

CS107, Lecture 8

C Generics – Void *

CS107 Topic 4: How can we use our knowledge of memory and data representation to write code that works with any data type?

Learning Goals

- Learn how to write C code that works with any data type.
- Learn about how to use void * and avoid potential pitfalls.

Lecture Plan

- **Overview:** Generics
- Generic Swap
- Generics Pitfalls
- Generic Array Swap
- Generic Stack

```
cp -r /afs/ir/class/cs107/lecture-code/lect8 .
```

Lecture Plan

- **Overview: Generics**
 - Generic Swap
 - Generics Pitfalls
 - Generic Array Swap
 - Generic Stack

Generics

- We always strive to write code that is as general-purpose as possible.
- Generic code reduces code duplication and means you can make improvements and fix bugs in one place rather than many.
- Generics is used throughout C for functions to sort any array, search any array, free arbitrary memory, and more.
- How can we write generic code in C?

Lecture Plan

- **Overview:** Generics
- **Generic Swap**
- Generics Pitfalls
- Generic Array Swap
- Generic Stack

Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

```
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
...	...
x 0xff14	2
y 0xff10	5
...	...

main()

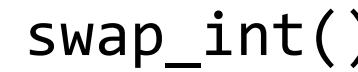
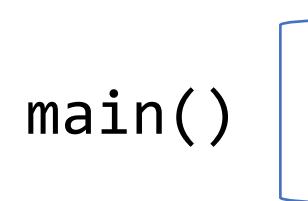


Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
x 0xff14	2
y 0xff10	5
b 0xf18	0xff10
a 0xf10	0xff14

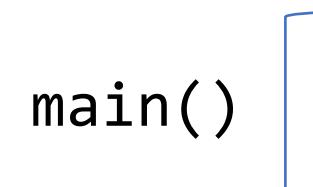


Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
...	...
x 0xff14	2
y 0xff10	5
...	...
b 0xf18	0xff10
a 0xf10	0xff14
temp 0xf0c	2
...	...



Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
x	0xff14
y	0xff10
b	0xf18
a	0xf10
temp	0xf0c

The diagram shows the state of the stack for both the main() and swap_int() functions. In the main() frame, variables x and y are at addresses 0xff14 and 0xff10 respectively, containing the values 5 and 2. In the swap_int() frame, the parameter b is at address 0xf18 and parameter a is at address 0xf10, both pointing to the same memory location 0xff10. A red arrow points from the value 5 in the x slot to the a slot, indicating that the swap operation has occurred.

Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
x	0xff14
y	0xff10
b	0xf18
a	0xf10
temp	0xf0c

The diagram shows the state of the stack during the execution of the swap operation. The stack grows from bottom to top. The current stack frame for `swap_int()` contains local variables `a` (value 0xff14), `b` (value 0xff10), and `temp` (value 0xf0c). The stack frame for `main()` contains local variables `x` (value 0xff14) and `y` (value 0xff10). Red arrows indicate the movement of values: one arrow points from `x` to `temp`, and another points from `y` to `a`.

Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

```
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
...	...
x 0xff14	5
y 0xff10	2
...	...

main()



Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

```
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
...	...
x 0xff14	5
y 0xff10	2
...	...

main()



Swap

You're asked to write a function that swaps two numbers.

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

```
int main(int argc, char *argv[]) {  
    int x = 2;  
    int y = 5;  
    swap_int(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

Stack Address	Value
...	...
x 0xff14	5
y 0xff10	2
...	...

main()



**“Oh, when I said ‘numbers’
I meant shorts, not ints.”**



Swap

```
void swap_short(short *a, short *b) {
    short temp = *a;
    *a = *b;
    *b = temp;
}

int main(int argc, char *argv[]) {
    short x = 2;
    short y = 5;
    swap_short(&x, &y);
    // want x = 5, y = 2
    printf("x = %d, y = %d\n", x, y);
    return 0;
}
```

Swap

```
void swap_short(short *a, short *b) {  
    short temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    short x = 2;  
    short y = 5;  
    swap_short(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %d, y = %d\n", x, y);  
    return 0;  
}
```

	Address	Stack Value
	x 0xff12	2
	y 0xff10	5
main()	b 0xf18	0xff10
	a 0xf10	0xff12
	temp 0xf0e	2

The diagram illustrates the state of the stack during the execution of the swap operation. The stack grows from higher addresses (top) to lower addresses (bottom). The variables x and y are at addresses 0xff12 and 0xff10 respectively. In the main() function, the arguments b=0xf18 and a=0xf10 are passed to the swap_short() function. Inside swap_short(), a temporary variable temp is created at address 0xf0e and assigned the value 2. Red arrows indicate the movement of values: one arrow points from the value 2 at address 0xf0e up to the variable temp; another arrow points from the value 5 at address 0xf10 up to the variable a.

