# Introduction to Computer Architecture 

# Lecture 8: Assembly Programming 

## Pooyan Jamshidi

CSCE 212: Introduction to Computer Architecture I Spring 2024 I https.//pooyanjamshidi. github.io/csce212/

Agenda for Today \& Next Few Lectures

- LC-3 and MIPS Instruction Set Architectures
- LC-3 and MIPS assembly and programming
- Introduction to microarchitecture and single-cycle microarchitecture
- Multi-cycle microarchitecture


## Required Readings

- This week
- Von Neumann Model, LC-3, and MIPS
- P\&P, Chapters 4, 5
- H\&H, Chapter 6
- P\&P, Appendices A and C (ISA and microarchitecture of LC-3)
- H\&H, Appendix B (MIPS instructions)
- Programming
- P\&P, Chapter 6
- Recommended: H\&H Chapter 5, especially 5.1, 5.2, 5.4, 5.5
- Next week
- Introduction to microarchitecture and single-cycle microarchitecture
- H\&H, Chapter 7.1-7.3
- P\&P, Appendices A and C
- Multi-cycle microarchitecture
- H\&H, Chapter 7.4
- P\&P, Appendices A and C


## What Will We Learn Today?

- Assembly Programming
- Programming constructs
- Debugging
- Conditional statements and loops in MIPS assembly
- Arrays in MIPS assembly
- Function calls
- The stack


## Recall: The Von Neumann Model



## Recall: LC-3: A Von Neumann Machine



## Recall: The Instruction Cycle



## Recall: The Instruction Set Architecture

- The ISA is the interface between what the software commands and what the hardware carries out
- The ISA specifies
- The memory organization
- Address space (LC-3: $2^{16}$, MIPS: $2^{32}$ )
- Addressability (LC-3: 16 bits, MIPS: 32 bits)
- Word- or Byte-addressable
- The register set
- R0 to R7 in LC-3
- 32 registers in MIPS

| Problem |
| :--- |
| Algorithm |
| Program |
| ISA |
| Microarchitecture |
| Circuits |
| Electrons |

- The instruction set
- Opcodes
- Data types
- Addressing modes

Our First LC-3 Program: Use of Conditional Branches for Looping

## An Algorithm for Adding Integers

- We want to write a program that adds 12 integers
- They are stored in addresses $0 \times 3100$ to $0 \times 310 \mathrm{~B}$
- Let us take a look at the flowchart of the algorithm


R1: initial address of integers
R3: final result of addition
R2: number of
integers left to be
added
Check if R2 becomes 0
(done with all
integers?)
Load integer in R4
Accumulate integer value in R3
Increment address R1
Decrement R2

## A Program for Adding Integers in LC-3

- We use conditional branch instructions to create a loop



The LC-3 Data Path Revisited

## The LC-3 Data Path

We highlight some
data path
components used in the execution of the instructions in the previous slides (not shown in the simplified data path)


## (Assembly) Programming

## Programming Constructs

- Programming requires dividing a task, i.e., a unit of work into smaller units of work
- The goal is to replace the units of work with programming constructs that represent that part of the task
- There are three basic programming constructs
- Sequential construct
- Conditional construct
a Iterative construct



## Sequential Construct

- The sequential construct is used if the designated task can be broken down into two subtasks, one following the other



## Conditional Construct

- The conditional construct is used if the designated task consists of doing one of two subtasks, but not both

- Either subtask may be "do nothing"
- After the correct subtask is completed, the program moves onward
- E.g., if-else statement, switch-case statement


## Iterative Construct

- The iterative construct is used if the designated task consists of doing a subtask a number of times, but only as long as some condition is true


Iterative


- E.g., for loop, while loop, do-while loop


## Constructs in an Example Program

- Let us see how to use the programming constructs in an example program
- The example program counts the number of occurrences of a character in a text file
- It uses sequential, conditional, and iterative constructs
- We will see how to write conditional and iterative constructs with conditional branches


## Counting Occurrences of a Character

- We want to write a program that counts the occurrences of a character in a file
- Character from the keyboard (TRAP instr.)
- The file finishes with the character EOT (End Of Text)
- That is called a sentinel
- In this example, EOT $=4$
- Result to the monitor (TRAP instr.)

