

Assignment: 1

CS101

To convert $-9/2$ to binary using the 8-bit floating-point notation, follow these steps:

1. Convert the absolute value of 9 to binary: 1001.
2. Convert $1/2$ to binary: 0.1.
3. Combine the integer and fractional parts: 1001.1.
4. Negate the result since the decimal fraction is negative: -1001.1.

To normalize the binary number, shift the decimal point to the left until there is only one non-zero digit to the left of the decimal point. This gives us -1.0011×2^3 .

Now, determine the sign, exponent, and mantissa:

- The sign bit is 1 since the number is negative.
- The exponent is 3, biased by adding 7 (excess-3) to get 10 (binary 1010).
- The mantissa is the binary fraction 0011, truncated to 4 bits.

Combining the sign, exponent, and mantissa, we get the 8-bit floating-point representation of $-9/2$ as 1 1010 0011.

The smallest negative value that can be represented using the 8-bit floating-point notation is when the sign bit is 1, the exponent is 000 (excess-3 notation), and the mantissa is 0000. This gives us the representation 1 000 0000.

For the largest positive value, assuming a standard format with 1 bit for the sign, 3 bits for the exponent (excess notation), and 4 bits for the mantissa, we set the sign bit to 0. The exponent bits are set to the maximum value, which is 111 (binary representation of 7), and the mantissa bits are set to all ones (1111). This results in the representation 0 111 1111, which represents the largest positive value in the 8-bit floating-point notation.