**“You know what, I goofed.
We’re going to use strings.
Could you write something
to swap those?”**



Swap

```
void swap_string(char **a, char **b) {
    char *temp = *a;
    *a = *b;
    *b = temp;
}

int main(int argc, char *argv[]) {
    char *x = "2";
    char *y = "5";
    swap_string(&x, &y);
    // want x = 5, y = 2
    printf("x = %s, y = %s\n", x, y);
    return 0;
}
```

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {
    char *temp = *a;
    *a = *b;
    *b = temp;
}

int main(int argc, char *argv[]) {
    char *x = "2";
    char *y = "5";
    swap_string(&x, &y);
    // want x = 5, y = 2
    printf("x = %s, y = %s\n", x, y);
    return 0;
}
```

	Address	Value
main()	x	0xff18
main()	y	0xff10
swap_string()	b	0xf18
swap_string()	a	0xf10
DATA SEGMENT	0xf	'\0'
DATA SEGMENT	0xe	'5'
DATA SEGMENT	0xd	'\0'
DATA SEGMENT	0xc	'2'
		...
		...
		...
		...
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
swap_string()	b	0xf18
	a	0xf10
	temp	0xf08
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
swap_string()	b	0xf18
	a	0xf10
	temp	0xf08
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
swap_string()	b	0xf18
	a	0xf10
	temp	0xf08
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x	0xff18
	y	0xff10
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

Swap

```
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
int main(int argc, char *argv[]) {  
    char *x = "2";  
    char *y = "5";  
    swap_string(&x, &y);  
    // want x = 5, y = 2  
    printf("x = %s, y = %s\n", x, y);  
    return 0;  
}
```

	Address	Value
main()	x 0xff18	...
	y 0xff10	0xe
DATA SEGMENT	0xf	'\0'
	0xe	'5'
	0xd	'\0'
	0xc	'2'
		...

The diagram illustrates the state of memory during the execution of the program. The top part shows the stack frame for the `main()` function, with local variables `x` and `y` pointing to memory locations 0xff18 and 0xff10 respectively. The bottom part shows the data segment, which contains five character values: '\0', '5', '\0', '2', and another '\0'. Red arrows indicate that the value at address 0xe (which is the value of `x`) has been copied to the memory location containing '5', and the value at address 0xc (which is the value of `y`) has been copied to the memory location containing '2'.

“Awesome! Thanks.”

“Awesome! Thanks. We also have 20 custom struct types. Could you write swap for those too?”



Generic Swap

What if we could write *one* function to swap two values of any single type?

```
void swap_int(int *a, int *b) { ... }  
void swap_float(float *a, float *b) { ... }  
void swap_size_t(size_t *a, size_t *b) { ... }  
void swap_double(double *a, double *b) { ... }  
void swap_string(char **a, char **b) { ... }  
void swap_mystruct(mystruct *a, mystruct *b) { ... }
```

...

Generic Swap

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
void swap_short(short *a, short *b) {  
    short temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

Generic Swap

```
void swap_int(int *a, int *b) {  
    int temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
void swap_short(short *a, short *b) {  
    short temp = *a;  
    *a = *b;  
    *b = temp;  
}  
  
void swap_string(char **a, char **b) {  
    char *temp = *a;  
    *a = *b;  
    *b = temp;  
}
```

All 3:

- Take pointers to values to swap
- Create temporary storage to store one of the values
- Move data at **b** into where **a** points
- Move data in temporary storage into where **b** points

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

```
int temp = *data1ptr;
```

4 bytes

```
short temp = *data1ptr;
```

2 bytes

```
char *temp = *data1ptr;
```

8 bytes

Problem: each type may need a different size temp!

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

