C (for those who know Java)

Jingpeng Li
http://www.cs.stir.ac.uk/~jli/

Why Learn C?

- Likely doing systems or general track
- Want to do well in OS, DB, Networks, …
- Interested in how real systems work
- Affinity for programming, not complexity theory

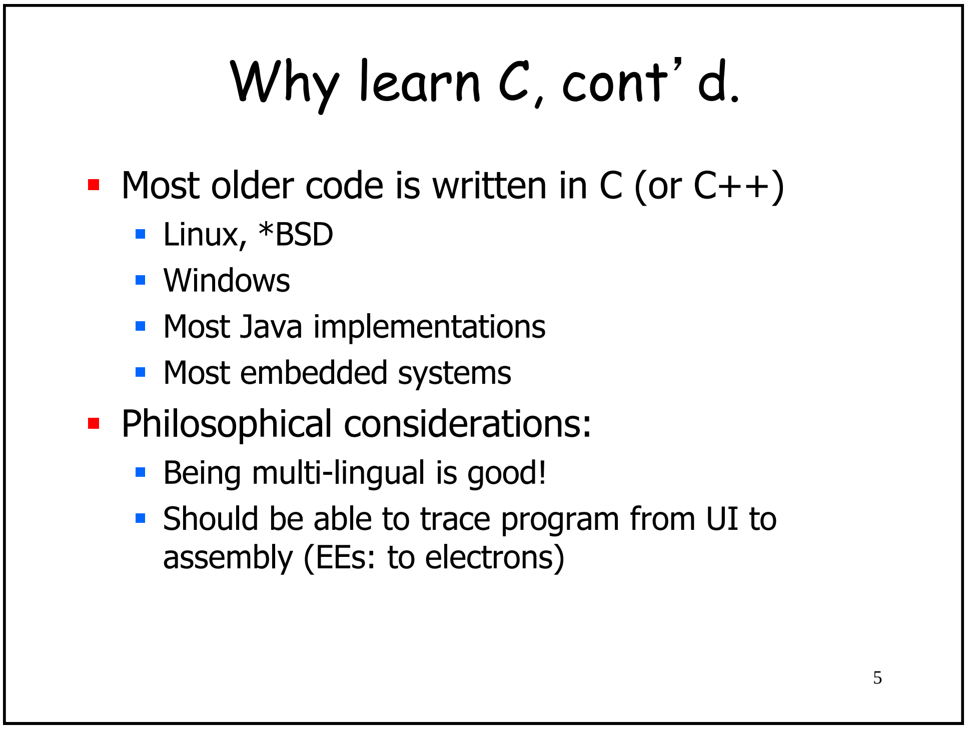
Overview

- Why learn C after Java?
- A brief background on C
- C preprocessor
- Modular C programs

Why learn C (after Java)?

- Both high-level and low-level language
  - OS: user interface to kernel to device driver
- Better control of low-level mechanisms
  - memory allocation, specific memory locations
- Performance *sometimes* better than Java
  - usually more predictable (also: C vs. C++)
- Java hides many details needed for writing OS code

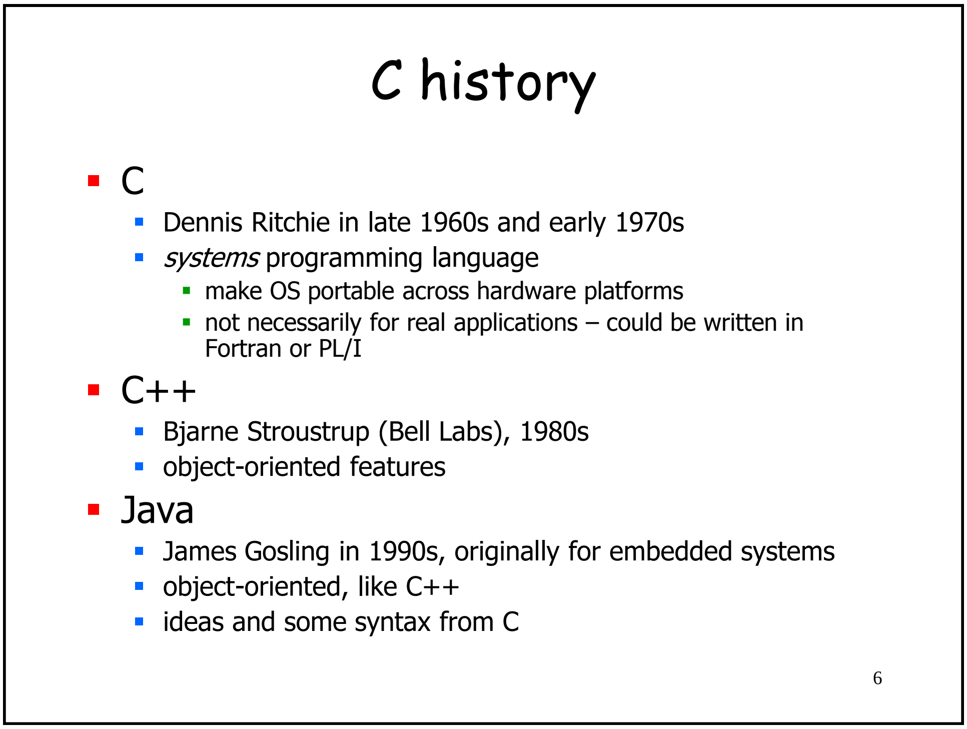
  But C comes with…
  - Memory management responsibility
  - Explicit initialization and error detection
  - generally, more lines for same functionality
  - More room for mistakes



Why learn C, cont'd.

