MAE384 HW1

Comments:

Prob.1
We will discuss this problem in class.

Prob. 2
Many of you attempted to divide 256.1875 by 265 before proceeding to the standard binary decomposition. In this case, the correct procedure is

Step 1: $$256.1875 = \frac{256.1875}{256} \times 256 = 1.000732421875 \times 2^8 \quad .$$

Step 2: $1.000732421875 = 2^1 + 2^{-11} + 2^{-12}$; $8 = 2^3$
$\Rightarrow$ $1.000732421875 = 1.000000000011 \times 2^{1000}$

In step 1, all digits in "1.000732421875" must be retained. Rounding the number shall lead to an incorrect answer. (Some of you obtained "1.000732422", which I realized was due to insufficient precision of your calculators.) For this problem, it is easier to first decompose 256.1875 into 256 + 0.125+ 0.0625 = $2^8+2^{-3}+2^{-4}$ = 100000000.0011 in binary form. Then, advancing the floating point by 8 places to the left, one immediately obtains the desirable final answer. See attached sample solution for yet another alternative.

Prob 3 and 4 are straightforward.

Sample solutions to Prob 2, 3, and 4 are attached.

Sample solution, Prob. 2 (Thanks to Anton Pestka)

256.1875 ⇒ CONVERT TO BINARY IN FLOATING POINT REPRESENTATION:

$256 = 2^8 = 100000000$ (BINARY)

NOW CONVERT .1875 INTO BINARY AND ADD:

$$\frac{.1875}{2^{-3}} \cdot 2^{-3} = 1.5 \times 2^{-3} = 1.1 \times 2^{-3} = .0011 \text{ (bin)}$$

$(2^{-1} = .5)$

NOW ADD THE TWO:

$$\begin{array}{r} 100000000. \\ + \quad .0011 \\ \hline = 100000000.0011 \end{array}$$

$$\underset{2^8 +}{1}00000000.00\underset{2^{-3} + 2^{-4}}{1\;1} \text{ (bin)} = 256 + .125 + .0625$$

$= 256.1875$

SO: $100000000.0011 = 256.1875$

$1000000000.0011 = 1.000000000011 \times 2^8$

$8 = 2^3 + 0 + 0 + 0 = 1000$ (BINARY)

SO NOW THE ANSWER IS:

$$\boxed{1.000000000011 \times 2^{1000}}$$

Sample solution, Prob 3 (Thanks to Aishwarya Stanley)

3) $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$

Evaluate e^{-2}

a) first four terms:

$e^{-2} = 1 - 2 + \frac{(-2)^2}{2!} + \frac{(-2)^3}{3!}$

$e^{-2} = 1 - 2 + \frac{4}{2} - \frac{8}{6} = \underline{\underline{-0.333333333}}$

b) first 6 terms:

$e^{-2} = 1 - 2 + \frac{(-2)^2}{2!} + \frac{(-2)^3}{3!} + \frac{(-2)^4}{4!} + \frac{(-2)^5}{5!}$

$e^{-2} = 1 - 2 + 2 - \frac{8}{6} + \frac{16}{24} - \frac{32}{120}$

$e^{-2} = 1 - \frac{4}{3} + \frac{4}{6} - \frac{8}{30}$

$e^{-2} = 1 - \frac{4}{3} + \frac{2}{3} - \frac{8}{30} = 1 - \frac{2}{3} - \frac{4}{15} = \frac{45-30-12}{45} = \frac{3}{45} = \underline{\underline{0.066666666}}$

c) first 8 terms

$e^{-2} = 1 - 2 + \frac{(-2)^2}{2!} + \frac{(-2)^3}{3!} + \frac{(-2)^4}{4!} + \frac{(-2)^5}{5!} + \frac{(-2)^6}{6!} + \frac{(-2)^7}{7!}$

$e^{-2} = 1 - 2 + 2 - \frac{8}{6} + \frac{16}{24} - \frac{32}{120} + \frac{64}{720} - \frac{128}{5040}$

