介绍，若对未提及的特性有需求，笔者可以考虑将其加入 `第八章`；
3. 笔者是一枚 `C++` 新手，所以水平有限，若有任何错误，欢迎指出，让我把这个教程做得更好；
4. 本教程有以下交流群，有兴趣的读者可以加入，加群请注明 `gitbook`：



群名称: Modern C++ 交流群
群 号: 306196433

GitHub

本书中涉及的相关代码可以在 GitHub 上查看：

<https://github.com/changkun/cpp1x-tutorial>

致谢

笔者感谢以下读者指出本书中出现的错误：

[Recolic Keghart](#), [sinomiko](#), [Wang Zenhua](#), [asm warrior](#)

版权声明



本教程系 [欧长坤\(欧龙崎\)](#) 为 [实验楼](#) 编写的基础课教程，采用[知识共享署名-非商业性使用-禁止演绎 4.0 国际许可协议](#)进行许可。

教程地址：<https://www.shiyanlou.com/courses/605>

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook
该文件修改时间：2017-11-16 09:30:52

第一章 C++11/14 简介

一、被弃用的特性

在学习 C++1x 之前，我们先了解一下从 C++11 开始，被弃用的主要特性：

注意：弃用不等于废弃，只是用于暗示程序员这些特性将从未来的标准中消失，应该尽量避免使用。但是，已弃用的特性依然是标准库的一部分，并且出于兼容性的考虑，这些特性其实会『永久』保留。

- 如果一个类有析构函数，为其生成拷贝构造函数和拷贝赋值运算符的特性被弃用了。
- 不再允许字符串字面值常量赋值给一个 `char *`。如果需要用字符串字面值常量赋值和初始化一个 `char *`，应该使用 `const char *` 或者 `auto`。

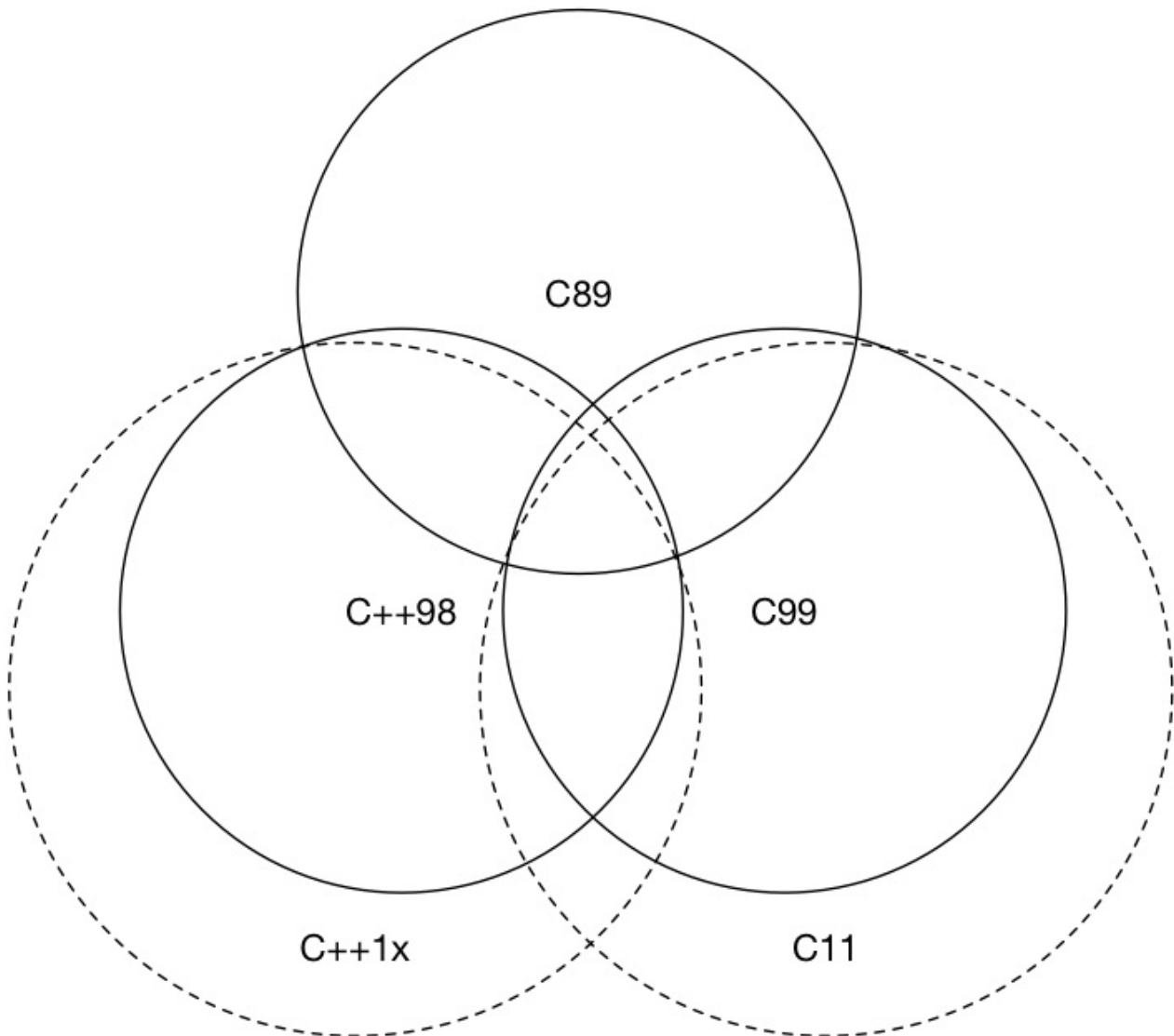
```
char *str = "hello world!"; // 将出现弃用警告
```

- C++98 异常说明、`unexcepted_handler`、`set_unexpected()` 等相关特性被弃用，应该使用 `noexcept`。
- `auto_ptr` 被弃用，应使用 `unique_ptr`。
- `register` 关键字被弃用。
- `bool` 类型的 `++` 操作被弃用。
- C 语言风格的类型转换被弃用，应该使用 `static_cast`、`reinterpret_cast`、`const_cast` 来进行类型转换。

还有一些其他诸如参数绑定（C++11 提供了 `std::bind` 和 `std::function`）、`export` 等特性也均被弃用。前面提到的这些特性如果你从未使用或者听说过，也请不要尝试去了解他们，应该向新标准靠拢，直接学习新特性。毕竟，技术是向前发展的。

二、与 C 的兼容性

出于一些不可抗力、历史原因，我们不得不在 C++ 中使用一些 C 语言代码（甚至古老的 C 语言代码），例如 Linux 系统调用。在 C++11 出现之前，大部分人当谈及『C 与 C++ 的区别是什么』时，普遍除了回答面向对象的类特性、泛型编程的模板特性外，就没有其他的看法了，甚至直接回答『差不多』，也是大有人在。下面的韦恩图大致上回答了 C 和 C++ 相关的兼容情况：



从现在开始，你的脑子里应该树立『**C++** 不是 **C** 的一个超集』这个观念（而且从一开始就不是，后面的进一步阅读的参考文献中给出了 C++98 和 C99 之间的区别）。在编写 C++ 时，也应该尽可能的避免使用诸如 `void*` 之类的程序风格。而在不得不使用 C 时，应该注意使用 `extern "C"` 这种特性，将 C 语言的代码与 C++ 代码进行分离编译，再统一链接这种做法，例如：

```
// foo.h
#ifdef __cplusplus
extern "C" {
#endif

int add(int x, int y);

#ifdef __cplusplus
}
#endif

// foo.c
int add(int x, int y) {
    return x+y;
}

// main.cpp
#include "foo.h"
int main() {
    add(1, 2);
    return 0;
}
```

应先使用 `gcc` 编译 C 语言的代码：

```
gcc -c foo.c
```

编译出 `foo.o` 文件，再使用 `g++` 将 C++代码和 `.o` 文件链接起来（或者都编译为 `.o` 再统一链接）：

```
g++ main.cpp foo.o -o main
```

进一步阅读的参考资料

1. C++ 语言导学. Bjarne Stroustrup
2. [C++ 历史](#)

- 3. C++ 1x 特性在 GCC/Clang 等编译器中的支持情况
- 4. C++98 与 C99 之间的区别

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by
Gitbook 该文件修改时间 : 2017-11-16 09:30:52

第二章 语言可用性的强化

一、本节内容

本节内容包括：

- 语言可用性的强化
 - `nullptr` 与 `constexpr`
 - 类型推导
 - `auto`
 - `decltype`
 - 尾返回类型、`auto` 与 `decltype` 配合
 - 区间迭代
 - 基于范围的 for 循环
 - 初始化列表
 - `std::initializer_list`
 - 统一初始化语法
 - 模板增强
 - 外部模板
 - 尖括号 `>`
 - 类型别名模板
 - 变长参数模板
 - 面向对象增强
 - 委托构造
 - 继承构造
 - 显式虚函数重载
 - `override`
 - `final`
 - 显式禁用默认函数
 - 强类型枚举
 - 总结

二、`nullptr` 与 `constexpr`

`nullptr`

`nullptr` 出现的目的是为了替代 `NULL`。在某种意义上来说，传统 C++ 会把 `NULL`、`0` 视为同一种东西，这取决于编译器如何定义 `NULL`，有些编译器会将 `NULL` 定义为 `((void*)0)`，有些则会直接将其定义为 `0`。

C++ 不允许直接将 `void *` 隐式转换到其他类型，但如果 `NULL` 被定义为 `((void*)0)`，那么当编译

```
char *ch = NULL;
```

时，`NULL` 只好被定义为 `0`。而这依然会产生问题，将导致了 C++ 中重载特性会发生混乱，考虑：

```
void foo(char *);  
void foo(int);
```

对于这两个函数来说，如果 `NULL` 又被定义为了 `0` 那么 `foo(NULL);` 这个语句将会去调用 `foo(int)`，从而导致代码违反直观。

为了解决这个问题，C++11 引入了 `nullptr` 关键字，专门用来区分空指针、`0`。`nullptr` 的类型为 `nullptr_t`，能够隐式的转换为任何指针或成员指针的类型，也能和他们进行相等或者不等的比较。

你可以尝试使用 `gcc` 和 `g++` 两个编译器同时编译下面的代码：

```
#include <iostream>
void foo(char *);
void foo(int);
int main() {

    if(NULL == (void *)0) std::cout << "NULL == 0" << std::endl;
    else std::cout << "NULL != 0" << std::endl;

    foo(0);
    // foo(NULL); // 编译无法通过
    foo(nullptr);

    return 0;
}
void foo(char *ch) {
    std::cout << "call foo(char*)" << std::endl;
}
void foo(int i) {
    std::cout << "call foo(int)" << std::endl;
}
```

将输出：

```
NULL == 0
call foo(int)
call foo(char*)
```

所以，当需要使用 `NULL` 时候，请养成直接使用 `nullptr` 的习惯。

constexpr

C++ 本身已经具备了常数表达式的概念，比如 `1+2`, `3*4` 这种表达式总是会产生相同的结果并且没有任何副作用。如果编译器能够在编译时就把这些表达式直接优化并植入到程序运行时，将能增加程序的性能。一个非常显著的例子就是在数组的定义阶段：

```
#define LEN 10

int len_foo() {
    return 5;
}

int main() {
    char arr_1[10];
    char arr_2[LEN];
    int len = 5;
    char arr_3[len];           // 非法
    const int len_2 = 10;
    char arr_4[len_2];         // 合法
    char arr_5[len_foo()+5];   // 非法

    return 0;
}
```

在 C++11 之前，可以在常量表达式中使用的变量必须被声明为 `const`，在上面代码中，`len_2` 被定义成了常量，因此 `len_2` 是一个常量表达式，所以能够合法的分配一个数组；

而对于 `arr_5` 来说，C++98 之前的编译器无法得知 `len_foo()` 在运行期实际上是返回一个常数，这也就导致了非法的产生。

C++11 提供了 `constexpr` 让用户显式的声明函数或对象构造函数在编译器会成为常数，这个关键字明确的告诉编译器应该去验证 `len_foo` 在编译器就应该是一个常数。

此外，`constexpr` 的函数可以使用递归：

```
constexpr int fibonacci(const int n) {
    return n == 1 || n == 2 ? 1 : fibonacci(n-1)+fibonacci(n-2);
}
```

