Map and Set

Key-Value Containers Concept, Maps, Sets, STL Algorithms

SoftUni Team

Technical Trainers

Software University

SoftUni

Software University

https://softuni.bg

Table of Contents

Have a Question?

Software University

sli.do

#cpp-advanced

Key-Value Containers

Key-Value Containers

- Real-World information is often "labeled" or "named"
  - Contacts usually have names and numbers/emails:

```
{George -> +359899123123} {NSA -> 1-301-688-6524}
{Bon Jovi -> [no info]}
```

- Labels can also be created by context – this is called "mapping"
  - E. g. numeric values mapped to their names

```
{1 -> "one"} {2 -> "two"} {3 -> "three"}
```

Key-Value Pairs (1)

- `std::pair<T1, T2>` can represent two values in one variable
  - `pair<string, int> namedNumber("five", 5);`
  - `#include<utility>`
  - `first` accesses the first value, `second` accesses the second value
  - `first` and `second` can be read and written directly, e.g.:
    `namedNumber.first="six"; namedNumber.second=6;`

```
pair<string, string> contact("George", "***@gmail.com");
contact.first = "George Georgiev";
cout << contact.first << " " << contact.second << endl;
```

Key-Value Pairs (2)

- Computer Science calls these labeled values "**key-value** pairs"
  - A "**key**" is the label, a "**value**" is the thing we have labeled
  - Accessing the value - through the key
- There are containers optimized for key-value operations
  - Called **associative containers**/**arrays**, **maps**, **dictionaries**, etc.
  - Fast access, insertion, and deletion by **key** – `O(log(N))` or `O(1)`

Key-Value Containers

LIVE DEMO

C++ Associative Containers

Maps, Sets, Ordered & Unordered

Associative Containers vs. Linear Containers

- **Associative containers** are arrays indexed by **keys**
  - A **key** can be anything – integer, string, or any other object
  - Linear containers can only have numeric indexing (array, vector)

Array or `std::vector`

key	0	1	2	3	4
value	8	-3	12	408	33

Associative array

key	value
John Smith	+1-555-8976
Lisa Smith	+1-555-1234
Sam Doe	+1-555-5030

C++ Associative Containers

- Saying just "C++ Associative Container" implies "ordered"
  - `std::map`, `std::set`, `std::multimap`, `std::multiset`
  - Keep elements ordered by key – iterating gives them sorted by key
  - `find()`, `insert()`, and `erase()` are fast – `O(log(N))`
- Ordered associative containers have requirements for the key
  - By default – must support `operator<` (`int`, `double`, `string`, …)

std::map – Initialization & Iteration

Software University

- Represents keys associated with values, ordered by key
  - Two type parameters – key and value `map<K, V>;`

```
map<string, int> cities = {
  pair<string, int> {"Gabrovo", 58950},
  pair<string, int> {"Sofia", 1307376},
  pair<string, int> {"Melnik", 385},
};
```

- Iterating – elements are `pairs`, ordered by `pair::first`

```
for (auto i = cities.begin(); i != cities.end(); i++)
{ cout << i->first << " " << i->second << endl; }
for (pair<string, int> element : cityPopulations)
{ cout << element.first << " " << element.second << endl; }
```

std::map – Access & Search

Software University

- `operator[]` by key, returns direct reference to the value
  - Accesses value, if no such element, creates it:

```
cities["X"]++; // adds {"X", 0}, returns int& (the 0), 0++ gives 1, so {"X", 1}
```

std::map – Access & Search

- Searching – `find()` by key, returns iterator to the pair

```
cout << cities.find("Lom")->second // prints 27294
```

if not found: `cities.find("Z") == cities.end();`

```
auto result = cities.find(searchCityName);
if (result != cities.end())
  cout <<result->first<<" "<<result->second;
else cout << "No information about " << searchCityName << endl;
```

std::map – insert & erase

Software University

- `insert()` adds an element (key-value pair) into the map
  - Position is determined automatically by the map

```
cities.insert(pair<string, int>("Melnik", 385));
```

- `erase()` can remove by key or by iterator

```
cities.erase("Melnik");

// almost the same as:
cities.erase(cities.find("Melnik"));
```

If "Melnik" key is in the map, otherwise there will be a runtime error in the second case

