Unit 12_010 – An Example

1. What data type does the program use for a floating point number?

2. What instructions are bing used for moving and adding floating point numbers?

3. What registers are you seeing used for floating point numbers?

Unit 12_020 – Decimal vs Binary Floating Point

4. Do integers always translate exactly between Decimal and Binary?

5. Do fractions expressed as decimals always translate exactly between Decimal and Binary? Are there some situations where they do correspond exactly? Give an example of each.

6. How do many financial calculations avoid the problem of amounts expressed in dollars not always corresponding in decimal and binary representations.

Unit 12_030 – IEEE 754

7. What does the IEEE 754 standard relate to?

8. How would 1234 be represented in scientific notation? How would 0.00001234 be represented in scientific notation?

9. In scientific notation, how many significant digits are to the left of the decimal?

10. In IEEE 754 notation, how many significant digits are to the left of the decimal?

11. What is the mantissa in a number such as 0.1234×10^{-4}

12. In the IEEE 754 standard, how many bits are used to represent the sign of the number in single precision? How many bits are used to represent the sign in double precision?

    single: _____   double: _____

13. In the IEEE 754 standard, how many bits are used to represent the exponent of the number in single precision? How many bits are used to represent the exponent in double precision?

    single: _____   double: _____

14. In the IEEE 754 standard, how many bits are used to represent the mantissa of the number in single precision? How many bits are used to represent the mantissa in double precision?

    single: _____   double: _____

Unit 12_040 – Generations

15. Traditional CPUs could only do **integer/floating point**(circle one) math operations in hardware. **Integer/Floating point**(circle one) operations had to be done with software.

16. Circle the operations that traditional CPUs could do before the introduction of the FPU.

    - addition
    - subtraction
    - multiplication
    - division
    - bit manipulation (and, or, xor, etc)

17. If CPUs could not do floating point arithmetic and complex math, how were math and floating point problems processed?

18. What did the 8087 coprocessor do?

19. What was the name of the floating point operation that was introduced on the Pentium processors?

_____

Unit 12_050 – Precision

Video Length

20. What data type is used in Java and C for single precision floating point numbers? _____ How many bits are used? _____

21. What data type is used in Java and C for double precision floating point numbers? _____ How many bits are used? _____

    In Java, Python, and modern C, what is the default or most common data type used with floating point numbers? _____

Unit 12_060 – Registers

22. How wide are the floating point registers in SSE2?

23. How many floating point registers are there in an x86-64 processor?

24. What are the names of the floating point registers in an x86-64 processor

25. What are lanes? How many lanes would be used if there are 4 32-bit numbers stored in the same register?

26. What lane (or lanes) do scalar values use?

27. What is the purpose of packing multiple values into one register?

28. What is the command to do each of the following single precision operations?

    • _____ move

    • _____ add

    • _____ subtract

    • _____ multiply

    • _____ divide

29. What is the command to do each of the following double precision operations?

    • _____ move

    • _____ add

    • _____ subtract

    • _____ multiply

    • _____ divide

30. In what ways are multiplication and division with floating point different than multiplication and division using integers? Which sounds easier?

Unit 12_080 – printf

31. What is the format specifier for floating point numbers in decimal notation?

32. What is the format specifier that forces scientific notation?

33. What is the format specifier that lets the printf command determine whether to use decimal or scientific notation?

34. How is the format string passed to the printf function with floating point numbers? How does it compare to passing integers?

35. If there are three floating point numbers passed to the printf, which registers would be used?

36. If there are three floating point values plus the format string, what number is placed in the EAX register before calling printf?

_____