

C Wrap-Up, File I/O

CSE 333 Summer 2019

Instructor: Aaron Johnston

Teaching Assistants:

Yifan Bai

Stuart Kol

Daniel Lyu

Kris Wong



About how long did Exercise 4 take?

- A. 0-3 Hours
- B. 3-4 Hours
- C. 4-5 Hours
- D. 5-6 Hours
- E. 6+ Hours
- F. I prefer not to say

Administrivia

- ❖ Exercise 5 out today, due Wednesday morning
- ❖ *No exercise due Friday!* Exercise 6 will be released on Thursday and due the following Monday (7/15)
- ❖ Homework 0 grades will be posted this afternoon
 - Run `git pull` to see feedback
- ❖ Homework 1 due Thursday (7/11) at 11:59 pm
 - Submit via **GitLab** (*i.e.* commit/push changes, then push tag)

Lecture Outline

- ❖ **Visibility of Symbols**
 - `extern, static`
- ❖ File I/O with the C standard library

Namespace Problem

- ❖ If we define a global variable named “counter” in one C file, is it visible in a different C file in the same program?
 - Yes, if you use *external linkage*
 - The name “counter” refers to the same variable in both files
 - The variable is *defined* in one file and *declared* in the other(s)
 - When the program is linked, the symbol resolves to one location
 - No, if you use *internal linkage*
 - The name “counter” refers to a different variable in each file
 - The variable must be *defined* in each file
 - When the program is linked, the symbols resolve to two locations

External Linkage

- ❖ `extern` makes a *declaration* of something externally-visible
 - For variables, tells the compiler not to allocate space for it

```
#include <stdio.h>

// A global variable, defined and
// initialized here in foo.c.
// It has external linkage by
// default.
int counter = 1;

int main(int argc, char** argv) {
    printf("%d\n", counter);
    bar();
    printf("%d\n", counter);
    return 0;
}
```

foo.c

```
#include <stdio.h>

// "counter" is defined and
// initialized in foo.c.
// Here, we declare it, and
// specify external linkage
// by using the extern specifier.
extern int counter;

void bar() {    doesn't create
    counter++;
    printf("(b): counter = %d\n",
           counter);
}
```

bar.c

Internal Linkage

- ❖ `static` (in the global context) restricts a definition to visibility within that file

```
#include <stdio.h>

// A global variable, defined and
// initialized here in foo.c.
// We force internal linkage by
// using the static specifier.
static int counter = 1;

int main(int argc, char** argv) {
    printf("%d\n", counter);
    bar();
    printf("%d\n", counter);
    return 0;
}
```

foo.c

```
#include <stdio.h>

// A global variable, defined and
// initialized here in bar.c.
// We force internal linkage by
// using the static specifier.
static int counter = 100;

void bar() {
    counter++;
    printf("(b): counter = %d\n",
           counter);
}
```

bar.c

Function Visibility

```
// By using the static specifier, we are indicating  
// that foo() should have internal linkage. Other  
// .c files cannot see or invoke foo().
```

```
static int foo(int x) {  
    return x*3 + 1;  
}
```

```
// Bar is "extern" by default. Thus, other .c files  
// could declare our bar() and invoke it.
```

```
int bar(int x) {  
    return 2*foo(x);  
}
```

bar.c

```
#include <stdio.h>
```

```
extern int bar(int x); // "extern" is default, usually omit
```

```
int main(int argc, char** argv) { (for functions)  
    printf("%d\n", bar(5));  
    return 0;  
}
```

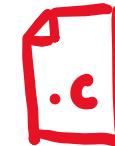
main.c

Linkage Issues

- ❖ Every global (variables and functions) is `extern` by default
 - Unless you add the `static` specifier, if some other module uses the same name, you'll end up with a collision!
 - Best case: compiler (or linker) error
 - Worst case: stomp all over each other



- ❖ It's good practice to:
 - Use `static` to “defend” your globals
 - Hide your private stuff!
 - Place external declarations in a module’s header file
 - Header is the public specification



(maybe: main)

all implementations

* you won't need to use this in CSE 333!
but it's good to be aware of.

Static Confusion...

- ❖ C has a *different* use for the word “**static**”: to create a persistent *local* variable
 - The storage for that variable is allocated when the program loads, in either the `.data` or `.bss` segment
 - Retains its value across multiple function invocations

```
void foo() {  
    static int count = 1;  
    printf("foo has been called %d times\n", count++);  
}  
  
void bar() {  
    int count = 1;  
    printf("bar has been called %d times\n", count++);  
}  
  
int main(int argc, char** argv) {  
    foo(); foo(); bar(); bar(); return 0;  
}  
static_extent.c
```

