# CMSC216: Introduction

Chris Kauffman

Last Updated: Thu Aug 29 09:17:44 AM EDT 2024

# CMSC216 1xx-2xx: Logistics

#### Introductions

- ▶ Prof Kauffman: profk@umd.edu
- ▶ Office Hours Tue 2-3pm / Wed 3-4pm in IRB 2226
- ▶ Slides: Linked from Canvas "Course Schedule/Materials"
- ▶ Static link: https://www.umd.edu/~profk/216/

NOTE: Videos posted for Course Mechanics, Coding Environment Setup posted, first lecture; Lab01 walk-through Video to come

#### Reading

#### **Assignments**

- Due Wed 04-Sep 11:59pm
- ▶ Bryant/O'Hallaron: Ch 1 ▶ C references: basic syntax,
- types, compilation

#### Goals

- ▶ Basic Model of Computation
- ▶ Begin discussion of C
- ▶ Coding Env Setup
- ▶ Lab01: Setup, submit to Gradescope
- ▶ HW01: Basics, online Gradescope Quiz
- How is  $Lab01$  going?

# "Von Kauffman" Model: CPU, Memory, Screen, Program

Most computers have 4 basic, physical components<sup>1</sup>

- 1. CPU: can execute "instructions"
- 2. CONTROL: CPU knows WHICH instruction to execute
- 3. MEMORY: data is stored and can change
- 4. Some sort of Input/Output device like a SCREEN (optional)

CPU understands some **set of instructions**; a sequence of instructions is a **program** that changes MEMORY and SCREEN

#### Example of a Running Computer Program



 $1$ Of course it's a *little* more complex than this but the addage, "All models are wrong but some are useful" applies here. This class is about asking "what is really happening?" and going deep down the resulting rabbit hole.

## Sample Run Part 1

CPU: at instruction 10: MEMORY: SCREEN:  $> 10:$  set #1024 to 195  $11:$  set  $\#1028$  to 21 12: sum #1024.#1028 into #1032 13: print #1024, "plus", #1028 14: print "is", #1032

CPU: at instruction 11: MEMORY: SCREEN: 10: set #1024 to 195  $> 11:$  set #1028 to 21 12: sum #1024,#1028 into #1032 13: print #1024, "plus", #1028 14: print "is", #1032

CPU: at instruction 12: MEMORY: SCREEN:  $10:$  set #1024 to 195  $11:$  set  $\#1028$  to 21  $> 12:$  sum #1024,#1028 into #1032 13: print #1024, "plus", #1028

#### 14: print "is", #1032







## Sample Run Part 2



12: sum #1024,#1028 into #1032 | #1032 | 216 | 13: print #1024, "plus", #1028 | #1028 | 21 |

14: print "is", #1032 | #1024 | 195 |

 $> 15: . . . .$ 



 $5$  plus  $21$ 

MEMORY: SCREEN: | Addr | Value | 195 plus 21

# Observations: CPU and Program Instructions

- ▶ Program instructions are usually small, simple operations:
	- ▶ Put something in a specific memory cell using its **address**
	- ▶ Copy the contents of one cell to another
	- ▶ Do arithmetic  $(+, -, *, /)$  on cells or constants
	- ▶ Print stuff to the screen
- ▶ The CPU keeps track of which instruction to execute next
- ▶ After executing an instruction, CPU advances to next instruction BUT **jumping** around to distant instructions is also possible: conditional and iterative execution
- ▶ Previous program is in **pseudocode** in which instructions can have any meaning understood by a human reader<sup>2</sup>
- ▶ Real machines require more precise instruction definitions as there are no smart humans to interpret them, only dumb physics to blindly execute them

<sup>2</sup>The pseudcode shown resembles a low-level **assembly language** rather than a high level language like C or Java