Programming constructs



## TRAP Instruction

- TRAP invokes an OS service call

LC-3 assembly
TRAP 0x23;
Machine Code


- $\mathrm{OP}=1111$
- trapvect8 = service call
- $0 \times 23=$ Input a character from the keyboard
- $0 \times 21=$ Output a character to the monitor
- $0 \times 25=$ Halt the program


## Counting Occurrences of a Char in LC-3

## - We use conditional branch instructions to create a loops and if statements

| Address | $\begin{array}{lllll}15 & 14 & 13 & 12\end{array}$ | $\begin{array}{lllllllllllll}11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0\end{array}$ |  |
| :---: | :---: | :---: | :---: |
|  | AND1 01 | 00 1 0 0 1 0 1 0 0 0 0 0 | R2 $=0$ // initialize counter |
| x3001 |  | 0            <br> 0 1 1 0 0 0 0 1 0 0 0 0 | R3 = M[0x3012] // initial address |
| x3002 | TRAP | 0            <br> 0 0 0 0 0 0 1 0 0 0 1 1 | TRAP 0x23 // input char to R0 |
| x3003 | DDR1 1 | 0 0 1 0 1 1 0 0 0 0 0 0 | R1 = M[R3] // char from file |
| x3004 | ADD0 0 | 1            <br> 1 0 0 0 0 1 1 1 1 1 0 0 | R4 = R1-4 // char - EOT |
| x3005 | BR 0 0 000 | 0            <br> 0 $\frac{2}{2}$ 0 0 0 0 0 0 1 0 0 0 | BRz 0x300E // check if end of file |
| x3006 | NOTO 0 | 0 0 1 0 0 1 1 1 1 1 1 1 | R1 = NOT(R1) // subtract char from |
| x3007 | ADD 0 |  | $\mathrm{R} 1=\mathrm{R} 1+1 \quad$ file from input char |
| x3008 | ADD 0 | 00 0 1 0 0 1 0 0 0 0 0 0 <br> 1 0  0 0 0 0 0 0 0 0  | R1 = R1 + R0 for comparison ? |
| x3009 | BR 0 | ค 100 | BRnp 0x300B ------ |
| x300A | ADD 0 | 0 1 0 0 1 0 1 0 0 0 0 1 | R2 = R2 + 1 // increment the counter |
| x300B | ADD0 0 | 00 1 1 0 1 1 1 0 0 0 0 1 | R3 $=$ R3 + 1 // increment address $\leftarrow^{\prime}{ }^{\prime}$ |
| x300C | ODR1 100 |  | R1 = M[R3] // char from file |
| x300D | BR 0000 | A $\frac{1}{2}$ $p$ 1 1 1 1 1 1 0 1 1 0 <br> 0 0            | BRnzp 0x3004 |
| x 300 E | OD 00 | 0 0 0 0 0 0 0 0 0 1 0 0 | R0 $=$ M [0x3013] $/$ output counter |
| x 300 F | ADD0 0 | 0 0 0 0 0 0 0 0 0 0 1 0 | $\mathrm{RO}=\mathrm{RO}+\mathrm{R} 2$ // output counter |
| x3010 | TRAP | 0            <br> 0 0 0 0 0 0 1 0 0 0 0 1 | TRAP 0x21 to monitor with |
| x3011 | AND1 | 0            <br> 0 0 0 0 0 0 1 0 0 1 0 1 | TRAP 0x25 TRAP |
| x3012 | Starting address of file |  |  |
| x3013 | ASCŶ TEMPDA | (1) 00 |  |

## Programming Constructs in LC-3

- Let us do some reverse engineering to identify conditional constructs and iterative constructs

