ANSWERS

CSci 5302   Homework 1   Due: 30 Jan 2018

• There are 15 subquestions over 3 questions worth a total of 100 points over 2 pages.

1. Assume you use the following 3 digit decimal arithmetic \( \pm X.YZ \times 10^{\pm W} \) where

- \( X \) is any decimal digit \( 1, \ldots, 9 \) (0 is allowed only for true zero)
- \( Y,Z,W \) are any decimal digits \( 0, \ldots, 9 \).

Assume that we use the “round-to-even” rule, that overflows become inf’s and that underflows become zeroes. The “round-to-even” rule says that any result that is not already representable should be rounded to the nearest representable number. If the result is exactly half way between two representable numbers, it should be rounded to the representable number that ends in an even digit. Answer the following questions

1a. (5 pts) Compute \( 2.04 \times 10^{+1} + 4.50 \times 10^{-1} \).

1b. (5 pts) Compute \( 9.04 \times 10^{+3} + 9.55 \times 10^{+2} \).

1c. (5 pts) Compute \( 9.04 \times 10^{+9} + 9.55 \times 10^{+8} \).

1d. (5 pts) Compute \( 1.44 \times 10^{-8} - 1.40 \times 10^{-8} \).

1e. (5 pts) Compute \( 1.44 \times 10^{-8} - 1.40 \times 10^{-8} \) in a modified system that allows for subnormal numbers.

1f. (5 pts) Give the largest positive floating point number \( s \) such that \( fl(9.99 \times 10^{-1} - s) = 9.99 \times 10^{-1} \).

1g. (5 pts) Give an example of a positive floating point number of the form \( a = 1.00 \times 10^k \) such that the floating product \( fl(a \times a) > 0 \), but \( fl(a \times a \times a) = 0 \).

1h. (5 pts) How many distinct positive numbers can be represented in this system?

1i. (5 pts) How many distinct additional positive numbers can be represented in a modified system that allows for subnormal numbers. Just count the numbers that can be represented in the modified system which cannot be represented in the original system as given above.

2. In the following, use the same 3 digit floating point system as in the previous question for all calculations.

2a. (12 pts) Compute the mathematically equivalent expressions \( y^2 - x^2 \) and \( [y-x] \cdot [y+x] \) with \( x = 2.22 \times 10^2 \) and \( y = 2.23 \times 10^2 \). Fill in the table of intermediate values

- (a) \( y^2 = 4.93 \times 10^4 \)
- (b) \( x^2 = 4.93 \times 10^4 \)
- (c) \( y^2 - x^2 = 4.98 \times 10^2 \)
- (d) \( y - x = 1.00 \times 10^0 \)
- (e) \( y + x = 4.45 \times 10^2 \)
- (f) \( (y-x)(y+x) = 4.98 \times 10^2 \)
2b. (8 pts) For the final results (c) and (f), answer the following

(i) Which of (c) (f) are accurate to the full accuracy of the underlying arithmetic, if any?

(ii) Between (c) & (f), treat the more accurate answer as the “true” answer and the other as the “computed” answer. Give the absolute error in the computed answer.

\[
|c(0.01) - f(0.01)| = 4.5 - 4.5 = 0.0001 \quad \text{Absolute Error} = 0.0001
\]

(iii) Give the corresponding relative error between the computed and the true answers;

\[
\text{Relative Error} = \frac{|c(0.01) - f(0.01)|}{c(0.01)} = \frac{0.0001}{4.5} \approx 0.0000222222
\]

(iv) the number of digits of accuracy in the computed answer.

\[
\# \text{digits accuracy} = 4
\]

3. Let \( p(x) = x^3 - 2.5x^2 + 2x - 0.501002 \).

3a. (10 pts) You buy a used calculator at a flea market which claims to be able to find roots of polynomials. When you use it to find the root of this \( p(x) \), the answer it gives is \( \hat{x} = 1.0 \). What is the backward error in this answer?

\[
p(1) = -0.01002 \quad \text{Backward Error} = 0.001002
\]

3b. (5 pts) Pick one of the following intervals you could use to start bisection to find the root of the polynomial of question 3: \([-1, 0], [0, 1], [1, 2]\), assuming your calculator is more accurate than the one from the flea market. Identify your chosen interval and show why you know the interval will work. Starting with your chosen interval, what would be the interval you’d have after 2 iterations of bisection?

\[
p(-1) = -6 \quad p(0) = -5.5 \quad p(1) = 0.001 \quad p(2) = 1.498 \quad \text{answer} = [1.52]
\]

3c. (10 pts) Apply 2 iterations of Newton’s method to find a root of \( p(x) = x^3 - 2.5x^2 + 2x - 0.501002 = 0 \) with the starting value \( x_0 = 1 \) and then again with \( x_0 = 0.9 \). Show the formula for Newton’s method and a short table of values

\[
\begin{array}{c|c|c|c}
 k & x_k & p(x_k) & p'(x_k) \\
 0 & x_0 & \vdots & .993 \quad .003 \quad -7e-2 \\
 1 & x_1 & \vdots & 1.9414 \quad 4.84 \quad -4.83e-2 \\
 2 & x_2 & \vdots & 1.95001 \quad 2.48e-4 \quad -4.5e-2 \\
\end{array}
\]

Use \( x_0 = 1 \) and again with \( x_0 = 0.9 \). If it fails, explain why. Stop after 2 iterations.

3d. (10 pts) Treat \( x_t = x_2 \) from one of the previous results as a true root pf \( p(x) \). Let \( p(x) = x^3 - 2.5x^2 + 2x - 0.5 \) be an approximate polynomial. Evaluate \( p(x_t), p(\hat{x}), p(\hat{x}_t), \) and \( p(\hat{x}) \). Use these four values to give an estimate of the condition number of the solution to \( p(x) = 0 \), pretending that \( p(x) \) is the exact polynomial and \( x_t \) is the correct answer.

\[
p(x_t) = p(1) = -1.1e-3 \\
p(\hat{x}) = p(1) = 0 \\
p(x_t) = 1.1e-3 \\
\text{For Error} = \left| x - x_t \right| = 0.05 \\
\text{Back Error} = \left| p(\hat{x}) - p(x_t) \right| = \left| 1.1e-3 \right| = 1.1e-3 \\
\text{Cond #} = \frac{\text{For Error}}{\text{Back Error}} = \frac{0.05}{0.005} = 45.
\]