Arrays

CS 115

Dr. Joseph Eremondi, adapted from Dr. Shakil Khan, Dr. Philip Fong, and Dr. Howard Hamilton

Last updated: January 15, 2025

One, two, and multi-dimensional arrays

• Print 1000 numbers in reverse order

```
int valueo;
int value1;
int value2;
// ...
int value999;
cin >> valueo;
cin >> value1;
// ...
cin >> value999;
cout << value999 << endl;
cout << value998 << endl;
// ...
cout << valueo << endl;</pre>
```

• How about 1000000 numbers?

- How about 1000000 numbers?
- Tedious, not scalable, and error prone

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - $\circ \ \ \text{homogenous components}$

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - homogenous components
 - indexing support

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - homogenous components
 - indexing support
 - o constant time access

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - homogenous components
 - indexing support
 - constant time access
 - random access

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - homogenous components
 - indexing support
 - constant time access
 - random access

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - homogenous components
 - indexing support
 - constant time access
 - random access

- How about 1000000 numbers?
- Tedious, not scalable, and error prone
- Solution: use aggregate data type
 - o homogenous components
 - indexing support
 - constant time access
 - o random access

```
int a[120000];  // Array declaration

for (int i = 0; i < 120000; i++)
    cin >> a[i];  // Array access
for (int i = 119999; i >= 0; i--)
    cout << a[i] << endl;</pre>
```

Array Operations

• Call the things we store in the array *elements*

Array Operations

- Call the things we store in the array elements
- Get the ith element's value: array[i]

Array Operations

- Call the things we store in the array elements
- Get the ith element's value: array[i]
- Set the ith element: array[i] = someValue;

• Array size must be a constant expression

- Array size must be a constant expression
- Easy to change size: just update N (the rest of the program remains intact)

```
int sumArray(int a[], unsigned int n) // Array argument
 int sum = 0;
 for (int i = 0; i < n; i++)
   sum += a[i];
 return sum;
int main()
 // Arrav initialization
 int a[] = { 3, 24, -88, 17, -1 };
 cout << sumArray(a, 5) << endl;</pre>
```

 Array size can be left unspecified in array initialization syntax

Array arguments are always automatically passed by reference

- Array arguments are always automatically passed by reference
- no special notation is require

- Array arguments are always automatically passed by reference
- no special notation is require

- Array arguments are always automatically passed by reference
- no special notation is require

- Array arguments are always automatically passed by reference
- no special notation is require

```
// int sumArray(int8 a[], unsigned int n) - INCORRECT
int sumArray(int a[], unsigned int n) // CORRECT
{
   ...
}
```

 Works for arrays of all sizes (size is passed as a separate argument)

- Array arguments are always automatically passed by reference
- no special notation is require

```
// int sumArray(int8 a[], unsigned int n) - INCORRECT
int sumArray(int a[], unsigned int n) // CORRECT
{
   ...
}
```

- Works for arrays of all sizes (size is passed as a separate argument)
- Interface not safe: can modify the content of A

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

• Use the following instead:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

• Use the following instead:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

• Use the following instead:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

• Use the following instead:

```
int sumArray(const int a[], unsigned int n)
```

 How to figure out array size when passing n if the size was left unspecified when declaring it?

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

```
int sumArray(const int a[], unsigned int n)
```

- How to figure out array size when passing n if the size was left unspecified when declaring it?
- use sizeof function:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

```
int sumArray(const int a[], unsigned int n)
```

- How to figure out array size when passing n if the size was left unspecified when declaring it?
- use sizeof function:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

```
int sumArray(const int a[], unsigned int n)
```

- How to figure out array size when passing n if the size was left unspecified when declaring it?
- use sizeof function:

```
int sumArray(int a[], unsigned int n)
// not safe, sumArray can modify A!
```

```
int sumArray(const int a[], unsigned int n)
```

- How to figure out array size when passing n if the size was left unspecified when declaring it?
- use sizeof function:

```
int a[] = {1,2,6,3,8};
int x = sumArray(a, sizeof(a) / sizeof(int));
```

• Check if integer array sorted

• Check if integer array sorted

• Check if integer array sorted

· Check if integer array sorted

```
bool arrayIsSorted(const int a[], unsigned int n){
   for (int i = 0; i < n-1; i++){
      if (a[i] > a[i+1])
        return false;
   }
   return true;
}
```

• Reversing items in integer array

• Reversing items in integer array

• Reversing items in integer array

· Reversing items in integer array

```
void swap(int &a, int &b) {
   int tmp = a;
   a = b;
   b = tmp;
}
// below a[] is not a constant as want to produce side-effect
void reverseArray(int a[], unsigned int n) {
   for (int i = 0; i < n/2; i++)
      swap(a[i], a[n - i - 1]);
}</pre>
```

