



SNS COLLEGE OF ENGINEERING

Kurumbapalayam (PO), Coimbatore - 641 107

Accredited by NAAC-UGC with 'A' Grade

Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

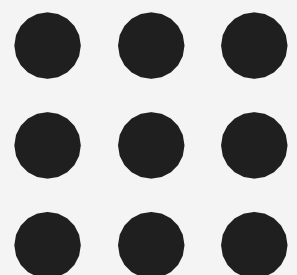
DEPARTMENT OF INFORMATION TECHNOLOGY
SNSCE / IT / III Sem / V Vaishnavee AP-IT

**COURSE NAME: 19IT301 COMPUTER ORGANIZATION
AND ARCHITECTURE**

II YEAR / III SEM

Unit 2 : ARITHMETIC OPERATIONS

**Topic 7: Floating-Point Numbers and
Operations**





Fractions

- If b is a binary vector, it can be interpreted as an unsigned integer by:

$$F(b) = b_{31} \cdot 2^{31} + b_{30} \cdot 2^{30} + b_{n-3} \cdot 2^{29} + \dots + b_1 \cdot 2^1 + b_0 \cdot 2^0$$

- This vector has an implicit binary point to its immediate right:

$$b_{31}b_{30}b_{29}\dots\dots\dots b_1b_0 \quad \text{implicit binary point}$$

- Suppose if the binary vector is interpreted with the implicit binary point is just left of the sign bit:

implicit binary point

$$.b_{31}b_{30}b_{29}\dots\dots\dots b_1b_0$$

The value of b is then given by:

$$F(b) = b_{31} \cdot 2^{-1} + b_{30} \cdot 2^{-2} + b_{29} \cdot 2^{-3} + \dots + b_1 \cdot 2^{-31} + b_0 \cdot 2^{-32}$$





11/18/2023

SNSCE / IT / V Sem / V Vaishnavee AP-IT

