9.1 Computer Arithmetic

1. Describe and explain the pros and cons of fixed and floating number system, particularly the representable range and precision. Illustrate using example of 32-bit fixed point and the 32-bit floating point number in single precision IEEE754 format.

2. An array consisting of the length of 256 wires is given by L[0], L[1], … , L[255]. Assume that user are not allowed to test for any overflow during computation, describe a scheme to compute the average length of the 256 wires that will yield a result with the highest precision based on the following specifications:

    - 16-bit registers are used for storing the data and result.
    - Only Single-Precision (16-bit) and Fixed-Point arithmetic is used.
    - Maximum possible length of each wire is 0x3FF and is an integer.

    No coding is required, illustrate your answer in the form of a mathematical expression and justify your answer.

9.2 Pipelines

3. Consider a processor (not ARM processor) with 4 pipeline stages: Fetch Instruction (F), Decode (D), Execute (E) and Store (S). Assume that

    - Branch target address is calculated at the execute stage
    - Instruction length for every instruction is one word long
    - Each pipeline stage take 1 cycle to complete
    - Delay Branching is not enabled.

    How many cycles does the code in Figure 9.2 take? You can ignore all pipeline conflicts you see in the code, just focus on the pipeline behaviour and execution cycles.

```
      MOV   R6, #0x900 ; I1
      MOV   R5, #0     ; I2
      MOV   R4, #0x800 ; I3
      MOV   R3, #0x300 ; I4
Loop  LDR   R0, [R3]   ; I5
      LDR   R1, [R4]   ; I6
      ADD   R2, R1, R0 ; I7
      ADD   R5, R5, #1 ; I8
      ADD   R4, R4, #1 ; I9
      ADD   R3, R3, #1 ; I10
      CMP   R5, #5     ; I11
      BNE   Loop       ; I12
      ADD   R4, R4, R2 ; I13
      STR   R2, [R6]   ; I14
```

Figure 9.2

1. Describe and explain the pros and cons of fixed and floating number system, particularly the representable range and precision. Illustrate using example of 32-bit fixed point and the 32-bit floating point number in single precision IEEE754 format.

Fixed linear ↑ wl —

2. An array consisting of the length of 256 wires is given by L[0], L[1], … , L[255]. Assume that user are not allowed to test for any overflow during computation, describe a scheme to compute the average length of the 256 wires that will yield a result with the highest precision based on the following specifications:

   - 16-bit registers are used for storing the data and result.
   - Only Single-Precision (16-bit) and Fixed-Point arithmetic is used.
   - Maximum possible length of each wire is 0x3FF and is an integer.

   No coding is required, illustrate your answer in the form of a mathematical expression and justify your answer.

3. Consider a processor (not ARM processor) with 4 pipeline stages: Fetch Instruction (F), Decode (D), Execute (E) and Store (S). Assume that

   - Branch target address is calculated at the execute stage
   - Instruction length for every instruction is one word long
   - Each pipeline stage take 1 cycle to complete
   - Delay Branching is not enabled.

How many cycles does the code in Figure 9.2 take? You can ignore all pipeline conflicts you see in the code, just focus on the pipeline behaviour and execution cycles.

```
      MOV   R6, #0x900 ; I1
      MOV   R5, #0     ; I2
      MOV   R4, #0x800 ; I3
      MOV   R3, #0x300 ; I4
Loop  LDR   R0, [R3]   ; I5
      LDR   R1, [R4]   ; I6
      ADD   R2, R1, R0 ; I7
      ADD   R5, R5, #1 ; I8
      ADD   R4, R4, #1 ; I9
      ADD   R3, R3, #1 ; I10
      CMP   R5, #5     ; I11
      BNE   Loop       ; I12
      ADD   R4, R4, R2 ; I13
      STR   R2, [R6]   ; I14
```

Figure 9.2

3 fill pipeline
4 to Loop start } 7

8+2 x 4

8

(d) Consider a processor with 4 pipeline stages: Fetch Instruction (F), Decode (D), Execute (E) and Store (S). Assume that

- Branch target address is calculated at the execute stage
- Instruction length for every instruction is one word long
- Each pipeline stage takes 1 cycle to complete
- No resource conflicts
- Delayed branching is enabled
- This processor is not an ARM processor

How many cycles does the CPU takes to execute the program in Figure Q1b?

(8 marks)

```
      MOV   R0, #0       ;I1
      MOV   R1, #0x400   ;I2
      MOV   R2, #0x200   ;I3
      MOV   R6, #0x300   ;I4
Loop  LDR   R3, [R2]     ;I5
      LDR   R4, [R3]     ;I6
      ADD   R4, R4, R3   ;I7
      ADD   R0, R0, #1   ;I8
      CMP   R0, #10      ;I9
      BNE   Loop         ;I10
      ADD   R6, R6, #1   ;I11
      ADD   R2, R2, #1   ;I12
      STR   R5, [R1]     ;I13
```

Figure Q1b

4. Consider a processor (not ARM processor) with 4 pipeline stages: Fetch Instruction (F), Decode (D), Execute (E) and Store (S). Assume that

   - Branch target address is calculated at the execute stage
   - Instruction length for every instruction is one word long
   - Each pipeline stage takes 1 cycle to complete
   - No Resource Conflicts
   - Delayed Branching is enabled

   Identify and describe ALL pipeline conflicts the code in Figure 9.3 has when executed in the pipeline processor above. Suggest workaround for pipeline conflicts identified.

```
      MOV   R6, #0x900 ; I1
      MOV   R5, #0     ; I2
      MOV   R4, #0x800 ; I3
      MOV   R3, #0x300 ; I4
Loop  LDR   R0, [R3]   ; I5
      LDR   R1, [R4]   ; I6
      ADD   R2, R1, R0 ; I7
      ADD   R5, R5, #1 ; I8
      CMP   R5, #5     ; I9
      BNE   Loop       ; I10
      ADD   R4, R4, #1 ; I11
      ADD   R3, R3, #1 ; I12
      ADD   R4, R4, R2 ; I13
      STR   R2, [R6]   ; I14
```

Figure 9.3

(Not necessary to be covered during tutorial)
[Optional, but students are encouraged to attempt these questions]

9.3 Rounding Error

5. You have been tasked to write a program that calculates the actual time based on a counter that is incremented once every 0.10 seconds. For example, if the counter value is 3,600,000, you would expect the actual time to be 100 hours ((3,600,000 x 0.1) / (60 x 60)).

   Suppose you have decided to use a 24-bit fixed point representation as shown in Figure 8.4 to store the value of 0.10 seconds ($2^{-4} + 2^{-5} + 2^{-8} + 2^{-9} + 2^{-12} + 2^{-13} + \ldots$).

0	▪	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0

Figure 9.3 – Fixed point representation of 0.1010

   a. Approximate the round-off error (in decimal) of 0.10_{10} due to the fixed-point representation.
   b. What is the effect of this round-off error on the time calculated if the counter value is 3,600,000?