- Most older code is written in C (or C++)
  - Linux, *BSD
  - Windows
  - Most Java implementations
  - Most embedded systems
- Philosophical considerations:
  - Being multi-lingual is good!
  - Should be able to trace program from UI to assembly (EEs: to electrons)

5



C history

- C
  - Dennis Ritchie in late 1960s and early 1970s
  - *systems* programming language
    - make OS portable across hardware platforms
    - not necessarily for real applications – could be written in Fortran or PL/I
- C++
  - Bjarne Stroustrup (Bell Labs), 1980s
  - object-oriented features
- Java
  - James Gosling in 1990s, originally for embedded systems
  - object-oriented, like C++
  - ideas and some syntax from C

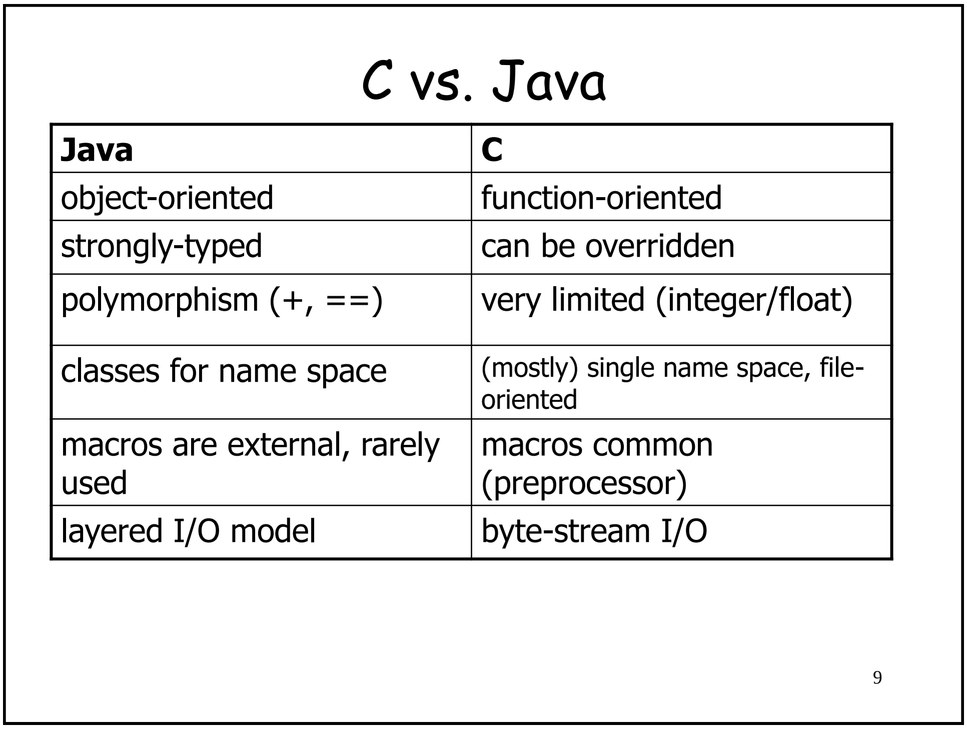
6

C for Java programmers

- Java is mid-90s high-level OO language
- C is early-70s *procedural* language
- C advantages:
  - Direct access to OS primitives (system calls)
  - Fewer library issues – just execute
- (More) C disadvantages:
  - language is portable, APIs are not
  - memory and "handle" leaks
  - preprocessor can lead to obscure errors

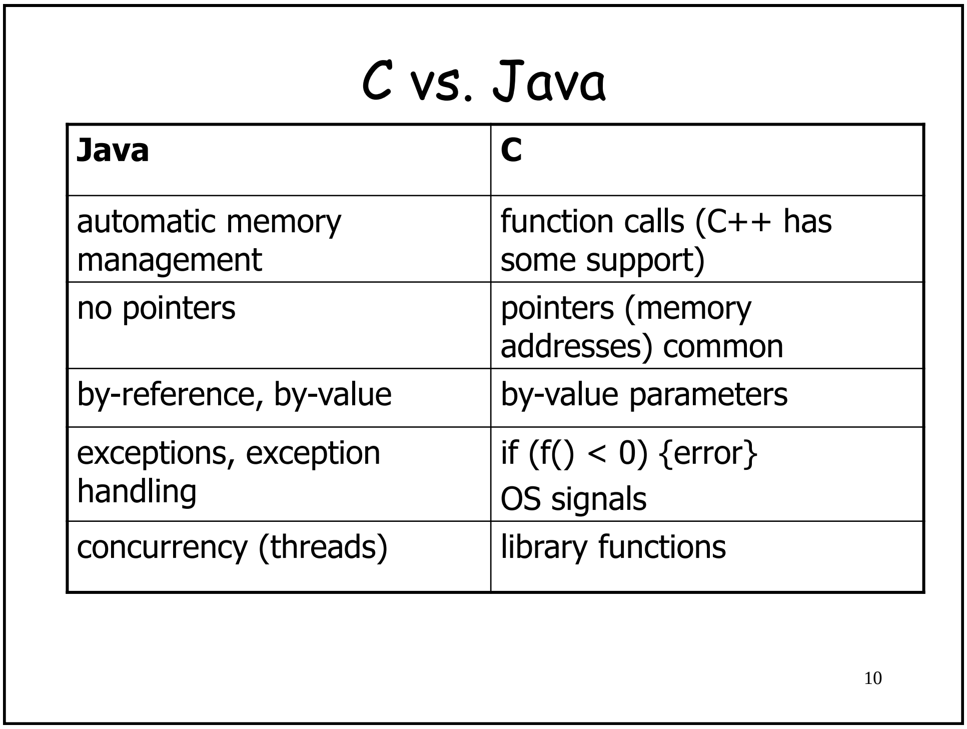
Aside: "generations" and abstraction levels

- Binary, assembly
- Fortran, Cobol
- PL/I, APL, Lisp, …
- C, Pascal, Ada
- C++, Java, Modula3
- Scripting: Perl, Tcl, Python, Ruby, …
- XML-based languages: CPL, VoiceXML