• Compute the sum of an array segment

• Compute the sum of an array segment

• Compute the sum of an array segment

· Compute the sum of an array segment

• Another way to do the same thing

• Another way to do the same thing

• Another way to do the same thing

Another way to do the same thing

```
// begin: index of first component in the subarray
// end : index of the last component in the subarray
int sumSubarray(const int a[],
                 unsigned int begin,
                 unsigned int end){
  assert(begin <= end);</pre>
  int sum = \odot;
  for (int i = begin; i <= end; i++)
    sum += a[i]:
  return sum;
```

• C++ does not check if array indices are within bound

- C++ does not check if array indices are within bound
- it's your responsibility

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

a = b // invalid

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

a = b // invalid

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

a = b // invalid

- C++ does not check if array indices are within bound
- it's your responsibility
- Array Copying

• Array Comparison

• Array Comparison

• Array Comparison

Array Comparison

• compare each pair of cells at a time

Array Comparison

```
if(a == b) // invalid
```

- compare each pair of cells at a time
- No need to return array as function output, uses call by reference anyway!

• C++ arrays are unsafe

- C++ arrays are unsafe
- This is *terrible* language design

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler
- C++ will never change

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler
- C++ will never change
 - Backwards compatibility

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler
- C++ will never change
 - Backwards compatibility
 - std::array is safe but isn't the default

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler
- C++ will never change
 - Backwards compatibility
 - std::array is safe but isn't the default
- Languages like Rust make sure that these errors are impossible

- C++ arrays are unsafe
- This is terrible language design
 - Billions of dollars and many security incidents caused by unsafe memory access
 - Error cost outweighs performance cost of checking array bounds
 - Most checks can be optimized out by the compiler
- C++ will never change
 - Backwards compatibility
 - std::array is safe but isn't the default
- Languages like Rust make sure that these errors are impossible
 - o Unless you explicitly disable safety

```
#include <iostream>
using namespace std;
int main(){
  char passwd[8] = "secret";
  char username[8] = "bob101";
  string toPrint = "";
 // Oops reading past end of array!
  for (int i = 0; i < 16; i++){
    toPrint += username[i];
 cout << toPrint << endl;</pre>
```

```
#include <iostream>
using namespace std;
int main(){
  char passwd[8] = "secret";
  char username[8] = "bob101";
  string toPrint = "";
 // Oops reading past end of array!
  for (int i = 0; i < 16; i++){
    toPrint += username[i];
 cout << toPrint << endl;</pre>
```

```
#include <iostream>
using namespace std;
int main(){
  char passwd[8] = "secret";
  char username[8] = "bob101";
  string toPrint = "";
 // Oops reading past end of array!
  for (int i = 0; i < 16; i++){
    toPrint += username[i];
  cout << toPrint << endl;</pre>
```

bob101secret

Two Dimensional Arrays

• Want to store quantity of different products sold in a store

- Want to store quantity of different products sold in a store
- but for multiple locations/regions

- Want to store quantity of different products sold in a store
- · but for multiple locations/regions
- Conceptually can store as a matrix, where rows represent different locations and columns represent different products

- Want to store quantity of different products sold in a store
- but for multiple locations/regions
- Conceptually can store as a matrix, where rows represent different locations and columns represent different products
- sales[2][1] are the total number of items sold for location 2 and product 1

- Want to store quantity of different products sold in a store
- but for multiple locations/regions
- Conceptually can store as a matrix, where rows represent different locations and columns represent different products
- sales[2][1] are the total number of items sold for location 2 and product 1
- recall item n is the (n+1)-th item

- Want to store quantity of different products sold in a store
- but for multiple locations/regions
- Conceptually can store as a matrix, where rows represent different locations and columns represent different products
- sales[2][1] are the total number of items sold for location 2 and product 1
- recall item n is the (n+1)-th item
 - index starts from 0!

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
sales[1][0] // recall, indices start from 0
```

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
sales[1][0] // recall, indices start from 0
```

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
sales[1][0] // recall, indices start from 0
```

```
const unsigned int NUM_OF_REGIONS = 4;
const unsigned int NUM_OF_PRODUCTS = 3;
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

 To access sales figure for first product in second region, use:

```
sales[1][0] // recall, indices start from 0
```

```
sales[1][0] = 500;
```

Populating and Accessing

Populating and Accessing

Populating and Accessing

```
// Read input stream
for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
  for (unsigned int product = 0; product < NUM_OF_PRODUCTS; product++
      cin >> sales[region][product];

// total sales for a particular product (product 0)
unsigned int total_sales = 0;

for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
      // add up sales from all regions for product 0
      total_sales += sales[region][0];</pre>
```

Can you compute total sales from region 1?

Passing 2D Arrays

Passing 2D Arrays

Passing 2D Arrays

```
unsigned int sumProductSales(
          unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS],
          unsigned int product)
{
   unsigned int total_sales = 0;
   for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
      total_sales += sales[region][product];
   return total_sales;
}</pre>
```

Can you implement a safer interface?

