Towards Non-Archimedean Artificial Intelligence

teaching material

Lorenzo Fiaschi and Marco Cococcioni

Course: Symbolic and Evolutionary Artificial Intelligence
University of Pisa

a.a. 2020-2021

What is a natural number?

- Modern mathematics is reductionist
- Everything is founded on set theory
- Numbers are classes of equivalent sets

$$0 = \emptyset, \qquad 1 = \{\emptyset\} = \{0\}, \qquad 2 = \{\emptyset, \{\emptyset\}\} = \{0, 1\}, \qquad \ldots$$

- Set operations defines arithmetical operations
- Let $\mathfrak{n}(\cdot)$ indicate the number of elements in a set

$$\begin{aligned} 3 + 2 &= \mathfrak{n}(\{A, B, C\}) + \mathfrak{n}(\{e, f\}) = \mathfrak{n}(\{A, B, C\} \cup \{e, f\}) = \\ &= \mathfrak{n}(\{A, B, C, e, f\}) = 5 \end{aligned}$$

$$\begin{aligned} 3 \cdot 2 &= \mathfrak{n}(\{A, B, C\}) \cdot \mathfrak{n}(\{e, f\}) = \mathfrak{n}(\{A, B, C\} \times \{e, f\}) = \\ &= \mathfrak{n}(\{(A, e), (A, f), (B, e), (B, f), (C, e), (C, f)\}) = 6 \end{aligned}$$

Facts on (finite) sets

Axiom (Hume)

Let A and B be two sets. If there exists a bijection $\Phi \colon A \to B$, then A and B have the same number of elements, i.e., $\mathfrak{n}(A) = \mathfrak{n}(B)$.

Axiom (Euclides - V)

Let A and B be two sets such that $A \subset B$. Then A contains less elements than B, i.e., $\mathfrak{n}(A) < \mathfrak{n}(B)$.

- The number of axioms in this slide is equal to the number of eyes on a human face (Hume)
- The number of characters in both the axioms is bigger than the ones in the first axiom only (Euclides)

To the infinity and beyond

- $\mathbb{N} = \{1, 2, 3, \ldots\}$ is a set
- So it must represent a number
- Which one? For sure an infinite one!
- The proof comes from Archimedes' Axiom
- How many infinite number exist?
- Does infinitesimal numbers exist?
- Do they behave like real numbers?
- Scientists and philosopher wondered a lot about these questions
- This debate lasts since a while: about 2500 years

- Ancient Greeks: Pythagoras, Archimedes, Euclides
    - Everything is a number
    - Everything can be counted
    - The part is smaller than the whole

Axiom (Archimedes)

Let $\mathcal{U}$ be any ordered set. Then:

$$\forall\, x, y \in \mathcal{U},\ |x| < |y|\, (x, y \Rightarrow \exists\, n \in \mathbb{N} : |nx| > |y|)$$

- Modern mathematicians and the infinity paradoxes: Galileo
    - The set of even numbers contains as many elements as $\mathbb{N}$?
    - The set of perfect squares contains as many elements as $\mathbb{N}$?

- Hume and Euclides cannot hold together with infinite sets

XVII century: the fathers of differential calculus

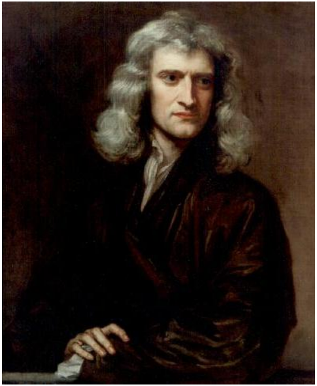

Figure: Isaac Newton (1642-1726)

Figure: Gottfried Leibniz (1646-1716)

- Harsh debate (1666-1684)
- Application to mechanics
- Infinitesimal variations of time

- Differential and integral
- Infinitesimal pieces of curves
- More prolific seminal work

Early adopters, Math growth and first critics

- Early adopters: Euler, Bernoulli brothers, Maria Agnesi
- Mathematics grew a lot for roughly 150 years
  - Prediction of celestial bodies and cannon balls trajectories
  - Computation of areas and volumes
  - Optimization of multivariate functions
- Critics from a philosophical point of view
- Infinitesimal numbers were not rigorously defined

