Mathematical and Logical Foundations of Computer Science
— Summary of Lecture 4 —

Achim Jung
School of Computer Science
University of Birmingham, UK

Autumn 2020

The real numbers

- We think of a real number as represented by its infinite decimal expansion, for example:

$$\pi = 3.14159265358979323846264338327950288419716939 9\ldots$$

- This can be done in any base b, for example, here is π again but represented in its binary expansion:

$$\begin{aligned}\pi \quad = \quad & 11.0010010000111111011010101000100010000101 1\\ & 0100011000010001101001100010011000110011 0\\ & 0010100010111000000011011100000111001101 0\\ & 00100101001000000100100110 1\ldots\end{aligned}$$

- The representation is very good but it also requires some element of identification, though not as much as we needed for $\mathbb{Q}$ or $\mathbb{Z}_m$. That's because

$$4.99999999\ldots \quad \text{is considered equal to} \quad 5.000000\ldots$$

- The set of real numbers is traditionally denoted with $\mathbb{R}$.

Scientific notation

- Each real number is an infinite object and therefore can not be stored in a computer.

- This is not a big problem because most of the time we are satisfied with a finite approximation, for example, 3.14159265 is a pretty good approximation to π.

- These finite approximations are written in scientific notation as in this example

  $$6.02214\,\mathrm{E}\,23$$

  The first part is called the mantissa and the second, the exponent.

- Floating point numbers in a computer (Java's `float` and `double`) are implementations of scientific notation. Apart from the fact that everything is done in base $b = 2$, this is exactly scientific notation.

Arithmetic with floating point numbers

- While the real numbers form a field, all operations with floating numbers incur rounding errors.
- Because of rounding errors, the usual rules of arithmetic may fail for floating point numbers, for example, associativity almost always fails

$$a + (b + c) \neq (a + b) + c$$

- The biggest rounding errors happen when
  - a large number is added to a small number;
  - two numbers of similar size are subtracted from each other.