Antal Jakovác

A Novel Approach to Artificial Intelligence

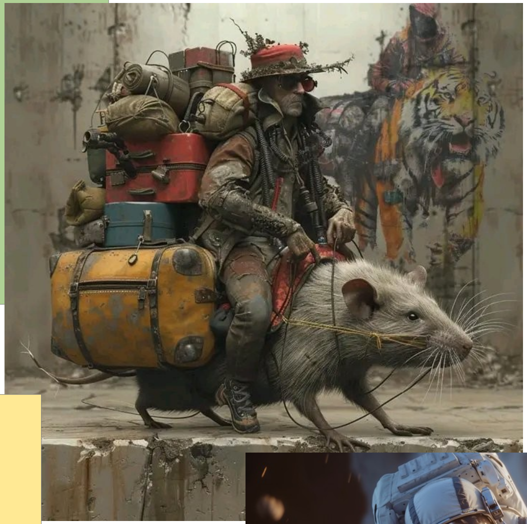
MT Kurbucz
P Pósfay
A Telcs
TS Biró

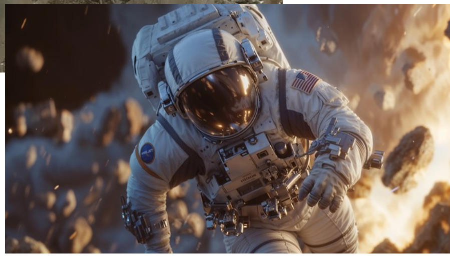
AIME 2024
Budapest, Hungary, November 21-22, 2024

Introduction: successes of AI

Neural networks provided a lot of magnificent achievements

- classification (dog breeds, faces, birdsong, flowers, etc.)
- text generation (chatGPT, Bing, Bard, Vicuna23B, etc.)
- image generation (midjourney, Dall-E, Dreamstudio, etc.)
- autonomous cars / AI driving assistants
- etc.



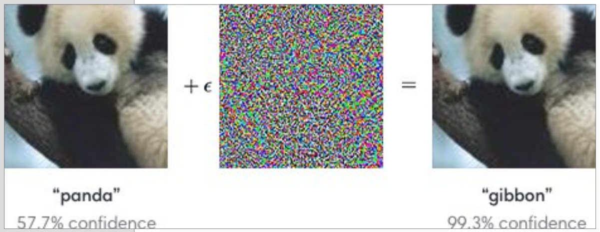


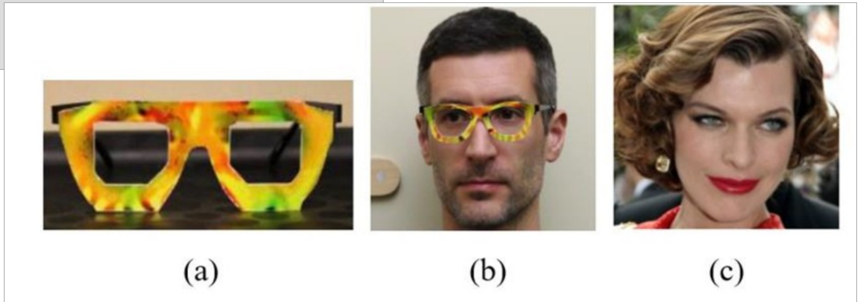
- Billion USD business (2023: ~ 200 bUSD)
- dangerous or advantageous?
  (jobs, information access and safety, decision making,…)
- cultivate or regulate?