Passing 2D Arrays

```
unsigned int sumProductSales(
          unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS],
          unsigned int product)
{
   unsigned int total_sales = 0;
   for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
      total_sales += sales[region][product];
   return total_sales;
}</pre>
```

- Can you implement a safer interface?
- As usual, can leave size of first dimension unspecified, e.g. int F(int arr[][SIZE])

Passing 2D Arrays

```
unsigned int sumProductSales(
          unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS],
          unsigned int product)
{
   unsigned int total_sales = 0;
   for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
      total_sales += sales[region][product];
   return total_sales;
}</pre>
```

- Can you implement a safer interface?
- As usual, can leave size of first dimension unspecified, e.g. int F(int arr[][SIZE])
- but not the second one (why?)

```
// Implement a function that returns
// the value of one element from the sales array
unsigned int getSales(
        const unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS],
        unsigned int r, unsigned int p){
  return sales[r][p];
// Implement a function that sets the value
// of one element from the sales array
void setSales(unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS],
              unsigned int r, unsigned int p, unsigned int v){
 sales[r][p] = v;
```

```
// too lazy to write long types? Use typedef instead!

typedef unsigned int Sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];

unsigned int sumSales(const Sales sales){
    ...
    }
```

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

versus

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

Issue: how to map between these two?

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

- Issue: how to map between these two?
 - o row-major vs. column-major order

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

- Issue: how to map between these two?
 - o row-major vs. column-major order
 - ∘ e.g. sales[i][j]

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

- Issue: how to map between these two?
 - o row-major vs. column-major order
 - ∘ e.g. sales[i][j]
 - same as _sales[i * NUM_OF_PRODUCTS + j] in row-major

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

- Issue: how to map between these two?
 - o row-major vs. column-major order
 - ∘ e.g. sales[i][j]
 - same as _sales[i * NUM_OF_PRODUCTS + j] in row-major
- Now you know why the size of the 2nd dimension can't be left unspecified!

```
unsigned int sales[NUM_OF_REGIONS][NUM_OF_PRODUCTS];
```

```
unsigned int _sales[NUM_OF_REGIONS * NUM_OF_PRODUCTS];
```

- Issue: how to map between these two?
 - o row-major vs. column-major order
 - ∘ e.g. sales[i][j]
 - same as _sales[i * NUM_OF_PRODUCTS + j] in row-major
- Now you know why the size of the 2nd dimension can't be left unspecified!
 - Can you write the formula for column-major order?

```
unsigned int totalSales = 0;
for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
for (unsigned int product = 0;
    product < NUM_OF_PRODUCTS;
    product++){
    totalSales += _sales[region * NUM_OF_PRODUCTS + product];
}</pre>
```

This is why we need to know the size of the second dimension

```
unsigned int totalSales = 0;
for (unsigned int region = 0; region < NUM_OF_REGIONS; region++)
for (unsigned int product = 0;
    product < NUM_OF_PRODUCTS;
    product++){
    totalSales += _sales[region * NUM_OF_PRODUCTS + product];
}</pre>
```

- This is why we need to know the size of the second dimension
 - To calculate offset

Multi-dimensional Arrays

Multi-dimensional Arrays

Multi-dimensional Arrays

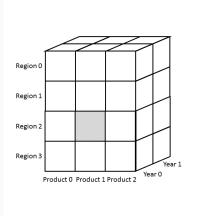
```
const unsigned int NUM_YEARS = 2;
const unsigned int NUM_REGIONS = 4;
const unsigned int NUM_PRODUCTS = 3;

typedef unsigned int Sales[NUM_YEARS][NUM_REGIONS][NUM_PRODUCTS];

unsigned int total_sales = 0;
for (unsigned int year = 0; year < NUM_YEARS; year++)
  for (unsigned int region = 0; region < NUM_REGIONS; region++)
  for (unsigned int product = 0; product < NUM_PRODUCTS; product++)
    total_sales += sales[year][region][product];</pre>
```

Simulating 3d with 1d

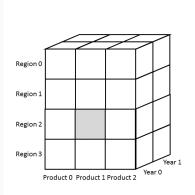
• Sales[year][region][product]



Mem-Pos	
0	YOROPO
1	YOROP1
2	YOROP2
3	YOR1PO
4	YOR1P1
5	YOR1P2
6	YOR2PO
7	YOR2P1
8	YOR2P2
9	YOR3PO
10	YOR3P1
11	YOR3P2
12	Y1ROPO
13	Y1ROP1
14	Y1ROP2
15	Y1R1P0
16	Y1R1P1
17	Y1R1P2
18	Y1R2P0
19	Y1R2P1
20	Y1R2P2
21	Y1R3P0
22	Y1R3P1
23	Y1R3P2

Simulating 3d with 1d

- Sales[year][region][product]
- vs_Sales[(year * NUM_REGS * NUM_PRODS) + (region * NUM_OF_PRODS) + product]



YOROPO
YOROP1
YOROP2
YOR1PO
YOR1P1
YOR1P2
YOR2PO
YOR2P1
YOR2P2
YOR3P0
YOR3P1
YOR3P2
Y1ROPO
Y1ROP1
Y1ROP2
Y1R1P0
Y1R1P1
Y1R1P2
Y1R2P0
Y1R2P1
Y1R2P2
Y1R3P0
Y1R3P1
Y1R3P2