Floating-Point Real Numbers:

Specification:

In FORTRAN → real

In C → float

Some precision required for floating-point numbers, in terms of the numbers of digits used in the decimal representation, may be specified by the programmer, as in Ada.

- Similar arithmetic operations, relational, and assignment operations as with integers are usually provided for real.

- Boolean operations has restrictions

- Equality between two Real no. is rarely achieved due to roundoff issues. Because program that checks for equality to exit a loop may never terminate.

- Some inbuilt functions like

  Sin: real x real → real

  and

  max: real x real → real
Implementation:

- Storage representation based on hardware representation in which a storage location is divided into Mantissa (i.e., significant digit of the no.) and an exponent.

- Any number $N$ can be expressed as $N = m \times 2^k$ form between 0 and 1 and for some integer $k$.

- A double-precision form of floating-point number is also often available, in which an additional memory word is used to store an extended mantissa.