STL: Библиотека алгоритмов

- В STL для работы с контейнерами имеется *библиотека алгоритмов* с функциями для сортировки, поиска, копирование . . .
- Алгоритмы глобальные функции оперирующие с итераторами
- Существуют «специализированные» версии алгоритмов, например list.sort(), map.find(), которые более эффективны по сравнению с «глобальными»
- Для настройки алгоритмов предусмотрен механизм «подключения» пользовательских функций
- Основная задача при работе с алгоритмами состоит в написании пользовательских функций

Алгоритмы: простейшая классификация

- Немодифицирующие алгоритмы: поиск, подсчёт числа элементов . . .
- Модифицирующие: удаление, вставка, замена . . .
- Алгоритмы сортировки и работы с сортированными последовательностями

Заголовочные файлы

```
#include <algorithm> // main header file
#include <numeric> // numeric operations
// auxiliary headers
```

Немодифицирующие алгоритмы

🕏 Функции которы	ве не меняют контейнер
<pre>for_each() for_each_n()</pre>	вызывает заданную функцию для каждого элемента для первых N элементов
<pre>find() find_if()</pre>	находит первый заданный элемент находит первый элемент удовлетворяющий условию
<pre>count() count_if()</pre>	число элементов равных заданному число элементов удовлетворяющих условию
<pre>min_element() max_element() minmax_element()</pre>	минимальный элемент максимальный элемент возвращает пару <i>min,max</i>
<pre>all_of() any_of() none_of()</pre>	проверяет, истинен ли предикат для всех, для любого, или ни для одного из элементов

for_each() u for_each_n()

- for_each(beg, end, UnaryFunction op)
 - «функция» op(elem) вызывается для всех элементов из [beg,end), а возвращаемое значение op() игнорируется
 - for_each возвращает копию функционального объекта ор() после выполнения последнего вызова
- ② for_each_n(beg, size_t N, UnaryFunction op), since C++17
 - то же, что for_each(beg,beg+N,op), но возвращает итератор на beg+N
 - ecли op(elem) модифицирует элемент то for_each() изменит содержимое контейнера
- for_each() это один из самых универсальных алгоритмов

print P = std::for_each(V.cbegin(), V.cend(), print(0));

cout << P.count << " items printed\n"; // 5 items printed</pre>

cout << endl: // 1 2 3 4 5

```
Pipumep для for_each_n()

// print N elements and return iterator to first+N

size_t N = 3;

auto itN = std::for_each_n(cbegin(V), N, // number of elements

[](const auto& x) {cout<<' '<<x:});</pre>
```

обратите внимание

cout << endl; // 1 2 3

print(0) – это вызов конструктора класса print

cout << "*itN=" << *itN << endl; // *itN=4

- for_each возвращает «функцию», for_each_n возвращает итератор
- проверка выхода за пределы контейнера не выполняется
- старайтесь везде использовать константные итераторы [cbegin,cend)

```
for_each() может менять элементы в контейнере «в старом стиле»
struct add_value {
   add_value(int v=1) : the_value(v) {}
   void operator() (int& elem) const {elem+=the_value;}
   int the_value; // value to add to elements
};
std::for_each(V.begin(), V.end(), add_value()); // add_value() and
std::for_each(V.cbegin(), V.cend(), print()); // print() are ctrs
cout << end1; // 2 3 4 5 6</pre>
```

```
• ...то же самое с лямбда функцией std::for_each( begin(V), end(V), [](int& x) {x+=10;}); std::for_each( cbegin(V), cend(V), [](int x) {cout<<x<' ';}); cout << endl; // 12 13 14 15 16
```

```
обратите внимание, что в модифицирующем варианте:
```

- надо использовать не константные итераторы [begin,end)
- тип аргумента модифицирующей функции: int&

Минимальный и максимальный элементы

```
min_element(), max_element(), minmax_element()
```

