**Float Worksheet**

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The name of the floating world: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

For all of these problems, you will be converting to the IEEE 754 32-bit format (F32), which has 1 sign bit (S) followed by an 8-bit exponent (E) which is biased by 127, and a 23-bit mantissa (M). You will convert each number in parts and show your work to receive full credit. Remember that decimal places work in binary, following the same pattern as usual: 11.11 = 2 + 1 + ½ + ¼ = 3.75.

**Problem 1: Convert 4 to F32.**

S (Sign bit: 0 = positive, 1 = negative):

Convert 4 to binary:

Shift it so that it has a 1 before the decimal place:

M (Mantissa; drop the leading 1):

Write the number of shifts here:

Write down E (127+right shifts or –left shifts):

Write down the 32-bit number in binary here (S E M):

**Problem 10: Convert -5.5 to F32**

S (Sign bit: 0 = positive, 1 = negative):

Convert 5.5 to binary:

Shift it so that it has a 1 before the decimal place:

M (Mantissa; drop the leading 1):

Write the number of shifts here:

Write down E (127+right shifts or –left shifts):

Write down the 32-bit number in binary here (S E M):

**Problem 11: Convert 0.0625 to F32.**

S (Sign bit: 0 = positive, 1 = negative):

Convert 0.0625 (1/16) to binary:

Shift it so that it has a 1 before the decimal place:

M (Mantissa; drop the leading 1):

Write the number of shifts here:

Write down E (127+right shifts or –left shifts):

Write down the 32-bit number in binary here (S E M):

**Problem 100: Convert -13.375 to F32.**

S (Sign bit: 0 = positive, 1 = negative):

Convert 13.375 to binary:

Shift it so that it has a 1 before the decimal place:

M (Mantissa; drop the leading 1):

Write the number of shifts here:

Write down E (127+right shifts or –left shifts):

Write down the 32-bit number in binary here (S E M):

**Problem 101: Convert -256.75 to F32**

S (Sign bit: 0 = positive, 1 = negative):

Convert 256.75 to binary:

Shift it so that it has a 1 before the decimal place:

M (Mantissa; drop the leading 1):

Write the number of shifts here:

Write down E (127+right shifts or –left shifts):

Write down the 32-bit number in binary here (S E M):

**Problem 110: Write down any NaN in F32:**