# C programming language

College of Saint Benedict & Saint John's University

# origins



Dennis Ritchie in 2011 / CC BY 2.0



Brian Kernighan in 2012 / CC BY 2.0

- Dennis Ritchie and Brian Kernighan creators of C circa 1972
- TODO: more thorough history

1

## hello, world

```
/* file: helloworld.c */

#include <stdio.h>

int main() {
   printf("hello, world\n");
   return 0;
}
```

```
$ gcc -o helloworld helloworld.c
$ ./helloworld
hello, world
```

- Remind students that not everyone has same background in C those with experience can help those without
- The tradition of using the phrase "Hello, world!" as a test message was influenced by an example program in the seminal book *The C Programming Language*
- Every statement in C exists in a function, starting with main
- · Refresh students memory on the *traditional* compilation process

### global variables

```
$ gcc -o figure2-4 figure2-4.c
$ ./figure2-4
M 419
N
424
```

Every C variable has three attributes:
 name an identifier determined arbitrarily by the programmer
 type specifies the kind of values it can have

#### value

- In C, variable declaration only reserves storage for the value; nothing can be assumed about the initial value
- What would you expect for input 'Z -3'?
- · What would you expect for input '9 a'?
- What would you expect for input '~ 2147483643'?

```
global variables are
declared here —
outside of any function

characters in C are
treated internally
like signed integers

#include <stdio.h>

char ch;
int j;

int main() {
    scanf("%c %d", &ch, &j);
    j += 5;
    rintf("%c\n%d\n", ch, j);
    return 0;
}
```

```
read data from stdin (the terminal)

print data to stdout (the terminal)
```

• C has no "built-in" functions; however, it does have a standard library that includes many useful utility functions.

```
#include <stdio.h>

char ch;
int j;

int main() {
    scanf("%c %d", &ch, &j); <----
    j += 5;
    ch++;
    printf("%c\n%d\n", ch, j);
    return 0;
}</pre>

**is the address of
    operator - scanf
    expects the address
    of the variables where
    the data will be stored

**include <stdio.h>

**include <stdio.h

**include <stdio.h>

**include <stdio.h

**include <stdio
```

· address here means the location in memory

#### conditions

```
if (<cond>) {
    /* ... */
}
else (<cond>) {
    /* ... */
}
else {
    /* ... */
}
```

- · C only expects that the expression yield an integer value
- 0 is false, non-zero is true

#### conditions

```
if (x) {
    /* ??? */
}

if (x-y) {
    /* ??? */
}

if (x=y) {
    /* ??? */
}

/* ??? */
}
```

· under what conditions will each of the above be executed?

- x != 0
- x != y
- y != 0

### switch

```
switch (<expr>) {
 break;
 default:
```

## loops

```
while (<cond>)
 while (<cond>);
```

- <init> is for initializing variables once, before loop begins execution
   normally the loop control variable
- Same rules apply for <cond>.
- <incr> is for assigning variables a value, happens once after the body
  of the loop has executed each iteration again, normally the loop
  control variable

### memory model — part i

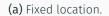
#### global variables

declared outside of any function and remain in place throughout the execution of the entire program. they are stored at a fixed location in memory.

#### local variables

declared within a function and come into existence when the function is called and cease to exist when the function terminates. they are stored on the run-time stack.







(b) Run-time stack.

- I will be using graphical notation consistent with that of the book.
- In this case, (a) and (b) represent the state of relevant memory for the previous program just before it terminates, i.e., in the process of executing line 15.

8

#### run-time stack a.k.a. "the stack"

#### run-time stack

stores information about the active functions of a C program, including:

- · return value.
- · actual parameters,
- · return address, and
- local variables

in that order.