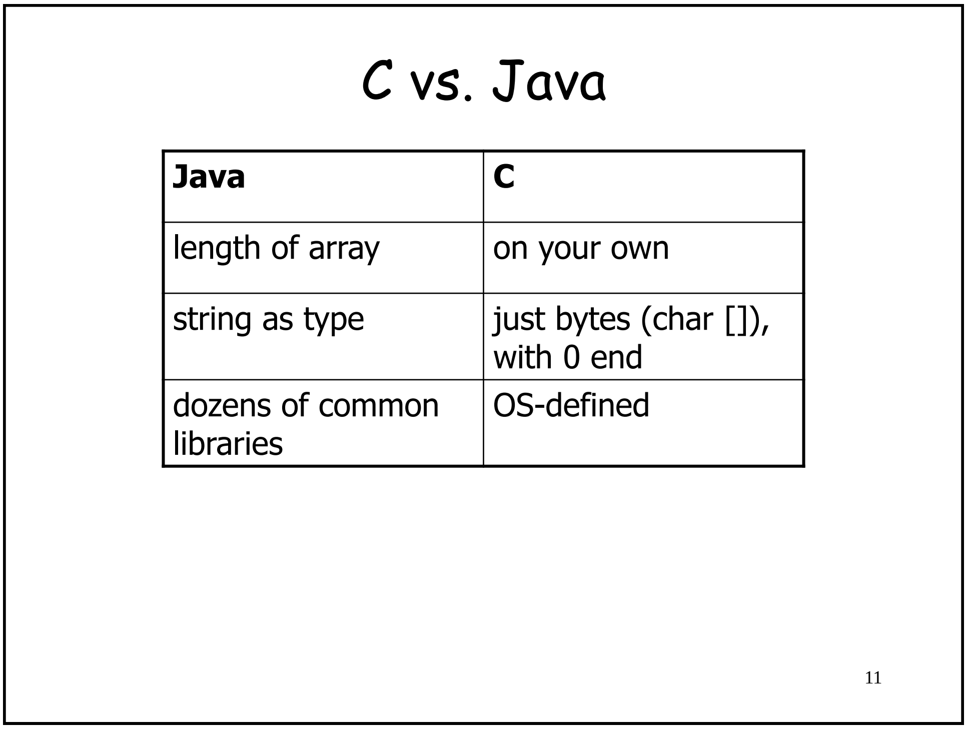
C vs. Java

Java	C
object-oriented	function-oriented
strongly-typed	can be overridden
polymorphism (+, ==)	very limited (integer/float)
classes for name space	(mostly) single name space, file-oriented
macros are external, rarely used	macros common (preprocessor)
layered I/O model	byte-stream I/O



C vs. Java

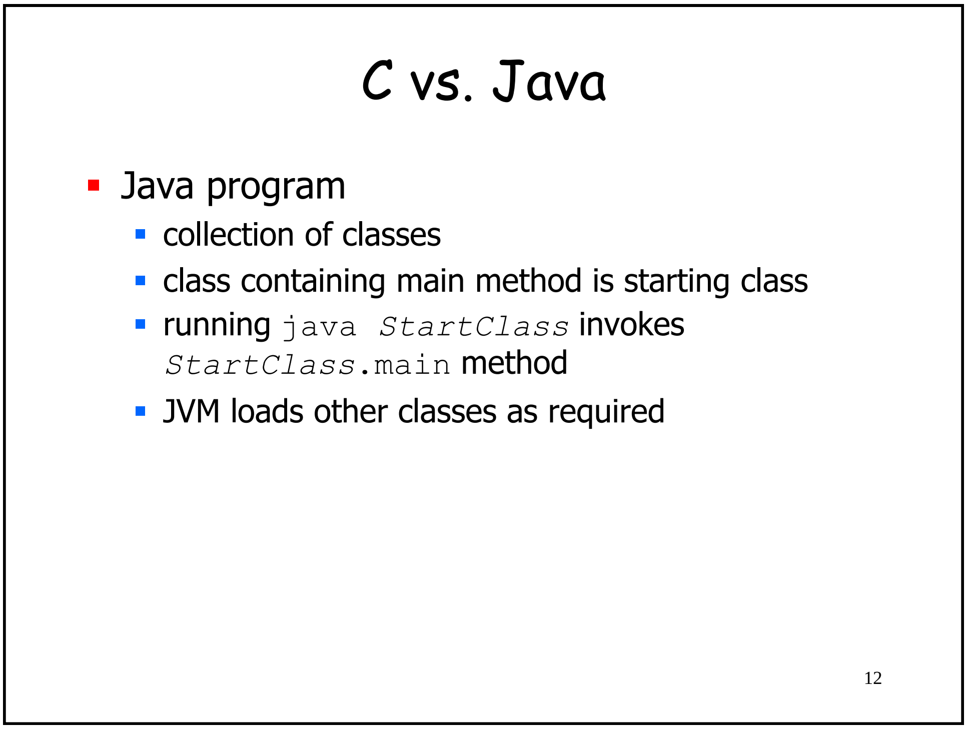
Java	C
automatic memory management	function calls (C++ has some support)
no pointers	pointers (memory addresses) common
by-reference, by-value	by-value parameters
exceptions, exception handling	if (f() < 0) {error} OS signals
concurrency (threads)	library functions



C vs. Java

Java	C
length of array	on your own
string as type	just bytes (char []), with 0 end
dozens of common libraries	OS-defined

11



C vs. Java

- Java program
  - collection of classes
  - class containing main method is starting class
  - running `java` *`StartClass`* invokes *`StartClass`*`.main` method
  - JVM loads other classes as required

12

C program

- collection of functions
- one function – `main()` – is starting function
- running executable (default name a.out) starts main function
- typically, single program with all user code linked in – but can be dynamic libraries (.dll, .so)

C vs. Java

```
public class hello
{
   public static void main
   (String args []) {
       System.out.println
       ("Hello world");
    }
}
```

```
#include <stdio.h>

int main(int argc, char *argv[])
{
  puts("Hello World");
  return 0;
}
```