从 C++14 开始，`constexpr` 函数可以在内部使用局部变量、循环和分支等简单语句，例如下面的代码在 C++11 的标准下是不能够通过编译的：


```
constexpr int fibonacci(const int n) {  
    if(n == 1) return 1;  
    if(n == 2) return 1;  
    return fibonacci(n-1)+fibonacci(n-2);  
}
```

三、类型推导

在传统 C 和 C++ 中，参数的类型都必须明确定义，这其实对我们快速进行编码没有任何帮助，尤其是当我们面对一大堆复杂的模板类型时，必须明确的指出变量的类型才能进行后续的编码，这不仅拖慢我们的开发效率，也让代码变得又臭又长。

C++11 引入了 `auto` 和 `decltype` 这两个关键字实现了类型推导，让编译器来操心变量的类型。这使得 C++ 也具有了和其他现代编程语言一样，某种意义上提供了无需操心变量类型的使用习惯。

auto

`auto` 在很早以前就已经进入了 C++，但是他始终作为一个存储类型的指示符存在，与 `register` 并存。在传统 C++ 中，如果一个变量没有声明为 `register` 变量，将自动被视为一个 `auto` 变量。而随着 `register` 被弃用，对 `auto` 的语义变更也就非常自然了。

使用 `auto` 进行类型推导的一个最为常见而且显著的例子就是迭代器。在以前我们需要这样来书写一个迭代器：

```
for(vector<int>::const_iterator itr = vec.cbegin(); itr != vec.c  
end(); ++itr)
```

而有了 `auto` 之后可以：

```
// 由于 cbegin() 将返回 vector<int>::const_iterator  
// 所以 itr 也应该是 vector<int>::const_iterator 类型  
for(auto itr = vec.cbegin(); itr != vec.cend(); ++itr);
```

一些其他的常见用法：

```
auto i = 5; // i 被推导为 int
auto arr = new auto(10) // arr 被推导为 int *
```

注意：`auto` 不能用于函数传参，因此下面的做法是无法通过编译的（考虑重载的问题，我们应该使用模板）：

```
int add(auto x, auto y);
```

此外，`auto` 还不能用于推导数组类型：

```
#include <iostream>

int main() {
    auto i = 5;

    int arr[10] = {0};
    auto auto_arr = arr;
    auto auto_arr2[10] = arr;

    return 0;
}
```

decltype

`decltype` 关键字是为了解决 `auto` 关键字只能对变量进行类型推导的缺陷而出现的。它的用法和 `sizeof` 很相似：

```
decltype(表达式)
```

有时候，我们可能需要计算某个表达式的类型，例如：

```
auto x = 1;
auto y = 2;
decltype(x+y) z;
```

尾返回类型、**auto** 与 **decltype** 配合

你可能会思考，`auto` 能不能用于推导函数的返回类型。考虑这样一个例子加法函数的例子，在传统 C++ 中我们必须这么写：

```
template<typename R, typename T, typename U>
R add(T x, U y) {
    return x+y
}
```

`typename` 和 `class` 在模板中没有区别，在 `typename` 这个关键字出现之前，都是使用 `class` 来定义模板参数的

这样的代码其实变得很丑陋，因为程序员在使用这个模板函数的时候，必须明确指出返回类型。但事实上我们并不知道 `add()` 这个函数会做什么样的操作，获得一个什么样的返回类型。

在 C++11 中这个问题得到解决。虽然你可能马上会反应出来使用 `decltype` 推导 `x+y` 的类型，写出这样的代码：

```
decltype(x+y) add(T x, U y)
```

但事实上这样的写法并不能通过编译。这是因为在编译器读到 `decltype(x+y)` 时，`x` 和 `y` 尚未被定义。为了解决这个问题，C++11 还引入了一个叫做尾返回类型（trailing return type），利用 `auto` 关键字将返回类型后置：

```
template<typename T, typename U>
auto add(T x, U y) -> decltype(x+y) {
    return x+y;
}
```

令人欣慰的是从 C++14 开始是可以直接让普通函数具备返回值推导，因此下面的写法变得合法：

```
template<typename T, typename U>
auto add(T x, U y) {
    return x+y
}
```

四、区间迭代

基于范围的 **for** 循环

终于，C++11 引入了基于范围的迭代写法，我们拥有了能够写出像 Python 一样简洁的循环语句：

```
int array[] = {1,2,3,4,5};
for(auto &x : array) {
    std::cout << x << std::endl;
}
```

最常用的 `std::vector` 遍历将从原来的样子：

```
std::vector<int> arr(5, 100);
for(std::vector<int>::iterator i = arr.begin(); i != arr.end();
++i) {
    std::cout << *i << std::endl;
}
```

变得非常的简单：

```
// & 启用了引用，如果没有则对 arr 中的元素只能读取不能修改
for(auto &i : arr) {
    std::cout << i << std::endl;
}
```

五、初始化列表

初始化是一个非常重要的语言特性，最常见的就是对对象进行初始化。在传统 C++ 中，不同的对象有着不同的初始化方法，例如普通数组、POD（plain old data，没有构造、析构和虚函数的类或结构体）类型都可以使用 `{}` 进行初始化，也就是我们所说的初始化列表。而对于类对象的初始化，要么需要通过拷贝构造、要么就需要使用 `()` 进行。这些不同方法都针对各自对象，不能通用。

```
int arr[3] = {1,2,3};    // 列表初始化

class Foo {
private:
    int value;
public:
    Foo(int) {}
};

Foo foo(1);              // 普通构造初始化
```

为了解决这个问题，C++11 首先把初始化列表的概念绑定到了类型上，并将其称之为 `std::initializer_list`，允许构造函数或其他函数像参数一样使用初始化列表，这就为类对象的初始化与普通数组和 POD 的初始化方法提供了统一的桥梁，例如：

```
#include <initializer_list>

class Magic {
public:
    Magic(std::initializer_list<int> list);
};

Magic magic = {1,2,3,4,5};

std::vector<int> v = {1, 2, 3, 4};
```

这种构造函数被叫做初始化列表构造函数，具有这种构造函数的类型将在初始化时被特殊关照。

初始化列表除了用在对象构造上，还能将其作为普通函数的形参，例如：

```
void foo(std::initializer_list<int> list);

foo({1,2,3});
```

其次，C++11 提供了统一的语法来初始化任意的对象，例如：

```
struct A {
    int a;
    float b;
};
struct B {

    B(int _a, float _b): a(_a), b(_b) {}
private:
    int a;
    float b;
};

A a {1, 1.1};      // 统一的初始化语法
B b {2, 2.2};
```

六、模板增强

外部模板

传统 C++ 中，模板只有在使用时才会被编译器实例化。换句话说，只要在每个编译单元（文件）中编译的代码中遇到了被完整定义的模板，都会实例化。这就产生了重复实例化而导致的编译时间的增加。并且，我们没有办法通知编译器不要出发模板实例化。

C++11 引入了外部模板，扩充了原来的强制编译器在特定位置实例化模板的语法，使得能够显式的告诉编译器何时进行模板的实例化：

```
template class std::vector<MagicClass>;          // 强行实例化
extern template class std::vector<MagicClass>;    // 不在该编译文件中
实例化模板
```

尖括号 ">"

在传统 C++ 的编译器中，`>>` 一律被当做右移运算符来进行处理。但实际上我们很容易就写出了嵌套模板的代码：

```
std::vector<std::vector<int>>> mtx;
```

这在传统 C++ 编译器下是不能够被编译的，而 C++11 开始，连续的右尖括号将变得合法，并且能够顺利通过编译。甚至于下下面这种写法都能够通过编译：

```
template<bool T> SuckType;  
std::vector<SuckType<(1>2)>> v; // 合法，但不建议写出这样的代码
```

类型别名模板

在了解类型别名模板之前，需要理解『模板』和『类型』之间的不同。仔细体会这句话：模板是用来产生类型的。在传统 C++ 中，`typedef` 可以为类型定义一个新的名称，但是却没有办法为模板定义一个新的名称。因为，模板不是类型。例如：

```
template<typename T, typename U>  
class SuckType;  
  
typedef SuckType<std::vector, std::string> NewType; // 不合法
```

C++11 使用 `using` 引入了下面这种形式的写法，并且同时支持对传统 `typedef` 相同的功效：

通常我们使用 `typedef` 定义别名的语法是：`typedef 原名称 新名称;`，但是对函数指针等别名的定义语法却不相同，这通常给直接阅读造成了一定程度的困难。

```
typedef int (*process)(void *); // 定义了一个返回类型为 int，参数为  
void* 的函数指针类型，名字叫做 process  
using process = void(*)(void *); // 同上，更加直观  
using NewType = SuckType<std::vector, std::string>;
```

默认模板参数

我们可能定义了一个加法函数：

```
template<typename T, typename U>
auto add(T x, U y) -> decltype(x+y) {
    return x+y
}
```

但在使用时发现，要使用 `add`，就必须每次都指定其模板参数的类型。

在 C++11 中提供了一种便利，可以指定模板的默认参数：

```
template<typename T = int, typename U = int>
auto add(T x, U y) -> decltype(x+y) {
    return x+y
}
```

变长参数模板

模板一直是 C++ 所独有的黑魔法（一起念：**Dark Magic**）之一。在 C++11 之前，无论是类模板还是函数模板，都只能按其指定的样子，接受一组固定数量的模板参数；而 C++11 加入了新的表示方法，允许任意个数、任意类别的模板参数，同时也不需要再定义时将参数的个数固定。

```
template<typename... Ts> class Magic;
```

模板类 `Magic` 的对象，能够接受不受限制个数的 `typename` 作为模板的形式参数，例如下面的定义：

```
class Magic<int,
            std::vector<int>,
            std::map<std::string,
                    std::vector<int>>>> darkMagic;
```


既然是任意形式，所以个数为0的模板参数也是可以的：`class Magic<> nothing;`。

如果不希望产生的模板参数个数为0，可以手动的定义至少一个模板参数：

```
template<typename Require, typename... Args> class Magic;
```

变长参数模板也能被直接调整到模板函数上。传统 C 中的 `printf` 函数，虽然也能达成不定个数的形参的调用，但其并非类别安全。而 C++11 除了能定义类别安全的变长参数函数外，还可以使类似 `printf` 的函数能自然地处理非自带类别的对象。除了在模板参数中能使用 `...` 表示不定长模板参数外，函数参数也使用同样的表示法代表不定长参数，这也就为我们简单编写变长参数函数提供了便捷的手段，例如：

```
template<typename... Args> void printf(const std::string &str, Args... args);
```

那么我们定义了变长的模板参数，如何对参数进行解包呢？

首先，我们可以使用 `sizeof...` 来计算参数的个数，：

```
template<typename... Args>
void magic(Args... args) {
    std::cout << sizeof...(args) << std::endl;
}
```

我们可以传递任意个参数给 `magic` 函数：

```
magic();           // 输出0
magic(1);          // 输出1
magic(1, "");      // 输出2
```

其次，对参数进行解包，到目前为止还没有一种简单的方法能够处理参数包，但有两种经典的处理手法：

1. 递归模板函数

递归是很容易想到的一种手段，也是最经典的处理方法。这种方法不断递归地向函数传递模板参数，进而达到递归遍历所有模板参数的目的：

```
#include <iostream>
template<typename T>
void printf(T value) {
    std::cout << value << std::endl;
}
template<typename T, typename... Args>
void printf(T value, Args... args) {
    std::cout << value << std::endl;
    printf(args...);
}
int main() {
    printf(1, 2, "123", 1.1);
    return 0;
}
```

2. 初始化列表展开

这个方法需要之后介绍的知识，读者可以简单阅读以下，将这个代码段保存，在后面的内容了解过了之后再回过头来阅读此处方法会大有收获。

递归模板函数是一种标准的做法，但缺点显而易见的在于必须定义一个终止递归的函数。

这里介绍一种使用初始化列表展开的黑魔法：

```
// 编译这个代码需要开启 -std=c++14
template<typename T, typename... Args>
auto print(T value, Args... args) {
    std::cout << value << std::endl;
    return std::initializer_list<T>{([&] {
        std::cout << args << std::endl;
    })(), value)...};
}

int main() {
    print(1, 2.1, "123");
    return 0;
}
```

在这个代码中，额外使用了 C++11 中提供的初始化列表以及 Lambda 表达式的特性（下一节中将提到），而 `std::initializer_list` 也是 C++11 新引入的容器（以后会介绍到）。

