University of Waterloo

CS241 - Winter 2024 - Course Notes

Author: Brandon Zhou

Course Code: CS241

Course Name: Foundations of Sequential Programs

Instructor: Chengnian Sun

Section: 001

Date Created: January 09, 2024

Final Exam Date: TBA

Disclamer: These course notes are intended to supplement primary instructional materials and facilitate learning. It's worth mentioning that some sections of these notes might have been influenced by ChatGPT, an OpenAI product. Segments sourced or influenced by ChatGPT, where present, will be clearly indicated for reference.

While I have made diligent efforts to ensure the accuracy of the content, there is a potential for errors, outdated information, or inaccuracies, especially in sections sourced from ChatGPT. I make no warranties regarding the completeness, reliability or accuracy of the notes contained in this notebook. It's crucial to view these notes as a supplementary reference and not a primary source.

Should any uncertainties or ambiguities arise from the material, I strongly advise consulting with your course instructors or the relevant course staff for comprehensive understanding. I apologize for any potential discrepancies or oversights.

Any alterations or modifications made to this notebook after its initial creation are neither endorsed nor recognized by me. For any doubts, always cross-reference with trusted academic resources.

Table of Contents

- Lecture 1
- Lecture 2
- Lecture 3
- Lecture 4
- Lecture 5
- Lecture 6
- Lecture 7
- Lecture 8
- Lecture 9
- Lecture 10
- Lecture 11
- Lecture 12
- Lecture 13
- Lecture 14
- Lecture 15
- Lecture 16
- Lecture 17
- Lecture 18
- Lecture 19
- Lecture 20
- Lecture 21
- Lecture 22
- Lecture 23
- Lecture 24
- Tutorial 1

- [Final Review]

Section 1: Lectures

Lecture 1

- Definition: A bit is a binary digit. That is, a 0 or 1 (off or on)
- Definition: A nibble is 4 bits.
  - Example: 1001
- Definition: A byte is 8 bits.
  - Example: 10011101
- in C/C++:
  - Char: 8 bits
  - Unsigned char: 8 bits
  - Short: 2 bytes/16 bits
  - int: 4 bytes
  - longlong: 16 bytes
- Definition: A word is a machine-specific grouping of bytes. For us, a word will be 4 bytes (32-bit architecture) though 8-byte (or 64-bit architectures) words are more common now.
- Definition (Hexadecimal Notation): The base-16 representation system is called the hexadecimal system. It consists of the numbers from 0 to 9 and the letters a, b, c, d, e, f (which convert to the numbers from 10 to 15 in decimal notation)
  - Sometimes we denote the base with a subscript like 10011101_2 and $9d_{16}$.
  - Also, for hexadecimal,you will routinely see the notation `0x9d`. (The `0x` denotes a hexadecimal representation in computer science).
  - Note that each hexadecimal character is a nibble (4 bits).
- Conversion Table
  - Note: upper case letters are also used for hexadecimal notation. Context should make things clear.

Binary	Decimal	Hex
0000	0	0
0001	1	1
0010	2	2
0011	3	3
0100	4	4
0101	5	5
0110	6	6
0111	7	7

Binary	Decimal	Hex
1000	8	8
1001	9	9
1010	10	a
1011	11	b
1100	12	c
1101	13	d
1110	14	e
1111	15	f

- Notation:
  - Binary: 0, 1
  - Decimal(Base 10): 0, ..., 9
  - Hexadecimal: 0, ..., 9, A(10), B(11), C(12), D(13), E(14), F(15)
- Examples:
  - 0000(base 2) -> 0x0(base 16)
  - 1111(base 2) -> 0xf(base 16)
- What do bytes represent?
  - Numbers
  - Characters
  - Garbage in memory
  - Instructions (Words, or 4 bytes, will correspond to a compute instruction in our computer system)
- Bytes as Binary Numbers
  - Unsigned (non-negative integers)
    - $b_7 ... b_0$ (base 2) = $b_7 \times 2^7 + ... + b_0 \times 2^0$ (base 10)
    - Example: $01010101 = 0 \times 2^7 + ... + 1 \times 2^0$
    - Converting to Binary:

- One way: Take the largest power of 2 less than the unsigned integer, subtract and repeat
- Another way is to constantly divide by 2, get the remainer for each division, reading *from the bottom to up* at the end, and that will be the binary representation of this unsigned integer
  - Example: 38

Number	Quotient	Remainder
38	19	0
19	9	1
9	4	1
4	2	0
2	1	0
1	0	1

  - Brief Explanation: Consider

$$N = b_0 + 2b_1 + 2^2 b_2 + \dots$$

The remainder when dividing N by 2 gives the b_0 value. After doing $\frac{N-b_0}{2}$, we end of with

$$\frac{N-b_0}{2} = b_1 + 2b_2 + 2^2 b_3 + \dots$$

and we can repeat the process. (This is why we have to read bottom-up as we get b_0 first, then b_1 ...)

