

Chapter Six: Arrays and Vectors I

C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

Slides by Evan Gallagher & Nikolay Kirov

Lecture Goals

- To become familiar with using arrays to collect values
- To learn about common algorithms for processing arrays
- To write functions that receive and return arrays



Mail, mail and more mail - how to manage it?

Using Vectors

- When you need to work with a large number of values all together, the vector construct is your best choice.
- By using a *vector* you
 - can conveniently manage collections of data
 - do not worry about the details of how they are stored
 - do not worry about how many are in the vector
 - a vector automatically grows to any desired size

Using Arrays

- Arrays are a lower-level construct
- The *array* is
 - less convenient
 - but sometimes required
 - for efficiency
 - for compatibility with older software



All the mail these days seems alike: *junk!*

In both vectors and arrays, the stored data is of the *same* type

Think of a sequence of data:

32 54 67.5 29 35 80 115 44.5 100 65

(all of the same type, of course) (storable as doubles)

32 54 67.5 29 35 80 115 44.5 100 65

Which is the largest in this set?

(You must look at every single value to decide.)

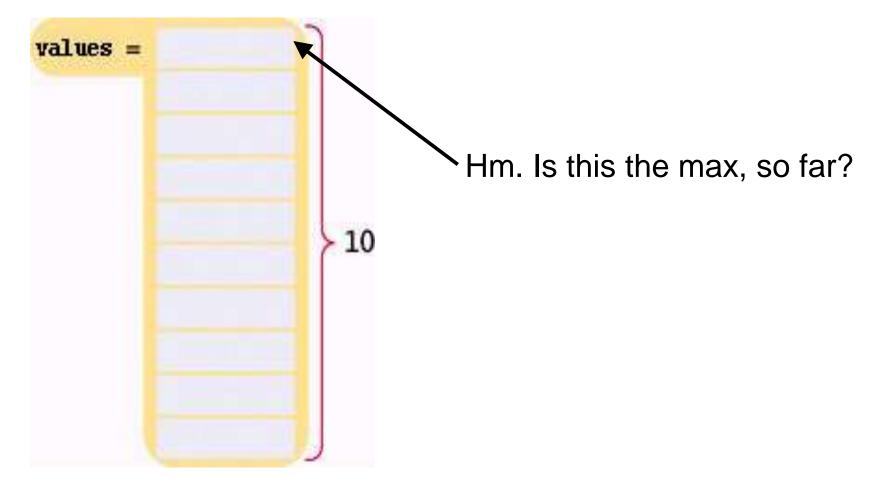
32 54 67.5 29 35 80 115 44.5 100 65

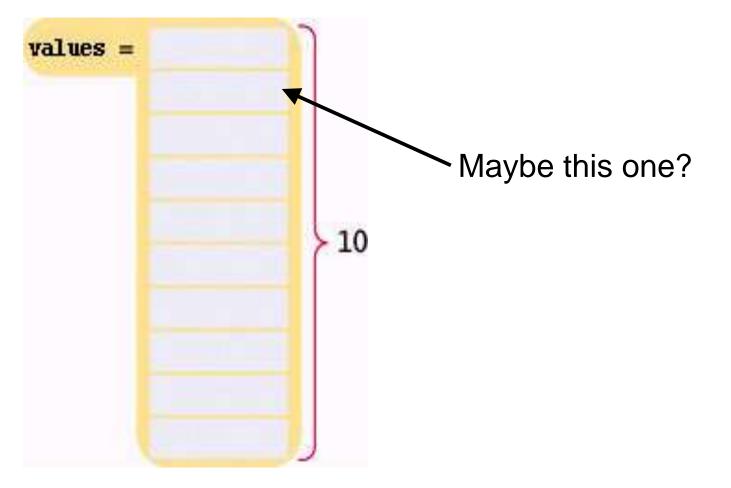
So you would create a variable for each, of course!

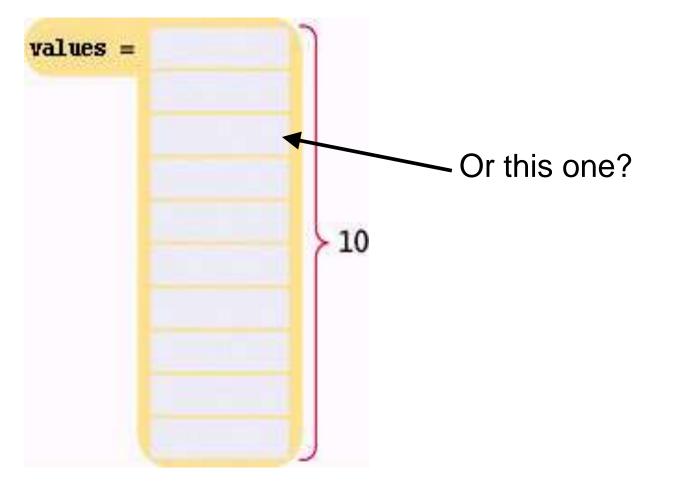
int n1, n2, n3, n4, n5, n6, n7, n8, n9, n10;

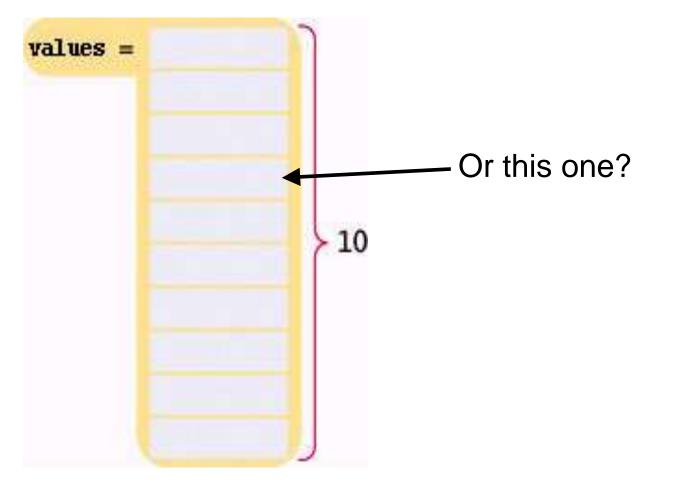
Then what ???

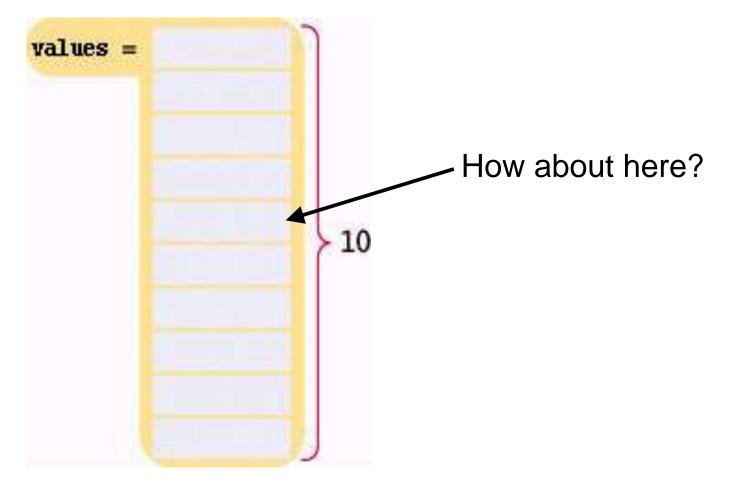
You can easily visit each element in an array or in a vector, checking and updating a variable holding the current maximum.