Additional C Topics

- ❖ Teach yourself!
 - 👉
 - **man pages** are your friend!
 - String library functions in the C standard library
 - `#include <string.h>`
 - `strlen()`, `strcpy()`, `strdup()`, `strcat()`, `strcmp()`, `strchr()`, `strstr()`, ...
 - `#include <stdlib.h>` or `#include <stdio.h>`
 - `atoi()`, `atof()`, `sprint()`, `sscanf()`
 - How to declare, define, and use a function that accepts a variable-number of arguments (`varargs`)
 - unions and what they are good for
 - enums and what they are good for
 - Pre- and post-increment/decrement
 - Harder: the meaning of the “`volatile`” storage class

Lecture Outline

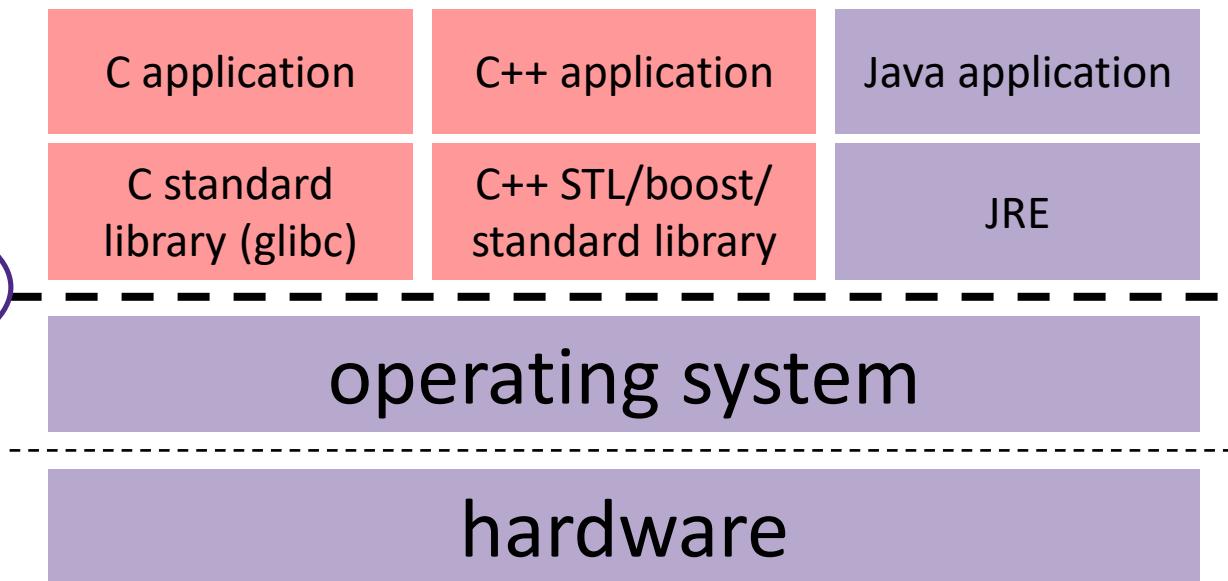
- ❖ Visibility of Symbols
 - `extern`, `static`
- ❖ File I/O with the C standard library

Remember This Picture?

A brief diversion...



HW/SW interface
(x86 + devices)



File I/O

- ❖ We'll start by using C's standard library
 - These functions are part of glibc on Linux
 - They are implemented using Linux system calls
- ❖ C's stdio defines the notion of a stream
 - A way of reading or writing a sequence of characters to and from a device
 - Can be either *text* or *binary*; Linux does not distinguish
 - Is *buffered* by default; libc reads ahead of your program
 - Three streams provided by default: stdin, stdout, stderr
 - You can open additional streams to read and write to files
 - C streams are manipulated with a FILE* pointer, which is defined in stdio.h

an abstraction
for a stream

C Stream Functions

* be sure to peruse the documentation!

- ❖ Some stream functions (complete list in `stdio.h`):

- `FILE* fopen(filename, mode);`
 - Opens a stream to the specified file in specified file access mode
- `int fclose(stream);`
 - Closes the specified stream (and file)
- `int fprintf(stream, format, ...);`
 - Writes a formatted C string

★ – `printf(...);` is equivalent to `fprintf(stdout, ...);`
- `int fscanf(stream, format, ...);`
 - Reads data and stores data matching the format string

C Stream Functions

- ❖ Some stream functions (complete list in `stdio.h`):

- `FILE* fopen(filename, mode);`
 - Opens a stream to the specified file in specified file access mode
 - `int fclose(stream);`
 - Closes the specified stream (and file)
 - `size_t fwrite(ptr, size, count, stream);`
 - Writes an array of *count* elements of *size* bytes from *ptr* to *stream*
 - `size_t fread(ptr, size, count, stream);`
 - Reads an array of *count* elements of *size* bytes from *stream* to *ptr*
- doesn't do formatting (plain bytes)

Error Checking/Handling

- Some error functions (complete list in stdio.h):

