Converting binary floating point numbers to denary and vice versa

Floating point binary numbers are expressed in two parts, a mantissa and an exponent. Before performing any conversion you need to know the format of the binary number, for example a 12 bit number may use 8 bits for the mantissa and 4 bits for the exponent. Both mantissa and exponent are expressed in 2s complement format.

Floating Point Binary to Denary

1. Identify the sign of the exponent

2. Calculate the denary value of the exponent, converting from 2s complement as necessary.

3. Identify the sign of the mantissa.

4. Write down the mantissa with a binary point immediately after the first digit.

5. Move the binary point in the mantissa by the number of places specified by the value of the exponent.
   If the exponent is +ve move the point to the right.
   If the exponent is -ve move the binary point to the left

6. Calculate the denary value of the mantissa, converting from 2s complement as necessary.

Examples

a) Convert 01101000 0100 to denary assuming that the mantissa is 8 bits and the exponent is 4 bits, both in 2s complement format.

1. Exponent 0100 is positive
2. The denary value of the exponent is +4.
3. Mantissa 01101000 is positive
4. Mantissa with the binary point 0**.**1101000
5. Mantissa with the binary point moved 4 places right (because its positive) 01101.000
6. mantissa converted to denary 1+4+8 = +13

   So the denary value of 01101000 0100 is 13.

b) Convert 01011010 0100 to denary assuming that the mantissa is 8 bits and the exponent is 4 bits, both in 2s complement format.

1. Exponent 0100 is positive
2. The denary value of the exponent is +4.
3. Mantissa 01011010 is positive
4. Mantissa with the binary point 0**.**1011010
5. Mantissa with the binary point moved 4 places right (because its positive) 01011.010
6. Mantissa converted to denary 1+2+8 + 1/4 = +11.25

   So the denary value of 01011010 0100 is +11.25

Denary to Floating Point Binary

1. Convert the denary value to 2s complement binary. It may help to split the denary value into whole and fractional parts. Write down the resulting binary value **including a binary point**. A positive number should start with “0” and a negative number with “1”

2. For a positive number move the binary point until it is in front of the leftmost “1”
   For a negative number move the binary point until it is in front of the leftmost “0”
   If necessary add trailing zeroes to make up the full width of the mantissa or discard digits that are shifted outside the width of the mantissa.
   The resulting binary value is the mantissa. For a positive number it should begin with “01”, for a negative number it should begin with “10”

3. The denary value of the exponent is equal to the number of places the binary point has been moved. The value is positive if the point has been moved to the left and negative if the point has been moved to the right.

4. Calculate the 2s complement value of the exponent.If necessary add leading zeroes to make up the full width of the exponent.

5. The final answer is the combined binary values of the mantissa and exponent.

Examples

a) Convert +13 to floating point binary, assuming that the mantissa is 8 bits and the exponent is 4 bits, both in 2s complement format.

1. +13 = 8+4+1 01101**.**000 (adding 3 trailing zeroes)
2. shift left 4 places 0**.**1101000
3. exponent is +4
4. 0100
5. Final answer 01101000 0100

b) Convert +11.25 to floating point binary, assuming that the mantissa is 8 bits and the exponent is 4 bits, both in 2s complement format.

1. +11.25 = 8+2+1+1/4 01011**.**010 (adding 1 trailing zero)
2. shift left 4 places 0**.**1011010
3. exponent is +4
4. 0100
5. Final answer 01011010 0100