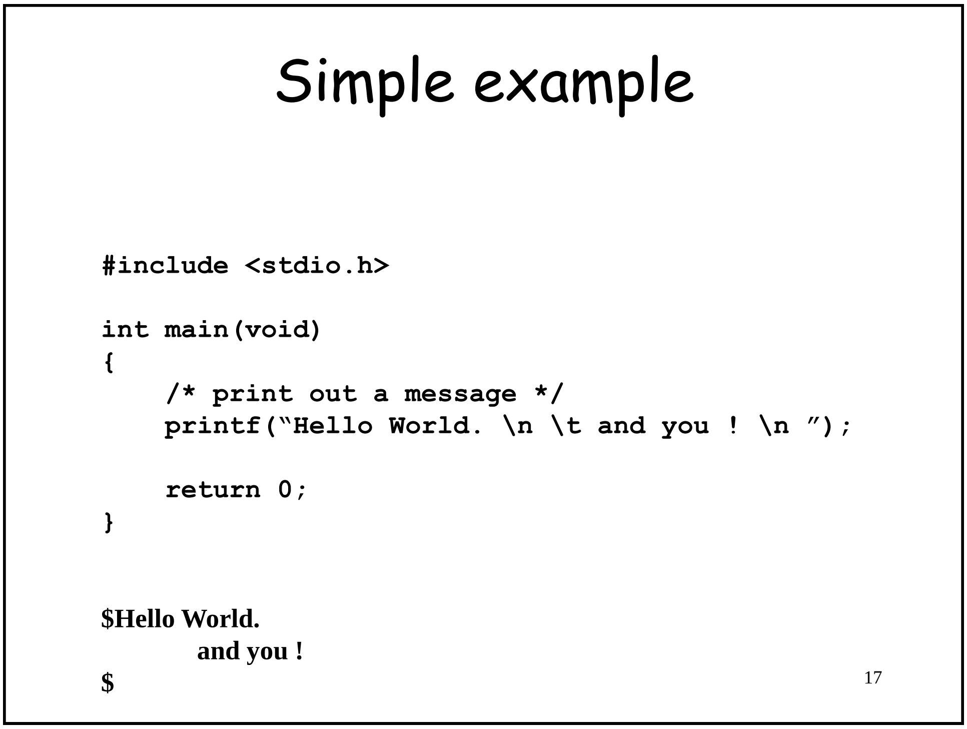
What does this C program do ?

```
#include <stdio.h>

struct list{int data; struct list *next};
struct list *start, *end;

void add(struct list *head, struct list *list, int data};
int delete(struct list *head, struct list *tail);

int main(void)
{
 start=end=NULL;
 add(start, end, 2); add(start, end, 3);
 printf(“First element: %d”, delete(start, end));
 return 0;
}

void add(struct list *head, struct list *tail, int data}
{
 if(tail==NULL){
  head=tail=malloc(sizeof(struct list));
  head->data=data; head->next=NULL;
 }
 else{
  tail->next= malloc(sizeof(struct list));
  tail=tail->next; tail->data=data; tail->next=NULL;
 }
}
```

Terrified ? Come back to this at the end of the slide set and work through it.

What does this C program, do – cont’d?

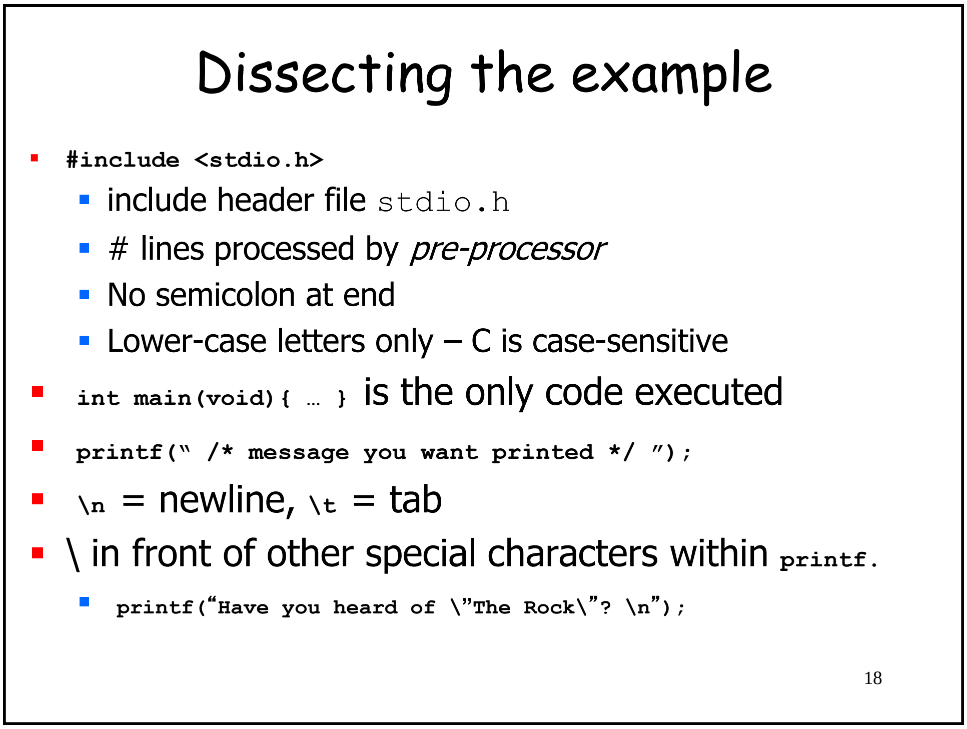
```
void delete (struct list *head, struct list *tail)
{
 struct list *temp;
 if(head==tail){
  free(head); head=tail=NULL;
 }
 else{
  temp=head->next; free(head); head=temp;
 }
}
```



Simple example

```
#include <stdio.h>

int main(void)
{
    /* print out a message */
    printf(“Hello World. \n \t and you ! \n ”);

    return 0;
}
```

```
$Hello World.
        and you !
$
```



Dissecting the example

- `#include <stdio.h>`
  - include header file `stdio.h`
  - # lines processed by *pre-processor*
  - No semicolon at end
  - Lower-case letters only – C is case-sensitive
- `int main(void){ … }` is the only code executed
- `printf(“ /* message you want printed */ ”);`
- `\n` = newline, `\t` = tab
- \ in front of other special characters within `printf`.
  - `printf(“Have you heard of \”The Rock\”? \n”);`

