Lesson 6, part 2) A Complete study of quarter-precision IEEE format

Recall in this lesson, we began our study of IEEE floating-point data formats. Out of the five formats we will study in this class including

- binary 8
- binary 16
- binary 32
- binary 64
- binary 128

the first one binary 8 is by far the easiest to analyze by hand. In this part of our lesson, we consider all possible 8-bit binary words and their corresponding interpreted values using the binary 8 encoding scheme.

Remember that in binary8, we encode a floating-point number using 8-bits of information:

sign	exponent field			significand field			
B_7	B_6	B_5	B_4	B_3	B_2	B_1	B_0
bit # 7	6	5	4	3	2	1	0

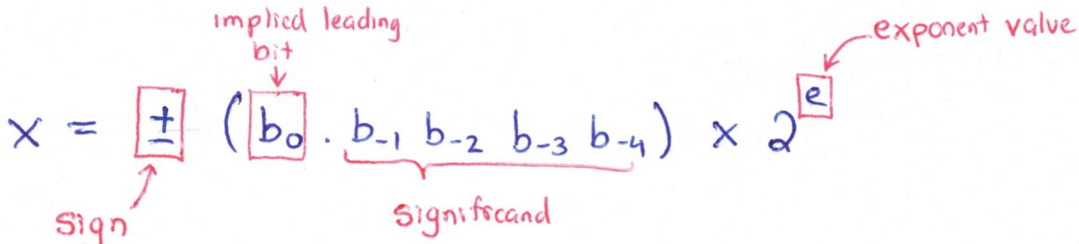
In this encoding, we have two centrally important concepts

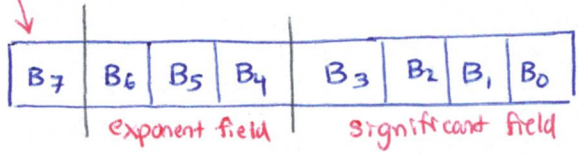
- $B = B_7 B_6 B_5 B_4 B_3 B_2 B_1 B_0$ ← we will call this 8-bit binary word the raw, uninterpreted 8-bit word. This represents a string of 8 binary digits each of whose values $B_K \in \{0,1\}$ for $K \in \{0,1,\ldots,7\}$.

- $x = \text{binary8}(B) \in \mathbb{Q}_I$ OR $x = \text{binary8}(B_7 B_6 B_5 B_4 B_3 B_2 B_1 B_0) \in \mathbb{Q}_I$ ← we call $x \in \mathbb{Q}_I$ the interpreted value of B resulting from the binary8 encoding scheme.

Recall the foundational idea behind floating-point representation is to use the intuition behind scientific notation to enable a limited number of bits to encode an expanded range of $x \in \mathbb{Q}_I$ by allowing the radix point to float.

To this end, we set



where if $B =$ 

, then

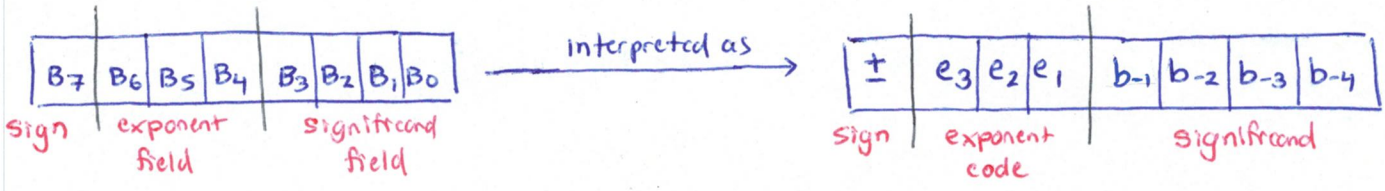
- sign bit $B_7 = \begin{cases} 0 & \text{iff } x > 0 \\ 1 & \text{iff } x < 0 \end{cases}$

- excess-K biased exponent value $e = u - K$ where $K = 3$

$$e = \text{uint3}(B_6 B_5 B_4) - K$$

$$\Rightarrow \text{uint3}(B_6 B_5 B_4) = e + K = e + 3 = e + 011$$

Let's suppose we interpret our raw bits $B = B_7\,B_6\,B_5\,B_4\,B_3\,B_2\,B_1\,B_0$ as

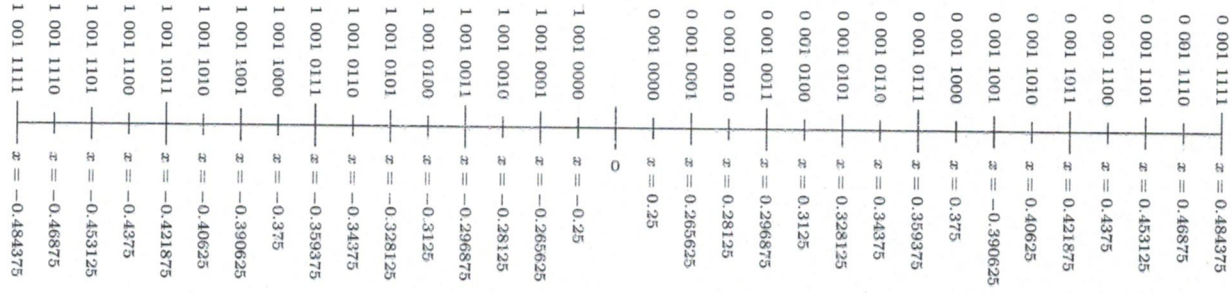


	Exponent code: if Raw, uninterpreted 3-bit exponent word $e_3\,e_2\,e_1$ is	exponent value: then the excess-K value of $e = U - K$ in decimal is	and the interpreted numerical value of $x = \text{binary8}(B)$ is
SUBNORMAL NUMBERS → 0	000	special case of subnormals	$x = \pm 0.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4}$
NORMALIZED NUMBERS → 1	001 (emin)	$e = \text{uint3}(001) - K = 1 - 3 = \boxed{-2}$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{\boxed{-2}}$
2	010	$e = \text{uint3}(010) - K = 2 - 3 = -1$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{-1}$
3	011	$e = \text{uint3}(011) - K = 3 - 3 = \boxed{0}$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{\boxed{0}}$
4	100	$e = \text{uint3}(100) - K = 4 - 3 = 1$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{1}$
5	101	$e = \text{uint3}(101) - K = 5 - 3 = \boxed{2}$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{\boxed{2}}$
6	110 (emax)	$e = \text{uint3}(110) - K = 6 - 3 = 3$	$x = \pm 1.b_{-1}\,b_{-2}\,b_{-3}\,b_{-4} \times 2^{3}$
7 →	111	special case of $\pm\infty$ or NaN	$x = \pm\infty$ if $b_{-1} = \cdots = b_{-4} = 0$ or $x = \text{NaN}$ otherwise

