Antal Jakovác

Representing Reality

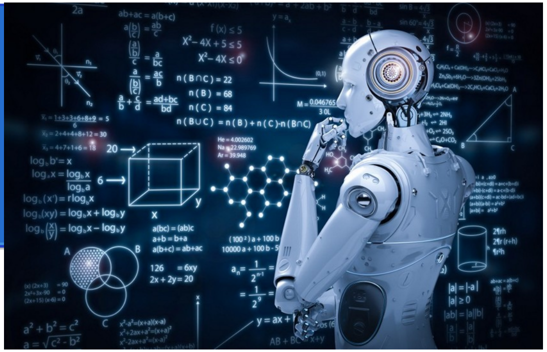
AIME 2025
Budapest, Hungary, November 27-28, 2025

Introduction: successes of AI

Neural networks → magnificent achievements

- classification (dog breeds, faces, birdsong, flowers, etc.)
- text generation (chatGPT, Bing, DeepSeek, etc.)
- image generation (midjourney, Dall-E, Dreamstudio, etc.)
- autonomous cars / AI driving assistants
- etc.
- Billion USD business (2023: ~ 200 bUSD)

Introduction: AI challenges



Challenges and limitations

- “I don’t know” category (balanced datasets are needed for training)
- generalization (“catastrophic forgetting”)
- prone to adversarial attacks
- no goals, planning
- hallucinations in generative networks

At the root of many AI-related problems lies an insufficient representation of reality.

Reality and AI Models

Aspects of reality: general and actual

- **general reality**
  - laws and generic algorithms describing how the world works
- examples
  - a map describing the streets → show the possibilities
  - in LLMs: possible tokens that can come from a context
- in AI: part of the (trained) hardware → represented by millions to ~100billions of parameters

Reality and AI Models

Aspects of reality: general and actual

- **actual reality**:
  - describes the actual environment, the objects, persons and phenomena that are actually active
- examples
  - in the map the actual traffic, road works, detours, obstacles, etc.
  - in LLMs: the actual question we want an answer to, prompt
- in AI: part of the input → ~million tokens

In AI applications, general reality is represented much more thoroughly.

Introduction: target accuracy

How can we shoot better?

- take a gun, shooting to a target: it is not accurate
- improve accuracy → general level
  - improve weapon (e.g. attach a scope)
  - change for better weapon (e.g. handgun → rifle)
- improve accuracy → actual level
  - better aiming
  - steadier hold

Introduction: target accuracy

How can we shoot better?

- take a gun, shooting to a target: it is not accurate
- improved accuracy: better tool, better aiming
  → still there is a distance where precision is lost
- **initial conditions (aiming) is a pure representation of reality**
- true for any systems: initial conditions & propagation → drift
  - physical systems → like mechanical projectiles
  - iterative computations → like generative neural networks

Introduction: target accuracy

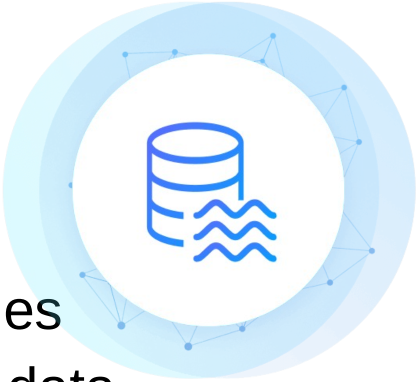
Maintain precision

- for long-term precision a control mechanism is necessary: compare modeled and actual reality
- weapons → frequent course corrections
- **in general**:
  - apply generation in small steps
  - compare state to the actual reality
  - apply correction before continuing generation
- in this way we can follow the plan, without drifting away from reality

Representing Reality

How can we characterize reality? → ask questions (measurements)
(A. Jakovac, A. Telcs, Mathematics, 2025)

what is a good question?



- **acquire information → data lakes**
  - big data approach, measure what we can
    → e.g. measure pixel colors, collect logs, store amplitudes
  - hope that all important detail can be recovered from the data
  - important details are deeply hidden in the data
  - often leads to an enormous, unstructured database with a lot of unnecessary/redundant information

Representing Reality

what is a good question?