  - Deletion by iterator (if you have it) is a bit faster

Quick Quiz

- What will the following code print?

```
map<int, string> numbers { {2, "two"} };
for (int i = 0; i < numbers.size(); i++) {
  cout << numbers[i] << ",";
}
```

a) **zero,one,two,**

b) **,,two,**

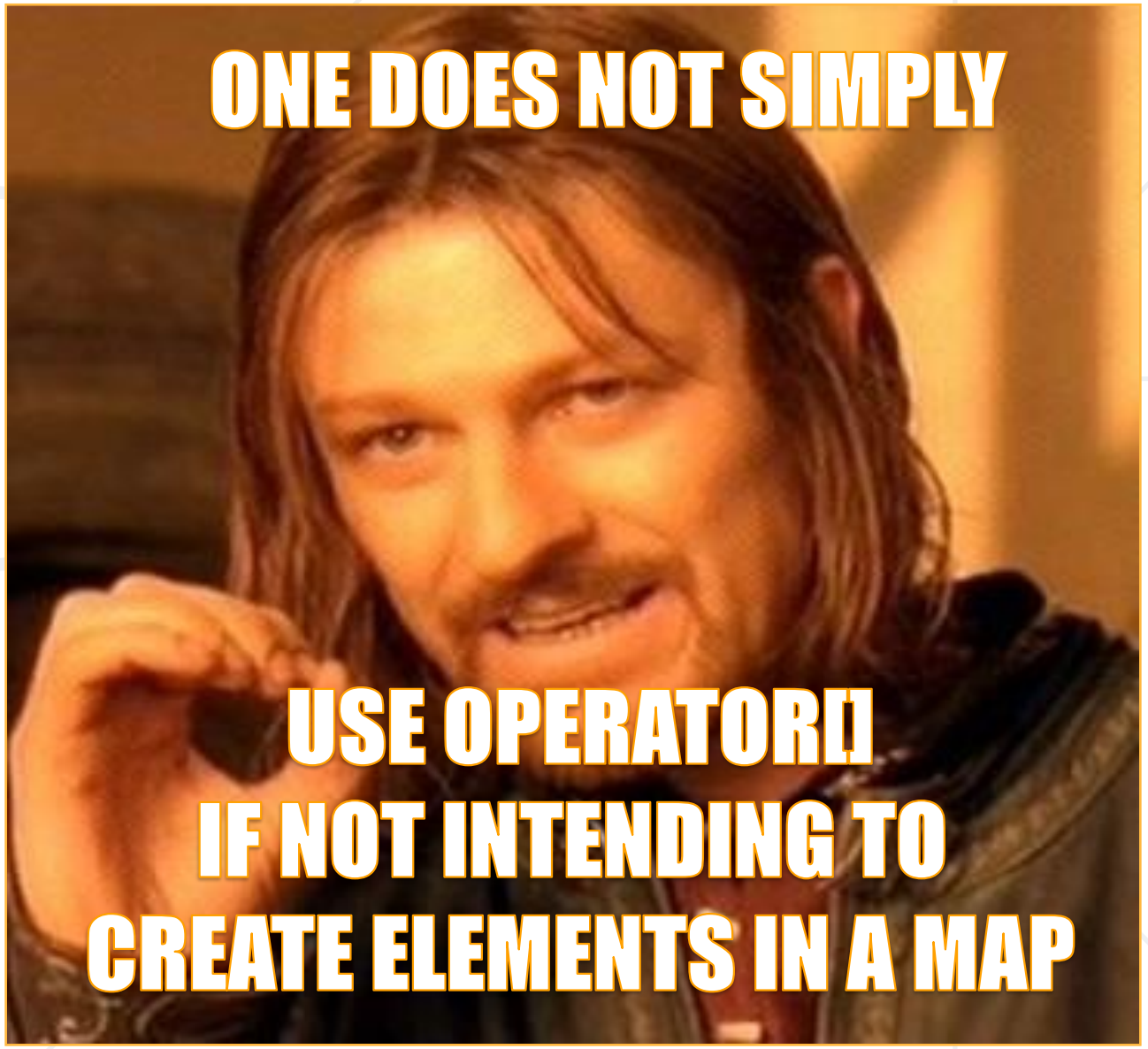
c) **two,**

d) There will be a runtime error

C++ PITFALL: MAP OPERATOR[] INSERTS NEW ELEMENT IF KEY NOT FOUND

If you use access elements with `operator[]`, without checking, in a loop, it is possible that you always add an element and increase the map's `size()`