- I will not distinguish between functions and procedures
  - functions that have no return value, i.e., return type void, just omit the first bullet point
- each program has one stack

9

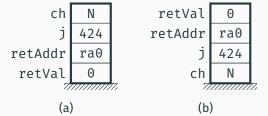
#### run-time stack a.k.a. "the stack"

#### run-time stack

stores information about the active functions of a C program, including:

- · return value,
- actual parameters,
- · return address, and
- local variables

in that order.



- I will not distinguish between functions and procedures
  - functions that have no return value, i.e., return type void, just omit the first bullet point
- · each program has one stack

9

 which visualization is correct, had ch and j been declared as local variables instead of global?

### functions

```
return value type

- --> void print_bar (int n) {

int k;

for (k=0; k<n; k++) {

printf("*");

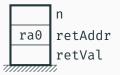
printf("\n");

printf("\n");
```

#### functions

```
list of formal
                                         parameters
                             function name
return value type
                      r> void print_bar (int n) {
return value type
                      +> int fact(int n) {
                                                                  type of <expr>
                           return f; <---------
                                                                 must match return
                                                                  type of function
```

```
int fact(int n) {
```



- you will have to reconcile this example a little bit with the book, since the solution is written slightly different
- assume nothing about values that are blank they are not 0

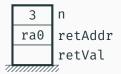
11

```
int fact(int n) {
```

```
3 n
ra0 retAddr
retVal
```

· assume the user inputs the number 3

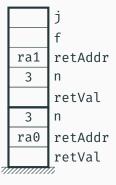
```
5 int fact(int n) {
6   int f, j;
7   f = 1;
8   for (j=1; j<=n; j++) {
9     f *= j;
10  }
11   return f;
12 }
13
14  int main() {
15   int n;
16   scanf("%d", &n);
17   printf("%d\n", fact(n)); // ra1
18   return 0;
19 }</pre>
```



• this statement requires that we evaluate the expression fact(n)

```
int fact(int n) {
   int f, j;
   f = 1;
   for (j=1; j<=n; j++) {
    f *= j;
   }
   return f;
}

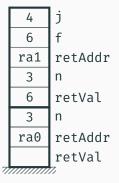
int main() {
   int n;
   scanf("%d", &n);
   printf("%d\n", fact(n)); // ra1
   return 0;
}</pre>
```



- fact(n) is a function call, so a new stack frame is pushed on to the stack in the sequence described before:
  - return value
  - actual parameters
  - return address
  - local variables

```
int fact(int n) {
   int f, j;
   f = 1;
   for (j=1; j<=n; j++) {
    f *= j;
   }
   return f;
}

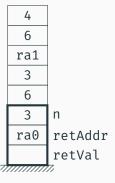
int main() {
   int n;
   scanf("%d", &n);
   printf("%d\n", fact(n)); // ra1
   return 0;
}</pre>
```



 the state of the run-time stack after fact function has completed, but before returning

```
int fact(int n) {
   int f, j;
   f = 1;
   for (j=1; j<=n; j++) {
    f *= j;
   }
   return f;
}

int main() {
   int n;
   scanf("%d\n", &n);
   printf("%d\n", fact(n)); // ra1
   return 0;
}</pre>
```

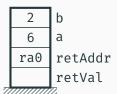


 after fact returns, its stack frame is deallocated, however, the values computed are still in memory — since return value is always first thing pushed on to stack, the main function knows exactly where to find the value returned by fact(n)

```
4 j
6 f
ral retAddr
3 n
6 retVal
3 n
ra0 retAddr
retVal
```

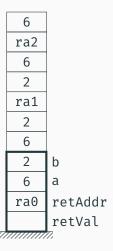
- assume the the first call to fact got lucky and the value 1 happended to be in the memory cell designated to the variable f, so that the calculation worked out correctly
- this would be the state of the stack at the beginning of execution of the second call to fact

## functions — call-by-"reference"



- in C, parameters are ALWAYS call-by-value, if you want the behavior of call-by-reference, then you pass the address of the variable instead of its value
- but this is still call-by-value, its just that now, the value happens to represent an address of a value
- you will have to reconcile this example a little bit with the book, since the solution is written slightly different
- · assume the user inputs the numbers 6 & 2
- what will the stack look like after the call to order, but before the call to printf?