通过初始化列表，`(lambda 表达式, value)...` 将会被展开。由于逗号表达式的出现，首先会执行前面的 lambda 表达式，完成参数的输出。唯一不美观的地方在于如果不使用 `return` 编译器会给出未使用的变量作为警告。

事实上，有时候我们虽然使用了变参模板，却不一定需要对参数做逐个遍历，我们可以利用 `std::bind` 及完美转发等特性实现对函数和参数的绑定，从而达到成功调用的目的。

关于这方面的使用技巧，可以通过项目课：[100 行 C++ 代码实现线程池](#) 进行进一步巩固学习。

七、面向对象增强

委托构造

C++11 引入了委托构造的概念，这使得构造函数可以在同一个类中一个构造函数调用另一个构造函数，从而达到简化代码的目的：

```
class Base {
public:
    int value1;
    int value2;
    Base() {
        value1 = 1;
    }
    Base(int value) : Base() { // 委托 Base() 构造函数
        value2 = 2;
    }
};

int main() {
    Base b(2);
    std::cout << b.value1 << std::endl;
    std::cout << b.value2 << std::endl;
}
```

继承构造

在传统 C++ 中，构造函数如果需要继承是需要将参数一一传递的，这将导致效率低下。C++11 利用关键字 `using` 引入了继承构造函数的概念：

```
class Base {
public:
    int value1;
    int value2;
    Base() {
        value1 = 1;
    }
    Base(int value) : Base() { // 委托 Base() 构造函数
        value2 = 2;
    }
};

class Subclass : public Base {
public:
    using Base::Base;          // 继承构造
};

int main() {
    Subclass s(3);
    std::cout << s.value1 << std::endl;
    std::cout << s.value2 << std::endl;
}
```

显式虚函数重载

在传统 C++ 中，经常容易发生意外重载虚函数的事情。例如：

```
struct Base {
    virtual void foo();
};

struct SubClass: Base {
    void foo();
};
```

`SubClass::foo` 可能并不是程序员尝试重载虚函数，只是恰好加入了一个具有相同名字的函数。另一个可能的情形是，当基类的虚函数被删除后，子类拥有旧的函数就不再重载该虚拟函数并摇身一变成为了一个普通的类方法，这将造成灾难性的后果。

C++11 引入了 `override` 和 `final` 这两个关键字来防止上述情形的发生。

override

当重载虚函数时，引入 `override` 关键字将显式的告知编译器进行重载，编译器将检查基函数是否存在这样的虚函数，否则将无法通过编译：

```
struct Base {  
    virtual void foo(int);  
};  
struct SubClass: Base {  
    virtual void foo(int) override; // 合法  
    virtual void foo(float) override; // 非法，父类没有此虚函数  
};
```

final

`final` 则是为了防止类被继续继承以及终止虚函数继续重载引入的。

```
struct Base {  
    virtual void foo() final;  
};  
struct SubClass1 final: Base {  
}; // 合法  
  
struct SubClass2 : SubClass1 {  
}; // 非法，SubClass1 已 final  
  
struct SubClass3: Base {  
    void foo(); // 非法，foo 已 final  
};
```

显式禁用默认函数

在传统 C++ 中，如果程序员没有提供，编译器会默认为对象生成默认构造函数、复制构造、赋值算符以及析构函数。另外，C++ 也为所有类定义了诸如 `new` `delete` 这样的运算符。当程序员有需要时，可以重载这部分函数。

这就引发了一些需求：无法精确控制默认函数的生成行为。例如禁止类的拷贝时，必须将赋值构造函数与赋值算符声明为 `private`。尝试使用这些未定义的函数将导致编译或链接错误，则是一种非常不优雅的方式。

并且，编译器产生的默认构造函数与用户定义的构造函数无法同时存在。若用户定义了任何构造函数，编译器将不再生成默认构造函数，但有时候我们却希望同时拥有这两种构造函数，这就造成了尴尬。

C++11 提供了上述需求的解决方案，允许显式的声明采用或拒绝编译器自带的函数。例如：

```
class Magic {  
public:  
    Magic() = default; // 显式声明使用编译器生成的构造  
    Magic& operator=(const Magic&) = delete; // 显式声明拒绝编译器  
    生成构造  
    Magic(int magic_number);  
}
```

八、强类型枚举

在传统 C++ 中，枚举类型并非类型安全，枚举类型会被视作整数，则会让两种完全不同的枚举类型可以进行直接比较（虽然编译器给出了检查，但并非所有），甚至枚举类型的枚举值名字不能相同，这不是我们希望看到的结果。

C++11 引入了枚举类（enumeration class），并使用 `enum class` 的语法进行声明：

```
enum class new_enum : unsigned int {  
    value1,  
    value2,  
    value3 = 100,  
    value4 = 100  
};
```

这样定义的枚举实现了类型安全，首先他不能够被隐式的转换为整数，同时也不能够将其与整数数字进行比较，更不可能对不同的枚举类型的枚举值进行比较。但相同枚举值之间如果指定的值相同，那么可以进行比较：

```
if (new_enum::value3 == new_enum::value4) {  
    // 会输出  
    std::cout << "new_enum::value3 == new_enum::value4" << std::endl;  
}
```

在这个语法中，枚举类型后面使用了冒号及类型关键字来指定枚举中枚举值的类型，这使得我们能够为枚举赋值（未指定时将默认使用 `int`）。

而我们希望获得枚举值的值时，将必须显式的进行类型转换，不过我们可以通过重载 `<<` 这个算符来进行输出，可以收藏下面这个代码段：

```
#include <iostream>  
template<typename T>  
std::ostream& operator<< (typename std::enable_if<std::is_enum<T>::value, std::ostream>::type& stream, const T& e)  
{  
    return stream << static_cast<typename std::underlying_type<T>::type>(e);  
}
```

这时，下面的代码将能够被编译：

```
std::cout << new_enum::value3 << std::endl
```

总结

本节介绍了 C++11/14 中对语言可用性的增强，其中笔者认为最为重要的几个特性是几乎所有人都需要了解并熟练使用的：

1. `auto` 类型推导
2. 范围 `for` 迭代
3. 初始化列表

4. 变参模板

进一步阅读的参考资料

1. 深入理解 C++11: C++11 新特性解析与应用. Michael Wong, IBM XL 编译器中国开发团队著
2. 深入应用 C++11: 代码优化与工程级应用. 祁宇著

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook
该文件修改时间 : 2017-11-16 09:30:52

第三章 语言运行期的强化

一、本节内容

本节内容包括：

- 语言运行期的强化
 - lambda 表达式
 - lambda 表达式基础
 - 值捕获
 - 引用捕获
 - 隐式捕获
 - 表达式捕获
 - 泛型 lambda
 - 函数对象包装器
 - `std::function`
 - `std::bind/std::placeholder`
 - 右值引用
 - 左值、右值的纯右值、将亡值、右值
 - 右值引用和左值引用
 - 移动语义
 - 完美转发

二、Lambda 表达式

Lambda 表达式是 C++11 中最重要的新特性之一，而 Lambda 表达式，实际上就是提供了一个类似匿名函数的特性，而匿名函数则是在需要一个函数，但是又不想费力去命名一个函数的情况下去使用的。这样的场景其实有很多很多，所以匿名函数几乎是现代编程语言的标配。

Lambda 表达式基础

Lambda 表达式的基本语法如下：

```
[捕获列表](参数列表) mutable(可选) 异常属性 -> 返回类型 {  
    // 函数体  
}
```

上面的语法规则除了 [捕获列表] 内的东西外，其他部分都很好理解，只是一般函数的函数名被略去，返回值使用了一个 `->` 的形式进行（我们在上一节前面的尾返回类型已经提到过这种写法了）。

所谓捕获列表，其实可以理解为参数的一种类型，`lambda` 表达式内部函数体在默认情况下是不能够使用函数体外部的变量的，这时候捕获列表可以起到传递外部数据的作用。根据传递的行为，捕获列表也分为以下几种：

1. 值捕获

与参数传值类似，值捕获的前期是变量可以拷贝，不同之处则在于，被捕获的变量在 `lambda` 表达式被创建时拷贝，而非调用时才拷贝：

```
void learn_lambda_func_1() {  
    int value_1 = 1;  
    auto copy_value_1 = [value_1] {  
        return value_1;  
    };  
    value_1 = 100;  
    auto stored_value_1 = copy_value_1();  
    // 这时, stored_value_1 == 1, 而 value_1 == 100.  
    // 因为 copy_value_1 在创建时就保存了一份 value_1 的拷贝  
}
```

2. 引用捕获

与引用传参类似，引用捕获保存的是引用，值会发生变化。

```
void learn_lambda_func_2() {  
    int value_2 = 1;  
    auto copy_value_2 = [&value_2] {  
        return value_2;  
    };  
    value_2 = 100;  
    auto stored_value_2 = copy_value_2();  
    // 这时, stored_value_2 == 100, value_1 == 100.  
    // 因为 copy_value_2 保存的是引用  
}
```

3. 隐式捕获

手动书写捕获列表有时候是非常复杂的，这种机械性的工作可以交给编译器来处理，这时候可以在捕获列表中写一个 `&` 或 `=` 向编译器声明采用 引用捕获或者 值捕获。

总结一下，捕获提供了lambda 表达式对外部值进行使用的功能，捕获列表的最常用的四种形式可以是：

- `[]` 空捕获列表
- `[name1, name2, ...]` 捕获一系列变量
- `[&]` 引用捕获, 让编译器自行推导捕获列表
- `[=]` 值捕获, 让编译器执行推导应用列表

4. 表达式捕获(C++14)

这部分内容需要了解后面马上要提到的右值引用以及智能指针

上面提到的值捕获、引用捕获都是已经在外层作用域声明的变量，因此这些捕获方式捕获的均为左值，而不能捕获右值。

C++14 给与了我们方便，允许捕获的成员用任意的表达式进行初始化，这就允许了右值的捕获，被声明的捕获变量类型会根据表达式进行判断，判断方式与使用

`auto` 本质上是相同的：

```
#include <iostream>
#include <utility>

int main() {
    auto important = std::make_unique<int>(1);
    auto add = [v1 = 1, v2 = std::move(important)](int x, int y)
    -> int {
        return x+y+v1+(*v2);
    };
    std::cout << add(3,4) << std::endl;
    return 0;
}
```

在上面的代码中，`important` 是一个独占指针，是不能够被捕获到的，这时候我们需要将其转移为右值，在表达式中初始化。

泛型 Lambda

上一节中我们提到了 `auto` 关键字不能够用在参数表里，这是因为这样的写法会与模板的功能产生冲突。但是 Lambda 表达式并不是普通函数，所以 Lambda 表达式并不能够模板化。这就为我们造成了一定程度上的麻烦：参数表不能够泛化，必须明确参数表类型。

幸运的是，这种麻烦只存在于 C++11 中，从 C++14 开始，Lambda 函数的形式参数可以使用 `auto` 关键字来产生意义上的泛型：

```
auto add = [](auto x, auto y) {
    return x+y;
};

add(1, 2);
add(1.1, 2.2);
```

二、函数对象包装器

这部分内容虽然属于标准库的一部分，但是从本质上来看，它却增强了 C++ 语言运行时的能力，这部分内容也相当重要，所以放到这里来进行介绍。

std::function

Lambda 表达式的本质是一个函数对象，当 Lambda 表达式的捕获列表为空时，Lambda 表达式还能够作为一个函数指针进行传递，例如：

```
#include <iostream>

using foo = void(int); // 定义函数指针, using 的使用见上一节中的别名语法
void functional(foo f) {
    f(1);
}

int main() {
    auto f = [](int value) {
        std::cout << value << std::endl;
    };
    functional(f); // 函数指针调用
    f(1);          // lambda 表达式调用
    return 0;
}
```

上面的代码给出了两种不同的调用形式，一种是将 Lambda 作为函数指针传递进行调用，而另一种则是直接调用 Lambda 表达式，在 C++11 中，统一了这些概念，将能够被调用的对象的类型，统一称之为可调用类型。而这种类型，便是通过 `std::function` 引入的。

