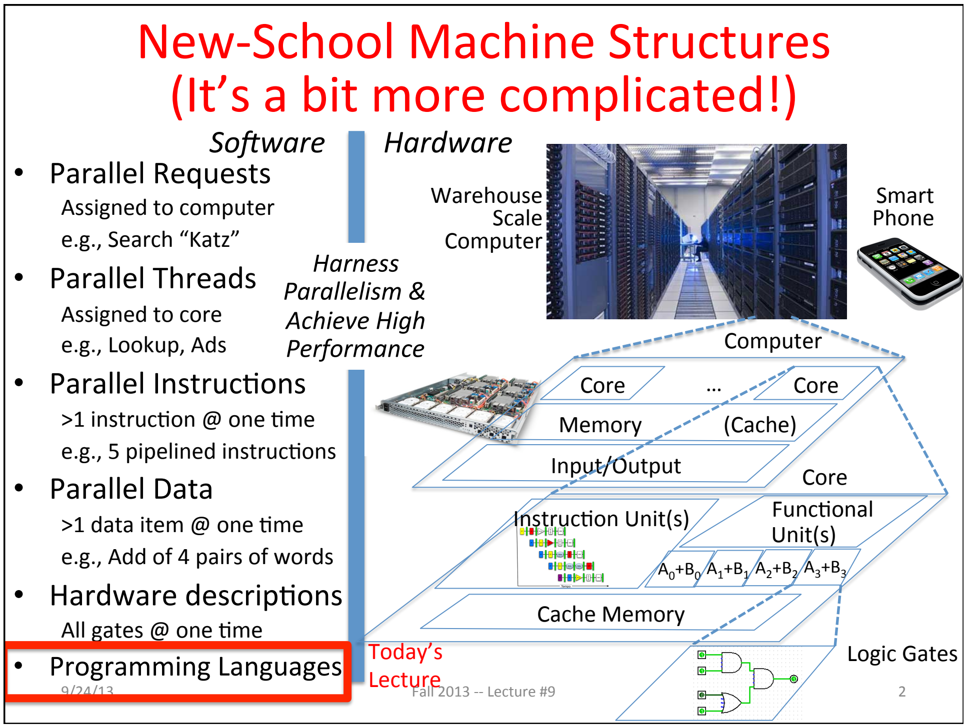
CS 61C:
Great Ideas in Computer Architecture
Assemblers, Linkers, and Compilers

Instructor:
Randy H. Katz
http://inst.eecs.Berkeley.edu/~cs61c/fa13



New-School Machine Structures (It’s a bit more complicated!)

Software | *Hardware*

- Parallel Requests
  Assigned to computer
  e.g., Search “Katz”
- Parallel Threads
  Assigned to core
  e.g., Lookup, Ads
- Parallel Instructions
  >1 instruction @ one time
  e.g., 5 pipelined instructions
- Parallel Data
  >1 data item @ one time
  e.g., Add of 4 pairs of words
- Hardware descriptions
  All gates @ one time
- Programming Languages

Harness Parallelism & Achieve High Performance

Warehouse Scale Computer

Smart Phone

Computer

Core ... Core

Memory (Cache)

Input/Output

Core

Instruction Unit(s)

Functional Unit(s)