- **useful information (feature) → characterize a phenomenon**
  - phenomena are equivalence classes of elementary data
  - equivalence classes are defined by labels (supervised) or by other means (e.g. temporal subsequence – unsupervised)
  - to represent equivalence classes we choose functions constant on them *(→ measurable functions)*

$$f(x)=f_C \quad \text{if} \quad x \in C \quad \Rightarrow \quad f(C)$$

  - features depend on the selected phenomena (context)
    (and any context have stable patterns)

Representing Reality

Concepts as class-constant functions

- **objects:** we observe solid states with definite shape → remains constant while changing place, time, rotate, etc. → in a gaseous environment, ”object” concept would be useless
- **angle:** consider two lines, observers see it in different position, rotation; the invariant property is the angle
- **laws:** persistent relation between measurable quantities:
  - Newton law: $ma - F = 0$ is true for all times
  - market model: price distribution $P(S, \dot{S}, \ldots)$

Representing Reality

Elements of a good representation

context: relevant objects – measurable functions – values

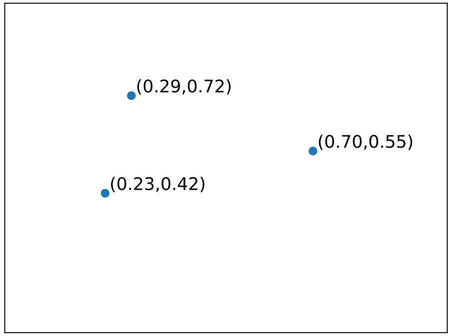
another (natural science vibe) wording:

context: objects – coordinates – values

- everything is context-dependent (no “general” reality)
- coordination can be general or object dependent
- world basically consists of discrete objects, but the raw input is often continuous

General coordinates

Examples for object independent coordination



- mathematics: points in a d-dimensional space
- we not have to name the coordinates → $p:(x_1, x_2, \ldots, x_d)$
- word embedding in LLMs
  - one-hot-encoding: all words represent a new dimension
  - vector embedding: ~thousand dimensional space
  - all coordinates represent a (fixed, abstract) feature, and we store the relation of the word to the given feature