# Observations: Memory Cells and the Screen

## Memory Cells

- ▶ Memory cells have Fixed **ADDRESS** Changeable **CONTENTS**
- ▶ Random Access Memory (RAM): the value in any memory cell can be retrieved FAST using its address
- ▶ My laptop has 16GB of memory  $= 4,294,967,296$ (4 billion) integer boxes (!)
- $\blacktriangleright$  Cell Address  $\#$ 's never change: always cell  $#1024$
- ▶ Cell Contents / Values often change: set #1024 to 42

## Screen versus Memory

- ▶ Nothing is on the screen until it is explicitly print-ed by the program
- ▶ Don't get to see memory while the program runs: **print stuff while debugging programs so you can see it**
- ▶ Forming a mental model of what values are in memory and how they relate to one another is a valuable skill which we will practice, often by drawing memory explicitly

# Variables are Named Memory Cells

- $\triangleright$  Dealing with raw memory addresses is tedious
- ▶ Any programming language worth its salt will have **variables**: symbolic names associated with memory cells
- ▶ **You pick variable names**; compiler/interpreter automatically translates to memory cell/address

```
PROGRAM ADDRESSES ONLY
CPII: at instruction 50: MEMORY:
> 50: copy #1024 to #1032
  51: copy #1028 to #1024
  52: copy #1032 to #1028
  53: print "first",#1024
  54: print "second",#1028 | #1024 | 42 |
```
PROGRAM WITH NAMED CELLS MEMORY:  $CPU:$  at instruction  $51:$  $> 50:$  copy x to temp  $51:$  copy y to x  $52:$  copy temp to  $y$ 53:  $print$  "first", $x$ 54: print "second",y





# Correspondence of C Programs to Memory

- ▶ C programs require memory cell names to be declared with the **type of data** they will hold (a novel idea when C was invented).
- $\blacktriangleright$  The equal sign (=) means

"store the result on the right in the cell named on the left"

 $\triangleright$  Creating a cell and giving it a value can be combined

**int** x; // need a cell named x, holds an integer  $x = 42$ ;  $// put 42 in cell x$ **int**  $y = 31$ ; // need a cell named y and put 31 in it **int tmp** =  $x + y$ ; // cell named tmp, fill with sum of x and y

### Other Rules

- $\triangleright$  C/Java compilers read whole functions to figure out how many memory cells are needed based on declarations like int a; and int c=20;
- ▶ Lines that only declare a variable do nothing except indicate a cell is needed to the compiler
- ▶ In C, uninitialized variables may have arbitrary crud in them making them dangerous to use: we'll find out why in this course

# Exercise: First C Snippet

- $\triangleright$  Lines starting with // are comments, not executed
- $\triangleright$  printf("%d %d\n",x,y) shows variable values on the screen as decimal integers



#### **With your nearby colleagues:**

- 1. Show what memory / screen look like after running the program
- 2. Correct the program if needed: make swapping work

I will chat with a couple folks about their answers which will earn participation credit leading to **Bonus Engagement Points**.