XIX century: the rationalization of Math

- In XIX century Mathematics started to be rationalized
- Infinitesimals faced a technical problem
- In 1821 Cauchy introduces the concept of limit
- Weierstrass formalized it at the end of the century without the use of infinitesimals
- It needed the existence of one infinite set: $\mathbb{N}$
- Rigorous researches on infinitesimals did not stop yet

XX century: the ideological war

- The debate upon infinitesimal rigorous definition became harsh

- Non-Archimedean alliance:
  - Du Bois-Reymond
  - De Morgan
  - Peirce
  - Veronese
  - Tullio Levi-Civita
  - Enriques
- Archimedean horde:
  - Bolzano
  - Cantor
  - Frege
  - Peano
  - Dedekind
  - Russel

"Abbiamo trovato che il calcolo differenziale e integrale non hanno bisogno dell'infinitesimale, e che [...] la forma più usuale [...] non è implicata [...] e risulta autocontraddittoria (B. Russell, 1901)"

"Infinitesimals as explaining continuity must be regarded as unnecessary, erroneous, and self-contradictory (B. Russell, The principles of Mathematics, 1903)"

Gödel, Robinson and a new era for infinitesimals

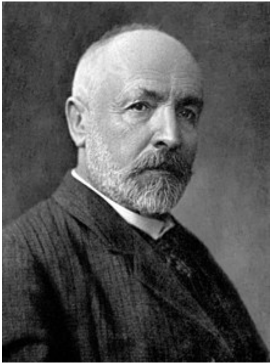

Figure: Georg Cantor (1845-1918)

- Axiomatization of $\mathbb{R}$ does not characterize a unique field
- They are infinite and the only Archimedean one is $\mathbb{R}$

Figure: A. Robinson (1918-1974)

- 1960/66 "Non-Standard Analysis"
- Non-Standard numbers are special non-Archimedean ones

Non-Standard analysis nowadays

Figure: B. De Finetti (1906-1985)

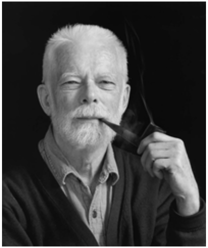

Figure: Edward Nelson (1932-2014)

Figure: Howard J. Keisler (1936 -)

Figure: Mauro Di Nasso (? -)

Non-Archimedean approaches: honorable mentions

Figure: John Conway (1937 - 2020)

- Surreal numbers
- Non-Archimedean field
- Sequential games are surreal numbers and vice versa
- Any surreal number is also the solution of that game

Figure: Yaroslav D. Sergeyev (1963 -)

- Grossone Methodology, ①
- Pioneer in numerical non-Archimedean research

Figure: Vieri Benci (1949 -)

- Inventor of Alpha Theory (with Di Nasso)
- Professor at Academia Nazionale dei Lincei
- Commendatore al Merito della Repubblica Italiana
- Ordine del Cherubino

Alpha Theory: an axiomatization

- Axiomatization is not the only way: constructive approaches exist.
- Axiomatization of a theory is never unique.
- Which is better? It depends, as always.
- Conciseness: just three axioms.
- Simplicity: logic-free.
- Application-oriented: easy to standardize an embedding.

First axiom

Axiom 1 (Existence)

Exists an ordered field $\mathbb{E} \supset \mathbb{R}$ whose numbers are called Euclidean numbers.

Field (informal)

A field $\mathbb{F}$ is a set equipped with two binary operations, namely sum and multiplication, which behave as they behave on rational ($\mathbb{Q}$) and real ($\mathbb{R}$) numbers, including the existence of their inverse operations (subtraction and division).

Definition (Field)

$(\mathbb{F}, +, \cdot)$ is a field $\overset{def}{\iff}$ $(\mathbb{F}, +)$ is an Abelian group with 0 as identity; $(\mathbb{F} \setminus \{0\}, \cdot)$ is an Abelian group with 1 as identity; and $\cdot$ distributes over $+$.

