18. Dynamic Data Structures I

Dynamic Memory, Addresses and Pointers, Const-Pointer Arrays, Array-based Vectors

Recap: vector<T>

Can be initialised with arbitrary size n

Recap: vector<*T*>

- Can be initialised with arbitrary size n
- Supports various operations:

```
e = v[i];  // Get element
v[i] = e;  // Set element
l = v.size();  // Get size
v.push_front(e);  // Prepend element
v.push_back(e);  // Append element
...
```

Recap: vector<*T*>

- Can be initialised with arbitrary size n
- Supports various operations:

```
e = v[i];  // Get element
v[i] = e;  // Set element
l = v.size();  // Get size
v.push_front(e); // Prepend element
v.push_back(e); // Append element
...
```

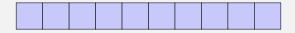
A vector is a dynamic data structure, whose size may change at runtime

Our Own Vector!

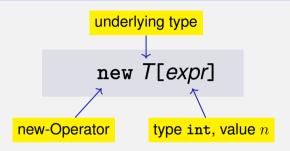
- Today, we'll implement our own vector: vec
- Step 1: vec<int> (today)
- Step 2: vec<*T*> (later, only superficially)

Vectors in Memory

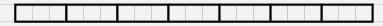
Already known: A vector has a contiguous memory layout

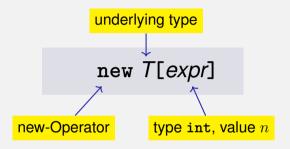


Question: How to *allocate* a chunk of memory of *arbitrary* size during runtime, i.e. *dynamically*?

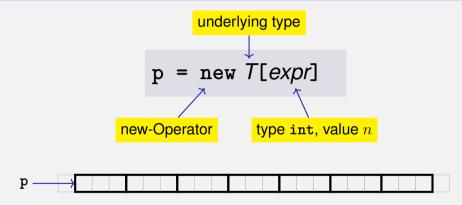


■ Effect: new contiguous chunk of memory *n* elements of type *T* is allocated

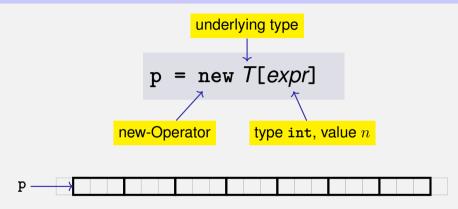




- Effect: new contiguous chunk of memory n elements of type T is allocated
- This chunk of memory is called an *array* (of length n)



Type: A pointer T* (more soon)



- **Type**: A pointer *T** (more soon)
- Value: the starting address of the memory chunk

new T[expr]

 So far: memory (local variables, function arguments) "lives" only inside a function call

new T[expr]

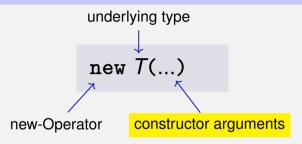
- So far: memory (local variables, function arguments) "lives" only inside a function call
- But now: memory chunk inside vector must not "die" before the vector itself

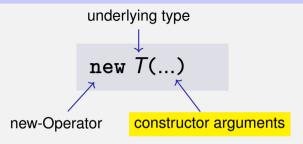
new T[expr]

- So far: memory (local variables, function arguments) "lives" only inside a function call
- But now: memory chunk inside vector must not "die" before the vector itself
- Memory allocated with new is not automatically deallocated (= released)

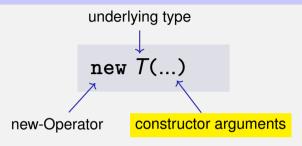
new T[expr]

- So far: memory (local variables, function arguments) "lives" only inside a function call
- But now: memory chunk inside vector must not "die" before the vector itself
- Memory allocated with new is not automatically deallocated (= released)
- Every new must have a matching delete that releases the memory explicitly → in two weeks

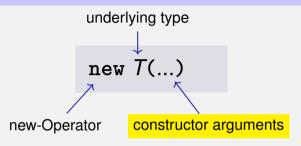




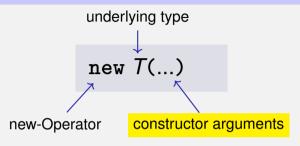
■ Effect: memory for a new object of type *T* is allocated . . .