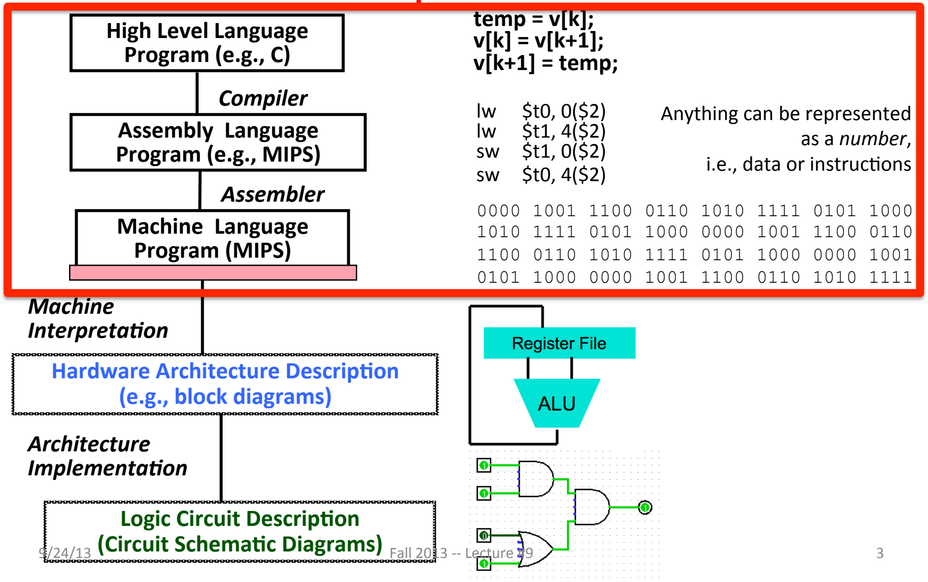
A_0+B_0 A_1+B_1 A_2+B_2 A_3+B_3

Cache Memory

Logic Gates

Today’s Lecture

Big Idea #1: Levels of Representation/ Interpretation



Today's Lecture

High Level Language Program (e.g., C)

Compiler

Assembly Language Program (e.g., MIPS)

Assembler

Machine Language Program (MIPS)

Machine Interpretation

Hardware Architecture Description (e.g., block diagrams)

Architecture Implementation

Logic Circuit Description (Circuit Schematic Diagrams)

```
temp = v[k];
v[k] = v[k+1];
v[k+1] = temp;
```

```
lw   $t0, 0($2)
lw   $t1, 4($2)
sw   $t1, 0($2)
sw   $t0, 4($2)
```

Anything can be represented as a *number*, i.e., data or instructions

```
0000 1001 1100 0110 1010 1111 0101 1000
1010 1111 0101 1000 0000 1001 1100 0110
1100 0110 1010 1111 0101 1000 0000 1001
0101 1000 0000 1001 1100 0110 1010 1111
```

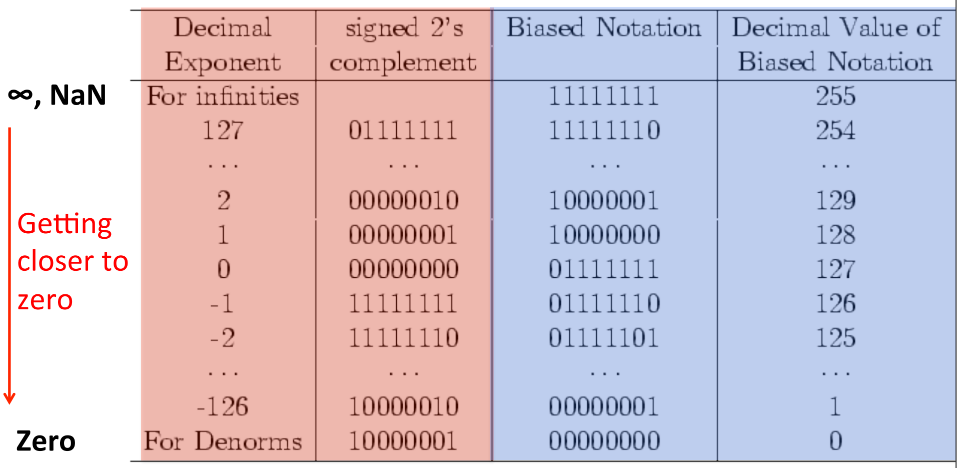
Agenda

- Review: Single + Double Precision FP
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

Agenda

- Review: Single + Double Precision FP
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

Bias Notation (+127)



	How it is interpreted		How it is encoded	
	Decimal Exponent	signed 2's complement	Biased Notation	Decimal Value of Biased Notation
∞, NaN	For infinities		11111111	255
	127	01111111	11111110	254
	...	...	...	...
	2	00000010	10000001	129
	1	00000001	10000000	128
	0	00000000	01111111	127
	-1	11111111	01111110	126
	-2	11111110	01111101	125
	...	...	...	...
	-126	10000010	00000001	1
Zero	For Denorms	10000001	00000000	0

Getting closer to zero

Review: Single Precision Floating Point Numbers

- 32-bit word has 2^{32} patterns, so must be approximation of real numbers ≥ 1.0, < 2
- IEEE 754 Floating Point Standard:
  - 1 bit for *sign (s)* of floating point number
  - 8 bits for *exponent (E)*
  - 23 bits for *fraction (F)* (get 1 extra bit of precision if leading 1 is implicit)

