Model Answers Hw1 - Chapter 2 & 3

2.11. Consider two different machines, with two different instruction sets, both of which have a clock rate of 200 MHz. The following measurements are recorded on the two machines running a given set of benchmark programs:

Instruction Type	Instruction Count (millions)	Cycles per Instruction
Machine A Arithmetic and logic Load and store Branch Others	8 4 2 4	1 3 4 3
Machine A Arithmetic and logic Load and store Branch Others	10 8 2 4	1 2 4 3

a.Determine the effective CPI, MIPS rate, and execution time for each machine.
b. Comment on the results.

$$CPI_A = \frac{\sum CPI_i \times I_i}{I_c} = \frac{(8\times 1 + 4\times 3 + 2\times 4 + 4\times 3)\times 10^6}{(8+4+2+4)\times 10^6} \approx 2.22$$

$$MIPS_A = \frac{f}{CPI_A \times 10^6} = \frac{200\times 10^6}{2.22\times 10^6} = 90$$

$$CPU_A = \frac{I_c \times CPI_A}{f} = \frac{18\times 10^6 \times 2.2}{200\times 10^6} = 0.2 \text{ s}$$

$$CPI_B = \frac{\sum CPI_i \times I_i}{I_c} = \frac{(10\times 1 + 8\times 2 + 2\times 4 + 4\times 3)\times 10^6}{(10+8+2+4)\times 10^6} \approx 1.92$$

$$MIPS_B = \frac{f}{CPI_B \times 10^6} = \frac{200\times 10^6}{1.92\times 10^6} = 104$$

$$CPU_B = \frac{I_c \times CPI_B}{f} = \frac{24\times 10^6 \times 1.92}{200\times 10^6} = 0.23 \text{ s}$$

b. Even though, machine B has a higher MIPS than machine A, it needs a longer CPU time to execute the similar set of benchmark programs (instructions).

2.12. Early examples of CISC and RISC design are the VAX 11/780 and the IBM RS/6000, respectively. Using a typical benchmark program, the following machine characteristics result:

Processor	Clock Frequency	Performance	CPU Time
VAX 11/780	5 MHz	1 MIPS	12 x seconds
IBM RS/6000	25 MHz	18 MIPS	x seconds

The final column shows that the VAX required 12 times longer than the IBM measured in CPU time.
a. What is the relative size of the instruction count of the machine code for this benchmark program running on the two machines?
b. What are the CPI values for the two machines?

Answer:
a. The MIPs rate could be computed as the following:

$[\text{(MIPS rate)} / 10^6] = I_c / T$
Thus that:
$I_c = T \times [\text{(MIPS rate)} / 10^6]$

Now by computing the ratio of the instruction count of the IBM RS/6000 to the VAX 11/780 which is:
$[x \times 18] / [12x \times 1] = 18x / 12x = 1.5$

b. Regarding to the VAX 11/780, the CPI = (5 MHz) / (1 MIPS) = 5
Regarding to the IBM RS/6000, the CPI = (25 MHz) / (18 MIPS) = 1.4

2.13. Four benchmark programs are executed on three computers with the following results:

	Computer A	Computer B	Computer C
Program 1	1	10	20
Program 2	1000	100	20
Program 3	500	1000	50
Program 4	100	800	100

The table shows the execution time in seconds, with 100,000,000 instructions executed in each of the four programs. Calculate the MIPS values for each computer for each program.

MIPS rate given as $MIPS = \frac{I_c}{T \times 10^6}$.

Arithmetic mean is $R_A = \frac{1}{m} \sum_{i=1}^{m} R_i$.

Harmonic mean is $R_H = \frac{m}{\sum_{i=1}^{m} \frac{1}{R_i}}$.

By applying $MIPS = I_c / (T \times 10^6) = 100{,}000{,}000/(T \times 10^6) = 100/T$. Therefore,

the MIPS values are:

	Computer A	Computer B	Computer C
Program 1	100	10	5
Program 2	0.1	1	5
Program 3	0.2	0.1	2
Program 4	2	0.125	1

	Arithmetic mean	Rank
Computer A	25.575	1
Computer B	2.80	3
Computer C	3.25	2

	Harmonic mean	Rank
Computer A	0.25	2
Computer B	0.21	3
Computer C	2.1	1

2.16. Consider the example in Section 2.5 for the calculation of average CPI and MIPS rate, which yielded the result of CPI=2.24 and MIPS rate=178. Now assume that the program can be executed in eight parallel tasks or threads with roughly equal number of instructions executed in each task. Execution is on an 8-core system with each core (processor) having the same performance as the single processor originally used. Coordination and synchronization between the parts adds an extra 25,000 instruction executions to each task. Assume the same instruction mix as in the example for each task, but increase the CPI for memory reference with cache miss to 12 cycles due to contention for memory.

a. Determine the average CPI.

b. Determine the corresponding MIPS rate.

c. Calculate the speedup factor.

d.Compare the actual speedup factor with the theoretical speedup factor determined by Amdhal's law.

