Assignment 1

Yuvateja - EE18BTECH11043

Download all latex-tikz codes from

https://github.com/yuvateja−ctrl/EE4013/blob/main/assignment1.tex

1 Problem

(Q 24) Consider the following representation of a number in IEEE-754 single precision floating point format with a bias of 127
S : 1, E : 10000001 , F : 11110000000000000000000
Here S, E and F denote the sign, exponent and fraction components of the floating point representation. The decimal value corresponding to the representation (rounded to 2 decimal places) is

```
#include<stdio.h>
#include<stdlib.h>
#include<math.h>

int main()
{
    char S[] = "1";
    char E[] = "10000001";
    char M[] = "11110000000000000000000";

    int e = sizeof(E)/sizeof(E[0]);
    int m = sizeof(M)/sizeof(M[0]);

    printf("%d .. %d\n",e,m);

    int exponent = 0;
    for(int i = 0;i<e−1;i++)
    {
        if(E[i] == '1')
        {
            exponent = exponent∗2 + 1;
        }
        else
        {
            exponent = exponent∗2 + 0;
        }
        printf("%d\n",exponent);
    }

    double fraction = 0;
    for(int i = 0;i<m−1;i++)
    {
        if(M[i] == '1')
        {
            fraction += 1∗(pow(2,−i−1));
        }
    }
    int s = (S[0] == '1')?−1:1;
    printf("%lf\n",fraction);
    double ans = s∗(1+fraction)∗(pow(2,exponent
        −127));
    printf("The decimal value corresponding to
        above representation is %lf \n",ans);


    return 0;
}
```

The output of the program upon execution is

2 Solution

Answer : -7.74

Explanation

In IEEE-754 single precision format a floating point number is represented in 32 bits.

1.Sign bit(MSB) - 1 bit

2.biased exponent (E') - 8 bits

3.Normalized mantissa (M) - 23 bits

Sign bit value 0 means positive number, 1 means negative number.

The floating point number can be obtained by formula $\pm 1.M * 2 * *(E - 127)$

Given sign bit is 1 , i.e, the number is negative

Biased exponent (E') is 10000001 = 128 + 1 = 129

Normalized Mantissa (M) is 11110000000000000000000 = 0.937

Therefore the decimal value representation is calculated by above mentioned formula

$\pm 1.M * 2 * *(E - 127) = -1.937 * (2 * *(129 - 127)) = -1.937 * (2 * *2) = -7.748 = -7.75$
(rounded to 2 decimals)