Introduction: AI challenges

We can not address certain problems…

- pen or not pen; “I don’t know” (they need balanced datasets)
- generalization (“catastrophic forgetting”)
- error control (prone to adversarial attacks, hallucinations)
- planning





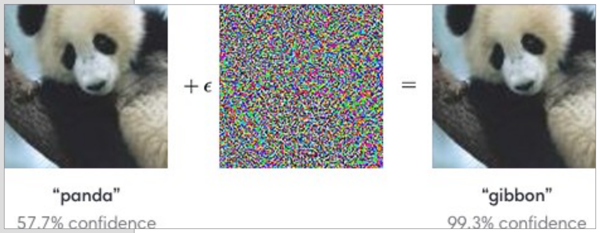
4: The eyeglass frames (a) used by Lujo Bauer (b) to impersonate Milla Jovovich (c)

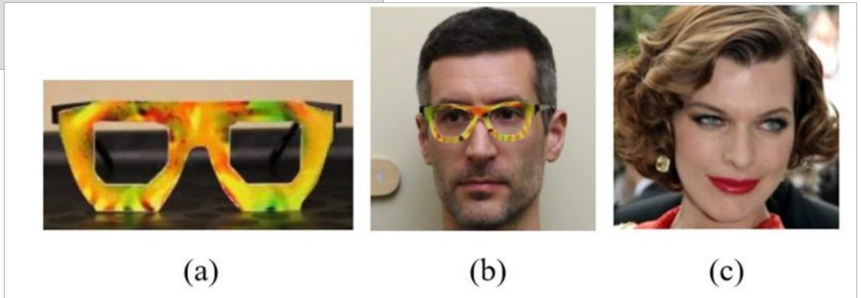
Intelligent from certain points of view,
but certainly not “thinking machines”

Introduction: AI challenges

We can not address certain problems…

- pen or not pen; “I don’t know” (they need balanced datasets)
- generalization (“catastrophic forgetting”)
- error control (prone to adversarial attacks, hallucinations)
- planning





4: The eyeglass frames (a) used by Lujo Bauer (b) to impersonate Milla Jovovich (c)

Intelligent from certain points of view, but certainly not “thinking machines”

How can we approach “thinking” – what was the question, if the answer is human intelligence? (Turing test, classification task, …)

Introduction: back to the origin…

What is the intelligence good for?

Motto:

πάντα ῥεῖ

(Ἡράκλειτος)

panta rhei: everything flows

(Heraclitus)

Introduction: what is a pen?

Motto:

πάντα ῥεῖ

(Ἡράκλειτος)

panta rhei: everything flows

(Heraclitus)

We get all information in time, but …

nothing remains the same in time:

(example: a pen)

- move position
- change shape
- change matter content
- change microstate or quantum state

What we refer to as “a pen” is in fact a set of states.

The important question is: what defines this set?

Introduction: what is a pen?

What defines this set? → the question to approach *intelligence*

present day (DNN) approach:

- **we need someone to tell** what a pen is (training, supervised learning)
- image recognition: we use universal function approximator (neural networks)
- seek a function that tells apart e.g. pens and pencils
- **technically:** adjust parameters of trial function until appropriate accuracy

$$f(x;\alpha_1 \ldots \alpha_n) \longrightarrow f(x;\alpha_1^* \ldots \alpha_n^*) \approx I_P(x)$$

Leads to the aforementioned problems

Persistent concepts

Is there another way?

- everything flows, but we want to ensure our existence for a longer time period
- we have to do predictions → need some stability!
- we need quantities that are the same in the future as in the past

We need to find **conserved quantities** (laws) in the observed data!

Persistent concepts

We need to find **conserved quantities** (laws) in the observed data!

Remarks:

- The important (relevant) information is what remains stable
  → laws (rules) are the **relevant features** of a phenomenon
- We characterize an item with a **lot of laws**
  (e.g. small, four-legged, furry, round eared, squeaking, etc. animal)
- laws are **inherent properties of the data** (data driven), no need to annotate (unsupervised → supervised learning, when the label is part of the time series)

Persistent concepts: science

Surprisingly robust – in fact all of our concepts come from conservation laws!

