ESO 208A
ESO 218

LECTURE 3

AUGUST 2, 2013

FLOATING POINT REPRESENTATION

- m*b^e
- m = the mantissa
- b = base
- e = exponent

156.78 = 0.15678*10^3

FLOATING POINT PRESENTATION

- MANTISSA normalized {for leading zero digits}
- 1/34 = 0.029411765…
- 0.0294*10^0
- 0.2941*10^-1
- ABSOLUTE VALUE OF ‘m’ IS LIMITED
- $1/b \le m < 1$

EXAMPLE

- CREATE A FLOATING-POINT NUMBER SET FOR A MACHINE THAT STORES INFORMATION USING 7-BIT WORDS. USE THE FIRST BIT FOR THE SIGN OF THE NUMBER, THE NEXT THREE FOR SIGN AND MAGNITUDE OF THE EXPONENT, AND THE LAST 3 FOR THE MAGNITUDE OF THE MANTISSA.

SOLUTION

- 7 bits
- 1: sign of number
- 2: sign of exponent
- 3-4: magnitude of exponent
- 5-7: mantissa

- Consider the smallest possible positive number
- 0 1 1 1 1 0 0 {1/b≤m<1}
- Exponent=-3
- Mantissa=2^{-1}=0.5
- +0.5*2^{-3}= 1*2^{-1}+0.0625

Next number

0111101=(2^{-1}+2^{-3})*2^{-3}=0.078125

- Numbers

0.062500; 0.078125; 0.093750; 0.109375

0.125000; 0.156250; 0.187500; ---------------

Characteristics:

1. Limited range of quantities {underflow/ overflow}
2. Finite number of quantities {chopping/ rounding}
3. Interval between numbers increases

   $\{\xi = b^{1-t}\}$

Arithmetic Manipulations

- $0.1557*10^{1} + 0.4381*10^{-1}$

0.1557	10^1
0.004381	10^1
0.160081	10^1
0.1600	10^1

Home work

- Addition of large and small number
- Subtraction
- Multiplication
- Division

Problems: 3.7 and 3.11 {pages 76 and 77}

Truncation errors

- Taylor Series

$f(x+h)=f(x)+f'(x)h+f''(x)h^2/2!+ \ldots.$

Taylor series approximation of a polynomial

- Use zer -4th order Taylor series expansion to approximate

$f(x)=-0.1x^4-0.15x^3-0.5x^2-0.25x+1.2$

Use x = 0 to 1

Solution

f(0)=1.2; f(1)=0.2

By Taylor Series,

f(1) = 1.2; 0.95; 0.45; 0.2

Differentiation

- Taylor Series

$f(x+h)=f(x)+f'(x)h+f''(x)h^2/2!+ . . . .$

- $f'(x)=\{f(x+h)-f(x)\}/h-$ ---O(h)

- Forward finite difference
- Backward finite difference
- Central finite difference

Error propagation

- $f(x)$
- x is with error
- How it affects $f(x)$

- If x is with error Δx, what is the error associated with $f(x)$

- Taylor Series

$f(x+\Delta x)=f(x)+f'(x)\Delta x+f''(x)\Delta x^2/2!+ \ldots.$

Error in the function $=f(x+\Delta x)-f(x)$

$=f'(x)\Delta x+f''(x)\Delta x^2/2!+ \ldots.$

$\Delta f \cong f'(x)\Delta x$

Multi-variable function $g(x, y, z)$

$\Delta g \cong \delta g/\delta x * \Delta x+\delta g/\delta y * \Delta y+\delta g/\delta z * \Delta z$

Example

- The deflection y of the top of a sailboat mast is $y = FL^4/(8EI)$

F = uniform side loading; L = height; E=modulous of elasticity; I=moment of inertia. Estimate the error in y given the following data.

F=750 errorF=30

L=9 errorL=0.03

E=7.5*10^9 errorE=5*10^7

I=0.0005 errorI=0.000005