CSC349a Assignment #2

Parm Johal
V00787710

1.) | S_m | S_e | e_1 | e_0 | q_1 | q_2 | q_3 | q_4 |

a) $-30_{10} = -(1\times4^2 + 3\times4^1 + 2\times4^0) = -132_4 = -0.1320\times4^3$ (16, 12)

∴ The computer representation is:

1	0	0	3	1	3	2	0

b) 7.1875 → 0.1875_{10}

$4\times0.1875 = 0.75$
$4\times0.75 = 3.0$
$= .03_4$

$7_{10} = 13_4$

∴ $7.1875 = 13.03_4 = 0.1303_4\times4^2$

∴ The computer representation is:

0	0	0	2	1	3	0	3

c) The smallest positive non-zero number will contain the following:

- the value of S_m is 0 to indicate a positive sign
- the vallue of S_e is 1 to indicate a negative exponent
- the magnitude of the exponent will be 33_4 (to indicate the highest base 4 exponent).
- the magnitude of the mantissa will be 1000

∴ The smallest positive non-zero number represented in this system is:

0	1	3	3	1	0	0	0

and it's value in decimal is calculated as follows:

$0.1000_4 \times 4^{-33_4} = \frac{1}{4}\times4^{-15}$

$= \frac{1}{4^{16}} \approx 2.328\times10^{-10}$

CSC349a Assignment #2

Parm Johal
V00787710

1 (continued)

d)

16_{10} can be represented as 100_4 in base 4, which can be represented in floating point as: $0.1000_4 \times 4^{3_4}$. In the system, this is shown as

0	0	0	3	1	0	0	0	0

The next consecutive number in this system can be shown as:

0	0	0	3	1	0	0	0	1

Which is represented as

$$0.1001_4 \times 4^{3_4} = 100.1_4 = (1\times4^2)+(0\times4^1)+(0\times4^0)+(1\times4^{-1})$$
$$= 16.25_{10}$$

in base 10

∴ The size of the gap between any 2 consecutive numbers in this system is 0.25_{10}.

Parm Johal
V00787710

CSC 349a Assignment #2

2) $f(x,y) = -x - \sqrt{x^2 - y}$

a) $x = 0.1234 \times 10^3$, $y = -0.1234 \times 10^1$

$fl(f(x,y)) = fl(-x - fl(\sqrt{fl(fl(x^2) - y)}))$

$fl(x^2) = fl(123.4^2) = 15220 = 0.1522 \times 10^5$

$fl(x^2 - y) = fl(15520 - (-1.234)) = 15520 = 0.1522 \times 10^5$

$fl(\sqrt{x^2 - y}) = 124.5 = 0.1245 \times 10^3$

$fl(-x - \sqrt{x^2 - y}) = -(123.4) - (124.5) = -247.9 = -0.2479 \times 10^3$

true Value $\cong -246.8049999....$

$$|\varepsilon_t| = \left|\frac{p - p^*}{p}\right| = \left|\frac{-246.8049999 - (-247.9)}{-246.8049999}\right| \cong 0.0044367 \approx 0.44\%$$

b) $x = -123.4 = -0.1234 \times 10^3$, $y = 1.234 = 0.1234 \times 10^1$

$fl(x^2) = 15220 = 0.1522 \times 10^5$

$fl(x^2 - y) = 15210 = 0.1521 \times 10^5$

$fl(\sqrt{x^2 - y}) = 123.3 = 0.1233 \times 10^3$

$fl(-x - \sqrt{x^2 - y}) = 0.1$; true Value $\cong 0.005000101$

$$|\varepsilon_t| = \left|\frac{p - p^*}{p}\right| = \left|\frac{0.005000101 - 0.1}{0.005000101}\right| \cong 18.9996 \approx 1900\%$$

Parm Johal
V00787710

CSC 349a Assignment #2

2 (cont'd)

c) The range of of values for x will be large negative values and small positive and negative values for y. These range of values for both x and y will always give inaccurate results for the computation of $fl(f(x,y))$.

CSC 349a Assignment #2

Parm Johal
V00787710

3) Note: Taylor's theorem is shown as

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!} + \dots + \frac{f^{(n)}(a)(x-a)^n}{n!} + R_n$$

a) $f(x) = \sqrt{x+3}$; $a = 1$, $n = 2$

$$f'''(x) = \frac{3}{8(x+3)^{5/2}}$$

$$f(a) = f(1) = \sqrt{1+3} = \sqrt{4} = 2$$

$$f'(x) = \frac{1}{2(x+3)^{1/2}} \rightarrow f'(1) = \frac{1}{4}$$

$$f''(x) = \frac{-1}{4(x+3)^{3/2}} \rightarrow f''(1) = \frac{-1}{4 \cdot 8} = \frac{-1}{32}$$

$$R_n = \frac{f^{(n+1)}(\xi)(x-a)^{n+1}}{(n+1)!} \Rightarrow R_2 = \frac{f'''(\xi)}{3!}(x-1)^3$$

$$= \frac{1}{16(\xi+3)^{5/2}} \cdot (x-1)^3$$

$$\therefore f(x) = \sqrt{x+3} \approx 2 + \frac{1}{4}(x-1) - \frac{1}{64}(x-1)^2 + \frac{1}{16(\xi+3)^{5/2}} \cdot (x-1)^3$$

b) $\sqrt{4.14} = 2.0346989949\dots$ (exact Value) $= p$

$$f(1.14) = 2 + \frac{1}{4}(0.14) - \frac{1}{64}(0.14)^2 = 2.03469 = p^*$$

$$|E_p| = |p - p^*| = 0.000008994 = 8.994 \times 10^{-6}$$

c) A good upper bound for the truncation error will be when the $(x-1)^3$ term is maximized and the $(\xi+3)^{5/2}$ is minimized. This can be achieved by having $x = 1.2$ and $\xi = 1$. This gives us

$$R_2 = \frac{1}{16(1+3)^{5/2}} \cdot (1.2-1)^3 \cong 1.5625 \times 10^{-5} > |E_p| \; (= 8.994 \times 10^{-6})$$