```
*data1Ptr = *data2ptr;
```

4 bytes

```
*data1Ptr = *data2ptr;
```

2 bytes

```
*data1Ptr = *data2ptr;
```

8 bytes

Problem: each type needs to copy a different amount of data!

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

*data2ptr = temp;

4 bytes

*data2ptr = temp;

2 bytes

*data2ptr = temp;

8 bytes

Problem: each type needs to copy a different amount of data!

**C knows the size of temp,
and knows how many bytes
to copy, because of the
variable types.**

**Is there a way to make a
version that doesn't care
about the variable types?**

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(pointer to data1, pointer to data2) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr) {  
    store a copy of data1 in temporary storage  
    copy data2 to location of data1  
    copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr) {  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr) {  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

If we don't know the data type, we don't know how many bytes it is. Let's take that as another parameter.

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

If we don't know the data type, we don't know how many bytes it is. Let's take that as another parameter.

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Let's start by making space to store the temporary value. How can we make **nbytes** of temp space?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    void temp; ???  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Let's start by making space to store the temporary value. How can we make **nbytes** of temp space?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

temp is **nbytes** of memory,
since each **char** is 1 byte!

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Now, how can we copy in what
data1ptr points to into **temp**?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    temp = *data1ptr; ???  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Now, how can we copy in what
data1ptr points to into **temp**?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    temp = *data1ptr; ???  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

We can't dereference a **void *** (or set an array equal to something). C doesn't know what it points to! Therefore, it doesn't know how many bytes there it should be looking at.

memcpy

memcpy is a function that copies a specified amount of bytes at one address to another address.

```
void *memcpy(void *dest, const void *src, size_t n);
```

It copies the next n bytes that src points to to the location contained in dest. (It also returns **dest**). It does not support regions of memory that overlap.

```
int x = 5;  
int y = 4;  
memcpy(&x, &y, sizeof(x)); // like x = y
```

memcpy must take **pointers** to the bytes to work with to know where they live and where they should be copied to.

memmove

memmove is the same as `memcpy`, but supports overlapping regions of memory. (Unlike its name implies, it still “copies”).

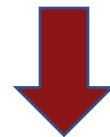
```
void *memmove(void *dest, const void *src, size_t n);
```

It copies the next `n` bytes that `src` points to to the location contained in `dest`. (It also returns `dest`).

memmove

When might memmove be useful?

1	2	3	4	5	6	7
---	---	---	---	---	---	---



4	5	6	7	5	6	7
---	---	---	---	---	---	---

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    temp = *data1ptr; ???  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

We can't dereference a **void ***. C doesn't know what it points to! Therefore, it doesn't know how many bytes there it should be looking at.

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    temp = *data1ptr; ???  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

How can **memcpy** or **memmove** help us here?

```
void *memcpy(void *dest, const void *src, size_t n);
```

```
void *memmove(void *dest, const void *src, size_t n);
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

We can copy the bytes ourselves into temp! This is equivalent to **temp = *data1ptr** in non-generic versions, but this works for *any* type of *any* size.

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    // copy data in temporary storage to location of data2  
}
```

How can we copy data2 to the location of data1?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    *data1ptr = *data2ptr; ???  
    // copy data in temporary storage to location of data2  
}
```

How can we copy data2 to the location of data1?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
}
```

How can we copy data2 to the location of data1?
memcpy!

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
}
```

How can we copy temp's data to the location of data2?

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
    memcpy(data2ptr, temp, nbytes);  
}
```

How can we copy temp's data to the location of data2? **memcpy!**

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
    memcpy(data2ptr, temp, nbytes);  
}
```

```
int x = 2;  
int y = 5;  
swap(&x, &y, sizeof(x));
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
    memcpy(data2ptr, temp, nbytes);  
}
```

```
short x = 2;  
short y = 5;  
swap(&x, &y, sizeof(x));
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
    memcpy(data2ptr, temp, nbytes);  
}
```

```
char *x = "2";  
char *y = "5";  
swap(&x, &y, sizeof(x));
```

Generic Swap

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    // store a copy of data1 in temporary storage  
    memcpy(temp, data1ptr, nbytes);  
    // copy data2 to location of data1  
    memcpy(data1ptr, data2ptr, nbytes);  
    // copy data in temporary storage to location of data2  
    memcpy(data2ptr, temp, nbytes);  
}
```

```
mystruct x = {...};  
mystruct y = {...};  
swap(&x, &y, sizeof(x));
```

C Generics

- We can use **void *** and **memcpy** to handle memory as generic bytes.
- If we are given where the data of importance is, and how big it is, we can handle it!