• параметры функций одинаковы:

```
min_element (beg, end, [CompFunc op])
max_element (beg, end, [CompFunc op])
minmax_element (beg, end, [CompFunc op])
```

- функции min_,max_ возвращают итератор, a minmax_ пару итераторов, на минимальный/максимальный элемент из [beg,end)
- функция сравнения op(elem1,elem2) должна возвращать true если elem1 < elem2; версия без op() использует operator<()

пример на minmax_ с лямбда функцией для сравнения

Модифицирующие алгоритмы

transform()	применяет функцию к элементам из заданного диапазона и сохраняет результат в целевом диапазоне
replace(),	заменяет элементы удовлетворяющие условию
<pre>copy() copy_n() copy_if()</pre>	копирует, начиная с первого элемента копирует ${\tt n}$ элементов копирует элементы удовлетворяющие условию
<pre>move(), remove(), unique()</pre>	перемещает элементы на новое место удаляет элементы удовлетворяющие условию удаляет соседние эквивалентный элементы
<pre>generate() fill() iota()</pre>	заменяет элементы результатом операции заполняет одним элементом заполняет возрастающей серией

transform()

- transform(srcBeg, srcEnd, destBeg, UnaryFunction op)
 - выполняет функцию op(elem) для элементов из [srcBeg,srcEnd) и записывает возвращаемое значение в [destBeg,...)
 - возвращает позицию последнего элемента в принимающем контейнере
 - srcBeg и destBeg могут принадлежать одному контейнеру



принимающий контейнер должен быть достаточно большим чтобы вместить входящие элементы

```
🖙 Вспомогательная функция (generic lambda) печати
```

```
auto prt = [](std::string_view msg, auto&& ss) {
   cout << msg << "{ ";
   for( const auto& s : ss ) {cout << s << " ";}
   cout << "}\n";
};</pre>
```

```
источник и целевой диапазоны совпадают
std::vector V {2,3,5,7};
```

// change the sign of all elements
std::transform(cbegin(V),cend(V), // source range
 begin(V), // destination
 [](int x) {return -x;}); // operation
prt("after: ",V); // after: { -2 -3 -5 -7 }

```
std::list<int> L( size(V) );  // list of required size
std::transform( cbegin(V),cend(V), // source range
                // destination
    begin(L).
     [](int x) {return x+8;}); // operation
prt("V= ",V); // V= { -2 -3 -5 -7 } без изменений
prt("L= ",L); // L= { 6 5 3 1 }
• итератор вставки в конец контейнера: back_inserter(Container)
std::transform( cbegin(V)+2,cend(V), // source range
     [](int x) {return -x/2;}); // operation
```

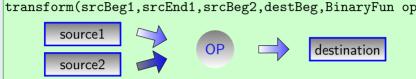
• сохранение в другом контейнере

prt("after inserter: L= ",L); // { 2 6 5 3 1 2 3 }

② transform(srcBeg1,srcEnd1, srcBeg2, destBeg, BinaryFun op)

В версия для двух источников

transform(srcBeg1,srcEnd1,srcBeg2,destBeg,BinaryFun op)



• Пример с двумя источниками

```
generate() u iota()
```

iota(beg, end, T startValue)

```
generate(beg, end, GeneratorFun op)
заменяет элементы контейнера на результат вызова ор() (без аргумента)
```

определена в <numeric>

```
заполняет контейнер приращениями: startValue, startValue+1,...
std::arrav<int.10> A:
// fill 'A' with random numbers from 0 to 99
std::generate( begin(A),end(A), // source range
      []() {return rand()%100;} ); // operation
prt("generate: A= ",A); // generate: A= { 7 49 73 58 30 72 44 78 23 9 }
// fill 'A' with sequentially increasing numbers
std::iota(begin(A),end(A), 1); // starting with 1
prt("iota: A= ",A); // iota: A= { 1 2 3 4 5 6 7 8 9 10 }
```

