

Solutions - Homework 4

(Due date: June 15th)

Presentation and clarity are very important! Show your procedure!

PROBLEM 1 (20 PTS)

- Calculate the result (provide the 32-bit result) of the following operations with 32-bit floating point numbers. Truncate the results when required. When doing fixed-point division, use $x = 4$ fractional bits. Show your procedure.

✓ C3FA8000 - C1E00000	✓ D0D80000 + D0FA0000	✓ 80C00000 × FAD00000	✓ 7B380000 ÷ C8A00000
-----------------------	-----------------------	-----------------------	-----------------------

- ✓ $X = C3FA8000 - C1E00000$:

C3FA8000: 1100 0011 1111 1010 1000 0000 0000 0000

 $e + bias = 10000111 = 135 \rightarrow e = 135 - 127 = 8$

Significand = 1.11110101

C3FA8000 = -1.11110101×2^8

C1E00000: 1100 0001 1110 0000 0000 0000 0000 0000

 $e + bias = 10000011 = 131 \rightarrow e = 131 - 127 = 4$

Significand = 1.11

C1E00000 = -1.11×2^4

$$X = -1.11110101 \times 2^8 + 1.11 \times 2^4 = -1.11110101 \times 2^8 + \frac{1.11}{2^4} \times 2^8$$

$$X = -(1.11110101 - 0.000111) \times 2^8$$

We perform unsigned subtraction: $X = -1.11011001 \times 2^8$

$$X = -1.11011001 \times 2^8, e + bias = 8 + 127 = 135 = 10000111$$

$$X = 1100 0011 1110 1100 1000 0000 0000 0000 = C3EC8000$$

$$\begin{array}{r} b_9=0 \quad b_8=0 \quad b_7=0 \quad b_6=0 \quad b_5=1 \quad b_4=1 \quad b_3=0 \quad b_2=0 \quad b_1=0 \quad b_0=0 \\ 1.1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 1 \ - \\ 0.0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \\ \hline 1.1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \end{array}$$

- ✓ $X = D0D80000 + D0FA0000$:

D0D80000: 1101 0000 1101 1000 0000 0000 0000 0000

 $e + bias = 10100001 = 161 \rightarrow e = 161 - 127 = 34$

Significand = 1.1011

D0D80000 = -1.1011×2^{34}

D0FA0000: 1101 0000 1111 1010 0000 0000 0000 0000

 $e + bias = 10100001 = 161 \rightarrow e = 161 - 127 = 34$

Significand = 1.1111010

D0FA0000 = -1.1111010×2^{34}

$$X = -1.1011 \times 2^{34} - 1.1111010 \times 2^{34} \text{ (unsigned addition)}$$

$$X = -11.101001 \times 2^{34} = -1.1101001 \times 2^{35}$$

 $e + bias = 35 + 127 = 162 = 10100010$

$$X = 1101 0001 0110 1001 0000 0000 0000 0000 = D1690000$$

$$\begin{array}{r} b_{31}=1 \quad b_{30}=1 \quad b_{29}=1 \quad b_{28}=1 \quad b_{27}=1 \quad b_{26}=0 \quad b_{25}=0 \quad b_{24}=0 \\ 1.1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ + \\ 1.1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \\ \hline 1 \ 1.1 \ 0 \ 1 \ 0 \ 0 \ 1 \end{array}$$

- ✓ $X = 80C00000 \times FAD00000$:

80C00000: 1000 0000 1100 0000 0000 0000 0000 0000

 $e + bias = 00000001 = 1 \rightarrow e = 1 - 127 = -126$

Significand = 1.1

80C00000 = -1.1×2^{-126}

FAD00000: 1111 1010 1101 0000 0000 0000 0000 0000

 $e + bias = 11110101 = 245 \rightarrow e = 245 - 127 = 118$

Significand = 1.101

FAD00000 = -1.101×2^{118}

$$X = (-1.1 \times 2^{-126}) \times (-1.101 \times 2^{118}) = 10.0111 \times 2^{-8} = 1.00111 \times 2^{-7}$$

 $e + bias = -7 + 127 = 120 = 01111000$

$$X = 0011 1100 0001 1100 0000 0000 0000 0000 = 3C1C0000$$

$$\begin{array}{r} 1.1 \ 0 \ 1 \ x \\ \hline 1.1 \\ 1 \ 1 \ 0 \ 1 \\ 1 \ 0 \ 1 \\ \hline 1 \ 0.0 \ 1 \ 1 \ 1 \end{array}$$

- ✓ $X = 7B380000 \div C8A00000$:

7B380000: 0111 1011 0011 1000 0000 0000 0000 0000

 $e + bias = 11110110 = 246 \rightarrow e = 246 - 127 = 119$

Significand = 1.0111

7B380000 = 1.0111×2^{119}

C8A00000: 1100 1000 1010 0000 0000 0000 0000 0000

$$e + \text{bias} = 10010001 = 145 \rightarrow e = 145 - 127 = 18$$

$$\text{Significand} = 1.01$$

$$C8A00000 = -1.01 \times 2^{18}$$

$$X = -\frac{1.0111 \times 2^{119}}{1.01 \times 2^{18}} = -\frac{1.0111}{1.01} \times 2^{101}$$

$$\begin{array}{r} 000010010 \\ 10100 \overline{) 101110000} \\ \underline{10100} \\ 11000 \\ \underline{10100} \\ 1000 \end{array}$$

