2016 C++及系统软件技术大会 C++ and System Software Summit

C++和函数式编程

写紧凑、表意、可组合的代码

吴咏炜 wuyongwei@gmail.com

Ⅰ讲在前面

- 点到即止
 - 内容太多......
 - 有可工作的代码作参考(放在 GitHub 上)
- 充足的参考资料
- 主要模拟 Python 而非 Haskell



■函数式编程的基本概念

- 函数是「一等公民」
 - 函数可以作参数用,也可以返回函数
- 函数引用透明;有返回值而没有副作用
 - 易于测试,易于重现错误
 - 易于并行执行

■函数式编程的风格

- 强调数据的关系,而不是计算的步骤
- 使用递归
 - 传统函数式编程语言通常有尾递归优化
 - Not Python
- 使用序列和循环
 - Python!
- 使用高阶函数(函数的函数)
- 易于组合
- 易于惰性求值

Ⅰ最基本的函数式编程模式

- Map
 - map f [A, B, C, D, ...] = [fA, fB, fC, fD, ...]
- Reduce
 - reduce (+) 0 [A, B, C, D, ...] = 0 + A + B + C + D ...
- Compose
 - (f . g) A = f (g A)
 - $h = f \cdot g \rightarrow h A = f (g A)$

C++的函数式语言特性

- 函数对象
- std::bind1st/bind2nd-
- std::for_each
- std::transform
- std::accumulate

C++11/14 新增

- Lambda 表达式
- std::bind
- 基于区间的 for 循环
- std::function
- 自动类型推导
- 变参模板

函数对象

```
定义
class plus_n {
    int n_;
public:
    plus_n(int n) : n_(n) {}
    int operator()(int x) const { return x + n_; }
};
使用
plus_n plus_1(1);
cout << plus_1(41) << endl;</pre>
cout << plus_n(2)(40) << endl;</pre>
```

std::transform

```
int a[5] = \{0, 1, 2, 3, 4\};
transform(a, a + 5, a, plus_n(1));
// a = \{1, 2, 3, 4, 5\}
vector<int> v{0, 1, 2, 3, 4};
transform(v.begin(), v.end(), v.begin(), plus_n(1));
// v = \{1, 2, 3, 4, 5\}
transform(v.begin(), v.end(), a, plus_n(1));
// a = \{2, 3, 4, 5, 6\}
```

std::accumulate

▮自动类型推导

```
auto result = f(*begin(inputs));
decltype(auto) result = f(*begin(inputs));
decltype(f(*begin(inputs))) result;
decltype(f(declval<T>())) result;
```

- auto 和 decltype 有不同的推导规则
 - auto 抛弃引用和 cv 修饰
 - · decltype 可保留引用(详见参考资料)
- decltype(auto) 使用 decltype 的推导规则
- std::declval 可宣称一个某类型的参数

Lambda 表达式和 std::function

```
function<int(int)> get_strange_func(int n)
{
    if (n > 0)
        return [n](int x) { return x + n; };
    else
        return [n](int x) { return x * n; };
}
...

auto plus_1 = [](int x) { return x + 1; }; // 唯一类型,只能用 auto 接收 cout << plus_1(41) << endl; // 42 cout << get_strange_func(2)(40) << endl; // 42 cout << get_strange_func(-2)(-21) << endl; // 42</pre>
```

- Lambda 表达式等价于构造了一个函数对象
- 捕捉表达式可以转换为构造函数和成员变量
- 泛型 lambda 表达式可以认为是带模板的 operator()
- function 可以保存符合调用规范的函数、函数对象和 lambda 表达式

std::bind

```
using namespace std::placeholders;
auto plus_1 = bind(plus<int>(), _1, 1);
cout << plus_1(41) << endl;
// 42</pre>
```

■ 基于区间的 for 循环

```
for (auto x : r) {
    ...
}
```

大致相当于

```
for (auto it = begin(r); it != end(r); ++it) {
    auto x = *it;
    ...
}
```

变参模板

```
template <typename T>
constexpr auto sum(T x)
    return x;
template <typename T1, typename T2, typename... Targ>
constexpr auto sum(T1 x, T2 y, Targ... args)
    return sum(x + y, args...);
```

• 编译期递归!

