Could the laws of physics explain not just the universe but also the workings of our minds?

The 2024 Nobel Prize in Physics

Prof Truyen Tran

Applied AI Institute (A2I2), Deakin University

truyen.tran@deakin.edu.au | truyentran.github.io

3/4/2025

“The Nobel Prize in Physics 2024 recognizes methods that lay the foundation for the development of artificial intelligence.”

Source: nobelprize.org

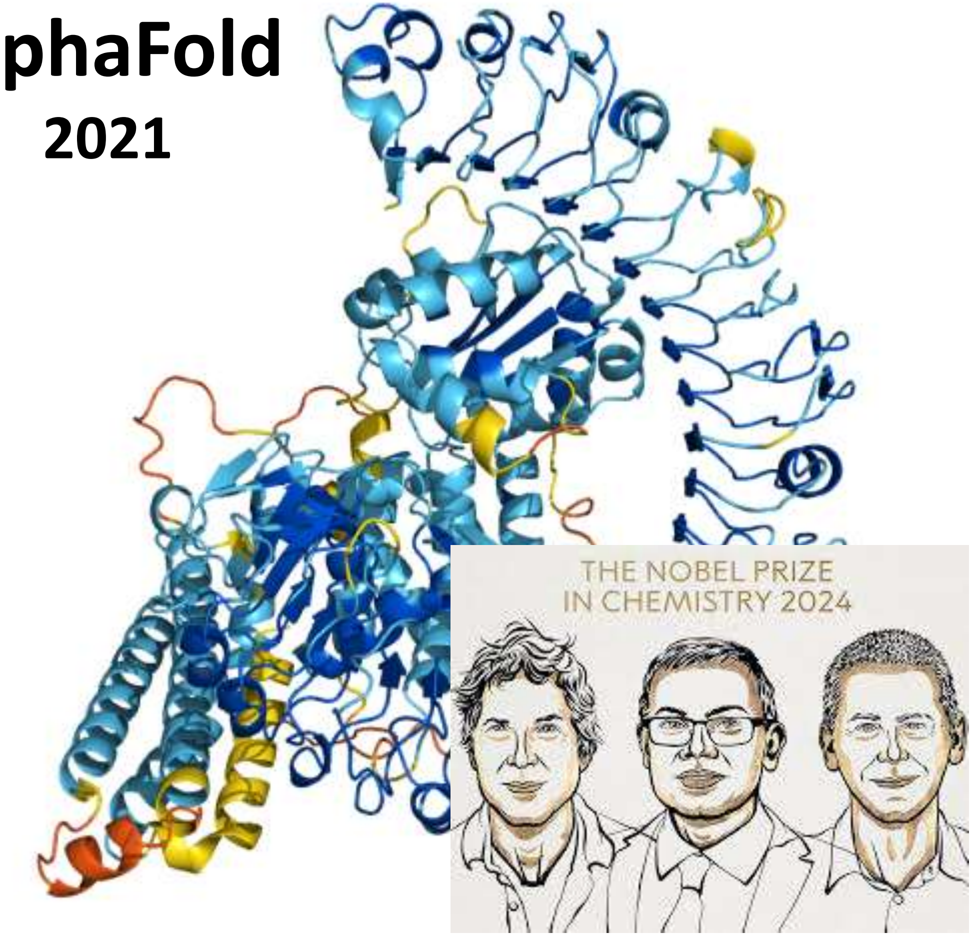
2012

AlphaGo

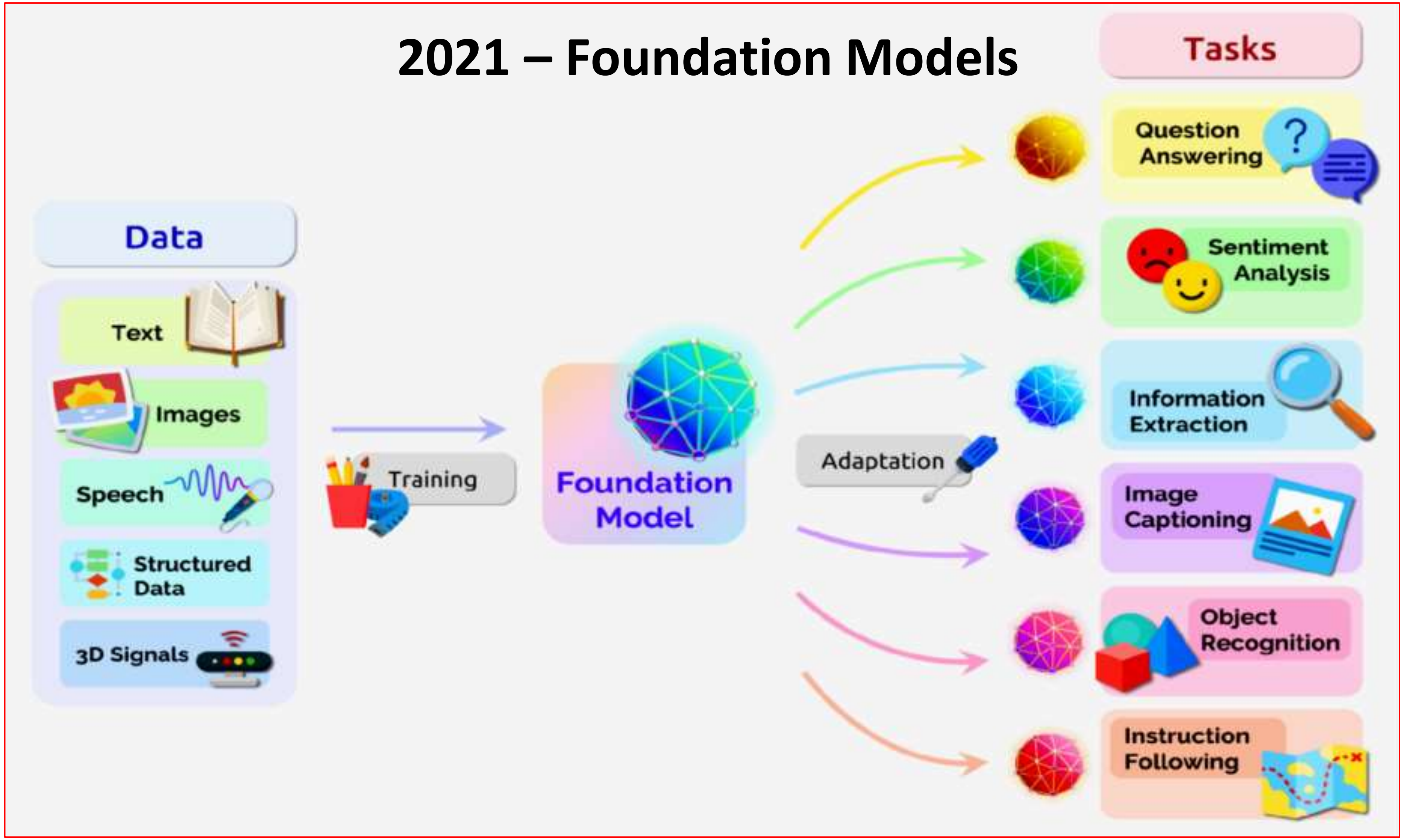
2016

AlphaFold

2021

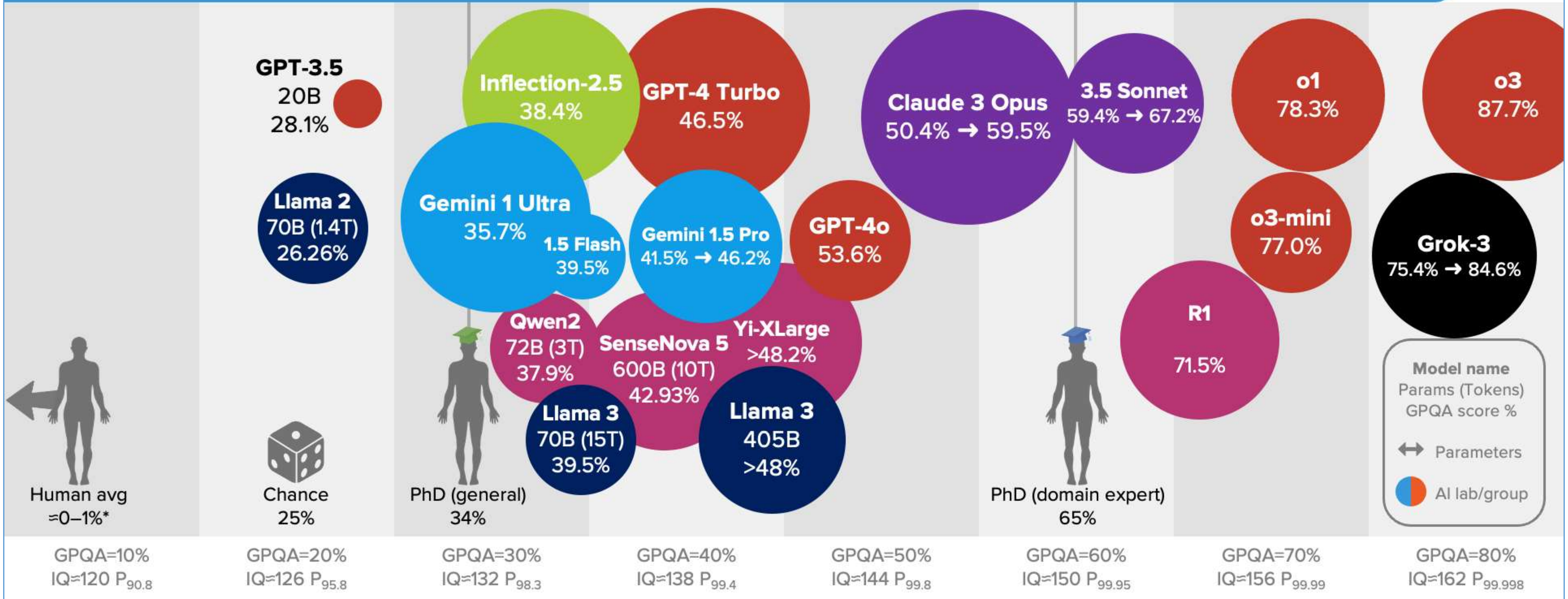


13 years snapshot



13 years snapshot

LARGE LANGUAGE MODELS + GPQA (FEB/2025)



Model sizes near to scale. * Estimates based on independent analysis. Selected highlights only. IQ correlation estimates only: http … //lifearchitect.ai/models-table/ Alan D. Thompson. 2025.

