

Hopfield-Ising networks and Hinton-Boltzmann machines

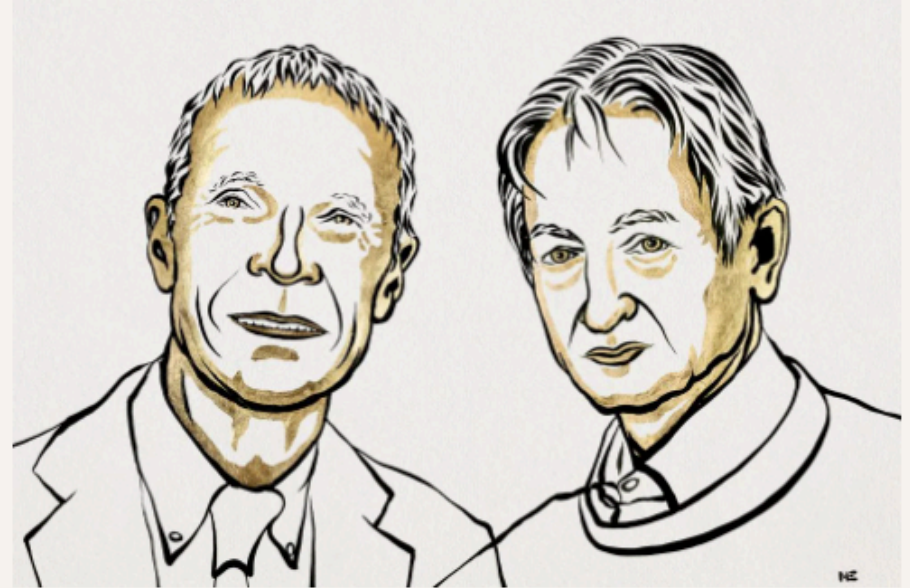
**How analogies and insights from statistical mechanics
created the AI revolution
and led to two Nobel prizes in Physics and Chemistry '24**

George Barbastathis
MIT Mechanical Engineering

The 2024 physics laureates

The Nobel Prize in Physics 2024 was awarded to John J. Hopfield and Geoffrey E. Hinton “for foundational discoveries and inventions that enable machine learning with artificial neural networks.”

John Hopfield created an associative memory that can store and reconstruct images and other types of patterns in data. Geoffrey Hinton invented a method that can autonomously find properties in data, and so perform tasks such as identifying specific elements in pictures.



John Hopfield and Geoffrey Hinton. Ill. Niklas Elmehed © Nobel Prize Outreach

The 2024 chemistry laureates

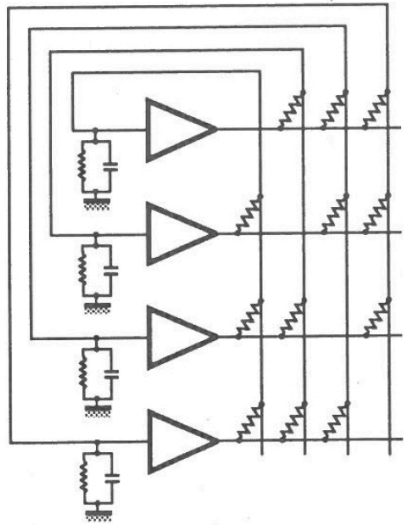
The Nobel Prize in Chemistry 2024 was awarded with one half to David Baker “for computational protein design” and the other half jointly to Demis Hassabis and John M. Jumper “for protein structure prediction”.

Demis Hassabis and John Jumper have successfully utilised artificial intelligence to predict the structure of almost all known proteins. David Baker has learned how to master life’s building blocks and create entirely new proteins.



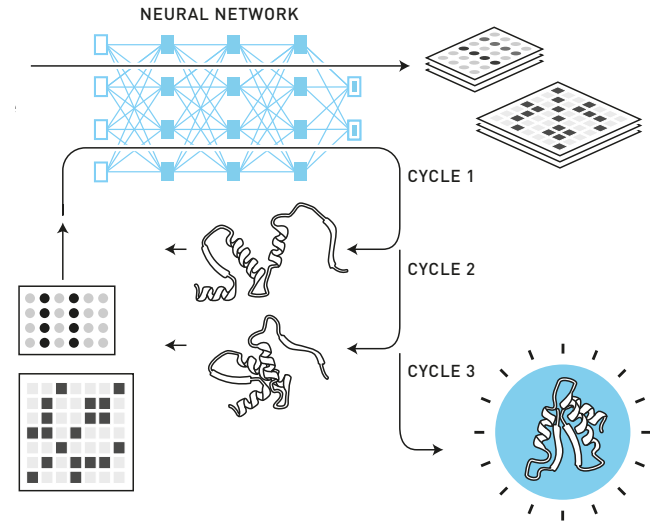
David Baker, Demis Hassabis and John Jumper. Ill. Niklas Elmehed © Nobel Prize Outreach