- Effect: memory for a new object of type T is allocated . . .
- ... and initialized by means of the matching constructor



- Effect: memory for a new object of type T is allocated . . .
- ... and initialized by means of the matching constructor
- Value: address of the new *T* object, Type: Pointer *T**



- Effect: memory for a new object of type T is allocated . . .
- ... and initialized by means of the matching constructor
- Value: address of the new T object, Type: Pointer T*
- Also true here: object "lives" until deleted explicitly (usefulness will become clearer later)

T* Pointer type for base type T

An expression of type T* is called *pointer (to T)*

T* Pointer type for base type T

An expression of type T* is called *pointer (to T)*

```
int* p; // Pointer to an int
std::string* q; // Pointer to a std::string
```

Value of a pointer to T is the address of an object of type T

Value of a pointer to T is the address of an object of type T

```
int* p = ...;
std::cout << p; // e.g. 0x7ffd89d5f7cc</pre>
```

Value of a pointer to T is the address of an object of type T

```
int* p = ...;
std::cout << p; // e.g. 0x7ffd89d5f7cc

int (e.g. 5)

addr
(e.g. 0x7ffd89d5f7cc)</pre>
```

Question: How to obtain an object's address?

■ Directly, when creating a new object via new

Question: How to obtain an object's address?

- Directly, when creating a new object via new
- For existing objects: via the address operator &

&
$$expr \leftarrow \frac{\text{I-value of type } T}{\text{I-value of type } T}$$

Question: How to obtain an object's address?

- Directly, when creating a new object via new
- For existing objects: via the address operator &

&expr
$$\leftarrow$$
 I-value of type T

■ Value of the expression: the address of object (I-value) expr

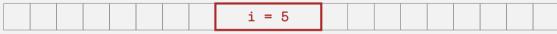
Question: How to obtain an object's address?

- Directly, when creating a new object via new
- For existing objects: via the address operator &

&
$$expr \leftarrow l$$
-value of type T

- Value of the expression: the *address* of object (I-value) *expr*
- **Type** of the expression: A pointer T* (of type T)

```
int i = 5; // i initialised with 5
```



addr

```
int i = 5; // i initialised with 5
int* p = &i; // p initialised with address of i

i = 5

p = &i = addr
addr
```

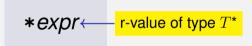
```
int i = 5; // i initialised with 5
int* p = &i; // p initialised with address of i

i = 5

p = &i = addr
addr
```

Next question: How to "follow" a pointer?

Answer: by using the dereference operator *



Answer: by using the dereference operator *

*
$$expr \leftarrow r$$
-value of type T^*

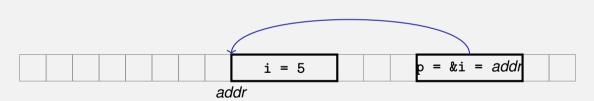
■ Value of the expression: the *value* of the object located at the address denoted by *expr*

Answer: by using the dereference operator *

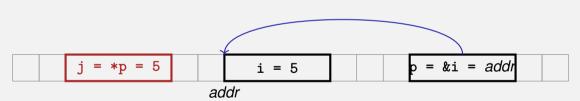
*
$$expr \leftarrow r$$
-value of type T^*

- Value of the expression: the *value* of the object located at the address denoted by *expr*
- **Type** of the expression: T

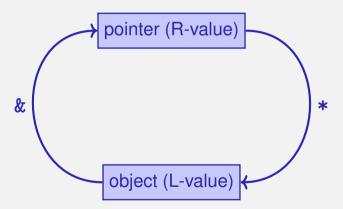
```
int i = 5;
int* p = &i; // p = address of i
```



```
int i = 5;
int* p = &i; // p = address of i
int j = *p; // j = 5
```



