Number Representation With Varying Number of Bits

Anuradha Choudhury, Md Ahsanul Haque, Saeefa Rubaiyet Nowmi, Ahmed Ann Noor Ryen, Sabrina Saika, and Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA
achoudhury@miners.utep.edu, mhaque3@miners.utep.edu, srnowmi@miners.utep.edu
aryen@miners.utep.edu, ssaika@miners.utep.edu, vladik@utep.edu

1. Formulation of the problem

- Computers usually use the same number of bits to store all real numbers: 64 bits, which is 8 bytes.
- This length potentially enables us to represent real numbers with relative precision up to 2^{-64}, which is approximately 10^{-19}.
- In some cases, we do need this precision.
- Sometimes, we even need double precision corresponding to 128 bits.
- However, in many practical situations, we process measurement results that are measured with precision 1% or even less.
- In such cases, we do not need that many bits, so additional bits are simply wasted.
- It is therefore reasonable to consider number representations with varying number of bits.
- Such representations have indeed been proposed.

2. Formulation of the problem (cont-d)

- This leads to several questions.
- The first question is: how much space do we save?
- On the one hand, we need fewer bits to store each number.
- On the other hand, we will need to store, for each number, information about its length – which also takes a few bits.
- The second question is related to the fact that in a computer, bits are usually organized into bytes.
- From this viewpoint:
  - it is easier to design a computer in which real numbers can use 1, 2, etc. bytes
  - than to allow also any number of bits, including the number of bits that does not divide by 8 (and thus, does not constitute bytes).

3. Formulation of the problem (cont-d)

- The bit-representation complication may be worth it, if it allows us to save a significant portion of memory.
- So, the second question is: how much memory do we save if we use bits and not bytes?

4. How we can answer these questions

- Strictly speaking, to answer these questions, we need to know how frequently we encounter numbers of different length.
- At present, this information is not available.
- In such situations, it makes sense to use what is called Laplace Indeterminacy Principle:
  - since we have no reason to believe that some lengths are more frequent than others,
  - it makes sense to assume that all possible lengths are equally frequent.
- Let us use this assumption to answer both questions.

5. Case of byte representation

- In the byte representation, a number can occupy 1, 2, ..., 8 bytes.
- Each of these cases has the same probability 1/8.
- So the average length is equal to

$$\frac{1+2+\ldots+8}{8} = 4.5 \text{ bytes} = 36 \text{ bits}.$$

- In addition, to store information about the length:
  - we need 3 bits,
  - since with 3 bits, we can describe $2^3 = 8$ different values of length.
- So overall, we need $36 + 3 = 39$ bits.
- This is much smaller than the current 64 bits – about 40% smaller.
- So, the answer to the first question is: yes, this is worth pursuing.

6. Case of bit representation

- In the bit representation, a number can occupy 1, 2, ..., 64 bits.
- Each of these cases has the same probability 1/64.
- So the average length is equal to

$$\frac{1+2+\ldots+64}{64} = 32.5 \text{ bits.}$$

- In addition, to store information about the length:
  - we need 6 bits,
  - since with 6 bits, we can describe $2^6 = 64$ different values of length.
- So overall, we need $32.5 + 6 = 38.5$ bits.
- The difference between this average length and 39 bits corresponding to byte representation is very small – about 1%.
- So the answer to the second question is: no, it is probably not worth doing.

7. Reference

- J. F. Gustafson, *The End of Error: Unum Computing*, Chapman & Hall/CRC, Boca Raton, FL, 2015.

8. Acknowledgments

This work was supported in part by:

- National Science Foundation grants 1623190, HRD-1834620, HRD-2034030, and EAR-2225395;
- AT&T Fellowship in Information Technology;
- program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- a grant from the Hungarian National Research, Development and Innovation Office (NRDI).