

## Study Questions (Chapter 02 – Part 3)

1. Pick an unsigned decimal number that has at least three digits after the decimal point.  
Convert this number into binary form.  $111.0101$   
Round this converted number to two binary digits after the binary point using  
a) round towards zero.  $111.01$   
b) round towards + infinity  $111.01$   
c) round to the nearest.  $111.01$   
Convert the rounded numbers back to decimal.  
Calculate the error in each case.  
Compare the results in these three cases.
2. Pick an unsigned decimal number with at least three digits after the decimal point and three digits before the decimal point.  
Convert this number into binary form.  
Normalize this converted number.  
Convert the result back to decimal to verify.  $E = 2^{-127}$
3. In a single-precision IEEE-754 floating-point format, what does it means to have the biased exponent = 0?
4. In a single-precision IEEE-754 floating-point format, what does it means to have the biased exponent = 255?  $\text{too big to be represented}$
5. In a single-precision IEEE-754 floating-point format, when is the excess-127 code used?
6. In a single-precision IEEE-754 floating-point format, when is the excess-126 code used?
7. In a single-precision IEEE-754 floating-point format, how many zeros can be represented?  $2$
8. In a single-precision, IEEE-754 floating-point format, how many unique +NaN numbers can be represented?  $2^{23}$
9. In a single-precision, IEEE-754 floating-point format, how many unique -NaN numbers can be represented?  $2^{23}$
10. In a single-precision, IEEE-754 floating-point format, how many unique NaN numbers can be represented?  $2^{24}$
11. In a single-precision IEEE-754 floating-point format, what is the largest normalized absolute number that can be represented?  $0 \ 111 \ 111 \ 011 \ 111 \ 111 \ 111 \ 111 \ 111 \ 111_2$
12. In a single-precision IEEE-754 floating-point format, what is the smallest normalized absolute number that can be represented?
13. Pick a decimal number.  
Convert this number into the single-precision IEEE-754 Floating-point format.  
Convert the number back to decimal to verify.