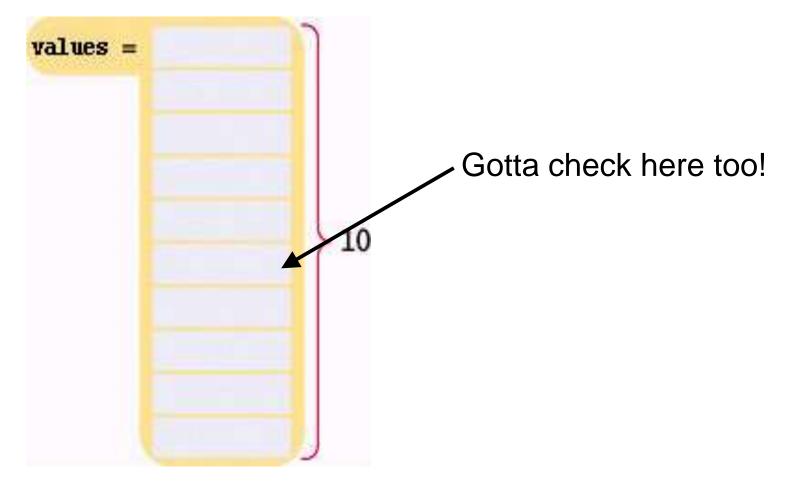


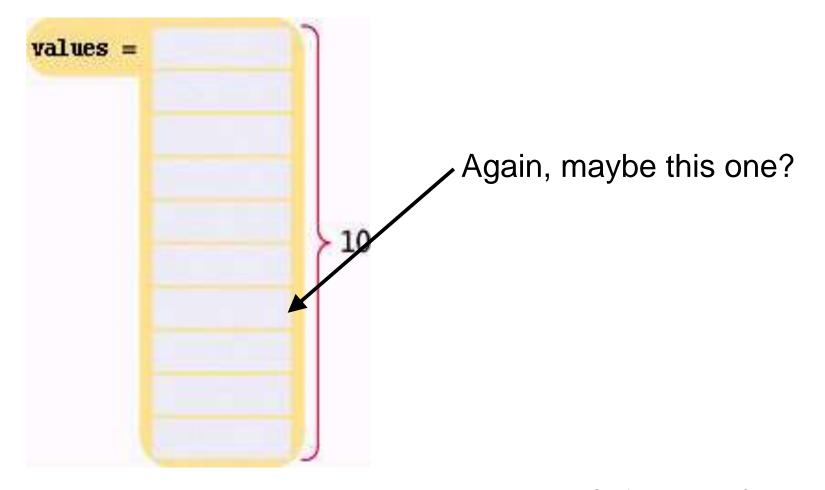


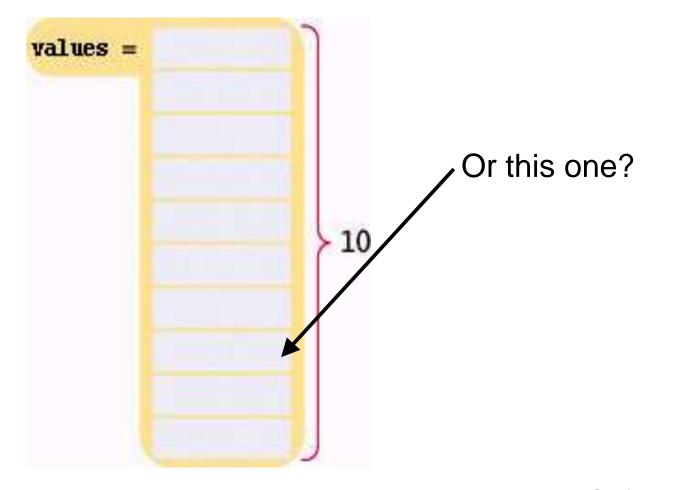


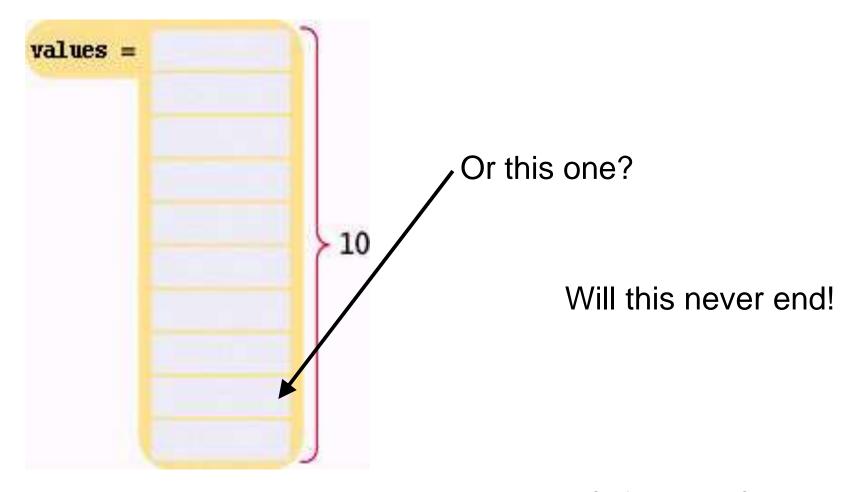


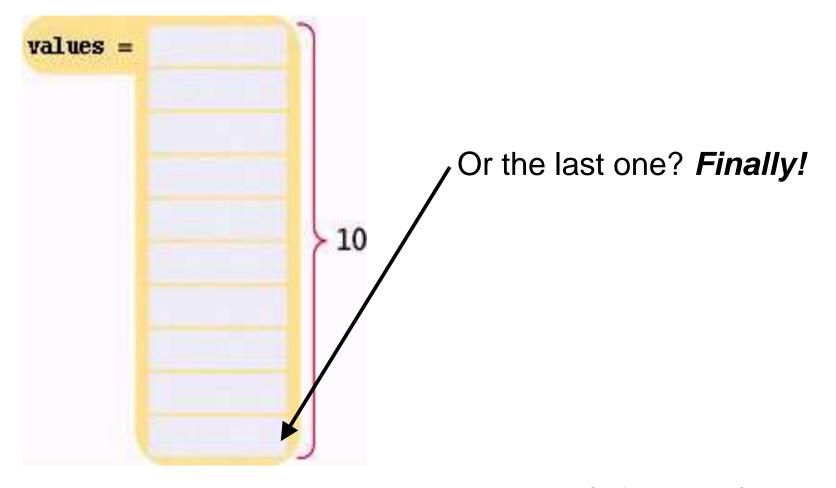












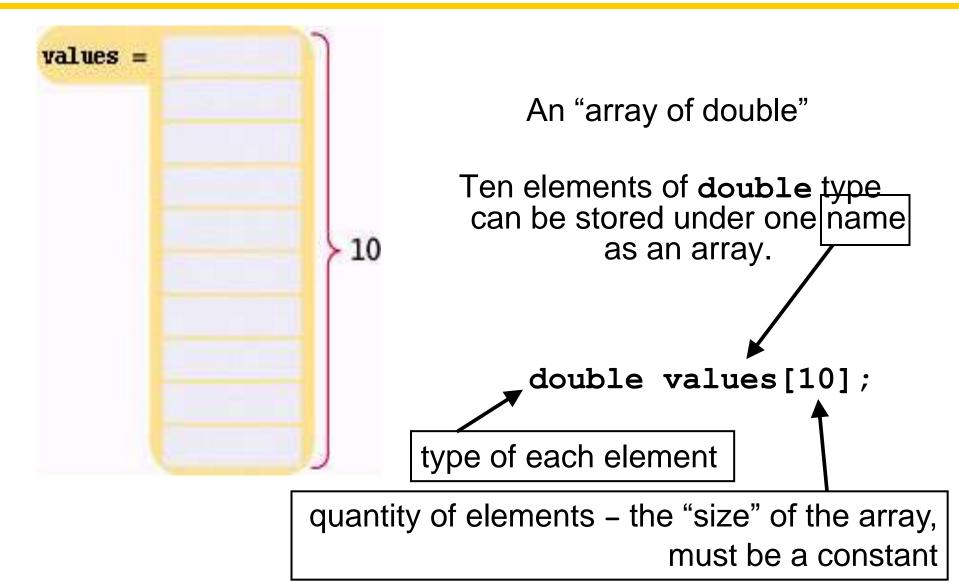
That would have been impossible with ten separate variables!

int n1, n2, n3, n4, n5, n6, n7, n8, n9, n10;

And what if there needed to be another double in the set?

ARGH!

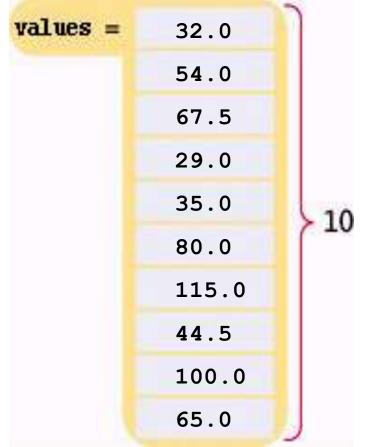
Defining Arrays



Defining Arrays with Initialization

When you define an array, you can specify the initial values:

double values[] = { 32, 54, 67.5, 29, 35, 80, 115, 44.5, 100, 65 };



An array element can be used like any variable.

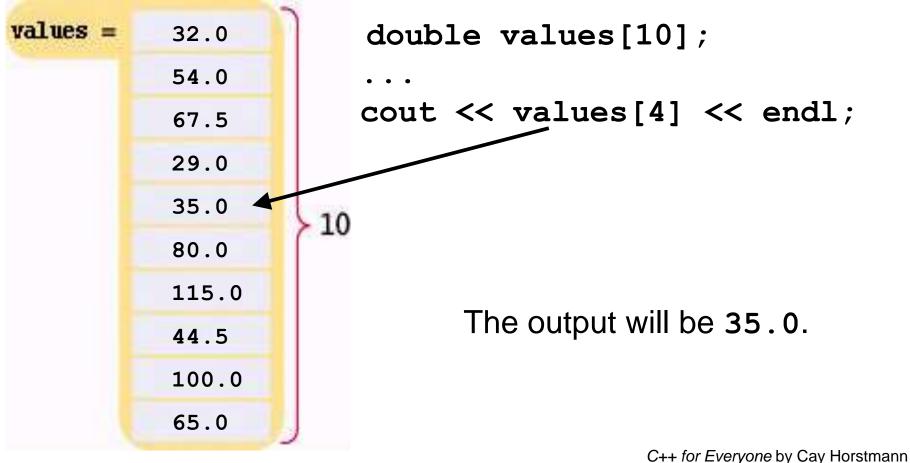
To access an array element, you use the notation:

values[i]

where i is the *index*.