```
void swap(void *data1ptr, void *data2ptr, size_t nbytes) {  
    char temp[nbytes];  
    memcpy(temp, data1ptr, nbytes);  
    memcpy(data1ptr, data2ptr, nbytes);  
    memcpy(data2ptr, temp, nbytes);  
}
```

Lecture Plan

- **Overview:** Generics
- Generic Swap
- **Generics Pitfalls**
- Generic Array Swap
- Generic Stack

Void * Pitfalls

- **void ***s are powerful, but dangerous - C cannot do as much checking!
- E.g. with **int**, C would never let you swap *half* of an int. With **void ***s, this can happen! (*How? Let's find out!*)

Demo: Void *s Gone Wrong



swap.c

Void * Pitfalls

- Void * has more room for error because it manipulates arbitrary bytes without knowing what they represent. This can result in some strange memory Frankenstein!



Lecture Plan

- **Overview:** Generics
- Generic Swap
- Generics Pitfalls
- **Generic Array Swap**
- Generic Stack

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers.

```
void swap_ends_int(int *arr, size_t nelems) {  
    int tmp = arr[0];  
    arr[0] = arr[nelems - 1];  
    arr[nelems - 1] = tmp;  
}
```

```
int main(int argc, char *argv[]) {  
    int nums[] = {5, 2, 3, 4, 1};  
    size_t nelems = sizeof(nums) / sizeof(nums[0]);  
    swap_ends_int(nums, nelems);  
    // want nums[0] = 1, nums[4] = 5  
    printf("nums[0] = %d, nums[4] = %d\n", nums[0], nums[4]);  
    return 0;  
}
```

Wait – we just wrote a generic swap function. Let's use that!

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers.

```
void swap_ends_int(int *arr, size_t nelems) {  
    swap(arr, arr + nelems - 1, sizeof(*arr));  
}  
  
int main(int argc, char *argv[]) {  
    int nums[] = {5, 2, 3, 4, 1};  
    size_t nelems = sizeof(nums) / sizeof(nums[0]);  
    swap_ends_int(nums, nelems);  
    // want nums[0] = 1, nums[4] = 5  
    printf("nums[0] = %d, nums[4] = %d\n", nums[0], nums[4]);  
    return 0;  
}
```

Wait – we just wrote a generic swap function. Let's use that!

Swap Ends

Let's write out what some other versions would look like (just in case).

```
void swap_ends_int(int *arr, size_t nelems) {
    swap(arr, arr + nelems - 1, sizeof(*arr));
}
```

```
void swap_ends_short(short *arr, size_t nelems) {
    swap(arr, arr + nelems - 1, sizeof(*arr));
}
```

```
void swap_ends_string(char **arr, size_t nelems) {
    swap(arr, arr + nelems - 1, sizeof(*arr));
}
```

```
void swap_ends_float(float *arr, size_t nelems) {
    swap(arr, arr + nelems - 1, sizeof(*arr));
}
```

The code seems to be the same regardless of the type!

Swap Ends

Let's write a version of `swap_ends` that works for any type of array.

```
void swap_ends(void *arr, size_t nelems) {  
    swap(arr, arr + nelems - 1, sizeof(*arr));  
}
```

Is this generic? Does this work?

Swap Ends

Let's write a version of `swap_ends` that works for any type of array.

```
void swap_ends(void *arr, size_t nelems) {
    swap(arr, arr + nelems - 1, sizeof(*arr));
}
```

Is this generic? Does this work?

Unfortunately not. First, we no longer know the element size. Second, pointer arithmetic depends on the type of data being pointed to. With a `void *`, we lose that information!

Swap Ends

Let's write a version of swap_ends that works for any type of array.

```
void swap_ends(void *arr, size_t nelems) {  
    swap(arr, arr + nelems - 1, sizeof(*arr));  
}
```

We need to know the element size, so let's add a parameter.

Swap Ends

Let's write a version of swap_ends that works for any type of array.

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, arr + nelems - 1, elem_bytes);  
}
```

We need to know the element size, so let's add a parameter.

Pointer Arithmetic

`arr + nelems - 1`

Let's say `nelems = 4`. How many bytes beyond `arr` is this?

If it's an array of...

Int?

Pointer Arithmetic

`arr + nelems - 1`

Let's say `nelems = 4`. How many bytes beyond `arr` is this?

If it's an array of...

Int: adds 3 places to `arr`, and $3 * \text{sizeof(int)} = 12$ bytes

Pointer Arithmetic

`arr + nelems - 1`

Let's say `nelems = 4`. How many bytes beyond `arr` is this?

If it's an array of...

Int: adds 3 places to `arr`, and $3 * \text{sizeof(int)} = 12$ bytes

Short?

Pointer Arithmetic

$\text{arr} + \text{nelems} - 1$

Let's say nelems = 4. How many bytes beyond arr is this?

If it's an array of...