realmin = smallest normalized floating-point number

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1} \ldots b_{-f}$		
1 001 1111	1	001	$x = -1.1111 \times 2^{-2}$	$x = -0.011111$	$x = -2^{-2} - 2^{-3} - 2^{-4} - 2^{-5} - 2^{-6}$	$x = -0.484375$
1 001 1110	1	001	$x = -1.1110 \times 2^{-2}$	$x = -0.011110$	$x = -2^{-2} - 2^{-3} - 2^{-4} - 2^{-5}$	$x = -0.46875$
1 001 1101	1	001	$x = -1.1101 \times 2^{-2}$	$x = -0.011101$	$x = -2^{-2} - 2^{-3} - 2^{-4} - 2^{-6}$	$x = -0.453125$
1 001 1100	1	001	$x = -1.1100 \times 2^{-2}$	$x = -0.011100$	$x = -2^{-2} - 2^{-3} - 2^{-4}$	$x = -0.4375$
1 001 1011	1	001	$x = -1.1011 \times 2^{-2}$	$x = -0.011011$	$x = -2^{-2} - 2^{-3} - 2^{-5} - 2^{-6}$	$x = -0.421875$
1 001 1010	1	001	$x = -1.1010 \times 2^{-2}$	$x = -0.011010$	$x = -2^{-2} - 2^{-3} - 2^{-5}$	$x = -0.40625$
1 001 1001	1	001	$x = -1.1001 \times 2^{-2}$	$x = -0.011001$	$x = -2^{-2} - 2^{-3} - 2^{-6}$	$x = -0.390625$
1 001 1000	1	001	$x = -1.1000 \times 2^{-2}$	$x = -0.011000$	$x = -2^{-2} - 2^{-3}$	$x = -0.375$
1 001 0111	1	001	$x = -1.0111 \times 2^{-2}$	$x = -0.010111$	$x = -2^{-2} - 2^{-4} - 2^{-5} - 2^{-6}$	$x = -0.359375$
1 001 0110	1	001	$x = -1.0110 \times 2^{-2}$	$x = -0.010110$	$x = -2^{-2} - 2^{-4} - 2^{-6}$	$x = -0.34375$
1 001 0101	1	001	$x = -1.0101 \times 2^{-2}$	$x = -0.010101$	$x = -2^{-2} - 2^{-4} - 2^{-6}$	$x = -0.328125$
1 001 0100	1	001	$x = -1.0100 \times 2^{-2}$	$x = -0.010100$	$x = -2^{-2} - 2^{-4}$	$x = -0.3125$
1 001 0011	1	001	$x = -1.0011 \times 2^{-2}$	$x = -0.010011$	$x = -2^{-2} - 2^{-5} - 2^{-6}$	$x = -0.296875$
1 001 0010	1	001	$x = -1.0010 \times 2^{-2}$	$x = -0.010010$	$x = -2^{-2} - 2^{-5}$	$x = -0.28125$
1 001 0001	1	001	$x = -1.0001 \times 2^{-2}$	$x = -0.010001$	$x = -2^{-2} - 2^{-6}$	$x = -0.265625$
1 001 0000	1	001	$x = -1.0000 \times 2^{-2}$	$x = -0.010000$	$x = -2^{-2}$	$x = -0.25$
0 001 0000	0	001	$x = +1.0000 \times 2^{-2}$	$x = +0.010000$	$x = 2^{-2}$	$x = 0.25$
0 001 0001	0	001	$x = +1.0001 \times 2^{-2}$	$x = +0.010001$	$x = 2^{-2} + 2^{-6}$	$x = 0.265625$
0 001 0010	0	001	$x = +1.0010 \times 2^{-2}$	$x = +0.010010$	$x = 2^{-2} + 2^{-5}$	$x = 0.28125$
0 001 0011	0	001	$x = +1.0011 \times 2^{-2}$	$x = +0.010011$	$x = 2^{-2} + 2^{-5} + 2^{-6}$	$x = 0.296875$
0 001 0100	0	001	$x = +1.0100 \times 2^{-2}$	$x = +0.010100$	$x = 2^{-2} + 2^{-4}$	$x = 0.3125$
0 001 0101	0	001	$x = +1.0101 \times 2^{-2}$	$x = +0.010101$	$x = 2^{-2} + 2^{-4} + 2^{-6}$	$x = 0.328125$
0 001 0110	0	001	$x = +1.0110 \times 2^{-2}$	$x = +0.010110$	$x = 2^{-2} + 2^{-4} + 2^{-6}$	$x = 0.34375$
0 001 0111	0	001	$x = +1.0111 \times 2^{-2}$	$x = +0.010111$	$x = 2^{-2} + 2^{-4} + 2^{-5} + 2^{-6}$	$x = 0.359375$
0 001 1000	0	001	$x = +1.1000 \times 2^{-2}$	$x = +0.011000$	$x = 2^{-2} + 2^{-3}$	$x = 0.375$
0 001 1001	0	001	$x = +1.1001 \times 2^{-2}$	$x = +0.011001$	$x = 2^{-2} + 2^{-3} + 2^{-6}$	$x = 0.390625$
0 001 1010	0	001	$x = +1.1010 \times 2^{-2}$	$x = +0.011010$	$x = 2^{-2} + 2^{-3} + 2^{-5}$	$x = 0.40625$
0 001 1011	0	001	$x = +1.1011 \times 2^{-2}$	$x = +0.011011$	$x = 2^{-2} + 2^{-3} + 2^{-5} + 2^{-6}$	$x = 0.421875$
0 001 1100	0	001	$x = +1.1100 \times 2^{-2}$	$x = +0.011100$	$x = 2^{-2} + 2^{-3} + 2^{-4}$	$x = 0.4375$
0 001 1101	0	001	$x = +1.1101 \times 2^{-2}$	$x = +0.011101$	$x = 2^{-2} + 2^{-3} + 2^{-4} + 2^{-6}$	$x = 0.453125$
0 001 1110	0	001	$x = +1.1110 \times 2^{-2}$	$x = +0.011110$	$x = 2^{-2} + 2^{-3} + 2^{-4} + 2^{-6}$	$x = 0.46875$
0 001 1111	0	001	$x = +1.1111 \times 2^{-2}$	$x = +0.011111$	$x = 2^{-2} + 2^{-3} + 2^{-4} + 2^{-5} + 2^{-6}$	$x = 0.484375$