C++11 `std::function` 是一种通用、多态的函数封装，它的实例可以对任何可以调用的目标实体进行存储、复制和调用操作，它也是对 C++ 中现有的可调用实体的一种类型安全的包裹（相对来说，函数指针的调用不是类型安全的），换句话说，就是函数的容器。当我们有了函数的容器之后便能够更加方便的将函数、函数指针作为对象进行处理。例如：

```
#include <functional>
#include <iostream>

int foo(int para) {
    return para;
}

int main() {
    // std::function 包装了一个返回值为 int, 参数为 int 的函数
    std::function<int(int)> func = foo;

    int important = 10;
    std::function<int(int)> func2 = [&](int value) -> int {
        return 1+value+important;
    };
    std::cout << func(10) << std::endl;
    std::cout << func2(10) << std::endl;
}
```

std::bind/std::placeholder

而 `std::bind` 则是用来绑定函数调用的参数的，它解决的需求是我们有时候可能并不一定能够一次性获得调用某个函数的全部参数，通过这个函数，我们可以将部分调用参数提前绑定到函数身上成为一个新的对象，然后在参数齐全后，完成调用。例如：

```
int foo(int a, int b, int c) {
    ;
}

int main() {
    // 将参数1,2绑定到函数 foo 上，但是使用 std::placeholders::_1 来
    // 对第一个参数进行占位
    auto bindFoo = std::bind(foo, std::placeholders::_1, 1, 2);
    // 这时调用 bindFoo 时，只需要提供第一个参数即可
    bindFoo(1);
}
```

提示：注意 `auto` 关键字的妙用。有时候我们可能不太熟悉一个函数的返回值类型，但是我们却可以通过 `auto` 的使用来规避这一问题的出现。

三、右值引用

右值引用是 C++11 引入的与 Lambda 表达式齐名的重要特性之一。它的引入解决了 C++ 中大量的历史遗留问题，消除了诸如 `std::vector`、`std::string` 之类的额外开销，也才使得函数对象容器 `std::function` 成为了可能。

左值、右值的纯右值、将亡值、右值

要弄明白右值引用到底是怎么一回事，必须要对左值和右值做一个明确的理解。

左值(**lvalue, left value**)，顾名思义就是赋值符号左边的值。准确来说，左值是表达式（不一定是赋值表达式）后依然存在的持久对象。

右值(**rvalue, right value**)，右边的值，是指表达式结束后就不再存在的临时对象。

而 C++11 中为了引入强大的右值引用，将右值的概念进行了进一步的划分，分为：纯右值、将亡值。

纯右值(**prvalue, pure rvalue**)，纯粹的右值，要么是纯粹的字面量，例如 `10`，`true`；要么是求值结果相当于字面量或匿名临时对象，例如 `1+2`。非引用返回的临时变量、运算表达式产生的临时变量、原始字面量、Lambda 表达式都属于纯右值。

将亡值(**xvalue, expiring value**)，是 C++11 为了引入右值引用而提出的概念（因此在传统 C++ 中，纯右值和右值是统一个概念），也就是即将被销毁、却能够被移动的值。

将亡值可能稍有些难以理解，我们来看这样的代码：

```
std::vector<int> foo() {  
    std::vector<int> temp = {1, 2, 3, 4};  
    return temp;  
}  
  
std::vector<int> v = foo();
```


在这样的代码中，函数 `foo` 的返回值 `temp` 在内部创建然后被赋值给 `v`，然而 `v` 获得这个对象时，会将整个 `temp` 拷贝一份，然后把 `temp` 销毁，如果这个 `temp` 非常大，这将造成大量额外的开销（这也就是传统 C++ 一直被诟病的问题）。在最后一行中，`v` 是左值、`foo()` 返回的值就是右值（也是纯右值）。

但是，`v` 可以被别的变量捕获到，而 `foo()` 产生的那个返回值作为一个临时值，一旦被 `v` 复制后，将立即被销毁，无法获取、也不能修改。

将亡值就定义了这样一种行为：临时的值能够被识别、同时又能够被移动。

右值引用和左值引用

需要拿到一个将亡值，就需要用到右值引用的申明：`T &&`，其中 `T` 是类型。右值引用的声明让这个临时值的生命周期得以延长、只要变量还活着，那么将亡值将继续存活。

C++11 提供了 `std::move` 这个方法将左值参数无条件的转换为右值，有了它我们就能够方便的获得一个右值临时对象，例如：

```
#include <iostream>
#include <string>

void reference(std::string& str) {
    std::cout << "左值" << std::endl;
}

void reference(std::string&& str) {
    std::cout << "右值" << std::endl;
}

int main()
{
    std::string lv1 = "string,";           // lv1 是一个左值
    // std::string&& r1 = s1;              // 非法，右值引用不能引用左值

    std::string&& rv1 = std::move(lv1);    // 合法，std::move可以将左
    值转移为右值
    std::cout << rv1 << std::endl;        // string,

    const std::string& lv2 = lv1 + lv1;    // 合法，常量左值引用能够延
    长临时变量的申明周期
    // lv2 += "Test";                    // 非法，引用的右值无法被修改

    std::cout << lv2 << std::endl;        // string,string

    std::string&& rv2 = lv1 + lv2;          // 合法，右值引用延长临时对
    象声明周期
    rv2 += "Test";                        // 合法，非常量引用能够修改
    临时变量
    std::cout << rv2 << std::endl;        // string,string,string,

    reference(rv2);                        // 输出左值

    return 0;
}
```

注意： `rv2` 虽然引用了一个右值，但由于它是一个引用，所以 `rv2` 依然是一个左值。

移动语义

传统 C++ 通过拷贝构造函数和赋值操作符为类对象设计了拷贝/复制的概念，但为了实现对资源的移动操作，调用者必须使用先复制、再析构的方式，否则就需要自己实现移动对象的接口。试想，搬家的时候是把家里的东西直接搬到新家去，而不是将所有东西复制一份（重买）再放到新家、再把原来的东西全部扔掉（销毁），这是非常反人类的一件事情。

传统的 C++ 没有区分『移动』和『拷贝』的概念，造成了大量的数据移动，浪费时间和空间。右值引用的出现恰好就解决了这两个概念的混淆问题，例如：

```
#include <iostream>
class A {
public:
    int *pointer;
    A():pointer(new int(1)) { std::cout << "构造" << pointer << std::endl; }
    A(A& a):pointer(new int(*a.pointer)) { std::cout << "拷贝" << pointer << std::endl; } // 无意义的对象拷贝
    A(A&& a):pointer(a.pointer) { a.pointer = nullptr; std::cout << "移动" << pointer << std::endl; }
    ~A(){ std::cout << "析构" << pointer << std::endl; delete pointer; }
};
// 防止编译器优化
A return_rvalue(bool test) {
    A a,b;
    if(test) return a;
    else return b;
}
int main() {
    A obj = return_rvalue(false);
    std::cout << "obj:" << std::endl;
    std::cout << obj.pointer << std::endl;
    std::cout << *obj.pointer << std::endl;

    return 0;
}
```

在上面的代码中：

1. 首先会在 `return_rvalue` 内部构造两个 `A` 对象，于是获得两个构造函数的输出；
2. 函数返回后，产生一个将亡值，被 `A` 的移动构造（`A(A&&)`）引用，从而延长生命周期，并将这个右值中的指针拿到，保存到了 `obj` 中，而将亡值的指针被设置为 `nullptr`，防止了这块内存区域被销毁。

从而避免了无意义的拷贝构造，加强了性能。再来看看涉及标准库的例子：

```
#include <iostream> // std::cout
#include <utility>    // std::move
#include <vector>     // std::vector
#include <string>     // std::string

int main() {

    std::string str = "Hello world.";
    std::vector<std::string> v;

    // 将使用 push_back(const T&), 即产生拷贝行为
    v.push_back(str);
    // 将输出 "str: Hello world."
    std::cout << "str: " << str << std::endl;

    // 将使用 push_back(const T&&), 不会出现拷贝行为
    // 而整个字符串会被移动到 vector 中，所以有时候 std::move 会用来减少
    // 拷贝出现的开销
    // 这步操作后，str 中的值会变为空
    v.push_back(std::move(str));
    // 将输出 "str: "
    std::cout << "str: " << str << std::endl;

    return 0;
}
```

完美转发

前面我们提到了，一个声明的右值引用其实是一个左值。这就为我们进行参数转发（传递）造成了问题：

```
void reference(int& v) {
    std::cout << "左值" << std::endl;
}
void reference(int&& v) {
    std::cout << "右值" << std::endl;
}
template <typename T>
void pass(T&& v) {
    std::cout << "普通传参:";
    reference(v);    // 始终调用 reference(int& )
}
int main() {
    std::cout << "传递右值:" << std::endl;
    pass(1);          // 1是右值，但输出左值

    std::cout << "传递左值:" << std::endl;
    int v = 1;
    pass(v);          // r 是左引用，输出左值

    return 0;
}
```

对于 `pass(1)` 来说，虽然传递的是右值，但由于 `v` 是一个引用，所以同时也是左值。因此 `reference(v)` 会调用 `reference(int&)`，输出『左值』。而对于 `pass(v)` 而言，`v` 是一个左值，为什么会成功传递给 `pass(T&&)` 呢？

这是基于引用坍缩规则的：在传统 C++ 中，我们不能够对一个引用类型继续进行引用，但 C++ 由于右值引用的出现而放宽了这一做法，从而产生了引用坍缩规则，允许我们对引用进行引用，既能左引用，又能右引用。但是却遵循如下规则：

函数形参类型	实参参数类型	推导后函数形参类型
T&	左引用	T&
T&	右引用	T&
T&&	左引用	T&
T&&	右引用	T&&

因此，模板函数中使用 `T&&` 不一定能进行右值引用，当传入左值时，此函数的引用将被推导为左值。更准确的讲，无论模板参数是什么类型的引用，当且仅当实参类型为右引用时，模板参数才能被推导为右引用类型。这才使得 `v` 作为左值的成功传递。

完美转发就是基于上述规律产生的。所谓完美转发，就是为了让我们在传递参数的时候，保持原来的参数类型（左引用保持左引用，右引用保持右引用）。为了解决这个问题，我们应该使用 `std::forward` 来进行参数的转发（传递）：

```
#include <iostream>
#include <utility>
void reference(int& v) {
    std::cout << "左值引用" << std::endl;
}
void reference(int&& v) {
    std::cout << "右值引用" << std::endl;
}
template <typename T>
void pass(T&& v) {
    std::cout << "普通传参:";
    reference(v);
    std::cout << "std::move 传参:";
    reference(std::move(v));
    std::cout << "std::forward 传参:";
    reference(std::forward<T>(v));
}
int main() {
    std::cout << "传递右值:" << std::endl;
    pass(1);

    std::cout << "传递左值:" << std::endl;
    int v = 1;
    pass(v);

    return 0;
}
```

输出结果为：

传递右值:

普通传参:左值引用

`std::move` 传参:右值引用

`std::forward` 传参:右值引用

传递左值:

普通传参:左值引用

`std::move` 传参:右值引用

`std::forward` 传参:左值引用

无论传递参数为左值还是右值，普通传参都会将参数作为左值进行转发，所以 `std::move` 总会接受到一个左值，从而转发调用了 `reference(int&&)` 输出右值引用。

唯独 `std::forward` 即没有造成任何多余的拷贝，同时完美转发(传递)了函数的实参给了内部调用的其他函数。

`std::forward` 和 `std::move` 一样，没有做任何事情，`std::move` 单纯的将左值转化为右值，`std::forward` 也只是单纯的将参数做了一个类型的转换，从是线上来看，`std::forward<T>(v)` 和 `static_cast<T&&>(v)` 是完全一样的。

总结

本节介绍了 C++11/14 中对语言可用性的增强，其中笔者认为本节中提到的所有特性都是值得掌握的：

1. Lambda 表达式
2. 函数对象容器 `std::function`
3. 右值引用

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook该文件修改时间：2017-11-16 09:30:52

第四章 对标准库的扩充：新增容器

一、本节内容

本节内容包括：

- 对标准库的扩充: 新增容器
 - `std::array`
 - `std::forward_list`
 - `std::unordered_set`
 - `std::unordered_map`
 - `std::tuple`
 - 基本操作
 - 运行期索引
 - 合并与迭代

二、`std::array` 和 `std::forward_list`

`std::array`

看到这个容器的时候肯定会出现这样的问题：

1. 为什么要引入 `std::array` 而不是直接使用 `std::vector` ？
2. 已经有了传统数组，为什么要用 `std::array` ？

先回答第一个问题，`std::vector` 太强大了，以至于我们没有必要为了去敲碎一个鸡蛋而用一个钉锤。使用 `std::array` 保存在栈内存中，相比堆内存中的 `std::vector`，我们就能够灵活的访问这里的元素，从而获得更高的性能；同时正式由于其堆内存存储的特性，有些时候我们还需要自己负责释放这些资源。