It is safer to use `find()` if you just want to access existing elements



std::set

Software University

- Similar to map, but only stores keys, without values
  - Single type parameter – `set<K>`, no `operator[]`
  - Useful for removing duplicates

```
set<int> nums { 4, 1, 4, 0, 6, 9, 1, 8, 6, 2, 3, 5, 6, 7 };
for (int n : nums {
  cout << n << " ";
} // 0 1 2 3 4 5 6 7 8 9
```

- Search, insertion, and deletion work the same as for `map`
  - `find()` returns iterator to key, or `end()` if not found
  - `insert()` only inserts if there is no such key

Unordered Associative Containers

- Same names but with `unordered_` prefix
  - E. g. `unordered_map`, `unordered_set`, etc.
- Same `operator[]` (for maps)
  - `find()`, `insert()`, `erase()`
- Faster *(usually)* – operations are `O(1)` instead of `O(log(N))`
- Elements are NOT ordered in any way

std::unordered_map

Software University

- Same operations, methods, initialization, etc. as **map**

```
unordered_map<string, int> cities = {
  pair<string, int>{"Gabrovo", 58950} };
cities.insert(pair<string, int>{"Sofia", 58950));
cities["Melnik"] = 385;
cities.erase("Gabrovo");
```

- Iteration order is not defined (random), same syntax

```
for (auto i = cities.begin(); i != cities.end(); i++) {
  cout << i->first << " " << i->second << endl; }
for (pair<string, int> element : cityPopulations) {
  cout << element.first << " " << element.second << endl; }
```

std::unordered_set

Software University

- Same as **set**, but no order for the keys

```
unordered_set<int> nums {
  4, 1, 4, 0, 6, 9, 1, 8, 6, 2, 3, 5, 6, 7
};
for (int n : nums) { cout << n << " "; }
// prints 0 1 2 3 4 5 6 7 8 9, but the order is unknown
```

  - Single type parameter – **set<K>**, no **operator[]**
- Useful when existence of elements needs to be checked
  - cases when no order information is needed
  - cases where output order will not match "natural" order

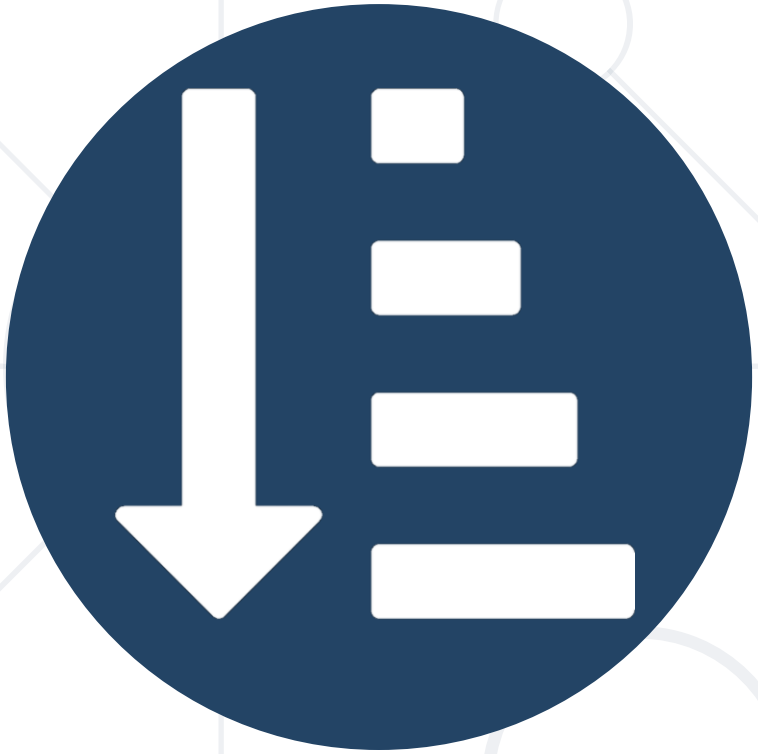
Multiple Values with Same Key

Software University

- A common case is keeping multiple values having the same key
- One approach is a map of vectors (or other linear container)
  - The key points to a `list`/`vector`/… of items
    e. g. `map<string, vector<int> > studentGrades;`
- Another approach (less common) – `multimap`/`multiset`
  - Allow duplicate keys & have operations for multiple equal keys

