Floating-Point Format

How do you represent numbers on a computer, using 32 on/off switches? Floating-point format, and is a map from a sequence of 32 (or 64) bits to real numbers.

*What is the maximum number of numbers you can represent using 32 bits? Using 64 bits?*

Approximation

How can you describe how good an approximation (A) is a number being approximated (B):

Error = B – A

Absolute Error = | B – A |

Relative Error = | B – A | / |A|

An approximation is accurate to n decimal places, if you can trust n digits after the decimal point.

Ex. 3.14 is accurate to 2 decimal places when measuring pi

An approximation is accurate to n significant digits, if you can trust n total digits after the first nonzero digit.

Ex. 3.14 is accurate to 3 significant digits when measuring pi

*If A approximate B to 4 decimal places, does that mean*

1. *|A – B| < 10^4 b) |A-B| < ½ 10^5*

* I don’t think there is a right answer