Let's graph this table on a number line to visualize the set of $x \in \mathbb{Q}_I$ that we can represent in `binary8` when we hold the exponent code at 001.



IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 010 1111	1	010	$x = -1.1111 \times 2^{-1}$	$x = -0.11111$	$x = -2^{-1} - 2^{-2} - 2^{-3} - 2^{-4} - 2^{-5}$	$x = -0.96875$
1 010 1110	1	010	$x = -1.1110 \times 2^{-1}$	$x = -0.11110$	$x = -2^{-1} - 2^{-2} - 2^{-3} - 2^{-4}$	$x = -0.9375$
1 010 1101	1	010	$x = -1.1101 \times 2^{-1}$	$x = -0.11101$	$x = -2^{-1} - 2^{-2} - 2^{-3} - 2^{-5}$	$x = -0.90625$
1 010 1100	1	010	$x = -1.1100 \times 2^{-1}$	$x = -0.11100$	$x = -2^{-1} - 2^{-2} - 2^{-3}$	$x = -0.875$
1 010 1011	1	010	$x = -1.1011 \times 2^{-1}$	$x = -0.11011$	$x = -2^{-1} - 2^{-2} - 2^{-4} - 2^{-5}$	$x = -0.84375$
1 010 1010	1	010	$x = -1.1010 \times 2^{-1}$	$x = -0.11010$	$x = -2^{-1} - 2^{-2} - 2^{-4}$	$x = -0.8125$
1 010 1001	1	010	$x = -1.1001 \times 2^{-1}$	$x = -0.11001$	$x = -2^{-1} - 2^{-2} - 2^{-5}$	$x = -0.78125$
1 010 1000	1	010	$x = -1.1000 \times 2^{-1}$	$x = -0.11000$	$x = -2^{-1} - 2^{-2}$	$x = -0.75$
1 010 0111	1	010	$x = -1.0111 \times 2^{-1}$	$x = -0.10111$	$x = -2^{-1} - 2^{-3} - 2^{-4} - 2^{-5}$	$x = -0.71875$
1 010 0110	1	010	$x = -1.0110 \times 2^{-1}$	$x = -0.10110$	$x = -2^{-1} - 2^{-3} - 2^{-4}$	$x = -0.6875$
1 010 0101	1	010	$x = -1.0101 \times 2^{-1}$	$x = -0.10101$	$x = -2^{-1} - 2^{-3} - 2^{-5}$	$x = -0.65625$
1 010 0100	1	010	$x = -1.0100 \times 2^{-1}$	$x = -0.10100$	$x = -2^{-1} - 2^{-3}$	$x = -0.625$
1 010 0011	1	010	$x = -1.0011 \times 2^{-1}$	$x = -0.10011$	$x = -2^{-1} - 2^{-4} - 2^{-5}$	$x = -0.59375$
1 010 0010	1	010	$x = -1.0010 \times 2^{-1}$	$x = -0.10010$	$x = -2^{-1} - 2^{-4}$	$x = -0.5625$
1 010 0001	1	010	$x = -1.0001 \times 2^{-1}$	$x = -0.10001$	$x = -2^{-1} - 2^{-5}$	$x = -0.53125$
1 010 0000	1	010	$x = -1.0000 \times 2^{-1}$	$x = -0.10000$	$x = -2^{-1}$	$x = -0.5$
0 010 0000	0	010	$x = +1.0000 \times 2^{-1}$	$x = +0.10000$	$x = 2^{-1}$	$x = 0.5$
0 010 0001	0	010	$x = +1.0001 \times 2^{-1}$	$x = +0.10001$	$x = 2^{-1} + 2^{-5}$	$x = 0.53125$
0 010 0010	0	010	$x = +1.0010 \times 2^{-1}$	$x = +0.10010$	$x = 2^{-1} + 2^{-4}$	$x = 0.5625$
0 010 0011	0	010	$x = +1.0011 \times 2^{-1}$	$x = +0.10011$	$x = 2^{-1} + 2^{-4} + 2^{-5}$	$x = 0.59375$
0 010 0100	0	010	$x = +1.0100 \times 2^{-1}$	$x = +0.10100$	$x = 2^{-1} + 2^{-3}$	$x = 0.625$
0 010 0101	0	010	$x = +1.0101 \times 2^{-1}$	$x = +0.10101$	$x = 2^{-1} + 2^{-3} + 2^{-5}$	$x = 0.65625$
0 010 0110	0	010	$x = +1.0110 \times 2^{-1}$	$x = +0.10110$	$x = 2^{-1} + 2^{-3} + 2^{-4}$	$x = 0.6875$
0 010 0111	0	010	$x = +1.0111 \times 2^{-1}$	$x = +0.10111$	$x = 2^{-1} + 2^{-3} + 2^{-4} + 2^{-5}$	$x = 0.71875$
0 010 1000	0	010	$x = +1.0000 \times 2^{-1}$	$x = +0.11000$	$x = 2^{-1} + 2^{-2}$	$x = 0.75$
0 010 1001	0	010	$x = +1.1001 \times 2^{-1}$	$x = +0.11001$	$x = 2^{-1} + 2^{-2} + 2^{-5}$	$x = 0.78125$
0 010 1010	0	010	$x = +1.1010 \times 2^{-1}$	$x = +0.11010$	$x = 2^{-1} + 2^{-2} + 2^{-4}$	$x = 0.8125$
0 010 1011	0	010	$x = +1.1011 \times 2^{-1}$	$x = +0.11011$	$x = 2^{-1} + 2^{-2} + 2^{-4} + 2^{-5}$	$x = 0.84375$
0 010 1100	0	010	$x = +1.1100 \times 2^{-1}$	$x = +0.11100$	$x = 2^{-1} + 2^{-2} + 2^{-3}$	$x = 0.875$
0 010 1101	0	010	$x = +1.1101 \times 2^{-1}$	$x = +0.11101$	$x = 2^{-1} + 2^{-2} + 2^{-3} + 2^{-5}$	$x = 0.90625$
0 010 1110	0	010	$x = +1.1110 \times 2^{-1}$	$x = +0.11110$	$x = 2^{-1} + 2^{-2} + 2^{-3} + 2^{-4}$	$x = 0.9375$
0 010 1111	0	010	$x = +1.1111 \times 2^{-1}$	$x = +0.11111$	$x = 2^{-1} + 2^{-2} + 2^{-3} + 2^{-4} + 2^{-5}$	$x = 0.96875$