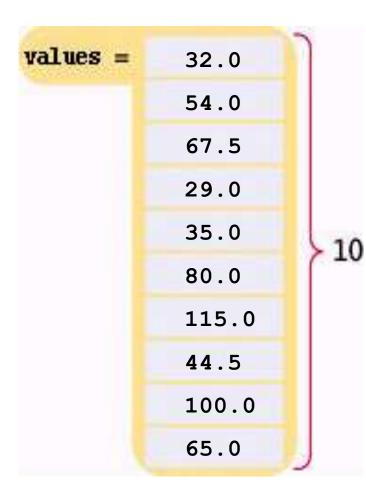


To access the element at index 4 using this notation: **values[4]** 4 is the *index*.



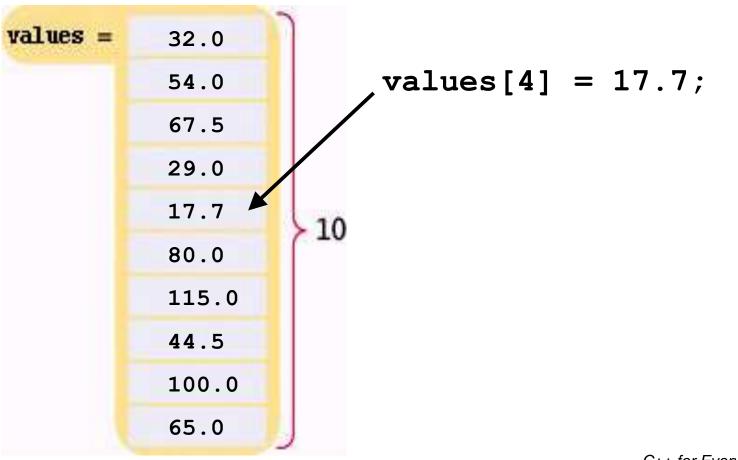
Copyright © 2012 by John Wiley & Sons. All rights reserved

The same notation can be used to change the element.

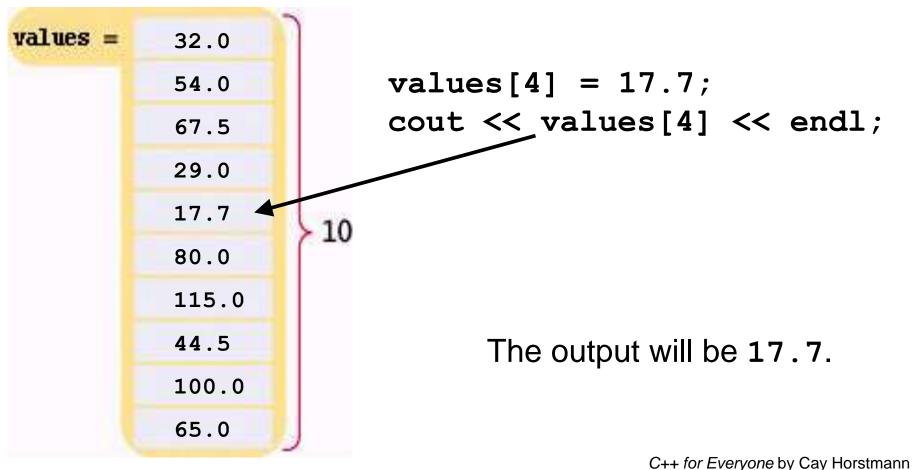


values[4] = 17.7;

The same notation can be used to change the element.

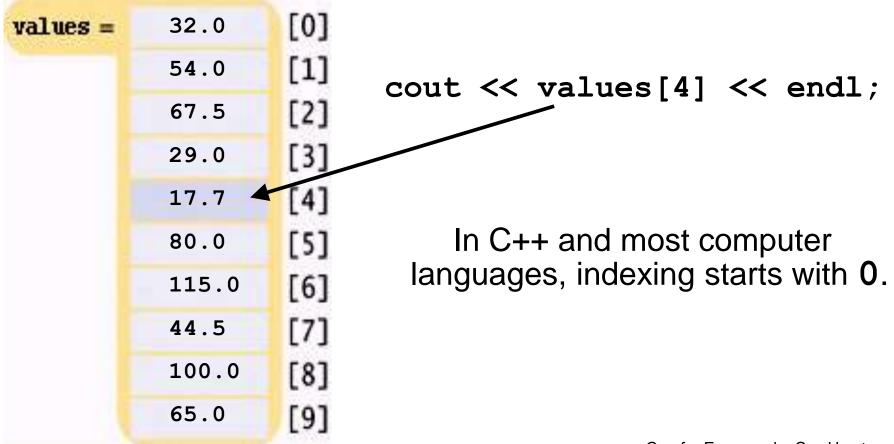


The same notation can be used to change the element.



Copyright © 2012 by John Wiley & Sons. All rights reserved

You might have thought those last two slides were wrong: **values[4]** is getting the data from the "fifth" element.



That is, the legal elements for the **values** array are:

values[0], the first element
values[1], the second element
values[2], the third element
values[3], the fourth element
values[4], the fifth element

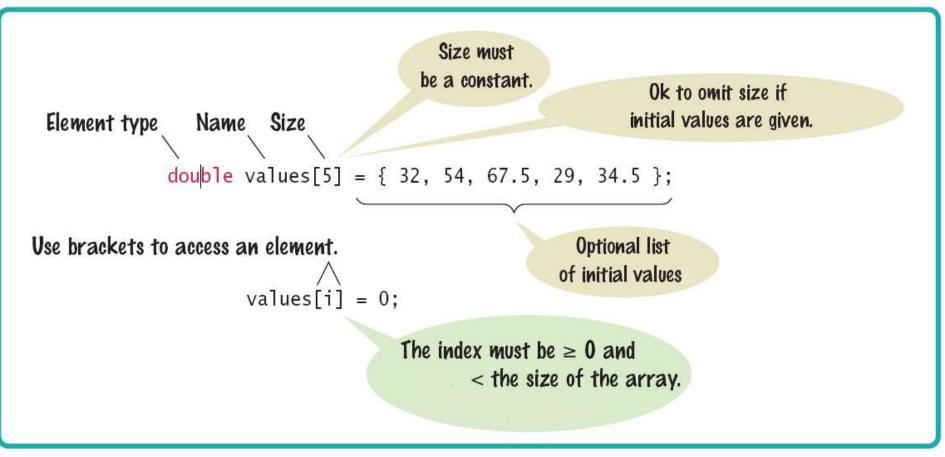
• • •

values[9], the tenth and last legal element
 recall: double values[10];

The index must be >= 0 and <= 9. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 is 10 numbers.

Array Syntax

Defining an Array

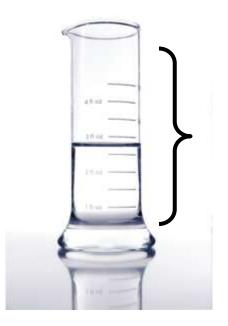


Array Syntax

Table 1 Defining Arrays	
int numbers[10];	An array of ten integers.
<pre>const int SIZE = 10; int numbers[SIZE];</pre>	It is a good idea to use a named constant for the size.
<pre>int size = 10; int numbers[size];</pre>	Caution: In standard C++, the size must be a constant. This array definition will not work with all compilers.
int squares[5] = { 0, 1, 4, 9, 16 };	An array of five integers, with initial values.
int squares[] = { 0, 1, 4, 9, 16 };	You can omit the array size if you supply initial values. The size is set to the number of initial values.
int squares[5] = { 0, 1, 4 };	If you supply fewer initial values than the size, the remaining values are set to 0. This array contains 0, 1, 4, 0, 0.
string names[3];	An array of three strings.
	C++ for Everyone by Cay Horstmann

Partially-Filled Arrays

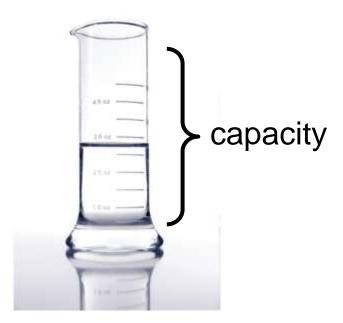
Suppose an array can hold 10 elements:



Does it always? Just look at that beaker. Guess not!

How many elements, at most, can an array hold?

We call this quantity the *capacity*.



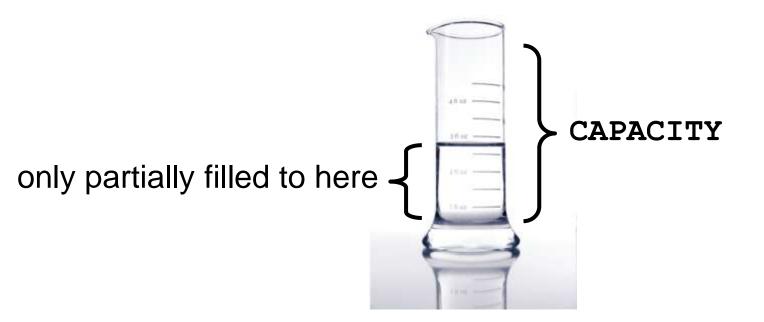
For example, we may decide for a particular problem that there are usually ten or 11 values, but never more than 100.