$(-1)^s \times (1 + F) \times 2^E$

- Ranges from 2.0×10^{-38} to 2.0×10^{38}

Review: Double Precision Floating Point Numbers

- What about bigger or smaller numbers?
- IEEE 754 Floating Point Standard: *Double Precision* (64 bits)
  - 1 bit for *sign (s)* of floating point number
  - 11 bits for *exponent (E)*
  - 52 bits for *fraction (F)* (get 1 extra bit of precision if leading 1 is implicit)

$(-1)^s \times (1 + F) \times 2^E$

- Range from 2.0×10^{-308} to 2.0×10^{308}

Floating Point Pitfalls

- A = (1000000.0 + 0.000001) - 1000000.0
- B = (1000000.0 - 1000000.0) + 0.000001
- In single precision floating point arithmetic, A does not equal B

    A = 0.000000, B = 0.000001
- Floating Point Addition is not Associative!
    - Integer addition is associative
- When does this matter?

Agenda

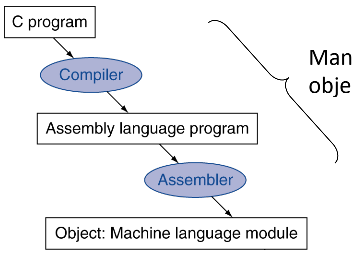
- Review: Single + Double Precision FP
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

Assembler

- Input: Assembly Language Code (e.g., `foo.s` for MIPS)
- Output: Object Code, information tables (e.g., `foo.o` for MIPS)
- Reads and Uses Directives
- Replace Pseudo-instructions
- Produce Machine Language
- Creates Object File

Translation

§2.12 Translating and Starting a Program



Many compilers produce object modules directly

Assembly and Pseudo-instructions

- Turning textual MIPS instructions into machine code called *assembly,* program called *assembler*
  - Calculates addresses, maps register names to numbers, produces binary machine language
  - Textual language called *assembly language*
- Can also accept instructions convenient for programmer but not in hardware
  - *Load immediate* (*li*) allows 32-bit constants, assembler turns into lui + ori (if needed)
  - *Load double* (*ld*) uses two lwc1 instructions to load a pair of 32-bit floating point registers
  - Called *Pseudo-Instructions*

Assembler Directives (P&H Appendix A)

- Give directions to assembler, but do not produce machine instructions

  `.text`: Subsequent items put in user text segment

  `.data`: Subsequent items put in user data segment

  `.globl sym`: declares `sym` global and can be referenced from other files

  `.asciiz str`: Store the string `str` in memory and null-terminate it

  `.word` $w_1 \ldots w_n$: Store the *n* 32-bit quantities in successive memory words

Assembler Pseudo-instructions

- Most assembler instructions represent machine instructions one-to-one
- Pseudo-instructions: figments of the assembler's imagination

```
move $t0, $t1      → add $t0, $zero, $t1
blt $t0, $t1, L    → slt $at, $t0, $t1
                     bne $at, $zero, L
```

  - `$at` (register 1): assembler temporary

More Pseudo-instructions

- Asm. treats convenient variations of machine language instructions as if real instructions

Pseudo: Real:

```
addu $t0,$t6,1        _______________
subu $sp,$sp,32       _______________
sd $a0,32($sp)        _______________
                      _______________
la $a0,str            _______________
```

Agenda

- Review: Single + Double Precision FP
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

Agenda

- Review: Single + Double Precision FP
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

Producing an Object Module

- Assembler (or compiler) translates program into machine instructions
- Provides information for building a complete program from the pieces
  - Header: described contents of object module
  - Text segment: translated instructions
  - Static data segment: data allocated for the life of the program
  - Relocation info: for contents that depend on absolute location of loaded program
  - Symbol table: global definitions and external refs
  - Debug info: for associating with source code

Separate Compilation and Assembly

- No need to compile all code at once
- How to put pieces together?