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 011 1111	1	011	$x = -1.1111 \times 2^0$	$x = -1.1111$	$x = -2^0 - 2^{-1} - 2^{-2} - 2^{-3} - 2^{-4}$	$x = -1.9375$
1 011 1110	1	011	$x = -1.1110 \times 2^0$	$x = -1.1110$	$x = -2^0 - 2^{-1} - 2^{-2} - 2^{-3}$	$x = -1.875$
1 011 1101	1	011	$x = -1.1101 \times 2^0$	$x = -1.1101$	$x = -2^0 - 2^{-1} - 2^{-2} - 2^{-4}$	$x = -1.8125$
1 011 1100	1	011	$x = -1.1100 \times 2^0$	$x = -1.1100$	$x = -2^0 - 2^{-1} - 2^{-2}$	$x = -1.75$
1 011 1011	1	011	$x = -1.1011 \times 2^0$	$x = -1.1011$	$x = -2^0 - 2^{-1} - 2^{-3} - 2^{-4}$	$x = -1.6875$
1 011 1010	1	011	$x = -1.1010 \times 2^0$	$x = -1.1010$	$x = -2^0 - 2^{-1} - 2^{-3}$	$x = -1.625$
1 011 1001	1	011	$x = -1.1001 \times 2^0$	$x = -1.1001$	$x = -2^0 - 2^{-1} - 2^{-4}$	$x = -1.5625$
1 011 1000	1	011	$x = -1.1000 \times 2^0$	$x = -1.1000$	$x = -2^0 - 2^{-1}$	$x = -1.5$
1 011 0111	1	011	$x = -1.0111 \times 2^0$	$x = -1.0111$	$x = -2^0 - 2^{-2} - 2^{-3} - 2^{-4}$	$x = -1.4375$
1 011 0110	1	011	$x = -1.0110 \times 2^0$	$x = -1.0110$	$x = -2^0 - 2^{-2} - 2^{-3}$	$x = -1.375$
1 011 0101	1	011	$x = -1.0101 \times 2^0$	$x = -1.0101$	$x = -2^0 - 2^{-2} - 2^{-4}$	$x = -1.3125$
1 011 0100	1	011	$x = -1.0100 \times 2^0$	$x = -1.0100$	$x = -2^0 - 2^{-2}$	$x = -1.25$
1 011 0011	1	011	$x = -1.0011 \times 2^0$	$x = -1.0011$	$x = -2^0 - 2^{-3} - 2^{-4}$	$x = -1.1875$
1 011 0010	1	011	$x = -1.0010 \times 2^0$	$x = -1.0010$	$x = -2^0 - 2^{-3}$	$x = -1.125$
1 011 0001	1	011	$x = -1.0001 \times 2^0$	$x = -1.0001$	$x = -2^0 - 2^{-4}$	$x = -1.0625$
1 011 0000	1	011	$x = -1.0000 \times 2^0$	$x = -1.0000$	$x = -2^0$	$x = -1$
0 011 0000	0	011	$x = +1.0000 \times 2^0$	$x = +1.0000$	$x = 2^0$	$x = 1$
0 011 0001	0	011	$x = +1.0001 \times 2^0$	$x = +1.0001$	$x = 2^0 + 2^{-4}$	$x = 1.0625$
0 011 0010	0	011	$x = +1.0010 \times 2^0$	$x = +1.0010$	$x = 2^0 + 2^{-3}$	$x = 1.125$
0 011 0011	0	011	$x = +1.0011 \times 2^0$	$x = +1.0011$	$x = 2^0 + 2^{-3} + 2^{-4}$	$x = 1.1875$
0 011 0100	0	011	$x = +1.0100 \times 2^0$	$x = +1.0100$	$x = 2^0 + 2^{-2}$	$x = 1.25$
0 011 0101	0	011	$x = +1.0101 \times 2^0$	$x = +1.0101$	$x = 2^0 + 2^{-2} + 2^{-4}$	$x = 1.3125$
0 011 0110	0	011	$x = +1.0110 \times 2^0$	$x = +1.0110$	$x = 2^0 + 2^{-2} + 2^{-3}$	$x = 1.375$
0 011 0111	0	011	$x = +1.0111 \times 2^0$	$x = +1.0111$	$x = 2^0 + 2^{-2} + 2^{-3} + 2^{-4}$	$x = 1.4375$
0 011 1000	0	011	$x = +1.0000 \times 2^0$	$x = +1.1000$	$x = 2^0 + 2^{-1}$	$x = 1.5$
0 011 1001	0	011	$x = +1.1001 \times 2^0$	$x = +1.1001$	$x = 2^0 + 2^{-1} + 2^{-4}$	$x = 1.5625$
0 011 1010	0	011	$x = +1.1010 \times 2^0$	$x = +1.1010$	$x = 2^0 + 2^{-1} + 2^{-3}$	$x = 1.625$
0 011 1011	0	011	$x = +1.1011 \times 2^0$	$x = +1.1011$	$x = 2^0 + 2^{-1} + 2^{-3} + 2^{-4}$	$x = 1.6875$
0 011 1100	0	011	$x = +1.1100 \times 2^0$	$x = +1.1100$	$x = 2^0 + 2^{-1} + 2^{-2}$	$x = 1.75$
0 011 1101	0	011	$x = +1.1101 \times 2^0$	$x = +1.1101$	$x = 2^0 + 2^{-1} + 2^{-2} + 2^{-4}$	$x = 1.8125$
0 011 1110	0	011	$x = +1.1110 \times 2^0$	$x = +1.1110$	$x = 2^0 + 2^{-1} + 2^{-2} + 2^{-3}$	$x = 1.875$
0 011 1111	0	011	$x = +1.1111 \times 2^0$	$x = +1.1111$	$x = 2^0 + 2^{-1} + 2^{-2} + 2^{-3} + 2^{-4}$	$x = 1.9375$