Analog electronic implementation of the Hopfield network, ~mid-1980s



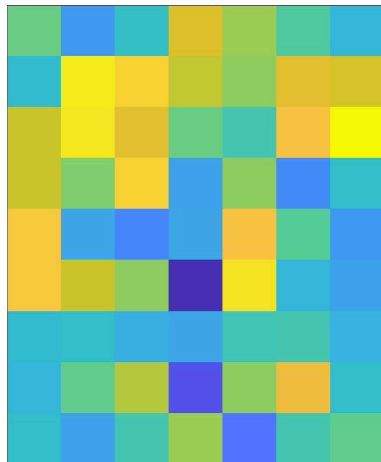
Hertz, Krogh
& Palmer

AlphaFold2 2021

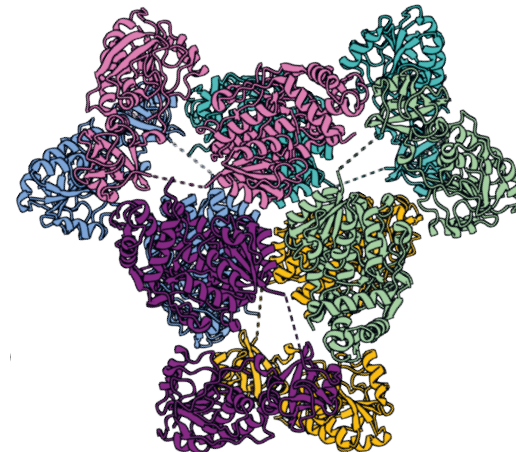


©Johan Jarnestad/The Royal Swedish Academy of Sciences

Associative Memory from noise to the “nearest” stored pattern



Enzyme causing antimicrobial resistance



©Terezia Kovalova/The Royal Swedish Academy of Sciences

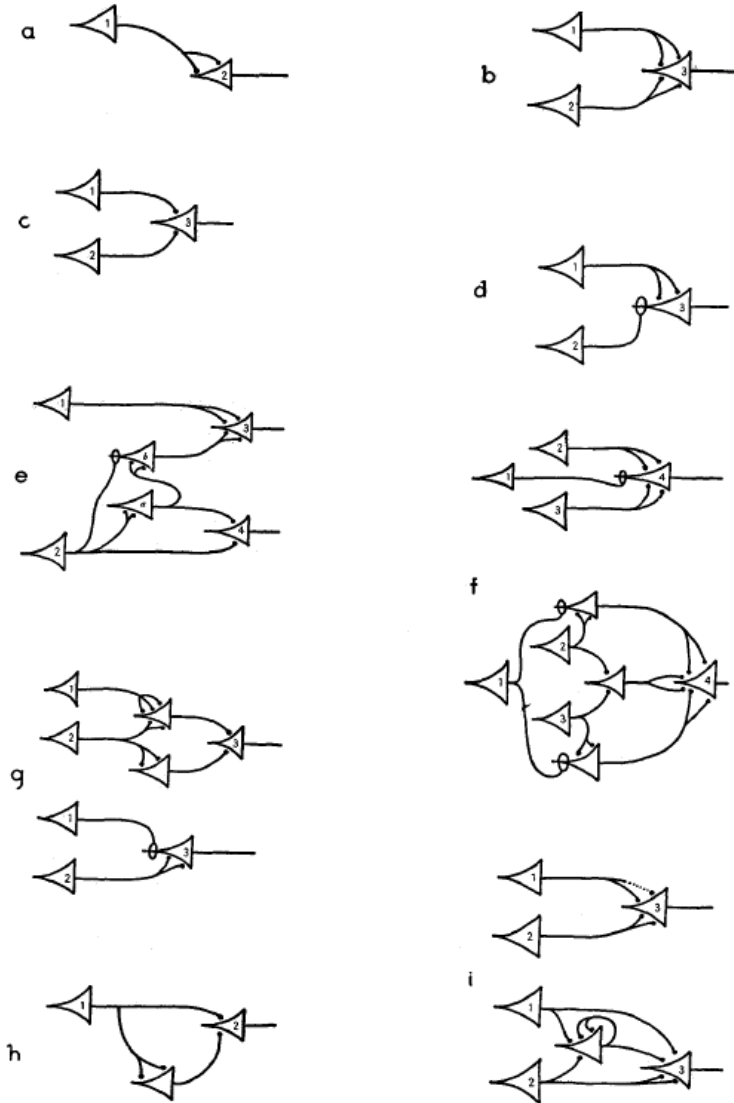


FIGURE 1

McCulloch and Pitts 1943

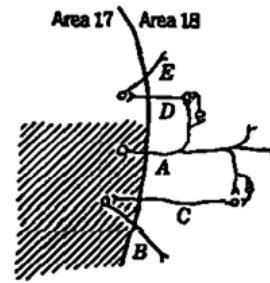


FIGURE 8

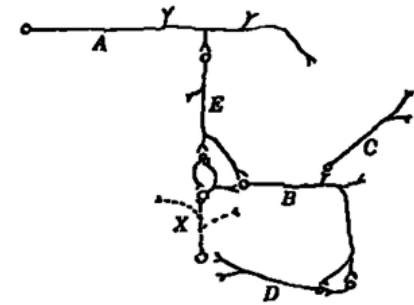


FIGURE 9

FIGURE 8 Cells A and B lie in a region of area 17 (shown by hatching) which is massively excited by an afferent stimulation. C is a cell in area 18 which leads back into 17. E is in area 17 but lies outside the region of activity. See text.

FIGURE 9 A, B, and C are cells in area 18 which are excited by converging fibers (not shown) leading from a specific pattern of activity in area 17. D, E, and X are, among the many cells with which A, B, and C have connections, ones which would contribute to an integration of their activity. See text.

Hebb 1949

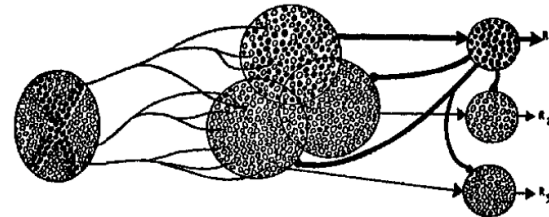


FIG. 3A. Predominant phase. Inhibitory connections are not shown. Solid black units are active.

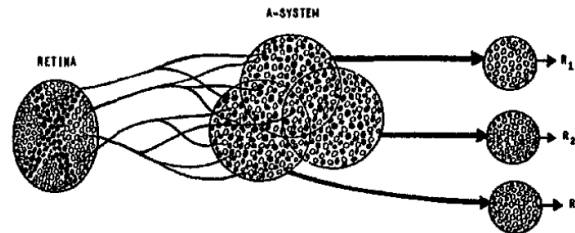
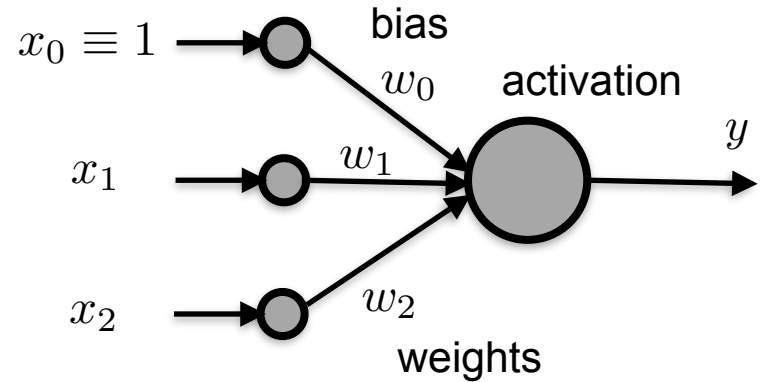
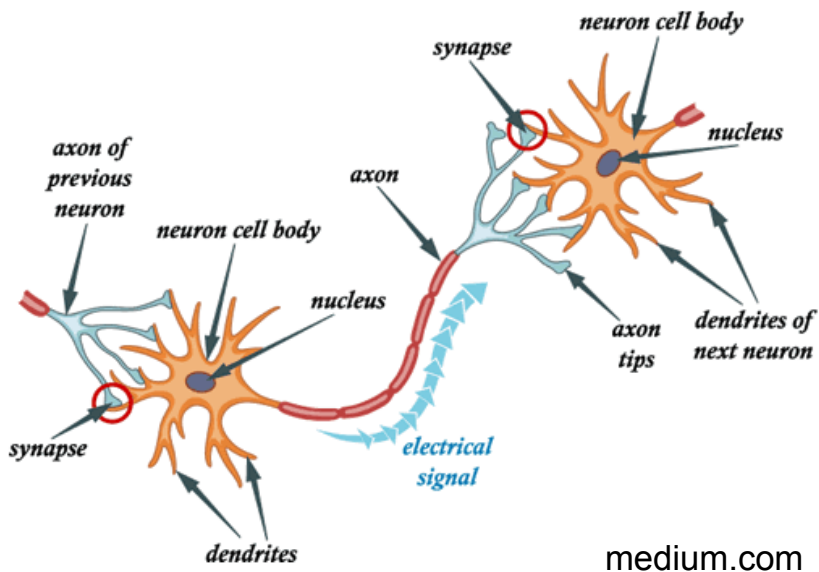


FIG. 3B. Postdominant phase. Dominant subset suppresses rival sets. Inhibitory connections shown only for R₁.

FIG. 3. Phases of response to a stimulus.

