



Batch: A3 Roll No.: 16010121051

Experiment / assignment / tutorial No. _10_

Grade: AA / AB / BB / BC / CC / CD / DD

Signature of the Staff In-charge with date

TITLE: Study of basic computer organisation and architecture concepts through Virtual lab

AIM: Understanding Virtual Lab concepts

Expected OUTCOME of Experiment:

The objective of this experiment is to learn the fundamentals of Floating Point Representation of Numbers.

Books/ Journals/ Websites referred:

<http://vlabs.iitb.ac.in/vlab/labscse.html>

<http://vlabs.iitb.ac.in/vlab/#>

<http://www.vlab.co.in/>

Pre Lab/ Prior Concepts:

The main aim of this experiment is to provide remote-access to Labs in various disciplines of Science and Engineering. These Virtual Labs would cater to students at the undergraduate level, post graduate level as well as to research scholars. Also, to enthuse students to conduct experiments by arousing their curiosity. This would help them in learning basic and advanced concepts through remote experimentation. It also provides a complete Learning Management System around the Virtual Labs where the students can avail the various tools for learning, including additional web-resources, video-lectures, animated demonstrations and self-evaluation. We can share costly equipment



and resources, which are otherwise available to limited number of users due to constraints on time and geographical distances

Salient Features:

1. Virtual Labs will provide to the students the result of an experiment by one of the following methods (or possibly a combination)

- Modeling the physical phenomenon by a set of equations and carrying out simulations to yield the result of the particular experiment. This can, at-the-best, provide an approximate version of the 'real-world' experiment.
- Providing measured data for virtual lab experiments corresponding to the data previously obtained by measurements on an actual system.
- Remotely triggering an experiment in an actual lab and providing the student the result of the experiment through the computer interface. This would entail carrying out the actual lab experiment remotely.

2. Virtual Labs will be made more effective and realistic by providing additional inputs to the students like accompanying audio and video streaming of an actual lab experiment and equipment.

Observations

Title of Study Experiment:

Floating Point Numbers Representation

Brief description of experiment under study:

Computers use computations with integers and real numbers. In computers we cannot precisely represent all the real numbers. Hence, there is a way to effectively represent them with only a little loss in precision.

Floating point numbers are numbers that contain floating decimal points. Computers recognize real numbers that contain fractions as floating point numbers. When a calculation includes a floating point number, it is called a "floating point calculation." Older computers used to have a separate floating point unit (FPU) that handled these calculations, but now the FPU is typically built into the computer's CPU.

This experiment is to understand the concept of Floating Point Numbers and how they are converted to and from decimal form.

Learning's recorded:



Exponent Field

8 - Bits Long

Determines The Range Of Numbers That Can be Represented

Increasing The Bits Will Increase The Range , Not Precision -> To Cover For -ve Numbers , $\text{exp} = 127 + \text{real exp}$

Sign Bit

1- Bit Long

Determines The +ve or -ve number -> 1 = -ve Number 0 = +ve Number

Mantissa Field

23 - Bits Long

Determines the precision of Numbers

Increasing bits Will Increase precision, not range.

Procedure for conversion:

- Convert the absolute value of the number to binary, perhaps with a fractional part after the binary point. This can be done by converting the integral and fractional parts separately.
- Normalize the number. Move the binary point so that it is one bit from the left. Adjust the exponent of two so that the value does not change. Place the mantissa into the mantissa field of the number. Omit the leading one and fill with zeros on the right.
- Add the bias to the exponent of two, and place it in the exponent field. The bias is $2^{k-1}-1$, where k is the number of bits in the exponent field. 6.
- For the eight-bit format, k-3, so the bias is 2^3-1-3 . For IEEE 32-bit, k-8, so the bias is $2^8-1-127$.



- Set the sign bit, 1 for negative, 0 for positive, according to the sign of the original number.

Screenshots from Simulator:

Simulations

DECIMAL NUMBER
123.125

BITS FOR EXPONENT
8

Submit **Reset**

RESULTS

8-bit binary
01000101

Binary Representation Of Integral Part
1111011

Binary Representation Of Fractional Part
001

Binary Representation of the Number
1111011.001

Normalised Representation of the Number
1. X 2 power 6

Sign
0

Bias
127

Mantiss
133

Expone
133

Handwritten:

Decimal number: 123.125

STEP 1: To Binary: 1111011.001

2 | 123 0.125 x 2 = 0.250 - 0
2 | 61 - 1 0.250 x 2 = 0.500 - 0
2 | 30 - 1 0.500 x 2 = 1.00 - 1
2 | 15 - 0
2 | 7 - 1
2 | 3 - 1
1 - 1

STEP 2: Normalization: 1.111011001 x 2⁶
power = 6.

STEP 3: Single Precision Double Precision
E = 127 = 6 E = 1023 = 6
E = 127 + 6 = 133 E = 6 + 1023 = 1029
In Binary: 10000101 In Binary: 1000000101

STEP 4: Representation:

Single Precision:

Sign	Biased Exponent	Mantissa
0	10000101	111011001000...00

23 bits

Double Precision:

Sign	Biased Exponent	Mantissa
0	1000000101	111011001000...0

52 bits



Knowledge gained / Inference Obtained :

The importance of Floating Point Representation was understood: Floating point representation makes numerical computation much easier.

The procedure to convert a decimal to floating point representation was also studied with the help of an example and verifying it manually as well.

Post Lab Descriptive Questions

1. What are the applications of the virtual lab case study / tool reviewed by you?

- **Tensor Processing Units (TPUs)**

Besides the 64-bit float we explored at length, there are also 32-bit floats (single precision) and 16-bit floats (half-precision) commonly available.

Google's Tensor Processing Units instead use a modified 16-bit format for multiplication as part of their many optimizations for deep-learning tasks.

- **HDR Images**

HDR image uses floating point numbers to represent the pixels! This allows a high “dynamic” range (the exponent can be high or low) while still maintaining relative precision across all brightness scales. Perfect for keeping the data from scenes with high contrast.