Ⅰ小串烧

```
int main()
{
    (void)sum<int, int, int, int>; // GCC requires this instantiation
    auto sqr = [](auto x) { return x * x; };
    auto sqr list = [sqr](auto x) { return fmap(sqr, x); };
    auto sum_list = [](auto x) { return reduce(plus<int>(), x); };
    vector<int> v{1, 2, 3, 4, 5};
    tuple<int, int, int, int, int> t\{1, 2, 3, 4, 5\};
    auto sum sqr list = compose(sum list, sqr list);
    cout << sum sqr list(v) << endl;</pre>
    // 55
    cout << pipeline(v, sqr list, sum list) << endl;</pre>
    // 55
    cout << apply(sum<int, int, int, int>, fmap(sqr, t)) << endl;</pre>
    // 55
    cout << reduce(plus<int>(), fmap(sqr, t), ∅) << endl;</pre>
    // 55
```



A Tale of Two Languages (I)

```
#!/usr/bin/env python
#coding: utf-8
import sys

def main():
    for line in sys.stdin:
        print(line.rstrip('\n'))

if __name__ == '__main__':
    main()
```

```
#include <iostream>
#include <string>
using namespace std;

int main()
{
    string line;
    for (;;) {
        getline(cin, line);
        if (!cin) {
            break;
        }
        cout << line << '\n';
    }
}</pre>
```

I istream_line_reader (C++98 兼容)

std::istream& stream_;

};

```
class istream_line_reader {
                                                          #include <iostream>
public:
   class iterator { // implements InputIterator
                                                          #include <string>
      typedef const std::string& reference;
                                                          #include "istream line reader.h"
                  erence operator*()
             *terator(std::istream& is) : stream_(&is)
                                                          using namespace std;
      { ++*this.
      reference operator assert(stream_ != nullptr);
                                                          int main()
         return line; return line;
      value_typq* operator->()
                                                               istream line reader reader(cin);
         assert(stream_ != NULL);
                                                               for (istream line reader::iterator
         return &line_;
                                                                            it = reader.begin();
      iterator& operator++()
                                                                      it != reader.end(); ++it) {
         getli
                                                                     cout << *it << '\n';
         return *this; getline(*stream_, line_);}
      iterator operator (ft) (!*stream ) {
         iterator temp(*this); stream = nullptr;
         return temp;
      bool operator == (coast iterator % rhs) const
      { return stream = return; *this; bool operator!=(cdrseturn; rhs)this;
      { return loperator==(rhs);
   private:
      std::istream* stream_;
      std::string line_;
  };
   explicit istream_line_reader(std::istream& is)
      : stream (is) {}
   iterator begin() { return iterator(stream_); }
   iterator end() const { return iterator(); }
private:
                                                                           2016 C++ and System Software Summit
```

■ 等效的 C++11 代码

```
C + +98
```

C + +11

```
#include <iostream>
#include <string>
#include "istream_line_reader.h"

using namespace std;

int main()
{
    for (auto& line :
        istream_line_reader(cin)) {
        cout << line << '\n';
    }
}</pre>
```

A Tale of Two Languages (II)

```
#include ...
#!/usr/bin/env python
#coding: utf-8
                                                  using namespace std;
import sys
                                                  template <typename C>
def backsort(lines):
                                                  template <typename C>
    pairs = [] 
    for line in lines:
        pairs.append(('.'.join(
            reversed(line.split('.'))), line))
    return map(lambda item: item[1], _
               sorted(pairs))
def main():
    result = backsort(map(
               lambda line: line.rstrip('\n')
               sys.stdin))
    for line in result:
        print(line)
                                                    _return result:
if name == ' main ':
    main()
```

```
vector<string> split(const string& str, char delim);
string join(const C& str_list, char delim);
vector<string> backsort(C&& lines)
  vector<pair<string, string>> pairs;
    for (auto& line : lines) {
        auto split line = split(line, '.');
        reverse(split line.begin(),
                split line.end());
        pairs.emplace back(
            make_pair(join(split_line, '.'), line));
   sort(pairs.begin(), pairs.end());
    vector<string> result(pairs.size());
   transform(pairs.begin(), pairs.end(),
              result.begin(),
              [](const auto& pr) { return pr.second; });
int main(int argc, char* argv[])
    auto result = backsort(istream_line_reader(cin));
   for (auto& item : result) {
        cout << item << endl;</pre>
```

▮ 简化后的 backsort

Old

New

A Tale of Two Languages (IIIa)

```
#!/usr/bin/env python
#coding: utf-8
import sys
def cat(files):
    for fn in files:
        with open(fn) as f:
            for line in f:
                yield line.rstrip('\n')
def backsort(lines):
    result = {}
    for line in lines:
        result['.'.join(
            reversed(line.split('.')))] = line
    return map(lambda item: item[1],
               sorted(result.items()))
def main():
    if sys.argv[1:]:
        result = backsort(cat(sys.argv[1:]))
    else:
        result = backsort(map(
            lambda line: line.rstrip('\n'),
            sys.stdin))
    for line in result:
        print(line)
if name == ' main ':
    main()
```

