Started on	Friday, 11 September 2020, 4:08 PM
State	Finished
Completed on	Friday, 11 September 2020, 5:01 PM
Time taken	53 mins 4 secs
Marks	10.00/14.00
Grade	**7.14** out of 10.00 (**71**%)

Question **1**

Correct

Mark 1.00 out of 1.00

It is always true that "a + b + c" and "c + b + a" will produce the same result for IEEE floating values "a", "b" and "c".

Select one:

- True
- False ✔

Question **2**

Correct

Mark 1.00 out of 1.00

It is always true that "a * b" and "b * a" will produce the same result for IEEE floating values "a" and "b".

Select one:

- True ✔
- False

Question **3**

Correct

Mark 1.00 out of 1.00

Convert the binary fractional number 0.00111_2 to a decimal representation. You can use fractions or floating point values, but your answer must be precise. You can use expressions if that's useful.

7/32

Your last answer was interpreted as follows: $\frac{7}{32}$

Correct answer, well done.

Question **4**

Correct

Mark 1.00 out of 1.00

Convert the binary fractional number 10.00011_2 to a decimal representation. You can use fractions or floating point values, but your answer must be precise. You can use expressions if that's useful.

2.09375

Your last answer was interpreted as follows: 2.09375

Correct answer, well done.

Question **5**

Partially correct

Mark 0.50 out of 1.00

Consider a 12-bit floating-point representation based on the IEEE floating-point format, with one sign bit, 5 exponent bits and 6 fraction bits. The exponent bias follows the IEEE standard.

Interpret the bitstring $1\_00001\_101101_2$ using this 12-bit floating-point representation and fill in the table below.

You must express your results precisely (e.g. using rationals (M/N) and exponents (M^N), rather than approximate decimal fractions such as 0.ABCD). You may use expressions if it's useful.

Field	Mean	Value
e	The value represented by considering the exponent field to be an unsigned integer (as a decimal value)	1 Your last answer was interpreted as follows: 1
E	The value of the exponent after biasing (as a decimal value)	-14 Your last answer was interpreted as follows: -14
2^E	The numeric weight of the exponent (as a decimal value)	2^(-14) Your last answer was interpreted as follows: 2^{-14}
f	The value of the fraction (as a fraction such as 3/4 or the exact floating point number)	109/64 Your last answer was interpreted as follows: $\frac{109}{64}$
M	The value of the significand (as a fraction such such as 7/4 or the exact floating point number)	45/64 Your last answer was interpreted as follows: $\frac{45}{64}$
$s*2^E * M$	The value of the number in decimal. The 's' is equal to +1 if the number is positive and -1 if it is negative.	-(45/46)*2^(-14) Your last answer was interpreted as follows: $\left(-\frac{45}{46}\right) \cdot 2^{-14}$

Your answer is partially correct.

Question **6**

Partially correct

Mark 0.50 out of 1.00

Consider a 10-bit floating-point representation based on the IEEE floating-point format, with one sign bit, 4 exponent bits and 5 fraction bits. The exponent bias follows the IEEE standard.

Interpret the bitstring $1\_0000\_11101_2$ using this 10-bit floating-point representation and fill in the table below.

You must express your results precisely (e.g. using rationals (M/N) and exponents (M^N), rather than approximate decimal fractions such as 0.ABCD). You may use expressions if it's useful.

Field	Mean	Value
e	The value represented by considering the exponent field to be an unsigned integer (as a decimal value)	0 Your last answer was interpreted as follows: 0
E	The value of the exponent after biasing (as a base-10 value)	-7 Your last answer was interpreted as follows: -7
2^E	The numeric weight of the exponent (as a base-10 value)	2^-7 Your last answer was interpreted as follows: 2^{-7}
f	The value of the fraction (as a ratio such as 3/4 or the exact floating point number)	29/32 Your last answer was interpreted as follows: $\frac{29}{32}$
M	The value of the significand (as a fraction such such as 7/4 or the exact floating point number)	29/32 Your last answer was interpreted as follows: $\frac{29}{32}$
$s*2^E * M$	The value of the number in decimal. The 's' is equal to +1 if the number is positive and -1 if it is negative.	-(29/32)*2^-7 Your last answer was interpreted as follows: $\left(-\frac{29}{32}\right) \cdot 2^{-7}$

Your answer is partially correct.

Question **7**

Partially correct

Mark 0.50 out of 1.00

Consider a 7-bit floating-point representation based on the IEEE floating-point format, with one sign bit, 3 exponent bits and 3 fraction bits. The exponent bias follows the IEEE standard.

Interpret the bitstring $1\_101\_001_2$ using this 7-bit floating-point representation and fill in the table below.