Alignment:

$$\frac{1.0111}{1.01} = \frac{1.0111}{1.0100} = \frac{10111}{10100}$$

$$\text{Append } x = 4 \text{ zeros: } \frac{101110000}{10100}$$

Integer division:

$$Q = 10010 \rightarrow Qf = 1.001$$

$$\text{Thus: } X = -1.001 \times 2^{101}, e + \text{bias} = 101 + 127 = 228 = 11100100$$

$$X = 1111 \ 0010 \ 0001 \ 0000 \ 0000 \ 0000 \ 0000 \ 0000 = F2100000$$

PROBLEM 2 (20 PTS)

- Complete the table for the following DFX formats (10 pts)

$$B = 2^{n-p_0-2}$$

$$\text{num0} \in [-2^{n-p_0-2}, 2^{n-p_0-2} - 2^{-p_0}]$$

$$\text{num1} \in [-2^{n-p_1-2}, -2^{n-p_0-2} - 2^{-p_1}] \cup [2^{n-p_0-2}, 2^{n-p_1-2} - 2^{-p_1}]$$

$$\text{Dynamic Range} = 20 \log_{10}(2^{n-2-p_1+p_0})$$

DFX format	p_0	p_1	Number of bits of significand	Boundary value	num0 range	num1 range	Dynamic Range (dB)
8_4_2	4	2	7	4	[-4,3.9375]	[-16,-4.25] U [4,15.75]	48.1648
16_8_4	8	4	15	64	[-64, 63.99609375]	[-1024, -64.0625] U [64, 1023.9375]	108.3708

- Convert the following signed fixed-point numbers in format [16 8] to the dual fixed-point format 16_8_4. (10 pts)

FX	3A.CD	9B.E6	7A.CE	CA.FE
DFX				

✓ 3A.CD:

$$0011 \ 1010.1100 \ 1101 \Rightarrow \text{To DFX 16_8_4 (num0): } 0011101011001101 = 3ACD$$

✓ 9B.E6:

$$1001 \ 1011.1110 \ 0110 \Rightarrow \text{To DFX 16_8_4 (num0): } 0001101111100110 = \text{not a num0!}$$

$$\Rightarrow \text{To DFX 16_8_4 (num1): } 1111100110111110 = F9BE$$

✓ 7A.CE:

$$0111 \ 1010.1100 \ 1110 \Rightarrow \text{To DFX 16_8_4 (num0): } 0111101011001110 = \text{not a num0!}$$

$$\Rightarrow \text{To DFX 16_8_4 (num1): } 1000011110101100 = 87AC$$

✓ CA.FE:

$$1100 \ 1010.1111 \ 1110 \Rightarrow \text{To DFX 16_8_4 (num0): } 0100101011111110 = 4AFE$$

PROBLEM 3 (20 PTS)

- Calculate the result of the following operations where the numbers are represented in dual fixed-point arithmetic. Note that the results must be in the same format. Include an overflow bit when necessary.

DFX Format: 8_4_2	Result	overflow		Result	overflow
EA+2E			EB-99		
D3+C5			65+FD		

✓ EA+2E:

$$\begin{array}{r}
 \begin{array}{cccccccc}
 & c_7 & c_6 & c_5 & c_4 & c_3 & c_2 & c_1 & c_0 \\
 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & + \\
 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & \\
 \hline
 11101.0110
 \end{array}
 \Rightarrow \text{To DFX 8_4_2 (num0): } 0101.0110 = 56 \quad \text{Overflow} = 0
 \end{array}$$

✓ D3 + C5:

$$\begin{array}{r}
 \begin{array}{cccccccc}
 & c_7 & c_6 & c_5 & c_4 & c_3 & c_2 & c_1 & c_0 \\
 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & + \\
 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & \\
 \hline
 100110.01
 \end{array}
 \Rightarrow \text{To DFX 8_4_2 (num0): } 01100000 = \text{not a num0!} \\
 \Rightarrow \text{To DFX 8_4_2 (num1): } 10011000 = \text{not a num1!} \quad \text{Overflow} = 1
 \end{array}$$

✓ EB-99:

$$\begin{array}{r}
 \begin{array}{cccccccc}
 & c_7 & c_6 & c_5 & c_4 & c_3 & c_2 & c_1 & c_0 \\
 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & - \\
 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & \\
 \hline
 10100.10
 \end{array}
 \Rightarrow \text{To DFX 8_4_2 (num0): } 01001000 = \text{not a num0!} \\
 \Rightarrow \text{To DFX 8_4_2 (num1): } 11010010 = D2 \quad \text{Overflow} = 0
 \end{array}$$

✓ 65+FD:

$$\begin{array}{r}
 \begin{array}{cccccccc}
 & c_7 & c_6 & c_5 & c_4 & c_3 & c_2 & c_1 & c_0 \\
 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & + \\
 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\
 \hline
 11101.1001
 \end{array}
 \Rightarrow \text{To DFX 8_4_2 (num0): } 01011001 = 59 \quad \text{Overflow} = 0
 \end{array}$$

PROBLEM 4 (40 PTS)

- Attach your Project Status Report (no more than 3 pages, single-spaced, 2 columns, only one submission per group). This report should contain the current status of your project. For formatting, use the provided template (Final Project - Report Template.docx). More details need to be provided:
 - ✓ Details, i.e., architecture of the AXI Interface
 - ✓ Allocation of tasks: i) software routine, and ii) reconfigurable hardware.
 - Software routine: provide top-level pseudo-code of your software application
 - If you plan to use run-time alterable hardware, indicate what tasks it will be doing.
 - ✓ Hardware Architecture: Include a Block Diagram with a complete I/O description (how many signals, how many bits per signal) and I/O mechanism.