EE 3755, Spring 04
Test 1 Solutions

Problem 1:

(a) Let X be $X = x_{n-1}x_{n-2}\cdots x_1x_0$ (1).

The additive inverse of X is

$-X$ = 2's compl. of $X = (\overline{x_{n-1}}\,\overline{x_{n-2}}\cdots\overline{x_1}\,\overline{x_0}) + 1$ (2)

(1), (2) $\Rightarrow$

$$
\begin{array}{rl}
X+(-X) = & x_{n-1}\,x_{n-2}\cdots x_1\,x_0 \\
 & \overline{x_{n-1}}\,\overline{x_{n-2}}\cdots\overline{x_1}\,\overline{x_0} \\
+) & \qquad\qquad\quad 1 \\
\hline
 & 1\;1\cdots 1\;1 \\
+) & \qquad\qquad\quad 1 \\
\hline
 & (1\,0\,0\cdots 0\,0)_2 = 2^n
\end{array}
$$

or $\boxed{X+(-X) = 2^n}$ (3)

On the other hand $\boxed{X+(-X) = 0}$ (4).

Eqs (3), (4) $\Rightarrow$ $\boxed{2^n = 0}$ (5).

But the carry out has a weight factor of 2^n which is equivalent to zero (see (5)).
Thus the carry out must be ignored.

(b). For addition 2 to produce a carry out it must be $z_{n-1} = z_{n-2} = \cdots = z_1 = z_0 = c = 1$. But this is impossible since the worst case addition 1 gives $\longrightarrow$ next page $\longrightarrow$

3755, Test 1 (2)
Solutions, Spr. 04

$$(X+Y)_{max} = \begin{array}{r} 111\cdots111 \\ +)\ 111\cdots111 \\ \hline 1111\cdots110 \end{array}$$

(c)

00000	1001	0

Initialization

11010 → 1×multiplicand / double shift

11010	1001	0

11110	1010	0

01100 → −2×multiplicand / double shift

01010	1010	0

00010	1010	1

→ product $= (00101010)_2 = (+42)_{10}$

(d) $A = A_1 \times 2^n + A_0$ (1)

$A = B \times Q + R$ (2)

(1), (2) $\Rightarrow$ $\boxed{A_1 \times 2^n + A_0 = B \times Q + R}$ (3)

Since $A_1 < B$, the maximum of left side of (3) is $A_{1\,max} \times 2^n + A_{0\,max} = (B-1)\times 2^n + 2^n - 1 =$
$= B\times 2^n - 2^n + 2^n - 1 = B\times 2^n - 1$ or

$\boxed{\text{max. left side of (3)} = B\times 2^n - 1}$ (4).

An n-bit quotient Q can achieve the same maximum of (4) for the rightside of (3). Just see that if Q is n-bit long, then $(B\times Q+R)_{max}$
$= B\times Q_{max} + R_{max} = B\times(2^n-1) + B - 1 = B\times 2^n - 1$.

Therefore, the length of n bits is sufficient for the quotient and overflow does not occur.

(e).

C	R | Q	
X	001 | 110	Initialization

1st cycle:

C	R | Q	
	011 | 10	shift left
	101	subtract B
1	000 | 101	

2nd cycle:

C	R | Q	
	001 | 01	shift left
	101	subtract B
0	110 | 010	

3rd cycle:

C	R | Q	
	100 | 10	shift left
	011	add B
0	111 | 100	

C	R | Q	
	011	restore
	010 | 100	

$R = (010)_2 = 2$; $Q = (100)_2 = 4$

(f). (i) → True ; (ii) → False ; (iii) → False.

(g) (i) $Q < 0$; (ii) $R > 0$.

Problem 2:

(a). The range of the fraction is $\boxed{0.5 \le f \le 1-2^{-17}}$

The range of the exponent is $-2^7 \le e \le 2^7-1$ or $\boxed{-128 \le e \le 127}$. Thus the positive floating point dynamic range is:

$$\boxed{0.5 \times 2^{-128} \le A^+ \le (1-2^{-17}) \times 2^{127}}$$

The negative floating point dynamic range is:

$$\boxed{-(1-2^{-17}) \times 2^{127} \le A^- \le -0.5 \times 2^{-128}}$$

(b). Align/Adjust

$e_1 - e_2 = e_1 + 2\text{'s compl. of } e_2 =$

$$\begin{array}{r} 1001 \\ +)\ 0110 \\ \hline 0\,1111 \end{array}$$

$\hookrightarrow c=0$ means $e_1 - e_2 < 0$

or $e_2 > e_1$. Then $e_2 - e_1 = 2\text{'s compl. of } (1111) =$
$= (0001)_2 = (1)_{10}$. Thus N_1:

s_1	e_2	f_1'
0	1010	01010

Subtract fractions: Here a true subtraction takes place. $f_1' - f_2 = f_1' + 2\text{'s compl. of } f_2 =$

$$\begin{array}{r} = 01010 \\ +)\ 01111 \\ \hline 0\,11001 \end{array}$$

$\hookrightarrow c=0$ means negative result. Thus $\Rightarrow$

EE 3755, Test 1 Solutions, Spr. 04 (5)

$N_3 = N_1 - N_2$ will be a negative number with fraction = 2's compl. of $(11001) = (00111)$

or N_3: $\boxed{1 \,|\, 1010 \,|\, 00111}$

Postnormalizing we get N_3: $\begin{array}{c|c|c} s_3 & e_3 & f_3 \\ \hline 1 & 1000 & 11100 \end{array}$

Exponent underflow did not occur since

$e_3 = (1000)_2 = (8)_{10} \in [0 \quad 15]$.

(c). Consider the integers A and B where

$A = a_{n-1} a_{n-2} \ldots a_1 a_0 \underbrace{000 \ldots 000}_{n \text{ zeroes}}.$

$B = b_{n-1} b_{n-2} \ldots b_1 b_0.$

Divide A by B to get an n-bit quotient Q and an n-bit remainder R. Let Q' be

$$Q' = \begin{cases} Q \text{ if } \frac{R}{B} < 1/2 \\ Q+1 \text{ if } \frac{R}{B} \geq 1/2 \end{cases}$$

Let Q' be represented in binary as

$Q' = q'_{n-1} q'_{n-2} \ldots q'_1 q'_0$. Then

$$f_3 = \frac{f_1}{f_2} \cong .q'_{n-1} q'_{n-2} \ldots q'_1 q'_0$$

(d). The dynamic range of 4-bit biased exponents is $DR=[0\ 15]$. Here $f_1 > f_2$ and thus alignment of dividend is needed. Thus, the exponent of the quotient is e_1+1-e_2+ $+\text{bias} = 5+1-14+8 = 0 \in [0\ 15]$. Therefore, neither exponent overflow nor exponent underflow occurs.

(e) The dynamic range of 4-bit biased exponents is $DR=[0\ 15]$. Here $f_1 \times f_2 = (.10000)_2 \times (.10000)_2 =$ $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 2^{-2} = (.01000)_2$. Postnormalization will be needed and the exponent of the product will be $e_1 + e_2 - \text{bias} - 1 =$ $= 11 + 13 - 8 - 1 = 15 \in [0\ 15]$. Therefore, neither exponent overflow nor exponent underflow occurs.

(f). Here $e_{biased} = (1011)_2 = 11$; $\text{bias} = 2^3 = 8$. Therefore $e_{unbiased} = e_{biased} - \text{bias} = 11 - 8 =$ $= 3$. Thus, $A_{value} = -(.111100)_2 \times 2^3 =$ $= -(111.100)_2 = -(7.5)_{10}$.