Address and Dereference Operator



Pointer Types

A T* must actually point to a T

```
int* p = ...; // p points to an int double* q = p; // but q to a double \rightarrow compiler error!
```

Mnenmonic Trick

The declaration

```
T* p; // p is of the type "pointer to T"
```

Mnenmonic Trick

The declaration

```
T* p; // p is of the type "pointer to T"
```

can be read as

```
T *p; // *p is of type T
```

Null-Pointer

- Special pointer value that signals that no object is pointed to
- represented b the literal nullptr (convertible to T*)

```
int* p = nullptr;
```

- Cannot be dereferenced (runtime error)
- Exists to avoid undefined behaviour

```
int* p; // p could point to anything
int* q = nullptr; // q explicitly points nowhere
```

```
T* p = new T[n]; // p points to first array element

p
size
of a T
```

How to point to rear elements?

```
T* p = new T[n]; // p points to first array element
```

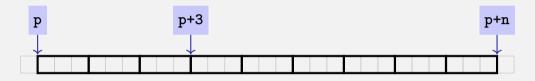
How to point to rear elements? \rightarrow *Pointer arithmetic*:

```
T* p = new T[n]; // p points to first array element
```

How to point to rear elements? \rightarrow *Pointer arithmetic*:

p yields the value of the first array element, *p its value

$$T* p = new T[n]; // p points to first array element$$



How to point to rear elements? \rightarrow *Pointer arithmetic*:

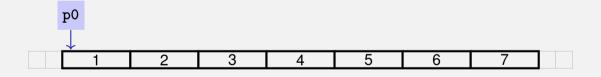
- p yields the *value* of the *first* array element, *p its *value*
- \blacksquare *(p + i) yields the value of the ith array element, for $0 \le i < n$

```
T* p = new T[n]; // p points to first array element
```

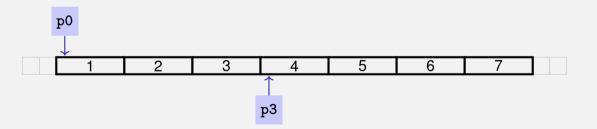
How to point to rear elements? \rightarrow *Pointer arithmetic*:

- p yields the *value* of the *first* array element, *p its *value*
- \blacksquare *(p + i) yields the value of the ith array element, for $0 \le i < n$
- *p is equivalent to *(p + 0)

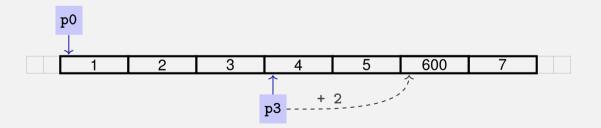
```
int* p0 = new int[7]{1,2,3,4,5,6,7}; // p0 points to 1st element
```



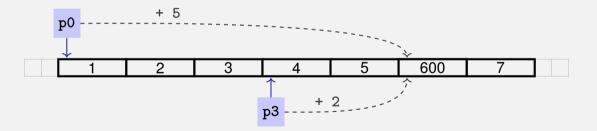
```
int* p0 = new int[7]{1,2,3,4,5,6,7}; // p0 points to 1st element
int* p3 = p0 + 3; // p3 points to 4th element
```



```
int* p0 = new int[7]{1,2,3,4,5,6,7}; // p0 points to 1st element
int* p3 = p0 + 3; // p3 points to 4th element
*(p3 + 2) = 600; // set value of 6th element to 600
std::cout << *(p0 + 5);</pre>
```



```
int* p0 = new int[7]{1,2,3,4,5,6,7}; // p0 points to 1st element
int* p3 = p0 + 3; // p3 points to 4th element
*(p3 + 2) = 600; // set value of 6th element to 600
std::cout << *(p0 + 5); // output 6th element's value (i.e. 600)</pre>
```



```
char* p = new char[3]{'x', 'y', 'z'};
p
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', ';
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x y
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x y
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x y
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x y z
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

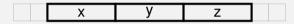
    std::cout << *it << ', '; // x y z
}</pre>
```

```
for (char* it = p;
    it != p + 3;
    ++it) {

    std::cout << *it << ', '; // x y z
}</pre>
```

```
char* p = new char[3]{'x', 'y', 'z'};
```