Int: adds 3 places to arr, and $3 * \text{sizeof(int)} = 12$ bytes

Short: adds 3 places to arr, and $3 * \text{sizeof(short)} = 6$ bytes

Pointer Arithmetic

`arr + nelems - 1`

Let's say `nelems = 4`. How many bytes beyond `arr` is this?

If it's an array of...

Int: adds 3 places to `arr`, and $3 * \text{sizeof(int)} = 12$ bytes

Short: adds 3 places to `arr`, and $3 * \text{sizeof(short)} = 6$ bytes

Char *: adds 3 places to `arr`, and $3 * \text{sizeof(char *)} = 24$ bytes

In each case, we need to know the element size to do the arithmetic.

Swap Ends

Let's write a version of `swap_ends` that works for any type of array.

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, arr + nelems - 1, elem_bytes);  
}
```

How many bytes past `arr` should we go to get to the last element?

$(nelems - 1) * elem_bytes$

Swap Ends

Let's write a version of swap_ends that works for any type of array.

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

How many bytes past arr should we go to get to the last element?

$(nelems - 1) * elem_bytes$

Swap Ends

Let's write a version of swap_ends that works for any type of array.

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

But C still can't do arithmetic with a void*. We need to tell it to not worry about it, and just add bytes. **How can we do this?**

Swap Ends

Let's write a version of swap_ends that works for any type of array.

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

But C still can't do arithmetic with a void*. We need to tell it to not worry about it, and just add bytes. **How can we do this?**

char * pointers already add bytes!

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers. Well, now it can swap for an array of anything!

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers. Well, now it can swap for an array of anything!

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);
}
```

```
int nums[] = {5, 2, 3, 4, 1};
size_t nelems = sizeof(nums) / sizeof(nums[0]);
swap_ends(nums, nelems, sizeof(nums[0]));
```

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers. Well, now it can swap for an array of anything!

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

```
short nums[] = {5, 2, 3, 4, 1};  
size_t nelems = sizeof(nums) / sizeof(nums[0]);  
swap_ends(nums, nelems, sizeof(nums[0]));
```

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers. Well, now it can swap for an array of anything!

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);
}
```

```
char *strs[] = {"Hi", "Hello", "Howdy"};
size_t nelems = sizeof(strs) / sizeof(strs[0]);
swap_ends(strs, nelems, sizeof(strs[0]));
```

Swap Ends

You're asked to write a function that swaps the first and last elements in an array of numbers. Well, now it can swap for an array of anything!

```
void swap_ends(void *arr, size_t nelems, size_t elem_bytes) {  
    swap(arr, (char *)arr + (nelems - 1) * elem_bytes, elem_bytes);  
}
```

```
mystruct structs[] = ...;  
size_t nelems = ...;  
swap_ends(structs, nelems, sizeof(structs[0]));
```

Demo: Void *s Gone Wrong



swap_ends.c

Lecture Plan

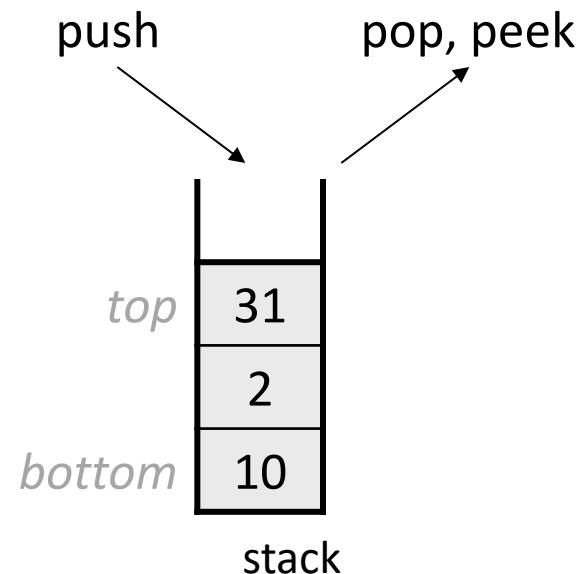
- **Overview:** Generics
- Generic Swap
- Generics Pitfalls
- **Announcements**
- Generic Array Swap
- **Generic Stack**

Stacks

- C generics are particularly powerful in helping us create generic data structures.
- Let's see how we might go about making a Stack in C.

Refresher: Stacks

- A **Stack** is a data structure representing a stack of things.
- Objects can be *pushed* on top of or *popped* from the top of the stack.
- Only the top of the stack can be accessed; no other objects in the stack are visible.
- Main operations:
 - **push(value)**: add an element to the top of the stack
 - **pop()**: remove and return the top element in the stack
 - **peek()**: return (but do not remove) the top element in the stack

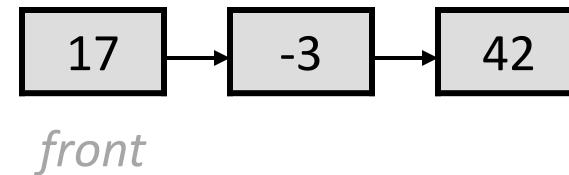


Refresher: Stacks

A stack is often implemented using a **linked list** internally.

- "bottom" = tail of linked list
- "top" = head of linked list (*why not the other way around?*)

