

Homework 1

(Due date: September 19th)

Presentation and clarity are very important! Show your procedure!

PROBLEM 1 (14 PTS)

- Calculate the result of the additions and subtractions for the following fixed-point numbers.

UNSIGNED (6 pts)		SIGNED (8 pts)	
1.10010 + 0.101001	1.10101 - 0.0100111	10.101 + 1.10101	0.0011 - 1.101101
10.1011 + 1.1101	100.1 + 0.0100101	1001.0111 - 111.01001	101.1101 + 1.1010011

PROBLEM 2 (22 PTS)

- Multiply the following signed fixed-point numbers: (10 pts)

01.101 × 1.101001	100.001 × 01.10001	110.000 × 10.10101	0.11001 × 11.11011
----------------------	-----------------------	-----------------------	-----------------------

- Get the division result (with $x = 4$ fractional bits) for the following signed fixed-point numbers: (12 pts).

101.0101 ÷ 1.101	1.1011 ÷ 1.01101	10.0101 ÷ 01.11	0.10100 ÷ 110.101
---------------------	---------------------	--------------------	----------------------

PROBLEM 3 (16 PTS)

- We want to represent numbers between -128.87 and 127.12 . What is the fixed point format that requires the fewest number of bits for a resolution better or equal than 0.0015 ? (5 pts).
- We want to represent numbers between -255.12 and 256.91 . What is the fixed point format that requires the fewest number of bits for a resolution better or equal than 0.0025 ? (5 pts).

- Represent these numbers in Fixed Point Arithmetic (signed numbers). Select the minimum number of bits in each case.

-128.1875	-78.1255	-34.625	207.3125
-----------	----------	---------	----------

PROBLEM 4 (12 PTS)

- Complete the table for the following fixed point formats (signed numbers): (3 pts)

Fractional bits	Integer Bits	FX Format	Range	Dynamic Range (dB)	Resolution
9	3				
11	5				
15	9				

- Complete the table for the following floating point formats (which resemble the IEEE-754 standard) with 16, 24, 48 bits. Only consider ordinary numbers. (9 pts)

Exponent bits (E)	Significant bits (p)	Min	Max	Range of e	Range of significand
8	6				
10	13				
15	32				

PROBLEM 5 (36 PTS)

- Calculate the decimal values of the following floating point numbers represented as hexadecimals. Show your procedure.

Single (32 bits)		Double (64 bits)	
✓ 90DBD800	✓ 7F85B0AC	✓ DECAFC0FFEE80000	✓ ACCEDE90BEAD5000
✓ 800BEEF0	✓ 70DECADE	✓ 49A5DEAF8FAD8000	✓ 8009BEBEFACE8000
✓ 3CF398C0	✓ 90BEE950	✓ FFFACADE19801990	✓ 70800FEDCAB09000