- Signed integers
  - Attempt 1: make the first bit a signed bit. This is called the "sign-magnitude" representation
    - Problems:
      - Two representations of 0(wasteful and awkward)
      - Arithmetic is tricky. Is the sum of a positive number and a negative number positive or negative? It depends!
  - Attempt 2: Two's complement form
    - Similar to "sign-magnitude" representation in spirit, first bit is 0 if non-negative, 1 if negative
    - Negate a value by just subtracting from zero and *letting it overflow*.
    - Decimal to Two's Compliment:
      - A trick to the same thing of negating a value:
        - Take the complement of all bits (flip the 0 bits to 1 and 1 bits to 0)
        - Add 1
      - A slightly faster trick is to locate the rightmost 1 bit and flip all the bits to the left of it
        - Example: 11011010 Negating: 00100110 = 00100101 + 1
        - Note: Flipping the bits and adding 1 is the same as
          - subtracting 1 and flipping the bits for non-zero numbers
          - subtracting from 0
        - Example: compute -38_{10} using this notation in one byte of space:
          - Step 1: $38_{10} = 00100110_2$
          - Step 2: take the complment of all the bits: 11011001_2
          - Step 3: plus 1: 11011010_2
    - Two's Compliment to Decimal
      - Let's compute -38_{10} using one-byte Two's complement. First, write 38 in binary: 38_{10} = 00100110_2. Next, take the complement of all the bits 11011001_2. Finally, add 1: 11011010_2. This last value is -38_{10}.
      - To convert 11011010_2, a number in Two's complement representation, to decimal, one method is to flip the bits and add 1: 00100110_2 = $2^5 + 2^2 + 2^1 = 38$. Thus, the corresponding positive number is 38 and so the original number is -38.

      - Another way to do this computation is to treat the original number 11011010_2 as an unsigned number, convert to decimal and subtract 2^8 from it (since we have 8 bits, and the first bit is a 1 meaning it should be a negative value). This also gives -38

$$\begin{aligned} 11011010_2 &= 2^7 + 2^6 + 2^4 + 2^3 + 2^1 - 2^8 \\ &= 128 + 64 + 16 + 8 + 2 - 256 \\ &= 218 - 256 \\ &= -38 \end{aligned}$$

      - The idea behind [one byte] Two's Complement notation is based on the following observations:
        - The range for unsigned integers is 0 to 255. Recall that 255 is 11111111_2. If we add 1 to 255, then, after discarding overflow bits, we get the number 0.
        - Thus, let's treat 2^8 as 0, i.e., let's work modulo $2^8 = 256$. In this vein, we set up a correspondence between the positive integer k and the unsigned integer $2^8 - k$. Since we are working modulo 2^8, subtracting a positive integer k from 0 is the same as subtracting it from 2^8.
        - In this cse, note that $255 = 2^8 - 1 = 2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0$ and in general

$$2^n - 1 = \sum_{i=0}^{n-1} 2^i$$

        - As an explicit example (which can be generalized naturally) take a number, say $38_{10} = 00100110_2 = 2^5 + 2^2 + 2^1$. What should the corresponding negative number be? Well, note that we've said subtracting a positive integer k from 0 is the same as subtracting it from 2^8:

$$\begin{aligned} 2^8 - 1 &= 2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 \\ 2^8 - 1 &= 38 + 2^7 + 2^6 + 2^4 + 2^3 + 2^0 \\ 2^8 - 38 &= 2^7 + 2^6 + 2^4 + 2^3 + 2^0 + 1 && \text{(flip the bits and add 1)} \end{aligned}$$

        - We mentioned that another method of negating a two's complement number is to flip the bits to the left of the rightmost 1. Justify why this technique works.
          - Every bit to the right of the rightmost 1 is a 0. When we "flip the bits", these become 1s. When we "add 1", the carry propagates up until the position of the "rightmost 1" (the "rightmost 1" is 0 after flip and will stop propagating when the carry reaches this point, and everything on the left of the "rightmost 1" is flipped).

