Errors

Computer Arithmetic

X = 0.1

While (x != 1.0)

{

x = x + 0.1

}

(0.1)10 = 0.00011001100110011…

(78.25)10

convert 78 to binary

78.25 = 78 + 0.25

convert each separately to binary

Keep dividing by 2 and take remainders

Byte = 8 bits

Word = 4 bytes = 32 bits

Dword = 8 bytes = 64 bits

8 bits exponent, 23 bit mantissa, -127 <= e <= 128

Absolute error = true value – approximate value

Relative error = absolute error / true value

Rel(x) = E / Xt

Rel(y) = N / Yt

Rel(2) = (E + N) / (Xt + Yt)