Student Name: Ahmet Uysal

Student ID: 60780

COMP 303 - Computer Architecture: HW1

Due: Oct 13, 2019, 6pm

Instructor: Didem Unat

Corresponding TAs: Aditya Sasongko

Notes: You may discuss the problems with your peers but the submitted work must be your own work. No late assignment will be accepted. Submit a SOFT copy of your assignment to the blackboard *AND* submit a HARD copy of it to the COMP 303 mailboxes. This assignment is worth 4% of your total grade.

Problem 1

(10 pts)

a) In MIPS assembly, write an assembly language version of the following C code segment. At the beginning of the segment, the only values in registers are the base address of arrays A and B in registers $\$a0$ and $\$a1$. To get credit, comment your code.

```
1  int A[40], B[40];
2  for (i=1; i < 40; i++) {
3       B[i] = A[i] + A[i-1];
4       A[i] = 5*B[i] ;
5  }
```

b) Compilers can optimize loops to reduce memory operations (loads and stores) for performance. Implement more optimized version of the above loop in assembly language that uses fewer memory instructions.

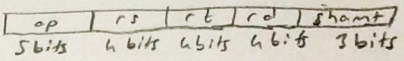
```
//write your code for (a) here
addi $s0, $zero, 1          # i = 0+1
addi $t0, $zero, 40         # temp0 = 0+40

Loop: bge $s0, $t0, Exit    # if i>= temp then jump to Exit
sll $t1, $s0, 2             #shift i left by two bits and save to temp1, temp1 = 4* i
add $t2, $a0, $t1           # temp 2 = address(A) + 4*i = Address of A[i]
lw $t3, 0($t2)              # Load temp3 = A[i]
lw $t4, -4($t2)             # Load temp4 = A[i-1]
add $t5, $t3,$t4            # temp5 = temp3(A[i]) +temp4 (A[i-1])
add $t6, $a1, $t1           # temp6 = address(B) + 4*i = Address of B[i]
sw $t5, 0($t6)              # save temp5 (A[i] + A[i-1]) to B[i]
lw $t7, 0($t6)              # Load temp7 = B[i]
addi $t8, $zero, 5          # temp8 = 0+5, will be used in multiplication

mult $t7, $t8               # temp7 * temp8 => Hi and Lo registers
mflo $t9                    # copy Lo to temp9
sw $t9, 0($t3)              # Save temp9 to A[i]
addi $s0, $s0, 1            # increment i, i++

j Loop                      # jump to Loop (control statement)
Exit: ...


//write your code for (b) here
addi $s0, $zero, 1          #  i = 0+1
addi $t0, $zero, 40         #  temp0 = 0+40
addi $t1, $zero, 5          #  temp1 = 0+5 (will be used for multiplication)

Loop: bge $s0, $t0, Exit    #  if i>= temp then jump to Exit
sll $t2, $s0, 2             #  shift i left by two bits and save to temp2
add $t3, $a0, $t2           #  temp3 = address(A) + 4*i = Address(A[i])
lw $t4, 0($t3)              #  Load temp4 = A[i]
lw $t5, -4($t3)             #  Load temp5 = A[i-1]
add $t6, $t4, $t5           #  temp6 = temp4(A[i]) + temp5(A[i-1])

add $t7, $a1, $t2           #  temp7 = address(B) + 4*i = address(B[i])
sw $t6, 0($t7)              #  save temp6(A[i]+A[i-1]) to B[i]

mult $t1, $t6               #  temp1(5) * temp6(A[i]+A[i-1]) => Hi and Lo registers
mflo $t8                    #  copy Lo register to temp8
sw $t8 0($t3)               #  save temp8 to A[i]

addi $s0, $s0, 1            #  increment i by 1, i++
j Loop                      #  jump to Loop

Exit: ...

.
```

Problem 2

(6 pts) Suppose that you are asked to design a chip for an embedded system that will be used for gym equipment. Because of the cost and space restrictions, you decided to have 20-bit wide instructions and have only 16 general-purpose registers. Similar to MIPS, this chip uses a three-address ISA, which takes two sources, performs an operation on these sources and stores the result back into a destination register.

a) How will your R-type instructions look like? Assume that there is no opcode extension and shift amount is represented in 3 bits.

we have 16 general-purpose registers, so 4 bits are required to encode register numbers rs, rt and rd will be 4 bits long.
shamt is given as 3 bits and there is no funct.
There are $20-(3*4+3)=5$ bits left for op.

