IEE 754

Oscar Lizarraga Pelayo

Hay dos formas de representar números, la representación en punto fijo y la representación de punto flotante este estándar trata sobre punto flotante.

La notación de punto flotante es una forma de expresar numero reales que es parecida a la notación científica estándar. En notación científica, un numero x se expresa como producto de un numero entre 1 y 10 por una potencia adecuada de 10, por ejemplo:





**El estándar 754** establece dos formas básicas para representar a los números reales en la computadora, el de simple precisión(float) y el de doble precisión(double).

Para el de precisión simple(float) para escribir un número real se usan 32 bits, 1 bit para el signo del número 23 bits para la mantisa y 8 bits para el exponente. El exponente en exceso se representa con un desplazamiento de 127

|  |  |  |
| --- | --- | --- |
| 1 bit | 8 bits | 23 bits |
| Signo | Exponente | Mantisa |

De manera similar la representación para la doble precisión(double) el bit más significativo es utilizado para el signo los siguientes 11 bits para el exponente y los últimos 52 bits para la mantisa, el exponte en exceso se representa con un desplazamiento de 1023.

|  |  |  |
| --- | --- | --- |
| 1 bit | 11 bits | 52 bits |
| Signo | Exponente | Mantisa |

Recordando que en ambos casos la mantisa se tiene una mantisa normalizada de la forma



Puesto que todas las mantisas tienen un 1 excepto por el número 0, seguido del punto binario, estos elementos nunca se guardan.

**Rango de valores representables**