## **Answer**: First C Snippet

CPU: at line 54 MEMORY: SCREEN: 50: int x;  $\begin{array}{ccc} 1 & 4 \text{ ddr} & 4 \text{ Name} & 0 \text{ Value} \\ 51: x = 42: & 0 \end{array}$ 51:  $x = 42$ ;<br>
52: int  $y = 31$ ;<br>
52: int  $y = 31$ ;<br>
51:<br>
51:<br>
531; 52: int  $y = 31$ ;  $\frac{141032 + y}{141032 + x}$  | 31 | 53: // swap x and y (?)  $\frac{141028 + x}{14028 + x}$  | 42 | | #1028 | x | 42 |
| #1024 | |  $> 54: x = v:$ 55:  $v = x$ : 56: printf("%d %d\n",x,y); CPU: at line 55 MEMORY: SCREEN:<br>50: int x: <br>and addr | Name | Value | 50: int x;  $\vert$  Addr  $\vert$  Name  $\vert$  Value  $\vert$ <br>51: x = 42;  $\vert$  -----------------------51:  $x = 42$ ;<br>
52: int  $y = 31$ ;<br>
52: int  $y = 31$ ;<br>
53:  $\frac{1}{2}$ <br>
53:  $\frac{1}{2}$ <br>
53:  $\frac{1}{2}$  $1 \# 1032 \mid v \mid 31 \mid$ 53: // swap x and y (?) | #1028 | x | 31 |  $54: x = y;$  | #1024 |  $> 55: v = x:$ 56: printf("%d %d\n",x,y); CPU: at line 57 MEMORY: SCREEN: 50: int x;  $\begin{array}{ccc} 50: & \text{int } x; \\ 51: & x = 42: \end{array}$  | Addr | Name | Value | 31 31 51:  $x = 42$ ;<br>
52: int  $y = 31$ ;<br>
52: int  $y = 31$ ;<br>
51:<br>
51:<br>
531:  $\frac{141032}{x}$  | 31 53: // swap x and y (?) | #1028 | y | 31 |  $54: x = v:$ 55:  $v = x$ : 56: printf("%d %d\n",x,y);  $> 57:$  ...

Clearly **incorrect**: how does one swap values properly? (fix swap\_main\_bad.c)

## First Full C Program: swap\_main.c



- ▶ Swaps variables using tmp space (exotic alternatives exist)
- ▶ Executables always have a main() function: starting point
- ▶ Note inclusion of stdio.h **header** to declare printf() exists, allusions to C's (limited and clunky) library system

Exercise: Functions in C, swap func.c

```
1 // C program which attempts to swap using a function.
 2 //
 3 // > gcc swap_func.c
 4 // > ./a.out5<br>6 #include <stdio.h>
                                        // declare existence printf()
 7 void swap(int a, int b); // function exists, defined below main
 8
 9 int main(int argc, char *argv[]){ // ENTRY POINT: start executing in main()
10 int x = 42;
11 int y = 31;<br>12 swap(x, y);
12 swap(x, y); \frac{1}{2} swap(x, y); \frac{1}{2} invoke function to swap x/y (?)<br>13 printf("%d %d\n",x,y); \frac{1}{2} print the values of x and y
                                           13 13 13 print the values of x and y
14 return 0;
15 }
16
17 // Function to swap (?) contents of two memory cells
18 void swap(int a, int b){ // arguments to swap
19 int tmp = a; \frac{1}{20} \frac{1}{20} a = b: \frac{1}{20} \frac{1}{4} \frac{1}{6} \frac{1}{6} \frac{1}{20} \frac{1}{20} \frac{1}{20} a \frac{1}{20} b
20 a = b; \frac{1}{2} // a <- b<br>21 b = tmp; \frac{1}{2} // b <- tr
                                            1/ b < - \text{tmp}=a22 return;
23 }
```
Does swap() "work"? Discuss with neighbors and justify why the code works or why not

# **Answers**: Swapping in a Function is Tricky

#### swap\_func.c will not print swapped values

- $\blacktriangleright$  If you thought the values would print swapped, you're about to learn something interesting
- ▶ If you were confident they would not print swapped but had difficulty articulating why, that's great: this class is here to give the vocab to do so
- ▶ If you knew the values wouldn't swap and also knew how to explain it well, tune in anyway as the subsequent explanation will introduce conventions used for the rest of the course

### Why No Swap??

Necessitates introducing the **Function Call Stack** which is where functions store their local variables and parameters

## **Answers**: The Function Call Stack and swap()



- ▶ **Caller function** main() and **Callee function** swap()
- ▶ Caller **pushes** a stack frame onto the **function call stack**
- ▶ Frame has space for All Callee parameters/locals vars
- ▶ Caller tracks where it left off to resume later
- ▶ Caller copies values to Callee frame for parameters
- $\blacktriangleright$  Callee begins executing at its first instruction