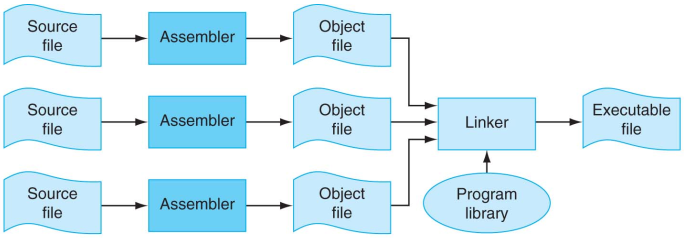
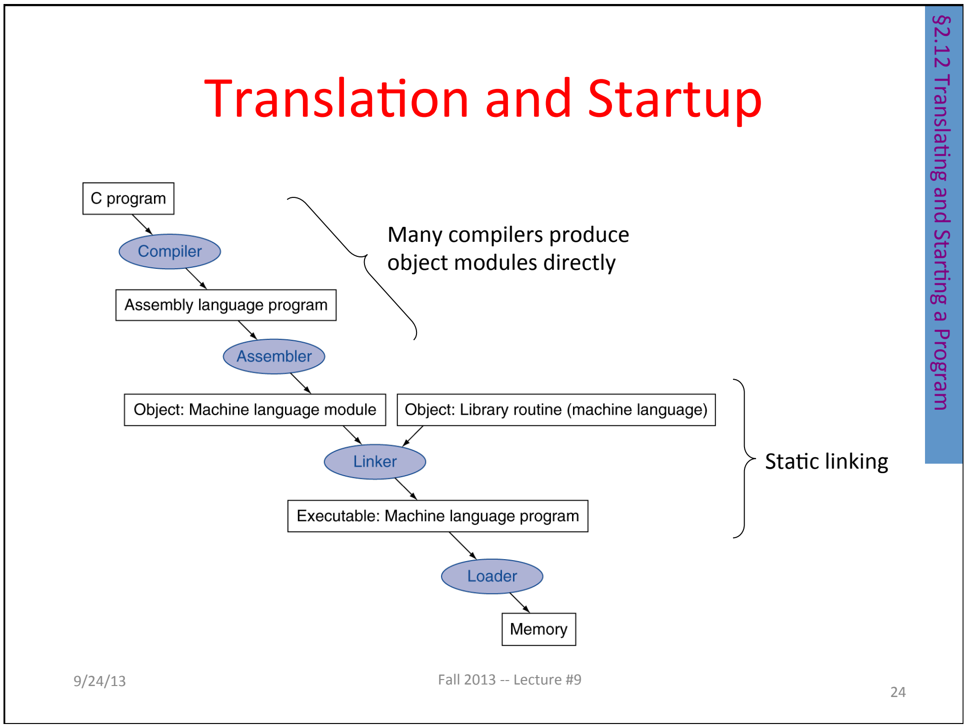


FIGURE B.1.1 The process that produces an executable file. An assembler translates a file of assembly language into an object file, which is linked with other files and libraries into an executable file. Copyright © 2009 Elsevier, Inc. All rights reserved.



Translation and Startup

§2.12 Translating and Starting a Program

C program

Compiler

Assembly language program

Assembler

Object: Machine language module

Object: Library routine (machine language)