Алгоритмы сортировки

сортировка		
sort()	сортировка в порядке возрастания	
<pre>partial_sort()</pre>	сортировка первых <u>n</u> -элементов	
stable_sort()	сортировка с сохранением порядка равных элементов	
разбиение		
partition()	делит элементы на две группы относительно заданно-	
	го предиката: (true-els,false-els)	
<pre>stable_partition()</pre>	partition c сохранением порядка равных элементов	
двоичный поиск в сортированном контейнере		
binary_search()	проверка элемента	
<pre>lower_bound()</pre>	позиция первого элемента не меньшего заданного	
upper_bound()	позиция первого элемента большего заданного	
equal_range()	возвращает пару: lower_bound(),upper_bound()	

sort() u partial_sort()

- sort(beg, end, Compare op)
 sort() сортирует элементы в промежутке [beg,end) используя
 op(elem1,elem2) как критерий сортировки
- ② partial_sort(beg, endSort, end, Compare op)

 в partial_sort() сортированным будет лишь промежуток
 [beg,endSort) (endSort ≤ end)

Эффективность алгоритмов

- sort(): $\sim n \times \log(n)$ вызовов op(), n = (end beg)
- ullet partial_sort(): $\sim n imes \log(m)$ вызовов op(), $m = (\mathsf{endSort} \mathsf{beg})$
- stable_sort(): $\sim n \times \log^2(n)$ вызовов op()

```
● Сортировка «по умолчанию»: в возрастающем порядке std::vector<int> V(5,0); std::generate(begin(V),end(V), []() {return rand()%100;});
```

```
prt("before: V= ",V); // { 7 49 73 58 30 }
std::sort(begin(V),end(V)); // sort in ascending order (default)
prt("sort(<): V= ",V); // { 7 30 49 58 73 }</pre>
```

```
• Сортировка в убывающем порядке
```

```
std::sort(begin(V),end(V), [](int x,int y) {return x>y;});
prt("sort(>): V= ",V); // { 73 58 49 30 7 }
```

```
• пример на partial_sort()
std::vector W {3,8,3,6,7,9,1,5,10,6};
// select first M largest elements
size_t M = 5; // must be < size(W)</pre>
std::partial_sort( begin(W),
                                // begin of the range
      begin(W)+M.
                                       // end of sorted range
      end(W),
                                       // end of full range
      [](int x,int y) return x>y;); // comparison function
cout << "sort in [beg,beg+"<<M<<"): ";</pre>
prt("",W); // sort in [beg,beg+5): { 10 9 8 7 6 3 1 3 5 6 }
```

Обратите внимание

выражения begin(W)+5 или end(W)-2 допустимы лишь для контейнеров с произвольным доступом: vector<>, array<>

📭 приращение итератора общего вида:

```
_____ increments iterator it by 5 elements _____
auto it = begin(W);
std::advance(it,5); // the same meaning as it+=5
```

partition()

- partition(beg, end, UnaryPredicate op)
 - упорядочивает элементы так, что сначала идут элементы для которых (op(el)==true), а затем те, для которых (op(el)==false)
 - возвращает итератор на начало второй группы

```
std::vector<int> U(10.0):
std::generate(begin(U),end(U), []() {return rand()%100;});
prt("U= ",U); // U= { 72 44 78 23 9 40 65 92 42 87 }
size_t mid = 50:
// reorders the vector so that (elements < mid) precede all others
auto itm = std::partition(begin(U),end(U),  // range
      [mid](const auto& e) {return e<mid;}); // predicate</pre>
prt("after: ",U); // after: { 42 44 40 23 9 78 65 92 72 87 }
prt("true: ",std::vector(begin(U),itm)); // { 42 44 40 23 9 }
prt("false: ",std::vector(itm, end(U))); // { 78 65 92 72 87 }
```

Параллельные алгоритмы в С++17

Политика выполнения алгоритмов (execution policy)

основная идея — с помощью дополнительного параметра указать, как какой либо алгоритм должен выполняться: последовательно, параллельно или векторизованно

```
std::algorithm_name(policy, /* normal args... */);
```

