Arithmetics
Lecture 7: Floating point
reference materials, internal use

Todor Milev

May 12, 2021

Example

An important example is the geometric series

$$a + ar + ar^2 + ar^3 + \cdots + ar^{n-1} + \cdots = \sum_{n=1}^{\infty} ar^{n-1}, \qquad a \neq 0$$

- If $r = 1$, then $s_n = a + a + \cdots + a = na \to \pm\infty$.
- Since $\lim_{n\to\infty} s_n$ doesn't exist, the series is divergent when $r = 1$.
- If $r \neq 1$, then

$$\begin{array}{rclllll} s_n & = & a & +ar & +ar^2 & +\cdots & +ar^{n-1} & \\ -\quad rs_n & = & & ar & +ar^2 & +\cdots & +ar^{n-1} & +ar^n \\ \hline s_n - rs_n & = & a - ar^n & & & & & \\ s_n & = & \frac{a(1-r^n)}{1-r} & & & & & \end{array}$$

- If $-1 < r < 1$, then $r^n \to 0$, so the geometric series is convergent and its sum is $a/(1-r)$.
- If $r > 1$ or $r \leq -1$, then r^n is divergent, so $\sum_{n=1}^{\infty} ar^{n-1}$ diverges.

This theorem summarizes the results of the previous example.

Theorem (Convergence of Geometric Series)

The geometric series

$$\sum_{n=1}^{\infty} ar^{n-1} = a + ar + ar^2 + \cdots$$

is convergent if $|r| < 1$ *and its sum is*

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}.$$

If $|r| \geq 1$*, the series is divergent.*
a *is called the first term and* r *is called the common ratio.*

For $|r| < 1$, recall that the sum of a geometric series is

$$a + ar + ar^2 + ar^3 + \dots = a\left(1 + r + r^2 + r^3 + \dots\right) = \frac{a}{1-r}$$

alternatively

$$\sum_{n=1}^{\infty} ar^{n-1} = \sum_{m=0}^{\infty} ar^m = a\sum_{m=0}^{\infty} r^m = \frac{a}{1-r}$$

Example

Find the sum of the geometric series $\quad -2 + \frac{6}{5} - \frac{18}{25} + \frac{54}{125} - \cdots$

- The first term is $a = -2$.
- The common ratio is $r = \frac{\frac{6}{5}}{-2} = -\frac{3}{5}$.
- Therefore the sum is

$$\sum_{n=1}^{\infty}(-2)\left(-\frac{3}{5}\right)^{n-1} = \frac{(-2)}{1-\left(-\frac{3}{5}\right)} = -\frac{2}{\frac{8}{5}} = -\frac{5}{4}$$

Example

Write the number $2.3\overline{17} = 2.3171717\ldots$ as a quotient of integers.

$$2.3171717\ldots = 2.3 + \frac{17}{10^3} + \frac{17}{10^5} + \frac{17}{10^7} + \cdots$$

- After the first term, we have a geometric series.
- $a = \frac{17}{10^3}$ and $r = \frac{1}{10^2}$.

$$\begin{aligned} 2.3171717\ldots &= 2.3 + \frac{\frac{17}{10^3}}{1 - \frac{1}{10^2}} = 2.3 + \frac{\frac{17}{1000}}{\frac{99}{100}} \\ &= \frac{23}{10} + \frac{17}{990} = \frac{1147}{495} \end{aligned}$$

Algorithm (Write a fraction to periodic base 10)

Input: a fraction $\frac{p}{q}$. Output: integer A and two groups of digits such that $\frac{p}{q} = A + 0.b_1 \dots b_m\overline{c_1 \dots c_n}$.