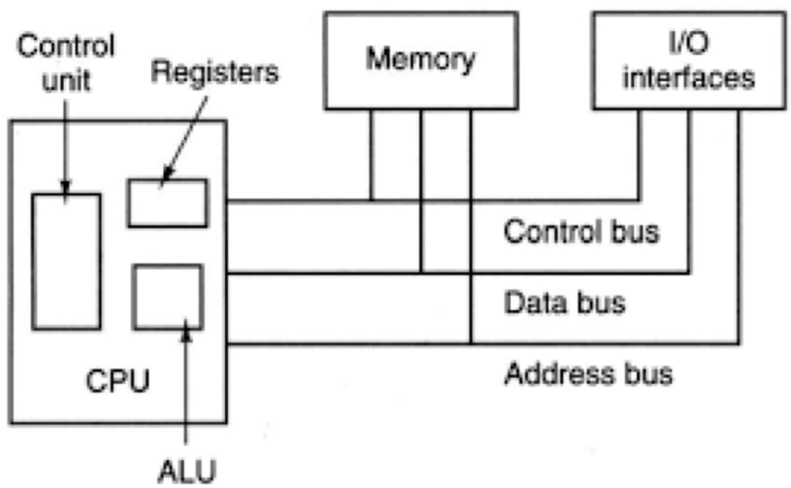
  - The main difference between signed and unsigned binary arithmetic is that we are now working modulo 256 (or, more generally, 2n in the case of n-bit Two's complement numbers)
  - When working in Two's complement, overflow occurs when adding numbers if the original two numbers have the same sign, but the result has a different sign.
  - Arithmetic of Signed Integers
    - All of the arithmetic works by ignoring overflow precisely because arithmetic works in Z_{256}!
  - What is the range of numbers expressible in one-byte Two's Complement notation?
    - $-128 \sim 127$
  - Definitions: The Most Significant Bit (MSB) is the left- most bit (highest value/sign bit); The Least Significant Bit (LSB) is the right-most bit (lowest value).

Lecture 2

- ASCII (American Standard Code for Information Interchange) uses 7 bits to represent characters.

```
char c = '0';
printf("%c", c);
// stdout: 0
printf("%d", (int)c);
// stdout: 48
```

- Bit-Wise Operators
  - Example: suppose we have `unsigned char a=5, b=3;`, which means $a = 00000101, b = 00000011$
    - Bitwise not $\sim a$, for example $c = \sim a$; gives $c = 11111010$
    - Bitwise and $\boxed{\&}$, for example $\boxed{c=a\&b;}$ gives $c = 00000001$
    - Bitwise or $|$, for example $c = a|b$; gives $c = 00000111$
    - Bitwise exclusive^(only true if the bit in a and b is different), for example $c = a\hat{}b$; gives $c = 00000110$
    - Bitwise shift right or left $>>$ and $<<$, for example
      - $c = a >> 2$; gives $c = 00000001$ and
      - $c = a << 3$; gives $c = 00101000$
    - $a << 1$ equivalent to $a * 2$
    - $a >> 1$ equivalent to $a/2$
    - These can even be combined with the assignment operator: `c &= 5;`



- CPU with Memory
- MIPS has 32 registers that are called "general purpose"
  - Some general-purpose registers are special:
    - `$0` is always 0
    - `$31` is for return address
    - `$30` is our stack pointer
    - `$29` is our frame pointer
- Problem: We only know from context what bits have what meaning, and in particular, which are instructions.
  - Solution: Convention is to set memory address 0 in RAM to be an instruction. `0x0: instruction i1`
- Problem: How does MIPS know what to do next?
  - Solution: Have a special register called the Program Counter (or PC for short) to tell us what instruction to do next.
- Problem: How do we put our program into RAM?
  - Solution: A program called a loader puts our program into memory and sets the PC to be the first address.
- Algorithm 1 Fetch-Execute Cycle

```
PC=0
while true do
    IR = MEM[PC]
    PC += 4
    Decode and execute instruction in IR
end while
```

- Write a program in MIPS that adds the values of registers `$8` and `$9` and stores the result in register `$3`.

```
add $d, $s, $t
0000 00ss ssst tttt dddd d000 0010 0000
```

- Why 5 bits for each? Because there are 32 registers `$0 - $31` and $2^5 = 32$
- Adds registers `$s` and `$t` and stores the sum in register `$d`. Important! The order of `$d`, `$s` and `$t` are shifted in the encoding.