• параметр policy может принимать значения:

• множество алгоритмов, но не все, имеют policy параметр; «поведение» некоторых алгоритмов немного отличается от стандартной версии

```
Последовательное, параллельное и векторизованное выполнение
// pseudocode example
auto mul = [](float x,float y) {return x*y;};
cbegin(x), cend(x), // 1st vector
                       // 2nd vector
             cbegin(y),
             begin(z), // output vector
             mul):
                          // operation
// sequential execution:
load x[i]; load y[i]; mul(x[i],y[i]); store into z[i]
// parallel execution: multiple threads
new_thread(load x[i]; load y[i]; mul(x[i],y[i]); store into z[i])
// vectorized execution: SIMD (Single-Instruction, Multiple-Data)
load x[i...i+3]; load v[i...i+3];
mul(four elements at once); store into z[i..i+3]
```

• пример с сортировкой

```
void TimeSort(std::vector<double> v, size_t ipolicy) {
   const auto start = std::chrono::steady_clock::now();
   std::string spolicy("Error");
   switch (ipolicy) {
      case 0: spolicy = "whitout policy";
         std::sort(begin(v), end(v)); break;
      case 1: spolicy = "sequenced policy";
         std::sort(std::execution::seq, begin(v), end(v)); break;
      case 2: spolicy = "parallel policy";
         std::sort(std::execution::par, begin(v), end(v)); break;
      case 3: spolicy = "parallel unsequenced policy";
         std::sort(std::execution::par_unseq, begin(v), end(v)); break;
   const auto finish = std::chrono::steady_clock::now();
   auto tm = std::chrono::duration_cast<std::chrono::milliseconds>
      (finish - start).count():
   std::cout << spolicy << ": " << tm << "ms\n";
};
```

• продолжение примера с сортировкой

```
std::vector<double> v(10'000'000);
std::generate(begin(v),end(v), []() {return double(rand())/RAND_MAX;});
for( size_t ipolicy : {0,1,2,3} ) {TimeSort(v,ipolicy);}
```

```
output: gcc(14.2.0) -std=c++20 -0 -ltbb (4-core i686) whitout policy: 845ms sequenced policy: 844ms parallel policy: 276ms parallel unsequenced policy: 277ms unsequenced policy: 843ms
```

Замечания

- Поддержка параллельных алгоритмов есть не у всех компиляторов
- Если выполнение невозможно в параллельном виде, все политики могут вернуться к последовательному выполнению
- При использовании параллельных алгоритмов на программисте лежит ответственность избегать data races и deadlocks

STL: Библиотека файловой системы

The Filesystem Library

since C++17

- Класс описывающий путь к файлу (path object)
 - манипуляции с путями, информация о пути
 - информация о типе объекта связанного с путем: файл, директория, ссылка . . .
 - манипуляции с файлами: создание, копирование, перемещение
- Итераторы по директории, записи directory_entry
- А так же: права доступа, временные файлы, специальные файлы . . .

```
Tpocтpaнство имен std::filesystem
#include <filesystem> // header
namespace fs = std::filesystem; // shortcut for std::filesystem
```

```
• пример операций с путем: fs::path
fs::path p1("/Users");
cout << p1 << endl; // "/Users"
p1.append("nefedov"); // adds a path with a directory separator
p1 /= "test"; // synonym for append()
cout << p1 << endl; // "/Users/nefedov/test"</pre>
p1.concat("/new_"); // only adds new characters to the end
p1 += "cpp/a.out"; // synonym for concat()
cout << p1 << endl; // "/Users/nefedov/test/new_cpp/a.out"</pre>
for (const auto& part : p1) { // path iteration
   cout << part << ", ";
}
cout << endl; // "/", "Users", "nefedov", "test", "new_cpp", "a.out"</pre>
PathInfo(p1); // следующий слайд
```