```
Stack<int> s;  
s.push(42);  
s.push(-3);  
s.push(17);
```



Problem: C is not object-oriented! We can't call methods on variables.

Demo: Int Stack



int_stack.c

**What modifications are
necessary to make a
generic stack?**

Stack Structs

```
typedef struct int_node {  
    struct int_node *next;  
    int data;  
} int_node;
```

```
typedef struct int_stack {  
    int nelems;  
    int_node *top;  
} int_stack;
```

How might we modify the Stack data representation itself to be generic?

Stack Structs

```
typedef struct int_node {  
    struct int_node *next;  
    int data;  
} int_node;
```

```
typedef struct int_stack {  
    int nelems;  
    int_node *top;  
} int_stack;
```

Problem: each node can no longer store the data itself, because it could be any size!

Generic Stack Structs

```
typedef struct int_node {  
    struct int_node *next;  
    void *data;  
} int_node;
```

```
typedef struct stack {  
    int nelems;  
    int elem_size_bytes;  
    node *top;  
} stack;
```

Solution: each node stores a pointer, which is always 8 bytes, to the data somewhere else. We must also store the data size in the Stack struct.

Stack Functions

- `int_stack_create()`: creates a new stack on the heap and returns a pointer to it
- `int_stack_push(int_stack *s, int data)`: pushes data onto the stack
- `int_stack_pop(int_stack *s)`: pops and returns topmost stack element

int_stack_create

```
int_stack *int_stack_create() {  
    int_stack *s = malloc(sizeof(int_stack));  
    s->nelems = 0;  
    s->top = NULL;  
    return s;  
}
```

How might we modify this function to be generic?

From previous slide:

```
typedef struct stack {  
    int nelems;  
    int elem_size_bytes;  
    node *top;  
} stack;
```

Generic stack_create

```
stack *stack_create(int elem_size_bytes) {
    stack *s = malloc(sizeof(stack));
    s->nelems = 0;
    s->top = NULL;
    s->elem_size_bytes = elem_size_bytes;
    return s;
}
```

int_stack_push

```
void int_stack_push(int_stack *s, int data) {
    int_node *new_node = malloc(sizeof(int_node));
    new_node->data = data;

    new_node->next = s->top;
    s->top = new_node;
    s->nelems++;
}
```

How might we modify this function to be generic?

From previous slide:

```
typedef struct stack {
    int nelems;
    int elem_size_bytes;
    node *top;
} stack;
```

```
typedef struct node {
    struct node *next;
    void *data;
} node;
```

Generic stack_push

```
void int_stack_push(int_stack *s, int data) {  
    int_node *new_node = malloc(sizeof(int_node));  
    new_node->data = data;  
  
    new_node->next = s->top;  
    s->top = new_node;  
    s->nelems++;  
}
```

Problem 1: we can no longer pass the data itself
as a parameter, because it could be any size!

Generic stack_push

```
void int_stack_push(int_stack *s, const void *data) {  
    int_node *new_node = malloc(sizeof(int_node));  
    new_node->data = data;  
  
    new_node->next = s->top;  
    s->top = new_node;  
    s->nelems++;  
}
```

Solution 1: pass a pointer to the data as a parameter instead.

Generic stack_push

```
void int_stack_push(int_stack *s, const void *data) {  
    int_node *new_node = malloc(sizeof(int_node));  
    new_node->data = data;  
  
    new_node->next = s->top;  
    s->top = new_node;  
    s->nelems++;  
}
```

Problem 2: we cannot copy the existing data pointer into new_node. The data structure must manage its own copy that exists for its entire lifetime. The provided copy may go away!

Generic stack_push

