

Homework Lecture 4

- 3.1
- 3.3
- 3.4

Exercise 1

Multiply the following twos complement numbers:
1011 and 0110

Exercise 2

There is not a unique floating point number system. DEC introduced a 32 bit floating point number system with the following properties (base 2):

- Fraction: 23 bits and additional 1 hidden bit. Point is left of hidden bit
- Exponent: 8 bits in excess 128 code
- Sign bit (0 is positive, 1 is negative).
- Number is not normalized if exponent field is filled with zero's. In that case the represented value (independent of sign and fraction field) is zero.
- Rounding style is truncation.

Questions:

For a) until f) the normalized numbers:

- a) Max decimal value of the mantissa (M_{\max})
- b) Min decimal value of the mantissa (M_{\min})
- c) Max decimal value of the exponent (E_{\max})
- d) Min decimal value of the exponent (E_{\min})
- e) Largest positive decimal value that can be representend (V_{\max})
- f) Smallest positive decimal value that can be represented (V_{\min})

- g) What is the smallest positive decimal numbers that can be represented?
And what is the next positive value that can be exactly represented?
- h) What is the representation of $2^{9/16}$?
- i) What is the representation of the decimal value 0.2 ?
- j) What decimal value is represented with the pattern:
1 00000111 110100000000000000000001

Exercise 3

The IEEE FP standard 754 also defines a "double precision" floating point number system. Properties:

- total number of bits 64
- sign: single bit (left most bit)
- exponent: 11 bits, excess 1023 code, all zero and all one is used for special numbers.
- fraction 52 bits. A hidden bit is used (not included in the 52 bits). Point is right of hidden bit.

Question: the same questions as a) to f) of previous exercise.