ESO 208A/ESO 218

LECTURE 2

JULY 31, 2013

ERRORS

- MODELING OUTPUTS
- QUANTIFICATION
- TRUE VALUE
- APPROXIMATE VALUE

ROUND-OFF ERROR

- LIMITATION OF A COMPUTER
- QUANTITY WITH FINITE NUMBER OF DIGITS

TRUNCATION ERROR

- NUMERICAL METHOD
- APPROXIMATION OF MATHEMATICAL OPERATIONS/QUANTITY

SIGNIFICANT FIGURES

- SIGNIFICANT DIGITS OF A NUMBER ARE THOSE THAT CAN BE USED WITH CONFIDENCE
- CERTAIN DIGITS+ESTIMATED DIGIT(1)
- ESTIMATED DIGIT = HALF OF SMALLEST SCALE DIVISION

- 9753.4 – 973.5 ➔ 9753.45
- 60 – 65 ➔ ?
- IMPLICATIONS OF SIGNIFICANT DIGITS

Significant figures

- 0.00001845 {4}
- 0.0001845 {4}
- 0.001845 {4}
- 4.53*10^4 {3}
- 4.530*10^4 {4}

ACCURACY

- HOW CLOSELY A COMPUTED VALUE AGREES WITH THE TRUE VALUE
- INACCURACY ALSO KNOWN AS BIAS

PRECISION

- HOW CLOSELY INDIVIDUAL COMPUTED VALUES AGREE WITH EACH OTHER
- IMPRECISSION IS ALSO UNCERTAINTY

TERMINOLOGY

- E_t =TRUE VALUE-APPROXIMATION
- ε_t = TRUE ERROR*100/TRUE VALUE
- ε_a = APPROXIMATE ERROR*100/APPROXIMATION

ERRORS

- ITERATIVE METHODS
- $Abs(\varepsilon_a) < \varepsilon_s$
- $\varepsilon_s = (0.5*10^{2-n})\%$

Example

- e^x =1+x+x^2/2!+x^3/3!+……+x^n/n!
- Estimate e^0.5
- Add terms until the relative approximate error falls below specified tolerance confirming to 3 significant figures

Solution

- ε_s = (0.5*10^{2-n})%
- Use n = 3
- We get ε_s = 0.05%
- True value = e^0.5 = 1.648721…..

terms	result	ε_t	ε_a
1	1	39.3	-
2	1.5	9.02	33.3
3	1.625	1.44	7.69
4	1.645833333	.175	1.27
5	1.648437500	.0172	.158
6	1.648697917	.00142	.0158

NUMBERS

- NUMBER SYSTEM: DECIMAL, OCTAL ,BINARY
- POSITIONAL NOTATION
- 86409 = 8*10^4+……..
- 10101101(BINARY)=
- 2^7+2^5+2^3+2^2+2^0
- 173 (DECIMAL)

INTEGER

- SIGNED MAGNITUDE METHOD
- . . . . . .
- S ----------NUMBER----------
- 1 0 0 0 0 0 0 0 1 0 1 0 1 1 0 1

EXAMPLE

- DETERMINE THE RANGE OF INTEGERS IN BASE 10 THAT CAN BE REPRESENTED ON A 16-BIT COMPUTER.

SOLUTION

- THE FIRST BIT HOLDS THE SIGN
- REMAINING 15 BITS CAN HAVE 0 TO 11111….
- HIGHEST NUMBER IS 2^14+2^13+…….
- 2^15-1 = 32767
- RANGE IS -32767 TO +32767
- CONSIDER: 0,00000000…. AND 1,000000….
- -ZERO BECOMES THE NEXT NEGATIVE NUMBER
- RANGE IS -32768 TO +32767

FLOATING POINT REPRESENTATION

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

156.78 = 0.15678*10^3

-

FLOATING POINT PRESENTATION

- . . . . . .
- sn <--exp---><-----mantissa---------->
- 16 bits
- 1: sign of number
- 2-4: exponent; 2: sign of exponent
- 5-16: mantissa

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
- $1/b \leq m < 1$