Answer:

a. Since we have the same instruction mix, that means the additional instructions for each task could be allocated appropriately between the instruction types. Therefore, the following table be gotten:

Instruction Type	CPI	Instruction Mix
Arithmetic and logic	1	60%
Load/store with cache hit	2	18%
Branch	4	12%
Memory reference with cache miss	12	10%

The average CPI = (1× 0.6) + (2 × 0.18) + (4 × 0.12) + (12 × 0.1) = 2.64. Therefore, the CPI has been increased since the time for memory access is also increased.

b. MIPS = 400/2.64 = 152. There is a corresponding drop in the MIPS rate.

c. The speedup factor equals to the ratio of the execution times. The execution time is calculated as the following: $T = I_c / (MIPS \times 10^6)$.
For the one processor, $T_1 = (2 \times 10^6) / (178 \times 10^6) = 11$ ms.
For the 8 processors, each processor executes 1/8 of the 2 million instructions plus the 25,000

$$T_8 = \frac{\frac{2 \times 10^6}{8} + 0.025 \times 10^6}{152 \times 10^6} = 1.8 \text{ ms}$$

Therefore we have

$$\text{Speedup} = \frac{\text{time to execute program on a single processor}}{\text{time to execute program on } N \text{ parallel processors}} = \frac{11}{1.8} = 6.11$$

d. **In fact, there are two inefficiencies in the parallel system.**
The first one is that there are more additional instructions which is added to coordinate between threads.
The second one is that there is contention for memory access. Thus, none of the code is inherently serial, and all of it is parallelizable but with scheduling overhead. It could be said that the memory access conflict means some extent memory reference instructions are not parallelizable.
By depending on the information given, it is not obvious how to quantify this effect in Amdahl's equation. Therefore, if it is supposed that the fraction of code ,which is parallelizable, is f = 1, then Amdahl's law decreases to Speedup = N = 8. Therefore, the actual speedup is only about 75% of the theoretical speedup.

3.1. The hypothetical machine of Figure 3.4 also has two I/O instructions:

0011 Load AC from I/O

0011 Store AC to I/O

In these cases, the 12-bit address identifies a particular I/O device. Show the program execution (using the format of Figure 3.5) for the following program:

1. Load AC from device 5.

2. Add contents of memory location 940.

3. Store AC to device 6.

Assume that the next value retrieved from device 5 is 3 and that location 940 contains a value of 2.

Memory

300	3 0 0 5
301	5 9 4 0
302	7 0 0 6
.	
.	
940	0 0 0 2
941	

We will assume that the memory (contents in hex) as the previous table:
300: 3005; 301: 5940; 302: 7006
Therefore, the steps will be as the following:

Step 1: 3005 → IR
Step 2: 3 → AC
Step 3: 5940 → IR
Step 4: 3 + 2 = 5 → AC
Step 5: 7006 → IR
Step 6: AC → Device 6

3.2. **Consider a hypothetical 32-bit microprocessor having 32-bit instructions composed of two fields: the first byte contains the opcode and the remainder the immediate operand or an operand address.**

a. What is the maximum directly addressable memory capacity (in bytes)?

b. Discuss the impact on the system speed if the microprocessor bus has

1. a 32-bit local address bus and a 16-bit local data bus, or

2. a 16-bit local address bus and a 16-bit local data bus.

c. How many bits are needed for the program counter and the instruction register?

Answer:

a. 2^(32-8) = 2^24 = 16,777,216 bytes = 16 MB ,(8 bits = 1 byte for he opcode).

b.1. a 32-bit local address bus and a 16-bit local data bus. Instruction and data transfers would take three bus cycles each, one for the address and two for the data. Since If the address bus is 32 bits, the whole address can be transferred to memory at once and decoded there; however, since the data bus is only 16 bits, it will require 2 bus cycles (accesses to memory) to fetch the 32-bit instruction or operand.

b.2. a 16-bit local address bus and a 16-bit local data bus. Instruction and data transfers would take four bus cycles each, two for the address and two for the data. Therefore, that will have the processor perform two transmissions in order to send to memory the whole 32-bit address; this will require more complex memory interface control to latch the two halves of the address before it performs an access to it. In addition to this two-step address issue, since the data bus is also 16 bits, the microprocessor will need 2 bus cycles to fetch the 32-bit instruction or operand.

c. For the PC needs 24 bits (24-bit addresses), and for the IR needs 32 bits (32-bit addresses).

3.3. Consider a hypothetical microprocessor generating a 16-bit address (for example, assume that the program counter and the address registers are 16 bits wide) and having a 16-bit data bus.
a. What is the maximum memory address space that the processor can access directly if it is connected to a "16-bit memory"?
b. What is the maximum memory address space that the processor can access directly if it is connected to an "8-bit memory"?
c. What architectural features will allow this microprocessor to access a separate "I/O space"?
d. If an input and an output instruction can specify an 8-bit I/O port number, how many 8-bit I/O ports can the microprocessor support? How many 16-bit I/O ports? Explain.

Answer:
a. The Maximum memory address space = 2^16 = 64 Kbytes.

b. The Maximum memory address space = 2^16 = 64 Kbytes.
Therefore, in (a) and (b), the microprocessor is to access 64K bytes, but the difference thing between them is that the access of 8-bit memory will transfer a 8 bits and the access of 16-bit memory may transfer 8 bits or 16 bits word.
c. Separate I/O instructions are needed because during its execution will generate separate its own signals I/O signals. That signals will be different from the memory signals which is generated during the execution for memory instructions. Therefore, one more output pin will be needed to carry I/O signals.

d. With an 8-bit I/O port number the microprocessor can support 2^8 = 256 8-bit input ports,
and 2^8 = 256 8-bit output ports.
With an 8-bit I/O port number the microprocessor can support 2^8 = 256 16-bit input ports,
and 2^8 = 256 16-bit output ports.
Thus, the size of the I/O port will not change the number of I/O ports since the number of I/O ports depends on the number of bits which is used to represent the I/O port number (equals to 8 bits in both cases).