Linker

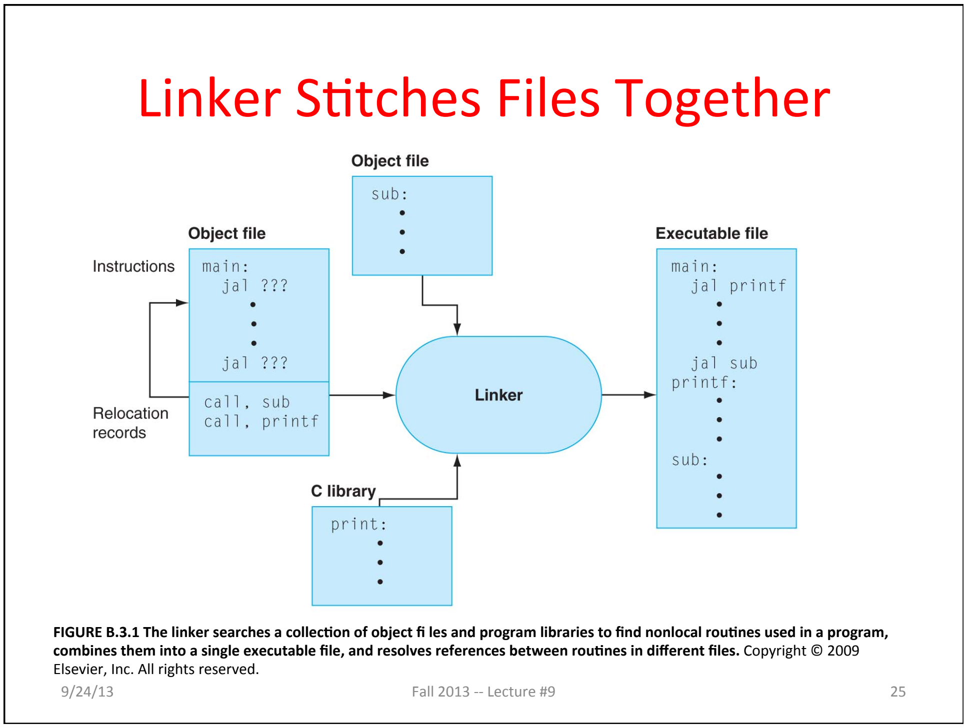
Executable: Machine language program

Loader

Memory

Many compilers produce object modules directly

Static linking



Linker Stitches Files Together

Object file

```
sub:
  •
  •
  •
```

Object file

Instructions

```
main:
  jal ???
  •
  •
  •
  jal ???
```

Relocation records

```
call, sub
call, printf
```

Linker

C library

```
print:
  •
  •
  •
```

Executable file

```
main:
  jal printf
  •
  •
  •
  jal sub
printf:
  •
  •
  •
sub:
  •
  •
  •
```

FIGURE B.3.1 The linker searches a collection of object files and program libraries to find nonlocal routines used in a program, combines them into a single executable file, and resolves references between routines in different files. Copyright © 2009 Elsevier, Inc. All rights reserved.

Linking Object Modules

- Produces an executable image
  1. Merges segments
  2. Resolve labels (determine their addresses)
  3. Patch location-dependent and external refs
- Often slower than compiling
  - All the machine code files must be read into memory and linked together

Loading a Program

- Load from image file on disk into memory
  1. Read header to determine segment sizes
  2. Create virtual address space (covered later in semester)
  3. Copy text and initialized data into memory
  4. Set up arguments on stack
  5. Initialize registers (including `$sp`, `$fp`, `$gp`)
  6. Jump to startup routine
     - Copies arguments to $a0, … and calls main
     - When main returns, do “exit” systems call

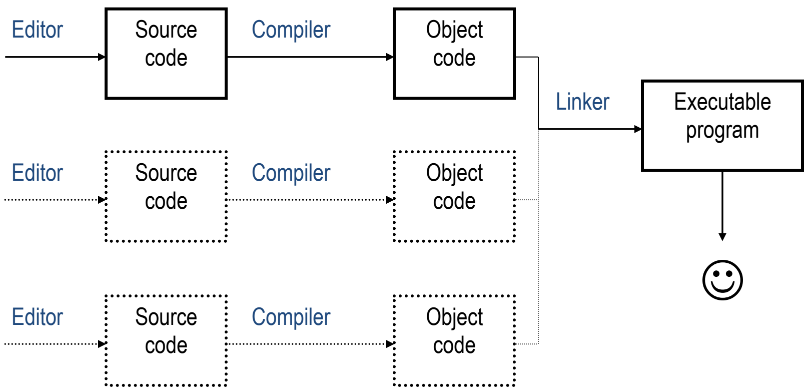
Agenda

- Review: Floating Point Numbers
- Assemblers
- Administrivia
- Linkers
- Compilers vs. Interpreters
- And in Conclusion, …

What’s a Compiler?

- *Compiler*: a program that accepts as input a program text in a certain language and produces as output a program text in another language, *while preserving the meaning of that text*.
- Text must comply with the syntax rules of whichever programming language it is written in.
- Compiler's complexity depends on the syntax of the language and how much abstraction that programming language provides.
  - A C compiler is much simpler than C++ Compiler
- Compiler executes *before* compiled program runs

Compiled Languages: Edit-Compile-Link-Run



Compiler Optimization

- gcc compiler options

-O1: the compiler tries to reduce code size and execution time, without performing any optimizations that take a great deal of compilation time

-O2: Optimize even more. GCC performs nearly all supported optimizations that do not involve a space-speed tradeoff. As compared to -O, this option increases both compilation time and the performance of the generated code

-O3: Optimize yet more. All -O2 optimizations and also turns on the -finline-functions, …

What is Typical Benefit of Compiler Optimization?