| Address | 15 $14 \begin{array}{lll}14 & 13 & 12\end{array}$ | $\begin{array}{lllllllllllll}11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0\end{array}$ |  |
| :---: | :---: | :---: | :---: |
| x3000 | AND1 01 | 0 1 0 0 1 0 1 0 0 0 0 0 |  |
| x3001 | OD 0110 | 00 1 1 0 0 0 0 1 0 0 0 0 |  |
| x3002 | TRAP | 0 0 0 0 0 0 1 0 0 0 1 1 |  |
| x3003 | DDR1 100 | 0 0 1 0 1 1 0 0 0 0 0 0 |  |
| x3004 | ADD0 01 | 1            <br> 1 0 0 0 0 1 1 1 1 1 0 0 | R4 = R1-4 // char - EOT |
| x3005 | BR 000 | 01 $z$ 0 0 0 0 0 0 1 0 0 0 | BRz 0x300E // check if end of file |
| x3006 | NOTO 0 | 0 0 1 0 0 1 1 1 1 1 1 1 | R1 = NOT(R1) // subtract char from |
| x3007 | ADD 0 |  | R1 $=$ R1 + 1 file from input char |
| x3008 | ADD 0 | 0            <br> 0 0 1 0 0 1 0 0 0 0 0 0 | R1 $=$ R1 + R0 for comparison |
| x3009 | BR 0000 |  | BRnp 0x300B - - - - |
| x300A | ADD0 0 | 0 1 0 0 1 0 1 0 0 0 0 1 | R2 = R2 + $1 / /$ increment the counter , / I |
| x300B | ADD0 0 | 0 1 1 0 1 1 1 0 0 0 0 1 | 1 |
| x300C | DDR1 100 |  |  |
| x300D | BR 0000 |  | BRnzp 0x3004 |
| x 300 E | OD 0010 | 0 0 0 0 0 0 0 0 0 1 0 0 | 4 |
| x 300 F | ADD0 0 | 0 0 0 0 0 0 0 0 0 0 1 0 |  |
| x3010 | TRAP 1 | 0 0 0 0 0 0 1 0 0 0 0 1 | if (R1 = = R0) \{ |
| x3011 | AND1 | 00 0 0 0 0 0 1 0 0 1 0 1 | // increment the counter |
| x3012 |  | Starting address of file | \} |
| x3013 | $\begin{array}{llll}0 & 0 & 0 & 0\end{array}$ | 0 0 0 0 0 0 1 1 0 0 0 0 |  |

## Debugging

## Debugging

- Debugging is the process of removing errors in programs
- It consists of tracing the program, i.e., keeping track of the sequence of instructions that have been executed and the results produced by each instruction
- A useful technique is to partition the program into parts, often referred to as modules, and examine the results computed in each module
- High-level language (e.g., C programming language) debuggers: dbx, gdb, Visual Studio debugger
- Machine code debugging: Elementary interactive debugging operations


## Interactive Debugging

- When debugging interactively, it is important to be able to
- 1. Deposit values in memory and in registers, in order to test the execution of a part of a program in isolation
- 2. Execute instruction sequences in a program by using
- RUN command: execute until HALT instruction or a breakpoint
- STEP N command: execute a fixed number ( N ) of instructions
- 3. Stop execution when desired
- SET BREAKPOINT command: stop execution at a specific instruction in a program
- 4. Examine what is in memory and registers at any point in the program


## Example: Multiplying in LC-3 (Buggy)

- A program is necessary to multiply, since LC-3 does not have multiply instruction
- The following program multiplies R4 and R5
- Initially, R4 $=10$ and R5 $=3$
- The program produces 40 . What went wrong?
- It is useful to annotate each instruction

| Address | $15 \quad 14$ | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| x3200 | AND 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | R2 = 0 // initialize register |
| x3201 | ADD 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |  |
| x3202 | ADD 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | R5 $=$ R $5-1$ |
| x3203 | BR 0 | 0 | 0 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | BRzp 0x3201 ----'? |
| x3204 | HALT 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | HALT // end program |

## Debugging the Multiply Program



- We examine the contents of all registers after the execution of each instruction

| PC | R2 | R4 | R5 |
| :---: | :---: | :---: | :---: |
| x3201 | 0 | 10 | 3 |
| x3202 | 10 | 10 | 3 |
| x3203 | 10 | 10 | 2 |
| x3201 | 10 | 10 | 2 |
| x3202 | 20 | 10 | 2 |
| x3203 | 20 | 10 | 1 |
| x3201 | 20 | 10 | 1 |
| x3202 | 30 | 10 | 1 |
| x3203 | 30 | 10 | 0 |
| x3201 | 30 | 10 | 0 |
| x3202 | 40 | 10 | 0 |
| x3203 | 40 | 10 | -1 |
| x3204 | 40 | 10 | -1 |
|  | 40 | 10 | -1 |

