MATH 351, MIDTERM EXAM, FALL 2019

Name	Student ID

- Write your solution to each problem in a readable manner. Circle your final results.
- Show all your work. Answers not supported by valid arguments will get little or no credit.
- Doing correctly Problems 1,2,3,4,5,6 will result in 100% credit of the exam. You can earn extra credit by doing Problem 7.

Problem	Possible points	Earned points
1	10	
2	10	
3	10	
4	10	
5	10	
6	5	
7	5	
Total	60	

Some formula:

$$n \geq \log_2\left(\frac{b-a}{\epsilon}\right) - 1,$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})},$$

$$|x_{n+1} - \alpha| \leq C|x_n - \alpha|^p.$$

[This page is intentionally left blank]

Consider a binary floating-point system described as follows. A number x is represented approximately as $x \approx \sigma \cdot \bar{x} \cdot 2^e$ where

- If $1 \le E \le 14$ then

$$\begin{aligned} \sigma &= \begin{cases} 1 & \text{if} \quad x \ge 0, \\ -1 & \text{if} \quad x < 0, \end{cases} \\ e &= E - 7, \\ \bar{x} &= (1.a_1a_2a_3)_2 \text{ (rounding to truncate)} \end{aligned}$$

- If $E = 0$ then $e = -6$ and $\bar{x} = (0.a_1a_2a_3)_2$ (rounding to truncate).
- If $E = 15$ then the bit sequence represents $\pm\infty$ (depending on the sign σ).

Problem 1. (10 points) Represent the number 2.2 in this floating-point system.

First, we convert 2.2 from decimal system to binary system.

$$2.2 = 2 + 0.2$$

$$2 = (10)_2$$

$$\left.\begin{aligned} 0.2 \times 2 &= 0.4 \longrightarrow 0 \\ 0.4 \times 2 &= 0.8 \longrightarrow 0 \\ 0.8 \times 2 &= 1.6 \longrightarrow 1 \\ 0.6 \times 2 &= 1.2 \longrightarrow 1 \\ 0.2 \times 2 &= 0.4 \longrightarrow 0 \\ &\cdots \end{aligned}\right\} \quad 0.2 = (0.\underbrace{0011}\underbrace{0011}\underbrace{0011}\cdots)_2$$

Thus, $2.2 = (10.\underbrace{0011}\underbrace{0011}\underbrace{0011}\cdots)_2$.

We have

$$2.2 = (1.00011 0011\cdots)_2 \times 2^1 \quad \text{(shifting the floating point)}$$

$$\approx (1.001)_2 \times 2^1 \quad \text{(rounding)}$$

Problem 2. (10 points) Let $x = (1.101)_2 \times 2^2$ and $y = (1.010)_2 \times 2^3$. Perform the operations $x + y$ and $x \cdot y$ in the floating-point system given on the previous page.

* Compute $x+y$:

$$x = (0.1101)_2 \times 2^3$$
$$y = (1.010)_2 \times 2^3$$

Summing the mantissa:

$$\begin{array}{r} 0.1101 \\ +\,1.0100 \\ \hline 10.0001 \end{array}$$

Then $x+y = (10.0001)_2 \times 2^3$

$= (1.00001)_2 \times 2^4$

$x+y \approx (1.000)_2 \times 2^4$

* Compute xy:

Multiplying the mantissa:

$$\begin{array}{r} 1.101 \\ \times\;\; 1.010 \\ \hline 0000 \\ 1101 \\ 0000 \\ 1101 \\ \hline 10.000010 \end{array}$$

Thus, $xy = (10.000010)_2 \times 2^{2+3}$

$= (1.000001)_2 \times 2^6$

$xy \simeq (1.000)_2 \times 2^6$

Problem 3. (10 points) Let $f(x) = x^3 + 3x - 1$.

(a) Use bisection method (by performing 4 iterations) to find a approximate root of f on the interval (0,1).

(b) How many iterations are needed to obtain an approximate root with error at most $\epsilon = 10^{-6}$?