C++ Associative Containers

LIVE DEMO

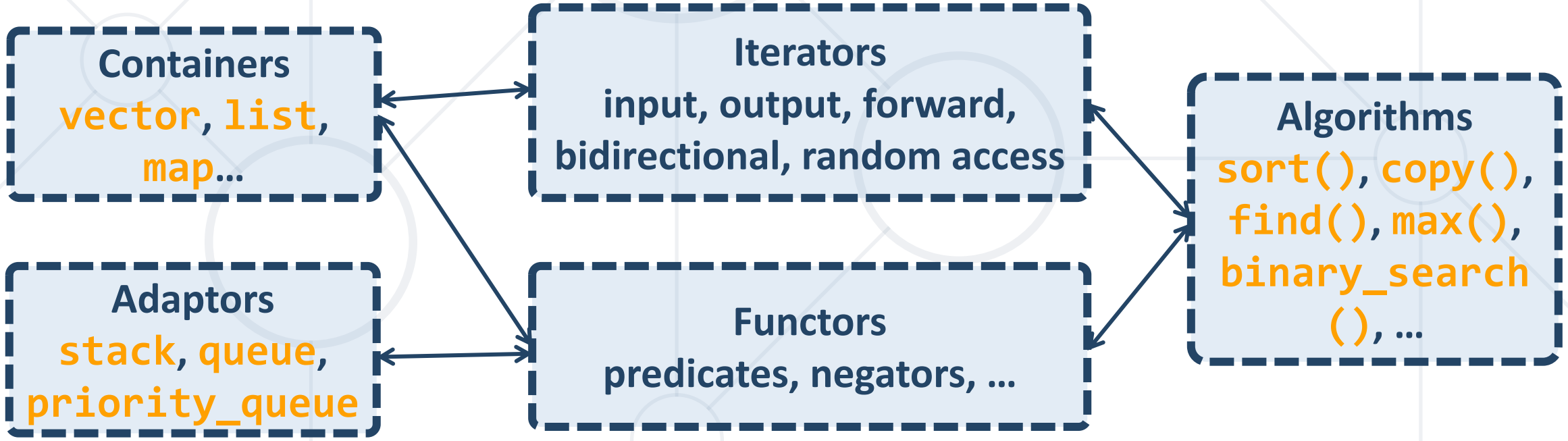


STL Algorithms

Sorting, Searching

STL Algorithms

- STL Provides common Computer Science algorithms
- Iterators define where to act (e.g. from `begin()` to `end()`)
- Functors define how to act (e.g. how to compare values)



Array Iterators

Software University

- Normal arrays can also be used in STL algorithms
  - The array's **name** acts is it's **begin()** iterator
  - Array iterators are random-access iterators
  - So, **array name + array size = array end()** iterator

```
string wordsArray[4] { "whales", "cats", "dogs", "fish" };
auto begin = wordsArray;
auto end = wordsArray + 4;
```

Sorting Array-Like Containers

Software University

- `std::sort(begin, end)`
  - Sorts the range `[begin, end)`, data must have `operator<`
  - Requires random-access iterators (`array`, `vector`, `deque`)

```
vector<int> numsVect { 61, 41, 231, 764, 45 };
sort(numsVect.begin(), numsVect.end());
string wordsArr[4] { "whales", "cats", "dogs", "fish" };
sort(wordsArr, wordsArr + 4);
```

- `std::greater<T>` additional parameter for descending sort

```
sort(numsVect.begin(), numsVect.end(), greater<int>());
```