Axiom 2 (Numerosity α)

Exists a function $\mathfrak{num}$, $\mathfrak{num} : \mathcal{U} \to \mathbb{E}$ *which satisfies*

- $\alpha = \mathfrak{num}(\mathbb{N})$
- $\mathfrak{num}(A \cup B) = \mathfrak{num}(A) + \mathfrak{num}(B) - \mathfrak{num}(A \cap B)$
- $\mathfrak{num}(A \times B) = \mathfrak{num}(A) \cdot \mathfrak{num}(B)$

Comments:

- $\alpha \in \mathbb{E}$
- α is infinite
- $\eta = \alpha^{-1} \in \mathbb{E}$ (existence of inverse)
- η is infinitesimal
- $\mathbb{E}$ contains algebraic manipulations of α, e.g., $\frac{\alpha^7 - \alpha^\pi + \eta^\alpha}{-3 + 5\eta^3}$

Definition (Infinite number)

$x \in \mathbb{E}$ *is infinite* $\overset{def}{\Longleftrightarrow} \nexists\, k \in \mathbb{N}$ *such that* $k > |x|$.

Definition (Finite number)

$x \in \mathbb{E}$ *is finite* $\overset{def}{\Longleftrightarrow} \exists\, k \in \mathbb{N}$ *such that* $k > |x| > \frac{1}{k}$.

Definition (Infinitesimal number)

$x \in \mathbb{E}$ *is infinitesimal* $\overset{def}{\Longleftrightarrow} \nexists\, k \in \mathbb{N}$ *such that* $|x| > \frac{1}{k}$.

Axiom 3 (Transfer principle - weak form)

Given a real function φ, $\exists!\,\varphi^*$ defined over $\mathbb{E}$ such that

- $\varphi(x) = \varphi^*(x)\ \forall\, x \in \mathbb{R}$
- $Id_{\mathbb{R}}^* = Id_{\mathbb{E}}$, *where* Id_A *is the identity function on* A
- *Any couple of real functions* φ, ψ *satisfies:*
  - $(\varphi + \psi)^* = \varphi^* + \psi^*$
  - $(\varphi \cdot \psi)^* = \varphi^* \cdot \psi^*$
  - $(\varphi \circ \psi)^* = \varphi^* \circ \psi^*$

Comments:

- Any function φ on $\mathbb{R}$ can be extended to $\mathbb{E}$
- φ^* satisfies all the properties of φ (continuity, differentiability, etc.).

Some examples and some questions

- The following relations hold on $\mathbb{E}$:

$$\alpha \cdot (\alpha + 2) = \alpha^2 + 2\alpha \qquad 0 < \frac{1}{\alpha} = \alpha^{-1} < \alpha^0 = 1 < \alpha^1 = \alpha < (\alpha + 1)$$

$$\frac{-10.0\alpha^2 + 16.0 + 42.0\eta^2}{5.0\alpha^2 + 7.0} = -2.0 + 6.0\eta^2$$

- Given $\varphi(x) = x^2$, $x \in \mathbb{R}$, $\exists \varphi^*(y) = y^2$, $y \in \mathbb{E}$, such that φ^* is continuous, quadratic, with minimum in 0, differentiable, its differential is linear, etc.
- Is everything that easy? Can you spot a case in which the Transfer Principle fails?
- Is $\mathbb{E}$ unique?
- $\exists x \in \mathbb{E}$ such that $x \neq \varphi^*(\alpha)\ \forall \varphi_{\mathbb{R}^{\mathbb{R}}}$?

Definition (Completeness)

X is complete $\overset{def}{\iff}$ every nonempty subset of X having an upper bound must have a least upper bound (or supremum) in X.

- Transfer Principle does not transfer $\mathbb{R}$ completeness. An example is the monad of zero, i.e., $\mu(0) := \{x \in \mathbb{E} \mid x \approx 0\}$.
- There are many other cases, e.g., linear systems solutions magnitude (we shall see it in the future).
- $\mathbb{E}$ is not unique at all. Further assumptions are needed, e.g., the absolute continuum axiom.
- It depends on how “poor” you choose $\mathbb{E}$. For our case of study the answer is yes.

Transfer Principle and Internal Sets

Theorem (Transfer principle)

For any object $\mathcal{A}$ defined over $\mathbb{R}$ $\exists! \mathcal{A}^*$ over $\mathbb{E}$ which satisfies all the first order properties of $\mathcal{A}$. Moreover, $\mathcal{A} \subseteq \mathcal{A}^*$ and $\mathcal{A} \equiv \mathcal{A}^* \iff \mathcal{A}$ is finite.

Definition (Internal set - informal)

Let $\mathcal{A}^*$ be a set defined over $\mathbb{E}$. Then, $\mathcal{A}^*$ is an internal set $\overset{def}{\iff} \exists \mathcal{A}$ over $\mathbb{R}$ for which the Transfer Principle holds true.