- What is a typical program?
- For now, try a toy program: BubbleSort.c

```
#define ARRAY_SIZE 20000
int main() {
  int iarray[ARRAY_SIZE], x, y, holder;
for(x = 0; x < ARRAY_SIZE; x++)
   for(y = 0; y < ARRAY_SIZE-1; y++)
     if(iarray[y] > iarray[y+1]) {
       holder = iarray[y+1];
       iarray[y+1] = iarray[y];
       iarray[y] = holder;
     }
}
```

Unoptimized MIPS Code

```
$L3:
    lw     $2,80016($sp)
    slt    $3,$2,20000
    bne    $3,$0,$L6
    j      $L4
$L6:
    .set   noreorder
    nop
    .set   reorder
    sw     $0,80020($sp)
$L7:
    lw     $2,80020($sp)
    slt    $3,$2,19999
    bne    $3,$0,$L10
    j      $L5
$L10:
    lw     $2,80020($sp)
    move   $3,$2
    sll    $2,$3,2
    addu   $3,$sp,16
```

```
addu   $2,$3,$2
lw     $4,80020($sp)
addu   $3,$4,1
move   $4,$3
sll    $3,$4,2
addu   $4,$sp,16
addu   $3,$4,$3
lw     $2,0($2)
lw     $3,0($3)
slt    $2,$3,$2
beq    $2,$0,$L9
lw     $3,80020($sp)
addu   $2,$3,1
move   $3,$2
sll    $2,$3,2
addu   $3,$sp,16
addu   $2,$3,$2
lw     $3,0($2)
sw     $3,80024($sp
```

```
lw     $3,80020($sp)
addu   $2,$3,1
move   $3,$2
sll    $2,$3,2
addu   $3,$sp,16
addu   $2,$3,$2
lw     $3,80020($sp)
move   $4,$3
sll    $3,$4,2
addu   $4,$sp,16
addu   $3,$4,$3
lw     $4,0($3)
sw     $4,0($2)
lw     $2,80020($sp)
move   $3,$2
sll    $2,$3,2
addu   $3,$sp,16
addu   $2,$3,$2
lw     $3,80024($sp)
sw     $3,0($2)
```

```
$L11:
$L9:
    lw     $2,80020($sp)
    addu   $3,$2,1
    sw
     $3,80020($sp)
    j      $L7
$L8:
$L5:
    lw     $2,80016($sp)
    addu   $3,$2,1
    sw
     $3,80016($sp)
    j      $L3
$L4:
$L2:
    li     $12,65536
    ori
     $12,$12,0x38b0
    addu   $13,$12,$sp
    addu   $sp,$sp,$12
    j      $31
```

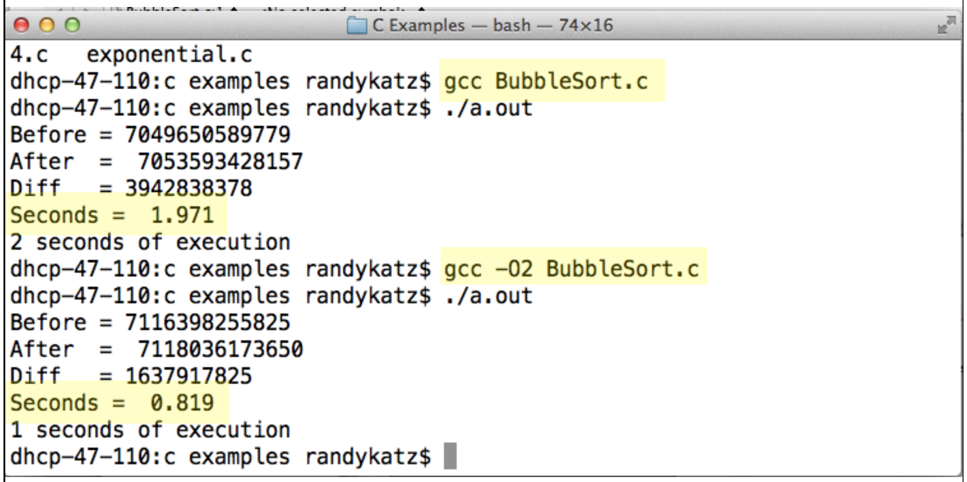
-O2 optimized MIPS Code

```
        li      $13,65536
        ori     $13,$13,0x3890
        addu    $13,$13,$sp
        sw      $28,0($13)
        move    $4,$0
        addu    $8,$sp,16
$L6:
        move    $3,$0
        addu    $9,$4,1
        .p2align 3
$L10:
        sll     $2,$3,2
        addu    $6,$8,$2
        addu    $7,$3,1
        sll     $2,$7,2
        addu    $5,$8,$2
        lw      $3,0($6)
        lw      $4,0($5)
        slt     $2,$4,$3
        beq     $2,$0,$L9
        sw      $3,0($5)
        sw      $4,0($6)
$L9:
        move    $3,$7
        slt     $2,$3,19999
        bne     $2,$0,$L10
        move    $4,$9
        slt     $2,$4,20000
        bne     $2,$0,$L6
        li      $12,65536
        ori     $12,$12,0x38a0
        addu    $13,$12,$sp
        addu    $sp,$sp,$12
        j       $31
        .
```