Sorting Linked-Lists

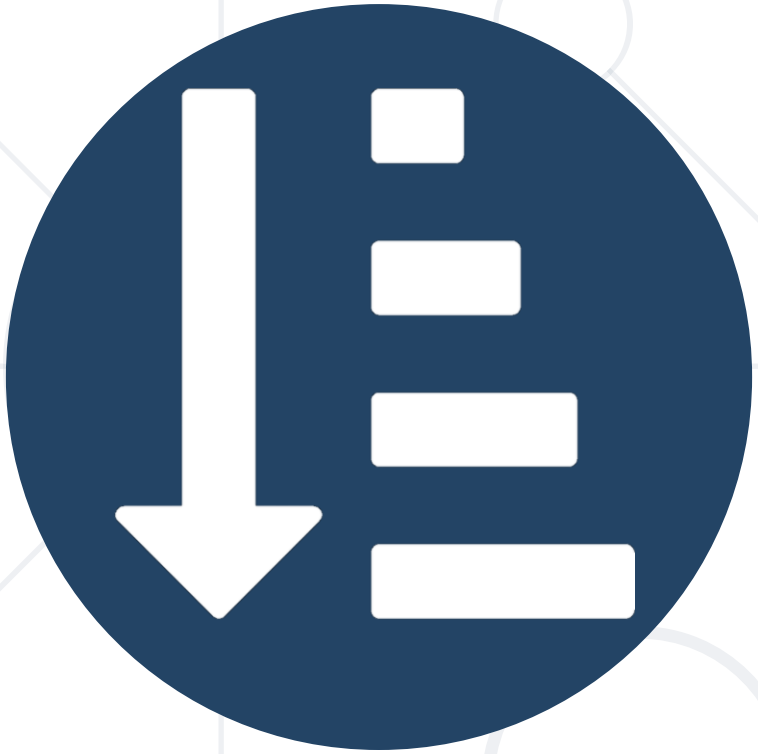
Software University

- `std::list` is not random-access
  - `std::sort` requires random-access iterators
- Lists have their own `sort` version
  - Called directly on a list, i.e. `someList.sort();`

```
list<int> nums { 61, 41, 231, 764, 45 };
nums.sort();
```

  - List sort can also be told to sort from greater to lesser values

```
nums.sort(std::greater<int>());
```



STL Algorithms - Sorting

LIVE DEMO

Searching – find

- `std::find(begin, end, value)`
  - Searches `[begin, end)` for `value`
  - Returns iterator to `value`, or `end` if `value` isn't found
  - If searching a `vector`/array, can subtract `begin()` to get index

```
vector<int> nums { 61, 41, 231, 764, 45 };
auto it = find(nums.begin(), nums.end(), 41);
if (it != nums.end()) {
  cout << "found " << *it << " at " << it - nums.begin() << endl;
} else {
  cout << "not found" << endl; }
```

Searching – min_element & max_element

- `std::min_element(begin, end)`
  - Searches `[begin, end)` for the minimum element
  - Returns iterator if range is not empty, `end` otherwise
  - Data must have `operator<`
  - `std::max_element` does the same for the maximum element

```
vector<int> nums { 61, 41, 231, 764, 45 };
cout << *min_element(nums.begin(), nums.end()) << endl; // 41
cout << *max_element(nums.begin(), nums.end()) << endl; // 764
```

Quick Quiz

- What will the following code print?