The branch condition codes were set wrong. The conditional branch should only be taken if R5 is positive

Correct instruction:
BRp \#-3 // BRp 0x3201

## Easier Debugging with Breakpoints

| Address | $15 \quad 14$ | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| x3200 | AND 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | R2 = 0 // initialize register |
| x3201 | ADD 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | R2 = R2 + R4 4 - - - |
| x3202 | ADD 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | R5 $=$ R $5-1$ |
| x3203 | BR 0 | 0 | 0 | 0 | 1 | b | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | BRzp 0x3201 - - - |
| x3204 | HALT 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | HALT // end program |

- We could use a breakpoint to save some work
- Setting a breakpoint in 0x3203 (BR) allows us to examine the results of each iteration of the loop

| PC | R2 | R4 | R5 |
| :---: | :---: | :---: | :---: |
| x3203 | 10 | 10 | 2 |
| x3203 | 20 | 10 | 1 |
| x3203 | 30 | 10 | 0 |
| x3203 | 40 | 10 | -1 |

## One last question:

 Does this program work if the initial value of R5 is 0 ?A good test should also consider the corner cases, i.e., unusual values that the programmer might fail to consider

## Conditional Statements and Loops in MIPS Assembly

## If Statement

- In MIPS, we create conditional constructs with conditional branches (e.g., beq, bne...)

High-level code


Branch not equal
Compares two values ( $\$ s 3=i, \$ s 4=j$ ) and jumps if they are different

## If-Else Statement

- We use the unconditional branch (i.e., j) to skip the "else" subtask if the "if" subtask is the correct one

High-level code

|  | $\begin{aligned} & \# \$ s 0=\mathrm{f}, \quad \$ \mathrm{~s} 1=\mathrm{g} \\ & \# \$ \mathrm{~s} 2=\mathrm{h} \\ & \# \$ \mathrm{~s} 3=\mathrm{i}, \$ \mathrm{~s} 4=\mathrm{j} \end{aligned}$ |
| :---: | :---: |
| $\begin{aligned} & \text { if } \quad(i==j) \\ & f=g+h ; \end{aligned}$ | bne \$s3, \$s4, L1 add \$s0, \$s1, \$s2 |
| else $\mathrm{f}=\mathrm{f}-\mathrm{i} ;$ <br> 1. Compare two values ( $\$ s 3=i, \$ s 4=j$ ) | $\begin{aligned} & \text { L1: } \ddagger \text { done } \\ & \text { done: } \$ s 0, \$ s 0, \$ s 3 \end{aligned}$ |
| and, if they are different, jump to $L 1$, to execute the "else" subtask | 2. Jump to done, after executing the "if" subtask |

## While Loop

- As in LC-3, the conditional branch (i.e., beq) checks the condition and the unconditional branch (i.e., j) jumps to the beginning of the loop

High-level code


## For Loop

- The implementation of the "for" loop is similar to the "while" loop

High-level code


1. Conditional branch to check if the condition still holds

MIPS assembly


## For Loop Using SLT

- We use slt (i.e., set less than) for the "less than" comparison

High-level code


Set less than
$\$ \mathrm{t} 1=\$ \mathrm{~s} 0$

MIPS assembly


Shift left logical

## Arrays in MIPS

- Accessing an array requires loading the base address into a register

- In MIPS, this is something we cannot do with one single immediate operation
- Load upper immediate + OR immediate

```
lui $s0, 0x1234
ori $s0, $s0, 0x8000
```


## Arrays: Code Example

- We first load the base address of the array into a register (e.g., \$s0) using lui and ori

High-level code
int array[5];
array[0] $=\operatorname{array[0]~} \quad \star 2 ;$
$\operatorname{array[1]}=\operatorname{array[1]} * 2 ;$

MIPS assembly

## Function Calls

## Function Calls

- Why functions (i.e., procedures)?
- Frequently accessed code
- Make a program more modular and readable
- Functions have arguments and return value
- Caller: calling function
- main()
- Callee: called function
- sum()

```
void main()
{
    int y;
    y = sum(42, 7);
    ...
}
int sum(int a, int b)
{
    return (a + b);
}
```