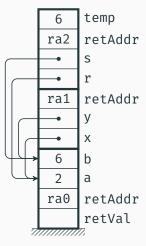
## functions — call-by-"reference"



- in C, parameters are ALWAYS call-by-value, if you want the behavior of call-by-reference, then you pass the address of the variable instead of its value
- but this is still call-by-value, its just that now, the value happens to represent an address of a value
- you will have to reconcile this example a little bit with the book, since the solution is written slightly different
- assume the user inputs the numbers 6 & 2
- what will the stack look like after the call to order, but before the call to printf?

### functions — call-by-"reference"

```
void swap(int *r, int *s) {
```



- this is what the stack will look like just before returning from the swap function
- now, instead of passing the values **a** and **b**, we will pass their addresses, so that we can update their values, i.e., pass them by "reference"

# pointers

• a pointer is a variable whose value is a memory address

```
int i = 0x1A;
int *ip = &i;
```

- $\upbeta \mathbf{i}$  evaluates to the address where the variable  $\mathbf{i}$  is stored in memory
- i is an int, so ip is a pointer to an int

### pointers cont.

```
printf("0x%X\n", i);  /* 0x1A */
printf("0x%#X\n", &i); /* 0x12A0 */
printf("0x%#X\n", ip); /* 0x12A0 */
printf("0x%#X\n", &ip); /* 0x???????? */
```

• so how can we use the pointer, ip, to access the value of i?

### pointer dereference

- \*ptr will
  - 1. treat the value of **ptr** as a memory address
  - 2. get the bytes of data located at that memory address
  - 3. interpret those bytes according to the type of pointer that **ptr** is

```
printf("0x<mark>%X\n", *ip);</mark> /* 0x1A */
```

the C compiler is "smart enough" to "know" that + X really means add
 X \* sizeof(\*ip) to ip

### pointer dereference

- \*ptr will
  - 1. treat the value of ptr as a memory address
  - 2. get the bytes of data located at that memory address
  - 3. interpret those bytes according to the type of pointer that **ptr** is

```
printf("0x%X\n", *ip); /* 0x1A */
```

 $\cdot ip[X] = *(ip + X)$ 

printf("0x%X\n", ip[0]); /\* 0x1A \*/

the C compiler is "smart enough" to "know" that + X really means add
 X \* sizeof(\*ip) to ip

### pointers cont.

pointers cont.

```
· cp is a pointer to a char
0x00004C80 | h | e | l | l | o | , |
                                | w | o | r |
0x??????? | 00 | 00 | 4C | 80 |
   printf("%c\n", cp[0]); /* h */
   printf("%c\n", cp[4]); /* o */
   printf("%c\n", *(cp+4)); /* 0 */
   printf("0x%#X\n", cp); /* 0x4C80 */
```

• why not say **cp** is a *pointer* to a **char** array?

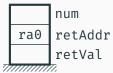
#### recursion

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

#### recursion

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

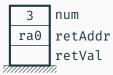
```
$ ./figure2-22
```



#### recursion

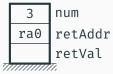
```
#include <stdio.h>
```

```
$ ./figure2-22
Enter a small integer: 3
```



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

```
$ ./figure2-22
Enter a small integer: 3
```



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

```
$ ./figure2-22
Enter a small integer: 3
```

	_
ra1	retAddr
3	n
	retVal
3	num
ra0	retAddr
	retVal
///////////////////////////////////////	7.

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

```
$ ./figure2-22
Enter a small integer: 3
```

```
ra1 retAddr
n retVal
num
ra0 retAddr
retVal
```

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

```
$ ./figure2-22
Enter a small integer: 3
```

		_
	ra2	retAddr
	2	n
		retVal
	ra1	retAddr
	3	n
		retVal
	3	num
	ra0	retAddr
		retVal
7	///////////////////////////////////////	77.