We would set the capacity with a **const**:

const int CAPACITY = 100; double values[CAPACITY];

Arrays will usually hold less than **CAPACITY** elements.

We call this kind of array a *partially filled array*:

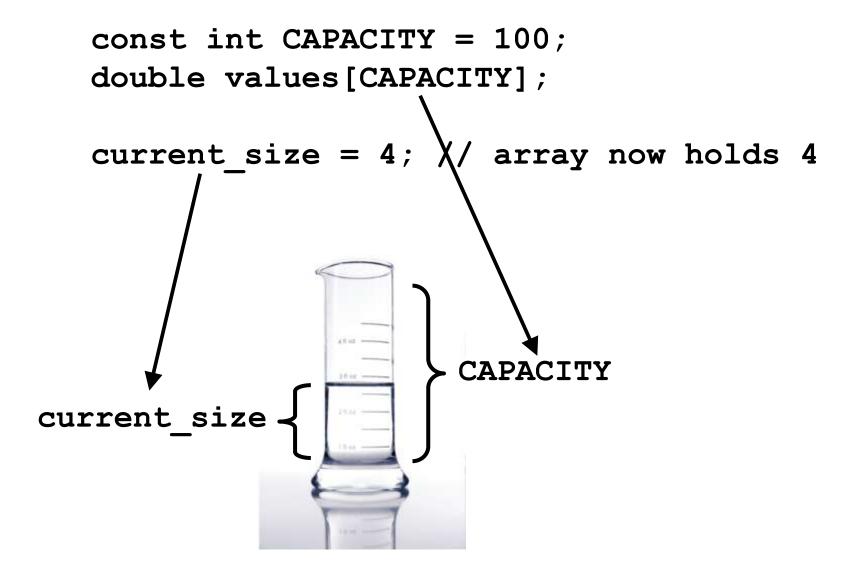


But how many actual elements are there in a partially filled array?

We will use a companion variable to hold that amount: const int CAPACITY = 100; double values[CAPACITY]; int current_size = 0; // array is empty

Suppose we add four elements to the array?

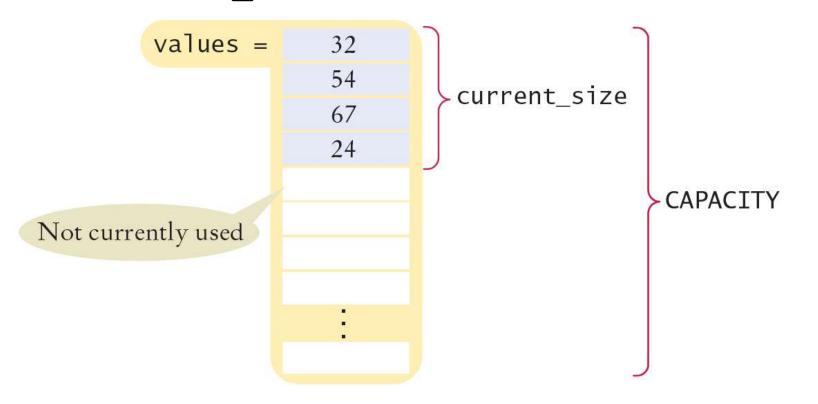
Partially-Filled Arrays – Companion Variable for Size



Partially-Filled Arrays – Companion Variable for Size

const int CAPACITY = 100; double values[CAPACITY];

current size = 4; // array now holds 4



Partially-Filled Arrays – Capacity

The following loop fills an array with user input. Each time the size of the array changes we update this variable: const int CAPAdITY = 100;double values [dAPACITY]; int size = 0;double input; while (cin >> input) ł if (size < CAPACITY) values[size] = x; size++;

Partially-Filled Arrays – Capacity

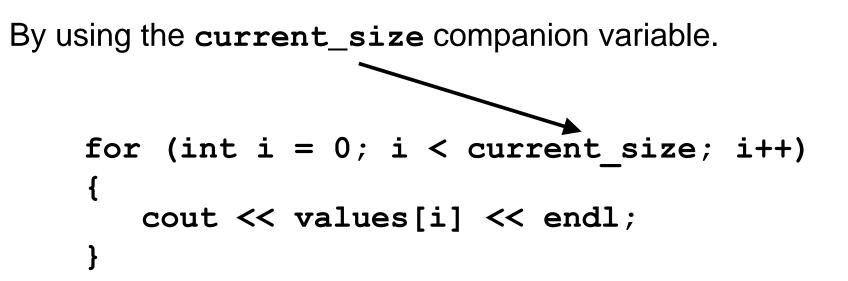
The following loop fills an array with user input. Each time the size of the array changes we update this variable: const int CAPACITY = 100;double values [CAPACITY]; int size = 0;double input; while (cin >> input) ł if (size < CAPACITY) values[size] = x; size+ C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

When the loop ends, the companion variable size has the number of elements in the array.

```
const int CAPACITY = 100;
double values[CAPACITY];
```

```
int size = 0;
double input;
while (cin >> input)
{
    if (size < CAPACITY)
    {
        values[size] = x;
        size++;
    }
}</pre>
```

How would you print the elements in a partially filled array?



}

To visit all elements of an array, use a variable for the index. A **for** loop's variable is best:

```
for (int i = 0; i < CAPACITY; i++)
{
    cout << values[i] << endl;</pre>
```

```
for (int i = 0; i < CAPACITY; i++)
{
    cout << values[i] << endl;
}
When i is 0,</pre>
```

When i is 0, values[i] is values[0], the first element.

```
for (int i = 0; i < CAPACITY; i++)
{
    cout << values[i] << endl;
}
When i is 0, values[i] is values[0], the first element.
When i is 1,</pre>
```

When i is 0, values[i] is values[0], the first element. When i is 1, values[i] is values[1], the second element.

```
for (int i = 0; i < CAPACITY; i++)
{
    cout << values[i] << endl;
}
When i is 0, values[i] is values[0], the first element.
When i is 1, values[i] is values[1], the second element.
When i is 2,</pre>
```

When i is 0, values[i] is values[0], the first element. When i is 1, values[i] is values[1], the second element. When i is 2, values[i] is values[2], the third element.

```
for (int i = 0; i < CAPACITY; i++)
{
    cout << values[i] << endl;
}
When i is 0, values[i] is values[0], the first element.
When i is 1, values[i] is values[1], the second element.
When i is 2, values[i] is values[2], the third element.</pre>
```

When i is 9,

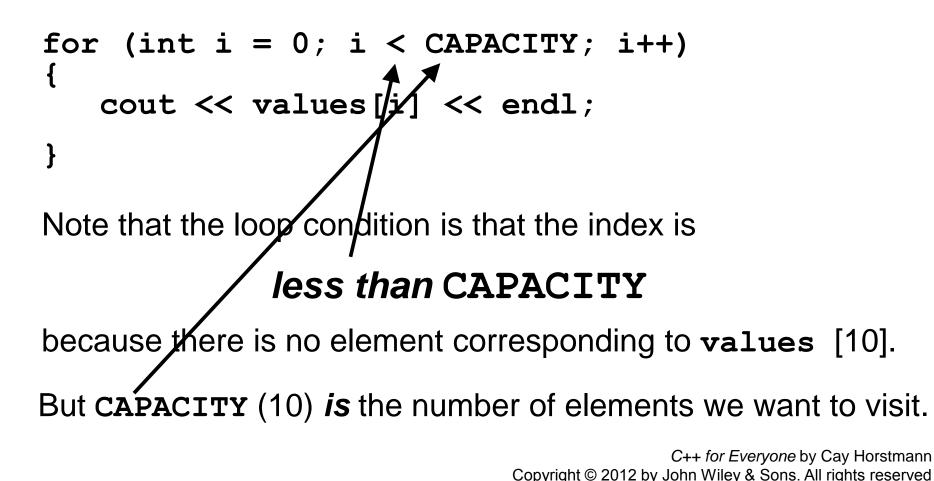
When i is 0, values[i] is values[0], the first element.

When i is 1, values[i] is values[1], the second element.

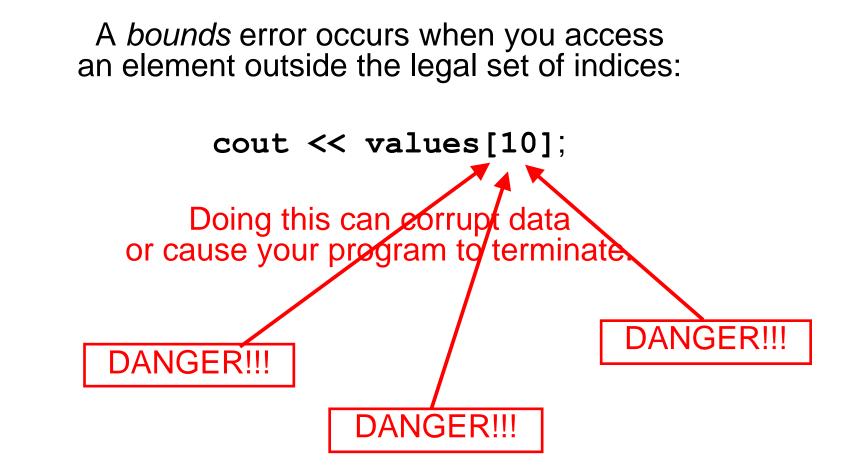
When i is 2, values[i] is values[2], the third element.