## **Answers**: Function Call Stack: Returning from swap()



- ▶ On finishing, Callee stack frame **pops** off, Control returns to Caller which resumes executing next instruction
- ▶ Callee may pass a return value to Caller but otherwise does not directly affect Caller stack frame on return
- $\triangleright$  swap() does NOT swap the variables  $x,y$  in main(), only its own local variables a,b

## Motivation for C

# Pure Abstraction



If this were Java, Python, many others, discussion would be over:

- ▶ Provide many safety and convenience features
- ▶ Insulate programmer from hardware for ease of use

C presents many CPU capabilities directly

- ▶ Very few safety features
- ▶ Little between programmer and hardware

You just have to know C. Why? Because for all practical purposes, every computer in the world you'll ever use is a **von Neumann machine**, and C is a lightweight, expressive syntax for the von Neumann machine's capabilities. –Steve Yegge, Tour de Babel <sup>17</sup>

# Von Neumann Machine Architecture (Wikip)

#### Processing

# $\blacktriangleright$  Wires/gates that accomplish

- Memory
- fundamental ops
- $\blacktriangleright$  +, -, \*, AND, OR, move, copy, shift, etc.
- ▶ Ops act on contents of memory cells to change them

#### **Control**

- ▶ Memory address of next instruction to execute
- $\blacktriangleright$  After executing, move ahead one unless instruction was to jump elsewhere

- ▶ Giant array of bits/bytes so **everything** is represented as 1's and 0's, including instructions
- ▶ Memory cells accessible by address number

### Input/Output

- ▶ Allows humans to interpret what is happening
- ▶ Often special memory locations for screen and keyboard

Wait, these items seem kind of familiar.  $\ldots$  18

## Exercise: C allows direct use of memory cell addresses



Where/how are these used in the code below?

```
1 // swap_pointer.c: swaps values using a function with pointer arguments.
\mathfrak{D}3 #include <stdio.h> // declare existence printf()
4 void swap_ptr(int *a, int *b); // function exists, defined below main
5
6 int main(int argc, char *argv[]){ // ENTRY POINT: start executing in main()
7 int x = 42;
8 int y = 31;
9 swap_ptr(\&x, \&y); // call swap() with addresses of x/y<br>10 printf("\&d \&d \&x",x,y); // print the values of x and y
                                    // print the values of x and y
11 return 0;
12 }
13
14 // Function to swap contents of two memory cells
15 void swap_ptr(int *a, int *b){ // a/b are addresses of memory cells<br>16 int tmp = *a: // go to address a, copy value int to
16 int tmp = *\mathbf{a};<br>
17 *\mathbf{a} = *b;<br>
17 *\mathbf{a} = *b;<br>
17 *\mathbf{a} = *b;
                                      1/ copy val at addr in b to addr in a
18 *b = tmp; \frac{1}{2} // copy temp into address in b
19 return;
20 } \hspace{1.5cm} 19
```
# Swapping with Pointers/Addresses: Call Stack



- ▶ Syntax &x reads "Address of cell associated with x" or just "Address of x". Ampersand & is the address-of operator.
- ▶ Swap takes int \*a: **pointer** to integer / memory address
- ▶ Values associated with a/b are the addresses of other cells

# Swapping with Pointers/Addresses: Dereference/Use

```
9: int main(...) { \qquad LINE 19 executed: tmp gets 42 \qquad 0: int x = 42; \qquad | FRAME | ADDR | NAME | VA
 10: int x = 42; | FRAME | ADDR | NAME | VALUE |
 11: int y = 31;<br>
12: swap ptr(&x, &y);<br>
1 main() | #2048 | x | 42 |
 12: swap_ptr(&x, &y); <br>
13: printf("%d %d\n",x,y); <br>
1 line:12 | #2044 | y | 31 |<-|-
 13: printf("%d %d\n",x,y); <br>14: return 0: <br>14: return 0: <br>1
 14: return 0; |---------+-------+------+-------| ||
                                      | swap ptr| #2036 | a | #2048 |--+|
                                      | line:20 | #2028 | b | #2044 |---+<br>| | #2024 | tmp | ?->42 |
 18: void swap ptr(int *a,int *b){ |
 19: int tmp = *a; // copy val at #2048 to #2024
>20: *a = *b:
21: *b = tmp:
22: return;
 23: }
```
- ▶ Syntax \*a reads "Dereference a to operate on the cell pointed to by a" or just "Deref a"
- ▶ Line 19 dereferences via \* operator:
	- ▶ Cell  $#2036$  (a) contains address  $#2048$ ,
	- ▶ Copy contents of  $\#2048$  (42) into  $\#2024$  (tmp)