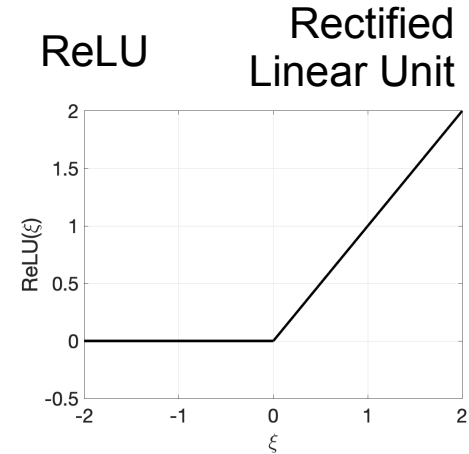
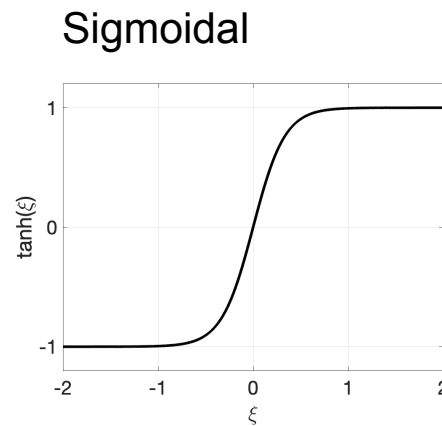
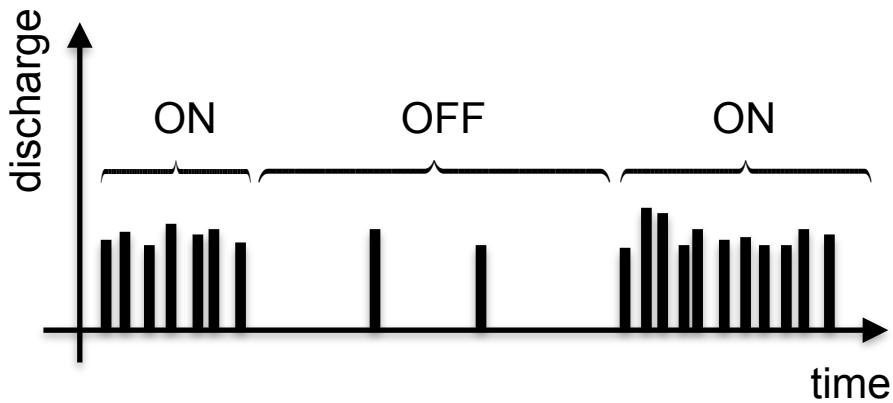
Rosenblatt
1958

Neurons: biological and artificial



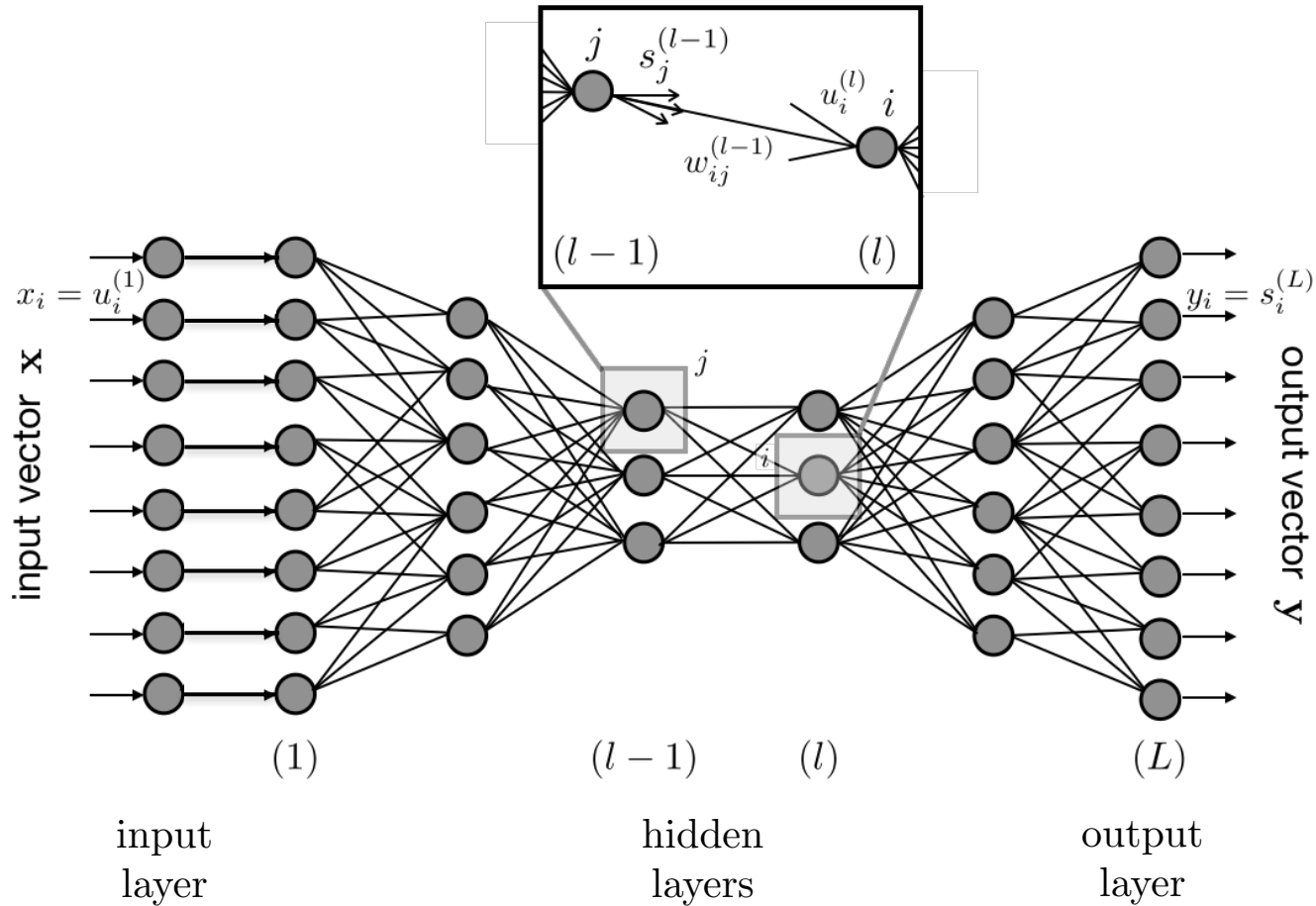
$$y = \sigma \left(\underbrace{w_0 + w_1x_1 + w_2x_2}_{\text{linear combination}} \right)$$

activation function



Linearly non-separable problem? => multi-layer perceptron

Deep Neural Networks



Training deep neural networks

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

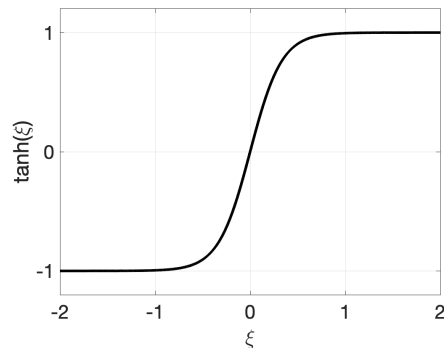
* Institute for Cognitive Science, C-015, University of California,
San Diego, La Jolla, California 92093, USA

† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

the “backprop” paper
Nature 323:533, 1986

essentially, the chain rule of differentiation!
progenitor to modern “automatic differentiation”

Sigmoidal activation



Rectified Linear Units Improve Restricted Boltzmann Machines

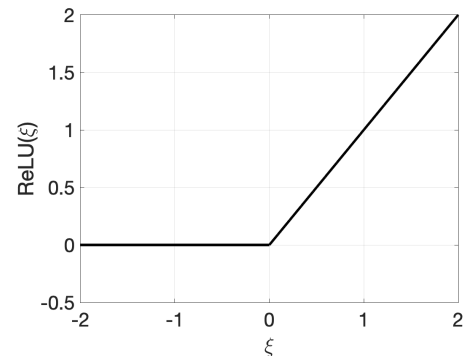
Vinod Nair
Geoffrey E. Hinton

Department of Computer Science, University of Toronto, Toronto, ON M5S 2G4, Canada

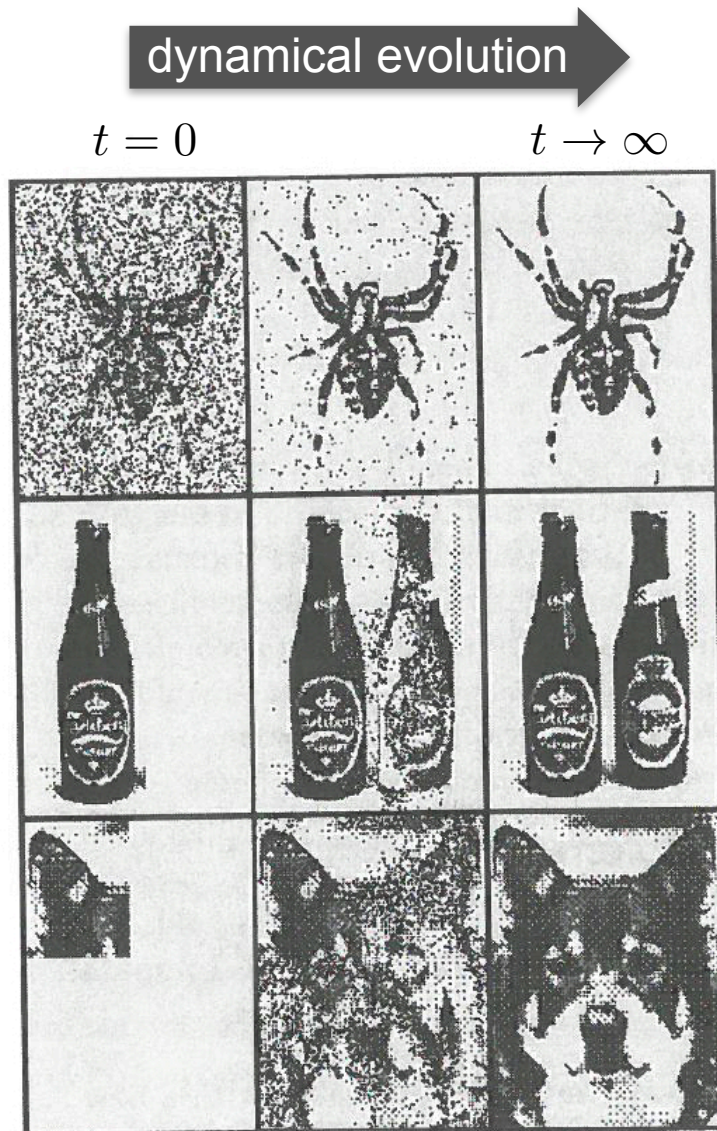
VNAIR@CS.TORONTO.EDU
HINTON@CS.TORONTO.EDU

the “ReLU” paper
Int. Conf. Machine Learning 2010

ReLU Rectified Linear Unit



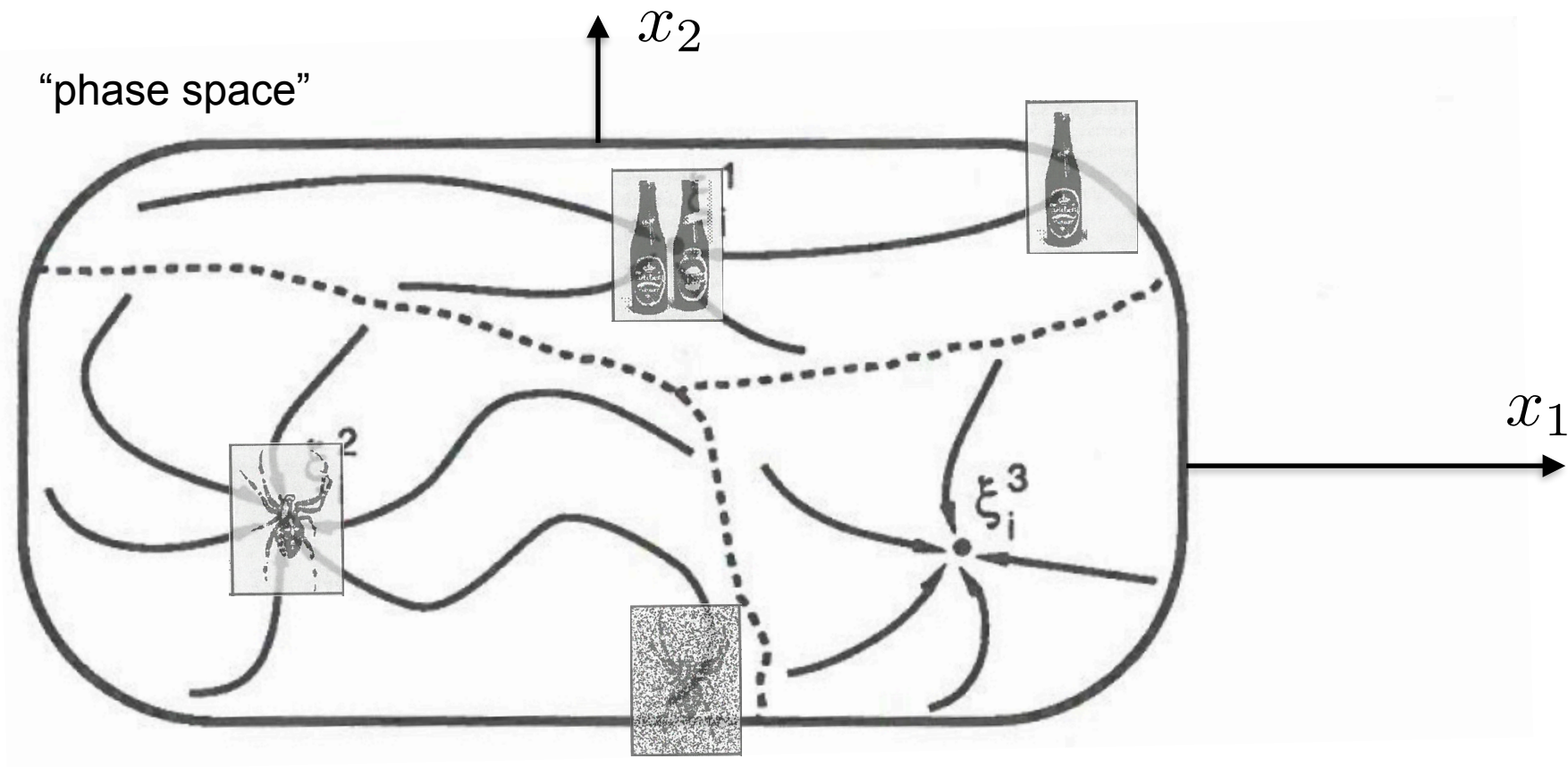
Dynamic associative memory



We can think of this memory as a *dynamical system*: it is initialized at $t=0$ with an imperfect version of the content; then, through dynamical interactions between pixels, as $t \rightarrow \infty$ it *evolves* to a final recall state that matches the content perfectly.