a) **764 at 3**

b) **45 at 4**

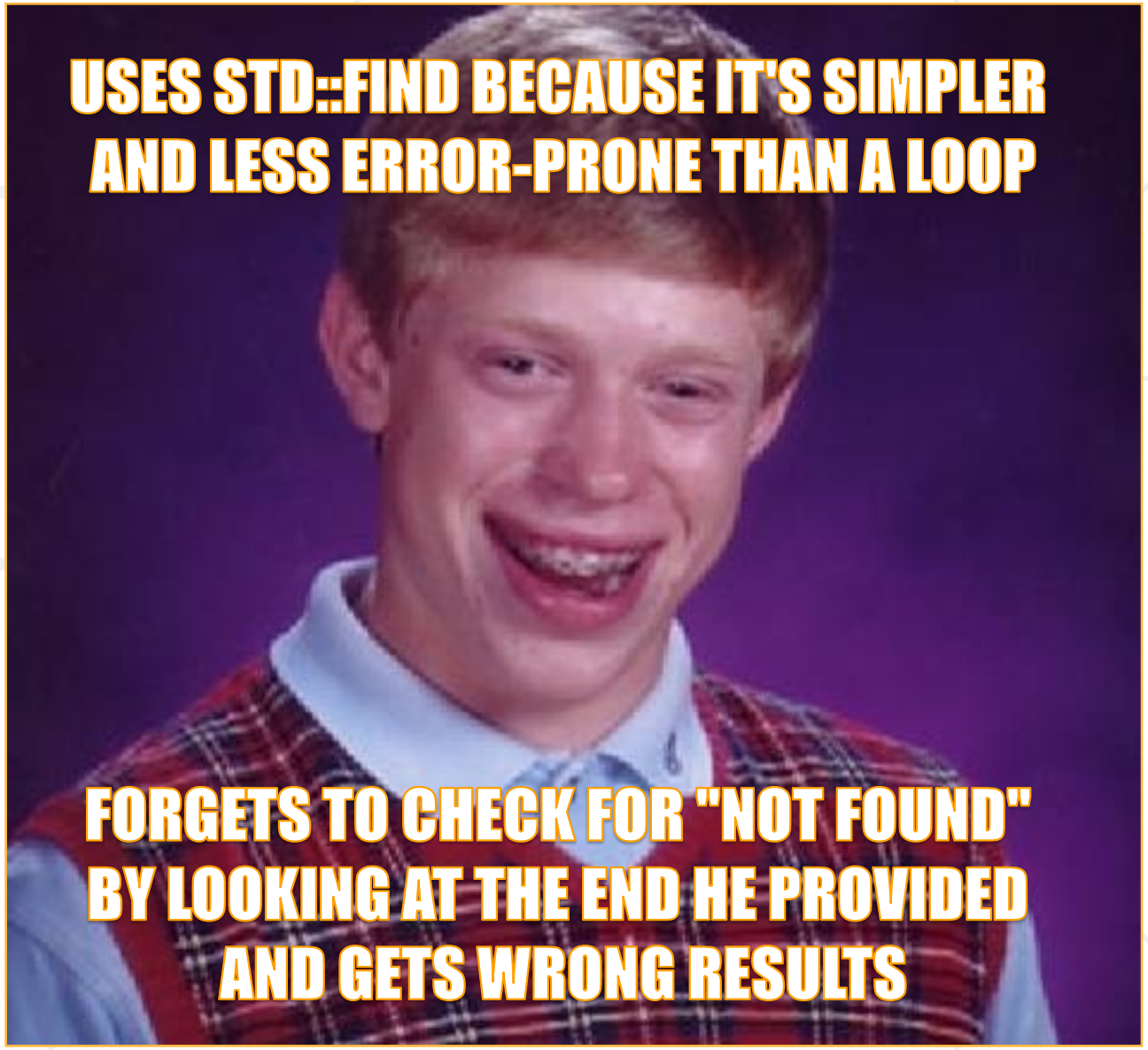
c) **not found**

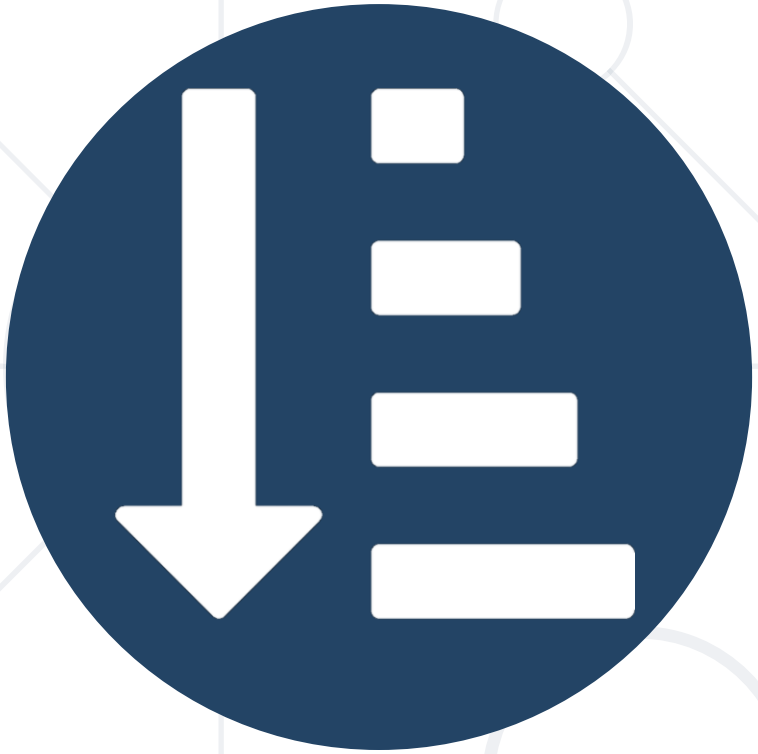
d) There will be a runtime error

```
vector<int> nums { 61, 41, 231, 764, 45 };
auto it = find(nums.begin(), nums.begin() + 3, 45);
if (it != nums.end()) {
  cout << *it << " at "
       << it - nums.begin() << endl;
} else {
  cout << "not found" << endl;
}
```

C++ PITFALL: WRONG "NOT FOUND" CHECK (DIFFERENCE IN END ITERATOR)