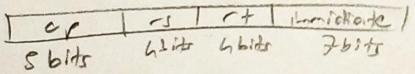


b) How many different instructions you can encode in this new ISA?

5 bits are reserved for op (operation code). So we can encode at most $2^5=32$ different instructions.

c) How many bits should the immediate field of an I-type instruction be to match the length on an R-type instruction?

I-type instruction consists of op code, two registers (rs & rt) and immidiate field. 5 bits are reserved for op code and 4 bits for each register. Immidiate field should be $20-(5+2*4)=7$ bits long to match 20 bit length.



Problem 3

(20 pts) 0-address machine uses a stack, where all operations are done using values stored on the stack. For example,

- PUSH *addr* - pushes the value stored at memory location *addr* onto the stack.
- POP *addr* - pops the stack and stores the value into memory location *addr*
- BOP *addr* - pops two values off the stack, performs the binary operation BOP on the two values, and pushes the result back onto the stack

For example, to compute A + B with 0-address machine, the following sequence of operations are necessary: PUSH A, PUSH B, ADD. After execution of ADD, A and B would no longer be on the stack, but the value A+B would be at the top of the stack.
Make the following assumptions for the next questions and the following code snippet.

- The opcode size is 1 byte (8 bits).
- All register operands are 1 byte (8 bits).
- All memory addresses are 2 bytes (16 bits).
- All data values are 4 bytes (32 bits).
- All instructions are an integral number of bytes in length.

```
1 X = Y + Z;
2 Y = X + Z;
3 T = X - Y;
```

For (a), (b), (c), use the provided table. I-bytes refers to instruction bytes, D-bytes refers to data bytes.

(a) What is the assembly code for the above code for 0-address machine? Assume initially the variables X, Y, Z and T are in memory. Be sure to store the contents of variables back into memory. Do not modify any other values in memory.

(b) What is the assembly code for the above code for MIPS?

(c) Calculate the instruction bytes fetched and the memory-data bytes transferred (read or written) for 0-address machine and for MIPS.

(d) Which ISA is most efficient as measured by code size? Why?

(e) Which ISA is most efficient as measured by total memory traffic (code + data)? Why?

Instruction Set Architecture	Opcode (e.g. Load, Store)	Operands (e.g. X, Y, Z, T)	I-bytes	D-bytes	Total Bytes
0-address Machine	PUSH	Y	3	4	7
	PUSH	Z	3	4	7
	ADD		1	12	13
	POP	X	3	4	7
	PUSH	X	3	4	7
	PUSH	Z	3	4	7
	ADD		1	12	13
	POP	Y	3	4	7
	PUSH	X	3	4	7
	PUSH	Y	3	4	7
	SUB		1	12	13
	POP	T	3	4	7
Total			30	72	102
MIPS	lw	$t0, 0($s1)	4	4	8
	lw	$t2, 0($s2)	4	4	8
	add	$t0, $t1, $t2	4	0	4
	sw	$t0, 0($s0)	4	4	8
	add	$t1, $t0, $t2	4	0	4
	sw	$t1, 0($s1)	4	4	8
	sub	$t3, $t0, $t1	4	0	4
	sw	$t3, 0($s3)	4	4	8
Total			32	20	52

d) 0-address machine ISA is more efficient based on code-size as it uses 30 bytes in total compared to 32 bytes of MIPS.

e) MIPS is almost twice more efficient compared to 0-adress Machine. The main difference is the effective use of registes. Also, I wrote an optimized MIPS implementation for the given code snippet. It doesn't load unused initial values of X & T and does not save Z as its value never changed. BCP instruction is too expensive as it requires two data read and one data write and costs 12 bytes.

