4CCS1CS1 Computer Systems

Tutorial

Josh Murphy
Questions? CS1 Discussion Forum

IEEE 754 format conversions

- Convert the following IEEE 754 single-precision numbers to the equivalent decimal numbers:
  - 11000001110101100100000000000000
  - 00111110010000000000000000000000
  - 11111111100000000000000000000000
- Convert the following decimal numbers to the equivalent IEEE 754 single-precision numbers:
  - 733_{10}
  - -17.875
  - 0_{10}
- **Want more practice?** `https://www.h-schmidt.net/FloatConverter/IEEE754.html`.

IEEE 754 format arithmetic

- Perform the following computations in IEEE 754 single-precision format. Give your answers in IEEE 754 single-precision format.
  - 11000011101010100000000000000000 + 11000010100100001000000000000000
  - 11000001100011000000000000000000 × 01000010111111000000000000000000

A new floating-point format

- For the following questions we will use a new floating-point format. The format is 8-bits: it has a sign bit, a 3-bit exponent, and a 4-bit significand. Unlike IEEE-754 format, the exponent is **not** stored using a bias, instead the exponent uses two's complement notation, and exponents of all 0s and all 1s are not interpreted as special values.
  - What is the largest positive and the smallest positive number that can be stored in this system if the significand is normalised in the same way as the significand is normalised in IEEE 754 format? What is the range?
  - If we wanted to use a bias representation for the exponent instead of using two's complement, what bias value should be used?
  - Add the following numbers assuming they are using the bias format for the exponent. Store the result in the same format. 01111000 + 01011001
  - Calculate the relative error of the above result, in any. **Hint:** you need to compare the value that is stored to the true value of the calculation.

Error

- In the lecture videos this week, we saw that the closest representation of $4{,}039{,}944{,}879_{10}$ in IEEE-754 single-precision was:

  0 10011110 11100001100110010101011

  However, this is not an exact representation. What is the relative error? **Hint:** convert the IEEE-754 representation back to decimal to the the stored value, and compare it to the true value.