Be careful with **`find(begin, end, value)`** – it returns whatever **`end`** you gave it, if it doesn't find the value. If you're looking in part of a vector, it will return an iterator to the end of that part – not to the end of the vector – if it doesn't find **`value`**



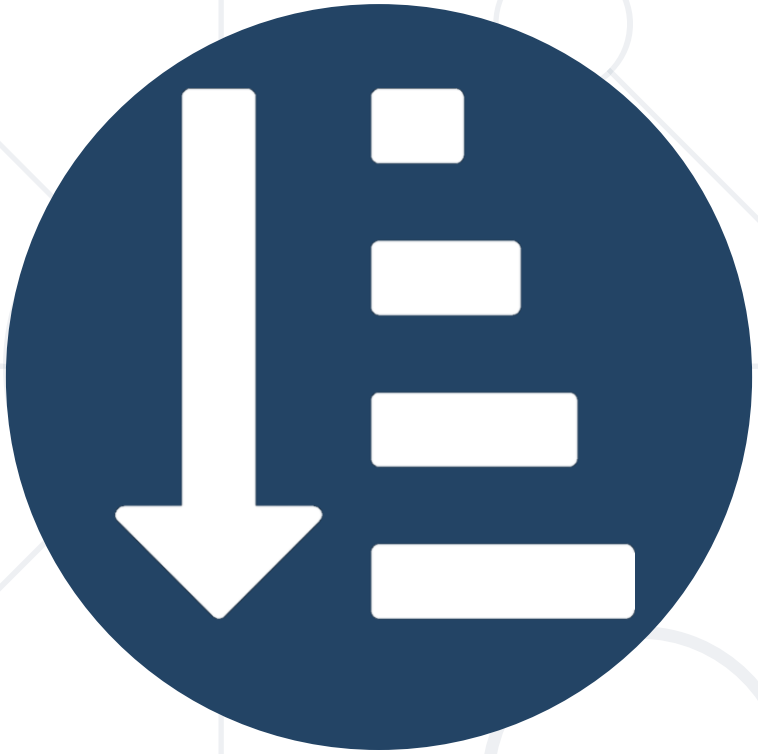


STL Algorithms - Searching

LIVE DEMO

Some Other Algorithms

- `std::lower_bound(begin, end, value)`
  - Requires `[begin, end)` to be sorted
  - Returns where `value` is, if it exists in `[begin, end)`
  - Returns where `value` should be if it doesn't exist
  - Fast – `O(log(N))`, vs. `O(N)` for `find()`
- There are many other algorithms
  - `upper_bound`, `copy`, `replace`, `remove`, `count`, `random_shuffle`, …
  - Look them up at http://en.cppreference.com/w/cpp/algorithm