Lecture 3

- **Putting values in registers** Load immediate and skip. This places the next value in RAM [an immediate] into $d and increments the program counter by 4 (it skips the next line which is usually not an instruction).

```
lis $d: 0000 0000 0000 0000 dddd d000 0001 0100
# What it really does is
$d = MEM[PC]
PC = PC + 4
```

- How do we get the value we care about into the next location in RAM?

```
.word i: iiii iiii iiii iiii iiii iiii iiii iiii
```

- The above is an assembler directive (not a MIPS instruction). The value i, as a two's complement integer, is placed in the correct memory location in RAM as it occurs in the code.
  - Can also use hexadecimal values: `0xi`
  - Decimal is also allowed.
- Example:

```
lis $1
.word 10
# At this moment, $1 = 10
add $1, $1, $0
```

- Example: Write a MIPS program that adds together 11 and 13 and stores the result in register $3.

```
lis $1
.word 11
list $2
.word 13
add $3, $1, $2
# Solution on the course notes
lis $8          0000 0000 0000 0000 0100 0000 0001 0100
.word 11        0000 0000 0000 0000 0000 0000 0000 1011
lis $9          0000 0000 0000 0000 0100 1000 0001 0100
.word 0xd       0000 0000 0000 0000 0000 0000 0000 1101
add $3,$8,$9    0000 0001 0000 1001 0001 1000 0010 0000
    # The code on the left is what we call Assembly Code.
    # The code on the right is what we call Machine Code.
```

- Jump Register. Sets the pc to be $s.

```
jr $s
0000 00ss sss0 0000 0000 0000 0000 1000
```

- For us, our return address will typically be in $31, so we will typically call the below. This command returns control to the loader.

```
jr $31
0000 0011 1110 0000 0000 0000 0000 1000
```

- So the complete example for the example is (so that the while loop will terminate)

```
lis $8          0000 0000 0000 0000 0100 0000 0001 0100
.word 11        0000 0000 0000 0000 0000 0000 0000 1011
lis $9          0000 0000 0000 0000 0100 1000 0001 0100
.word 0xd       0000 0000 0000 0000 0000 0000 0000 1101
add $3,$8,$9    0000 0001 0000 1001 0001 1000 0010 0000
jr $31          0000 0011 1110 0000 0000 0000 0000 1000
```

- To multiply two words, we need to use the two special registers `hi` and `lo`.
    - `hi` is most significant 4 bytes
    - `lo` is least significant 4 bytes

```
mult $s, $t
0000 00ss ssst tttt 0000 0000 0001 1000
```

- The above performs the multiplication and places the most significant word (largest 4 bytes) in `hi` and the least significant word in `lo`.
- `div $s, $t` performs integer division and places the quotient $s/$t in `lo` [lo quo] and the remainder st in `hi`. Note the sign of the remainder matches the sign of the divisor stored in $s.

```
div $s, $t
0000 00ss ssst tttt 0000 0000 0001 1010
```

- Multiplication and division happen on these special registers hi and lo. How can I access the data?
    - Move from register `hi` into register `$d`.

    ```
    mfhi $d
    0000 0000 0000 0000 dddd d000 0001 0000
    ```

    - Move from register `lo` into register `$d`.

    ```
    mflo $d
    0000 0000 0000 0000 dddd d000 0001 0010
    ```

- RAM
    - Large[r] amount of memory stored off the CPU.
    - RAM access is slower than register access (but is larger, as a tradeoff).
    - Data travels between RAM and the CPU via the bus.
    - Modern day RAM consists of in the neighbourhood of 10^{10} bytes.
    - Instructions occur in RAM starting with address 0 and increase by the word size (in our case 4).
        - But, this simplification will vanish later...
    - Each memory block in RAM has an address; say from 0 to $n-1$
    - Words occur every 4 bytes, starting with byte 0. Indexed by $0, 4, 8, ...n{-}4$.
    - Words are formed from consecutive, aligned (usually) bytes.
    - Cannot directly use the data in the RAM. Must transfer first to registers.
- Load word. Takes a word from RAM and places it into a register. Specifically, load the word in `MEM[$s + i]` and store in `$t`.

```
lw $t, i($s)
1000 11ss ssst tttt iiii iiii iiii iiii
# which is equivalent to
$t = MEM[$s + i]
# Example
lw $1, -4($30)
# which means
$1 <- MEM[$30 - 4]
```

- Store word. Takes a word from a register and stores it into RAM. Specifically, load the word in `$t` and store it in `MEM[$s + i]`.

```
sw $t, i($s)
1010 11ss ssst tttt iiii iiii iiii iiii
```

- Note that `i` must be an immediate, NOT another register! It is a 16-bit Two's complement immediate
- Example: Suppose that `$1` contains the address of an array of words, and `$2` takes the number of elements in this array (assume less than 220). Place the number 7 in the last possible spot in the array.

```
# First element in the array is arr, then the second element is arr + 4 ... the last element in the array is
arr + 4 * (length - 1)
lis $8 ; 7
.word 7
lis $9 ; 4
```

```
.word 4
mult $2, $9 ; length * 4
mflo $3 ; length * 4
add $3, $3, $1 ; arr + length 4
sw $8, -4($3) ; MEM[$3 - 4] = $8
jr $31
```