- The expression *(p + i)
- can also be written as p[i]



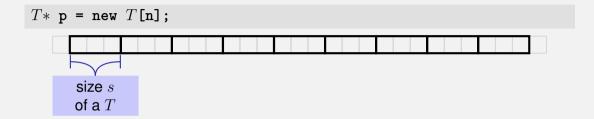
- The expression *(p + i)
- can also be written as p[i]
- E.g. p[1] == *(p + 1) == 'y'

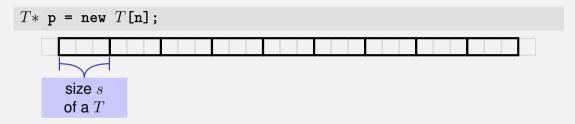
iteration over an array via indices and random access:

```
char* p = new char[3]{'x', 'y', 'z'};

for (int i = 0; i < 3; ++i)
   std::cout << p[i] << ' ';</pre>
```

But: this is less efficient than the previously shown sequential access via pointer iteration





■ Access p[i], i.e. *(p + i), "costs" computation $p + i \cdot s$

```
T* p = new T[n];
size s
of a T
```

- Access p[i], i.e. *(p + i), "costs" computation $p + i \cdot s$
- Iteration via *random access* (p[0], p[1], ...) costs one addition and one multiplication per access

```
T* p = new T[n];
size s
of a T
```

- Access p[i], i.e. *(p + i), "costs" computation $p + i \cdot s$
- Iteration via random access (p[0], p[1], ...) costs one addition and one multiplication per access
- Iteration via *sequentiall access* (++p, ++p, ...) costs only one addition per access

```
T* p = new T[n];
size s
of a T
```

- Access p[i], i.e. *(p + i), "costs" computation $p + i \cdot s$
- Iteration via random access (p[0], p[1], ...) costs one addition and one multiplication per access
- Iteration via sequentiall access (++p, ++p, ...) costs only one addition per access
- Sequential access is thus to be preferred for iterations

Reading a book ... with random access

Random Access

- open book on page 1
- close book
- open book on pages 2-3
- close book
- open book on pages 4-5
- close book
- **....**

Reading a book

... with sequential access

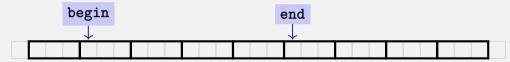
Random Access

- open book on page 1
- close book
- open book on pages 2-3
- close book
- open book on pages 4-5
- close book
-

Sequential Access

- open book on page 1
- turn the page
- **...**

 $C++\mbox{\it covention}$: arrays (or a segment of it) are passed using two pointers

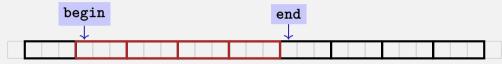


 $\mathrm{C}++\emph{covention}$: arrays (or a segment of it) are passed using two pointers



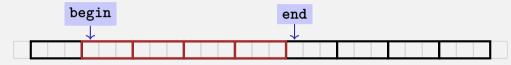
- begin: Pointer to the first element
- end: Pointer past the last element

C++ covention: arrays (or a segment of it) are passed using two pointers



- begin: Pointer to the first element
- end: Pointer past the last element
- [begin, end) Designates the elements of the segment of the array

C++ covention: arrays (or a segment of it) are passed using two pointers



- begin: Pointer to the first element
- end: Pointer past the last element
- [begin, end) Designates the elements of the segment of the array
- [begin, end) is empty if begin == end
- [begin, end) must be a valid range, i.e. a (pot. empty) array segment