What's an Interpreter?

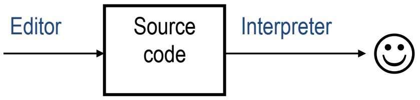
- Reads and executes source statements executed one at a time
  - No linking
  - No machine code generation, so more portable
- Starts executing quicker, but runs much more slowly than compiled code
- Performing the actions straight from the text allows better error checking and reporting to be done
- Interpreter stays around during execution
  - Unlike compiler, some work is done *after* program starts
- Writing an interpreter is much less work than writing a compiler

gcc BubbleSort.c with and without Optimization



```
4.c    exponential.c
dhcp-47-110:c examples randykatz$ gcc BubbleSort.c
dhcp-47-110:c examples randykatz$ ./a.out
Before = 7049650589779
After  =  7053593428157
Diff   = 3942838378
Seconds =  1.971
2 seconds of execution
dhcp-47-110:c examples randykatz$ gcc -O2 BubbleSort.c
dhcp-47-110:c examples randykatz$ ./a.out
Before = 7116398255825
After  =  7118036173650
Diff   = 1637917825
Seconds =  0.819
1 seconds of execution
dhcp-47-110:c examples randykatz$ 
```

Interpreted Languages: Edit-Run


Editor

Source code

Interpreter

☺

Compiler vs. Interpreter Advantages

Compilation:

- Faster Execution
- Single file to execute
- Compiler can do better diagnosis of syntax and semantic errors, since it has more info than an interpreter (Interpreter only sees one line at a time)
- Can find syntax errors *before* run program
- Compiler can optimize code

Interpreter:

- Easier to debug program
- Faster development time

Compiler vs. Interpreter Disadvantages

Compilation:

- Harder to debug program
- Takes longer to change source code, recompile, and relink

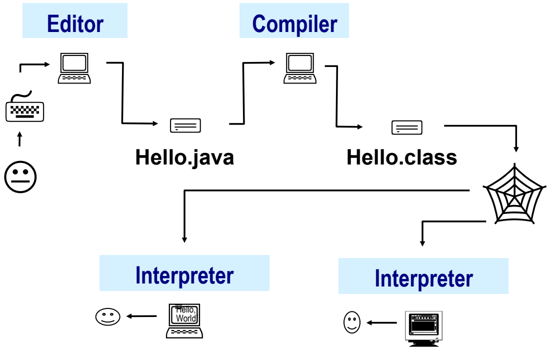
Interpreter:

- Slower execution times
- No optimization
- Need all of source code available
- Source code larger than executable for large systems
- Interpreter must remain installed while the program is interpreted

Java's Hybrid Approach: Compiler + Interpreter

- A Java compiler converts Java source code into instructions for the *Java Virtual Machine (JVM)*
- These instructions, called *bytecodes*, are same for any computer / OS
- A CPU-specific Java interpreter interprets bytecodes on a particular computer

Java's Compiler + Interpreter



Why Bytecodes?

