

CSCI 341: Computer Organization  
WS 2: Floating Point Numbers

1	<p>Given a 32 bit register, show where each part of a floating point number will be. Fill out the table below.</p> <table><tr><td>3</td><td>3</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table> <p>Solution:</p> <table><tr><td>3</td><td>3</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>S</td><td>E</td><td>E</td><td>E</td><td>E</td><td>E</td><td>E</td><td>E</td><td>E</td><td>E</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td><td>F</td></tr></table>	3	3	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1	0	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0																																											3	3	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1	0	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0											S	E	E	E	E	E	E	E	E	E	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
3	3	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1	0																																																																																																																																																																		
1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0																																																																																																																																																																												
3	3	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1	0																																																																																																																																																																		
1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0																																																																																																																																																																												
S	E	E	E	E	E	E	E	E	E	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F																																																																																																																																																																		
2	<p>Write down the formula for converting a binary floating point number to decimal.</p> <p>Solution: <math>(-1)^S(1 + F)(2^{E-b})</math></p>																																																																																																																																																																																																
3	<p>Write down the formula for converting a denormal binary floating point number to decimal.</p> <p>Solution: <math>(-1)^S(0 + F)(2^{1-b})</math></p>																																																																																																																																																																																																
4	<p>Convert the following floating point numbers to decimal</p> <p>1. 1001 0010 0100 1000 0000 0000 0000 0000</p> <p>2. 0111 1011 0100 1000 0000 0000 0000 0000</p> <p>Solution:</p> <p style="text-align: center;"><math>1\ 00100100\ 100100000000000000000000</math> <math>(-1)^1(1 + 2^{-1} + 2^{-4})(2^{36-127}) = -6.31089 * 10^{-28}</math> <math>0\ 11110110\ 100100000000000000000000</math> <math>(-1)^0(1 + 2^{-1} + 2^{-4})(2^{246-127}) = 1.03846 * 10^{36}</math></p>																																																																																																																																																																																																
5	<p>Convert the following floating point numbers to decimal. Show enough decimal places for the numbers to be distinguishable.</p> <p>1. 1000 0000 0100 1000 0000 0000 0000 0000</p> <p>2. 0000 0000 0100 1000 0000 0000 0000 0001</p> <p>Solution:</p>																																																																																																																																																																																																

1 00000000 100100000000000000000000

$$(-1)^1(0 + 2^{-1} + 2^{-4})(2^{1-127}) = -6.61215572 * 10^{-39}$$

0 00000000 1001000000000000000000001

$$(-1)^0(0 + 2^{-1} + 2^{-4} + 2^{-23})(2^{1-127}) = 6.61215712 * 10^{-39}$$