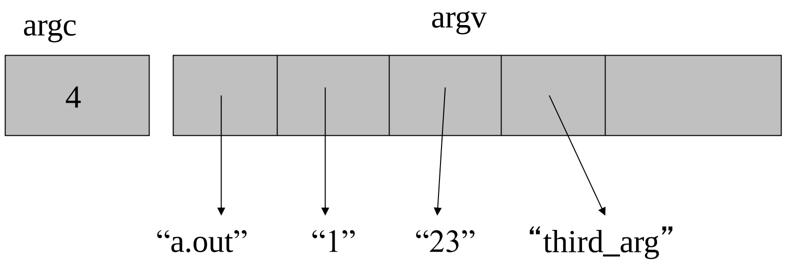
Executing the C program

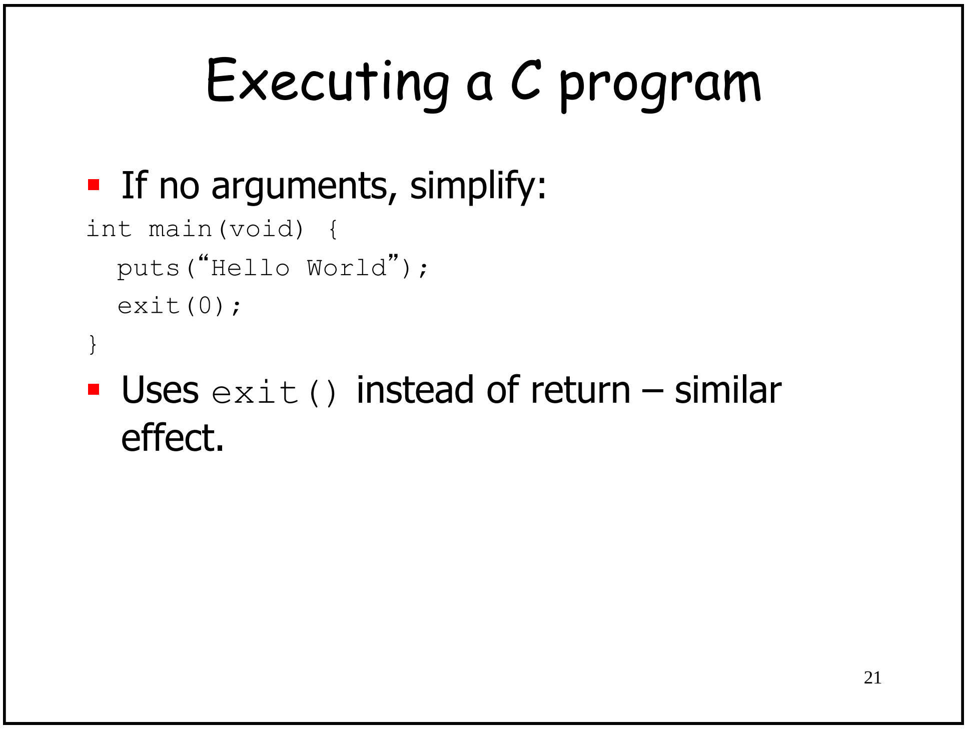
```
int main(int argc, char argv[])
```

- argc is the argument count
- argv is the argument vector
  - array of strings with command-line arguments
- the `int` value is the return value
  - convention: 0 means success, > 0 some error
  - can also declare as void (no return value)

Executing a C program

- Name of executable + space-separated arguments
- $ a.out 1 23 third_arg



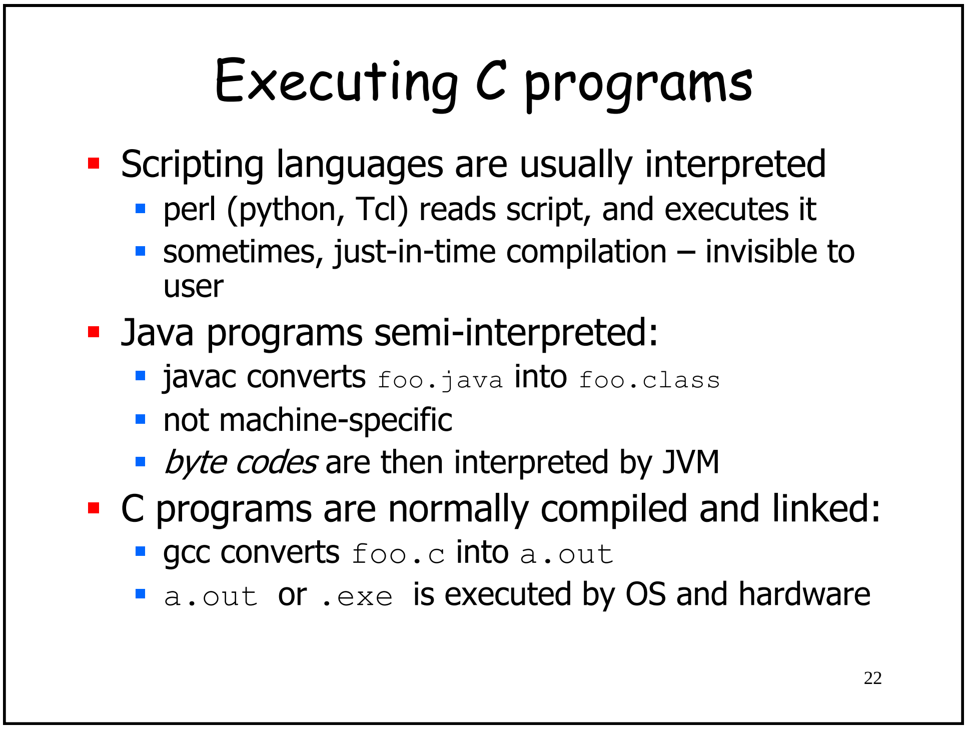


Executing a C program

- If no arguments, simplify:

```
int main(void) {
  puts(“Hello World”);
  exit(0);
}
```

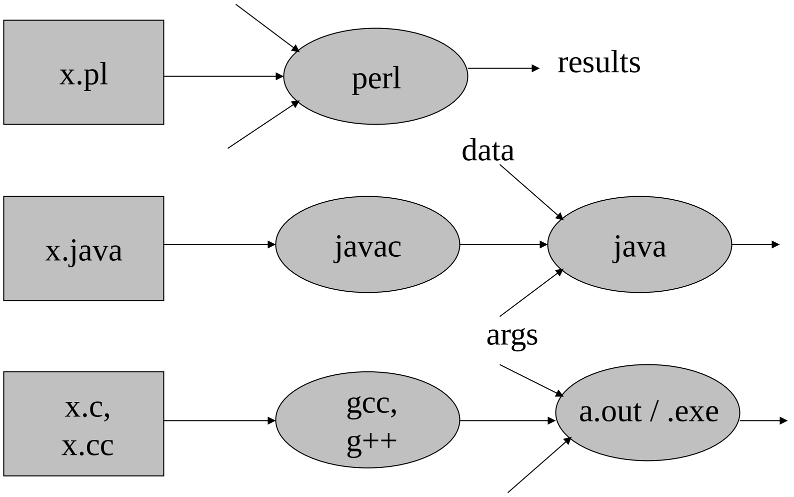
- Uses `exit()` instead of return – similar effect.