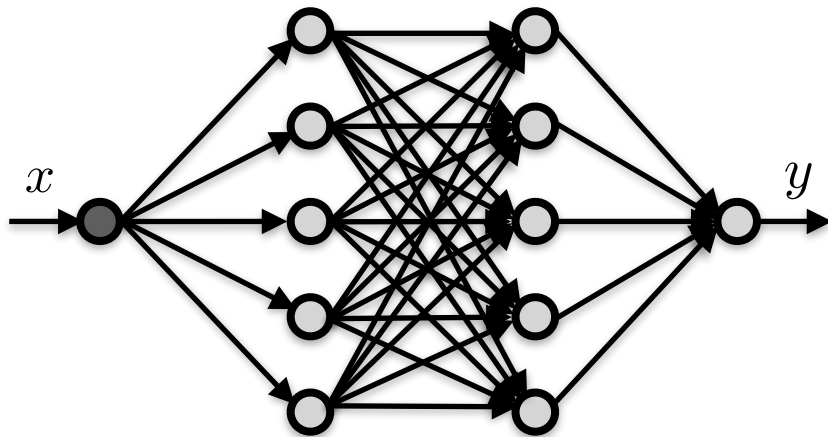
FIGURE 2.1 Example of how an associative memory can reconstruct images. These are binary images with 130×180 pixels. The images on the right were recalled by the memory after presentation of the corrupted images shown on the left. The middle column shows some intermediate states. A sparsely connected Hopfield network with seven stored images was used.

Associative Memory as a Dynamical Attractor



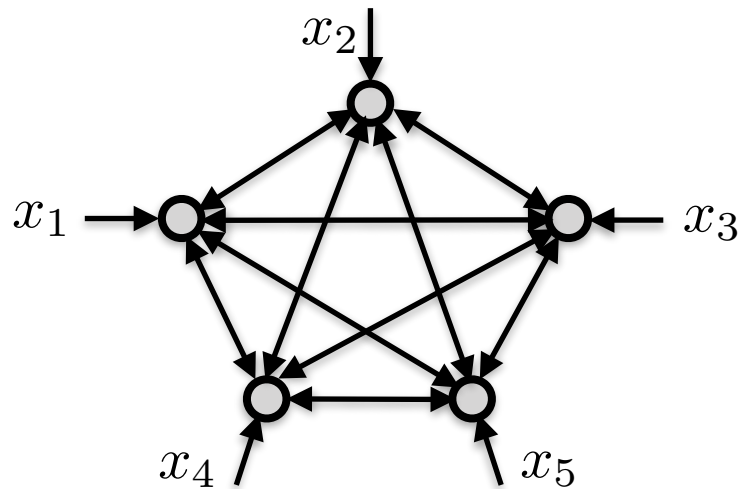
The recurrent neural network is initialized with a partial (or distorted) memory, then follows a trajectory to an equilibrium representing the correct complete memory.

Neural Nets: feedforward vs recurrent



Feedforward - Static

Input x propagates through, produces output y

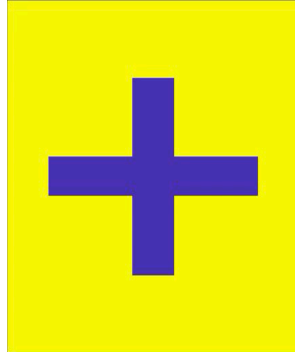


Recurrent - Dynamical

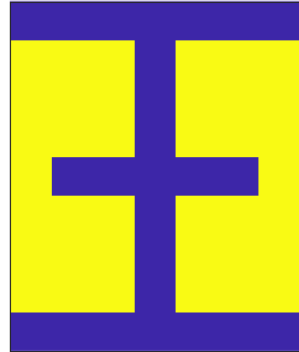
Inputs are initialized as states x_1, \dots, x_5 , dynamical interactions evolve, outputs are the same units' states x_1, \dots, x_5 , after sufficiently long time (*i.e.*, in equilibrium)

Continuous binary associative memory: Hopfield network dynamics in a 63-state system

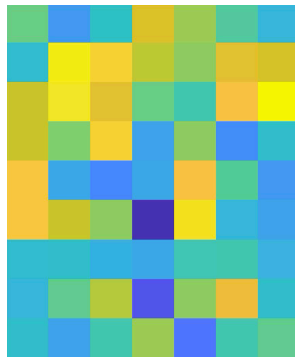
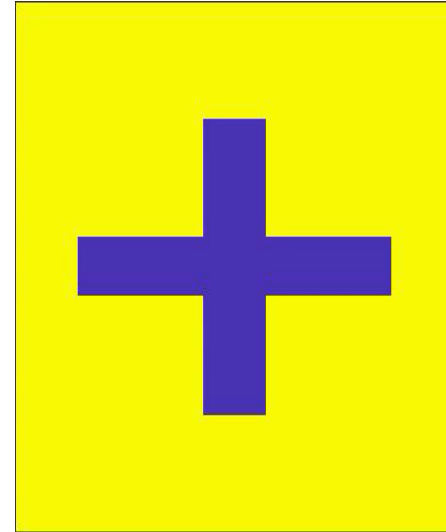
states $N = 9 \times 7 = 63$, $p = 6$.