. . .

```
When i is 9, values[i] is values[9],
the last legal element.
```



Illegally Accessing an Array Element – Bounds Error



Recall that the type of every element must be the same. That implies that the "meaning" of each stored value is the same.

int scores[NUMBER_OF_SCORES];

Clearly the meaning of each element is a score.

(even if it is a bad score, it's still a score)

But an array could be used *improperly*:

```
double personal_data[3];
personal_data[0] = age;
personal_data[1] = bank_account;
personal_data[2] = shoe size;
```

Clearly these doubles do not have the same meaning!

Use Arrays for Sequences of Related Values

But worse:

personal_data[] = new_shoe_size;

Use Arrays for Sequences of Related Values

But *worse*:

Oh dear!

Which position was I using for the shoe size?

Arrays should be used when the meaning of each element is the same.

Common Array and Vector Algorithms

There are many typical things that are done with sequences of values.

There many common algorithms for processing values stored in both arrays and vectors.

(We will get to vectors a bit later but the algorithms are the same)

> Did someone mention algorithms?

This loop fills an array with zeros:

```
for (int i = 0; i < SIZE Of values; i++)
{
    values[i] = 0;
}</pre>
```

Here, we fill the array with squares (0, 1, 4, 9, 16, ...).

Note that the element with index 0 will contain 0², the element with index 1 will contain 1², and so on.

```
for (int i = 0; i < SİZE Of squares; i++)
{
    squares[i] = i * i;
}</pre>
```

Consider these two arrays:

```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky numbers[5];
```

How can we copy the values from squares to lucky numbers?

Let's try what seems right and easy...

squares = lucky_numbers;

...and wrong!

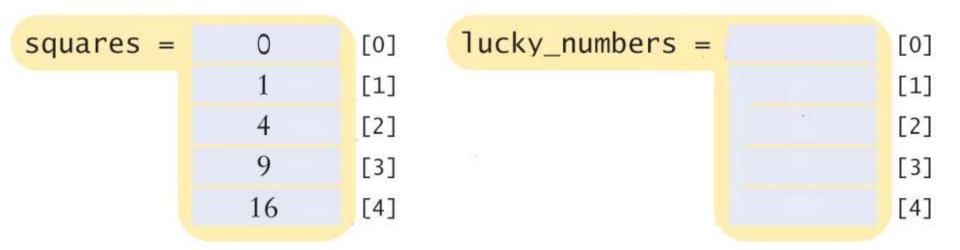
You cannot assign arrays!

You will have to do your own work, son.

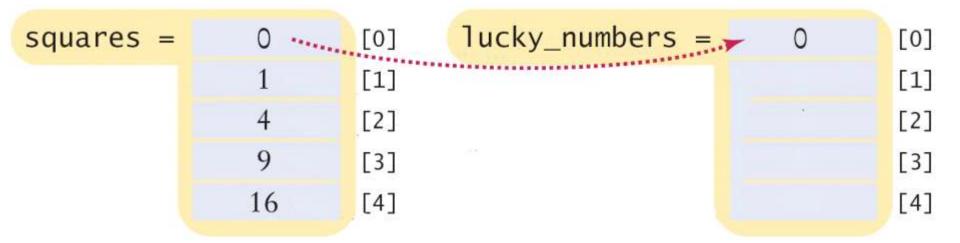
```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}</pre>
```

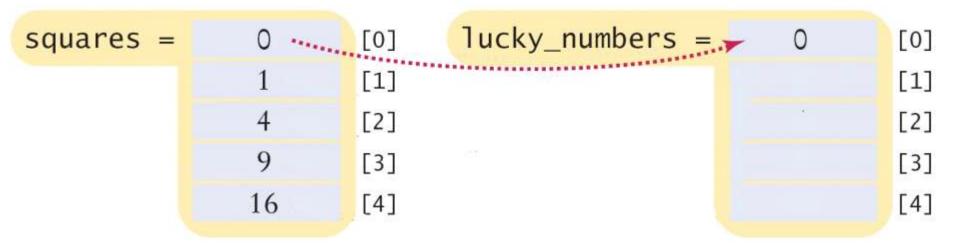
when i is 0



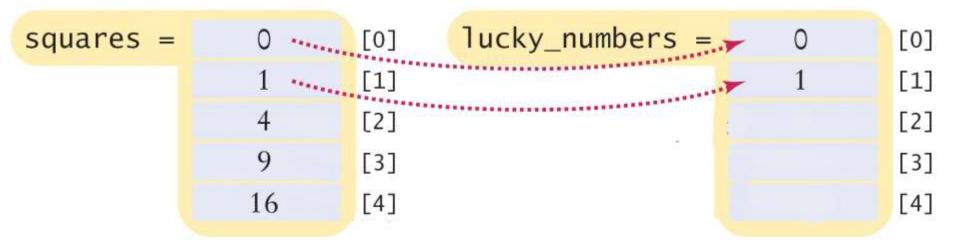
```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
for (int i = 0; i < 5; i++)
{
lucky_numbers[i] = squares[i];
}
when i is 0
```



```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}
  when i is 1
```

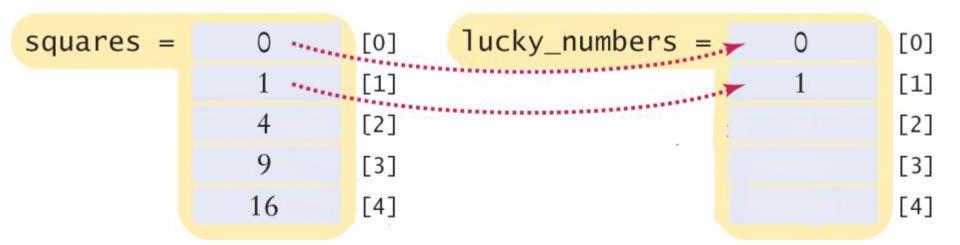


```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
for (int i = 0; i < 5; i++)
{
lucky_numbers[i] = squares[i];
}
when i is 1
```



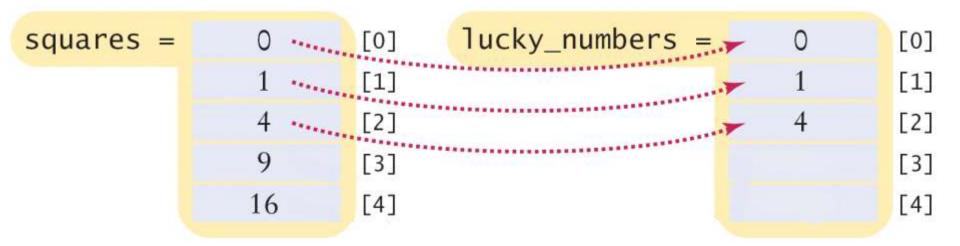
```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}
when i is 2</pre>
```



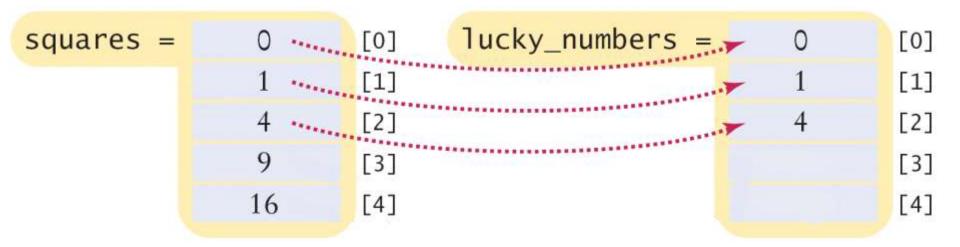
```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}
  when i is 2</pre>
```



```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

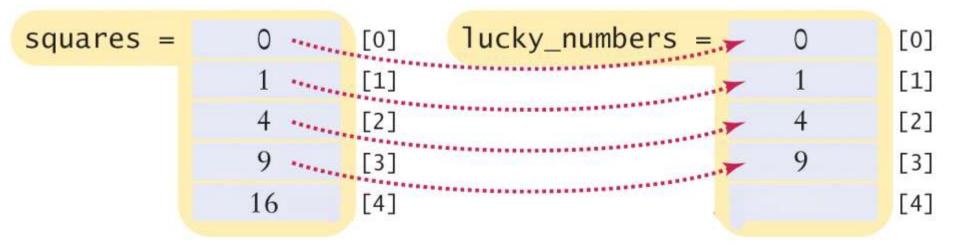
```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}
when i is 3</pre>
```



Common Algorithms – Copying

```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}
  when i is 3</pre>
```

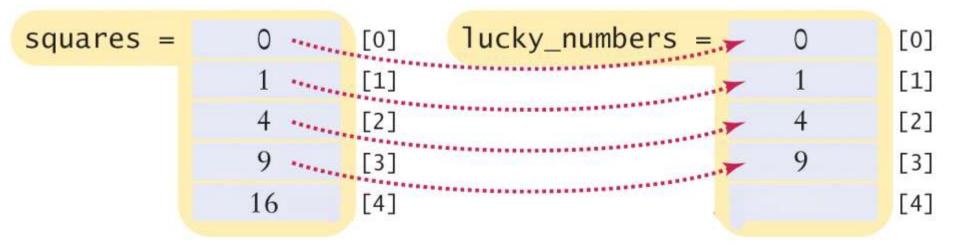


Common Algorithms – Copying