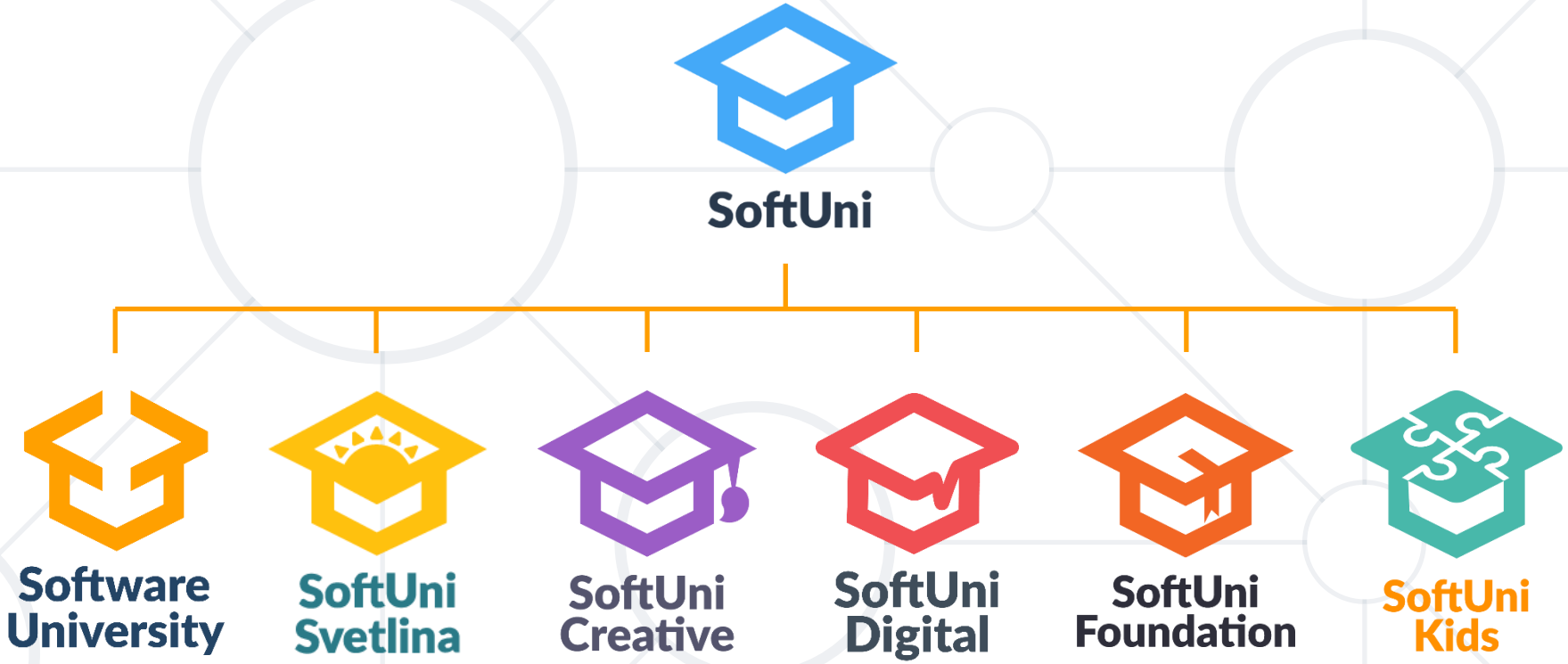
STL Algorithms

LIVE DEMO

Summary

- Associative containers map kesy to values
  - Fast access/lookup/insertion/removal by key
- Maps contain key-value pairs
  - `map`, `unordered_map`, `multimap`, `unordered_multimap`
- Sets only contain keys
  - `set`, `unordered_set`
  - Good for extracting unique elements
- The `<algorithm>` library provides many common algorithms

Questions?

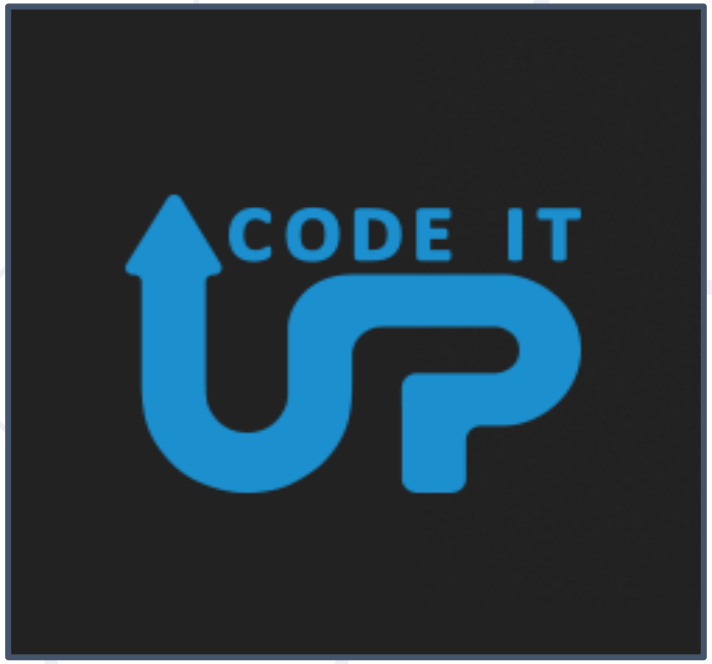


SoftUni Diamond Partners

Software University

Educational Partners

Software University



VIRTUAL RACING SCHOOL

License

- This course (slides, examples, demos, exercises, homework, documents, videos and other assets) is **copyrighted content**
- Unauthorized copy, reproduction or use is illegal
- © SoftUni – https://about.softuni.bg/
- © Software University – https://softuni.bg

Trainings @ Software University (SoftUni)

- Software University – High-Quality Education, Profession and Job for Software Developers
  - softuni.bg, about.softuni.bg
- Software University Foundation
  - softuni.foundation
- Software University @ Facebook
  - facebook.com/SoftwareUniversity
- Software University Forums
  - forum.softuni.bg