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 100 1111	1	100	$x = -1.1111 \times 2^1$	$x = -11.111$	$x = -2^1 - 2^0 - 2^{-1} - 2^{-2} - 2^{-3}$	$x = -3.875$
1 100 1110	1	100	$x = -1.1110 \times 2^1$	$x = -11.110$	$x = -2^1 - 2^0 - 2^{-1} - 2^{-2}$	$x = -3.75$
1 100 1101	1	100	$x = -1.1101 \times 2^1$	$x = -11.101$	$x = -2^1 - 2^0 - 2^{-1} - 2^{-3}$	$x = -3.625$
1 100 1100	1	100	$x = -1.1100 \times 2^1$	$x = -11.100$	$x = -2^1 - 2^0 - 2^{-1}$	$x = -3.5$
1 100 1011	1	100	$x = -1.1011 \times 2^1$	$x = -11.011$	$x = -2^1 - 2^0 - 2^{-2} - 2^{-3}$	$x = -3.375$
1 100 1010	1	100	$x = -1.1010 \times 2^1$	$x = -11.010$	$x = -2^1 - 2^0 - 2^{-2}$	$x = -3.25$
1 100 1001	1	100	$x = -1.1001 \times 2^1$	$x = -11.001$	$x = -2^1 - 2^0 - 2^{-3}$	$x = -3.125$
1 100 1000	1	100	$x = -1.1000 \times 2^1$	$x = -11.000$	$x = -2^1 - 2^0$	$x = -3$
1 100 0111	1	100	$x = -1.0111 \times 2^1$	$x = -10.111$	$x = -2^1 - 2^{-1} - 2^{-2} - 2^{-3}$	$x = -2.875$
1 100 0110	1	100	$x = -1.0110 \times 2^1$	$x = -10.110$	$x = -2^1 - 2^{-1} - 2^{-2}$	$x = -2.75$
1 100 0101	1	100	$x = -1.0101 \times 2^1$	$x = -10.101$	$x = -2^1 - 2^{-1} - 2^{-3}$	$x = -2.625$
1 100 0100	1	100	$x = -1.0100 \times 2^1$	$x = -10.100$	$x = -2^1 - 2^{-1}$	$x = -2.5$
1 100 0011	1	100	$x = -1.0011 \times 2^1$	$x = -10.011$	$x = -2^1 - 2^{-2} - 2^{-3}$	$x = -2.375$
1 100 0010	1	100	$x = -1.0010 \times 2^1$	$x = -10.010$	$x = -2^1 - 2^{-2}$	$x = -2.25$
1 100 0001	1	100	$x = -1.0001 \times 2^1$	$x = -10.001$	$x = -2^1 - 2^{-3}$	$x = -2.125$
1 100 0000	1	100	$x = -1.0000 \times 2^1$	$x = -10.000$	$x = -2^1$	$x = -2$
0 100 0000	0	100	$x = +1.0000 \times 2^1$	$x = +10.000$	$x = 2^1$	$x = 2$
0 100 0001	0	100	$x = +1.0001 \times 2^1$	$x = +10.001$	$x = 2^1 + 2^{-3}$	$x = 2.125$
0 100 0010	0	100	$x = +1.0010 \times 2^1$	$x = +10.010$	$x = 2^1 + 2^{-2}$	$x = 2.25$
0 100 0011	0	100	$x = +1.0011 \times 2^1$	$x = +10.011$	$x = 2^1 + 2^{-2} + 2^{-3}$	$x = 2.375$
0 100 0100	0	100	$x = +1.0100 \times 2^1$	$x = +10.100$	$x = 2^1 + 2^{-1}$	$x = 2.5$
0 100 0101	0	100	$x = +1.0101 \times 2^1$	$x = +10.101$	$x = 2^1 + 2^{-1} + 2^{-3}$	$x = 2.625$
0 100 0110	0	100	$x = +1.0110 \times 2^1$	$x = +10.110$	$x = 2^1 + 2^{-1} + 2^{-2}$	$x = 2.75$
0 100 0111	0	100	$x = +1.0111 \times 2^1$	$x = +10.111$	$x = 2^1 + 2^{-1} + 2^{-2} + 2^{-3}$	$x = 2.875$
0 100 1000	0	100	$x = +1.0000 \times 2^1$	$x = +11.000$	$x = 2^1 + 2^0$	$x = 3$
0 100 1001	0	100	$x = +1.1001 \times 2^1$	$x = +11.001$	$x = 2^1 + 2^0 + 2^{-3}$	$x = 3.125$
0 100 1010	0	100	$x = +1.1010 \times 2^1$	$x = +11.010$	$x = 2^1 + 2^0 + 2^{-2}$	$x = 3.25$
0 100 1011	0	100	$x = +1.1011 \times 2^1$	$x = +11.011$	$x = 2^1 + 2^0 + 2^{-2} + 2^{-3}$	$x = 3.375$
0 100 1100	0	100	$x = +1.1100 \times 2^1$	$x = +11.100$	$x = 2^1 + 2^0 + 2^{-1}$	$x = 3.5$
0 100 1101	0	100	$x = +1.1101 \times 2^1$	$x = +11.101$	$x = 2^1 + 2^0 + 2^{-1} + 2^{-3}$	$x = 3.625$
0 100 1110	0	100	$x = +1.1110 \times 2^1$	$x = +11.110$	$x = 2^1 + 2^0 + 2^{-1} + 2^{-2}$	$x = 3.75$
0 100 1111	0	100	$x = +1.1111 \times 2^1$	$x = +11.111$	$x = 2^1 + 2^0 + 2^{-1} + 2^{-2} + 2^{-3}$	$x = 3.875$