.

$$a_0 = 0, \qquad b_0 = 1, \qquad c_0 = 0.5$$

$$\rightsquigarrow \quad f(a_0) = -1 < 0 \qquad f(b_0) = 3 > 0 \qquad f(c_0) = 0.625 > 0$$

a_0	c_0	b_0
$-$	$+$	$+$

$$\rightsquigarrow \quad a_1 = a_0, \quad b_1 = c_0 = 0.5, \quad c_1 = \frac{1}{2}(a_1 + b_1) = \frac{1}{2}(0 + 0.5) = 0.25$$

$$f(c_1) = -0.234375 < 0$$

a_1	c_1	b_1
$-$	$-$	$+$

$$\rightsquigarrow \quad a_2 = c_1, \quad b_2 = b_1, \quad c_2 = \frac{1}{2}(a_2 + b_2) = \frac{1}{2}(0.25 + 0.5) = 0.375$$

a_2	c_2	b_2
$-$	$+$	$+$

$$f(c_2) = 0.177\ldots > 0$$

$$\rightsquigarrow \quad a_3 = a_2, \quad b_3 = c_2, \quad c_3 = \frac{1}{2}(a_3 + b_3) = \frac{1}{2}(0.25 + 0.375) = 0.3125$$

a_3	c_3	b_3
$-$	$-$	$+$

$$f(c_3) = -0.0319\ldots < 0$$

$$\rightsquigarrow \quad a_4 = c_3, \quad b_4 = b_3, \quad c_4 = \frac{1}{2}(a_4 + b_4) = \frac{1}{2}(0.3125 + 0.375) = 0.34375.$$

Therefore, the real root is about 0.34375.

To get an error at most $\varepsilon = 10^{-6}$, we need

$$n \geq \log_2\left(\frac{b_0 - a_0}{\varepsilon}\right) - 1 = \log_2\left(\frac{1-0}{10^{-6}}\right) - 1 \approx 18.93\ldots$$

Thus, $n \geq 19$.

Problem 4. (10 points) Let f be the function given in Problem 3.
(a) Write the iteration formula of Newton's method.
(b) With $x_0 = 1$, use Newton's method to find the approximate root after 4 iterations.

$$f(x) = x^3 + 3x - 1$$

$$f'(x) = 3x^2 + 3$$

Iteration formula: $$x_{n+1} = x_n - \frac{x_n^3 + 3x_n - 1}{3x_n^2 + 3}$$

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

n	x_n
0	1
1	$3/6 = 0.5$
2	$0.333\ldots$ $(=1/3)$
3	$29/90$
4	$0.32218\ldots$

Problem 5. (10 points) Consider the function $g(x) = 3x - x^2$.

(a) Solve for all fixed points of g.

(b) Write the iteration formula for the fixed point method for function g.

(c) With $x_0 = 1.8$, find x_5.

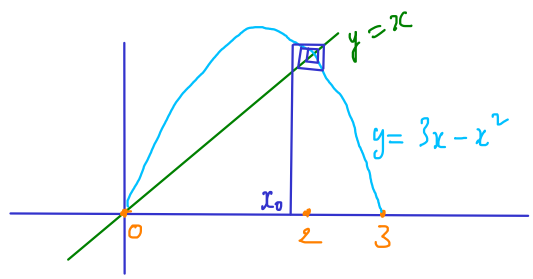
Set $x = g(x)$. Then $x = 3x - x^2$, which is equiv. to $2x - x^2 = 0$.

This equation has two solutions : $x = 0$ and $x = 2$. These are two fixed points of g.

Iteration formula: $x_{n+1} = g(x_n) = 3x_n - x_n^2$.

n	x_n
0	1.8
1	2.16
2	1.8144
3	2.15115...
4	1.82600...
5	2.14372...

Problem 6. (5 points) Draw a cobweb diagram that illustrates the fixed point method in Problem 5, Part (c).



Problem 7. (5 points) Given that the sequence x_n in Problem 5 converges to 2 as $n \to \infty$, find the order of convergence.

$$x_{n+1} = 3x_n - x_n^2$$

Subtract 2 from both sides:

$$x_{n+1} - 2 = 3x_n - x_n^2 - 2 = -(x_n - 2)(x_n - 1)$$

$$\approx -(x_n - 2)(2 - 1)$$

$$= -(x_n - 2)$$

Thus, $|x_{n+1} - 2| \lesssim |x_n - 2|$.

The order of convergence is therefore $p = 1$.