13 years snapshot

LifeArchitect.ai/iq-testing-ai

Five things in modern AI

Representation

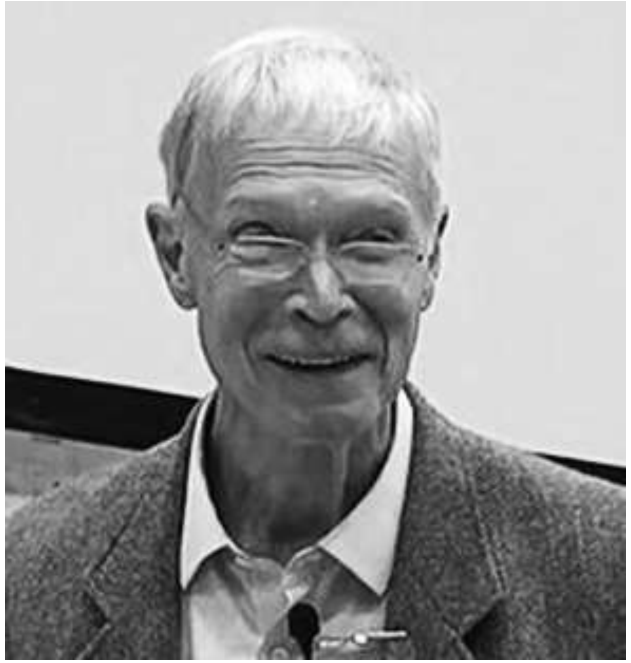
Reasoning

Learning

Noise tolerance

Scale with compute

What was the role of physics in shaping modern AI?



John Hopfield (born 1933)

"How mind emerges from brain is to me the deepest question posed by our humanity."

"Early AI was mainly based on logic. You're trying to make computers that reason like people.

The second route is from biology: You're trying to make computers that can perceive and act and adapt like animals."

Geoffrey Hinton (born 1947)

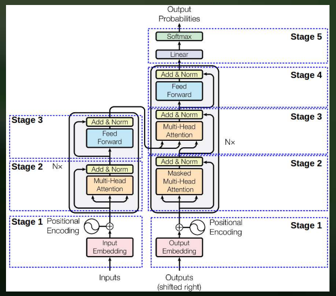
Credits: Wikipedia, Zdnet

BRAIN, MIND, COMPUTATION

HOPFIELD NETWORKS

BOLTZMANN MACHINES

AI: DEEP LEARNING APPROACH



Two major concepts

Associationism: Humans learn through association

- 400BC-1900AD: Plato, John Locke, David Hume, David Hartley, James Mill, John Stuart Mill, Ivan Pavlov.

Connectionism: The network that does computation

- Mid 1800s: The brain is a mass of interconnected neurons
- **Alexander Bain (1873)**: The information is in the connections.

Alexander Bain
(1818 –1903)

The original quest of AI

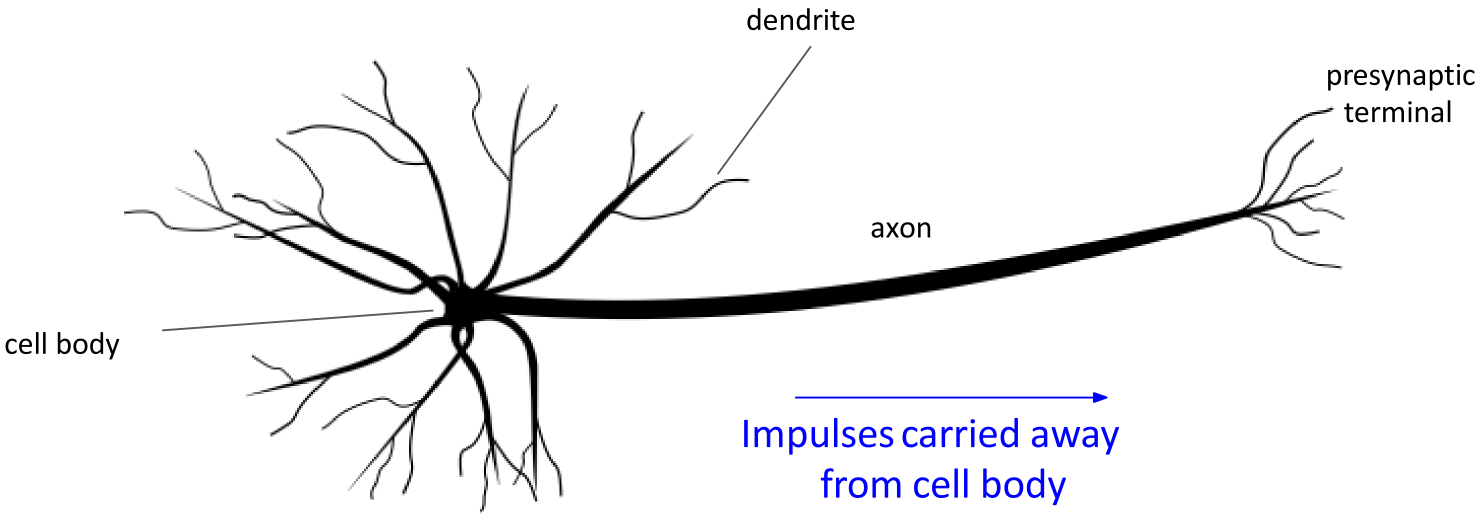
Among the most challenging scientific questions of our time are the corresponding analytic and synthetic problems:

- How does the brain function?
- Can we design a machine which will simulate a brain?

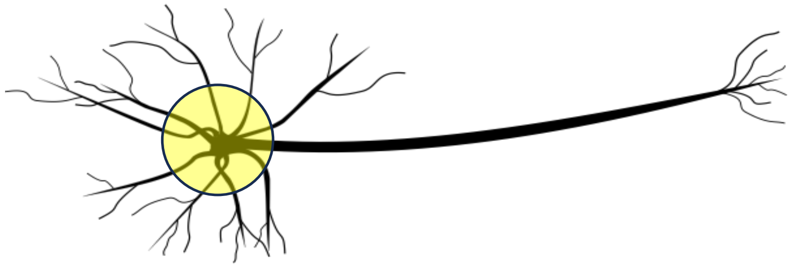
-- *Automata Studies*, 1956.

Image credit: Dickson Wu

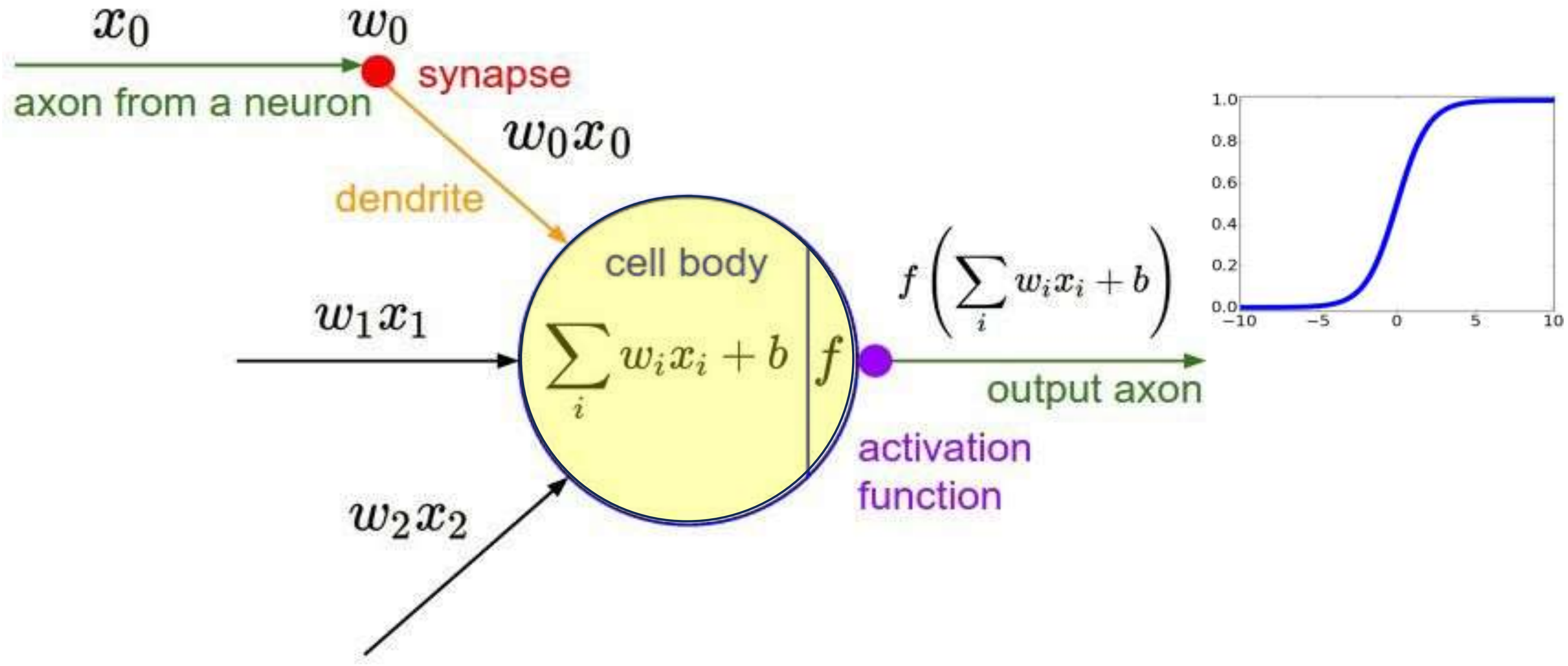
Neurons



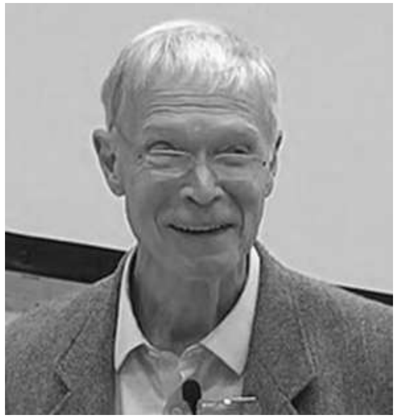
Credits: Fei-Fei Li, Ranjay Krishna, Danfei Xu