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 101 1111	1	101	$x = -1.1111 \times 2^2$	$x = -111.11$	$x = -2^2 - 2^1 - 2^0 - 2^{-1} - 2^{-2}$	$x = -7.75$
1 101 1110	1	101	$x = -1.1110 \times 2^2$	$x = -111.10$	$x = -2^2 - 2^1 - 2^0 - 2^{-1}$	$x = -7.5$
1 101 1101	1	101	$x = -1.1101 \times 2^2$	$x = -111.01$	$x = -2^2 - 2^1 - 2^0 - 2^{-2}$	$x = -7.25$
1 101 1100	1	101	$x = -1.1100 \times 2^2$	$x = -111.00$	$x = -2^2 - 2^1 - 2^0$	$x = -7$
1 101 1011	1	101	$x = -1.1011 \times 2^2$	$x = -110.11$	$x = -2^2 - 2^1 - 2^{-1} - 2^{-2}$	$x = -6.75$
1 101 1010	1	101	$x = -1.1010 \times 2^2$	$x = -110.10$	$x = -2^2 - 2^1 - 2^{-1}$	$x = -6.5$
1 101 1001	1	101	$x = -1.1001 \times 2^2$	$x = -110.01$	$x = -2^2 - 2^1 - 2^{-2}$	$x = -6.25$
1 101 1000	1	101	$x = -1.1000 \times 2^2$	$x = -110.00$	$x = -2^2 - 2^1$	$x = -6.$
1 101 0111	1	101	$x = -1.0111 \times 2^2$	$x = -101.11$	$x = -2^2 - 2^0 - 2^{-1} - 2^{-2}$	$x = -5.75$
1 101 0110	1	101	$x = -1.0110 \times 2^2$	$x = -101.10$	$x = -2^2 - 2^0 - 2^{-1}$	$x = -5.5$
1 101 0101	1	101	$x = -1.0101 \times 2^2$	$x = -101.01$	$x = -2^2 - 2^0 - 2^{-2}$	$x = -5.25$
1 101 0100	1	101	$x = -1.0100 \times 2^2$	$x = -101.00$	$x = -2^2 - 2^0$	$x = -5$
1 101 0011	1	101	$x = -1.0011 \times 2^2$	$x = -100.11$	$x = -2^2 - 2^{-1} - 2^{-2}$	$x = -4.75$
1 101 0010	1	101	$x = -1.0010 \times 2^2$	$x = -100.10$	$x = -2^2 - 2^{-1}$	$x = -4.5$
1 101 0001	1	101	$x = -1.0001 \times 2^2$	$x = -100.01$	$x = -2^2 - 2^{-2}$	$x = -4.25$
1 101 0000	1	101	$x = -1.0000 \times 2^2$	$x = -100.00$	$x = -2^2$	$x = -4$
0 101 0000	0	101	$x = +1.0000 \times 2^2$	$x = +100.00$	$x = 2^2$	$x = 4$
0 101 0001	0	101	$x = +1.0001 \times 2^2$	$x = +100.01$	$x = 2^2 + 2^{-2}$	$x = 4.25$
0 101 0010	0	101	$x = +1.0010 \times 2^2$	$x = +100.10$	$x = 2^2 + 2^{-1}$	$x = 4.5$
0 101 0011	0	101	$x = +1.0011 \times 2^2$	$x = +100.11$	$x = 2^2 + 2^{-1} + 2^{-2}$	$x = 4.75$
0 101 0100	0	101	$x = +1.0100 \times 2^2$	$x = +101.00$	$x = 2^2 + 2^0$	$x = 5$
0 101 0101	0	101	$x = +1.0101 \times 2^2$	$x = +101.01$	$x = 2^2 + 2^0 + 2^{-2}$	$x = 5.25$
0 101 0110	0	101	$x = +1.0110 \times 2^2$	$x = +101.10$	$x = 2^2 + 2^0 + 2^{-1}$	$x = 5.5$
0 101 0111	0	101	$x = +1.0111 \times 2^2$	$x = +101.11$	$x = 2^2 + 2^0 + 2^{-1} + 2^{-2}$	$x = 5.75$
0 101 1000	0	101	$x = +1.0000 \times 2^2$	$x = +110.00$	$x = 2^2 + 2^1$	$x = 6$
0 101 1001	0	101	$x = +1.1001 \times 2^2$	$x = +110.01$	$x = 2^2 + 2^1 + 2^{-2}$	$x = 6.25$
0 101 1010	0	101	$x = +1.1010 \times 2^2$	$x = +110.10$	$x = 2^2 + 2^1 + 2^{-1}$	$x = 6.5$
0 101 1011	0	101	$x = +1.1011 \times 2^2$	$x = +110.11$	$x = 2^2 + 2^1 + 2^{-1} + 2^{-2}$	$x = 6.75$
0 101 1100	0	101	$x = +1.1100 \times 2^2$	$x = +111.00$	$x = 2^2 + 2^1 + 2^0$	$x = 7$
0 101 1101	0	101	$x = +1.1101 \times 2^2$	$x = +111.01$	$x = 2^2 + 2^1 + 2^0 + 2^{-2}$	$x = 7.25$
0 101 1110	0	101	$x = +1.1110 \times 2^2$	$x = +111.10$	$x = 2^2 + 2^1 + 2^0 + 2^{-1}$	$x = 7.5$
0 101 1111	0	101	$x = +1.1111 \times 2^2$	$x = +111.11$	$x = 2^2 + 2^1 + 2^0 + 2^{-1} + 2^{-2}$	$x = 7.75$

