Floating Point Representation

Convert the fixed point binary numbers below into floating point representation that uses 16 bits for the mantissa (including the signed bit) and 8 bits two's complement for the exponent.

Example 1

Fixed Point	Signed bit (1 bit	Mantissa (15 bits)	Exponent (two's complement)	Notes
1000101.101	0	100010110100000	00000111	• Signed bit is 0 because the number is positive (no negative sign) • The decimal point is moved 7 places to the left so that it is in front of the first 1 therefore it is positive 7
Final answer	0100010110100000000000111			

Example 2

Fixed Point	Signed bit	Mantissa	Exponent	Notes
-0.00010111	1	101110000000000	11111101	• Signed bit is 1 because the number is negative • The decimal point has moved 3 place to the right therefore the exponent is -3 (convert to 8 bit two's complement)
Final answer	11011100000000000111111101			

Exercise 1

Use the method above to convert the following numbers to floating point representation that uses 16 bits for the mantissa including the signed bit and 8 bit two's complement for the exponent.

	Fixed Point	Signed bit	Mantissa	Exponent	Final Answer
1.	**1110101.1011**				
2.	**-10110.101**				
3.	**0.00101011**				
4.	**-0.0001101**				
5.	**11011.01101**				
6.	**-0.0101001**				
7.	**-0.00000111**				
8.	**1101101.1**				
9.	**011101.101**				
10.	**-1101.010101**				

Convert the floating point numbers below into fixed point binary. The floating point numbers use 16 bits (including this signed bit) to store the mantissa and 8 bit two's complement to store the mantissa.

Example 3

Floating Point Number Signed bit (1 bit	Signed bit	Mantissa (15 bits)	Exponent (two's complement)	Final answer
0100010110100000000000111 0 100010110100000 00000111 (rewrite to separate signed bit, mantissa and exponent)	0 +ve number	010001011010000000000111 Remove extra 0's 1000101101	00000111 Convert to decimal = 7 Decimal point moves 7 steps from first 1	1000101.101

Example 4

Floating Point Number Signed bit (1 bit	Signed bit	Mantissa (15 bits)	Exponent (two's complement)	Final answer
1101110000000000011110101 1 101110000000000 11110101	1 -ve number	101110000000000 10111	11111101 Convert to decimal = -3	-0.00010111

Exercise 2

Use the method above to convert the following numbers to floating point representation that uses 16 bits for the mantissa including the signed bit and 8 bit two's complement for the exponent.

	Floating Point	Signed bit	Mantissa	Exponent	Final Answer
1.	**10111010000000000001001**				
2.	**01100100100000011111010**				
3.	**11101001010000011111011**				
4.	**11101100000000011111100**				
5.	**01101010010000000000111**				

Exercise 3

1. Describe the effect if the number of bits used to store the mantissa in increased from 16 bits to 24 bits.

2. Describe the effect if the number of bits used to store the exponent is reduced to 4.

3. For each floating number below, state whether the number is positive or negative. Explain your answer. The floating point representation uses 16 bits to store the mantissa (including the signed bit) and 8 bits to store the exponent.

(a) 1101110010000000011110101

(b) 0100010110100000000000111