```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}</pre>
```

when i is 4

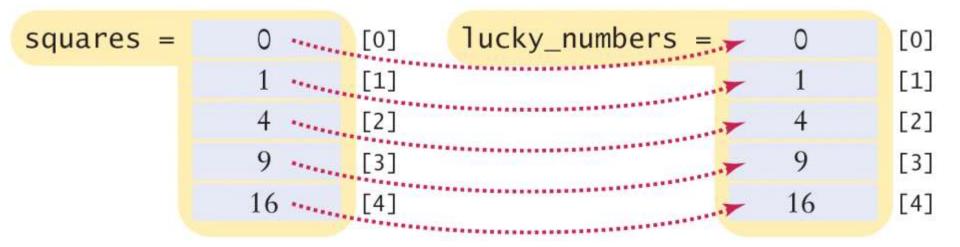


Common Algorithms – Copying

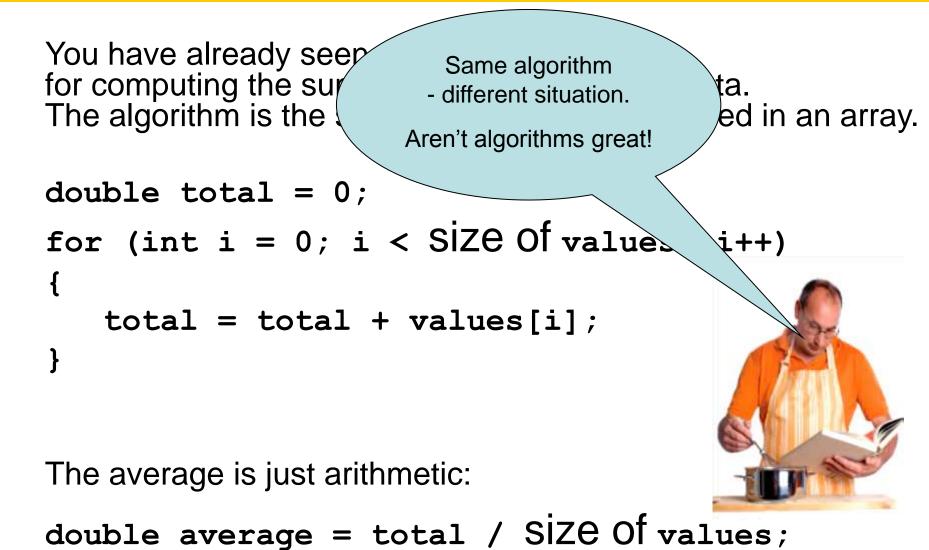
```
int squares[5] = { 0, 1, 4, 9, 16 };
int lucky_numbers[5];
```

```
for (int i = 0; i < 5; i++)
{
  lucky_numbers[i] = squares[i];
}</pre>
```

when i is 4



Common Algorithms – Sum and Average Value

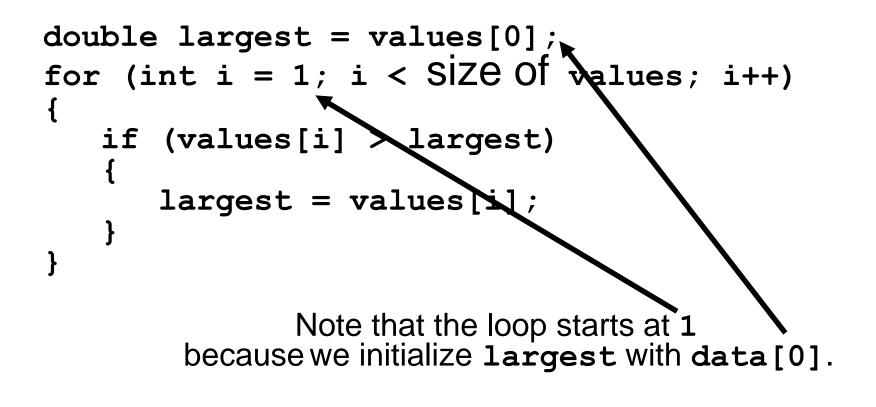


To compute the largest value in a vector, keep a variable that stores the largest element that you have encountered, and update it when you find a larger one.

```
double largest = values[0];
for (int i = 1; i < SiZE Of values; i++)
{
    if (values[i]>largest)
      {
      largest = values[i];
    }
}
```

Common Algorithms – Maximum and Minimum

To compute the largest value in a vector, keep a variable that stores the largest element that you have encountered, and update it when you find a larger one.



Common Algorithms – Maximum and Minimum

For the minimum, we just reverse the comparison.

```
double smallest = values[0];
for (int i = 1; i < SiZE Of values; i++)
{
    if (values[i] < smallest)
        {
        smallest = values[i];
    }
}
```

These algorithms require that the array contain at least one element.

Common Algorithms – Element Separators

When you display the elements of a vector, you usually want to separate them, often with commas or vertical lines, like this:

1 | 4 | 9 | 16 | 25

Note that there is one fewer separator than there are numbers.

To print five elements, you need *four* separators.



Common Algorithms – Element Separators

Print the separator before each element *except the initial one* (with index 0):

Find the position of a certain value, say 100, in an array:

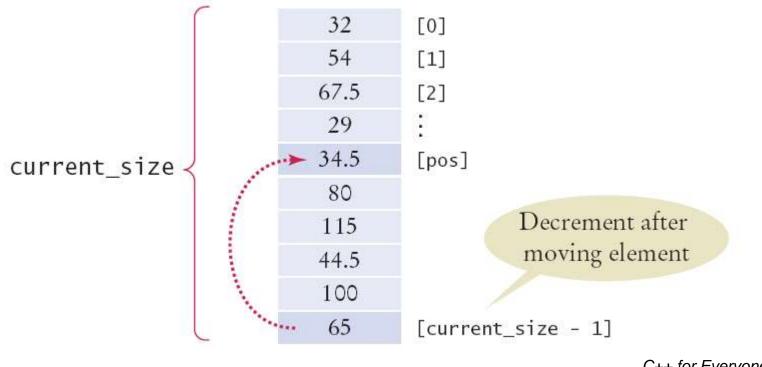
```
int pos = 0;
bool found = false;
while (pos < SIZE Of values \&\& !found)
   if (values[pos] = 100) // looking for 100
      found = true;
                            Don't get these tests
   else
                            in the wrong order!
      pos++;
```

Common Algorithms – Removing an Element, Unordered

Suppose you want to remove the element at index i.

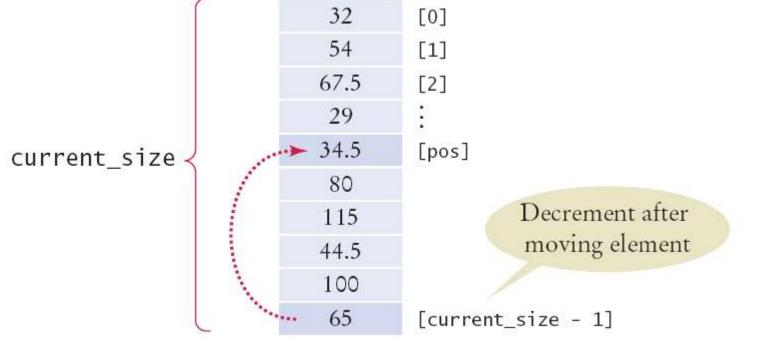
If the elements in the vector are not in any particular order, that task is easy to accomplish.

Simply overwrite the element to be removed with the *last* element of the vector, then shrink the size of the vector by removing the value that was copied.



C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

values[pos] = values[current_size - 1]; current_size--;



C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

Common Algorithms – Removing an Element, Ordered

The situation is more complex if the order of the elements matters.

Then you must move all elements following the element to be removed "down" (to a lower index), and then shrink the size of the vector by removing the last element.