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 1.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 110 1111	1	110	$x = -1.1111 \times 2^3$	$x = -1111.1$	$x = -2^3 - 2^2 - 2^1 - 2^0 - 2^{-1}$	$x = -15.5$
1 110 1110	1	110	$x = -1.1110 \times 2^3$	$x = -1111.0$	$x = -2^3 - 2^2 - 2^1 - 2^0$	$x = -15$
1 110 1101	1	110	$x = -1.1101 \times 2^3$	$x = -1110.1$	$x = -2^3 - 2^2 - 2^1 - 2^{-1}$	$x = -14.5$
1 110 1100	1	110	$x = -1.1100 \times 2^3$	$x = -1110.0$	$x = -2^3 - 2^2 - 2^1$	$x = -14$
1 110 1011	1	110	$x = -1.1011 \times 2^3$	$x = -1101.1$	$x = -2^3 - 2^2 - 2^0 - 2^{-1}$	$x = -13.5$
1 110 1010	1	110	$x = -1.1010 \times 2^3$	$x = -1101.0$	$x = -2^3 - 2^2 - 2^0$	$x = -13$
1 110 1001	1	110	$x = -1.1001 \times 2^3$	$x = -1100.1$	$x = -2^3 - 2^2 - 2^{-1}$	$x = -12.5$
1 110 1000	1	110	$x = -1.1000 \times 2^3$	$x = -1100.0$	$x = -2^3 - 2^2$	$x = -12.$
1 110 0111	1	110	$x = -1.0111 \times 2^3$	$x = -1011.1$	$x = -2^3 - 2^1 - 2^0 - 2^{-1}$	$x = -11.5$
1 110 0110	1	110	$x = -1.0110 \times 2^3$	$x = -1011.0$	$x = -2^3 - 2^1 - 2^0$	$x = -11$
1 110 0101	1	110	$x = -1.0101 \times 2^3$	$x = -1010.1$	$x = -2^3 - 2^1 - 2^{-1}$	$x = -10.5$
1 110 0100	1	110	$x = -1.0100 \times 2^3$	$x = -1010.0$	$x = -2^3 - 2^1$	$x = -10$
1 110 0011	1	110	$x = -1.0011 \times 2^3$	$x = -1001.1$	$x = -2^3 - 2^0 - 2^{-1}$	$x = -9.5$
1 110 0010	1	110	$x = -1.0010 \times 2^3$	$x = -1001.0$	$x = -2^3 - 2^0$	$x = -9$
1 110 0001	1	110	$x = -1.0001 \times 2^3$	$x = -1000.1$	$x = -2^3 - 2^{-1}$	$x = -8.5$
1 110 0000	1	110	$x = -1.0000 \times 2^3$	$x = -1000.0$	$x = -2^3$	$x = -8$
0 110 0000	0	110	$x = +1.0000 \times 2^3$	$x = +1000.0$	$x = 2^3$	$x = 8$
0 110 0001	0	110	$x = +1.0001 \times 2^3$	$x = +1000.1$	$x = 2^3 + 2^{-1}$	$x = 8.5$
0 110 0010	0	110	$x = +1.0010 \times 2^3$	$x = +1001.0$	$x = 2^3 + 2^0$	$x = 9$
0 110 0011	0	110	$x = +1.0011 \times 2^3$	$x = +1001.1$	$x = 2^3 + 2^0 + 2^{-1}$	$x = 9.5$
0 110 0100	0	110	$x = +1.0100 \times 2^3$	$x = +1010.0$	$x = 2^3 + 2^1$	$x = 10$
0 110 0101	0	110	$x = +1.0101 \times 2^3$	$x = +1010.1$	$x = 2^3 + 2^1 + 2^{-1}$	$x = 10.5$
0 110 0110	0	110	$x = +1.0110 \times 2^3$	$x = +1011.0$	$x = 2^3 + 2^1 + 2^0$	$x = 11$
0 110 0111	0	110	$x = +1.0111 \times 2^3$	$x = +1011.1$	$x = 2^3 + 2^1 + 2^0 + 2^{-1}$	$x = 11.5$
0 110 1000	0	110	$x = +1.0000 \times 2^3$	$x = +1100.0$	$x = 2^3 + 2^2$	$x = 12$
0 110 1001	0	110	$x = +1.1001 \times 2^3$	$x = +1100.1$	$x = 2^3 + 2^2 + 2^{-1}$	$x = 12.5$
0 110 1010	0	110	$x = +1.1010 \times 2^3$	$x = +1101.0$	$x = 2^3 + 2^2 + 2^0$	$x = 13$
0 110 1011	0	110	$x = +1.1011 \times 2^3$	$x = +1101.1$	$x = 2^3 + 2^2 + 2^0 + 2^{-1}$	$x = 13.5$
0 110 1100	0	110	$x = +1.1100 \times 2^3$	$x = +1110.0$	$x = 2^3 + 2^2 + 2^1$	$x = 14$
0 110 1101	0	110	$x = +1.1101 \times 2^3$	$x = +1110.1$	$x = 2^3 + 2^2 + 2^1 + 2^{-1}$	$x = 14.5$
0 110 1110	0	110	$x = +1.1110 \times 2^3$	$x = +1111.0$	$x = 2^3 + 2^2 + 2^1 + 2^0$	$x = 15$
0 110 1111	0	110	$x = +1.1111 \times 2^3$	$x = +1111.1$	$x = 2^3 + 2^2 + 2^1 + 2^0 + 2^{-1}$	$x = 15.5$