而第二个问题就更加简单，使用 `std::array` 能够让代码变得更加现代，且封装了一些操作函数，同时还能够友好的使用标准库中的容器算法等等，比如 `std::sort`。

`std::array` 会在编译时创建一个固定大小的数组，`std::array` 不能够被隐式的转换成指针，使用 `std::array` 很简单，只需指定其类型和大小即可：

```
std::array<int, 4> arr= {1,2,3,4};

int len = 4;
std::array<int, len> arr = {1,2,3,4}; // 非法, 数组大小参数必须是常
量表达式
```

当我们开始用上了 `std::array` 时, 难免会遇到要将其兼容 C 风格的接口, 这里有三种做法:

```
void foo(int *p, int len) {
    return;
}

std::array<int, 4> arr = {1,2,3,4};

// C 风格接口传参
// foo(arr, arr.size());           // 非法, 无法隐式转换
foo(&arr[0], arr.size());
foo(arr.data(), arr.size());

// 使用 `std::sort`
std::sort(arr.begin(), arr.end());
```

std::forward_list

`std::forward_list` 是一个列表容器, 使用方法和 `std::list` 基本类似, 因此我们就不花费篇幅进行介绍了。

需要知道的是, 和 `std::list` 的双向链表的实现不同, `std::forward_list` 使用单向链表进行实现, 提供了 $O(1)$ 复杂度的元素插入, 不支持快速随机访问 (这也是链表的特点), 也是标准库容器中唯一一个不提供 `size()` 方法的容器。当不需要双向迭代时, 具有比 `std::list` 更高的空间利用率。

三、无序容器

我们已经熟知了传统 C++ 中的有序容器 `std::map` / `std::set`，这些元素内部通过红黑树进行实现，插入和搜索的平均复杂度均为 $O(\log(\text{size}))$ 。在插入元素时候，会根据 `<` 操作符比较元素大小并判断元素是否相同，并选择合适的位置插入到容器中。当对这个容器中的元素进行遍历时，输出结果会按照 `<` 操作符的顺序来逐个遍历。

而无序容器中的元素是不进行排序的，内部通过 Hash 表实现，插入和搜索元素的平均复杂度为 $O(\text{constant})$ ，在不关心容器内部元素顺序时，能够获得显著的性能提升。

C++11 引入了两组无序容

器：`std::unordered_map` / `std::unordered_multimap` 和
`std::unordered_set` / `std::unordered_multiset`。

它们的用法和原有的

`std::map` / `std::multimap` / `std::set` / `std::multiset` 基本类似，由于这些容器我们已经很熟悉了，便不一一举例，我们直接来比较一下 `std::map` 和 `std::multimap`：

```
#include <iostream>
#include <string>
#include <unordered_map>
#include <map>

int main() {
    // 两组结构按同样的顺序初始化
    std::unordered_map<int, std::string> u = {
        {1, "1"},
        {3, "3"},
        {2, "2"}
    };
    std::map<int, std::string> v = {
        {1, "1"},
        {3, "3"},
        {2, "2"}
    };

    // 分别对两组结构进行遍历
    std::cout << "std::unordered_map" << std::endl;
    for( const auto & n : u)
        std::cout << "Key:[" << n.first << "]" Value:[" << n.second << "]\n";

    std::cout << std::endl;
    std::cout << "std::map" << std::endl;
    for( const auto & n : v)
        std::cout << "Key:[" << n.first << "]" Value:[" << n.second << "]\n";
}
```

最终的输出结果为：

```
std::unordered_map
Key:[2] Value:[2]
Key:[3] Value:[3]
Key:[1] Value:[1]

std::map
Key:[1] Value:[1]
Key:[2] Value:[2]
Key:[3] Value:[3]
```

四、元组 **std::tuple**

了解过 Python 的程序员应该知道元组的概念，纵观传统 C++ 中的容器，除了 `std::pair` 外，似乎没有现成的结构能够用来存放不同类型的数据（通常我们会自己定义结构）。但 `std::pair` 的缺陷是显而易见的，只能保存两个元素。

元组基本操作

关于元组的使用有三个核心的函数：

1. `std::make_tuple` : 构造元组
2. `std::get` : 获得元组某个位置的值
3. `std::tie` : 元组拆包

```
#include <tuple>
#include <iostream>

auto get_student(int id)
{
    // 返回类型被推断为 std::tuple<double, char, std::string>

    if (id == 0)
        return std::make_tuple(3.8, 'A', "张三");
    if (id == 1)
        return std::make_tuple(2.9, 'C', "李四");
    if (id == 2)
        return std::make_tuple(1.7, 'D', "王五");
    return std::make_tuple(0.0, 'D', "null");
    // 如果只写 0 会出现推断错误, 编译失败
}

int main()
{
    auto student = get_student(0);
    std::cout << "ID: 0, "
    << "GPA: " << std::get<0>(student) << ", "
    << "成绩: " << std::get<1>(student) << ", "
    << "姓名: " << std::get<2>(student) << '\n';

    double gpa;
    char grade;
    std::string name;

    // 元组进行拆包
    std::tie(gpa, grade, name) = get_student(1);
    std::cout << "ID: 1, "
    << "GPA: " << gpa << ", "
    << "成绩: " << grade << ", "
    << "姓名: " << name << '\n';
}
```

`std::get` 除了使用常量获取元组对象外, C++14 增加了使用类型来获取元组中的对象:

```
std::tuple<std::string, double, double, int> t("123", 4.5, 6.7, 8);
std::cout << std::get<std::string>(t) << std::endl;
std::cout << std::get<double>(t) << std::endl;    // 非法, 引发编译
期错误
std::cout << std::get<3>(t) << std::endl;
```

运行期索引

如果你仔细思考一下可能就会发现上面代码的问题，`std::get<>` 依赖一个编译期的常量，所以下面的方式是不合法的：

```
int index = 1;
std::get<index>(t);
```

那么要怎么处理？答案是，标准库做不到。这里介绍一个使用 `boost::variant` 配合变长模板参数的黑魔法：

```
#include <boost/variant.hpp>
template <size_t n, typename... T>
boost::variant<T...> _tuple_index(size_t i, const std::tuple<T...>& tpl) {
    if (i == n)
        return std::get<n>(tpl);
    else if (n == sizeof...(T) - 1)
        throw std::out_of_range("越界.");
    else
        return _tuple_index<(n < sizeof...(T)-1 ? n+1 : 0)>(i, tpl);
}
template <typename... T>
boost::variant<T...> tuple_index(size_t i, const std::tuple<T...>& tpl) {
    return _tuple_index<0>(i, tpl);
}
```

这样我们就能：

```
int i = 1;
std::cout << tuple_index(i, t) << std::endl;
```

元组合并与遍历

还有一个常见的需求就是合并两个元组，这可以通过 `std::tuple_cat` 来实现：

```
auto new_tuple = std::tuple_cat(get_student(1), std::move(t));
```

马上就能够发现，应该如何快速遍历一个元组？但是我们刚才介绍了如何在运行期通过非常数索引一个 `tuple` 那么遍历就变得简单了，首先我们需要知道一个元组的长度，可以：

```
template <typename T>
auto tuple_len(T &tpl) {
    return std::tuple_size<T>::value;
}
```

这样就能够对元组进行迭代了：

```
// 迭代
for(int i = 0; i != tuple_len(new_tuple); ++i)
    // 运行期索引
    std::cout << tuple_index(i, new_tuple) << std::endl;
```

总结

本节简单介绍了 C++11/14 中新增的容器，它们的用法和传统 C++ 中已有的容器类似，相对简单，可以根据实际场景丰富的选择需要使用的容器，从而获得更好的性能。

`std::tuple` 虽然有效，但是标准库提供的功能有限，没办法满足运行期索引和迭代的需求，好在我们还有其他的方法可以自行实现。

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by
Gitbook该文件修改时间 : 2017-11-16 09:30:52

第五章 对标准库的扩充：引用计数与智能指针

一、本节内容

本节内容包括：

- 对标准库的扩充: 智能指针和引用计数
 - RAII 与引用计数
 - `std::shared_ptr`
 - `std::weak_ptr`
 - `std::unique_ptr`

二、RAII 与引用计数

了解 `Objective-C / Swift` 的程序员应该知道引用计数的概念。引用计数这种计数是为了防止内存泄露而产生的。基本想法是对于动态分配的对象，进行引用计数，每当增加一次对同一个对象的引用，那么引用对象的引用计数就会增加一次，每删除一次引用，引用计数就会减一，当一个对象的引用计数减为零时，就自动删除指向的堆内存。

在传统 `C++` 中，『记得』手动释放资源，总不是最佳实践。因为我们很有可能就忘记了去释放资源而导致泄露。所以通常的做法是对于一个对象而言，我们在构造函数的时候申请空间，而在析构函数（在离开作用域时调用）的时候释放空间，也就是我们常说的 `RAII` 资源获取即初始化技术。

凡事都有例外，我们总会有需要将对象在自由存储上分配的需求，在传统 `C++` 里我们只好使用 `new` 和 `delete` 去『记得』对资源进行释放。而 `C++11` 引入了智能指针的概念，使用了引用计数的想法，让程序员不再需要关心手动释放内存。这些智能指针就包括

`std::shared_ptr` / `std::unique_ptr` / `std::weak_ptr`，使用它们需要包含头文件 `<memory>`。

注意：引用计数不是垃圾回收，引用技术能够尽快收回不再被使用的对象，同时在回收的过程中也不会造成长时间的等待，更能够清晰明确的表明资源的生命周期。

三、std::shared_ptr

`std::shared_ptr` 是一种智能指针，它能够记录多少个 `shared_ptr` 共同指向一个对象，从而消除显示的调用 `delete`，当引用计数变为零的时候就会将对象自动删除。

但还不够，因为使用 `std::shared_ptr` 仍然需要使用 `new` 来调用，这使得代码出现了某种程度上的不对称。

`std::make_shared` 就能够用来消除显示的使用 `new`，所以 `std::make_shared` 会分配创建传入参数中的对象，并返回这个对象类型的 `std::shared_ptr` 指针。例如：

```
#include <iostream>
#include <memory>

void foo(std::shared_ptr<int> i)
{
    (*i)++;
}

int main()
{
    // auto pointer = new int(10); // 非法，不允许直接赋值
    // 构造了一个 std::shared_ptr
    auto pointer = std::make_shared<int>(10);
    foo(pointer);
    std::cout << *pointer << std::endl; // 11

    // 离开作用域前，shared_ptr 会被析构，从而释放内存
    return 0;
}
```

`std::shared_ptr` 可以通过 `get()` 方法来获取原始指针，通过 `reset()` 来减少一个引用计数，并通过 `get_count()` 来查看一个对象的引用计数。例如：

```

auto pointer = std::make_shared<int>(10);
auto pointer2 = pointer;    // 引用计数+1
auto pointer3 = pointer;    // 引用计数+1
int *p = pointer.get();     // 这样不会增加引用计数

std::cout << "pointer.use_count() = " << pointer.use_count() <<
std::endl;    // 3
std::cout << "pointer2.use_count() = " << pointer2.use_count() <
< std::endl;    // 3
std::cout << "pointer3.use_count() = " << pointer3.use_count() <
< std::endl;    // 3

pointer2.reset();
std::cout << "reset pointer2:" << std::endl;
std::cout << "pointer.use_count() = " << pointer.use_count() <<
std::endl;    // 2
std::cout << "pointer2.use_count() = " << pointer2.use_count() <
< std::endl;    // 0, pointer2 已 reset
std::cout << "pointer3.use_count() = " << pointer3.use_count() <
< std::endl;    // 2

pointer3.reset();
std::cout << "reset pointer3:" << std::endl;
std::cout << "pointer.use_count() = " << pointer.use_count() <<
std::endl;    // 1
std::cout << "pointer2.use_count() = " << pointer2.use_count() <
< std::endl;    // 0
std::cout << "pointer3.use_count() = " << pointer3.use_count() <
< std::endl;    // 0, pointer3 已 reset