- Platform-independent
- Load from the Internet faster than source code
- Interpreter is faster and smaller than it would be for Java source
- Source code is not revealed to end users
- Interpreter performs additional security checks, screens out malicious code

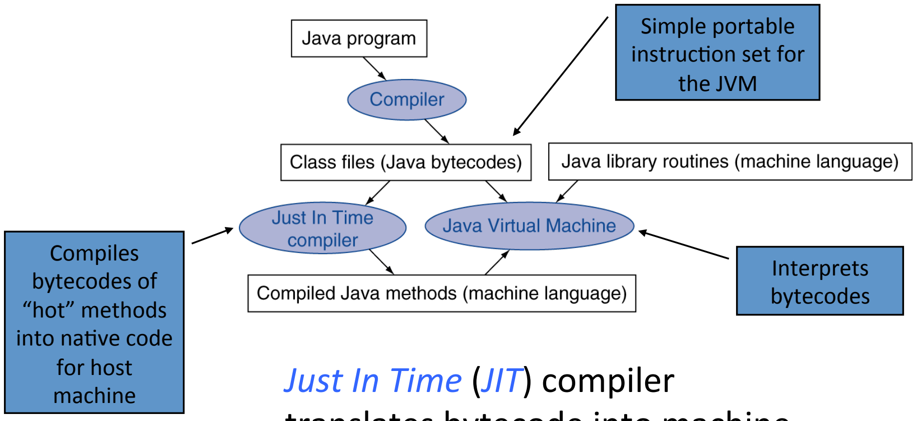
JVM uses Stack vs. Registers

```
a = b + c;
=>
iload b   ;  push b onto Top Of Stack (TOS)
iload c   ;  push c onto Top Of Stack (TOS)
iadd      ;  Next to top Of Stack (NOS) =
          ;  Top Of Stack (TOS) + NOS
istore a ;  store TOS into a and pop stack
```

Java Bytecodes (Stack) vs. MIPS (Reg.)

Category	Operation	Java bytecode	Size (bits)	MIPS Instr.	Meaning
Arithmetic	add	iadd	8	add	NOS=TOS+NOS; pop
	subtract	isub	8	sub	NOS=TOS–NOS; pop
	increment	iinc I8a I8b	8	addi	Frame[I8a]= Frame[I8a] + I8b
Data transfer	load local integer/address	iload I8/aload I8	16	lw	TOS=Frame[I8]
	load local integer/address	iload_/aload_{0,1,2,3}	8	lw	TOS=Frame[{0,1,2,3}]
	store local integer/address	istore I8/astore I8	16	sw	Frame[I8]=TOS; pop
	load integer/address from array	iaload/aaload	8	lw	NOS=*NOS[TOS]; pop
	store integer/address into array	iastore/aastore	8	sw	*NNOS[NOS]=TOS; pop2
	load half from array	saload	8	lh	NOS=*NOS[TOS]; pop
	store half into array	sastore	8	sh	*NNOS[NOS]=TOS; pop2
	load byte from array	baload	8	lb	NOS=*NOS[TOS]; pop
	store byte into array	bastore	8	sb	*NNOS[NOS]=TOS; pop2
	load immediate	bipush I8, sipush I16	16, 24	addi	push; TOS=I8 or I16
	load immediate	iconst_{–1,0,1,2,3,4,5}	8	addi	push; TOS={–1,0,1,2,3,4,5}
Logical	and	iand	8	and	NOS=TOS&NOS; pop
	or	ior	8	or	NOS=TOS|NOS; pop
	shift left	ishl	8	sll	NOS=NOS << TOS; pop
	shift right	iushr	8	srl	NOS=NOS >> TOS; pop
Conditional branch	branch on equal	if_icompeq I16	24	beq	if TOS == NOS, go to I16; pop2
	branch on not equal	if_icompne I16	24	bne	if TOS != NOS, go to I16; pop2
	compare	if_icomp{lt,le,gt,ge} I16	24	slt	if TOS {<,<=,>,>=} NOS, go to I16; pop2
Unconditional jump	jump	goto I16	24	j	go to I16
	return	ret, ireturn	8	jr	
	jump to subroutine	jsr I16	24	jal	go to I16; push; TOS=PC+3

Starting Java Applications



Just In Time (*JIT*) compiler translates bytecode into machine language just before execution

And, in Conclusion, …

- Assemblers can enhance machine instruction set to help assembly-language programmer
- Translate from text that easy for programmers to understand into code that machine executes efficiently: Compilers, Assemblers
- Linkers allow separate translation of modules
- Interpreters for debugging, but slow execution
- Hybrid (Java): Compiler + Interpreter to try to get best of both
- Compiler Optimization to relieve programmer