Executing C programs

- Scripting languages are usually interpreted
  - perl (python, Tcl) reads script, and executes it
  - sometimes, just-in-time compilation – invisible to user
- Java programs semi-interpreted:
  - javac converts `foo.java` into `foo.class`
  - not machine-specific
  - *byte codes* are then interpreted by JVM
- C programs are normally compiled and linked:
  - gcc converts `foo.c` into `a.out`
  - `a.out` or `.exe` is executed by OS and hardware

Executing C programs



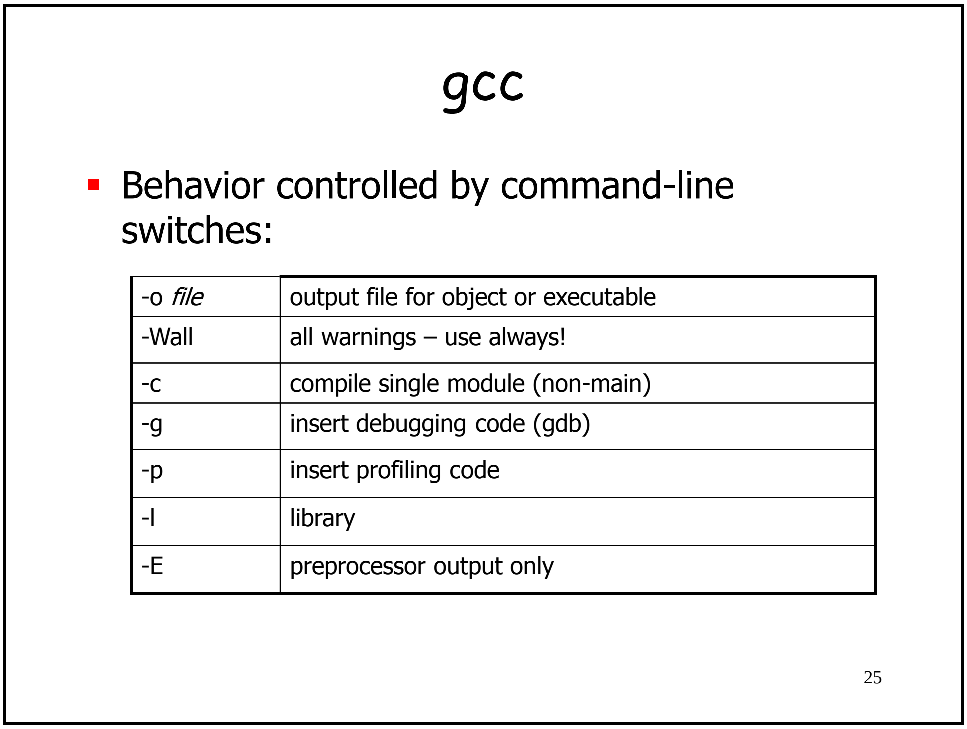
23

The C compiler gcc

- gcc invokes C compiler
- gcc translates C program into executable for some target
- default file name a.out
- also “cross-compilation”

```
$ gcc hello.c
$ a.out
Hello, World!
```

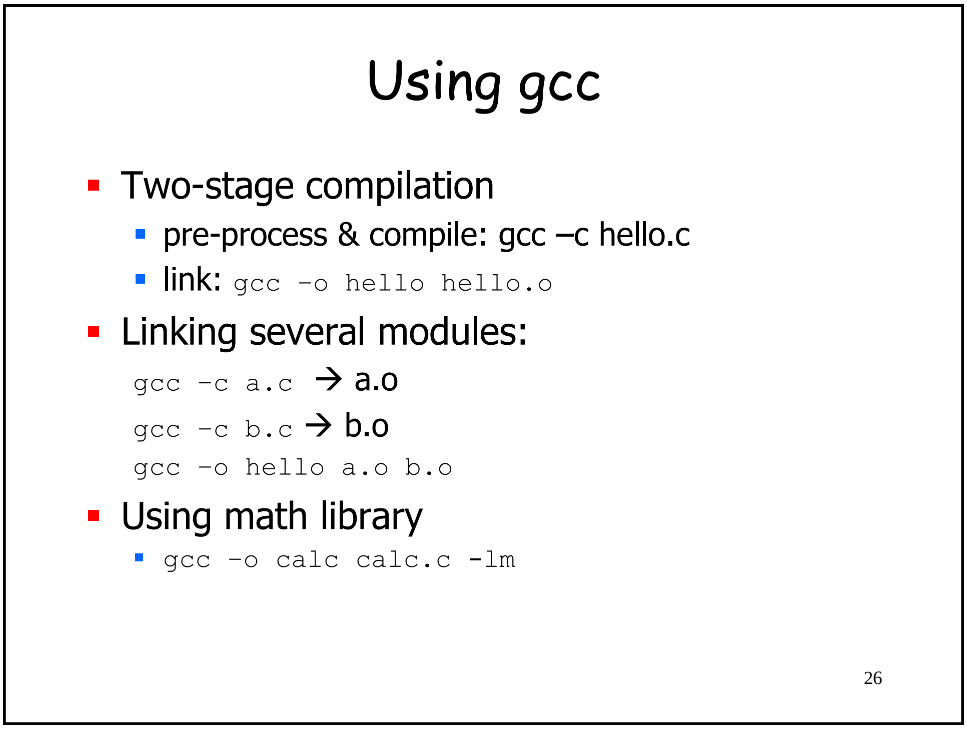
24



gcc

- Behavior controlled by command-line switches:

-o *file*	output file for object or executable
-Wall	all warnings – use always!
-c	compile single module (non-main)
-g	insert debugging code (gdb)
-p	insert profiling code
-l	library
-E	preprocessor output only