```
int main() {
    stack *int_stack = stack_create(sizeof(int));
    add_one(int_stack);
    // now stack stores pointer to invalid memory for 7!
}

void add_one(stack *s) {
    int num = 7;
    stack_push(s, &num);
}
```

Generic stack_push

```
void stack_push(stack *s, const void *data) {  
    node *new_node = malloc(sizeof(node));  
    new_node->data = malloc(s->elem_size_bytes);  
    memcpy(new_node->data, data, s->elem_size_bytes);  
  
    new_node->next = s->top;  
    s->top = new_node;  
    s->nelems++;  
}
```

Solution 2: make a heap-allocated copy
of the data that the node points to.

int_stack_pop

```
int int_stack_pop(int_stack *s) {
    if (s->nelems == 0) {
        error(1, 0, "Cannot pop from empty stack");
    }
    int_node *n = s->top;
    int value = n->data;

    s->top = n->next;

    free(n);
    s->nelems--;
}

return value;
```

How might we modify this function to be generic?

From previous slide:

```
typedef struct stack {
    int nelems;
    int elem_size_bytes;
    node *top;
} stack;
```

```
typedef struct node {
    struct node *next;
    void *data;
} node;
```

Generic stack_pop

```
int int_stack_pop(int_stack *s) {
    if (s->nelems == 0) {
        error(1, 0, "Cannot pop from empty stack");
    }
    int_node *n = s->top;
    int value = n->data;

    s->top = n->next;

    free(n);
    s->nelems--;
    return value;
}
```

Problem: we can no longer return the data itself, because it could be any size!

Generic stack_pop

```
void *int_stack_pop(int_stack *s) {
    if (s->nelems == 0) {
        error(1, 0, "Cannot pop from empty stack");
    }
    int_node *n = s->top;
    void *value = n->data;

    s->top = n->next;
    free(n);
    s->nelems--;
}

return value;
```

While it's possible to return the heap address of the element, this means the client would be responsible for freeing it. Ideally, the data structure should manage its own memory here.

Generic stack_pop

```
void stack_pop(stack *s, void *addr) {
    if (s->nelems == 0) {
        error(1, 0, "Cannot pop from empty stack");
    }
    node *n = s->top;
    memcpy(addr, n->data, s->elem_size_bytes);
    s->top = n->next;

    free(n->data);
    free(n);
    s->nelems--;
}
```

Solution: have the caller pass a memory location as a parameter and copy the data to that location.

Using Generic Stack

```
int_stack *intstack = int_stack_create();
for (int i = 0; i < TEST_STACK_SIZE; i++) {
    int_stack_push(intstack, i);
}
```

We must now pass the *address* of an element to push onto the stack, rather than the element itself.

Using Generic Stack

```
stack *intstack = stack_create(sizeof(int));  
for (int i = 0; i < TEST_STACK_SIZE; i++) {  
    stack_push(intstack, &i);  
}
```

We must now pass the *address* of an element to push onto the stack, rather than the element itself.

Using Generic Stack

```
int_stack *intstack = int_stack_create();
int_stack_push(intstack, 7);
```

We must now pass the *address* of an element to push onto the stack, rather than the element itself.

Using Generic Stack

```
stack *intstack = stack_create(sizeof(int));  
int num = 7;  
stack_push(intstack, &num);
```

We must now pass the *address* of an element to push onto the stack, rather than the element itself.

Using Generic Stack

```
// Pop off all elements
while (intstack->nelems > 0) {
    printf("%d\n", int_stack_pop(intstack));
}
```

We must now pass the *address* of where we would like to store the popped element, rather than getting it directly as a return value.

Using Generic Stack