- Branch on equal.If `$s==$t` then `pc += i*4`. That is, skip ahead i many instructions if `$s` and `$t` are equal.

```
beq $s, $t, i
0001 00ss ssst tttt iiii iiii iiii iiii
# It is like
if ($s == $t) {
    PC += i*4
}
```

- Branch on not equal.If `$s!=$t` then `pc+=i*4`. That is, skip ahead i many instructions if `$s` and `$t` are not equal.

```
bne $s, $t, i
0001 01ss ssst tttt iiii iiii iiii iiii
# It is like
if ($t != $s) {
    PC += i*4
}
```

- Example:

```
beq $0, $0, 0 ; This executes the next instruction as PC has been updated to +4 already
beq $0, $0, 1 ; This executes the second next instruction
```

Lecture 4

Lecture 5

Lecture 6

Lecture 7

Lecture 8

Lecture 9

Lecture 10

Lecture 11

Lecture 12

Lecture 13

Lecture 14

Lecture 15

Lecture 16

Lecture 17

Lecture 18

Lecture 19

Lecture 20

Lecture 21

Lecture 22

Lecture 23

Lecture 24

Section 2: Tutorials

Tutorial 1

- What is Binary?
  - Binary - ways our machines encode info, and $b \in 0, 1$
- Eg. what is 1000 be
  - $2^3 = 8$ unsigned magnitude
  - -8 2's complement
  - $[T, F, F, F]$ array of bools
  - "backspace char" in ASCII
  - **Representation matters**
- Unsigned binary: n-bit binary number is represented as $b_{n-1}, b_{n-2}, \ldots, b_0$, where $b \in 0, 1$
- To convert to decimal: $2^{n-1} \times b_{n-1} + \ldots + 2^0 \times b_0$
- Decimal to binary
  - Idea 1: take the highest powers of 2 from the decimal (inefficient)
  - Idea 2: Repeatedly divide the number by 2, tracking the quotient & remainder
  - Eg. Convert 23 to binary

Number	Quotient	Remainder
23/2	11	1
11/2	5	1
5/2	2	1
2/2	1	0
1/2	0	1

    - and read from bottom to up, that will be $10111_2 = 23_{10}$
- 2's complement
  - Range of values for n-ary
    - Unsigned Binary $0 \sim 2^n - 1$
    - 2's complement $-2^{n-1} \sim 2^{n-1} - 1$
- Convert decimal to 2's complement
  - if number is ≥ 0: use the unsigned representation
  - if number is < 0:

        - Get the binary rep of the positive number
        - flip the bits
        - then add 1
- Convert from 2's complement to decimal
    - Method 1: if $b_{n-1} = 1$, then flip the bits, add 1 and negate the positive decimal
    - Method 2: Treat $b_{n-1} - > -2^{n-1}b_{n-1}$, and add the rest as unsigned representation
- Assembly
    - CS 241: MIPS
        - Runs programs $ stores its data all in MEM (RAM)
        - 32 bits system where instructions are encoded as 4 bytes (1 word)
        - Registers hold 1 word of info
        - Special registers
            - 0 = 0$, immutable
            - 31$, end address in RAM, `jr $31` means return address
            - 3, 29,30$
            - $iii\ldots$ <- 2's complement number
                - `divu, multu, addi ...` treats the register values as unsigned binary
    - **Programs live in the same sapce in MEM (RAM) as as the data they operate on**
        - PC cannot distinguish the two
    - Fetch-Execute Cycle

```
PC = 0x00
while True do:    // until PC = $31
    IR = MEM[PC]
    PC = PC + 4
    ... decode $ execute IR ...
done
```

    - Constant Values
        - Use the Load Immediate Skip command (`lis $s`) followed by an instruction to save into `$s`.
            - `lis $s` -> skip the next instruction -> store that instruction into `$s`
            - Use with `.word i` to store `i` into `$s`
            - Eg. Store 10 into `$5`

```
lis $5
.word 10
; the above two lines of commands skips the .word 10 & sets $5 = 10
```

        - Machine Code (based on the order of machine code horizontally for `lis $5`)
            - 000000 (operating code)
            - 00000 (`$s`)
            - 00000 (`$t`)
            - 00101 (`$d`)
            - 00000 (dead code, always be 0)
            - 010100 (function code)
        - Machine Code (`.word 10`)
            - `00000 ....... 000 001010`

Section 3: Reviews

Final Review