```

四、std::unique_ptr

`std::unique_ptr` 是一种独占的智能指针，它禁止其他智能指针与其共享同一个对象，从而保证代码的安全：

```
std::unique_ptr<int> pointer = std::make_unique<int>(10);    // make_unique 从 C++14 引入
std::unique_ptr<int> pointer2 = pointer;    // 非法
```

make_unique 并不复杂，C++11 没有提供 std::make_unique，可以自行实现：

```
template<typename T, typename ...Args>
std::unique_ptr<T> make_unique( Args&& ...args ) {
    return std::unique_ptr<T>( new T( std::forward<Args>(args)... ) );
}
```

至于为什么没有提供，C++ 标准委员会主席 Herb Sutter 在他的[博客](#)中提到原因是因为『被他们忘记了』。

既然是独占，换句话说就是不可复制。但是，我们可以利用 std::move 将其转移给其他的 unique_ptr，例如：

```
#include <iostream>
#include <memory>

struct Foo {
    Foo()      { std::cout << "Foo::Foo" << std::endl; }
    ~Foo()     { std::cout << "Foo::~~Foo" << std::endl; }
    void foo() { std::cout << "Foo::foo" << std::endl; }
};

void f(const Foo &) {
    std::cout << "f(const Foo&)" << std::endl;
}

int main() {
    std::unique_ptr<Foo> p1(std::make_unique<Foo>());

    // p1 不空，输出
    if (p1) p1->foo();
}
```

```
std::unique_ptr<Foo> p2(std::move(p1));

// p2 不空, 输出
f(*p2);

// p2 不空, 输出
if(p2) p2->foo();

// p1 为空, 无输出
if(p1) p1->foo();

p1 = std::move(p2);

// p2 为空, 无输出
if(p2) p2->foo();
std::cout << "p2 被销毁" << std::endl;
}
// p1 不空, 输出
if (p1) p1->foo();

// Foo 的实例会在离开作用域时被销毁
}
```

五、std::weak_ptr

如果你仔细思考 `std::shared_ptr` 就会发现依然存在着资源无法释放的问题。看下面这个例子：

```

struct A;
struct B;

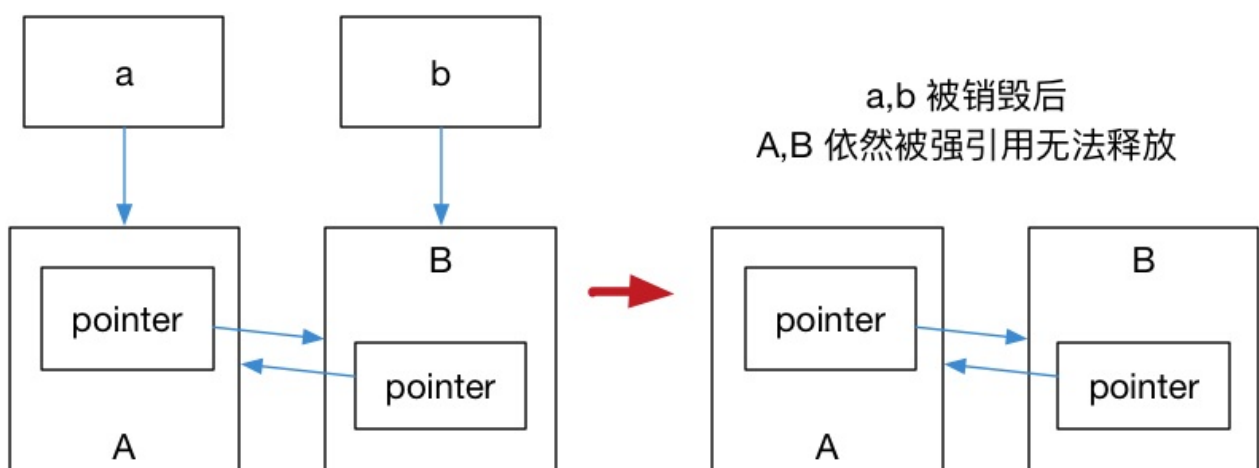
struct A {
    std::shared_ptr<B> pointer;
    ~A() {
        std::cout << "A 被销毁" << std::end;
    }
};

struct B {
    std::shared_ptr<A> pointer;
    ~B() {
        std::cout << "B 被销毁" << std::end;
    }
};

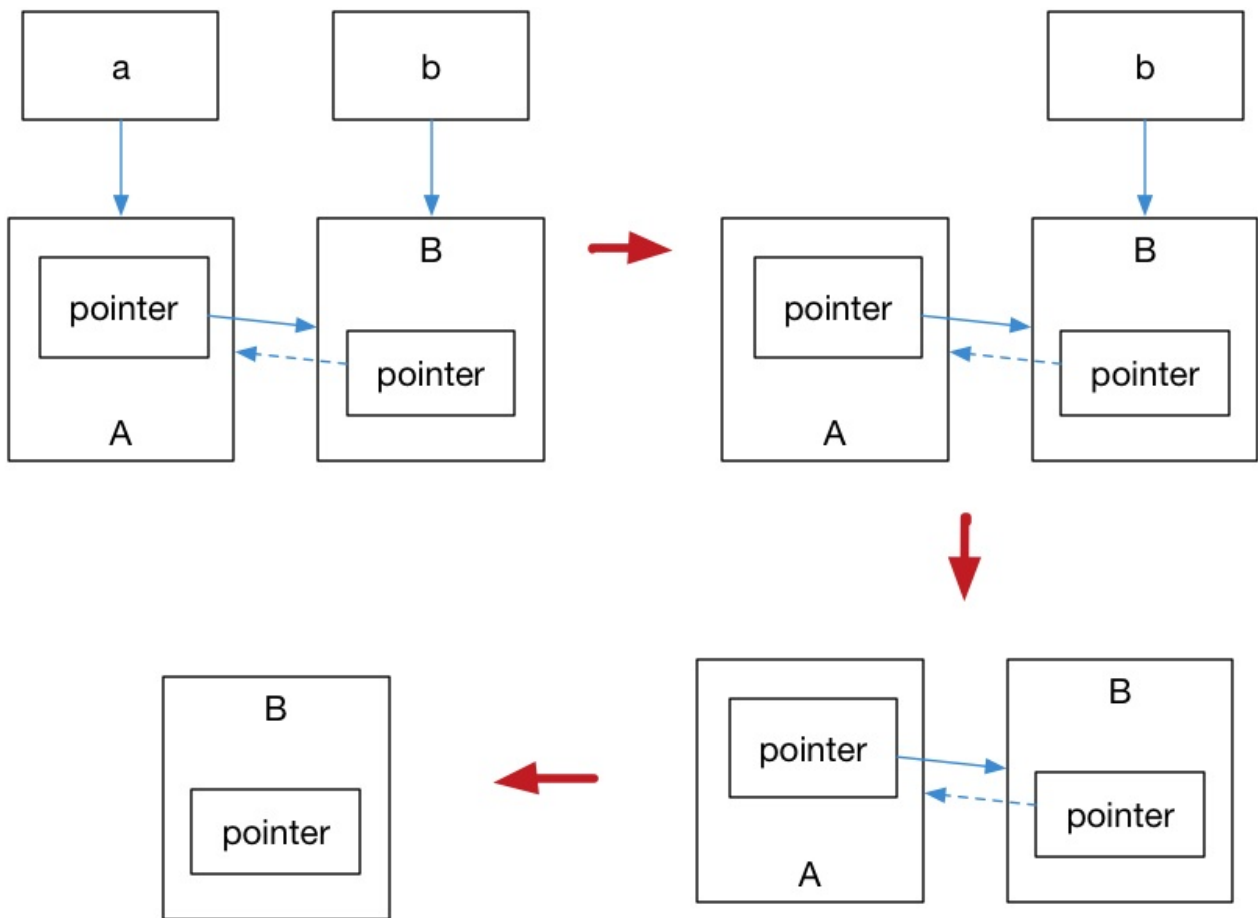
int main() {
    auto a = std::make_shared<A>();
    auto b = std::make_shared<B>();
    a.pointer = b;
    b.pointer = a;
}

```

运行结果是 A, B 都不会被销毁，这是因为 a,b 内部的 pointer 同时又引用了 a,b，这使得 a,b 的引用计数均变为了 2，而离开作用域时，a,b 智能指针被析构，却智能造成这块区域的引用计数减一，这样就导致了 a,b 对象指向的内存区域引用计数不为零，而外部已经没有办法找到这块区域了，也就造成了内存泄露，如图所示：



解决这个问题的办法就是使用弱引用指针 `std::weak_ptr`，`std::weak_ptr` 是一种弱引用（相比较而言 `std::shared_ptr` 就是一种强引用）。弱引用不会引起引用计数增加，当换用弱引用时候，最终的释放流程如下图所示：



在上图中，最后一步只剩下 `B`，而 `B` 并没有任何智能指针引用它，因此这块内存资源也会被释放。

`std::weak_ptr` 没有 `*` 运算符和 `->` 运算符，所以不能够对资源进行操作，它的唯一作用就是用于检查 `std::shared_ptr` 是否存在，`expired()` 方法在资源未被释放时，会返回 `true`，否则返回 `false`。

总结

智能指针这种技术并不新奇，在很多语言中都是一种常见的技术，`C++1x` 将这项技术引进，在一定程度上消除了 `new / delete` 的滥用，是一种更加成熟的编程范式。

进一步阅读的参考资料

1. [stackoverflow](#) 上关于『C++11为什么没有 `make_unique`』的讨论

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by
Gitbook该文件修改时间： 2017-11-16 09:30:52

第六章 正则表达式库

一、本节内容

本节内容包括：

- 对标准库的扩充: 正则表达式库
 - 正则表达式简介
 - 普通字符
 - 特殊字符
 - 限定符
 - `std::regex` 及其相关
 - `std::regex`
 - `std::regex_match`
 - `std::match_results`

二、正则表达式简介

正则表达式不是 C++ 语言的一部分，这里仅做简单的介绍。

正则表达式描述了一种字符串匹配的模式。一般使用正则表达式主要是实现下面三个需求：

1. 检查一个串是否包含某种形式的子串；
2. 将匹配的子串替换；
3. 从某个串中取出符合条件的子串。

正则表达式是由普通字符（例如 **a** 到 **z**）以及特殊字符组成的文字模式。模式描述在搜索文本时要匹配的一个或多个字符串。正则表达式作为一个模板，将某个字符模式与所搜索的字符串进行匹配。

普通字符

普通字符包括没有显式指定为元字符的所有可打印和不可打印字符。这包括所有大写和小写字母、所有数字、所有标点符号和一些其他符号。

特殊字符

特殊字符是正则表达式里有特殊含义的字符，也是正则表达式的核心匹配语法。参见下表：

特别字符	描述
\$	匹配输入字符串的结尾位置。
(,)	标记一个子表达式的开始和结束位置。子表达式可以获取供以后使用。
*	匹配前面的子表达式零次或多次。
+	匹配前面的子表达式一次或多次。
.	匹配除换行符 <code>\n</code> 之外的任何单字符。
[标记一个中括号表达式的开始。
?	匹配前面的子表达式零次或一次，或指明一个非贪婪限定符。
\	将下一个字符标记为或特殊字符、或原义字符、或向后引用、或八进制转义符。例如， <code>n</code> 匹配字符 <code>n</code> 。 <code>\n</code> 匹配换行符。序列 <code>\\</code> 匹配 <code>'\'</code> 字符，而 <code>\(</code> 则匹配 <code>'('</code> 字符。
^	匹配输入字符串的开始位置，除非在方括号表达式中使用，此时它表示不接受该字符集合。
{	标记限定符表达式的开始。
	指明两项之间的一个选择。