可能的解

- 像 istream_line_reader 一样实现个辅助类
 - 前者还算通用,这个.....好复杂啊
- 实现类似于 cat 的函数, 然后回调/调用 backsort
 - 哈, backsort 得改, 这个接口没法用
- · 把文件内容全读进来,再传给 backsort
 - 好吧,这次可以;如果函数做的事情不需要在内存保存所有数据,文件又超过内存大小呢?
- 实现 generator 模式!
 - 推荐 Boost.Coroutine2

■ 简化 istream_line_reader (示意)

```
class istream_line_reader {
public:
    class iterator { // implements InputIterator
        typedef const std::string& reference;
        typedef std::string value type;
        iterator() : stream (NULL) {}
        explicit iterator(std::istream& is) : stream_(&is)
        { ++*this; }
        reference operator*()
            assert(stream_ != NULL);
            return line ;
        value type* operator->()
            assert(stream != NULL);
            return &line_;
        iterator& operator++()
            getline(*stream_, line_);
            if (!*stream_) {
                stream = NULL;
            return *this;
        iterator operator++(int)
            iterator temp(*this);
            ++*this;
            return temp;
        bool operator==(const iterator& rhs) const
        { return stream_ == rhs.stream_; }
        bool operator!=(const iterator& rhs) const
        { return !operator==(rhs); }
    private:
        std::istream* stream_;
        std::string line_;
    };
    explicit istream_line_reader(std::istream& is)
        : stream (is) {}
    iterator begin() { return iterator(stream_); }
    iterator end() const { return iterator(); }
private:
    std::istream& stream_;
};
```

```
typedef boost::coroutines2::coroutine<const</pre>
string&> coro t;
void read istream(coro t::push type& yield,
                  std::istream& is)
{
    for (;;) {
        string line;
        if (getline(is, line)) {
            yield(line);
        } else {
            break;
auto istream line reader(std::istream& is)
    return coro t::pull type(
        boost::coroutines2::fixedsize_stack(),
        bind(read istream, 1, std::ref(is)));
```

■ 改造 backsort

```
typedef boost::coroutines2::coroutine<const string&> coro t;
void cat(coro_t::push_type& yield, int files_cnt, char* file_names[])
    for (int i = 0; i < files cnt; ++i) {</pre>
        ifstream ifs(file names[i]);
        for (auto& line : istream_line_reader(ifs)) {
            yield(line);
        }
int main(int argc, char* argv[])
    vector<string> result;
    if (argc > 1) {
        result = backsort(coro t::pull type(
            boost::coroutines2::fixedsize stack(),
            bind(cat, 1, argc - 1, argv + 1));
```

Pythagorean Triples (Haskell)

Pythagorean Triples (Python)

```
#!/usr/bin/env python
#coding: utf-8
def gen pythagorean triple():
    7 = 1
    while True:
        z += 1
        yield [(x, y, z)] for x in range(1, z)
                          for y in range(x, z)
                          if x^{**2} + y^{**2} == z^{**2}
def main():
    generator = gen_pythagorean_triple()
    triples = []
    while len(triples) < 10:</pre>
        triples += next(generator);
    print(triples)
if name == ' main ':
    main()
```

Pythagorean Triples (C++ Ranges)

```
#include <iostream>
#include <range/v3/all.hpp>
using namespace ranges;
int main()
    // Lazy ranges for generating integer sequences
    auto const intsFrom = view::iota;
    auto const ints = [=](int i, int j) { return view::take(intsFrom(i), j - i + 1); };
    // Define an infinite range of all the Pythagorean triples:
    auto triples = view::for each(intsFrom(1), [=](int z)
            return view::for each(ints(1, z), [=](int x)
                return view::for_each(ints(x, z), [=](int y)
                    return yield if(x*x + y*y == z*z, std::make tuple(x, y, z));
                });
            });
        });
    // Display the first 10 triples
    for (auto triple : triples | view::take(10)) {
        std::cout << '('
            << std::get<0>(triple) << ','
            << std::get<1>(triple) << ',
            << std::get<2>(triple) << ')' << '\n';
                                                             2016 C++ and System Software Summit
```