Arrays in (mutating) Functions: fill

```
// PRE: [begin, end) is a valid range
// POST: Every element within [begin, end) was set to value
void fill(int* begin, int* end, int value) {
  for (int* p = begin; p != end; ++p)
    *p = value;
}
```

Arrays in (mutating) Functions: fill

```
// PRE: [begin, end) is a valid range
// POST: Every element within [begin, end) was set to value
void fill(int* begin, int* end, int value) {
 for (int* p = begin; p != end; ++p)
   *p = value;
int* p = new int[5];
fill(p, p+5, 1); // Array at p becomes {1, 1, 1, 1, 1}
```

Functions with/without Effect

Pointers can (like references) be used for functions with effect. Example: fill

Functions with/without Effect

- Pointers can (like references) be used for functions with effect. Example: fill
- But many functions don't have an effect, they only read the data
- ⇒ Use of const

Functions with/without Effect

- Pointers can (like references) be used for functions with effect. Example: fill
- But many functions don't have an effect, they only read the data
- ⇒ Use of const
- So far, for example:

```
const int zero = 0;
const int& nil = zero;
```

Positioning of Const

const T is equivalent to T const (and can be written like this):

Positioning of Const

const T is equivalent to T const (and can be written like this):

Both keyword orders are used in praxis

Read the declaration from right to left

```
int const p; p is a constant integer
```

Read the declaration from right to left

```
int const p; p is a constant integer
int const* p; p is a pointer to a constant integer
```

Read the declaration from right to left

```
int const p; p is a constant integer
int const* p; p is a pointer to a constant integer
int* const p; p is a constant pointer to an integer
```

Read the declaration from right to left

624

Non-mutating Functions: print

```
// PRE: [begin, end) is a valid range
// POST: The values in [begin, end) were printed
void print(
   int const* const begin,
   const int* const end) {
 for (int const* p = begin; p != end; ++p)
   std::cout << *p << ' ';
```

Non-mutating Functions: print

```
// PRE: [begin, end) is a valid range
// POST: The values in [begin, end) were printed
void print(
                                  Const pointer to const int
   int const* const begin, 
   const int∗ const end) ←
                                  Likewise (but different keyword order)
 for (int const* p = begin; p != end; ++p)
   std::cout << *p << ' ':
```

Non-mutating Functions: print

```
// PRE: [begin, end) is a valid range
// POST: The values in [begin, end) were printed
void print(
                                   Const pointer to const int
   int const* const begin, 
   const int* const end)
                                   Likewise (but different keyword order)
 for (int const* p = begin; p != end; ++p)
    std::cout << *p <<
                                   Pointer. not const. to const int
```

Arrays are contiguous chunks of memory of statically unknown size

- Arrays are contiguous chunks of memory of statically unknown size
- new T[n] allocates a T-array of size n

- Arrays are contiguous chunks of memory of statically unknown size
- \blacksquare new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element

- Arrays are contiguous chunks of memory of statically unknown size
- lacktriangle new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements

- Arrays are contiguous chunks of memory of statically unknown size
- lacktriangle new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements
- Sequentially iterating over arrays via pointers is more efficient than random access

- Arrays are contiguous chunks of memory of statically unknown size
- \blacksquare new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements
- Sequentially iterating over arrays via pointers is more efficient than random access
- $lacktriangleq new \ T$ allocates memory for (and initialises) a single T-object, and yields a pointer to it

- Arrays are contiguous chunks of memory of statically unknown size
- new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements
- Sequentially iterating over arrays via pointers is more efficient than random access
- $lacktriangleq new \ T$ allocates memory for (and initialises) a single T-object, and yields a pointer to it
- Pointers can point to something (not) const, and they can be (not) const themselves