限定符

限定符用来指定正则表达式的一个给定的组件必须要出现多少次才能满足匹配。见下表：

字符	描述
<code>*</code>	匹配前面的子表达式零次或多次。例如， <code>foo*</code> 能匹配 <code>fo</code> 以及 <code>foooo</code> 。 <code>*</code> 等价于 <code>{0,}</code> 。
<code>+</code>	匹配前面的子表达式一次或多次。例如， <code>foo+</code> 能匹配 <code>foo</code> 以及 <code>foooo</code> ，但不能匹配 <code>fo</code> 。 <code>+</code> 等价于 <code>{1,}</code> 。
<code>?</code>	匹配前面的子表达式零次或一次。例如， <code>Your(s)?</code> 可以匹配 <code>Your</code> 或 <code>Yours</code> 中的 <code>Your</code> 。 <code>?</code> 等价于 <code>{0,1}</code> 。
<code>{n}</code>	<code>n</code> 是一个非负整数。匹配确定的 <code>n</code> 次。例如， <code>f{2}</code> 不能匹配 <code>for</code> 中的 <code>o</code> ，但是能匹配 <code>foo</code> 中的两个 <code>o</code> 。
<code>{n,}</code>	<code>n</code> 是一个非负整数。至少匹配 <code>n</code> 次。例如， <code>f{2,}</code> 不能匹配 <code>for</code> 中的 <code>o</code> ，但能匹配 <code>foooooo</code> 中的所有 <code>o</code> 。 <code>o{1,}</code> 等价于 <code>o+</code> 。 <code>o{0,}</code> 则等价于 <code>o*</code> 。
<code>{n,m}</code>	<code>m</code> 和 <code>n</code> 均为非负整数，其中 <code>n</code> 小于等于 <code>m</code> 。最少匹配 <code>n</code> 次且最多匹配 <code>m</code> 次。例如， <code>o{1,3}</code> 将匹配 <code>foooooo</code> 中的前三个 <code>o</code> 。 <code>o{0,1}</code> 等价于 <code>o?</code> 。注意，在逗号和两个数之间不能有空格。

有了这三张表，我们通常就能够读懂几乎所有的正则表达式了。

三、std::regex 及其相关

对字符串内容进行匹配的最常见手段就是使用正则表达式。可惜在传统 C++ 中正则表达式一直没有得到语言层面的支持，没有纳入标准库，而 C++ 作为一门高性能语言，在后台服务的开发中，对 URL 资源链接进行判断时，使用正则表达式也是工业界最为成熟的普遍做法。

一般的解决方案就是使用 `boost` 的正则表达式库。而 C++11 正式将正则表达式的处理方法纳入标准库的行列，从语言级上提供了标准的支持，不再依赖第三方。

C++11 提供的正则表达式库操作 `std::string` 对象，模式 `std::regex` (本质是 `std::basic_regex`) 进行初始化，通过 `std::regex_match` 进行匹配，从而产生 `std::smatch` (本质是 `std::match_results` 对象)。

我们通过一个简单的例子来简单介绍这个库的使用。考虑下面的正则表达式

- `[a-z]+\..txt`：在这个正则表达式中，`[a-z]` 表示匹配一个小写字母，`+` 可以使前面的表达式匹配多次，因此 `[a-z] +` 能够匹配一个小写字母组成的

字符串。在正则表达式中一个 `.` 表示匹配任意字符，而 `\.` 则表示匹配字符 `.`，最后的 `txt` 表示严格匹配 `txt` 则三个字母。因此这个正则表达式的所要匹配的内容就是由纯小写字母组成的文本文件。

`std::regex_match` 用于匹配字符串和正则表达式，有很多不同的重载形式。最简单的一个形式就是传入 `std::string` 以及一个 `std::regex` 进行匹配，当匹配成功时，会返回 `true`，否则返回 `false`。例如：

```
#include <iostream>
#include <string>
#include <regex>

int main() {
    std::string fnames[] = {"foo.txt", "bar.txt", "test", "a0.txt", "AAA.txt"};
    // 在 C++ 中 `.` 会被作为字符串内的转义符，为使 `.` 作为正则表达式
    // 传递进去生效，需要对 `.` 进行二次转义，从而有 `.`
    std::regex txt_regex("[a-z]+\\.txt");
    for (const auto &fname: fnames)
        std::cout << fname << ": " << std::regex_match(fname, txt_regex) << std::endl;
}
```

另一种常用的形式就是依次传入 `std::string` / `std::smatch` / `std::regex` 三个参数，其中 `std::smatch` 的本质其实是 `std::match_results`，在标准库中，`std::smatch` 被定义为了

`std::match_results<std::string::const_iterator>`，也就是一个子串迭代器类型的 `match_results`。使用 `std::smatch` 可以方便的对匹配的结果进行获取，例如：

```
std::regex base_regex("[a-z+)]\\.txt");
std::smatch base_match;
for(const auto &fname: fnames) {
    if (std::regex_match(fname, base_match, base_regex)) {
        // sub_match 的第一个元素匹配整个字符串
        // sub_match 的第二个元素匹配了第一个括号表达式
        if (base_match.size() == 2) {
            std::string base = base_match[1].str();
            std::cout << "sub-match[0]: " << base_match[0].str()
            << std::endl;
            std::cout << fname << " sub-match[1]: " << base << s
            td::endl;
        }
    }
}
```

以上两个代码段的输出结果为：

```
foo.txt: 1
bar.txt: 1
test: 0
a0.txt: 0
AAA.txt: 0
sub-match[0]: foo.txt
foo.txt sub-match[1]: foo
sub-match[0]: bar.txt
bar.txt sub-match[1]: bar
```

总结

本节简单介绍了正则表达式本身，然后根据使用正则表达式的主要需求，通过一个实际的例子介绍了正则表达式库的使用。

本节提到的内容足以让我们开发编写一个简单的 Web 框架中关于URL匹配的功能。

关于这方面的开发和细节，可以通过项目课：[C++ 开发 Web 服务框架](#) 进行进一步学习。

进一步阅读的参考资料

1. [知乎『如何评价 GCC 的 C++11 正则表达式？』中原库作者 Tim Shen 的回答](#)
2. [正则表达式库文档](#)
3. [C++ 开发 Web 服务框架](#)

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by
Gitbook该文件修改时间： 2017-11-16 09:30:52

第七章 语言级线程支持

一、本节内容

本节内容包括：

- 对标准库的扩充: 语言级线程支持
 - `std::thread`
 - `std::mutex/std::unique_lock`
 - `std::future/std::packaged_task`
 - `std::condition_variable`

二、`std::thread`

`std::thread` 用于创建一个执行的线程实例，所以它是一切并发编程的基础，使用时需要包含 `<thread>` 头文件，它提供了很多基本的线程操作，例如 `get_id()` 来获取所创建线程的线程 ID，例如使用 `join()` 来加入一个线程等等，例如：

```
#include <iostream>
#include <thread>
void foo() {
    std::cout << "hello world" << std::endl;
}
int main() {
    std::thread t(foo);
    t.join();
    return 0;
}
```

三、`std::mutex`, `std::unique_lock`

我们在操作系统的相关知识中已经了解过了有关并发技术的基本知识，`mutex` 就是其中的核心之一。C++11 引入了 `mutex` 相关的类，其所有相关的函数都放在 `<mutex>` 头文件中。

`std::mutex` 是 C++11 中最基本的 `mutex` 类，通过实例化 `std::mutex` 可以创建互斥量，而通过其成员函数 `lock()` 可以仅此能上锁，`unlock()` 可以进行解锁。但是在实际编写代码的过程中，最好不去直接调用成员函数，因为调用成员函数就需要在每个临界区的出口处调用 `unlock()`，当然，还包括异常。这时候 C++11 还为互斥量提供了一个 RAII 语法的模板类 `std::lock_guard`。RAII 在不失代码简洁性的同时，很好的保证了代码的异常安全性。

在 RAII 用法下，对于临界区的互斥量的创建只需要在作用域的开始部分，例如：

```
void some_operation(const std::string &message) {
    static std::mutex mutex;
    std::lock_guard<std::mutex> lock(mutex);

    // ...操作

    // 当离开这个作用域的时候，互斥锁会被析构，同时unlock互斥锁
    // 因此这个函数内部的可以认为是临界区
}
```

由于 C++ 保证了所有栈对象在声明周期结束时会被销毁，所以这样的代码也是异常安全的。无论 `some_operation()` 正常返回、还是在中途抛出异常，都会引发堆栈回退，也就自动调用了 `unlock()`。

而 `std::unique_lock` 则相对于 `std::lock_guard` 出现的，`std::unique_lock` 更加灵活，`std::unique_lock` 的对象会以独占所有权（没有其他的 `unique_lock` 对象同时拥有某个 `mutex` 对象的所有权）的方式管理 `mutex` 对象上的上锁和解锁的操作。所以在并发编程中，推荐使用 `std::unique_lock`。例如：

```
#include <iostream>
#include <thread>
#include <mutex>

std::mutex mtx;

void block_area() {
    std::unique_lock<std::mutex> lock(mtx);
    //...临界区
}

int main() {
    std::thread thd1(block_area);

    thd1.join();

    return 0;
}
```

std::future, std::packaged_task

`std::future` 则是提供了一个访问异步操作结果的途径，这句话很不好理解。为了理解这个特性，我们需要先理解一下在 C++11 之前的多线程行为。

试想，如果我们的主线程 A 希望新开辟一个线程 B 去执行某个我们预期的任务，并返回我一个结果。而这时候，线程 A 可能正在忙其他的事情，无暇顾及 B 的结果，所以我们会很自然的希望能够在某个特定的时间获得线程 B 的结果。

在 C++11 的 `std::future` 被引入之前，通常的做法是：创建一个线程 A，在线程 A 里启动任务 B，当准备完毕后发送一个事件，并将结果保存在全局变量中。而主函数线程 A 里正在做其他的事情，当需要结果的时候，调用一个线程等待函数来获得执行的结果。

而 C++11 提供的 `std::future` 简化了这个流程，可以用来获取异步任务的结果。自然地，我们很容易能够想象到把它作为一种简单的线程同步手段。

此外，`std::packaged_task` 可以用来封装任何可以调用的目标，从而用于实现异步的调用。例如：

```

#include <iostream>
#include <future>
#include <thread>

int main()
{
    // 将一个返回值为7的 lambda 表达式封装到 task 中
    // std::packaged_task 的模板参数为要封装函数的类型
    std::packaged_task<int()> task([](){return 7;});
    // 获得 task 的 future
    std::future<int> result = task.get_future();    // 在一个线程
    中执行 task
    std::thread(std::move(task)).detach();    std::cout << "Wait
    ing...";
    result.wait();
    // 输出执行结果
    std::cout << "Done!" << std::endl << "Result is " << result
    .get() << '\n';
}

```

在封装好要调用的目标后，可以使用 `get_future()` 来获得一个 `std::future` 对象，以便之后实施线程同步。

std::condition_variable

`std::condition_variable` 是为了解决死锁而生的。当互斥操作不够用而引入的。比如，线程可能需要等待某个条件为真才能继续执行，而一个忙等待循环中可能会导致所有其他线程都无法进入临界区使得条件为真时，就会发生死锁。所以，`condition_variable` 实例被创建出现主要就是用于唤醒等待线程从而避免死锁。`std::condition_variable` 的 `notify_one()` 用于唤醒一个线程；`notify_all()` 则是通知所有线程。下面是一个生产者和消费者模型的例子：

```

#include <condition_variable>
#include <mutex>
#include <thread>
#include <iostream>
#include <queue>

```

```
#include <chrono>

int main()
{
    // 生产者数量
    std::queue<int> produced_nums;
    // 互斥锁
    std::mutex m;
    // 条件变量
    std::condition_variable cond_var;
    // 结束标志
    bool done = false;
    // 通知标志
    bool notified = false;

    // 生产者线程
    std::thread producer([&]() {
        for (int i = 0; i < 5; ++i) {
            std::this_thread::sleep_for(std::chrono::seconds(1))
;

            // 创建互斥锁
            std::unique_lock<std::mutex> lock(m);
            std::cout << "producing " << i << '\n';
            produced_nums.push(i);
            notified = true;
            // 通知一个线程
            cond_var.notify_one();
        }
        done = true;
        cond_var.notify_one();
    });

    // 消费者线程
    std::thread consumer([&]() {
        std::unique_lock<std::mutex> lock(m);
        while (!done) {
            while (!notified) { // 循环避免虚假唤醒
                cond_var.wait(lock);
            }
            while (!produced_nums.empty()) {
```

```
        std::cout << "consuming " << produced_nums.front  
( ) << '\n';  
        produced_nums.pop();  
    }  
    notified = false;  
}  
});  
  
producer.join();  
consumer.join();  
}
```

总结

C++11 语言层提供了并发编程的相关支持，本节简单的介绍了

`std::thread` / `std::mutex` / `std::future` 这些并发编程中不可回避的重要工具。

本节提到的内容足以让我们使用不超过 100 行代码编写一个简单的线程池库

关于这方面的使用技巧，可以通过项目课：[100 行 C++ 代码实现线程池](#) 进行进一步巩固学习。

进一步阅读的参考资料

1. [C++ 并发编程\(中文版\)](#)
2. [线程支持库文档](#)
3. [100 行 C++ 代码实现线程池](#)