I 用 ranges 优化 backsort

Old

New



III. 总结

【C++ 里的函数式编程

优点

- 提供更清晰、更好的抽象
- 代码更紧凑
- 代码更容易组合
- 惰性求值可导致更好的性能

缺点

- 造轮子比较辛苦
- 类型系统比大多数语言复杂
 - POD 值类型, 非 POD 值类型, 各种不同的引用
- 出错信息可能让人发疯

【C++ 的坑

```
backsort1c.cpp: In instantiation of
                                                                                    'std::vector<std:: cxx11::basic string<char> > backsort(C&&)
                                                                                    [with C = istream line reader]':
                                                                                    backsort1c.cpp:61:52: required from here
         const?
                                              constexpr?
                                                                                    backsort1c.cpp:55:16: error: no matching function for call to
                                                                                    'fmap(backsort(C&&) [with C =
                                                                                    istream line reader]::<lambda(auto:4&)>,
                                                                                    std::vector<std::pair<std:: cxx11::basic string<char>,
                                                                                    std:: cxx11::basic string<char> > >&)'
                                                                                         return fmap([](auto& pr) { return pr.second; },
                                                                                                    pairs):
                                                                  noexcept?../nywa/nywa/functional.h:282:16: note: candidate:
                                                                                    template<class _Fn, class ... _Targs> constexpr auto
                                                                                    nvwa; fmap( Fn, const nvwa::optional< Targs>& ...)
                                                                                     constexpr auto fmap( Fn f, const optional < Targs>&... args)
                         expression
                                                                                    ../nvwa/nvwa/functional.h:282:16: note: template argument
                                                                                    deduction/substitution failed:
                                                                                    backsort1c.cpp:55:16: note:
                                                                                    'std::vector<std::pair<std:: cxx11::basic string<char>,
                                                                                    std:: cxx11::basic string<char> > ' is not derived from
                                                                                    'const nvwa::optional<_Targs>'
                                                                                         return fmap([](auto& pr) { return pr.second; },
                                                                                                    pairs):
             glvalue
                                           rvalue
                                                                                    In file included from backsort1c.cpp:6:0:
                                                                                    ../nvwa/nvwa/functional.h:303:16: note: candidate:
                                                                                    template<class Fn, class ... Targs> constexpr auto
                                                                                    nvwa::fmap(_Fn, nvwa::optional<_Targs>&& ...)
                                                                                     constexpr auto fmap(_Fn f, optional<_Targs>&&... args)
                                                                                    ../nvwa/nvwa/functional.h:303:16: note: template argument
                                                                                    deduction/substitution failed:
                                                                                    backsort1c.cpp:55:16: note:
Ivalue
                                                         prvalue
                            xvalue
                                                                                    'std::vector<std::pair<std:: cxx11::basic string<char>,
                                                                                    std::__cxx11::basic_string<char> > >' is not derived from
                                                                                    'nvwa::optional< Targs>'
                                                                                         return fmap([](auto& pr) { return pr.second; },
                                                                                                    pairs);
```

In file included from backsort1c.cpp:6:0: Software Summit

template<template<class, class> class OutCont, template<class>

../nvwa/nvwa/functional.h:329:16: note: candidate:

■参考资料 I

C++ 编程

- cppreference.com, C++ Reference, http://en.cppreference.com/w/ (有中文版)
- Eric Niebler, "Range Comprehensions", http://ericniebler.com/2014/04/27/range-comprehensions/
- Eric Niebler, "D4128: Ranges for the Standard Library: Revision 1", <u>https://ericniebler.github.io/std/wg21/D4128.html</u>
- Eric Niebler, "Range-v3", https://ericniebler.github.io/range-v3/
- Oliver Kowalke, "Boost.Coroutine2", http://www.boost.org/doc/libs/release/libs/coroutine2/
- Thomas Becker, "C++ Rvalue References Explained", <u>http://thbecker.net/articles/rvalue_references/section_01.html</u>
- Thomas Becker, "C++ auto and decltype Explained", <u>http://thbecker.net/articles/auto_and_decltype/section_01.html</u>

▮参考资料 II

函数式编程

- 阮─峰,《函数式编程初探》, http://www.ruanyifeng.com/blog/2012/04/functional_programming.html
- 陈皓,《函数式编程》, http://coolshell.cn/articles/10822.html
- Slava Akhmechet, "Functional Programming For The Rest of Us", http://www.defmacro.org/ramblings/fp.html (有中文版)

我的函数式参考代码

- fmap、reduce、compose、pipeline、apply、curry、optional 和 file_line_reader 的实现, https://github.com/adah1972/nvwa/tree/master/nvwa (functional.h, file_line_reader.h/cpp)
- 我的博客文章, http://wyw.dcweb.cn/index.htm#articles
- 演讲文件和代码, https://github.com/adah1972/cpp_conf_china_2016