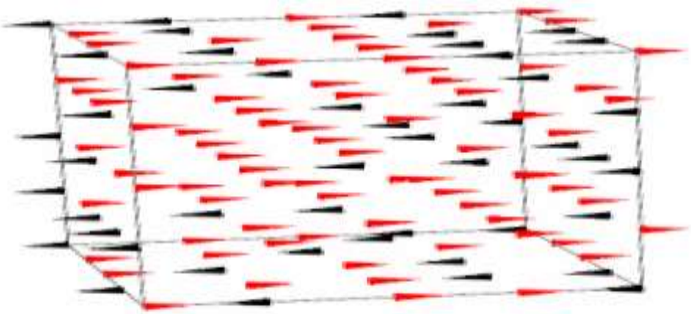
Integrate-and-fire neuron
in silico



Credits: Fei-Fei Li, Ranjay Krishna, Danfei Xu

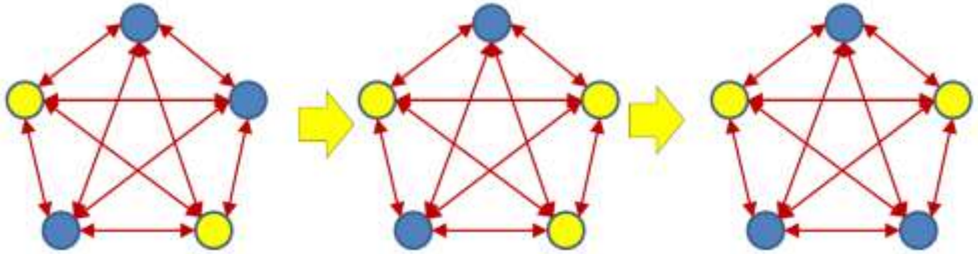


Hopfield network



A spin glass system

- Symmetric weights
- No self-connection
- Update either asynchronous or synchronous
- Neurons attract or repel each other
- The dynamic is deterministic



$$s_i \leftarrow \begin{cases} +1 & \text{if } \sum_j w_{ij} s_j \geq \theta_i, \\ -1 & \text{otherwise.} \end{cases}$$

Credits: Wikipedia, Bhiksha Raj, Rita Singh

Questions

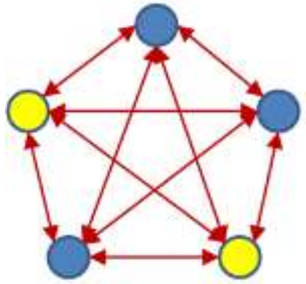
How do we teach the network to store *a specific* pattern or set of patterns?

How many patterns can we store?

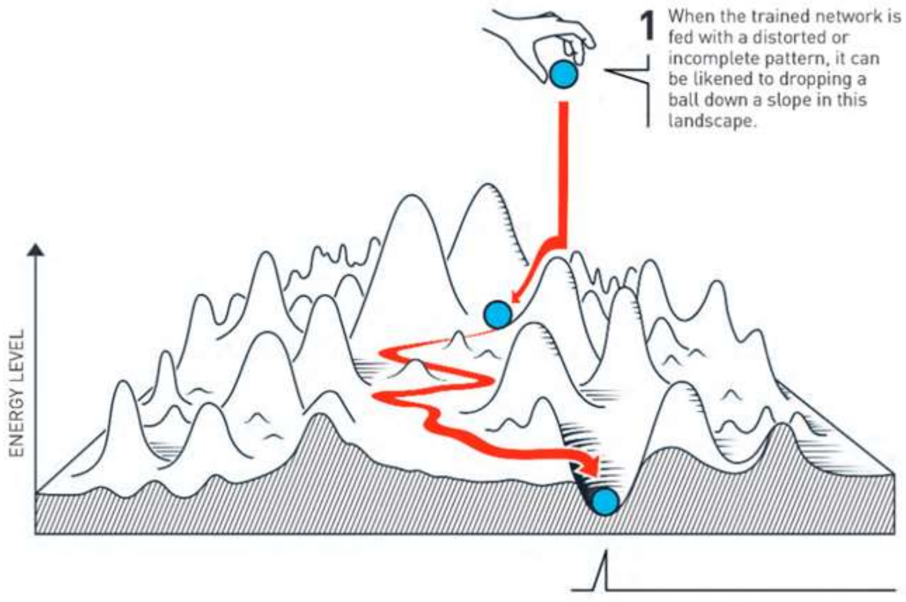
How to “retrieve” patterns better.

Update as energy minimisation

- Start at some initial pattern (configuration)
- Let the network "runs"
- The convergence is a local attractor (stable)
- The basis for associative memory recall!
- The operation is essentially Iterated Local Mode (ICM) known in spatial statistics, 1975!



$$E = -\frac{1}{2}\sum_{i,j} w_{ij} s_i s_j - \sum_i \theta_i s_i$$



Credits: Wikipedia, Bhiksha Raj, Rita Singh, nobelprize.org

How to store patterns

"Neurons that fire together, wire together"

$$w_{ji} = y_j y_i$$ *Hebbian learning rule*

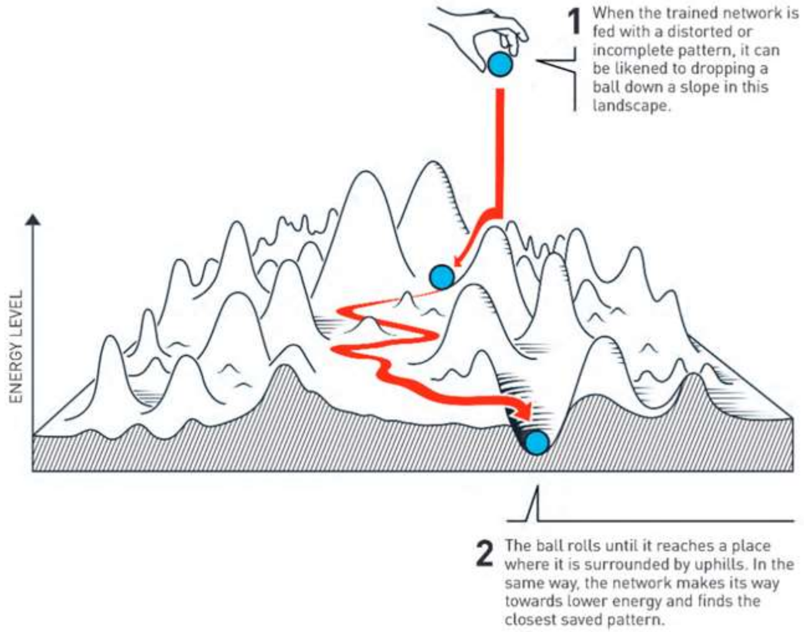
$$E = -\sum_i \sum_{j<i} w_{ji} y_j y_i = -\sum_i \sum_{j<i} y_i^2 y_j^2$$

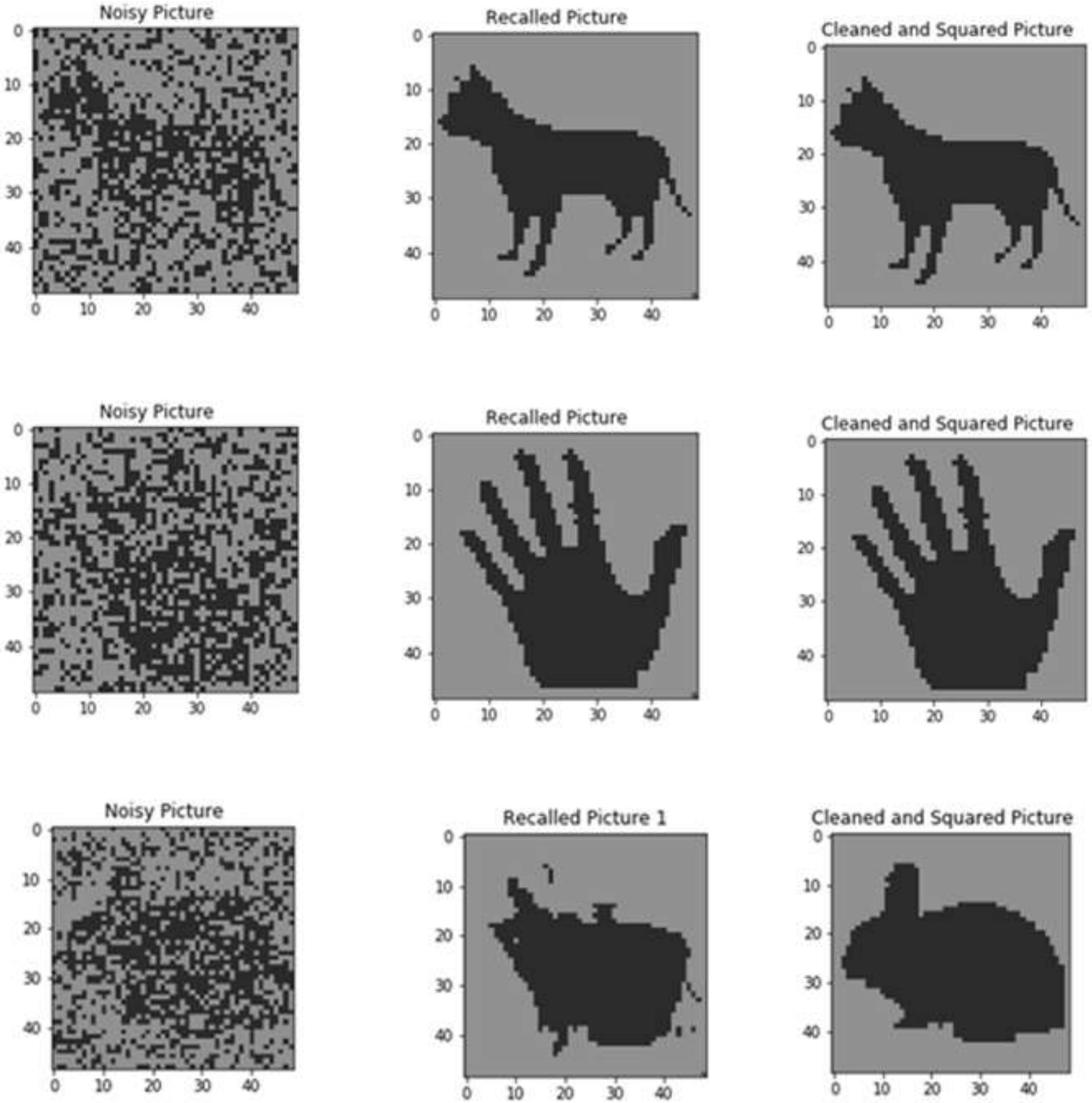
$$= -\sum_i \sum_{j<i} 1 = -0.5N(N-1)$$

This is the global minimum

For a network of N neurons can store up to ~0.15N patterns through Hebbian learning

Examples: Image denoising





Boltzmann machine

A stochastic Hopfield network with hidden nodes

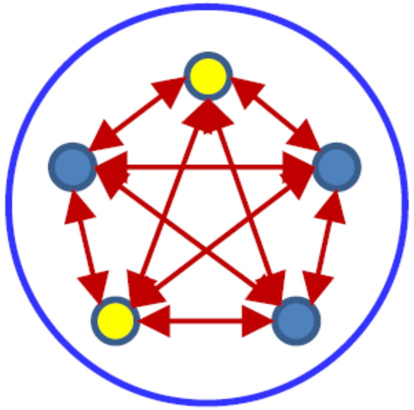
- Capacity of Hopfield network can be vastly increased by introducing hidden nodes
- Stochasticity gives principled ways to handle uncertainty, randomness and statistical properties
- Hopfield net is a special case when T → 0.