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

```
$ ./figure2-22
Enter a small integer: 3
```

```
retAddr
 ra2
     retVal
 ra1
     retAddr
  3
     n
     retVal
  3
     num
 ra0
     retAddr
     retVal
```

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

```
$ ./figure2-22
Enter a small integer: 3
```

	ra2	retAddr
	1	n
		retVal
	ra2	retAddr
	2	n
		retVal
	ra1	retAddr
	3	n
		retVal
	3	num
	ra0	retAddr
		retVal
7		7.

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

```
$ ./figure2-22
Enter a small integer: 3
```

```
retAddr
 ra2
      retVal
 ra2
     retAddr
      retVal
     retAddr
 ra1
  3
      n
      retVal
  3
      num
 ra0
     retAddr
      retVal
```

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

```
$ ./figure2-22
Enter a small integer: 3
```

	ra2	
	1	
	1	
	ra2	retAddr
	2	n
	2	retVal
	ra1	retAddr
	3	n
		retVal
	3	num
	ra0	retAddr
		retVal
7	///////////////////////////////////////	7.

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

```
$ ./figure2-22
Enter a small integer: 3
```

	ra2	
	1	
	1	
	ra2	
	2	
	2	
	ra1	retAddr
	3	n
	6	retVal
	3	num
	ra0	retAddr
		retVal
7	///////////////////////////////////////	77,

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

```
$ ./figure2-22
Enter a small integer: 3
Its factorial is: 6
```

```
ra2
 ra2
 ra1
  3
  6
  3
     num
 ra0
     retAddr
     retVal
```

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

```
$ ./figure2-22
Enter a small integer: 3
Its factorial is: 6
```

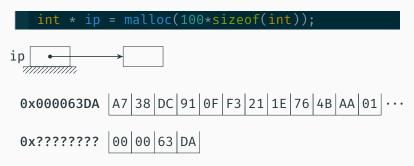
```
ra2
 ra2
 ra1
  3
  6
  3
     num
 ra0
     retAddr
  0
     retVal
```

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

```
$ ./figure2-22
...
$
```

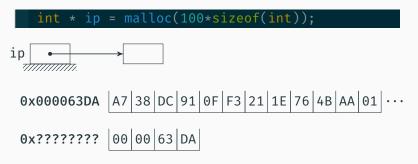
## heap memory

 designate a block of memory to store value(s) of a particular data type



## heap memory

 designate a block of memory to store value(s) of a particular data type



 $\cdot$  release a block of memory back to system to be used elsewhere

# free(ip);

## heap memory cont.

## heap memory cont.

```
ip[0] = 0x7; /* *ip = 0x7; */
0x000063DA | 00 | 00 | 07 | 0F | F3 | 21 | 1E | 76 | 4B | AA | 01 | · · ·
0x??????? | 00 | 00 | 63 | DA |
0x000063DA | 00 | 00 | 07 | 00 | 00 | 00 | 0A | 76 | 4B | AA | 01 | · · ·
```

#### structures

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

#### structures cont.

# comparison

Java	С
object-oriented	procedural
interpreted	compiled
String	<b>char</b> array
condition (boolean)	condition (int)
garbage-collected	no memory management
references	pointers
exceptions	error codes

- in Java, everything is a method that is called on an object
- · in C, everything is a function
- in Java, source code is compiled to byte code, which is then interpreted by Java VM
- in C, source code is compiled into binary machine code
- in Java, String is a class
- in C, a string is just an array of char values which ends with the char '\0'
- in Java, the Java VM takes care of deallocating memory used
- $\cdot$  in C, any memory you allocate, you must also deallocate



except where otherwise noted, this worked is licensed under creative commons attribution-sharealike 4.0 international license