Numbers with a Fractional Part

L.O.: To understand how numbers with a fractional part are represented in binary.

Numbers with fractional parts can be represented as:

* Fixed Point numbers
* Floating Point numbers

# Fixed Point

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 23 | 22 | 21 | 20 | 2-1 | 2-2 | 2-3 | 2-4 |
| 8 | 4 | 2 | 1 | 1/2 | 1/4 | 1/8 | 1/16 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |

= 4 + 2 + 1/4 + 1/8­ = 6.375

* Moving the binary point to the left allows for more accurate decimals, but reduces the range of numbers available
* Regardless of binary point position, an 8-bit number can represent 256 different numbers.

1. *What’s the largest number that could be represented using 8-bit fixed point binary with the binary point for places to the right of the least significant bit?*

255 / 24 = 15.9375

1. *What’s the smallest number above 0 that could be represented using 8-bit fixed point binary with the binary point in the same place?*

0.0625

1. *What is the largest number that can be represented if the binary point is two places to the left of the least significant bit?*

255 / 22 = 63.75

1. *What is the smallest number above 0 that can be represented if the binary point is two places to the left of the least significant bit?*

0.25

1. *Where would you fix the binary point in order to achieve the greatest accuracy of the fractional part?*

To the right of the least significant bit

1. *What would be the advantage, and disadvantage of fixing the binary point 7 places to the left of the least significant bit?*

It would allow the greatest accuracy of the fractional part, but the number represented cannot be 2 or greater

1. *With the binary point at 4 places to the left of the least significant bit, convert the following numbers to denary:*
2. 1001.0100 = 8 + 1 + 0.5 = **9.5**
3. 1100.1111 = 8 + 4 + 0.5 + 0.25 + 0.125 + 0.0625 = **12.9375**
4. 0100.1101 = 4 + 0.5 + 0.25 + 0.0625 = **4.8125**
5. *With the binary point at 4 places to the left of the least significant bit, convert the following numbers to binary:*
6. 2.5 = 2 + 0.5 = **0010.1000**
7. 5.825 can’t be done (5.825 x 24 = 93.2, not whole) Closest = 4 + 1 + 0.5 + 0.25 + 0.0625 (remainder 0.0125) = **0101.1101 (remainder 0.012510­)**
8. 1.5625 = 1 + 0.5 + 0.0625 = **0001.1001**

# Negative Fixed Point

* Two’s Compliment can be applied to fixed point numbers

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **2-3** | 22 | 21 | 20 | 2-1 | 2-2 | 2-3 | 2-4 |
| **-8** | 4 | 2 | 1 | 1/2 | 1/4 | 1/8 | 1/16 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |

1. *Using the negative fixed point representation above, convert the following numbers to denary:*
2. 0111.0100 = 4 + 2 + 1 + 0.25 = 7.25
3. 0001.1100 = 1 + 0.5 + 0.25 = 1.75
4. 1001.1100
5. 1010.0011 = 8 + 2 + 0.125 + 0.0625 = 10.1875