The Helmholtz Free Energy of a System

$$F_T = \sum_s P_T(s)\, E_s + kT \sum_s P_T(s) \log P_T(s)$$

$$P_T(s) = \frac{1}{Z} exp\left(\frac{-E_s}{kT}\right)$$

$$E(S) = -\sum_{i<j} w_{ij} s_i s_j$$



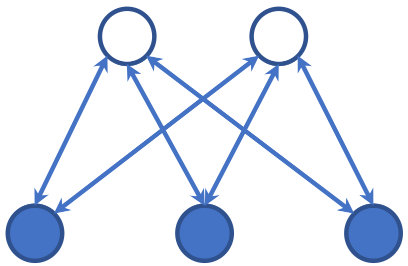
Gibbs distribution

Credits: Wikipedia, Bhiksha Raj, Rita Singh

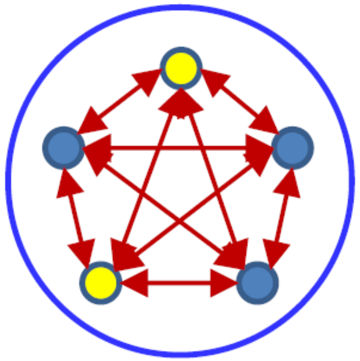
Restricted Boltzmann machine (RBM)

>Hidden variables to denote underlying unobserved processes

>RBM encourages data representation!



Restricted Boltzmann Machine (~1994, 2001)



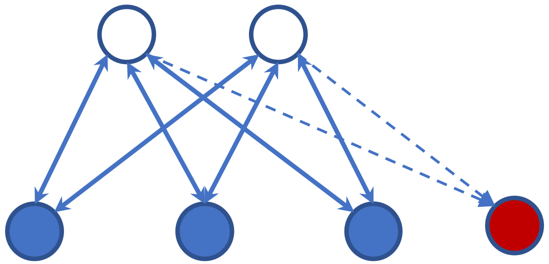
Boltzmann Machine (1985)

visible

$$p(\mathbf{v}, \mathbf{h}; \psi) \propto \exp\left[-E(\mathbf{v}, \mathbf{h}; \psi)\right]$$

hidden

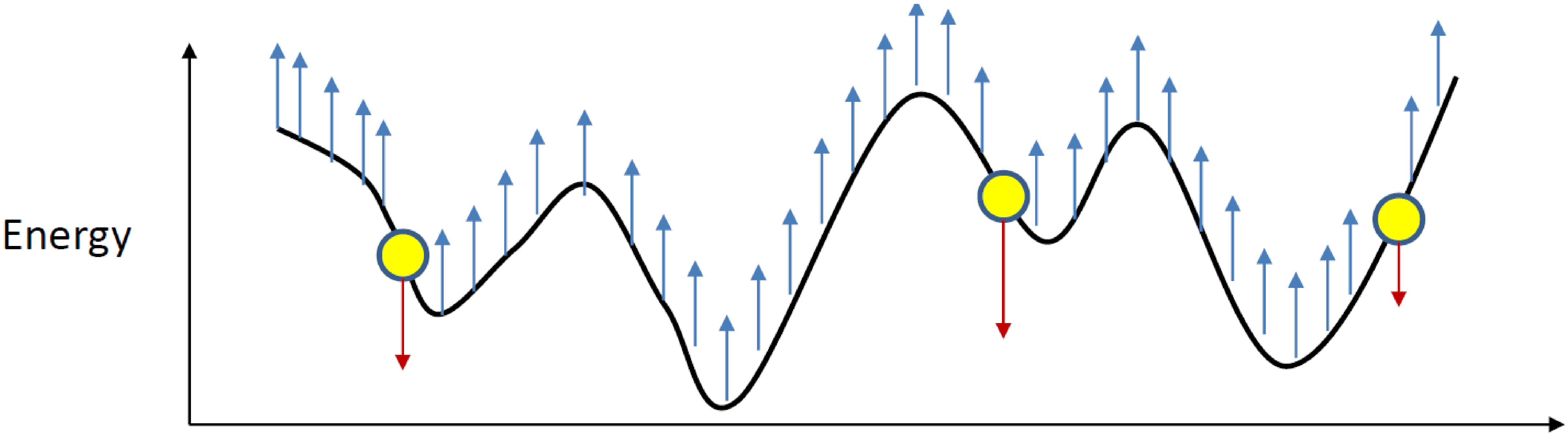
energ
y



RBM for prediction

Training a Boltzmann machine

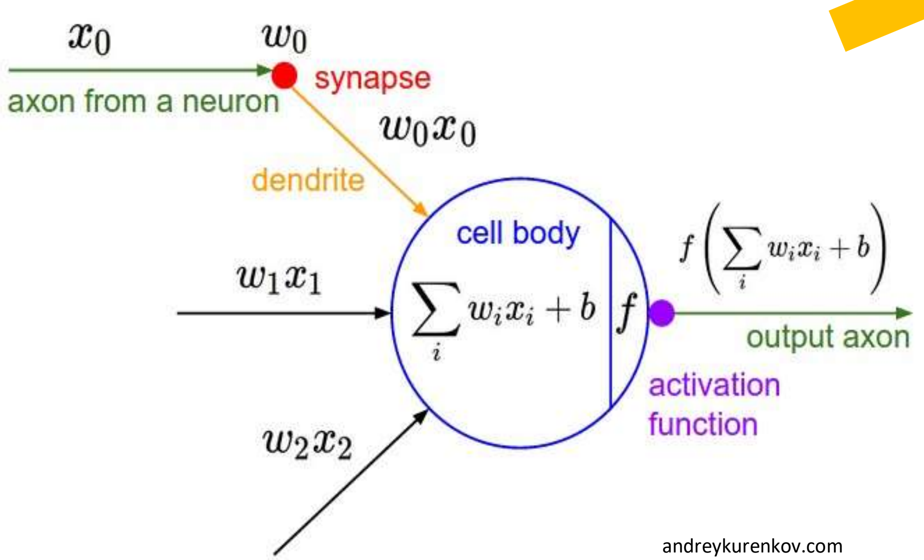
By minimising the energy of the observed configurations and raising the energies where else.



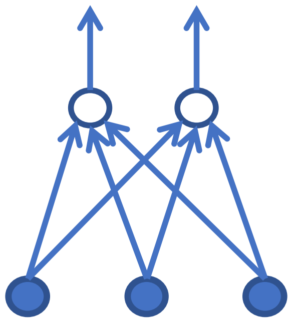
Models that can do these five things win!

- Representation
- Reasoning
- Learning
- Noise tolerance
- Scale with compute

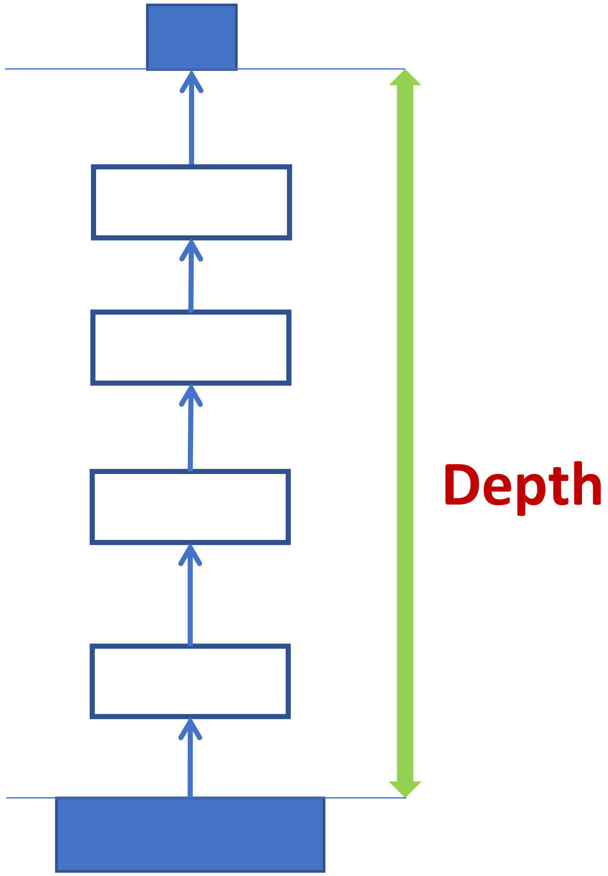
Deep neural networks



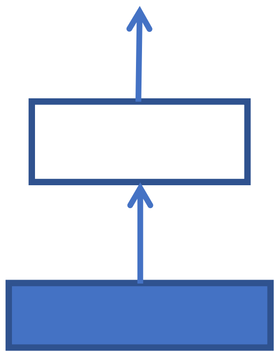
Integrate-and-fire neuron



Feature detector

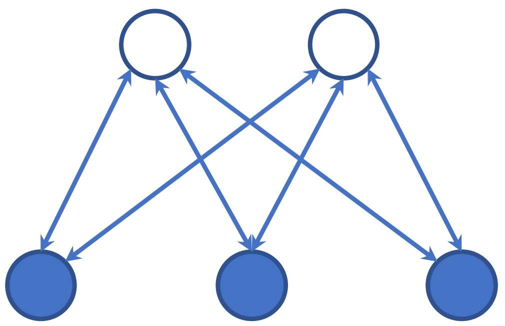


Deep network

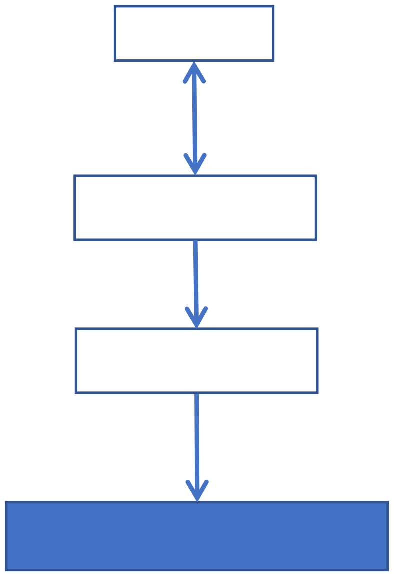


Block representation

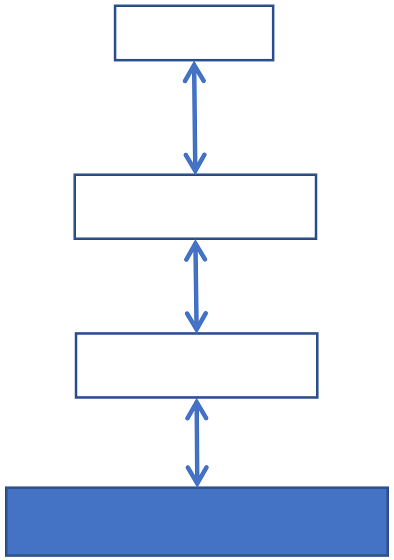
Layer-wise training



Restricted Boltzmann Machine (~1994, 2001)



