

## Class 4: Representing Information: Floating Point

- Floating Point Representation
- Floating Point Arithmetic
- IEEE Floating Point Formats
- ASCII

## Simple Floating Point Format

- Scientific notation:

## Floating Point Addition and Subtraction

- Review:

## Floating Point Multiplication and Division

- Multiplication:

- Division:

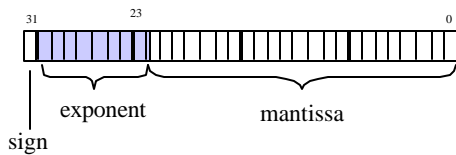
## Comparing Floating Point Numbers

- Be careful!

## IEEE Floating Point Formats

- Three floating point formats:
  - 32 bit single precision
    - 1 sign bit
    - 8 exponent bits
    - 24 mantissa bits
  - 64 bit double precision
    - 1 sign bit
    - 11 exponent bits
    - 53 mantissa bits
  - 80 bit bit extended precision
    - 1 sign bit
    - 15 exponent bits
    - 64 mantissa bits

## Single Precision Floating Point



- Sign –
- Exponent –
- Mantissa –

## Normalization

- How can you store a 24 bit mantissa in 23 bits?

## Normalization, cont.

- How does it work?

## Converting to Single Precision

1. Convert to binary (ignore sign for now)
2. Normalize
3. Compute Exponent using excess 127
4. Throw away high bit of mantissa and extend to 23 bits
5. Set (or clear) the sign bit as appropriate

## Converting to Single Precision, cont.

- Example: 27.4 decimal

## Another example

- 16.2 decimal

## Converting from Single Precision into Decimal Floating Point

1. Convert to Binary
2. Pull out the components:
  - Sign
  - Exponent
  - Mantissa
3. Convert the exponent
4. Convert the mantissa
5. Create the binary result
6. Convert into decimal
7. Add the sign

## Converting from Single Precision into Decimal Floating Point

- Example: BD500000h

- BE400000h in IEEE?

## ASCII

- ASCII (American Standard Code for Information Interchange) is commonly used to represent characters sent to a display
- This is what is used to display information (letters, symbols, numbers, and control characters).
- Examples:
  - 'A' = 41h = 65<sub>10</sub>
  - 'a' = 61h = 97<sub>10</sub>
  - '!' = 21h = 33<sub>10</sub>
  - '1' = 31h = 49<sub>10</sub>
  - CR (carriage return) = 0Dh = 13<sub>10</sub>