## Function Calls: Conventions

## - Conventions

- Caller
- passes arguments
- jumps to callee
- Callee
- performs the procedure
- returns the result to caller
- returns to the point of call
- must not overwrite registers or memory needed by the caller


## Function Calls in MIPS and LC-3

- Conventions in MIPS and LC-3
- Call procedure
- MIPS: Jump and link (jal)
- LC-3: Jump to Subroutine (JSR, JSRR)
- Return from procedure
- MIPS: Jump register (jr)
- LC-3: Return from Subroutine (RET)
- Argument values
- MIPS: \$a0-\$a3
- Return value
- MIPS: \$v0


## Function Calls: Simple Example

High-level code

```
int main() {
    simple();
    a = b + c;
}
void simple() {
    return;
}
```

MIPS assembly

```
0x00400200 main: jal simple
0x00400204 add $s0,$s1,$s2
0x00401020 simple: jr $ra
```

- jal jumps to simple() and saves PC+4 in the return address register (\$ra)
- \$ra = 0x00400204
- In LC-3, JSR(R) put the return address in R7
- jr \$ra jumps to address in \$ra (LC-3 uses RET instruction)


## Function Calls: Code Example

High-level code

```
int main()
```

int main()
{
{
int y;
int y;
// 4 arguments
// 4 arguments
y = diffofsums(2, 3, 4, 5);
y = diffofsums(2, 3, 4, 5);
}
}
int diffofsums(int f, int g,
int diffofsums(int f, int g,
int h, int i)
int h, int i)
{
{
int result;
int result;
result = (f + g) - (h + i);
result = (f + g) - (h + i);
// return value
// return value
return result;
return result;
}

```
}
```


## Function Calls: Need for the Stack

MIPS assembly

```
diffofsums:
    add $t0, $a0, $a1
    add $t1, $a2, $a3
    sub $s0, $t0, $t1
    add $v0, $s0, $0
    jr $ra
```

- What if the main function was using some of those registers?
- \$t0, \$t1, \$s0
- They could be overwritten by the function
- We can use the stack to temporarily store registers


## The Stack

- The stack is a memory area used to save local variables
- It is a Last-In-First-Out (LIFO) queue
- The stack pointer (\$sp) points to the top of the stack - It grows down in MIPS



## The Stack: Code Example

MIPS assembly

```
diffofsums:
    addi $sp, $sp, -12 # allocate space on stack to store 3 registers
    sw $s0, 8($sp) # save $s0 on stack
    sw $t0, 4($sp) # save $t0 on stack
    sw $t1, 0($sp) # save $t1 on stack
    add $t0, $a0, $a1
    add $t1, $a2, $a3
    sub $s0, $t0, $t1
    add $v0, $s0, $0
    lw $t1, 0($sp) # restore $t1 from stack
    lw $t0, 4($sp) # restore $t0 from stack
    lw $s0, 8($sp) # restore $s0 from stack
    addi $sp, $sp, 12 # deallocate stack space
    jr $ra # return to caller
```

- Saving and restoring all registers requires a lot of effort
- In MIPS, there is a convention about temporary registers (i.e., $\$ t 0-\$ t 9)$ : There is no need to save them
- Programmers can use them for temporary/partial results


## MIPS Stack: Register Saving Convention

MIPS assembly

```
diffofsums:
    addi $sp, $sp, -4 # allocate space on stack to store 1 register
    sw $s0, 0($sp) # save $s0 on stack
    add $t0, $a0, $a1
    add $t1, $a2, $a3
    sub $s0, $t0, $t1
    add $v0, $s0, $0
    lw $s0, 0($sp)
    addi $sp,$sp, 4 # deallocate stack space
    jr $ra
# restore $s0 from stack
# return to caller
```

- Temporary registers \$t0-\$t9 are nonpreserved registers. They are not saved, thus, they can be overwritten by the function
- Registers \$s0-\$s7 are preserved (saved; callee-saved) registers


## Lecture Summary

Assembly Programming

- Programming constructs
- Debugging
- Conditional statements and loops in MIPS assembly
- Arrays in MIPS assembly
- Function calls
- The stack