In science:

- **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
- **physics laws:** persistent relation between measurable quantities:
  e.g. Newton law: $ma - F = 0$ is true for all times
- gas molecules take different configurations, but what matters macroscopically are the (quasi) conserved quantities: **volume, temperature, particle number**, etc

Persistent concepts: finance

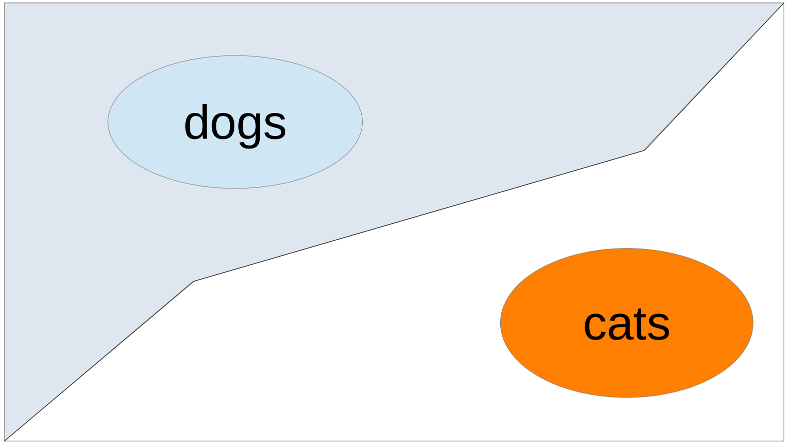
All of our concepts come from conservation laws:

- **Finance**: fluctuating prices $S(t)$ make predictions difficult; the goal is to find some persistent concepts
  - price distribution is more stable $P(S, \dot{S}, \ldots)$ → **market model**
  - to achieve persistency we combine assets → **hedging, indexes**
  - lack of arbitrage, extensive hedging → **risk neutral** market, **martingales** and **pricing**
  - for better predictive power we need to find more conserved quantities!
- to characterize animals we consider persistent properties → **species, breeds**
  taxonomy collects these properties in a hierarchical system

Laws and classification

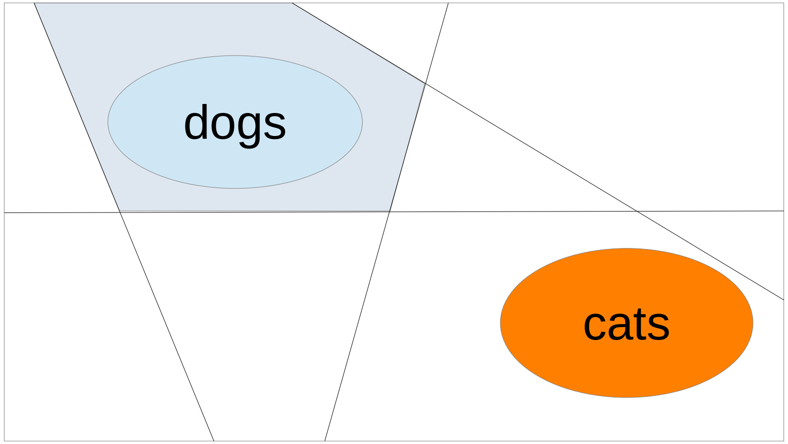
Classification strategies:

neural networks



System 1: few (but complicated) laws, fast evaluation, specific, not controlled

using laws



System 2: more (but simpler) laws, slower evaluation, not specific, control each other

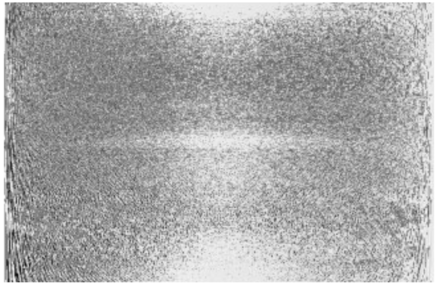
Daniel Kahneman: Thinking, fast and slow – System 1,2 in human psychology

