

Name: **Owen Sullivan**

Circle your lab: 2 **3:30** 4:30 5:30

Introduction

The goal of this lab is to introduce and give you practice IEEE 754 Floating Point conversion.

Due:

Monday, November 21, 2022

Submission: Print and bring your submission to class on Monday. Late submissions will not be accepted

Instructions

While I know this would be much easier to do by hand. I have found grading to be easier when your answers are typed. Therefore, you **must** type your answers print and bring the document to class. If you are sick and cannot come to class you must email the homework to me by 9:00 a.m on November 21. Please read the entire document. Points will be deducted if you do not follow directions.

Part 1:

Watch one of the two sets of videos pertaining to Floating Point conversion from decimal to binary and binary to decimal. I found both of these to be helpful, so feel free to watch both.

https://www.youtube.com/watch?v=tx-M_rghuUA

<https://www.youtube.com/watch?v=4DfXdJdaNYs>

or

<https://www.youtube.com/watch?v=8afbTaA-gQQ>

<https://www.youtube.com/watch?v=LXF-wcoeT0o>

Part 2: 50 points

Following the instructions in the videos above. Convert the following floating-point number to binary.

79.35

Show your work. Also, explain what you are doing each step of the way. Your explanation does not have to be a long explanation. Only enough to let me know you understand what you are doing. If you do not show and explain your work, you will receive a 0 for the question.

Convert left side of decimal point to binary by doing normal integer decimal to binary conversion:

$$79 \text{ to binary} = 64+8+4+2+1 = 01001111$$

Convert right side of decimal point to binary using a multiplication method:

$$0.35 * 2 = 0.7$$

$$0.7 * 2 = 1.4 \text{ } \leftarrow \text{ignore the 1 in front of the decimal when multiplying again}$$

$$0.4 * 2 = 0.8$$

$$0.8 * 2 = 1.6$$

$$0.6 * 2 = 1.2$$

$$0.2 * 2 = 0.4$$

$$0.4 * 2 = 0.8 \text{ } \leftarrow \text{this is the point where we start seeing repetition}$$

New result:

$$1001111.0101100 \text{ } \leftarrow \text{the last four bits are repeating}$$

Move decimal point to after the first bit, similar to scientific notation:

$$1.0011110101100 * 2^6 \text{ } \leftarrow \text{decimal place moved 6 places}$$

$$s = 0 \text{ } \leftarrow \text{positive}$$

$$E = 6 \text{ } \leftarrow \text{exponent}$$

$$m = 0011110101100 \text{ } \leftarrow \text{mantissa}$$

Drawing:



Add 127 to exponent and convert total to binary:

$$\text{Exp} = 6 + 127 = 133 \text{ } \longrightarrow 128 + 4 + 1 = 10000101$$

Final result (extending mantissa with repeating bits until 23 bit cap):

$$0 \ 10000101 \ 00111101011001100110011$$

Part 3: 50 points

Following the instructions in the above video. Convert the following binary representation to a floating point number.

0100 0010 1010 0111 1000 0000 0000 0000

Show your work. Also, explain what you are doing each step of the way. Your explanation does not have to be a long explanation. Only enough to let me know you understand what you are doing. If you do not show and explain your work, you will receive a 0 for the question.

Sign is the first bit of the binary representation:

$s = 0$ <— positive

Exponent is the next eight bits of the binary representation:

$E = 1000\ 0101$

Mantissa is the remaining 23 bits of the binary representation:

$m = 01001111000000000000000$

Actual exponent for decimal notation is without bias:

$E = 1000\ 0101$ base 2 = 133 base 10

$e = 133 - 127 = 6$

Actual mantissa value for decimal notation is using 2^{-1} , 2^{-2} , 2^{-3} and so on depending on bit layout:

$m = 0100\ 1111\ 0000\ 0000\ 0000\ 000$

$m = 2^{-2} + 2^{-5} + 2^{-6} + 2^{-7} + 2^{-8} =$ (using calculator) 0.30859375

Plug values into formula:

$-1^s * (1 + m) * 2^e = -1^0 * (1 + 0.30859375) * 2^6 = -1.30859375 * 2^6$

Final result:

$1.30859375 * 2^6$

OR

83.75

The following link is a nifty tool you can use to check your work. You should understand that sometime online tools like this will round which could change the last one or two bits on the tool. So, if your answer has a different bit on the end that is perfectly fine. I am not saying this will be the case only letting you know this could happen.

<https://evanw.github.io/float-toy/>

Submission:

Print and bring you submission to class on Monday, November 21, 2022. I will not accept late submissions.