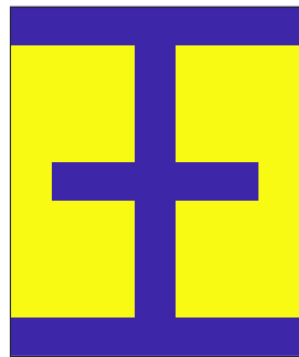
$\mathbf{V}(t = 0)$



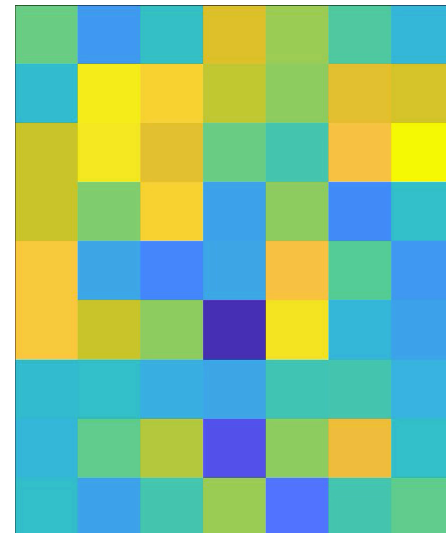
$\mathbf{V}(t \rightarrow \infty)$



$\mathbf{u}(t = 0)$



$\mathbf{u}(t \rightarrow \infty)$



Statistical mechanics of Hopfield networks

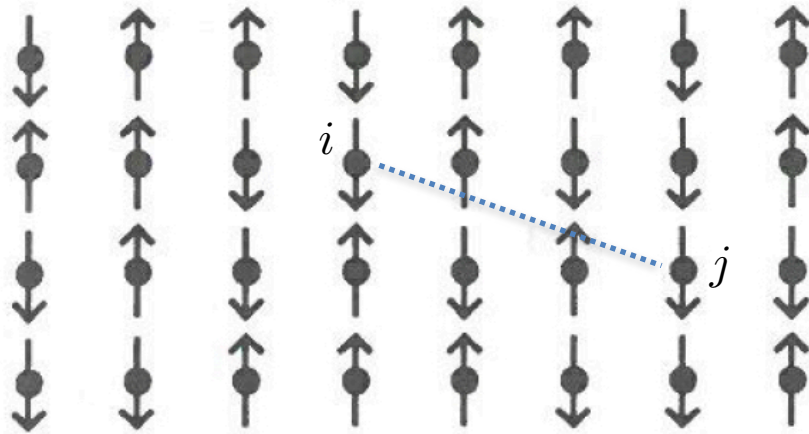


FIGURE 2.7 A very simplified picture of a magnetic material described by an Ising model.

w_{ij} : interaction strength between magnets i and j

in physical Ising model: $w_{ij} \sim \frac{1}{r_{ij}}$; ... and, by convention, nearest neighbors only.

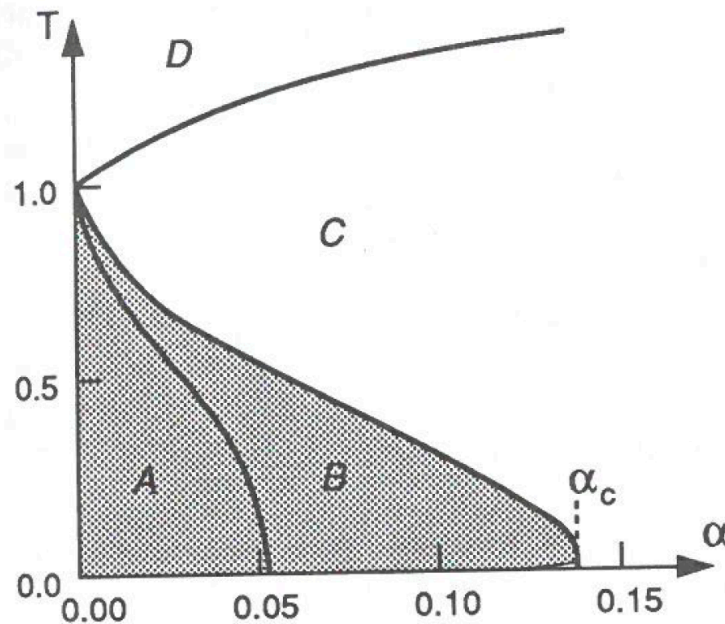
in Hopfield model: w_{ij} learned from the stored patterns

equilibrium condition:

magnet spin $x_i = \tanh \left(\beta \sum_j w_{ij} x_j \right)$ sigmoidal! $\beta = \frac{1}{T}$, T : “temperature”

Hertz, Krogh and Palmer p. 25

Capacity of the Hopfield network



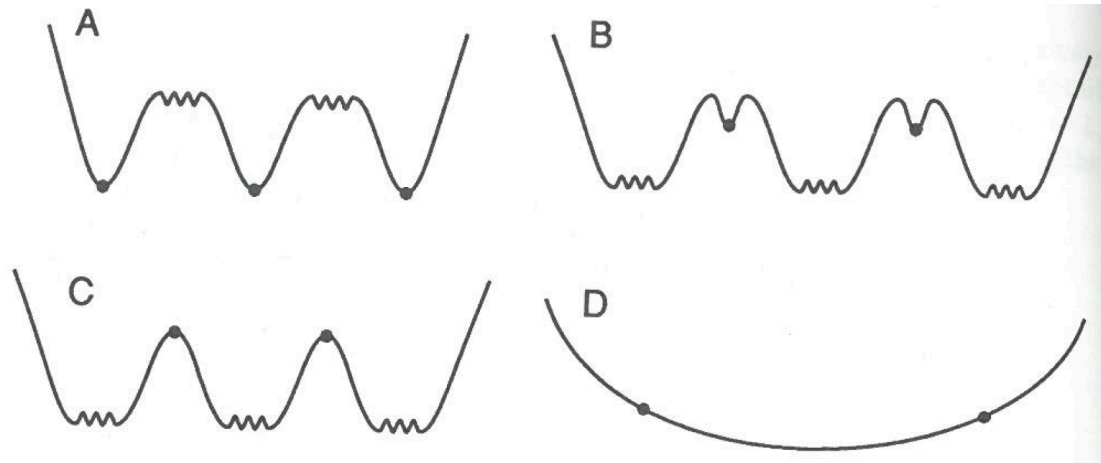
p patterns, N neurons

$$\alpha = \frac{p}{N} \quad \text{network capacity}$$

$\alpha_c \approx 0.138$ maximum capacity

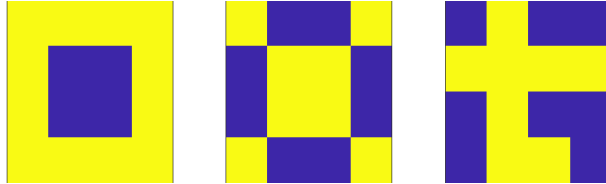
FIGURE 2.17 The phase diagram obtained by Amit et al.. The desired memory states are only stable in the shaded region.

energy
landscapes



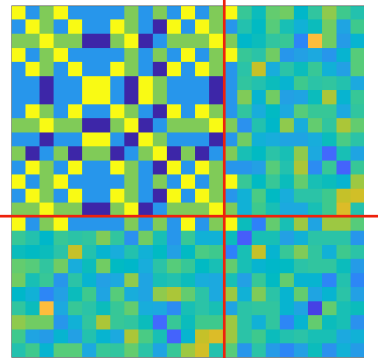
Boltzmann machine

Intent: starting from 4x4 pattern of random noise, and following Ising stochastic dynamics, to converge to one of the three patterns below with given probabilities p_o , p_x , p_t



Implemented as a Boltzmann machine with 16 visible units, 9 hidden units, and $p_o = p_x = p_t = 1/3$. Temperature $T=1$.