Persistent concepts: representation

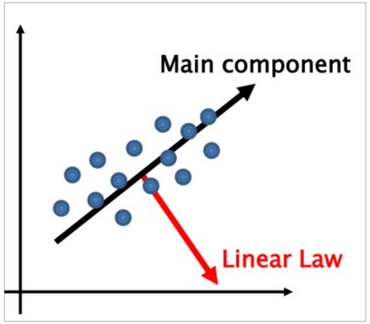
mathematically a persistent concept or **conservation law** can be expressed as

$$f(x)=0, \ \forall \ x \in X$$

- how can we find the appropriate law? assume digital input, i.e. $x \in X = \{0,1\}^N, \ |X| = 2^N$
- then the number of possible $X \rightarrow \{0,1\}$ functions is $2^{|X|}$: **impossible to explore**!
- we *single out some functional space*:
  - shapes, textures → humans are good
  - equivalent representations can be not recognizable → AI may help

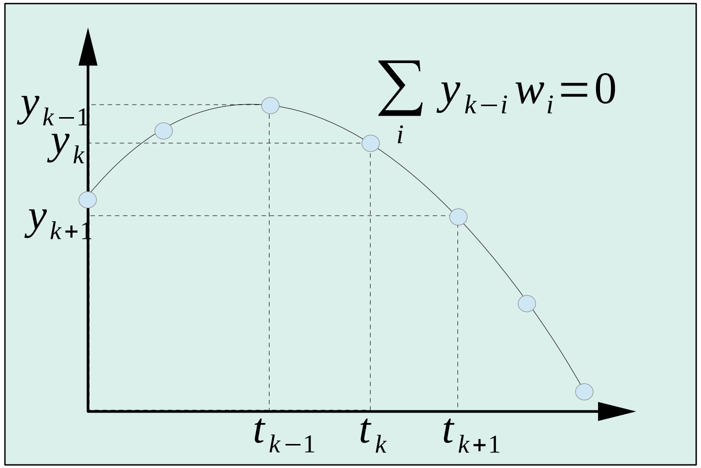


Laws in time series: training



Mathematical procedure: LLT MT Kurbucz, P Pósfay, A Jakovác - Scientific Reports, 2022

- find rules in sub-samples, e.g. linear relation
  - database: samples from time series with uniform steps $Y_{ki} = y((k-i)\Delta t), \quad i = 1 \ldots l$
  - for subsamples: best linear relation $F(Y_k) = \sum_i Y_{ki} w_i = (Yw)_k = \text{minimal}$
- representation of nonlinear laws: $y = \sin(\omega t + \phi) \Longleftrightarrow y_{k+1} - 2 y_k \cos(\omega \Delta t) + y_{k-1} = 0$
  - locally linear
  - multiple laws
  - continue finding local laws until the training set can be represented appropriately
- result: each element of the training set is represented by some of the laws



Laws in time series: classification

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%

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.

Persistent concepts: thinking

- **relevant concepts serve as coordinates:** we characterize the observed items with concepts (e.g. *small, four-legged, furry, round eared, squeaking, etc. animal*)
- often used combinations promote to standalone concepts (e.g. chair)
- **concepts are related**
  - coexistence, consecutiveness, causality → new laws (e.g. “fire – hot = 0” or “[rain now] – [wet street later] = 0”)
  - humans tend to establish these relations easily → not always correct, needs later refinement! (e.g. “[wet street now]-[rain earlier]=0” not necessarily true)
- **context**: relevant concepts depend on the environment (e.g. scale, or role) → change relevant concepts (renormalization group)
- **model:** net of concepts form a dynamical model of the reality → thinking

Conclusions

- we need to maintain ourselves (give predictions) in an ever flowing environment, thus we need persistent quantities (laws, relevant features)
- thus we need to find persistent phenomena (laws): these are our **concepts**
  - **DNN:** a single law for each class, fast, uncontrolled, specific → System 1
  - **generic laws:** data driven, not specific, slower, controlled → System 2
- we need a functional space to seek laws → LLT: multilinear feature transformation
- generic laws allow abstract, general, structured, dynamic modeling → **thinking**

The end