Deep Belief Net (2006)



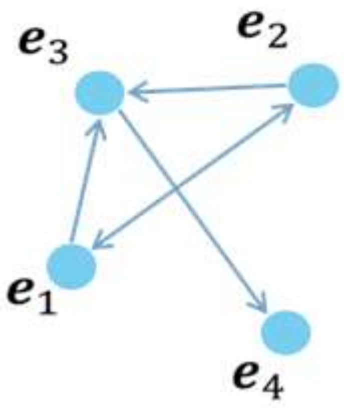
Deep Boltzmann Machine (2009)

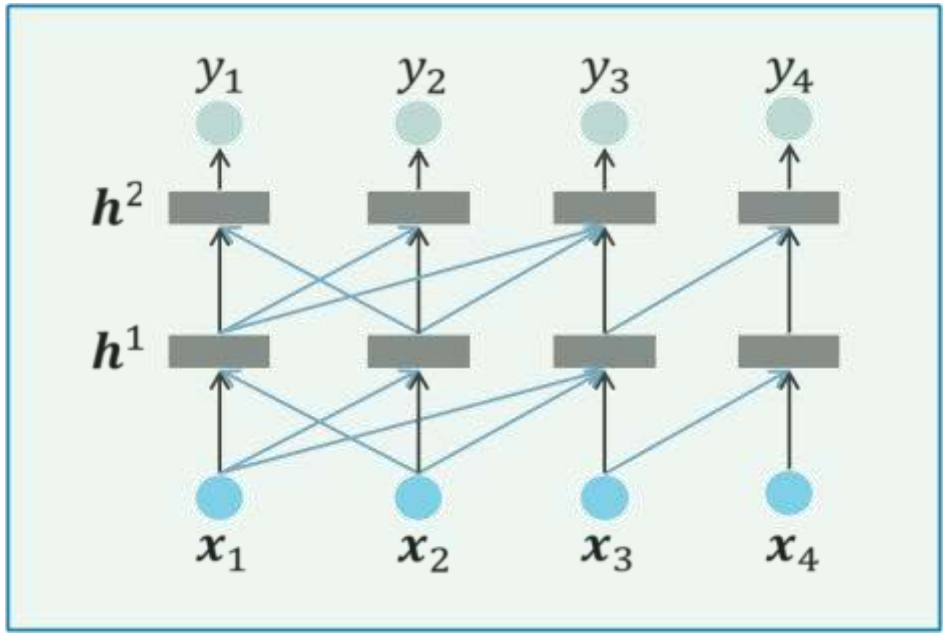
Spoil alert: Stack of RBMs is akin to renormalization trick in physics

Modelling the world with neural networks

Column Networks inspired from columnar structure of brain

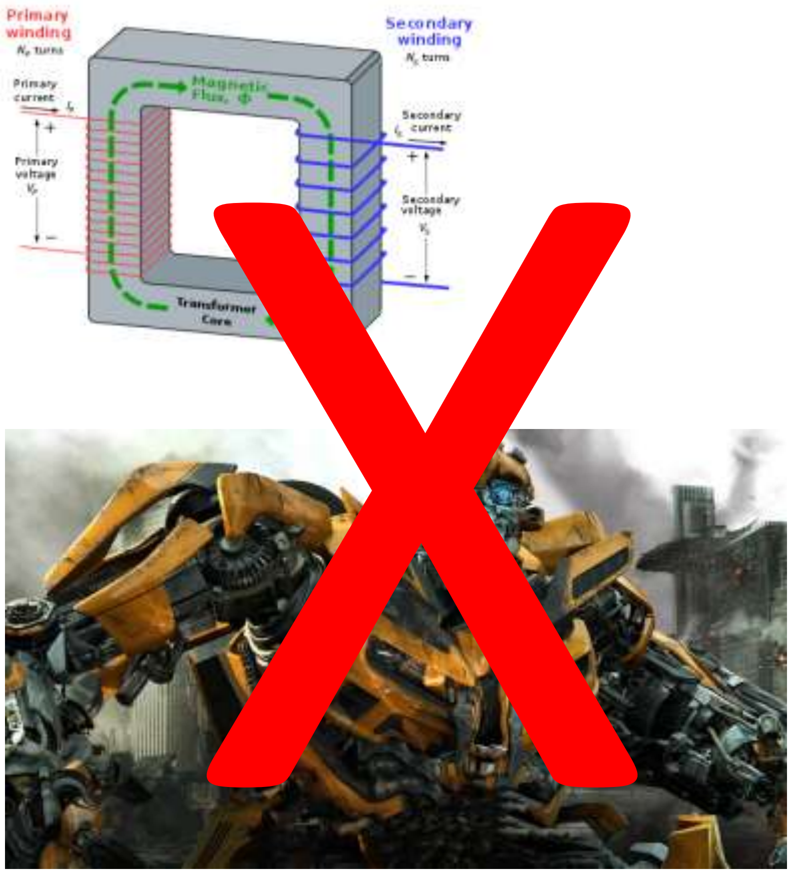
Relation graph

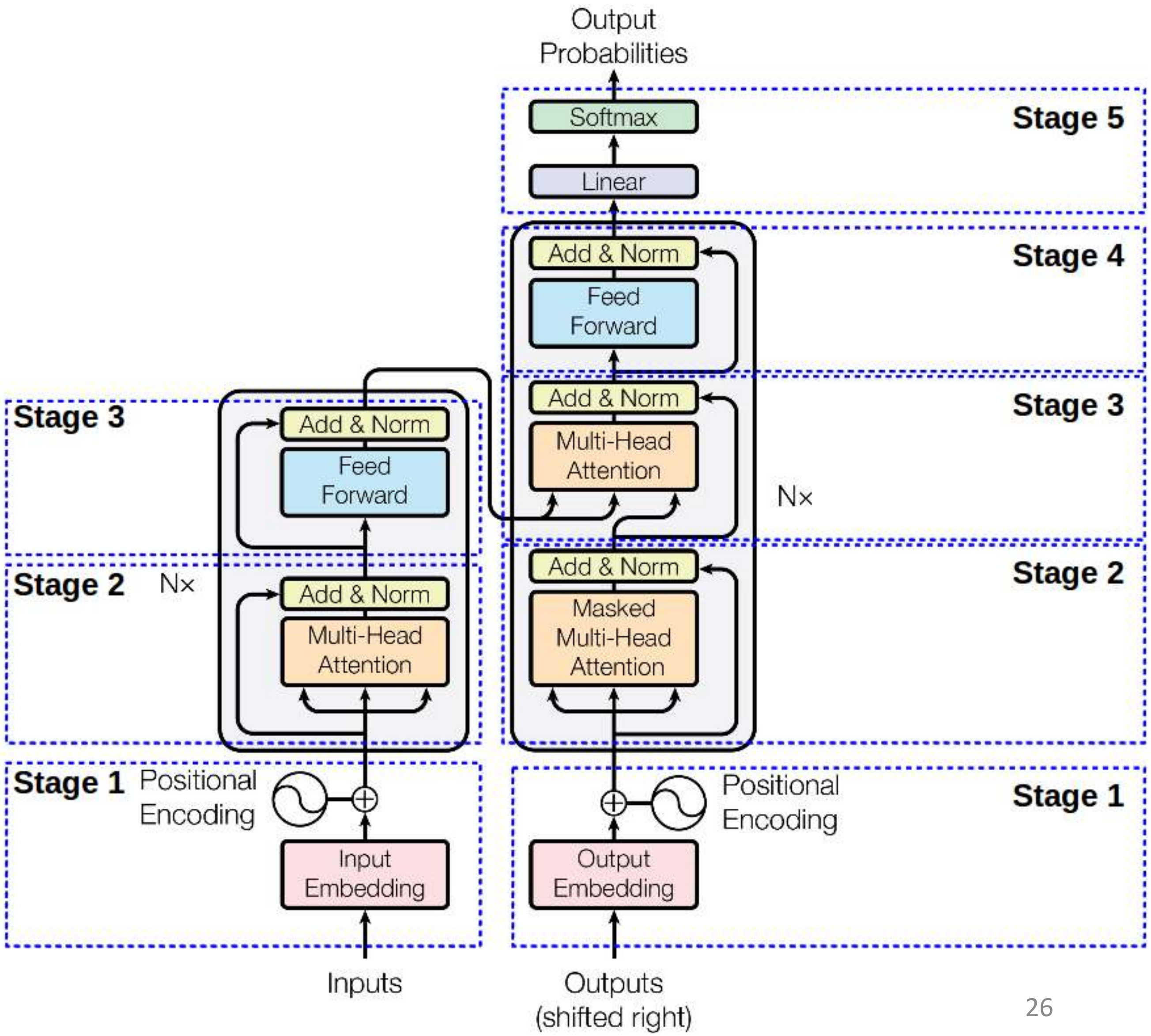




#REF: Pham, Trang, et al. "Column Networks for Collective Classification." *AAAI*. 2017.