- Arrays are contiguous chunks of memory of statically unknown size
- new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements
- Sequentially iterating over arrays via pointers is more efficient than random access
- $lacktriangleq new \ T$ allocates memory for (and initialises) a single T-object, and yields a pointer to it
- Pointers can point to something (not) const, and they can be (not) const themselves
- Memory allocated by new is not automatically released (more on this soon)

- Arrays are contiguous chunks of memory of statically unknown size
- new T[n] allocates a T-array of size n
- T* p = new T[n]: pointer p points to the first array element
- Pointer arithmetic enables accessing rear array elements
- Sequentially iterating over arrays via pointers is more efficient than random access
- $lacktriangleq new \ T$ allocates memory for (and initialises) a single T-object, and yields a pointer to it
- Pointers can point to something (not) const, and they can be (not) const themselves
- Memory allocated by new is not automatically released (more on this soon)
- Pointers and references are related, both "link" to objects in memory. See also additional the slides pointers.pdf)

■ Vectors . . . that somehow rings a bell



- Wir implementieren unseren eigenen Vektor: vec
- Schritt 1: vec<int> (heute)
- Schritt 2: vec<T> (später, nur kurz angeschnitten)

Vectors . . . that somehow rings a bell



Now we know how to allocate memory chunks of arbitrary size . . .

- Wir implementieren unseren eigenen Vektor: vec
- Schritt 1: vec<int> (heute)
- Schritt 2: vec<T> (später, nur kurz angeschnitten)

- Vectors . . . that somehow rings a bell
- Now we know how to allocate memory chunks of arbitrary size . . .
- ... we can implement a vector, based on such a chunk of memory

- Wir implementieren unseren eigenen Vektor: vec
- Schritt 1: vec<int> (heute)
- Schritt 2: vec<T> (später, nur kurz angeschnitten)

- Vectors . . . that somehow rings a bell
- Now we know how to allocate memory chunks of arbitrary size . . .
- ... we can implement a vector, based on such a chunk of memory
- avec an array-based vector of int elements

- Wir implementieren unseren eigenen Vektor: vec
- Schritt 1: vec<int> (heute)
- Schritt 2: vec<T> (später, nur kurz angeschnitten)

```
class avec {
 // Private (internal) state:
 int* elements: // Pointer to first element
 unsigned int count; // Number of elements
public: // Public interface:
 avec(unsigned int size); ←
                                     Constructor
 unsigned int size() const;
 int& operator[](int i);
 void print(std::ostream& sink) const;
```

```
class avec {
 // Private (internal) state:
 int* elements: // Pointer to first element
 unsigned int count; // Number of elements
public: // Public interface:
 avec(unsigned int size); // Constructor
 unsigned int size() const; ← Size of vector
 int& operator[](int i);
 void print(std::ostream& sink) const;
```

Array-based Vector avec: Class Signature

```
class avec {
 // Private (internal) state:
 int* elements: // Pointer to first element
 unsigned int count; // Number of elements
public: // Public interface:
 avec(unsigned int size); // Constructor
 unsigned int size() const; // Size of vector
 void print(std::ostream& sink) const;
```

Array-based Vector avec: Class Signature

```
class avec {
 // Private (internal) state:
 int* elements: // Pointer to first element
 unsigned int count; // Number of elements
public: // Public interface:
 avec(unsigned int size); // Constructor
 unsigned int size() const; // Size of vector
 int& operator[](int i); // Access an element
 void print(std::ostream& sink) const; ←
                                     Output elements
```

Array-based Vector avec: Class Signature

```
class avec {
 // Private (internal) state:
 int* elements: // Pointer to first element
 unsigned int count; // Number of elements
public: // Public interface:
 avec(unsigned int size); // Constructor
 unsigned int size() const; // Size of vector
 int& operator[](int i); // Access an element
 void print(std::ostream& sink) const; // Output elems.
```

Constructor avec::avec()

Constructor avec::avec()

Constructor avec::avec()