Examples:

- The set of even numbers on $\mathbb{E}$ is internal.
- The set of Euclidean quadratic functions is internal.
- The set of numbers not infinitely bigger than zero is external.
- The class of sets which contain one infinite number is external.

Is α even or odd? And some other similar questions

One may wonder:

- Is α even?
- More generally, $\frac{\alpha}{n} \in \mathbb{N}^* \; \forall n \in \mathbb{N}$?
- Similarly: $\sqrt{\alpha} \in \mathbb{N}^*$? $\sqrt[n]{\alpha} \in \mathbb{N}^* \; \forall n \in \mathbb{N}$?

Answers:

- It is up to you provided the coherency of the properties of α
- $\frac{\alpha}{2} \in \mathbb{N}^* \quad \wedge \quad \sqrt{\alpha} \in \mathbb{N}^*$ is fine
- $\frac{\alpha}{2} \in \mathbb{N}^* \quad \wedge \quad \alpha$ prime is not
- $\alpha \leq 0$ is not too
- Neither $\sin(\alpha) = 0$

Algorithmic fields

- $\mathbb{R}$ reference set for theory about data science, machine learning, market analysis, etc.
- In practice, $\mathbb{R}$ is "too large" to fit in a computer, e.g., $\frac{1}{3}$ which is a repeating decimal
- Numbers representation is not always *exact*, field properties do not hold anymore as well
- Common computers work with floating point numbers (IEEE 754 standard), which are a finite-dimension encoding
- *Algorithmic Field*: set of all the numbers which can be represented exactly within a machine
- It may contain symbols not present in the original field it approximates: `NaN`, $\pm$`Inf`.

An example, lack of associativity

With finite precision it happens that $(-2^{127} + 2^{127}) + 1 = 1$ and $-2^{127} + (2^{127} + 1) = 0$, as reported in the following Matlab code:

```
>> m1 = single(-2^127);
>> m2 = single(2^127);
>> m3 = single(1);
```

```
>> (m1 + m2) + m3
>> ans = 1
>> m1 + (m2 + m3)
>> ans = 0
```

"One of the most surprising aspects of algorithmic fields is that even if they are not closed with respect to any algebraic operations, even if they are discrete sets, and even if computations doing over them may suffer by numerical instabilities, they turn out to work extremely well for the vast majority of practical problems"

The importance of fixed length

- Fixed-length vs. Symbolic computations, e.g., Matlab vs. Wolfram Mathematica
- Symbolic:
  - Numbers variable-size typically slows a lot the computations
  - “Infinite precision”
  - Iterative schemes makes numbers inner representation grow at each iteration
  - Programs tend to get slower and slower during time
- Fixed-length:
  - Standardized operations, e.g., CPU instructions
  - Faster code
  - Deterministic time consumption
  - Possibility to build hardware accelerators, e.g., FPUs
  - Noisy operations which may notably affect the results

Algorithmic numbers (ANs)

Definition (Algorithmic Number)

$\xi \in \mathbb{E}$ is an Algorithmic Number $\overset{def}{\Longleftrightarrow} \xi = \sum_{k=0}^{\ell} r_k \alpha^{s_k}$ where $r_k \in \mathbb{R}$, $s_k \in \mathbb{Q}$, $s_k > s_{k+1}$. its normal form is $\xi = \alpha^p P\left(\eta^{\frac{1}{m}}\right)$ where $p \in \mathbb{Q}$, $m \in \mathbb{N}$ and $P(x)$ is a polynomial with real coefficients such that $P(0) = r_0 \neq 0$.

- Subset of $\mathbb{E}$ which can be better manipulated by computers
- Two main issues:
  - Not close with respect to inversion, e.g., $(\alpha + 1)^{-1}$
  - Variable length
- Both solvable by truncation

Truncation and Bounded ANs (BANs)

Definition (Truncation)

Let $P(x) = p_0 x^{z_0} + \ldots + p_m x^{z_m}$, $z_{i-1} < z_i$, $i = 1, \ldots, m$. *Then*

$$\mathfrak{tr}_n \left[P\left(x \right) \right] := \begin{cases} P(x) & n \geq m \\ p_0 x^{z_0} + \ldots + p_n x^{z_n} & n < m \end{cases}$$

Definition (BAN)

$\xi = \alpha^p P\left(\eta^{\frac{1}{m}}\right)$, $\xi \in \mathbb{E}$ *is a BAN* $\overset{def}{\iff}$ ξ *is an AN and* $p \in \mathbb{Z}$.