Workhorse of modern AI: Transformer





Transformer: Key ideas

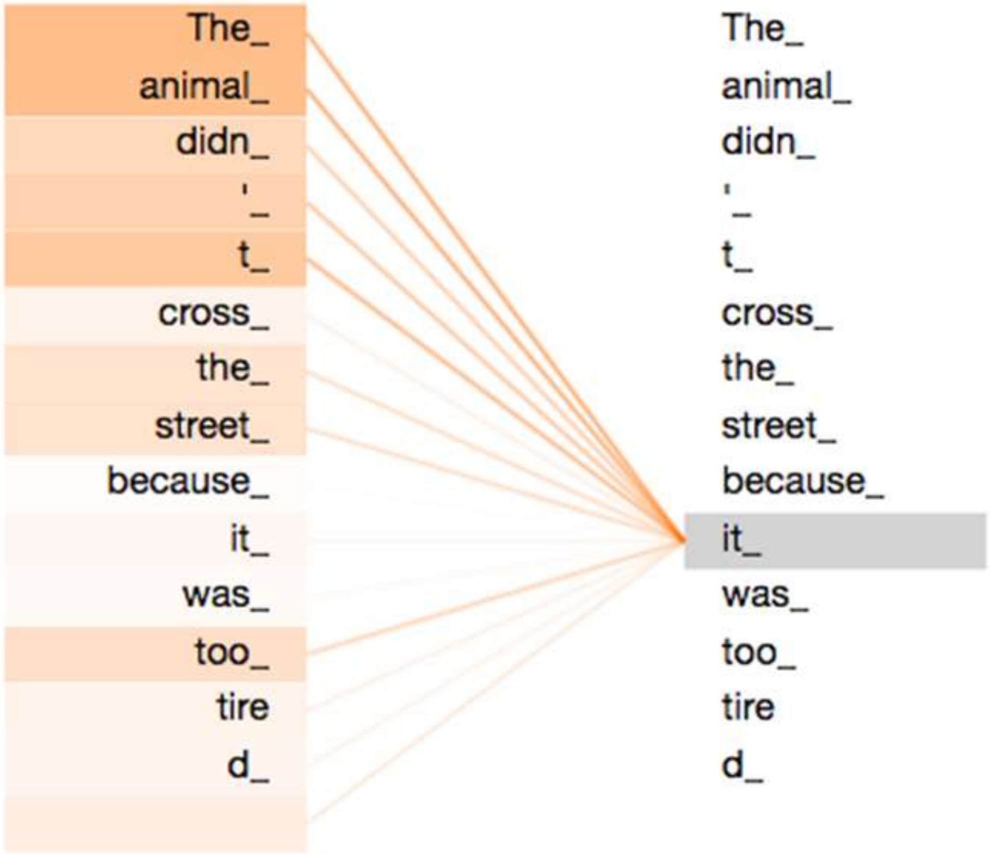
>A special case of Graph Neural Networks

>Everything is a set of tokens (e.g., words)

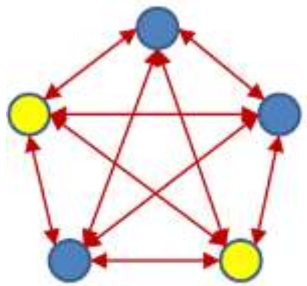
>Tokens are "embedded" into a high-dim vector (a set of neurons)

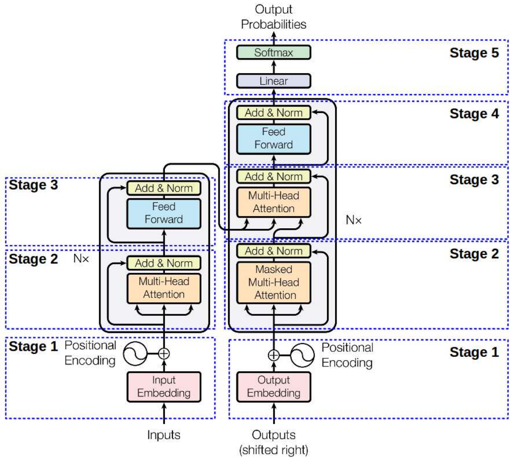
>Tokens in the same context are jointly considered

>Token embedding is "shaped" by their relationship with other tokens.



Transformers are Hopfield nets!



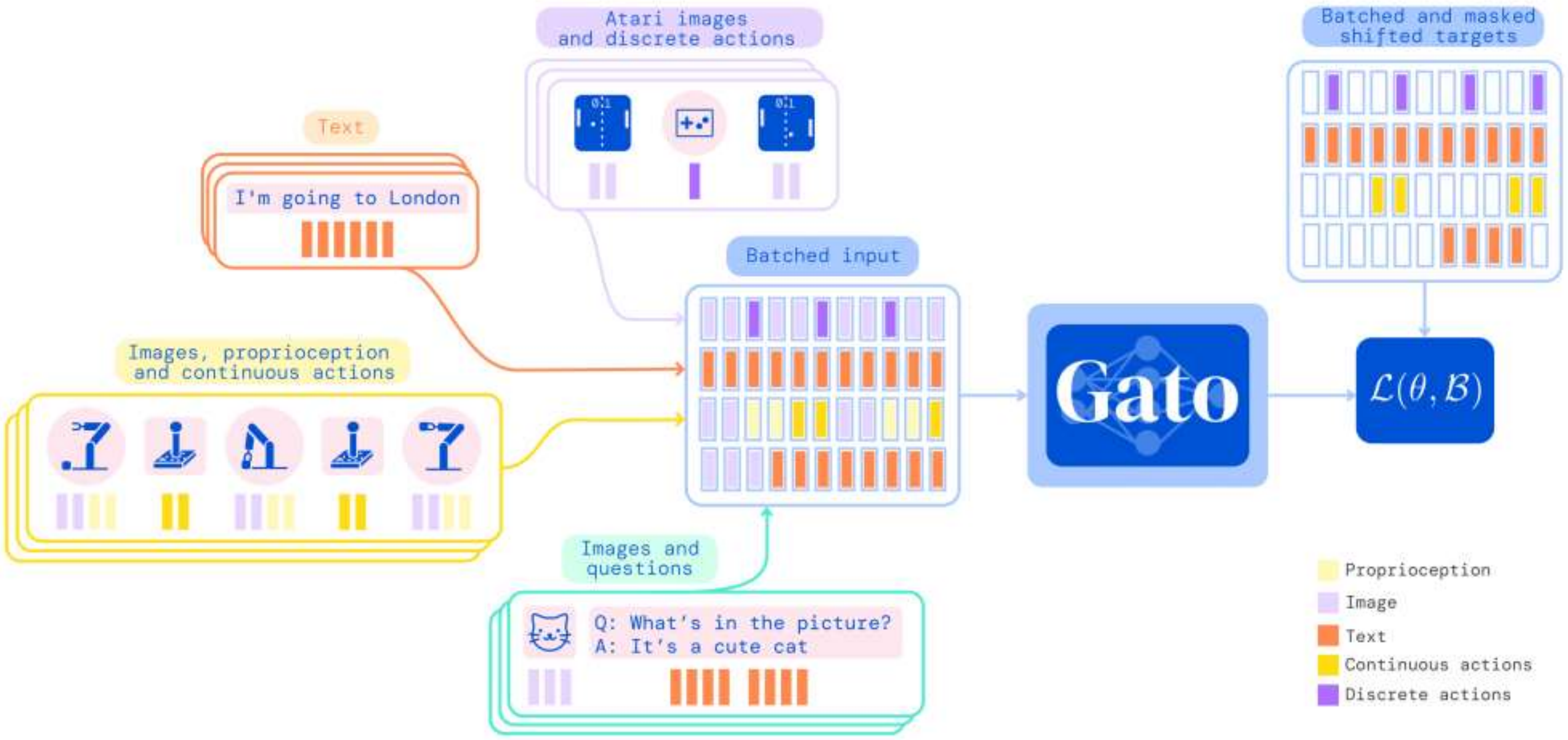


Hopfield Energy		New Energy		Update Rule		Transformer
$-\exp\left(\mathrm{lse}\left(1, \boldsymbol{\xi}\boldsymbol{X}^T\right)\right)$	$=$	$-\mathrm{lse}(\beta, \boldsymbol{\xi}\boldsymbol{X}^T) + \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi} + c$	$\Rightarrow$	$\mathrm{softmax}\left(\beta\ \boldsymbol{\xi}\boldsymbol{X}^T\right)\boldsymbol{X}$	$=$	$\mathrm{softmax}\left(\frac{1}{\sqrt{d_k}}\ \boldsymbol{Q}\boldsymbol{K}^T\right)\boldsymbol{V}$

Ramsauer, Hubert, et al. "Hopfield networks is all you need." *arXiv:2008.02217* (2020).

Credit: DeepMind

Convergence: One-model-for-all – the case of Gato (2022)



Reed et al. "A generalist agent." *arXiv:2205.06175* (2022).

Why one-model-for-all possible?

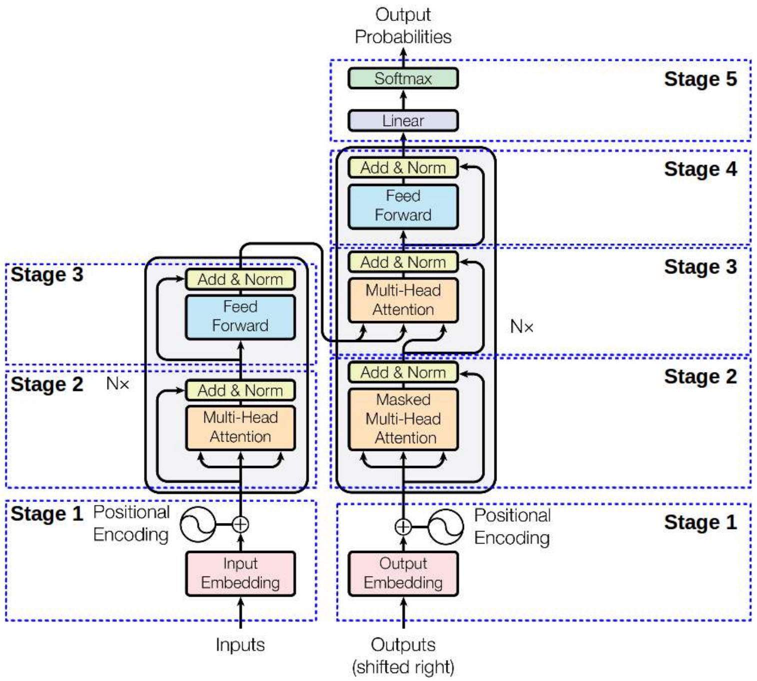
The world is regular: Rules, patterns, motifs, grammars, recurrence

- World models are learnable from data!

