OMEGA ACADEMY, NUMERICAL METHODS COURSE.

Erika Jissel Gutiérrez Beltrán

Daniel Fernandez Delgado

Frank Edward Daza González

Johanna Arias

Freddy Sebastian Garcia

Teacher:

Walter German Magaña

Matter:

Numerical Methods

Universidad de San Buenaventura Cali

2014

Guide numerical methods.
Multimedia Engineering and Systems Engineering

Section two

Convert binary number to IEEE 754 floating-point format 32 bits computer memory.

In order to transform a binary number into IEEE 754, we first must know its structure, in this case how it behaves on a 32 bit computer memory.

It's divided in 3 parts: sign bit (1), exponent width (8 bits), mantissa (23 bits).

Sign bit: determines the sign of the number, if it 0 if positive, 1 if negative.

Exponent width: determines what the exponent is going to be

Mantissa: expresses the fraction 2^23.

Example:

Convert the binary number 110010010.1 into IEEE 754 floating point format

in computer memory of thirty-two (32) bits.

1. First step is to transform the binary number 110010010.0 into its scientific notation, moving the comma to the first position.

110010010.1 = 1.100100101 2^8

The result is 1.100100101 2^8, considering that 8 are the positions that the comma moved.

```
110010010.1      1110010010.1      1.100100101*2^8
```

2. Being that 8 is the number of times the comma moved, it is the exponent. Add 8 to 127
   8+127=135

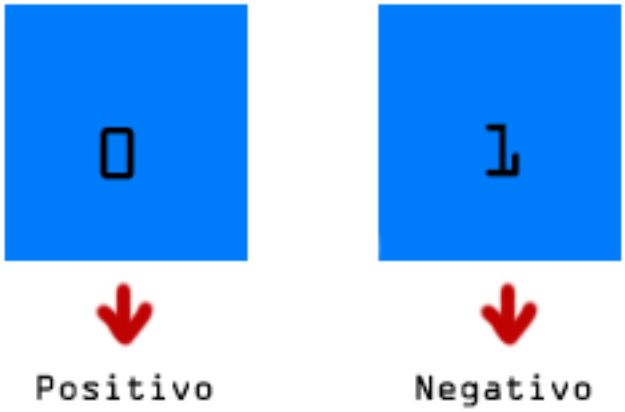
3. The value obtained is converted into binary, so that way we can position the number in place of the exponent (8 bits).

4. 135/2 = 67.5 → 1
   67/2 = 33.5 → 1
   33/2 = 16.5 → 1
   16/2 = 8 → 0
   8/2 = 4 → 0
   4/2 = 2 → 0
   2/2 = 1 → 0
   1/2 = 0.5 → 1
   0/2 = 0

   We read the number from the bottom up (it must have 8 bits)
   135=1110000

5. 

6. Let's not forget that the first bit is for the sign 0=positive, 1=negative; in this case the number is 0 so it's positive.



7. Getting the final result in IEEE 754 floating point format or for value 110010010.1 follows in computer memory of thirty-two (32) bits.