You must express your results precisely (e.g. using rationals (M/N) and exponents (M^N), rather than approximate decimal fractions such as 0.ABCD). You may use expressions if it's useful.

Field	Mean	Value
e	The value represented by considering the exponent field to be an unsigned integer (as a decimal value)	5 Your last answer was interpreted as follows: 5
E	The value of the exponent after biasing (as a decimal value)	2 Your last answer was interpreted as follows: 2
2^E	The numeric weight of the exponent (as a decimal value)	2^2 Your last answer was interpreted as follows: 2^2
f	The value of the fraction (as a fraction such as 3/4 or the exact floating point number)	41/8 Your last answer was interpreted as follows: $\frac{41}{8}$
M	The value of the significand (as a fraction such such as 7/4 or the exact floating point number)	1/8 Your last answer was interpreted as follows: $\frac{1}{8}$
$s*2^E * M$	The value of the number in decimal. The 's' is equal to +1 if the number is positive and -1 if it is negative.	-(1/8)*2^-7 Your last answer was interpreted as follows: $\left(-\frac{1}{8}\right) \cdot 2^{-7}$

Your answer is partially correct.

Question **8**

Partially correct

Mark 0.50 out of 1.00

Consider a 6-bit floating-point representation based on the IEEE floating-point format, with one sign bit, 3 exponent bits and 2 fraction bits. The exponent bias follows the IEEE standard.

Interpret the bitstring $1\_000\_10_2$ using this 6-bit floating-point representation and fill in the table below.

You must express your results precisely (e.g. using rationals (M/N) and exponents (M^N), rather than approximate decimal fractions such as 0.ABCD). You may use expressions if it's useful.

Field	Mean	Value
e	The value represented by considering the exponent field to be an unsigned integer (as a decimal value)	0 Your last answer was interpreted as follows: 0
E	The value of the exponent after biasing (as a decimal value)	2 Your last answer was interpreted as follows: 2
2^E	The numeric weight of the exponent (as a decimal value)	2^2 Your last answer was interpreted as follows: 2^2
f	The value of the fraction (as a fraction such as 3/4 or the exact floating point number)	1/2 Your last answer was interpreted as follows: $\frac{1}{2}$
M	The value of the significand (as a fraction such such as 7/4 or the exact floating point number)	1/2 Your last answer was interpreted as follows: $\frac{1}{2}$
$s*2^E * M$	The value of the number in decimal. The 's' is equal to +1 if the number is positive and -1 if it is negative.	-(1/2)*2^2 Your last answer was interpreted as follows: $\left(-\frac{1}{2}\right) \cdot 2^2$

Your answer is partially correct.

Question **9**

Correct

Mark 1.00 out of 1.00

Show how the following binary fractional values would be rounded to the nearest half (one 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 and the bit pattern after rounding.

Finally, please DO NOT USE fractions and type the decimal values very precisely (accurate to the last decimal place).

Binary Pattern	Value before rounding	Bit pattern rounded to the nearest half	Value after rounding
10.010	2.25 ✔	10.0 ✔	2 ✔
10.111	2.875 ✔	11.0 ✔	3 ✔
10.110	2.75 ✔	11.0 ✔	3 ✔
11.001	3.125 ✔	11.0 ✔	3 ✔

Question **10**

Partially correct

Mark 3.00 out of 5.00

Work through question number 2.88 in the *Computer Systems: A Programmer's Perspective, 2nd Edition* textbook:

We are running programs on a machine where values of type int have a 32- bit two's-complement representation. Values of type float use the 32-bit IEEE format, and values of type double use the 64-bit IEEE format. We generate arbitrary integer values x, y, and z, and convert them to values of type double as follows:

```
/* Create some arbitrary values */
int x = random();
int y = random();
int z = random();
/* Convert to double */
double dx = (double) x;
double dy = (double) y;
double dz = (double) z;
```

For each of the following C expressions, you are to indicate whether or not the expression always yields 1. If it always yields 1, describe the underlying mathematical principles. Otherwise, give an example of arguments that make it yield 0. Note that you cannot use an IA32 machine running gcc to test your answers, since it would use the 80-bit extended-precision representation for both float and double.

A. (float) x == (float) dx

Sometimes, depends on what the values of x and dx are ✖

B. dx-dy == (double) (x-y) Yes, double can completely hold value of a int ✖

C. (dx + dy) + dz == dx + (dy + dz) Yes, values of integer sums can be represented exactly ✔

D. (dx * dy) * dz == dx * (dy * dz) No, floating point multiplication not associative. ✔

E. dx/dx == dz/dz No, let x=0 and z = 1d ✔