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook 该文件修改时间 : 2017-11-16 09:30:52

第八章 其他杂项

一、本节内容

本节内容包括：

- 对标准库的扩充: 其他杂项
 - 新类型
 - `long long int`
 - `noexcept` 的修饰和操作
 - 字面量
 - 原始字符串字面量
 - 自定义字面量

二、新类型

`long long int`

`long long int` 并不是 C++11 最先引入的，其实早在 C99，`long long int` 就已经被纳入 C 标准中，所以大部分的编译器早已支持。C++11 的工作则是正式把它纳入标准库，规定了一个 `long long int` 类型至少具备 64 位的比特数。

三、`noexcept` 的修饰和操作

C++ 相比于 C 的一大优势就在于 C++ 本身就定义了一套完整的异常处理机制。然而在 C++11 之前，几乎没有人去使用在函数名后书写异常声明表达式，从 C++11 开始，这套机制被弃用，所以我们不去讨论也不去介绍以前这套机制是如何工作如何使用，你更不应该主动去了解它。

C++11 将异常的声明简化为以下两种情况：

1. 函数可能抛出任何异常
2. 函数不能抛出任何异常

并使用 `noexcept` 对这两种行为进行限制，例如：

```
void may_throw();           // 可能抛出异常
void no_throw() noexcept;   // 不可能抛出异常
```

使用 `noexcept` 修饰过的函数如果抛出异常，编译器会使用 `std::terminate()` 来立即终止程序运行。

`noexcept` 还能够做操作符，用于操作一个表达式，当表达式无异常时，返回 `true`，否则返回 `false`。

```
#include <iostream>

void may_throw() {
    throw true;
}

auto non_block_throw = []{
    may_throw();
};

void no_throw() noexcept {
    return;
}

auto block_throw = []() noexcept {
    no_throw();
};

int main()
{
    std::cout << std::boolalpha
    << "may_throw() noexcept? " << noexcept(may_throw()) << std::e
endl
    << "no_throw() noexcept? " << noexcept(no_throw()) << std::e
endl
    << "lmay_throw() noexcept? " << noexcept(non_block_throw())
<< std::endl
    << "lno_throw() noexcept? " << noexcept(block_throw()) << std
::endl;

    return 0;
}
```

noexcept 修饰完一个函数之后能够起到封锁异常扩散的功效，如果内部产生异常，外部也不会触发。例如：


```
try {
    may_throw();
} catch (...) {
    std::cout << "捕获异常, 来自 my_throw()" << std::endl;
}

try {
    non_block_throw();
} catch (...) {
    std::cout << "捕获异常, 来自 non_block_throw()" << std::endl;
}

try {
    block_throw();
} catch (...) {
    std::cout << "捕获异常, 来自 block_throw()" << std::endl;
}
```

最终输出为：

```
捕获异常, 来自 my_throw()
捕获异常, 来自 non_block_throw()
```

四、字面量

原始字符串字面量

传统 C++ 里面要编写一个充满特殊字符的字符串其实是非常痛苦的一件事情，比如一个包含 HTML 本体的字符串需要添加大量的转义符，例如一个 Windows 上的文件路径经常会：`C:\\\\What\\\\The\\\\Fxxk`。

C++11 提供了原始字符串字面量的写法，可以在一个字符串前方使用 `R` 来修饰这个字符串，同时，将原始字符串使用括号包裹，例如：

```
#include <iostream>
#include <string>

int main() {
    std::string str = R"(C:\\What\\The\\Fxxk)";
    std::cout << str << std::endl;
    return 0;
}
```

自定义字面量

C++11 引进了自定义字面量的能力，通过重载双引号后缀运算符实现：

```
// 字符串字面量自定义必须设置如下的参数列表
std::string operator"" _wow1(const char *wow1, size_t len) {
    return std::string(wow1)+"woooooooooow, amazing";
}

std::string operator"" _wow2 (unsigned long long i) {
    return std::to_string(i)+"woooooooooow, amazing";
}

int main() {
    auto str = "abc"_wow1;
    auto num = 1_wow2;
    std::cout << str << std::endl;
    std::cout << num << std::endl;
    return 0;
}
```

自定义字面量支持四种字面量：

1. 整型字面量：重载时必须使用 `unsigned long long` 、 `const char *` 、模板字面量算符参数，在上面的代码中使用的是前者；
2. 浮点型字面量：重载时必须使用 `long double` 、 `const char *` 、模板字面量算符；

3. 字符串字面量：必须使用 `(const char *, size_t)` 形式的参数表；
4. 字符字面量：参数只能是 `char` , `wchar_t` , `char16_t` , `char32_t` 这几种类型。

总结

本节介绍的几个特性是从仍未介绍的 C++1x 新特性里使用频次较靠前的特性了，`noexcept` 是最为重要的特性，它的一个功能在于能够阻止异常的扩散传播，有效的让编译器最大限度的优化我们的代码。

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook
该文件修改时间：2017-11-16 09:30:52

第九章 扩展主题：C++17 简介

一、本节内容

本节内容包括：

- 扩展主题: C++17 简介
 - 主要入选特性
 - 非类型模板参数的 `auto`
 - `std::variant<>`
 - 结构化绑定(Structured bindings)
 - 变量声明的强化
 - 未入选特性
 - Concepts

本节对即将到来的 C++17 进行介绍，几个月前（2016 年），目前为止，还没有一个正式发布的编译器来编译 C++17 特性的代码，本节作为扩展主题，供对 C++ 的历史进程及其未来发展感兴趣的读者阅读。

二、主要入选特性

非类型模板参数的 `auto`

模板参数分为两种，一种是类型模板参数，也是我们用得最多的一种：

```
template <typename T, typename U>
auto add(T t, U u) {
    return t+u;
}
```

里面的 `T` 和 `U` 都是类型模板参数。另一种是非类型模板参数，它可以让不同的字面量成为模板的参数：

```
template <typename T, int BufSize>
class buffer_t {
public:
    T& alloc();
    void free(T& item);
private:
    T data[BufSize];
}

buffer_t<int, 100> buf; // 100 作为模板参数
```

遗憾的是我们在编写模板的时候就必须明确非类型模板参数的具体类型，C++17 打破了这一限制，让我们能够在非类型模板参数中使用 `auto` 关键字，从而让编译器推导具体的类型：

```
template <auto value> void foo() {
    return;
}

foo<10>(); // value 被推导为 int 类型
```

std::variant<>

熟悉 `boost` 的人应该很早就听说过 `variant<>` 了。`variant<>` 可以用于存储和操作不同类型的对象。我们在前面（[对标准库的扩充：新增容器](#)）对于迭代 `std::tuple` 时，简单使用了 `boost::variant<>`。提供给 `variant<>` 的类型模板参数可以让一个 `variant<>` 从而容纳提供的几种类型的变量（在其他语言（例如 Python/JavaScript 等）表现为动态类型）。

C++17 正式将 `variant<>` 纳入标准库，摇身一变成为 `std::variant<>`，有了它之后，我们可以将前面的代码更改为：

```

#include <variant>
template <size_t n, typename... Args>
std::variant<Args...> _tuple_index(size_t i, const std::tuple<Ar
gs...>& tpl) {
    if (i == n)
        return std::get<n>(tpl);
    else if (n == sizeof...(Args) - 1)
        throw std::out_of_range("越界.");
    else
        return _tuple_index<(n < sizeof...(Args)-1 ? n+1 : 0)>(i
, tpl);
}
template <typename... Args>
std::variant<Args...> tuple_index(size_t i, const std::tuple<Arg
s...>& tpl) {
    return _tuple_index<0>(i, tpl);
}

```

结构化绑定(Structured bindings)

结构化绑定提供了类似其他语言中提供的多返回值的功能。到目前为止，我们可以通过 `std::tuple` 来构造一个元组，囊括多个返回值。但缺陷是显而易见的，我们没有一种简单的方法直接从元组中拿到并定义元组中的元素，尽管我们可以使用 `std::tie` 对元组进行拆包，但我们依然必须非常清楚这个元组包含多少个对象，各个对象是什么类型。

C++17 给出的结构化绑定可以让我们写出这样的代码：

```

std::tuple<int, double, std::string> f() {
    return std::make_tuple(1, 2.3, "456");
}
auto [x, y, z] = f(); // x, y, z 分别被推导为int, double, std::string

```

变量声明的强化

变量的声明在虽然能够位于任何位置，甚至于 `for` 语句内能够声明一个临时变量 `int`，但始终没有办法在 `if` 和 `switch` 语句中声明一个临时的变量。例如：

```
auto p = map_container.try_emplace(key, value);
if(!p.second) {
    //...
} else {
    //...
}
```

C++17 消除了这一限制，使得我们可以：

```
if (auto p = m.try_emplace(key, value); !p.second) {
    //...
} else {
    //...
}
```

三、未入选特性

C++ 组委员会在讨论投票最终确定 C++17 有很多提案，诸如

Concepts/Ranges/Module 等等，其中最受关注的就是 **Concepts**，可惜这一提案最终被拒，作为技术规范(Technical Specifications, TS) 将其发布。

Concepts TS

Concepts 是对 C++ 模板编程的进一步增强扩展。简单来说，**Concepts** 是一种编译期的特性，它能够让编译器在编译期时对模板参数进行判断，从而大幅度增强我们在 C++ 中模板编程的体验。使用模板进行编程时候我们经常会遇到各种令人发指的错误，这是因为到目前为止我们始终不能够对模板参数进行检查与限制，例如下面简单的两行代码会造成大量的几乎不可读的编译错误：

```
#include <list>
#include <algorithm>
int main() {
    std::list<int> l = {1, 2, 3};
    std::sort(l.begin(), l.end());
    return 0;
}
```

而这段代码出现错误的根本原因在于，`std::sort` 对排序容器必须提供随机迭代器，否则就不能使用，而我们知道 `std::list` 是不支持随机访问的。用 **Concepts** 的话来说就是：`std::list` 中的迭代器不满足 `std::sort` 中随机迭代器这个 **Concepts**（概念）的 **requirements**（要求）。有了 **Concepts**，我们就可以这样：

```
template <typename T>
requires Sortable<T>    // Sortable 是一个 concept
void sort(T& c);
```

缩写为：

```
template<Sortable T>    // T 是一个 Sortable 的类型名
void sort(T& c)
```

甚至于直接将其作为类型来使用：

```
void sort(Sortable& c); // c 是一个 Sortable 类型的对象
```

遗憾的是，C++组委会没有将 **Concepts** 纳入新标准，而是将其作为TS正式发布（其实早在 C++11 最终定案之前就已经有 **Concepts** 的呼声了，但 **Concepts TS** 是2015年才完整正式发布），也就是我们现在看到的 **Concepts TS**。C++组委会拒绝将 **Concepts** 纳入新标准的原因其实很简单，并不是技术层面上的原因，纯粹是觉得它还不够成熟。

Concepts TS 的发布到最后一次 C++17 的讨论会只相隔了不到四个月的时间，**Concepts** 的（唯一）实现只存在于一个未发布的 gcc 版本中。而 gcc 中关于 **Concepts** 的实现就是由撰写 **Concepts TS** 的人开发的，虽然它能够进行相关测

试，但还没有认真讨论过这份 TS 会产生哪些不良后果，更何况这份 TS 都没有被测试过。此外，已知的 **Concepts** 的一个明显的作用就是去辅助实现 Ranges TS 等提案，但实际上它们也没有被选入 C++17，所以可以把 **Concepts** 继续延后。

总结

总的来说，类似于 Concepts/Ranges/Modules 这些令人兴奋的特性并没有入选至 C++17，这注定了 C++17 某种意义上来说相较于 C++11/14 依然只是小幅度更新，但我们有望在 C++2x 中看到这些东西的出现，这些内容对于一门已经三十多岁『高龄』的编程语言，依然是充满魅力的。

进一步阅读的参考资料

1. [Final features of C++17](#)
2. [C++17: will it be great or just ok?](#)
3. [Why Concepts didn't make C++17?](#)
4. [C++11/14/17 编译器支持情况](#)

Copyright © 2016 Changkun Ou at shiyanlou.com all right reserved , powered by Gitbook
该文件修改时间：2017-11-16 09:30:52