# Swapping with Pointers/Addresses: Dereference/Assign

```
9: int main(...) { LINE 20 executed: alters x using a<br>0: int x = 42: \qquad | FRAME | ADDR | NAME | VALUE |
 10: int x = 42; | FRAME | ADDR | NAME | VALUE |
 11: int y = 31; |---------+-------+------+-------|
                                          | main() | #2048 | x   |42->31 |<-+<br>| line:12 | #2044 | y   |   31 |<-|+
 13: printf("%d %d\n",x,y);<br>14: return 0;
 14: return 0; |---------+-------+------+-------| ||
                                          | swap_ptr| #2036 | a | #2048 |--+|
                                          1ine:21 | #2028 | b | #2044 |---+<br>| #2024 | tmp | 42 |
 18: void swap_ptr(int *a,int *b){ | | #2024 | tmp | 42 | 19: int tmp = *a:
 19: int tmp = *a;<br>20: *a = *b;
                        \frac{1}{2} copy val at #2044 (31) to #2048 (was 42)
>21: *b = tmp:
 22: return;
 23: }
```
- ▶ Pointer Deref on Right Side **fetches** a value from a pointer location
- ▶ Pointer Deref on Left Side stores a value at a pointer location
- ▶ Line 20: Deref on both Left and right side of assignment
	- ▶ a and b contain pointers, not changed
	- $\triangleright$  x and y are pointed at, can change

## Swapping with Pointers/Addresses: Deref 2



▶ Line 21: dereference on left-hand side

 $*b = \ldots$ 

stores new value at address #2044

▶ Use of variable **bare name** always retrieves value it that cell

▶ tmp retrieves an int like 42

▶ a retrieves a pointer like #2048

# Swapping with Pointers/Addresses: Returning

```
9: int main(...) { LINE 22: prior to return<br>10: int x = 42: \qquad \qquad | FRAME | ADDR | NAME
   10: int x = 42; <br>11: int y = 31; <br>11: int y = 31; <br>11: int y = 31;
                                            |----------+-------+------+--------|<br>| main()  | #2048 | x     |     31 |<-+
12: \text{swap\_ptr}(\&x, \&y);<br>+->13: \text{print}("\&d \&d\n",x,y);13: printf("%d %d\n",x,y); <br>14: return 0: <br>14: return 0: <br>1
   | 14: return 0; |---------+-------+------+-------| || | | | | |
                                            | swap_ptr| #2036 | a | #2048 |--+|
                                             | line:22 | #2028 | b | #2044 |---+<br>|        | #2024 | tmp |     42 |
   | 18: void swap_ptr(int *a,int *b){ | | #2024 | tmp | 42 |
   19: int tmp = *a;<br>20: *a = *b:
   20: *a = *b;<br>
21: *b = tmp;<br>
21: kb = tmp;<br>
21: kb = temp;<br>
21: kb = temp;
| 21: *b = tmp; | FRAME | ADDR | NAME | VALUE |
   +-<22: return; |---------+-------+------+-------|
                                            23: } | main() | #2048 | x | 31 |
                                              line: 13 \mid #2044 \mid y \mid 42 \mid|---------+-------+------+-------|
```


 $\triangleright$  Variables x, y in main() have changed due to use of references to them.