• пример: некоторые «информационные» функции для fs::path

```
void PathInfo(const fs::path& TestPath) {
   cout << "Path info for: " << TestPath << endl:</pre>
   cout << "canonical path " << fs::canonical(TestPath) << endl;</pre>
   cout << "exists= " << fs::exists(TestPath) << endl;</pre>
   cout << "root_name= " << TestPath.root_name() << endl;</pre>
   cout << "root_path=" " << TestPath.root_path() << endl;</pre>
   cout << "relative_path= " << TestPath.relative_path() << endl;</pre>
   cout << "parent_path= " << TestPath.parent_path() << endl;</pre>
   if (fs::is_regular_file(TestPath) ) {
      cout << "regular file= " << TestPath.filename() << endl;</pre>
      cout << " stem= " << TestPath.stem() << endl;</pre>
      cout << " extension= " << TestPath.extension() << endl;</pre>
      cout << " size(byte)= " << fs::file_size(TestPath) << endl;</pre>
      cout << " time= " << FileTime(TestPath) << endl:</pre>
   } // else if ( fs::is_directory(TestPath) ) { ... skip
```

output of PathInfo(p1)

```
Path info for: "/Users/nefedov/test/new_cpp/a.out"
canonical path "/Volumes/CaseSensitive/test/cpp/MyLectures/new_cpp/a.out"
exists=
root name=
root_path=
               11 / 11
relative_path= "Users/nefedov/test/new_cpp/a.out"
parent_path=
              "/Users/nefedov/test/new_cpp"
regular file=
               "a.out"
               " 2 "
  stem=
  extension= ".out"
  size(byte) = 41272
              Wed Feb 12 21:08:36 2025
  time=
```

Обратите внимание

- функции члены, такие как TestPath.filename(): быстрые, дешевые
- автономные (free-standing) функции, как fs::exists(TestPath): затратные, обычно обращаются к реальной файловой системе
- функция FileTime() для красивой печати времени, следующий слайд

• время последней модификации файла std::string FileTime(const fs::path& FileName) { std::string ret; // empty string for result value if(fs::exists(FileName) && fs::is_regular_file(FileName)) { auto ftime = fs::last write time(FileName): #if __cplusplus >= 202002L using namespace std::chrono_literals; ftime += 3h; // timezone UTC -> MSK ret = std::format(":%c", ftime); // C++20 #else // C++17 namespace cr = std::chrono; auto nowF = fs::file_time_type::clock::now(); auto dsec = cr::duration cast<cr::seconds>(nowF-ftime).count(): const auto now = cr::system_clock::now(); std::time_t tm = cr::system_clock::to_time_t(now) - dsec; ret = std::string(std::ctime(&tm)); // C-time #endif return ret:

if (fs::is_directory(entry.status())) fname += "/";

cout << " " << fname << endl:

cout << "== " << Nf << " items in total ==\n":

Nf++;

std::string fname = entry.path().filename(); // get filename

```
тест для директории
fs::path p2("/Users/nefedov/test/new_cpp");
PathInfo(p2);
output of PathInfo(p2)
Path info for: "/Users/nefedov/test/new_cpp"
canonical path "/Volumes/CaseSensitive/test/cpp/MyLectures/new_cpp"
exists=
               11 11
root_name=
root_path=
               11 / 11
relative_path= "Users/nefedov/test/new_cpp"
parent_path= "/Users/nefedov/test"
directory= "new_cpp" contains the following items:
```

```
root_path= "/"
relative_path= "Users/nefedov/test/new_cpp"
parent_path= "/Users/nefedov/test"
directory= "new_cpp" contains the following items:
    BitManip.cpp
    файл для теста.txt
    a.out
    dir_test/
== 4 items in total ==
```

Дополнительные слайды

- nth_element(beg, end, IterNth, Compare op)
 - ставит на нужное «сортированное» место E_n элемент и упорядочивает элементы E_k таким образом, что ор (E_j, E_i) ==false если i<n и j>n
 - op(E1,E2): функция сравнения, как в sort()
 - позиция n-th элемента задается итератором IterNth