```
// Pop off all elements
int popped_int;
while (intstack->nelems > 0) {
    int_stack_pop(intstack, &popped_int);
    printf("%d\n", popped_int);
}
```

We must now pass the *address* of where we would like to store the popped element, rather than getting it directly as a return value.

Demo: Generic Stack



generic_stack.c

Recap

- **void *** is a variable type that represents a generic pointer “to something”.
- We cannot perform pointer arithmetic with or dereference a **void ***.
- We can use **memcpy** or **memmove** to copy data from one memory location to another.
- To do pointer arithmetic with a **void ***, we must first cast it to a **char ***.
- **void *** and generics are powerful but dangerous because of the lack of type checking, so we must be extra careful when working with generic memory.

Recap

- **Overview:** Generics
- Generic Swap
- Generics Pitfalls
- Generic Array Swap
- Generic Stack

Next time: More Generics, and Function Pointers

Question Break!

Post any questions or topics you'd like to review on the Ed thread for today's lecture.

Extra Practice

Exercise: Array Rotation

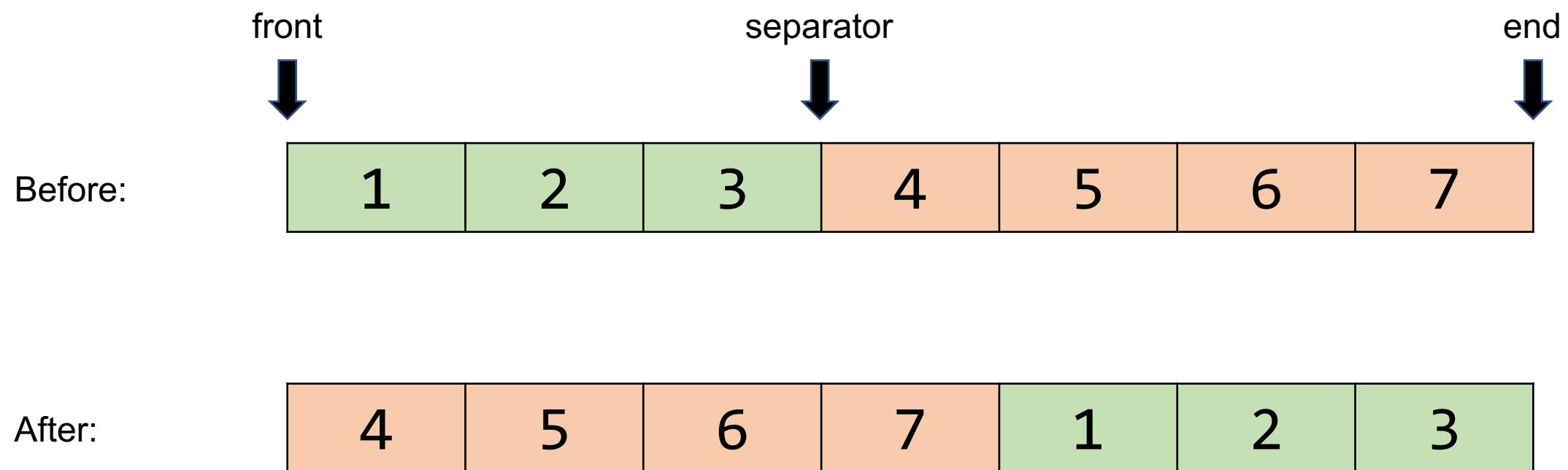
Exercise: You're asked to provide an implementation for a function called **rotate** with the following prototype:

```
void rotate(void *front, void *separator, void *end);
```

The expectation is that **front** is the base address of an array, **end** is the past-the-end address of the array, and **separator** is the address of some element in between. **rotate** moves all elements in between **front** and **separator** to the end of the array, and all elements between **separator** and **end** move to the front.

Exercise: Array Rotation

```
int array[7] = {1, 2, 3, 4, 5, 6, 7};  
rotate(array, array + 3, array + 7);
```



Exercise: Array Rotation

Exercise: A properly implemented **rotate** will prompt the following program to generate the provided output.

```
int main(int argc, char *argv[]) {  
    int array[10] = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10};  
    print_int_array(array, 10); // intuit implementation ☺  
    rotate(array, array + 5, array + 10);  
    print_int_array(array, 10);  
    rotate(array, array + 1, array + 10);  
    print_int_array(array, 10);  
    rotate(array + 4, array + 5, array + 6);  
    print_int_array(array, 10);  
    return 0;  
}
```

Output:

```
myth52:~/lect8$ ./rotate  
Array: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10  
Array: 6, 7, 8, 9, 10, 1, 2, 3, 4, 5  
Array: 7, 8, 9, 10, 1, 2, 3, 4, 5, 6  
Array: 7, 8, 9, 10, 2, 1, 3, 4, 5, 6  
myth52:~/lect8$
```

Demo: Array Rotation



rotate.c

Exercise: Array Rotation

Exercise: A properly implemented **rotate** will prompt the following program to generate the provided output.

And here's that properly implemented function!

```
void rotate(void *front, void *separator, void *end) {  
    int width = (char *)end - (char *)front;  
    int prefix_width = (char *)separator - (char *)front;  
    int suffix_width = width - prefix_width;  
  
    char temp[prefix_width];  
    memcpy(temp, front, prefix_width);  
    memmove(front, separator, suffix_width);  
    memcpy((char *)end - prefix_width, temp, prefix_width);  
}
```