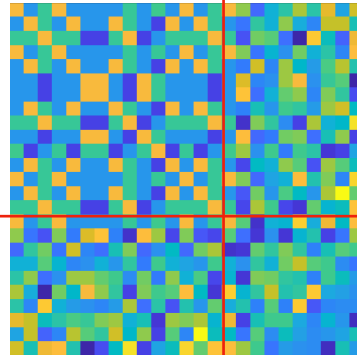
visible to visible visible to hidden



visible to hidden hidden to hidden

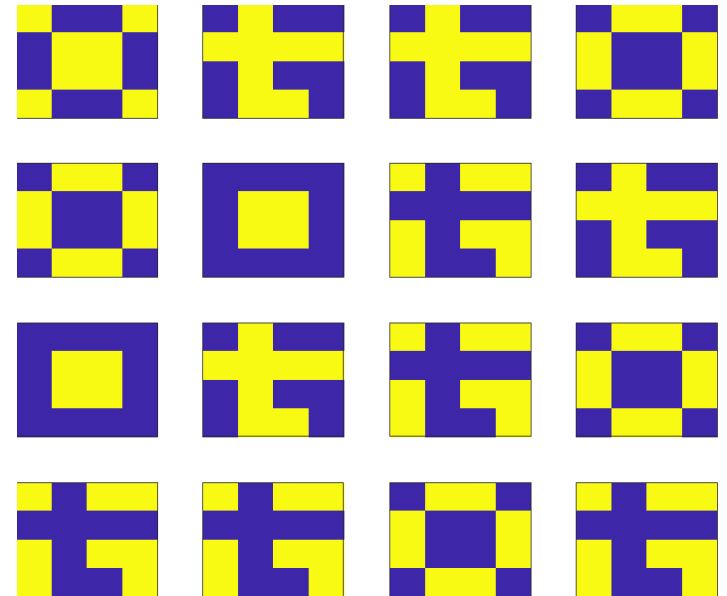
Initial w_{ij}

visible to visible visible to hidden



visible to hidden hidden to hidden

Trained w_{ij}



visible to visible: Hopfield-like
others: white but symmetric

Sixteen runs, initialized with white noise on visible + hidden units, shown to stabilize randomly at one of the three stored patterns

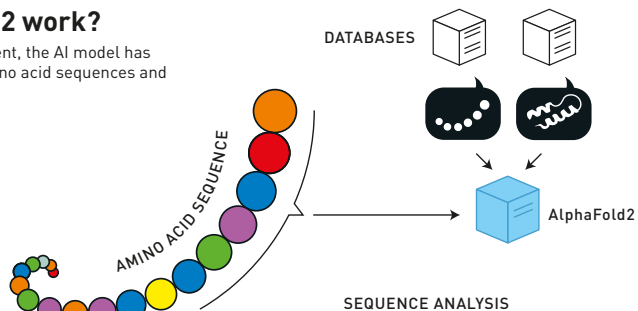
AlphaFold

How does AlphaFold2 work?

As part of AlphaFold2's development, the AI model has been trained on all the known amino acid sequences and determined protein structures.

1. DATA ENTRY AND DATABASE SEARCHES

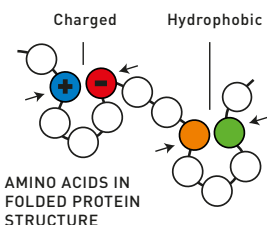
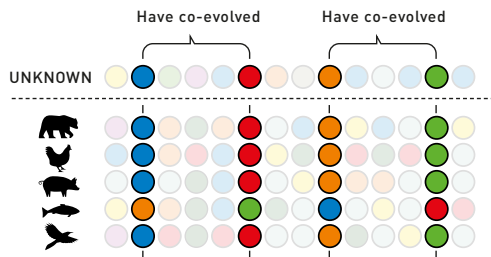
An amino acid sequence with unknown structure is fed into AlphaFold2, which searches databases for similar amino acid sequences and protein structures.



2. SEQUENCE ANALYSIS

The AI model aligns all the similar amino acid sequences – often from different species – and investigates which parts have been preserved during evolution.

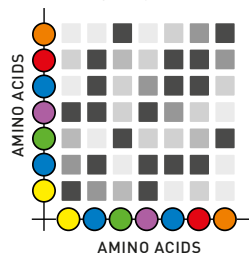
In the next step, AlphaFold2 explores which amino acids could interact with each other in the three-dimensional protein structure. Interacting amino acids co-evolve. If one is charged, the other has the opposite charge, so they are attracted to each other. If one is replaced by a water-repellent (hydrophobic) amino acid, the other also becomes hydrophobic.



Using this analysis, AlphaFold2 produces a distance map that estimates how close amino acids are to each other in the structure.

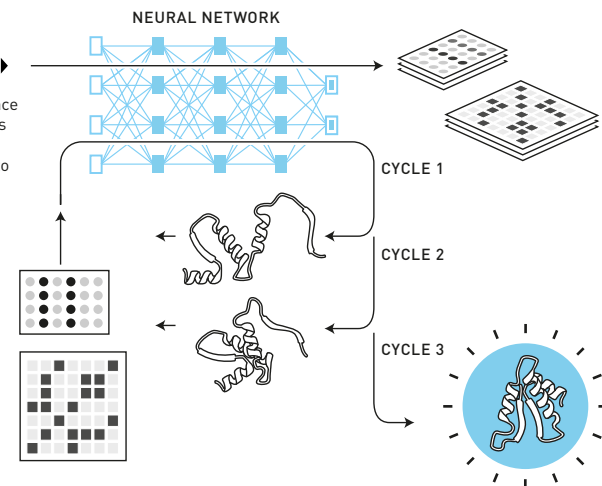


DISTANCE MAP



3. AI ANALYSIS

Using an iterative process, AlphaFold2 refines the sequence analysis and distance map. The AI model uses neural networks called transformers, which have a great capacity to identify important elements to focus on. Data about other protein structures – if they were found in step 1 – is also utilised.



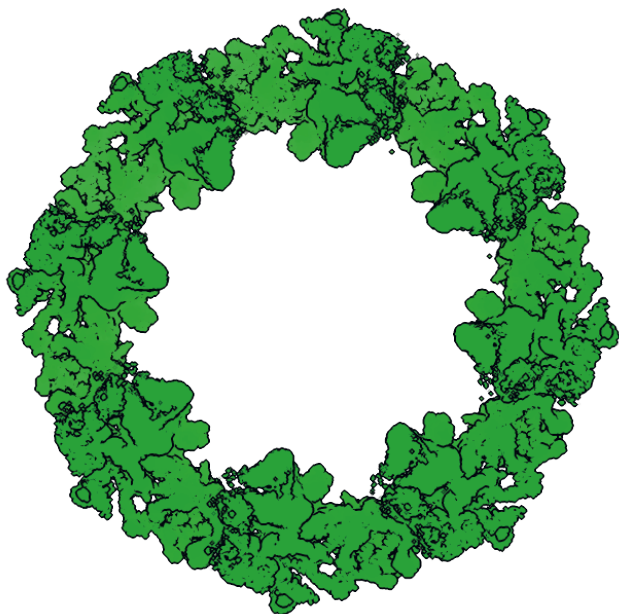
4. HYPOTHETICAL STRUCTURE

AlphaFold2 puts together a puzzle of all the amino acids and tests pathways to produce a hypothetical protein structure. This is re-run through step 3. After three cycles, AlphaFold2 arrives at a particular structure. The AI model calculates the probability that different parts of this structure correspond to reality.