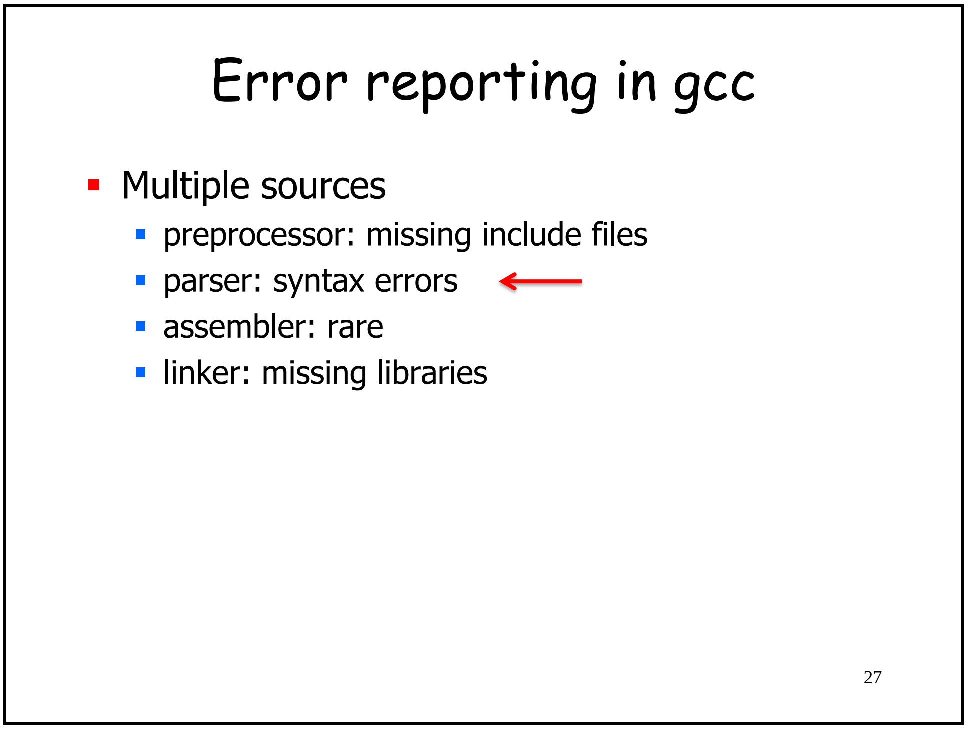


Using gcc

- Two-stage compilation
  - pre-process & compile: gcc –c hello.c
  - link: `gcc –o hello hello.o`
- Linking several modules:

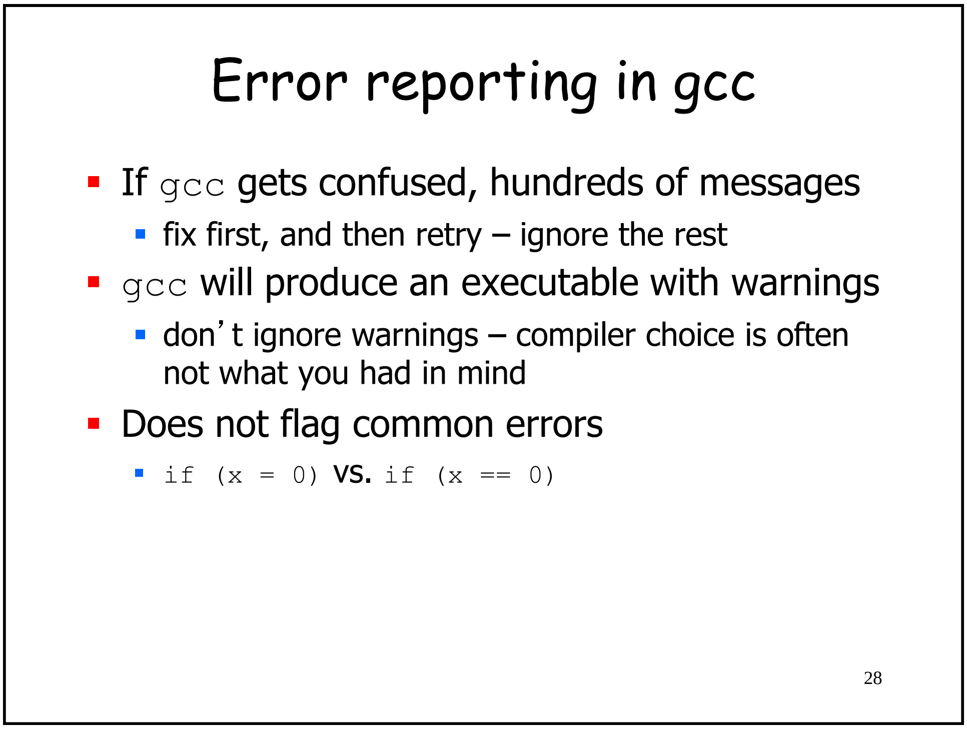
```
gcc –c a.c  → a.o
gcc –c b.c → b.o
gcc –o hello a.o b.o
```

- Using math library
  - `gcc –o calc calc.c -lm`



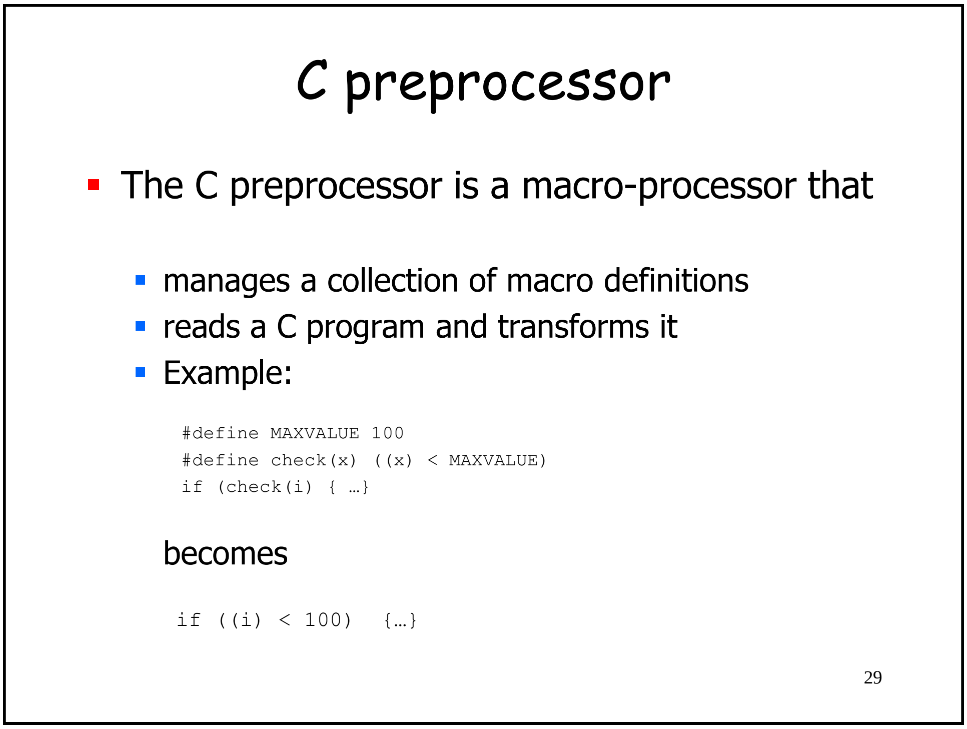
Error reporting in gcc

- Multiple sources
  - preprocessor: missing include files
  - parser: syntax errors ←
  - assembler: rare
  - linker: missing libraries