Side remark: vector is not initialised with a default value

```
avec::avec(unsigned int size): count(size) {
  elements = new int[size];
}
```

elements is a member variable of our avec instance

```
avec::avec(unsigned int size): count(size) {
  elements = new int[size];
}
```

- elements is a member variable of our avec instance
- That instance can be accessed via the *pointer* this

```
avec::avec(unsigned int size): count(size) {
   (*this).elements = new int[size];
}
```

- elements is a member variable of our avec instance
- That instance can be accessed via the pointer this
- elements is a shorthand for (*this).elements

```
avec::avec(unsigned int size): count(size) {
  this->elements = new int[size];
}
```

- elements is a member variable of our avec instance
- That instance can be accessed via the *pointer* this
- elements is a shorthand for (*this).elements
- Equivalent, but shorter: this—>elements

```
avec::avec(unsigned int size): count(size) {
  this->elements = new int[size];
}
```

- elements is a member variable of our avec instance
- That instance can be accessed via the pointer this
- elements is a shorthand for (*this).elements
- Equivalent, but shorter: this—>elements
- Mnemonic trick: "Follow the pointer to the member variable"

Function avec::size()

Function avec::size()

Usage example:

```
avec v = avec(7);
assert(v.size() == 7); // ok
```

Function avec::operator[]

Function avec::operator[]

```
int& avec::operator[](int i) {
  return this->elements[i];
}
```

Element access with index check:

```
int& avec::at(int i) const {
  assert(0 <= i && i < this->count);

return this->elements[i];
}
```

Function avec::operator[]

```
int& avec::operator[](int i) {
  return this->elements[i];
}
```

Usage example:

```
avec v = avec(7);
std::cout << v[6]; // Outputs a "random" value
v[6] = 0;
std::cout << v[6]; // Outputs 0</pre>
```

```
void avec::print(std::ostream& sink) const {
 p != this->elements + this->count;
    (\alpha++
  sink << *p << ' ';
```

```
void avec::print(std::ostream& sink) const {
 for (int* p = this->elements;
      p != this->elements + this->count: 4
      ++p)
                                             Abort iteration if
                                             past last element
   sink << *p << ' ';
```

```
void avec::print(std::ostream& sink) const {
 for (int* p = this->elements;
      p != this->elements + this->count;
                        Advance pointer element-wise
   sink << *p << ' ';
```

```
void avec::print(std::ostream& sink) const {
 for (int* p = this->elements;
     p != this->elements + this->count;
     (\alpha++
   Output current element
```

Finally: overload output operator:

Finally: overload output operator:

Finally: overload output operator:

Observations:

Constant reference to vec, since unchanged

Finally: overload output operator:

Observations:

- Constant reference to vec, since unchanged
- But not to sink: Outputing elements equals change

Finally: overload output operator:

Observations:

- Constant reference to vec, since unchanged
- But not to sink: Outputing elements equals change
- sink is returned to enable output chaining, e.g.
 std::cout << v << '\n'</pre>

Further Functions?

Further Functions?

Commonalities: such operations need to change the vector's size

An allocated block of memory (e.g. new int[3]) cannot be resized later on

An allocated block of memory (e.g. new int[3]) cannot be resized later on

2 1 7

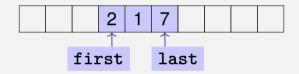
An allocated block of memory (e.g. new int[3]) cannot be resized later on



Possibility:

Allocate more memory than initially necessary

An allocated block of memory (e.g. new int[3]) cannot be resized later on



Possibility:

- Allocate more memory than initially necessary
- Fill from inside out, with pointers to first and last element



■ But eventually, all slots will be in use



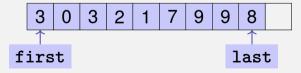
- But eventually, all slots will be in use
- Then unavoidable: Allocate larger memory block and copy data over



Deleting elements requires shifting (by copying) all preceding or following elements



Deleting elements requires shifting (by copying) all preceding or following elements





Deleting elements requires shifting (by copying) all preceding or following elements



Similar: inserting at arbitrary position