其他

 Guido van Rossum, "Tail Recursion Elimination", http://neopythonic.blogspot.com.au/2009/04/tail-recursion-elimination.html



IV. 备用材料

Currying

- 看起来很美
 - plus_1 = (+) 1
 - plus_1 2 \rightarrow (+) 1 2 \rightarrow 3
- C++ 里可以手工用 bind 或 lambda 实现
 - auto plus_1 = bind(plus<int>(), _1, 1);
- 可以用库实现通用的方案(产品项目中不推荐)
 - auto curried_plus = nvwa::make_curry<int(int, int)>(plus<int>());
 - auto plus_1 = curried_plus(1);
 - curried_plus(1)(1) → 2
 - plus_1(1) \rightarrow 2

I fmap (simplified)

```
template <class T1, class T2>
struct can reserve
    static const bool value = (sizeof(reserve<T1>(nullptr)) == sizeof(good) &&
                               sizeof(size<T2> (nullptr)) == sizeof(good));
};
template <class T1, class T2>
void try reserve(T1&, const T2&, false type)
template <class T1, class T2>
void try reserve(T1& dest, const T2& src, true type)
    dest.reserve(src.size());
template <template <typename, typename> class OutCont = vector,
          template <typename> class Alloc = allocator,
          typename Fn, class Cont>
constexpr auto fmap(Fn&& f, const Cont& inputs)
    typedef decay t<decltype(f(*begin(inputs)))> result type;
    OutCont<result type, Alloc<result type>> result;
    trv reserve(result, inputs,
        integral constant<bool, can reserve<decltype(result), Cont>::value>());
    for (const auto& item : inputs)
        result.push back(f(item));
    return result:
                                                        2016 C++ and System Software Summit
```

■ 惰性的 fmap (示例)

```
template <typename Rs, typename Fn, typename Rng>
auto fmap view impl(
    typename boost::coroutines2::coroutine<Rs>::push_type& yield,
    Fn&& f, Rng&& inputs)
    for (auto& x : inputs) {
        yield(f(x));
template <typename Fn, typename Rng>
auto fmap view(Fn&& f, Rng&& inputs)
    typedef decltype(f(*begin(inputs))) result type;
    return typename boost::coroutines2::coroutine<result type>::pull type(
        boost::coroutines2::fixedsize stack(),
        bind(fmap view impl<result type, Fn, Rng>,
             forward<Fn>(f),
             forward<Rng>(inputs)));
cout << sum list(fmap view(sqr, v)) << endl;</pre>
                                                      2016 C++ and System Software Summit
```

Ⅰ不使用异常的错误返回方式

- optional 模板 (将出现在 C++17 中)
- lift_optional 函数 (nvwa)

I用 file_line_reader 优化读行

Old

New

2016 C++ and System Software Summit

```
void cat(coro t::push type& yield,
                                                         void cat(coro t::push type& yield,
        int files cnt,
                                                                  int files cnt,
        char* file names[])
                                                                  char* file names[])
                                                             for (int i = 0; i < files cnt; ++i) {</pre>
   for (int i = 0; i < files cnt; ++i) {</pre>
       ifstream ifs(file_names[i]);
                                                                 FILE* fp = fopen(file_names[i], "r");
       for (auto& line : istream_line_reader(ifs)) {
                                                                 if (!fp) continue;
                                                                 for (auto& line : file_line_reader(fp)) {
           yield(line);
                                                                     yield(line);
                                                                 fclose(fp);
      根据环境不同, file_line_reader 性能最多可能有 istream_line_reader 的 15 倍
template <typename C>
                                                         template <typename C>
vector<string> backsort(C&& lines)
                                                         vector<string> backsort(C&& lines)
   vector<pair<string, string>> pairs;
                                                             vector<pair<string, string>> pairs;
                                                             for (const string& line : lines) {
   for (auto& line : lines) {
       auto split_line = action::split(line, '.');
                                                                 auto split line = action::split(line, '.');
                                                                 pairs.emplace back(make_pair(
        pairs.emplace back(make pair(
           to <string>(split line
                                                                     to <string>(split line
                       view::reverse
                                                                                view::reverse
                       view::join('.')),
                                                                                view::join('.')),
           line));
                                                                     line));
   ranges::sort(pairs);
                                                             ranges::sort(pairs);
   return pairs | view::transform([](const auto& pr)
                                                             return pairs | view::transform([](const auto& pr)
                      return pr.second;
                                                                                return pr.second;
                                                                           });
```