Error reporting in gcc

- If `gcc` gets confused, hundreds of messages
  - fix first, and then retry – ignore the rest
- `gcc` will produce an executable with warnings
  - don’t ignore warnings – compiler choice is often not what you had in mind
- Does not flag common errors
  - `if (x = 0)` vs. `if (x == 0)`



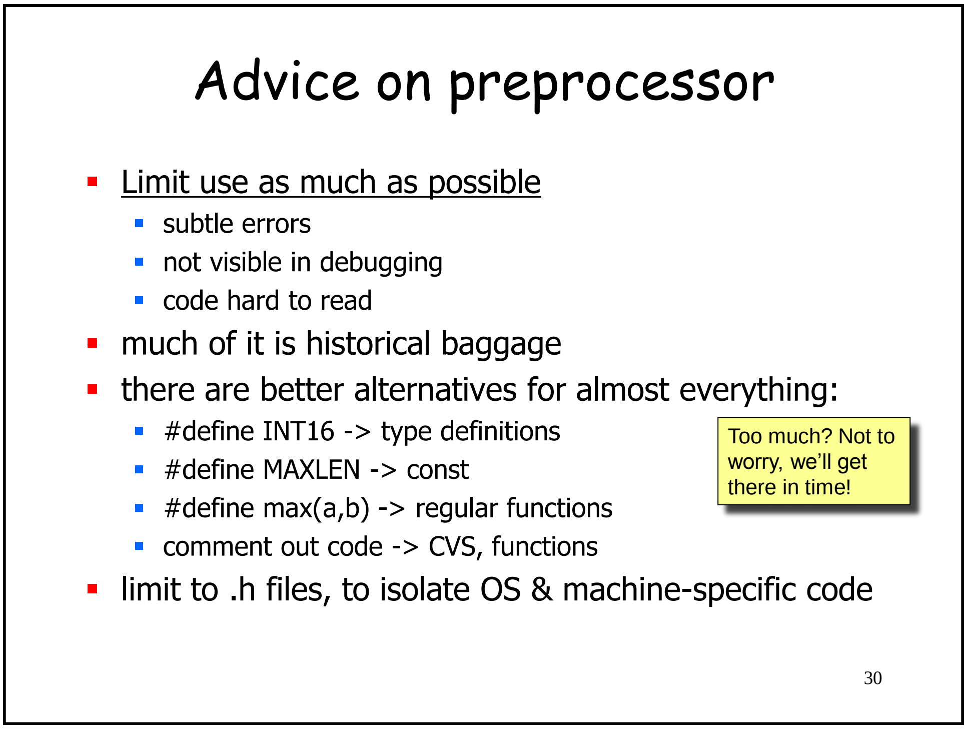
C preprocessor

- The C preprocessor is a macro-processor that
  - manages a collection of macro definitions
  - reads a C program and transforms it
  - Example:

```
#define MAXVALUE 100
#define check(x) ((x) < MAXVALUE)
if (check(i) { …}
```

becomes

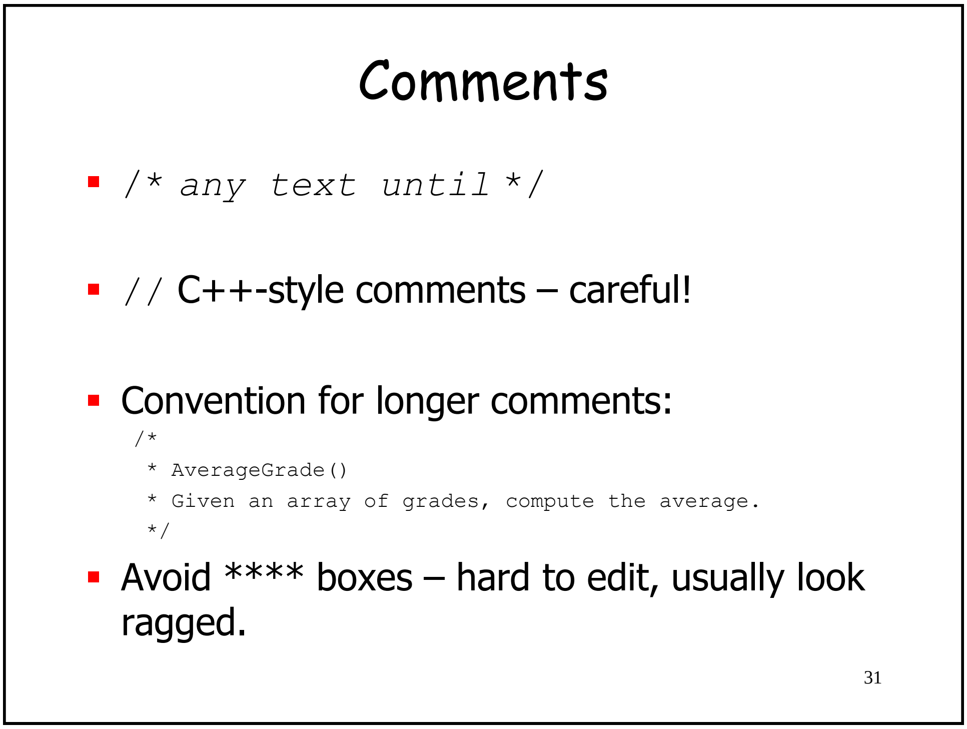
```
if ((i) < 100)  {…}
```



Advice on preprocessor

- Limit use as much as possible
  - subtle errors
  - not visible in debugging
  - code hard to read
- much of it is historical baggage
- there are better alternatives for almost everything:
  - #define INT16 -> type definitions
  - #define MAXLEN -> const
  - #define max(a,b) -> regular functions
  - comment out code -> CVS, functions
- limit to .h files, to isolate OS & machine-specific code

Too much? Not to worry, we'll get there in time!



Comments

- */* any text until */*
- // C++-style comments – careful!
- Convention for longer comments:

```
/*
 * AverageGrade()
 * Given an array of grades, compute the average.
 */
```

- Avoid **** boxes – hard to edit, usually look ragged.