University of Moratuwa

Circuits And System Design

EN3030

ASSIGNMENT : WEEK 01

Team : INFOS

Yasarathna D.D.K.B.	190719V	Ex - 01
Tharindu.O.K.D	190622R	Ex - 02
S.Sanjith	190562G	Ex - 03
K.Kajhanan	190286M	Ex - 04

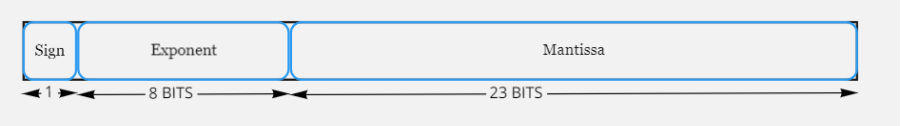
July 31, 2022

Contents

Executable codes can be found here

1 Floating Point Addition/Subtraction

1.1 IEEE-754 Single Precision Format



For the verification purposes h-schmidt, an online floatconverter is used.

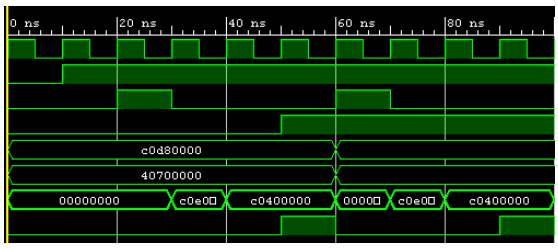
1.2 Algorithm

1. Compare the exponents of two numbers and find the absolute difference.
2. Shift the number with the smaller exponent to right by the amount of absolute difference to make the exponents of both numbers equal. Assign the prominent exponent as the exponent of the sum.
3. Check the signs of two numbers after performing sign conversion according to operation
   - If both are of the same sign add two corresponding mantissa and assign it to sum_mantissa.
   - Otherwise, subtract the matissa of number with 1 as sign bit from the other and assign the magnitude of it as sum_mantissa.
4. Assign the sum_sign according to sign of two numbers and above derived answer.
5. Shift the mantissa left until and 24th-bit position of the mantissa is one and reduce the exponent of the sum accordingly(normalization)
6. Now discard the implicit 1 and consider only the lower 23-bits of sum_mantissa.
7. Final Answer = {sum_sign, sum_exponent, sum_mantissa}

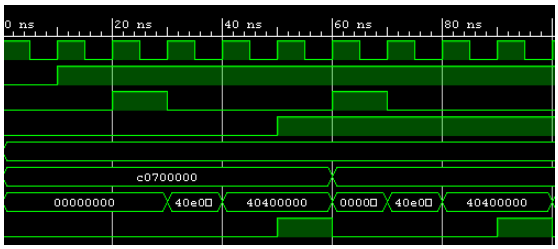
Visual implementation of the given algorithm as a flowchart is attached to this report in the end.

1.3 Results

The following results obtained from illustrated conditions ensure proper functioning of the unit.



-6.75+3.75 = -3 & -6.75-(-3.75) = -3



6.75-3.75 = 3 & 6.75+(-3.75) = 3

ROW 1 : Clock
ROW 3 : Validity of Answer
ROW 5 : Number 1
ROW 6 : Number 2
ROW 7 : Sum

Addition and Subtraction indicated above are choosen in a way to return the same result. During the true value of valid entry simulation returns same result for each operation.

A FP Add/Sub Implementation