- binary_search(beg, end, Value, Compare op)
 - проверяет наличие Value в отсортированном контейнере

• ор должен совпадать с алгоритмом который применялся для сортировки

auto lgt = [](int x,int y) {return x>y;}; // comparison function

```
std::sort(begin(W), end(W), lgt); // decreasing sort
prt("W= ",W); // W= { 10 9 8 7 6 6 5 3 3 1 }
// binary search must use exactly the same function as sorting
if( std::binary_search(cbegin(W),cend(W),3, lgt) ) {
```

} else {
 cout << "W does not contain 3" << endl;
} // W contains 3</pre>

cout << "W contains 3" << endl:

equal_range()

- equal_range(beg, end, Value, Compare op)
 - возвращает диапазон итераторов begV, endV, таких, что вставка Value перед ними не нарушает сортировку
 - ор должен совпадать с алгоритмом который применялся для сортировки

```
auto lgt = [](int x,int y) {return x>y;}; // Compare function
std::sort(begin(W), end(W), lgt);
prt("W= ",W); // W= { 10 9 8 7 6 6 5 3 3 1 }
for( size t i : {4.3} ) {
   auto [pL,pU] = std::equal_range(cbegin(W),cend(W),i,lgt);
   cout << i << " can be inserted in interval ["</pre>
         << std::distance(cbegin(W),pL) << ","
         << std::distance(cbegin(W),pU) << "]\n";
// 4 can be inserted in interval [7,7]
// 3 can be inserted in interval [7.9]
```

```
🖙 data-raice в параллельных вычислениях (см. пример с сортировкой)
// calculating the sum of the elements of V (uniformly random on [0,1))
std::atomic<double> s {0.}:
// double s {0.}; // ERROR: data raice for parallel execution
auto acc = [\&s] (const auto& v) { s += v;};
   case 1: spolicy = "sequenced policy";
      std::for_each(std::execution::seq,cbegin(V),cend(V),acc);
      break:
   case 2: spolicy = "parallel policy";
      std::for_each(std::execution::par,cbegin(V),cend(V),acc);
      break:
std::cout<<spolicy<<": sum="<<sum/size(V)<<" T="<<tm<<"ms\n";</pre>
```

```
sequenced policy: sum=0.500034 T=99ms
```

std::atomic<double> s

parallel policy: sum=0.500034 T=739ms par-unseq policy: sum=0.500034 T=739ms

sequenced policy: sum=0.500034 T=22ms
parallel policy: sum=0.169102 T=8ms
par-unseq policy: sum=0.0804907 T=8ms

double s // ERROR: data raice

• Печать дерева файлов

```
void DirTree(const fs::path& pathToDir, int level=0) {
   if (fs::exists(pathToDir) && fs::is_directory(pathToDir) ) {
      if( level == 0 ) { cout << "> " << pathToDir << "\n"; }</pre>
      auto shift = std::string(2+level*5,' ');
      for (const auto& entry : fs::directory_iterator(pathToDir)) {
         auto filename = entry.path().filename();
         if (fs::is_directory(entry.status()) ) {
            cout << shift << "[+] " << filename << "\n";</pre>
            DirTree(entry, level + 1);
         } else if ( fs::is_regular_file(entry.status()) ) {
            cout << shift << filename << "\n":</pre>
         } else {
            cout << shift << " [?]" << filename << "\n";</pre>
```

вызов DirTree() для рабочей директории

const fs::path wd { fs::current_path() }; // working directory DirTree(wd);

```
output DirTree(wd)
```

```
> "/Volumes/CaseSensitive/test/cpp/MyLectures/new_cpp"
 "BitManip.cpp"
  "файл для теста.txt"
 "a.out"
  [+] "dir_test"
       [+] "dir2"
            "file2"
            [+] "dir3"
                 "file3"
       "file1"
```