Aside: Star/Asterisk \* has **3 uses** in C

1. Multiply numbers as in

 $w = c*d$ :

2. **Declare** a pointer variable as in

int \*x; // pointer to integer(s) int b=4;  $x = kb$ ; // point x at b int  $**r$ ; // pointer to int pointer(s)

3. **Dereference** a pointer variable as in

int  $p = \frac{2}{x}$ ; // x must be an int pointer // retrieve contents at address

Three different context sensitive meanings for the same symbol makes \* hard on humans to parse, a BAD move by K&R.

int  $z = \ast x \ast \ast y + \ast (p+2)$ ; // standard, 'unambiguous' C The duck is ready to eat. // English is more ambiguous

# Some Common Examples and Errors

- ▶ Learning syntax and semantics of pointers requires some practice, get started with below examples
- ▶ Won't go through these in much detail YET but over next couple weeks will discuss at length

```
// pointer_examples.c
// 1: proper pointer assignment
int a4 = 17;
int a1 = 11;
int *p1 = & a1; // cool
int b1 = 55;
p1 = %b1; // cool
// 2: improper pointer assignment
int *p5 = &a5;
int a2 = 13;
int *p2 = a2; // ERROR
// 3: proper pointer copying
int a3 = 15;
int *p3 = &a3;
int *q3 = p3; // cool
                                // 4: proper pointer deref
                                int *p4 = %a4;int b4 = *p4; // cool// 5: improper int assign (no deref)
                                int a5 = 19;
                                int b5 = p5; // ERROR
```
# Important Principle: Non-local Changes

- ▶ Pointers allow functions to change variables associated with other running functions
- ▶ Common beginner example: scanf() family which is used to read values from terminal or files

```
▶ Snippet from scanf demo.c
```

```
1 int main(...){
2 int num = -1;
3 scanf("%d", &num); // addr
4 printf("%d\n",num); // val
4 return 0;
5 }
```
▶ See scanf\_error.c : forgetting & yields great badness



# Uncle Ben Said it Best. . .



All of these apply to our context..

- ▶ **Pointers allow any line of C programs to modify any of its data**
- ▶ A BLESSING: fine control of memory *→* efficiency, machine's true capability
- ▶ A CURSE: opens up many **errors** not possible in Java/Python which restrict use of memory

1972 - Dennis Ritchie invents a powerful gun that shoots both forward and backward simultaneously. Not satisfied with the number of deaths and permanent maimings from that invention he invents C and Unix. – A Brief, Incomplete, and Mostly Wrong History of Programming Languages

# Beneath the C

C is "high-level" as it abstracts away from a real machine. It must be translated to lower levels to be executed.

### Assembly Language

- ▶ Specific to each CPU architecture (Intel, etc)
- $\blacktriangleright$  Still "human readable" but fairly directly translated to binary using Assemblers



## Binary Opcodes

 $\blacktriangleright$  1's and 0's, represent the digital signal of the machine

```
▶ Codes corresponds to
                                     instructions directly
                                      understood by processor
INTEL x86-64 ASSEMBLY HEXADECIMAL/BINARY OPCODES
                                1124: 83 f9 011127: 7e 1e = 0111 1110 0001 1110
                                1129: be 02 00 00 00
                                112e: 89 c81130: 48.991132: f7 f91134: 83 fa 01
                                1137: 75 07
```
Looks like **fun**, right? You bet it is! Assembly coding is 6 weeks away. . .

# CMSC216: Course Goals

- ▶ Basic proficiency at C programming
- ▶ Knowledge of running programs in physical memory including the stack, heap, global, and text areas of memory
- ▶ Understanding of the essential elements of assembly languages
- ▶ Knowledge of the correspondence between high-level program constructs.
- ▶ Ability to use a symbolic debugger
- ▶ Basic understanding of how data is encoded in binary
- $\triangleright$  Understanding the process abstraction of running programs, ability to create and manipulate processes
- ▶ Basic understanding of execution threads, their relation to processes, the ability to create and manipulate threads