Human brain gives an example

- One brain processes all modalities, doing plenty of tasks, and learning from different kind of training signals.
- Thinking at high level is independent of input modalities and task-specific skills.

Deep neural networks | Electronics



Neuron → SENSOR, FILTER

Gate → AND gate, Transistor, Resistor

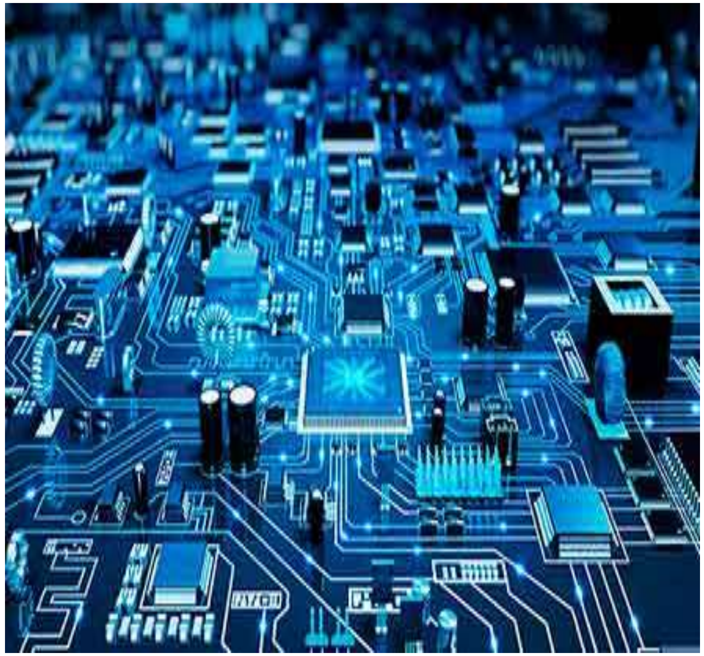
Attention mechanism → SWITCH gate

Memory + forgetting → Capacitor + leakage

Skip-connection → Short circuit

Computational graph → Circuit

Compositionality → Modular design



The current wave:

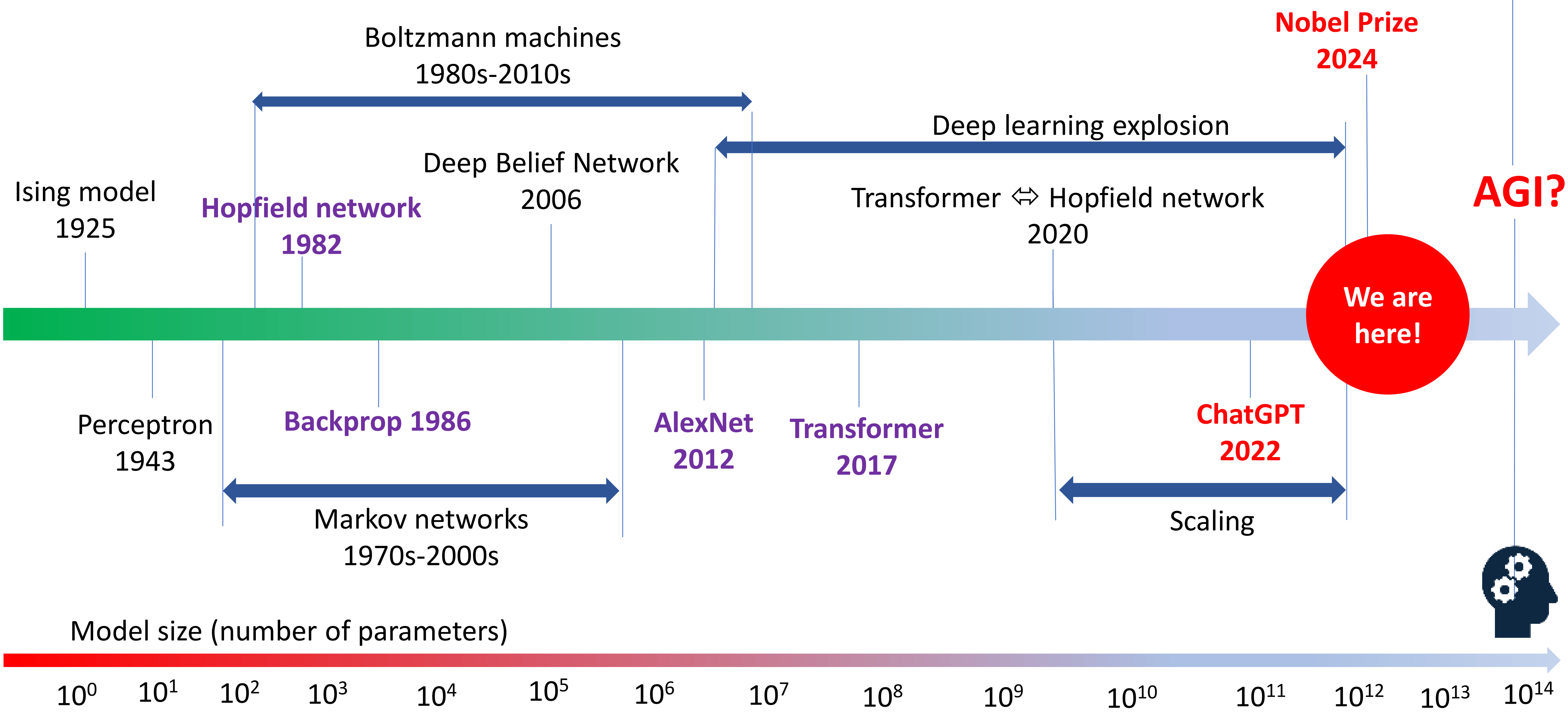
Generative AI

(e.g., ChatGPT, Midjourney, Sora, Veo)

Gen AIs are compression engine	Prompting is conditioning for the (preference-guided) decompression.
Gen AIs are approximate program database	Prompting is retrieving an approximate program that takes input and delivers output.
Gen AIs are World Model	We can live entirely in simulation!

The 100 years of making …

AGI = Artificial General Intelligence



Source: pastorboller.com

Recognition of the 2024 Nobel Prize in Physics

The interdisciplinary nature of modern physics, crossing with:

- Computation & information theory
- Cognition.

Concepts foundational to AI stem from physics

- Reshaping the discipline's boundaries.
- Foundational work => practical AI systems.

Innovation: AI has pervasive societal impact.

Looking into the future

Giorgio Parisi, 2021 Nobel in Physics for complex systems

AI as a discovery tool: e.g., quantum mechanics, materials science, and complex systems.

Interdisciplinary: Physics + AI + cognitive sciences for study of universe and human cognition.

Philosophical implications: Informational fabric => the nature of consciousness, intelligence, and the universe's computational structure.

Future-ready scientists: Technology + science.

Credit: nobelprize.org

AI as physics

> Intelligence as self-organizing phenomena: reducing ignorance/entropy

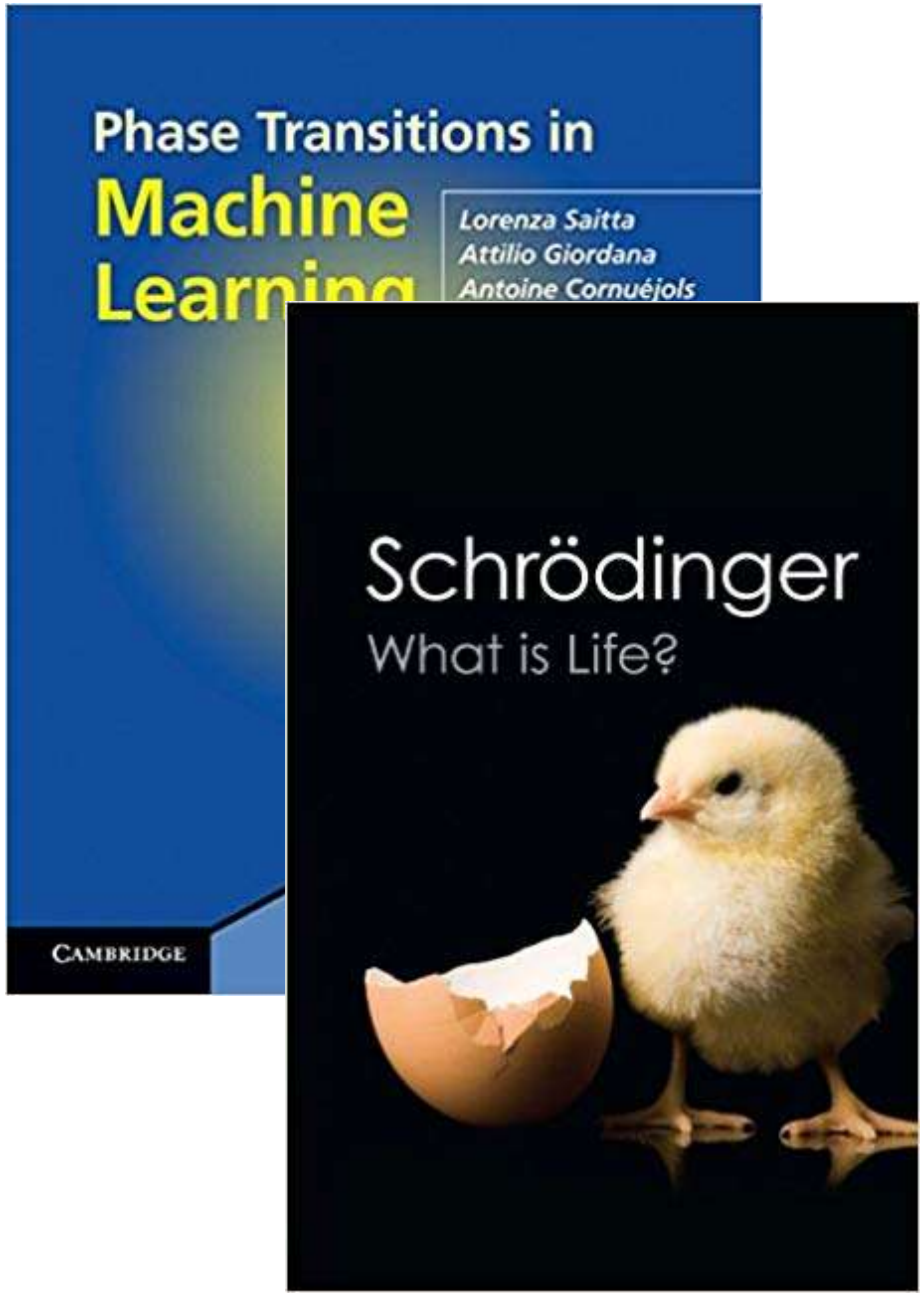
> Neural networks as a statistical mechanical system