```
for (int i = pos + 1; i < current_size; i++)
{
    values[i - 1] = values[i];
}
current size--;</pre>
```

Common Algorithms – Removing an Element, Ordered

```
for (int i = pos + 1; i < current size; i++)
ł
    values[i - 1] = values[i];
}
current size--;
                              32
                                    [0]
                              54
                                    [1]
                             67.5
                                    [2]
                              29
                                    [pos]
                              80
   current size
                             115
                                            Decrement after
                             44.5
                                            moving elements
                             100
                              65
                              65
                                    [current_size - 1]
```

Common Algorithms – Inserting an Element Unordered

If the order of the elements does not matter, in a partially filled array (which is the only kind you can insert into), you can simply insert a new element at the end.

if (current size < CAPACITY) current size++; values[current size - 1] = new element; 32 [0] 54 [1] 67.5 [2] 29 current_size 34.5 Insert new element here 80 115 Incremented before 44.5 inserting element 100 [current_size - 1]

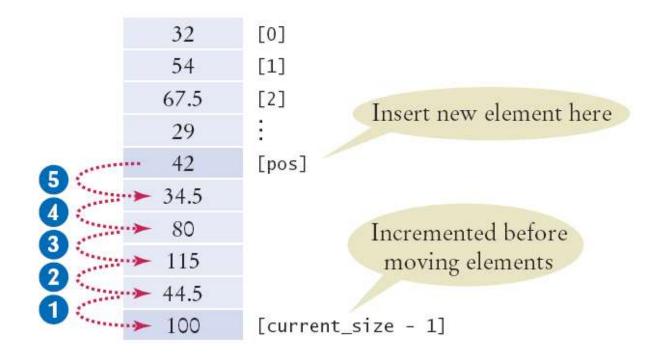
C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

Common Algorithms – Inserting an Element Ordered

If the order of the elements *does* matter, it is a bit harder.

To insert an element at position i, all elements from that location to the end of the vector must be moved "up".

After that, insert the new element at the now vacant position [i].



First, you must make the array one larger by incrementing current_size.

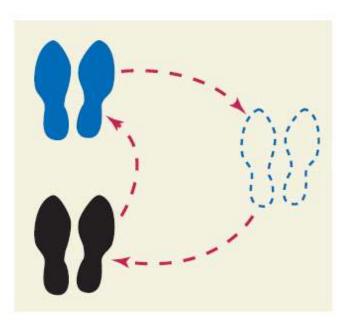
Next, move all elements above the insertion location to a higher index.

Finally, insert the new element in the place you made for it. **Common Algorithms – Inserting an Element Ordered**

```
(current size < CAPACITY)
if
ł
    current size++;
    for (int i = current size - 1; i > pos; i--)
         values[i] = values[i - 1];
    values[pos] = new element;
                       [0]
                32
                54
                       [1]
                67.5
                       [2]
                               Insert new element here
                29
                42
                       [pos]
                34.5
                80
                               Incremented before
                                moving elements
                115
                       [current size - 1]
                                                     C++ for Everyone by Cay Horstmann
                                         Copyright © 2012 by John Wiley & Sons. All rights reserved
```

Swapping two elements in an array is an important part of sorting an array.

To do a swap of two things, you need *three* things!



Suppose we need to swap the values at positions i and j in the array. Will this work?

```
values[i] = values[j];
values[j] = values[i];
```

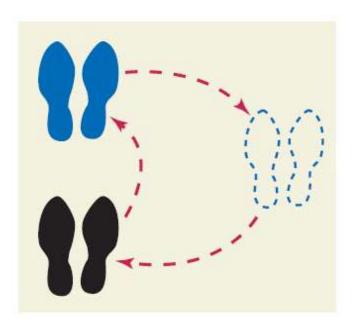
Look closely!

In the first line you lost – forever! – the value at i, replacing it with the value at j.

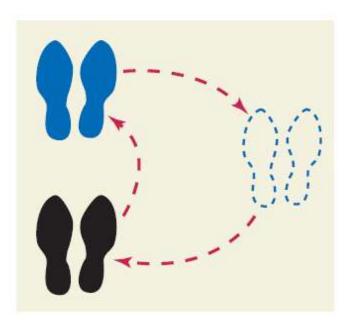
Then what? Put' j's value back in j in the second line?

ARGHHH!

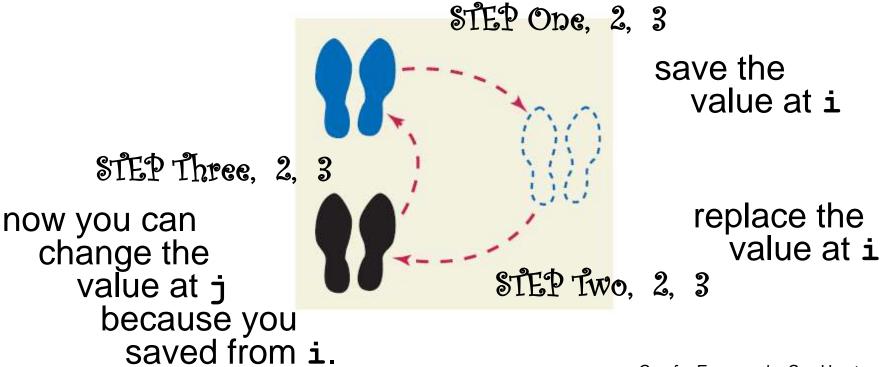
You need a *third* dance partner!



Let's Waltz! 1, 2, 3, go, 2, 3,



```
double temp = values[i];
values[i] = values[j];
values[j] = temp;
```



Common Algorithms – Reading Input

If the know how many input values the user will supply, you can store them directly into the array:

```
double values[NUMBER_OF_INPUTS];
for (i = 0; i < NUMBER_OF_INPUTS; i++)
{
    cin >> values[i];
}
```

Common Algorithms – Reading Input

When there will be an arbitrary number of inputs, things get more complicated. But not hopeless.

Add values to the end of the array until all inputs have been made. Again, the companion variable will have the number of inputs.

```
double values[CAPACITY];
int current size = 0;
double input;
while (cin >> input)
   if (current size < CAPACITY)
   ł
      values[current size] = input;
      current size++;
```

Unfortunately it's even more complicated:

Once the array is full, we allow the user to keep entering!

Because we can't change the size of an array after it has been created, we'll just have to give up for now.

Now back to where we started:

How do we determine the largest in a set of data?

```
#include <iostream>
```

```
using namespace std;
```

```
int main()
```

```
{
```

```
const int CAPACITY = 1000;
double values[CAPACITY];
int current_size = 0;
```

```
cout << "Please enter values, Q to quit:" << endl;
double input;
while (cin >> input)
{
    if (current_size < CAPACITY)
      {
      values[current_size] = input;
      current_size++;
    }
}</pre>
```

```
double largest = values[0];
for (int i = 1; i < current_size; i++)
{
    if (values[i] > largest)
      {
        largest = values[i];
      }
}
```

```
for (int i = 0; i < current_size; i++)
{
    cout << values[i];
    if (values[i] == largest)
    {
        cout << " <== largest value";
    }
    cout << endl;
}</pre>
```

return 0;

}

Getting data into order is something that is often needed.

An alphabetical listing.

A list of grades in descending order.

In C++, you call the sort function to do your sorting for you. But the syntax is new to you:

Recall our **values** array with the companion variable **current_size**.

sort(values, values + current_size);

To sort the elements into ascending numerical order, you call the **sort** algorithm:

You will need to write:

#include <algorithm>

in order to use the **sort** function.

sort(values, values + current_size);

Notice also that you must tell the **sort** function where to begin: **values**, (which is the start of the array) and where to end: **values + current_size**, (which is one *after* the last element in the array).



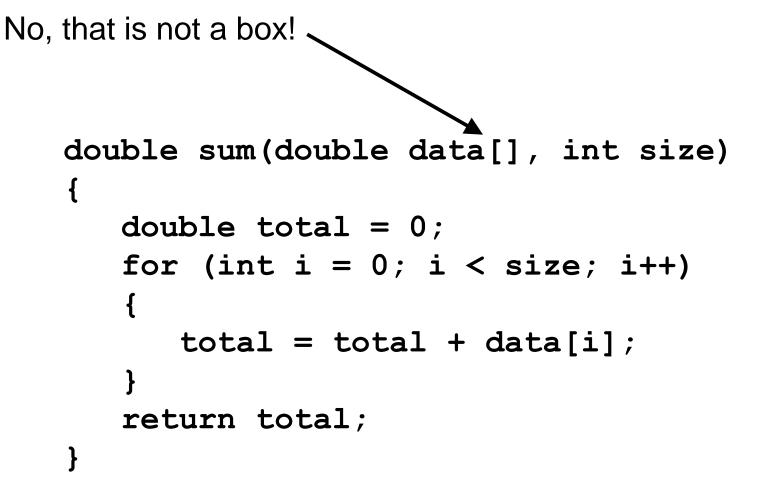
Recall that when we work with arrays we use a companion variable.

The same concept applies when using arrays as parameters:

You must pass the size to the function so it will know how many elements to work with.

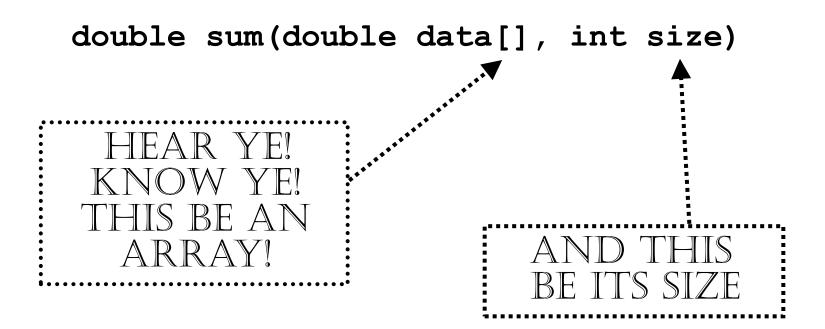
Here is the **sum** function with an array parameter: Notice that to pass one array, it takes two parameters.

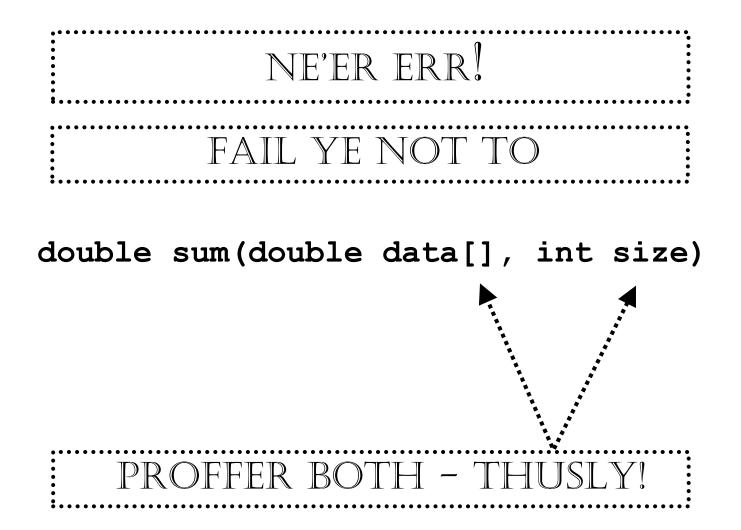
```
double sum(double data[], int size)
{
    double total = 0;
    for (int i = 0; i < size; i++)
    {
        total = total + data[i];
    }
    return total;
}</pre>
```



It is an empty pair of square brackets.

You use an empty pair of square brackets after the parameter variable's name to indicate you are passing an array.



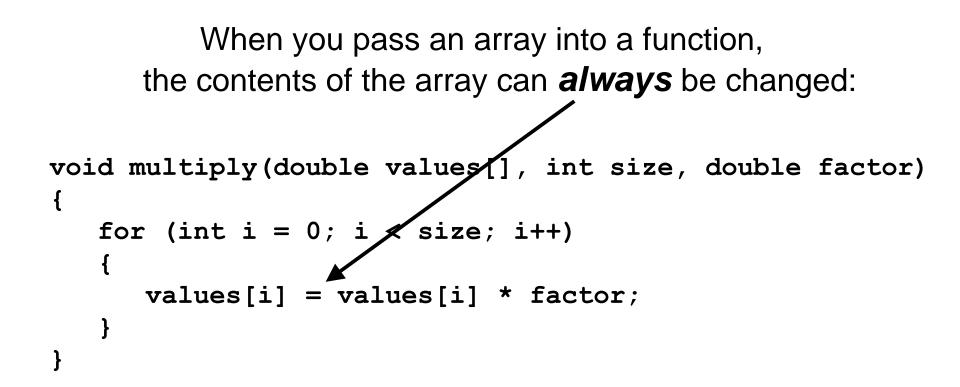


When you call the function, supply both the name of the array and the size: double NUMBER_OF_SCORES = 10; double scores[NUMBER_OF_SCORES] = { 32, 54, 67.5, 29, 34.5, 80, 115, 44.5, 100, 65 }; double total score = sum(scores, NUMBER OF SCORES);

You can also pass a smaller size to the function:

```
double partial_score = sum(scores, 5);
```

This will sum over only the first five **doubles** in the array.



```
And writing an ampersand is always an error.
void multiply1(double& values[], int size, double factor)
   for (int i = 0; i < size; i + +)
      values[i] = values[i] * factor;
}
void multiply2(double values[]&, int size, double factor)
   for (int i = 0; i < size; i++)
      values[i] = values[i] * factor;
```

And writing an ampersand is *always* an error:

```
void multiply1(double values[], int size, double factor)
   for (int i = 0; i < size; i++)
      values[i] = values[i] * factor;
}
void multiply2(double values[] int size, double factor)
   for (int i = 0; i < size; i++)
      values[i] = values[i] * factor;
```

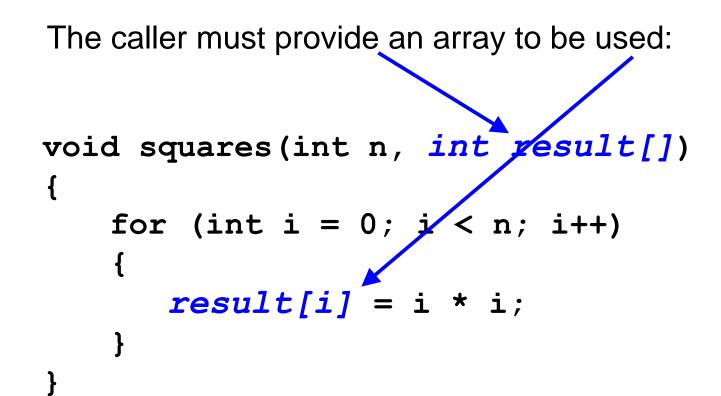
You can pass an array into a function

but

you cannot return an array.

If you cannot return an array, how can the caller get the data?

```
??? squares(int n)
{
    int result[]
    for (int i = 0; i < n; i++)
    {
        result[i] = i * i;
    }
    return result; // ERROR
}</pre>
```

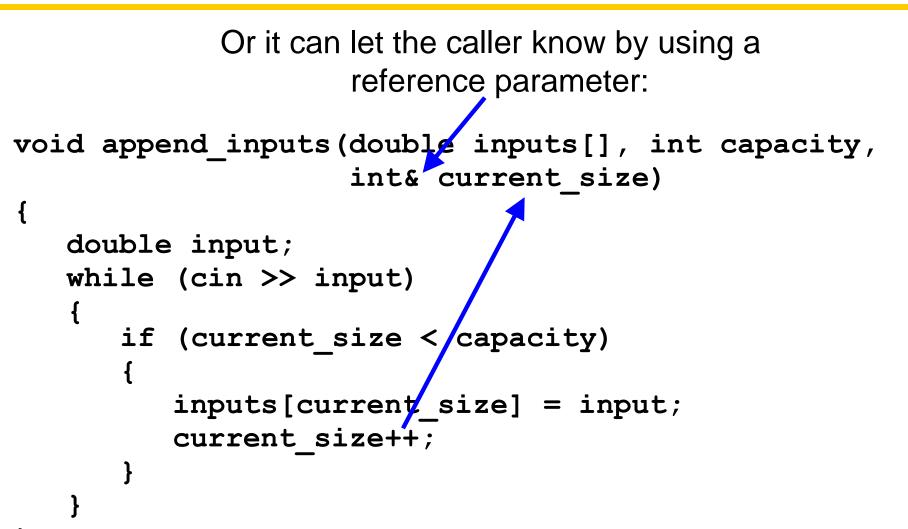


```
A function can change the size of an array.
           It should let the caller know of any change
                   by returning the new size.
int read_inputs(double inputs[], int capacity)
   int current size = Q;
   double input;
   while (cin >> inpu/t)
      if (current size < capacity)
          inputs[current size] = input;
          current size++;
   return current size;
```

Here is a call to the function:

```
const int MAXIMUM_NUMBER_OF_VALUES = 1000;
double values[MAXIMUM_NUMBER_OF_VALUES];
int current_size =
    read_inputs(values, MAXIMUM_NUMBER_OF_VALUES);
```

After the call, the current_size variable specifies how many were added.



Here is a call to the reference parameter version of append_inputs:

> As before, after the call, the current_size variable specifies how many are in the array.

The following program uses the preceding functions to read values from standard input, double them, and print the result.

- The **read_inputs** function fills an array with the input values. It returns the number of elements that were read.
- The multiply function modifies the contents of the array that it receives, demonstrating that arrays can be changed inside the function to which they are passed.
- The print function does not modify the contents of the array that it receives.

#include <iostream>
using namespace std;

ch06/functions.cpp

/**

Reads a sequence of floating-point numbers. **@param** inputs an array containing the numbers Oparam capacity the capacity of that array @return the number of inputs stored in the array */ int read inputs (double inputs [], int capacity) { int current size = 0; cout << "Please enter values, Q to quit:" << endl; bool more = true; while (more)

```
double input;
                                       ch06/functions.cpp
   cin >> input;
   if (cin.fail())
      more = false;
   else if (current size < capacity)</pre>
      inputs[current size] = input;
      current size++;
return current size;
```

ch06/functions.cpp

```
/**
Multiplies all elements of an array by a factor.
@param values a partially filled array
Oparam size the number of elements in values
Oparam factor the value with which each element is
  multiplied
*/
void multiply(double values[], int size,
              double factor)
{
   for (int i = 0; i < size; i++)
   {
      values[i] = values[i] * factor;
   }
```

ch06/functions.cpp

/** Prints the elements of a vector, separated by commas. **@param** values a partially filled array Oparam size the number of elements in values */ void print(double values[], int size) Ł for (int i = 0; i < size; i++) ł if (i > 0) { cout << ", "; } cout << values[i];</pre> cout << endl;

return 0;

ch06/functions.cpp

```
int main()
{
    const int CAPACITY = 1000;
    double values[CAPACITY];
    int size = read_inputs(values, CAPACITY);
    multiply(values, size, 2);
    print(values, size);
```



End Arrays and Vectors I

C++ for Everyone by Cay Horstmann Copyright © 2012 by John Wiley & Sons. All rights reserved

Slides by Evan Gallagher & Nikolay Kirov