realmax = largest floating-point number that does not overflow

Here is a table of all subnormal numbers for binary8

IEEE Quarter-Precision Raw 8-Bit Word	Sign bit	Exponent Code	Numerical Value (floating-point binary)	Numerical Value (fixed-point binary)	Numerical Value as a sum in Decimal	Numerical Value in Decimal
$B_7\ B_6B_5B_4\ B_3B_2B_1B_0$	B_7	$B_6B_5B_4$	$\pm 0.b_1b_2b_3b_4 \times 2^e$	$b_\ell \ldots b_0.b_{-1}\ldots b_{-f}$		
1 000 1111	1	000	$x = -0.1111 \times 2^{-2}$	$x = -0.001111$	$x = -2^{-3} - 2^{-4} - 2^{-5} - 2^{-6}$	$x = -0.234375$
1 000 1110	1	000	$x = -0.1110 \times 2^{-2}$	$x = -0.001110$	$x = -2^{-3} - 2^{-4} - 2^{-5}$	$x = -0.21875$
1 000 1101	1	000	$x = -0.1101 \times 2^{-2}$	$x = -0.001101$	$x = -2^{-3} - 2^{-4} - 2^{-6}$	$x = -0.203125$
1 000 1100	1	000	$x = -0.1100 \times 2^{-2}$	$x = -0.001100$	$x = -2^{-3} - 2^{-4}$	$x = -0.1875$
1 000 1011	1	000	$x = -0.1011 \times 2^{-2}$	$x = -0.001011$	$x = -2^{-3} - 2^{-5} - 2^{-6}$	$x = -0.171875$
1 000 1010	1	000	$x = -0.1010 \times 2^{-2}$	$x = -0.001010$	$x = -2^{-3} - 2^{-5}$	$x = -0.15625$
1 000 1000	1	000	$x = -0.1001 \times 2^{-2}$	$x = -0.001001$	$x = -2^{-3} - 2^{-6}$	$x = -0.140625$
1 000 1000	1	000	$x = -0.1000 \times 2^{-2}$	$x = -0.001000$	$x = -2^{-3}$	$x = -0.125$
1 000 0111	1	000	$x = -0.0111 \times 2^{-2}$	$x = -0.000111$	$x = -2^{-4} - 2^{-5} - 2^{-6}$	$x = -0.109375$
1 000 0110	1	000	$x = -0.0110 \times 2^{-2}$	$x = -0.000110$	$x = -2^{-4} - 2^{-6}$	$x = -0.09375$
1 000 0101	1	000	$x = -0.0101 \times 2^{-2}$	$x = -0.000101$	$x = -2^{-4} - 2^{-6}$	$x = -0.078125$
1 000 0100	1	000	$x = -0.0100 \times 2^{-2}$	$x = -0.000100$	$x = -2^{-4}$	$x = -0.0625$
1 000 0011	1	000	$x = -0.0011 \times 2^{-2}$	$x = -0.000011$	$x = -2^{-5} - 2^{-6}$	$x = -0.046875$
1 000 0010	1	000	$x = -0.0010 \times 2^{-2}$	$x = -0.000010$	$x = -2^{-5}$	$x = -0.03125$
1 000 0001	1	000	$x = -0.0001 \times 2^{-2}$	$x = -0.000001$	$x = -2^{-6}$	$x = -0.015625$
1 000 0000	1	000	$x = -0.0000 \times 2^{-2}$	$x = -0.000000$	$x = -0$	$x = -0$
0 000 0000	0	000	$x = +0.0000 \times 2^{-2}$	$x = +0.000000$	$x = 0$	$x = 0$
0 000 0001	0	000	$x = +0.0001 \times 2^{-2}$	$x = +0.000001$	$x = 2^{-6}$	$x = 0.015625$
0 000 0010	0	000	$x = +0.0010 \times 2^{-2}$	$x = +0.000010$	$x = 2^{-5}$	$x = 0.03125$
0 000 0011	0	000	$x = +0.0011 \times 2^{-2}$	$x = +0.000011$	$x = 2^{-5} + 2^{-6}$	$x = 0.046875$
0 000 0100	0	000	$x = +0.0100 \times 2^{-2}$	$x = +0.000100$	$x = 2^{-4}$	$x = 0.0625$
0 000 0101	0	000	$x = +0.0101 \times 2^{-2}$	$x = +0.000101$	$x = 2^{-4} + 2^{-6}$	$x = 0.078125$
0 000 0110	0	000	$x = +0.0110 \times 2^{-2}$	$x = +0.000110$	$x = 2^{-4} + 2^{-6}$	$x = 0.09375$
0 000 0111	0	000	$x = +0.0111 \times 2^{-2}$	$x = +0.000111$	$x = 2^{-4} + 2^{-5} + 2^{-6}$	$x = 0.109375$
0 000 1000	0	000	$x = +0.1000 \times 2^{-2}$	$x = +0.001000$	$x = 2^{-3}$	$x = 0.125$
0 000 1001	0	000	$x = +0.1001 \times 2^{-2}$	$x = +0.001001$	$x = 2^{-3} + 2^{-6}$	$x = 0.140625$
0 000 1010	0	000	$x = +0.1010 \times 2^{-2}$	$x = +0.001010$	$x = 2^{-3} + 2^{-5}$	$x = 0.15625$
0 000 1011	0	000	$x = +0.1011 \times 2^{-2}$	$x = +0.001011$	$x = 2^{-3} + 2^{-5} + 2^{-6}$	$x = 0.171875$
0 000 1100	0	000	$x = +0.1100 \times 2^{-2}$	$x = +0.001100$	$x = 2^{-3} + 2^{-4}$	$x = 0.1875$
0 000 1101	0	000	$x = +0.1101 \times 2^{-2}$	$x = +0.001101$	$x = 2^{-3} + 2^{-4} + 2^{-6}$	$x = 0.203125$
0 000 1110	0	000	$x = +0.1110 \times 2^{-2}$	$x = +0.001110$	$x = 2^{-3} + 2^{-4} + 2^{-6}$	$x = 0.21875$
0 000 1111	0	000	$x = +0.1111 \times 2^{-2}$	$x = +0.001111$	$x = 2^{-3} + 2^{-4} + 2^{-5} + 2^{-6}$	$x = 0.234375$