1 *Initialize* ***digits*** $= ()$ *as the empty sequence.*

2 *Initialize* ***remainders*** $= ()$ *as the empty sequence.*

2 *Divide p by q with remainder r and set A to be the quotient. Append r to remainders.*

3 *While r is not equal to zero:*

- 3.1 *Multiply r by 10.*
- 3.2 *Divide the result by q with remainder r' and quotient d.*
- 3.3 *If r' belongs to* ***remainders*** *with first occurrence at position $m+1$, slice digits into two sequences $b_1, \dots b_m$ and $c_1, \dots c_n$. Return A and $b_1, ..., b_m$, $c_1, \dots, c_n$ as the desired digits.*
- 3.4 *Append d to* ***digits*** *and r to* ***remainders****.*
- 3.5 *Set $r = r'$ and go back to Step 3.*

4 *If r attained the value 0 in the execution of the loop, the fraction $\frac{p}{q}$ has a finite decimal representation given by A and* ***digits****.*

Algorithm (Write a fraction to periodic base X)

Input: a fraction $\frac{p}{q}$. Output: integer A and two groups of digits such that $\frac{p}{q} = A + 0.b_1 \dots b_m\overline{c_1 \dots c_n}$.

1 *Initialize **digits** $= ()$ as the empty sequence.*

2 *Initialize **remainders** $= ()$ as the empty sequence.*

2 *Divide p by q with remainder r and set A to be the quotient. Append r to remainders.*

3 *While r is not equal to zero:*

  3.1 *Multiply r by X.*

  3.2 *Divide the result by q with remainder r' and quotient d.*

  3.3 *If r' belongs to **remainders** with first occurrence at position $m+1$, slice digits into two sequences $b_1, \dots b_m$ and $c_1, \dots c_n$. Return A and $b_1, ..., b_m$, $c_1, \dots, c_n$ as the desired digits.*

  3.4 *Append d to **digits** and r to **remainders**.*

  3.5 *Set $r = r'$ and go back to Step 3.*

4 *If r attained the value 0 in the execution of the loop, the fraction $\frac{p}{q}$ has a finite decimal representation given by A and **digits**.*

Example

Convert $\frac{86}{7}$ to repeating decimal notation.
Divide with 86 by 7 with remainder. Repeat using remainder ...

$$
\begin{array}{rcrcl|rcl}
 & & 86 & = & 12\cdot 7+2 & \frac{86}{7} & = & 12+\dots \\
2\cdot 10 & = & 20 & = & 2\cdot 7+6 & & = & 12.2+\dots \\
6\cdot 10 & = & 60 & = & 8\cdot 7+4 & & = & 12.28+\dots \\
4\cdot 10 & = & 40 & = & 5\cdot 7+5 & & = & 12.285+\dots \\
5\cdot 10 & = & 50 & = & 7\cdot 7+1 & & = & 12.2851+\dots \\
1\cdot 10 & = & 10 & = & 1\cdot 7+3 & & = & 12.28571+\dots \\
3\cdot 10 & = & 30 & = & 4\cdot 7+2 & & = & 12.285714+\dots \\
2\cdot 10 & = & 20 & = & \dots & & = & 12.\overline{285714}+\dots
\end{array}
$$

... until the remainder repeats. Answer: $\frac{86}{7} = 12.\overline{285714}$.

Example

Convert $\frac{2}{13}$ to repeating decimal notation.

$$\begin{array}{rclcl|rcl}
 & & 2 & = & 0\cdot 13+2 & \frac{2}{13} & = & 0+\dots \\
2\cdot 10 & = & 20 & = & 1\cdot 13+7 & & = & 0.1+\dots \\
7\cdot 10 & = & 70 & = & 5\cdot 13+5 & & = & 0.15+\dots \\
5\cdot 10 & = & 50 & = & 3\cdot 13+11 & & = & 0.153+\dots \\
11\cdot 10 & = & 110 & = & 8\cdot 13+6 & & = & 0.1538+\dots \\
6\cdot 10 & = & 60 & = & 4\cdot 13+8 & & = & 0.15384+\dots \\
8\cdot 10 & = & 80 & = & 6\cdot 13+2 & & = & 0.153846+\dots \\
2\cdot 10 & = & 20 & = & 1\cdot 13+7 & & = & 0.1538461+\dots \\
\dots & & & & & & = & 0.\overline{153846}+\dots
\end{array}$$

Answer: $\frac{2}{13} = 0.\overline{153846}$