Object dependent coordinates

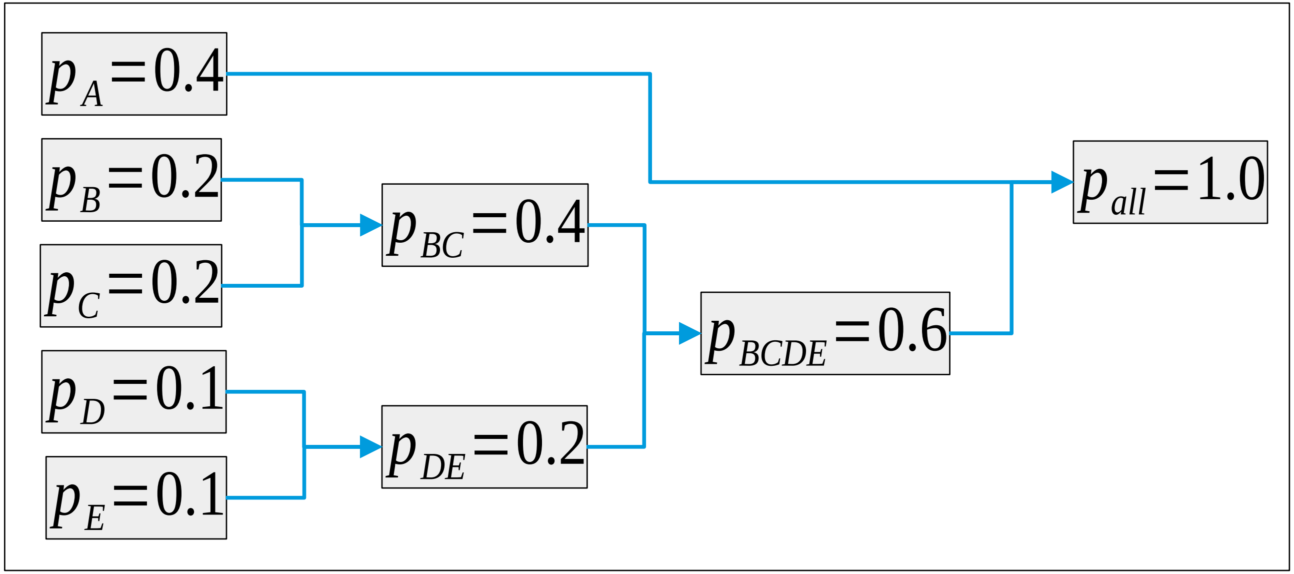
Optimal coordination

- assume we have discrete classes (objects) characterized by class-constant functions
- which coordination is the optimal?
  - we arrive at the given class with a minimal number of features
  - information theoretical question (Shannon entropy, Huffmann coding)
- optimal coding is **hierarchical**: properties of the parents are valid properties of the children → **generalization**
- usually much less property/feature is enough than in the general case → *c.f. twenty questions (barchoba)*

Object dependent coordinates

Shannon entropy and amount of information

- example for Huffmann coding
- average questions:
  0.4+3x0.6=2.2
- Shannon entropy:
  2.12



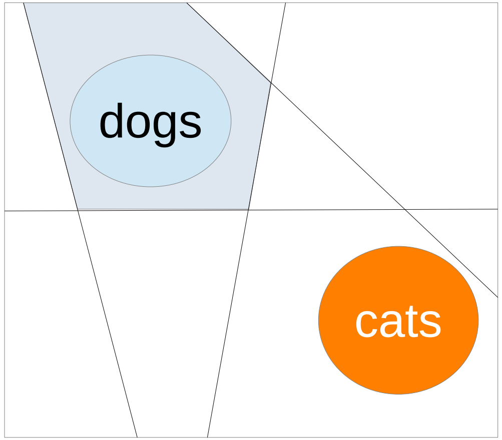
Continuous inputs

Sensory inputs are (quasi) continuous → discrete features

- we have to define the context (labels, or temporal subsequence)
- we have to define a parametrizable functional space (e.g. DNN, LLT, …)
- optimize parameters to find class-constant or conserved (time-constant) functions → necessary conditions
- collect conditions to fully represent a class

$$\text{object} : \left(\text{feature}_1 = x_1,\ \text{feature}_2 = x_2, \ldots\right)$$

- values usually discrete (discrete classes)
  → *for regression they can be continuous*



Coordination and objects

Technical solutions for representing reality

- time series analysis: LLT/ALT method
  (Kurbucz et al. arXiv 2025, Scientific Reports 2022)
- reusable computer vision applications (RESNET): last layer provides good features for natural images
- databases and Retrieval Augmented Generation (RAG)
  - relational databases → fixed coordinates
  - nonrelational databases → document databases (JSON)
  - graph databases → Resource Description Framework

Coordination and objects

Technical solutions for coordination

- time series analysis: LLT/ALT method
  (Kurbucz et al. arXiv 2025, Scientific Reports 2022)
  - linear laws for time segments
  - collection of laws for a given class with ‘OR’ relation
  - if rules do not fit to either class → “I don’t know” category (outlier)
  - adaptive version to find best scales

LLT/ALT classification results

Mathematical procedure:

- new sample arrives
  - try the laws of both classes
  - keep the bests (top 5%)
  - predicted class that performs the best
    use standard classifiers (SVM, KNN)
  - if no laws work: **none of the classes** → outlier analysis (e.g. ECG signal analysis)
- benchmarking: publicly available databases
- applications: mechanical motions, ECG signal processing, Bitcoin price prediction, etc.

Dataset	Accuracy	Training time	Benchmark
Ford A	97.5%	916 sec	97-98%
Ford B	94.3%	3070 sec	83-92%
AReM	100%	10 sec	99.6%
Gun_Point	96.7%	8 sec	100%

Resource Description Framework

Triplet representation of reality

- knowledge graph representation
- each entry is a triplet:
  (subject, predicate, object) → (name, property, value)
- example: Dorothy : son = Bill
- own development: TriosDB → lightweight RDF solution
  - values are vectors
  - circular representation for efficient queries
    Bill : Dorothy = son, son : Bill = Dorothy
  - context as data modules

Conclusion

Representation of reality

- correct representation of reality is the main goal of the intelligence
- general vs actual reality
  - general reality is well represented in AI solutions
  - actual reality is poorly and ineffectively represented
- key element: find class-constant functions
  - continuous input → discrete representation (e.g. with LLT)
  - discrete objects → knowledge graphs, hierarchical structure

The end