**Họ và tên: ……………………………………………. MSSV: ……………….**

**WEEK 03 LAB – FLOATING POINT**

**Thực hành**

**Practice Problem 2.45**Fill in the missing information in the following table:

|  |  |  |
| --- | --- | --- |
| Fractional value | Binary representation | Decimal representation |
| 1/8 | 0.001 | 0.125 |
| 3/4 | 0.110 | 0.75 |
| 25/16 | 1.1001 | 1.5625 |
| 43/16 | 10.1011 | 2.6875 |
| 9/8 | 1.001 | 1.125 |
| 47/8 | 101.111 | 5.875 |
| 51/16 | 11.0011 | 3.1875 |

Practice Problem 2.47Consider a 5-bit floating-point representation based on the IEEE floating-point format, with one sign bit, two exponent bits (k = 2), and two fraction bits (n = 2). The exponent bias is 22-1 - 1 = 1.  
The table that follows enumerates the entire nonnegative range for this 5-bit floating-point representation. Fill in the blank table entries using the following directions:

e: The value represented by considering the exponent field to be an unsigned integer  
E: The value of the exponent after biasing  
2E: The numeric weight of the exponent  
f : The value of the fraction

*M*: The value of the significand  
*M* × 2*E*: The (unreduced) fractional value of the number  
*V* : The reduced fractional value of the number  
Decimal: The decimal representation of the number  
Express the values of 2*E*, *f* , *M*, *M* × 2*E*, and *V* either as integers (when possible) or as fractions of the form *x/y*, where *y* is a power of 2. You need not fill in entries marked “—”.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Bit* | *e* | *E* | *2E* | *f* | *M* | *Mx2E* | *V* | *Decimal* |
| 0 00 00 | 0 | 0 | 0 | 0/4 | 0/4 | 0/4 | 0 | 0 |
| 0 00 01 | 0 | 0 | 1 | 1/4 | 1/4 | 1/4 | 1/4 | 0.25 |
| 0 00 10 | 0 | 0 | 1 | 3/4 | 3/4 | 3/4 | 3/4 | 0.75 |
| 0 01 00 | 0 | 0 | 1 | 0/4 | 0/4 | 4/4 | 1 | 1.0 |
| 0 01 01 | 1 | 0 | 1 | 1/4 | 5/4 | 5/4 | 5/4 | 1.25 |
| 0 01 10 | 1 | 0 | 1 | 2/4 | 6/4 | 6/4 | 3/2 | 1.5 |
| 0 01 11 | 1 | 0 | 1 | 3/4 | 7/4 | 7/4 | 7/4 | 1.75 |
| 0 10 00 | 2 | 1 | 2 | 0/4 | 4/4 | 8/4 | 2 | 2.0 |
| 0 10 01 | 2 | 1 | 2 | 1/4 | 5/4 | 10/4 | 5/2 | 2.5 |
| 0 10 10 | 2 | 1 | 2 | 2/4 | 6/4 | 12/4 | 3 | 3.0 |
| 0 10 11 | 2 | 1 | 2 | 3/4 | 7/4 | 14/4 | 7/2 | 3.5 |
| 0 11 00 | — | — | — | — | — | — |  | — |
| 0 11 01 | — | — | — | — | — | — | NaN | — |
| 0 11 10 | — | — | — | — | — | — | NaN | — |
| 0 11 11 | — | — | — | — | — | — | NaN | — |

**Practice Problem 2.48**As mentioned in Problem 2.6, the integer 3,510,593 has hexadecimal representation 0x00359141, while the single-precision, floating-point number 3510593.0 has hexadecimal representation 0x4A564504. Derive this floating-point representation and explain the correlation between the bits of the integer and floating-point representations.

B1 (đổi sang binary)

110101 1001 0001 0100 0001 . 0

B2 normalize

(-1)^s \* M \* 2^e

M = 1.10101 1001 0001 0100 0001 0 = 1.f

B3 exp, s, fraction

S = 0

Exp = E + bias = 21 + 127 = 148 = 1001 0100

Frac = 10101 1001 0001 0100 0001 00

BB4 IEEE754

0x4A564504

**Practice Problem 2.50**Show how the following binary fractional values would be rounded to the nearest half (1 bit to the right of the binary point), according to the round-to-even rule. In each case, show the numeric values, both before and after rounding.

|  |  |  |  |
| --- | --- | --- | --- |
| **Original** | | **Rounded** | |
| 10.0102 | 2x(1/4) | 10.0 | 2 |
| 10.0112 | 2x(3/8) | 10.1 | 2(1/2) |
| 10.1102 | 2x(3/4) | 11.0 | 3 |
| 11.0012 | 3x(1/8) | 11.0 | 3 |