$e^{-2} = 1 - \frac{4}{3} + \frac{2}{3} - \frac{8}{30} + \frac{4}{45} - \frac{8}{315}$

$e^{-2} = 1 - \frac{2}{3} - \frac{4}{15} + \frac{4}{45} - \frac{8}{315} = \underline{0.13015873}$

d) Evaluate e^{-2} using built-in exponential function

$\underline{\underline{e^{-2} = 0.135335283}}$ —— True value

Evaluate truncation errors for (a)-(c)

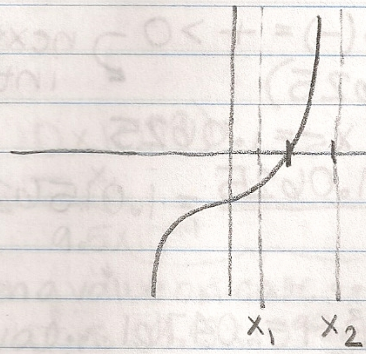
(a) $E^{TR} = 0.135335283 - (-0.333333333) = \underline{\underline{0.468668616}}$

(b) $E^{TR} = 0.135335283 - 0.066666666 = \underline{\underline{0.068668617}}$

(c) $E^{TR} = 0.135335283 - 0.13015873 = \underline{\underline{5.176553837 \times 10^{-3}}}$

Sample solution, Prob 4 (Thanks to Jennifer Gamboa)

4. $x^3-1=0$



① $x_1=0.5$ $x_2=2$

$x_3=\frac{2+0.5}{2}=\frac{2.5}{2}=1.25$

$f(x_1)=(0.5)^3-1=-.875$

$f(x_3)=(1.25)^3-1=.953125$

$f(x_1)\cdot f(x_3)=-.834\ldots<0$ next interval

$(0.5, 1.25)$

② $x_1=0.5$ $x_3=1.25$

$x_4=\frac{1.25+0.5}{2}=.875$

$f(x_1)=-.875$

$f(x_4)=(.875)^3-1=-.33007\ldots$

$f(x_1)\cdot f(x_4)=(-.875)(-.33007\ldots)=.2888\ldots>0$ next interval

$(0.875, 1.25)$

③ $x_3=1.25$ $x_4=0.875$

$x_5=\frac{1.25+0.875}{2}=1.0625$

$f(x_3)=.953125$

$f(x_5)=(1.0625)^3-1=.1994628\ldots$

$f(x_3)\cdot f(x_5)=(.953125)(.1994628\ldots)=.190113\ldots>0$ next interval

$(0.875, 1.0625)$

④ $x_4=0.875$ $x_5=1.0625$

$x_6=\frac{0.875+1.0625}{2}=.96875$

$f(x_4)=-.33007\ldots$

$f(x_6)=(.96875)^3-1=-.09085\ldots$

Continued to next page

Prob 4, continued

$f(x_4) \cdot f(x_6) = (-) \cdot (-) = + > 0$ next interval

$(0.96875, 1.0625)$

5 $x_6 = 0.96875 \quad x_5 = 1.0625$

$$x_7 = \frac{0.96875 + 1.0625}{2} = 1.015625$$

$f(x_6) = -.09085\ldots$

$f(x_7) = (1.015625)^3 - 1 = .04761\ldots$

$f(x_6) \cdot f(x_7) = (-.09085) \cdot (.04761) = -.004325\ldots < 0$ next interval

$(0.96875, 1.015625)$

$$x_8 = \frac{0.96875 + 1.015625}{2} = \boxed{0.9921875}$$

True error: $\text{True error} = x_{TS} - x_{NS}$

$= 1 - 0.9921875$

$= \boxed{0.0078125}$

$$\text{True relative error} = \left| \frac{x_{TS} - x_{NS}}{x_{TS}} \right|$$

$$= \left| \frac{1 - 0.9921875}{1} \right| = \boxed{0.0078125}$$

where $x_{TS} = 1$ and $x_{NS} = 0.9921875$