- `void perror(message);`

- Prints message followed by error message related to `errno` to `stderr`

uses `errno`

Global

`errno` 12

describes the last
error that
happened in any
library function

- `int ferror(stream);`

- Checks if the error indicator associated with the specified stream is set, returning 1 if so

uses error bit

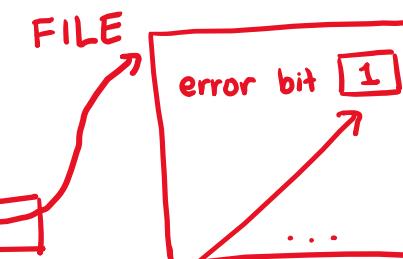
In each Stream

- `int clearerr(stream);`

- Resets error and eof indicators for the specified stream

, `FILE * fin`

| true/false: did an error
| happen in this specific stream?





pollev.com/cse333

Which of the following lines could be inserted to correctly indicate an error?

perror:
print info
ferror:
check for
error

- A. `perror("fwrite failed");` X don't know if an error happened
- B. `perror("fwrite failed");` ✓
- C. `ferror(stream);` X we already know an error happened
`fclose(stream);`
 X errno may have been changed by fclose
- D. `perror("fwrite failed");` return EXIT_FAILURE;
 }
- E. We're lost... ...

C Streams Example

cp_example.c

```
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#define READBUFSIZE 128

int main(int argc, char** argv) {
    FILE *fin, *fout;
    char readbuf[READBUFSIZE];
    size_t readlen;

    if (argc != 3) {
        fprintf(stderr, "usage: ./cp_example infile outfile\n");
        return EXIT_FAILURE;           // defined in stdlib.h
    }

    // Open the input file
    fin = fopen(argv[1], "rb");     // "rb" -> read, binary mode
    if (fin == NULL) {
        perror("fopen for read failed");
        return EXIT_FAILURE;
    }
    ...
}
```

C Streams Example

cp_example.c

```
int main(int argc, char** argv) {  
    ... // previous slide's code  
  
    // Open the output file  
    fout = fopen(argv[2], "wb"); // "wb" -> write, binary mode  
    if (fout == NULL) {  
        perror("fopen for write failed");  
        fclose(fin);  
        return EXIT_FAILURE;  
    }  
  
    // Read from the file, write to fout  
    while ((readlen = fread(readbuf, 1, READBUFSIZE, fin)) > 0) {  
        if (fwrite(readbuf, 1, readlen, fout) < readlen) {  
            perror("fwrite failed");  
            fclose(fin);  
            fclose(fout);  
            return EXIT_FAILURE;  
        }  
    }  
    ... // next slide's code  
}
```

C Streams Example

cp_example.c

```
int main(int argc, char** argv) {  
    ... // two slides ago's code  
    ... // previous slide's code  
  
    // Test to see if we encountered an error while reading  
    if (ferror(fin)) {  
        perror("fread failed");  
        fclose(fout);  
        return EXIT_FAILURE;  
    }  
  
    fclose(fin);  
    fclose(fout);  
  
    return EXIT_SUCCESS;  
}
```

Extra Exercise #1

- ❖ Write a program that:

- Uses argc/argv to receive the name of a text file
- Reads the contents of the file a line at a time
- Parses each line, converting text into a uint32_t
- Builds an array of the parsed uint32_t's
- Sorts the array
- Prints the sorted array to stdout

- ❖ Hint: use man to read about
getline, sscanf, realloc,
and qsort

```
bash$ cat in.txt
1213
3231
000005
52
bash$ ./extra1 in.txt
5
52
1213
3231
bash$
```

Extra Exercise #2

- ❖ Write a program that:
 - Loops forever; in each loop:
 - Prompt the user to input a filename
 - Reads a filename from `stdin`
 - Opens and reads the file
 - Prints its contents to `stdout` in the format shown:

```
00000000 50 4b 03 04 14 00 00 00 00 00 00 9c 45 26 3c f1 d5
00000010 68 95 25 1b 00 00 25 1b 00 00 0d 00 00 00 43 53
00000020 45 6c 6f 67 6f 2d 31 2e 70 6e 67 89 50 4e 47 0d
00000030 0a 1a 0a 00 00 00 0d 49 48 44 52 00 00 00 91 00
00000040 00 00 91 08 06 00 00 00 c3 d8 5a 23 00 00 00 09
00000050 70 48 59 73 00 00 0b 13 00 00 0b 13 01 00 9a 9c
00000060 18 00 00 0a 4f 69 43 43 50 50 68 6f 74 6f 73 68
00000070 6f 70 20 49 43 43 20 70 72 6f 66 69 6c 65 00 00
00000080 78 da 9d 53 67 54 53 e9 16 3d f7 de f4 42 4b 88
00000090 80 94 4b 6f 52 15 08 20 52 42 8b 80 14 91 26 2a
000000a0 21 09 10 4a 88 21 a1 d9 15 51 c1 11 45 45 04 1b
... etc ...
```

- ❖ Hints:
 - Use `man` to read about `fgets`
 - Or, if you're more courageous, try `man 3 readline` to learn about `libreadline.a` and Google to learn how to link to it