©Johan Jarnestad/The Royal Swedish Academy of Sciences

<https://www.nobelprize.org/uploads/2024/10/advanced-chemistryprize2024.pdf>

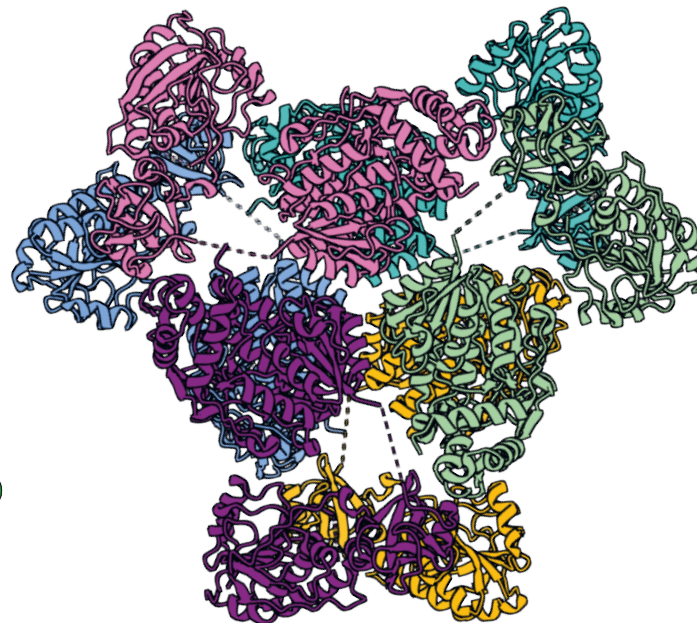
Natural and novel proteins & enzymes



2022: Part of a huge molecular structure in the human body. More than a thousand proteins form a pore through the membrane surrounding the cell nucleus.



2022: Natural enzymes that can decompose plastic. The aim is to design proteins that can be used to recycle plastic.

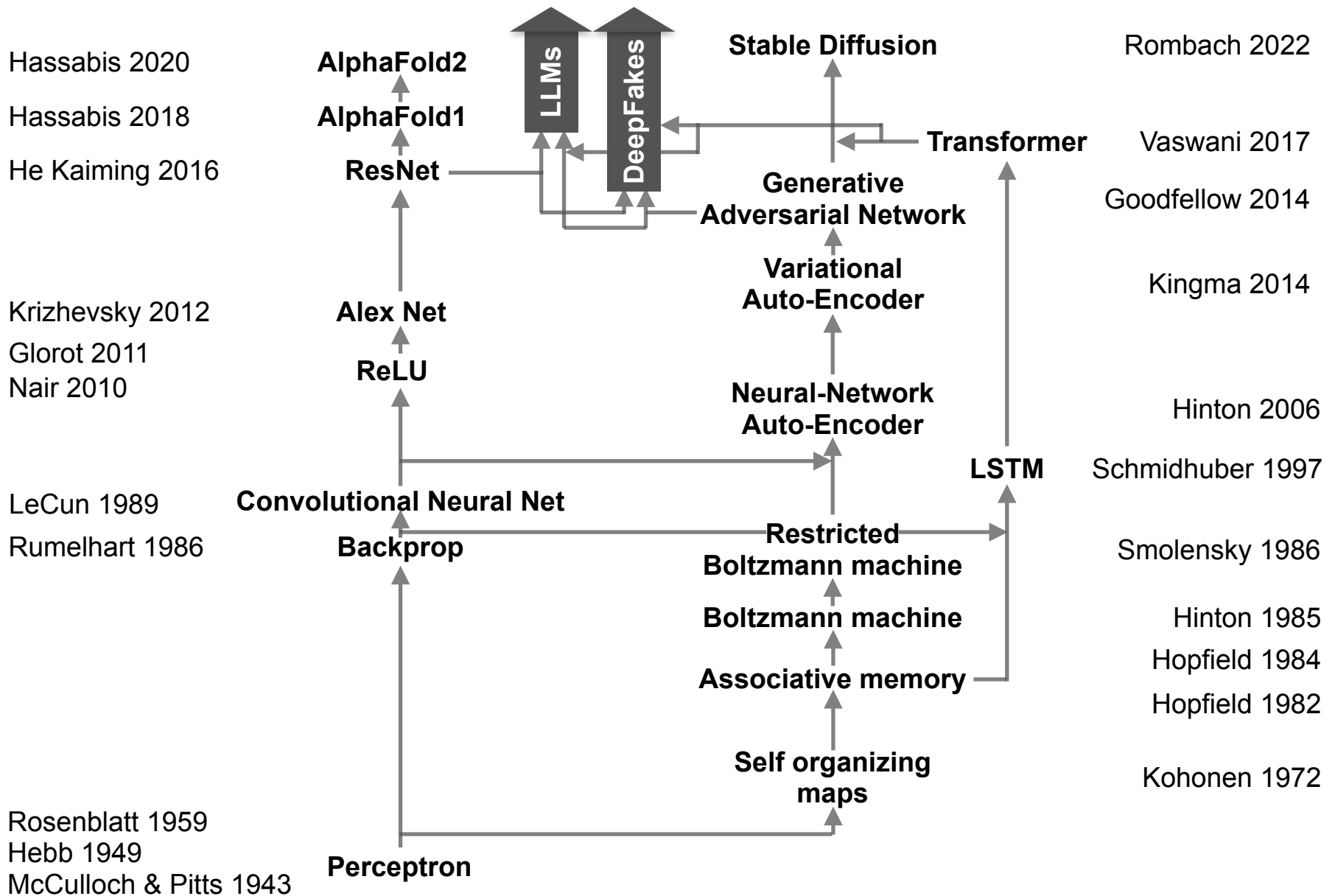


2023: A bacterial enzyme that causes antibiotic resistance. The structure is important for discovering ways of preventing antibiotic resistance.

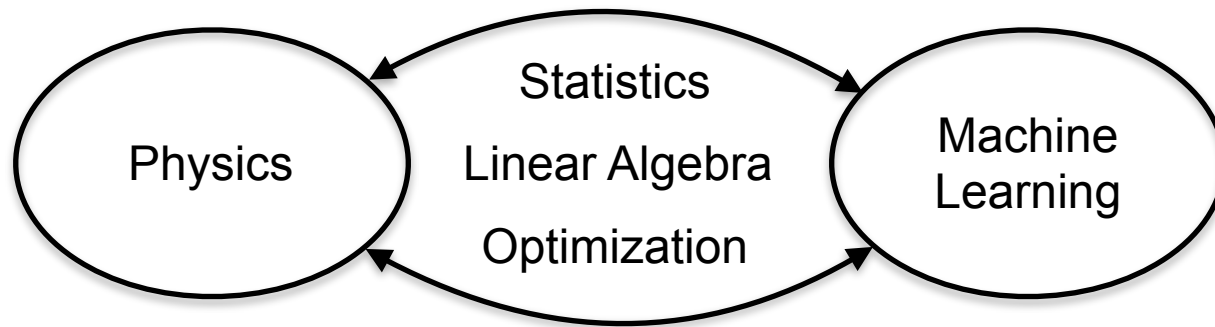
©Terezia Kovalova/The Royal Swedish Academy of Sciences

<https://www.nobelprize.org/uploads/2024/10/advanced-chemistryprize2024.pdf>

some partial lineages for Machine Learning



the 2.c01/2.c51 compass



- Physics principles to augment our understanding of how Machine Learning algorithms are designed and function
- Machine Learning algorithms to augment our ability to model, predict, and control physical systems