Abracadabra DB - Math Basics

Abracadabra DB is based on algebraic system. This presentation shows mathematical basics used by TSDB.

Michał Widera

December 10, 2019

Streams Lab
Zabrze, Poland
https://www.linkedin.com/in/michalwidera
https://abracadabradb.com

Algebra

الجبر - *Reunion of broken parts*

In its most general form, algebra is the study of mathematical symbols and the rules for manipulating these symbols.

An **algebraic system** is a set with operations and relations defined on it.

Examples:

1. Elementary
2. Boolean
3. Relation

Number Theory - Samuel Beatty (1881–1970)

Theorem (Beatty - 1926)
If p and q are positive irrationals such that $\frac{1}{p} + \frac{1}{q} = 1$ then the sequences $\{\lfloor np \rfloor\}_{n=1}^{\infty} = \lfloor p \rfloor, \lfloor 2p \rfloor, \lfloor 3p \rfloor, \ldots$ and $\{\lfloor nq \rfloor\}_{n=1}^{\infty} = \lfloor q \rfloor, \lfloor 2q \rfloor, \lfloor 3q \rfloor, \ldots$ constitute a partition of the set of positive integers.

The research in literature is based on the following sequence called Beatty sequence (floor - 1).

$$\mathcal{B}(\alpha, \alpha') := \left(\left\lfloor \frac{n - \alpha'}{\alpha} \right\rfloor \right)_{n=1}^{\infty} \tag{1}$$

or Beatty sequence (ceiling - 2):

$$\mathcal{B}^{(c)}(\alpha, \alpha') := \left(\left\lceil \frac{n - \alpha'}{\alpha} \right\rceil \right)_{n=1}^{\infty} \tag{2}$$

S. Beatty. Problem 3173. Amer. Math. Monthly, 33:159, 1926.

Number Theory - Aviezri Siegmund Fraenkel (1929)

Theorem (Fraenkel - 1969)
The sequences $\mathcal{B}(\alpha, \alpha')$ *oraz* $\mathcal{B}(\beta, \beta')$ *partition* $\mathbb{N}$ *if and only if the following five conditions are satisfied.:*

1. $0 < \alpha < 1$.
2. $\alpha + \beta = 1$.
3. $0 \leq \alpha + \alpha' \leq 1$.
4. *If* α *is irrational, then* $\alpha' + \beta' = 0$ *i* $k\alpha + \alpha' \notin \mathbb{Z}$ *for* $2 \leq k \in \mathbb{N}$.
5. *If* α *is rational, then (say* $q \in \mathbb{N}$ *is minimal with* $q\alpha \in \mathbb{N}$ *) then* $\frac{1}{q} \leq \alpha + \alpha'$ *and* $\lceil q\alpha' \rceil + \lceil q\beta' \rceil = 1$.

A. S. Fraenkel. The bracket function and complementary sets of integers. Canad. J.Math, 21:6–27, 1969.

Theorem (Widera - 2006)
The interlace operation assures a sequential cover of both sets containing elements of data streams.

$$\begin{aligned} \Delta_C &= \frac{\Delta_A \Delta_B}{\Delta_A + \Delta_B} \\ c_n &= \begin{cases} b_{n-\lfloor nz \rfloor} & \lfloor nz \rfloor = \lfloor (n+1)\, z \rfloor \\ a_{\lfloor nz \rfloor} & \lfloor nz \rfloor \neq \lfloor (n+1)\, z \rfloor \end{cases} , z = \frac{\Delta_B}{\Delta_A + \Delta_B} \end{aligned} \tag{3}$$

M.Widera. Deterministic method of data sequence processing. Annales Universitatis Mariae Curie-Skłodowska, Sectio AI, Informatica, Vol. 4 (2006), pages 314-331

Theorem (Widera - 2006)
The deinterlace operation is an instance of Fraenkel Partition Theorem.

$$a_n = c_{n+\left\lceil \frac{(n+1)\Delta_A}{\Delta_B} \right\rceil}, \;\; \Delta_A = \frac{\Delta_C \Delta_B}{|\Delta_C - \Delta_B|} \tag{4}$$

residue:

$$b_n = c_{n+\left\lfloor \frac{n\Delta_B}{\Delta_A} \right\rfloor}, \;\; \Delta_B = \frac{\Delta_C \Delta_A}{|\Delta_C - \Delta_A|} \tag{5}$$

M.Widera. Deterministic method of data sequence processing. Annales Universitatis Mariae Curie-Skłodowska, Sectio AI, Informatica, Vol. 4 (2006), pages 314-331

Sum

$$\Delta_C = \min\left(\Delta_A, \Delta_B\right)$$
$$c_n = \begin{cases} a_n | b_{\left\lfloor \frac{n\Delta_A}{\Delta_B} \right\rfloor} & \Delta_C == \Delta_A \\ a_{\left\lfloor \frac{n\Delta_B}{\Delta_A} \right\rfloor} | b_n & \Delta_C == \Delta_B \end{cases} \tag{6}$$

and Difference

$$a_n = \begin{cases} c_n & \Delta_B \geqslant \Delta_A \\ c_{\left\lceil \frac{n\Delta_A}{\Delta_B} \right\rceil} & \Delta_B < \Delta_A \end{cases} \tag{7}$$

M.Widera. Deterministic method of data sequence processing. Annales Universitatis Mariae Curie-Skłodowska, Sectio AI, Informatica, Vol. 4 (2006), pages 314-331

Proposed Time Series Algebra

Data model:

$S ::= (s_n, \Delta)$ where $\Delta \in \mathbb{Z} > 0$ is a time dimension, s_n is a set of measurments indexed by n.

Operations:

1. Projection
2. Sum and Difference
3. Interlace and Deinterlace
4. Agregation and Serialization
5. Shift

End

Thank you for your attention.

Streams Lab
Michał Widera