- High modelling power
- Substantial informative value
- Light and easy-to-manage embedding

An application: Markov chains

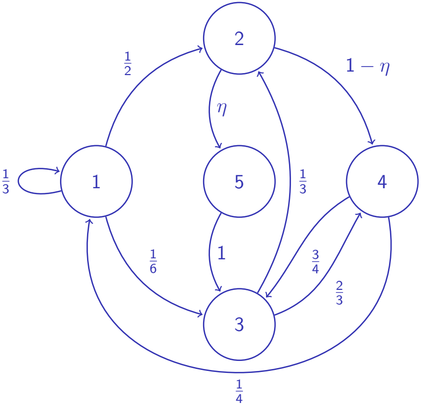

Figure: Markov Chain where 5 is a quasi-unreachable state

$$\pi = \begin{bmatrix} \frac{18-6\eta}{127+9\eta} \\ \frac{22}{127+9\eta} \\ \frac{39+9\eta}{127+9\eta} \\ \frac{48-16\eta}{127+9\eta} \\ \frac{22\eta}{127+9\eta} \end{bmatrix}$$

$$\pi = \begin{bmatrix} 0.142 - 0.057\eta + 0.004\eta^2 \\ 0.173 - 0.012\eta + 0.001\eta^2 \\ 0.307 - 0.049\eta + 0.003\eta^2 \\ 0.378 - 0.153\eta + 0.011\eta^2 \\ 0.173\eta + 0.012\eta^2 \end{bmatrix}$$

An application: Markov chains

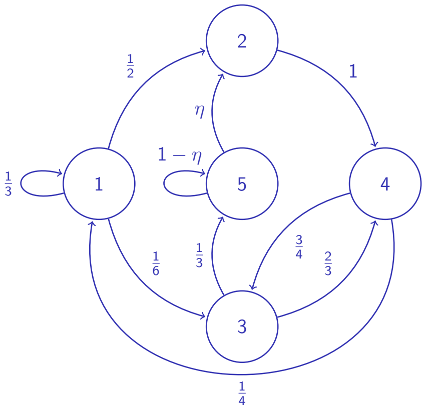

Figure: Markov Chain where 5 is a quasi-absorbing state

$$\pi = \begin{bmatrix} \frac{18\eta}{13+127\eta} \\ \frac{22\eta}{13+127\eta} \\ \frac{39\eta}{13+127\eta} \\ \frac{48\eta}{13+127\eta} \\ \frac{13}{13+127\eta} \end{bmatrix}$$

$$\pi = \begin{bmatrix} 1.385\eta - 13.527\eta^2 \\ 1.692\eta - 16.533\eta^2 \\ 3\eta - 29.308\eta^2 \\ 3.692\eta - 36.071\eta^2 \\ 1 - 9.769\eta + 95.438\eta^2 \end{bmatrix}$$

An application: Markov chains

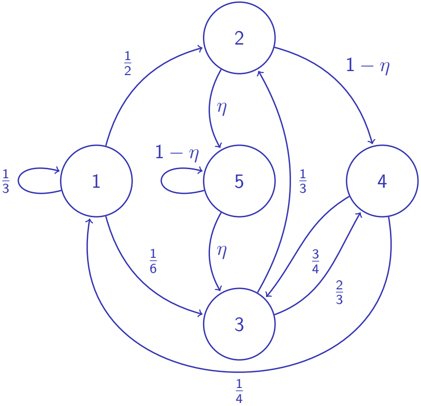

Figure: Markov Chain where 5 is both a quasi-unreachable state and a quasi-absorbing one

$$\pi = \begin{bmatrix} \frac{18-6\eta}{149-13\eta} \\ \frac{22}{149-13\eta} \\ \frac{39+9\eta}{149-13\eta} \\ \frac{48-16\eta}{149-13\eta} \\ \frac{22}{149-13\eta} \end{bmatrix}$$

$$\pi = \begin{bmatrix} 0.121 - 0.03\eta - 0.003\eta^2 \\ 0.148 + 0.013\eta + 0.001\eta^2 \\ 0.262 + 0.083\eta + 0.007\eta^2 \\ 0.322 - 0.079\eta - 0.007\eta^2 \\ 0.148 + 0.013\eta - 0.001\eta^2 \end{bmatrix}$$