Problem 4

(10 pts) MIPS is just one of the Instruction Set Architectures and it is a simple one. Consider a more complex instruction set, called MAPS. One instruction in MAPS can perform the function of many instructions in MIPS. In this question, you are ask to implement the MIPS equivalent for a single instruction in MAPS, namely *short_int_CPYX*, which is defined as follows:

This instruction copies short integers from one address to another. The instruction uses three registers: CNT (count), SRC (source), and DST (destination). The CNT indicates the number of iterations. Each iteration, the instruction moves a short integer (two bytes data) from memory at address SRC to memory at address DST, and then increments both address pointers by two bytes. Thus, the instruction copies in total 2*CNT many bytes from source to destination.

(a) Write the corresponding MIPS code for *short_int_CPYX* that performs the same function. Assume $\$s1 = CNT, \$s2 = SRC, \$s3 = DST$. Try to minimize code size as much as possible.

```
1  . addi  $t1, $zero, 0        # temp1 = 0, used for storing current iteration number
2  .
3  . Loop: beq  $t0, $s1, Exit  # if temp1 (iteration number) == CNT, jump to Exit
4  .
5  . lh  $t2, 0($s2)            # Load half-word (2 bytes) of SRC to temp2
6  . sh  $t2, 0($s3)            # Save half-word temp2 to DST
7  . addi $s2, $s2, 2           # SRC = SRC+2
8  .
9  . addi $s3, $s3, 2           # DST = DST+2
10 . addi $t1, $t1, 1           # temp1 = temp1 + 1, increment current iteration count
11 .
12 . j Loop
13 . Exit: ...
14 .
15 .
16 .
17 .
18 .
```

b) Compare the code sizes of MIPS and MAPS for the same function in bytes. Which one is shorter?

MIPS code contains 8 instructions (excluding Exit label). Each instruction is 4 bytes length in MIPS-32. So, code size is 32 byte for MIPS.

MAPS implements this function in only one instruction. We don't know how many bytes one MAPS instruction takes but it is surely smaller than 32 bytes and MAPS has a smaller code size.

Problem 5

4 pts (This topic will be covered in the class soon)

Instruction Type	Machine M1 Cycles/Instruction	Machine M2 Cycles/Instruction	Instruction Frequency
Loads/Stores	4	4	30%
ALU Operations	1	2	60%
Branches	2	1	10%

Consider two different implementations, Machine M1 and M2, of the same instruction set. There are three types of instructions (loads/stores, ALU operations and branches) in the instruction set. M1 has a clock rate of 800 MHz and M2 has a clock rate of 1.25 GHz. The average number of cycles for each instruction type and their frequencies for an image processing algorithm are provided in the table above.

a) Calculate the average CPI for each machine, M1 and M2.

$CPI_{M1} = CPI_{M1,Load/Store} * f_{Load/Store} + CPI_{M1,ALU} * f_{ALU} + CPI_{M1,Branch} * f_{Branch}$

$= 4 * 0.3 + 1 * 0.6 + 2 * 0.1 = 2$

$CPI_{M2} = CPI_{M2,Load/Store} * f_{Load/Store} + CPI_{M2,ALU} * f_{ALU} + CPI_{M2,Branch} * f_{Branch}$

$4 * 0.3 + 2 * 0.6 + 1 * 0.1 = 2.5$

b) Compare the CPU time of each machine for the image processing algorithm.

$$CPU\ time = \frac{\#\ instructions * CPI}{Clock\ Rate}$$

instruction is same since both executes the same program,

$$CPU\ time_{M1} \propto \frac{2}{800\ MHz} \qquad CPU\ time_{M2} \propto \frac{2.5}{1.25\ GHz}$$

$CPU\ time_{M1} > CPU\ time_{M2}$

$$\frac{CPU\ time_{M1}}{CPU\ time_{M2}} = \frac{5}{4}$$