

## Worksheet: Scientific Notation in Decimal and Binary

You have used scientific notation many times, but it is worth being precise now. It is a scheme for representing real numbers using decimal digits, and it allows you to write very large or very small numbers with good relative accuracy. Recalling how it works will help you understand the binary scheme on the other side of this sheet, which is an example of the real “floating point” systems described in Chapter 5.

**System D.** Suppose “ $\diamond$ ” stands for either + or −, “ $\hat{\square}$ ” stands for a single *positive* decimal digit, and “ $\square$ ” stands for any decimal digit.<sup>1</sup> Consider this (standardized) 3-digit system of scientific notation:

$$\diamond\hat{\square}.\square\square \times 10^{\diamond\square}$$

- (a) What are the largest and smallest numbers you can represent?
- (b) What is the smallest *positive* number you can represent?
- (c) How many *distinct* numbers can be represented?
- (d) How can zero be represented? (*It will have to be a special case, outside of the system.*)

---

<sup>1</sup>That is,  $\hat{\square}$  is from the set  $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$  and  $\square$  is from the set  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ .

**System B16.** On actual computers, real numbers are represented in binary, i.e. base 2. Consider the following scheme which uses 16 positions, each containing 1 bit.<sup>2</sup>

$$\diamond \hat{\square} . \square \square \square \square \square \square \square \square \square \square \square \square \square \square \square \square \times 2^{\diamond \square \square \square \square \square \square \square \square \square \square \square \square \square \square \square \square}$$

Similar to System D on the previous page, “ $\diamond$ ” stands for either + or –, “ $\hat{\square}$ ” stands for a single positive binary digit, and “ $\square$ ” stands for any binary digit.

(a) What are the largest and smallest numbers you can represent?

(b) What is the smallest *positive* number you can represent?

(c) How many *distinct* numbers can be represented?

(d) How could zero be represented? (*You will have to go outside of the system.*)

(e) How is a bit  $\{0, 1\}$  used to represent either + or – in the “ $\diamond$ ” locations?

(f) What is the gap between one and the smallest representable number larger than one? (*It is common to call this number “ $\epsilon$ ”.*)

(g) Can you make the representation slightly more efficient by not “wasting” a bit which never carries information?

---

<sup>2</sup>This system is close to the IEEE 754 “half precision” standard, also called **binary16**. MATLAB, and nearly every other modern computer system, uses IEEE double precision with 64 bits, i.e. **binary64**, also called **double**, but that would be tedious to use here. It appears on homework.