> Learning as variational optimization

> Reasoning free-energy minimization

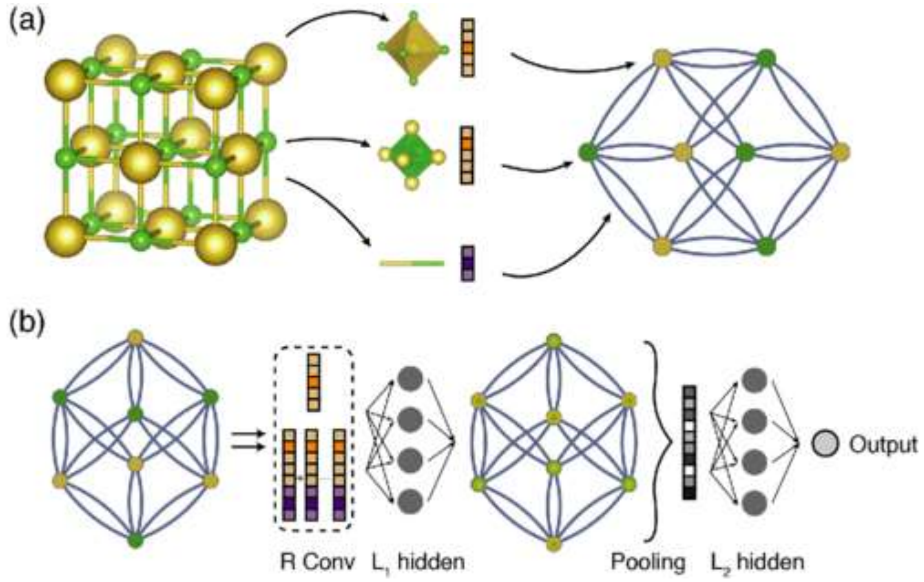
> Phase transition may occur in AI systems

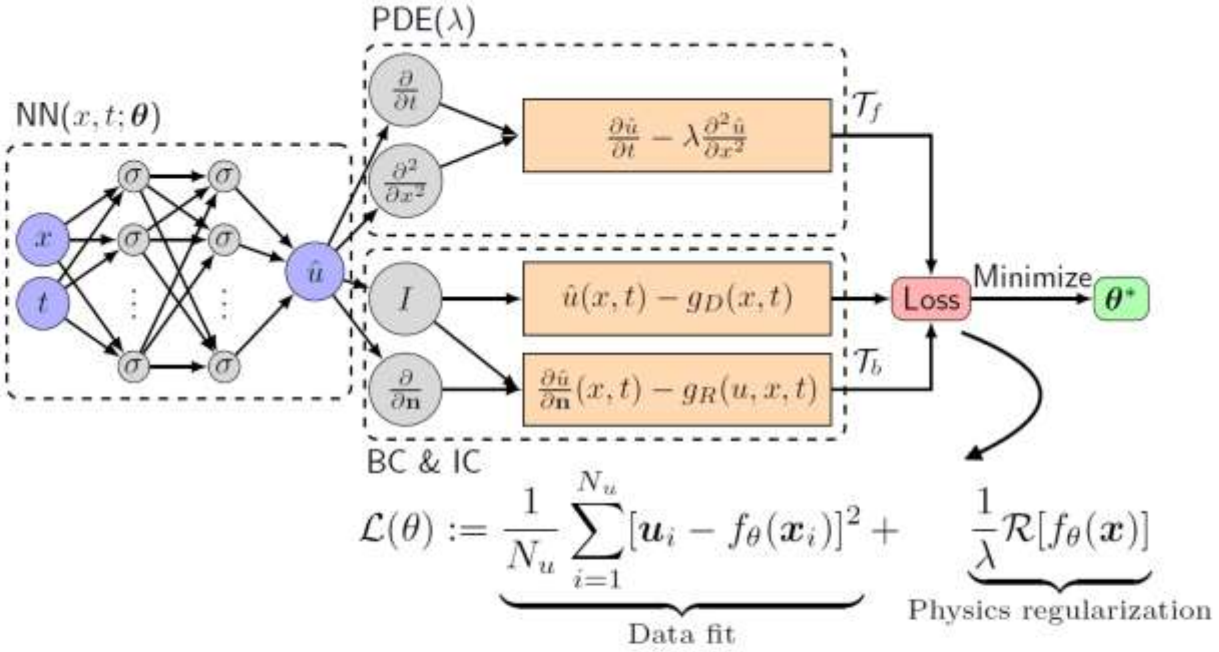
> Ultimate AI must solve the **consciousness problem**, which may require new physics.



AI for Physics

Physics-informed AI



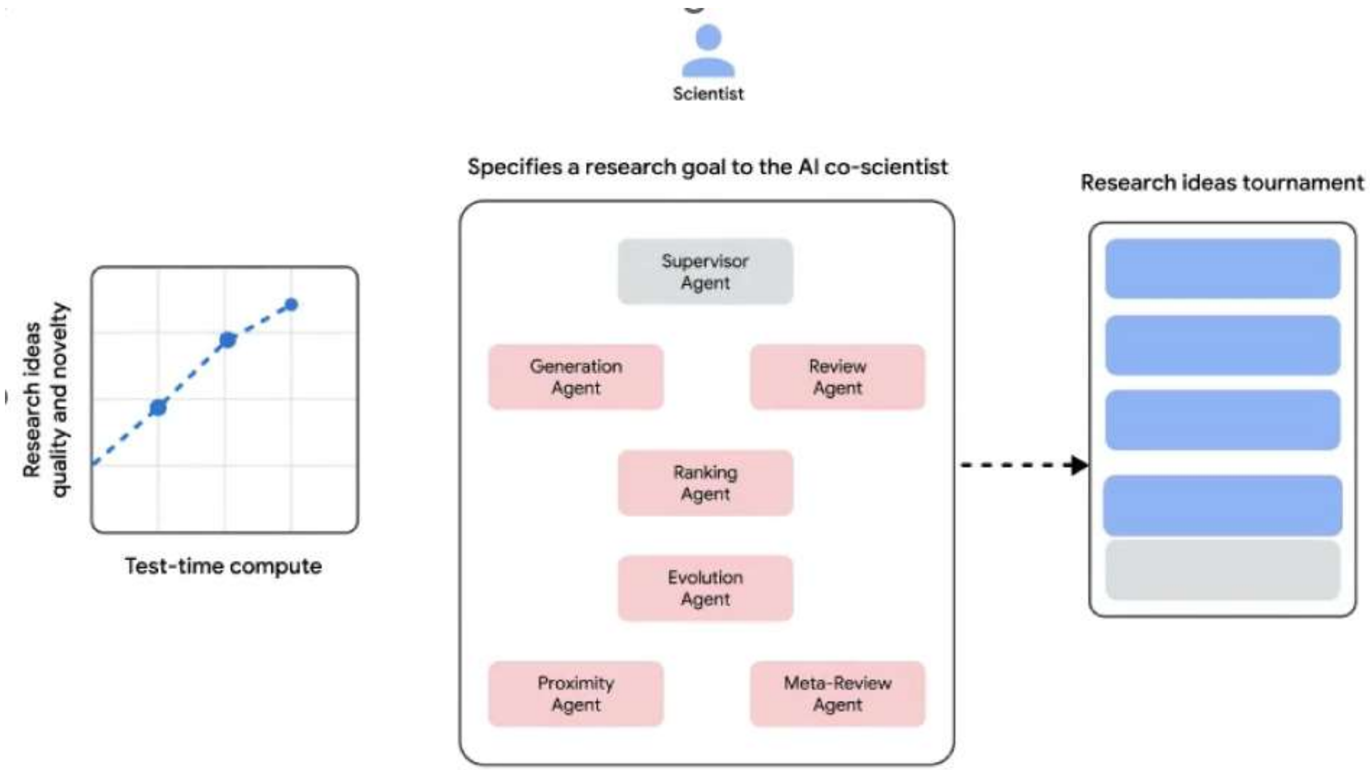


Xie, Tian, and Jeffrey C. Grossman. "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties." *Physical review letters* 120.14 (2018): 145301.

AI for automating chemistry

AI to accelerate medicinal discovery

AI Co-Scientist, Google DeepMind



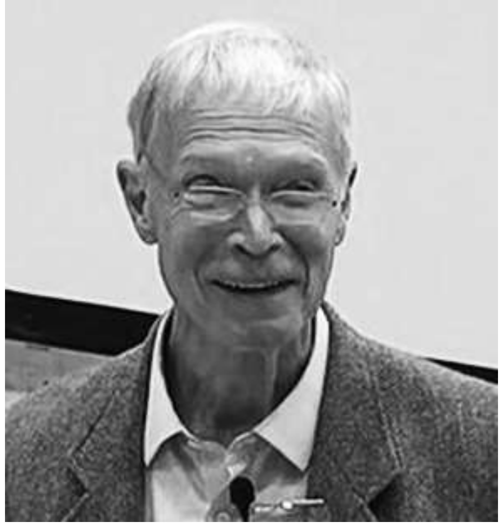
Why is AI scientist possible?

The world is regular: Rules, patterns, motifs, grammars, recurrence

> World models are learnable from data (real or simulated)!

> "Any pattern that can be generated or found in nature can be efficiently discovered or modelled by a classical learning algorithm".

(***Demis Hassabis***, *World Chess Championship, CEO of Google DeepMind, Nobel Laureate in Chemistry, 2024*)



“Physics is a point of view that the world around us is, with effort, ingenuity, and adequate resources, understandable in a predictive and reasonably quantitative fashion."

(John Hopfield)

My full-loop journey from physics to AI and back

Me and my teammates representing Vietnam at the International Physics Olympiad 1997, Canada.

Physics
Chemistry
Materials science
Drug discovery

Physics

Computer science

Statistical machine learning

Artificial intelligence

Crossing the boundaries can take us really far.

Image source: pobble365


A church in my home village in Vietnam in style of Buddhist temple

Fusing traditions can be very beautiful