Working towards human-level intelligence…

Dileep George

Vicarious

Intelligence:

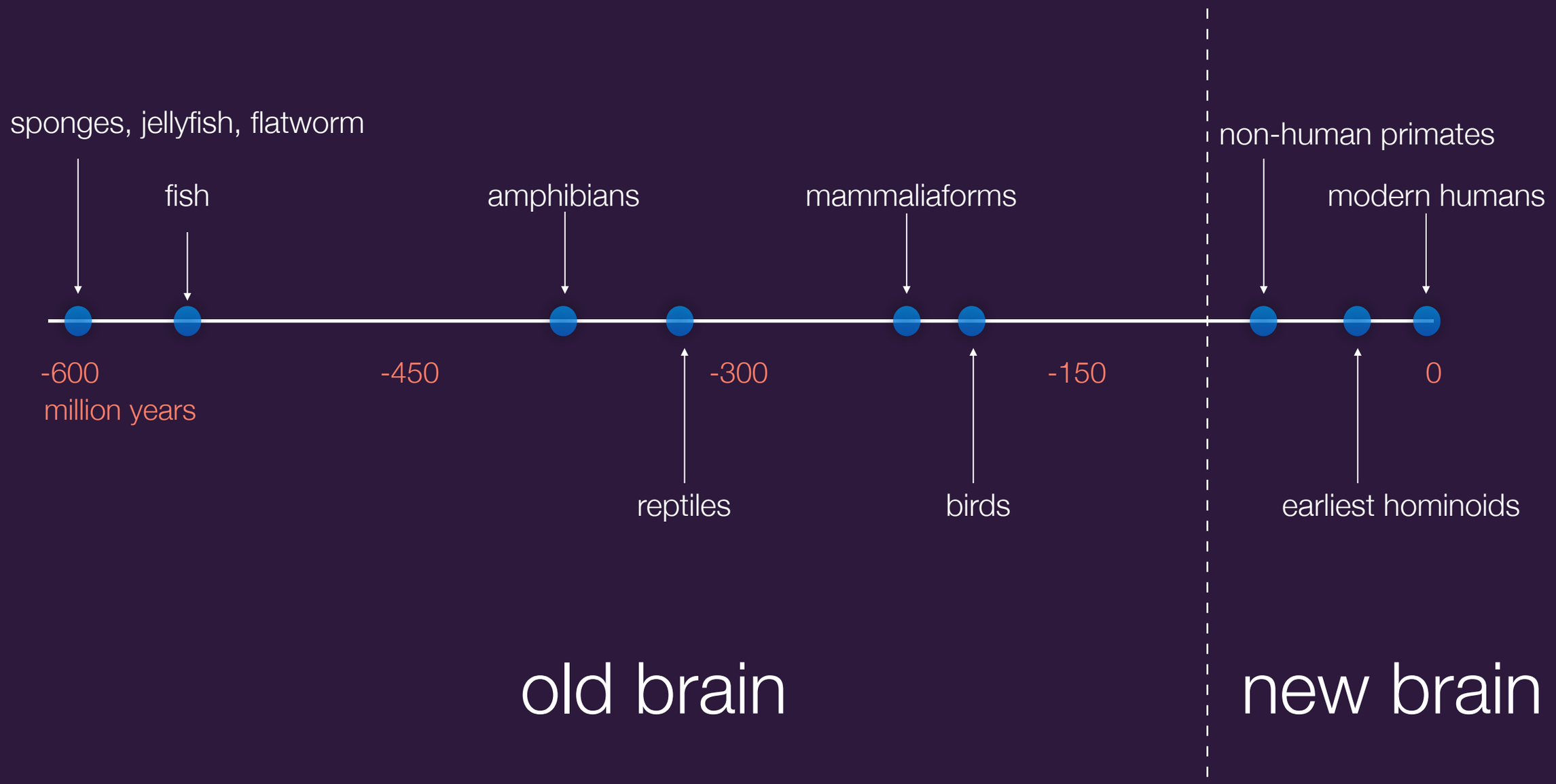
Ability to model the world and to act on it

Intelligence:

Ability to model the world and to act on it

It is possible to act on the world without modeling it

Old brain


sponges, jellyfish, flatworm
fish
amphibians
mammaliaforms
non-human primates
modern humans
-600
million years
-450
-300
-150
0
reptiles
birds
earliest hominoids
old brain
new brain

Dinosaurs appeared

Dinosaurs became extinct

Homo sapiens appeared

Jan 1st

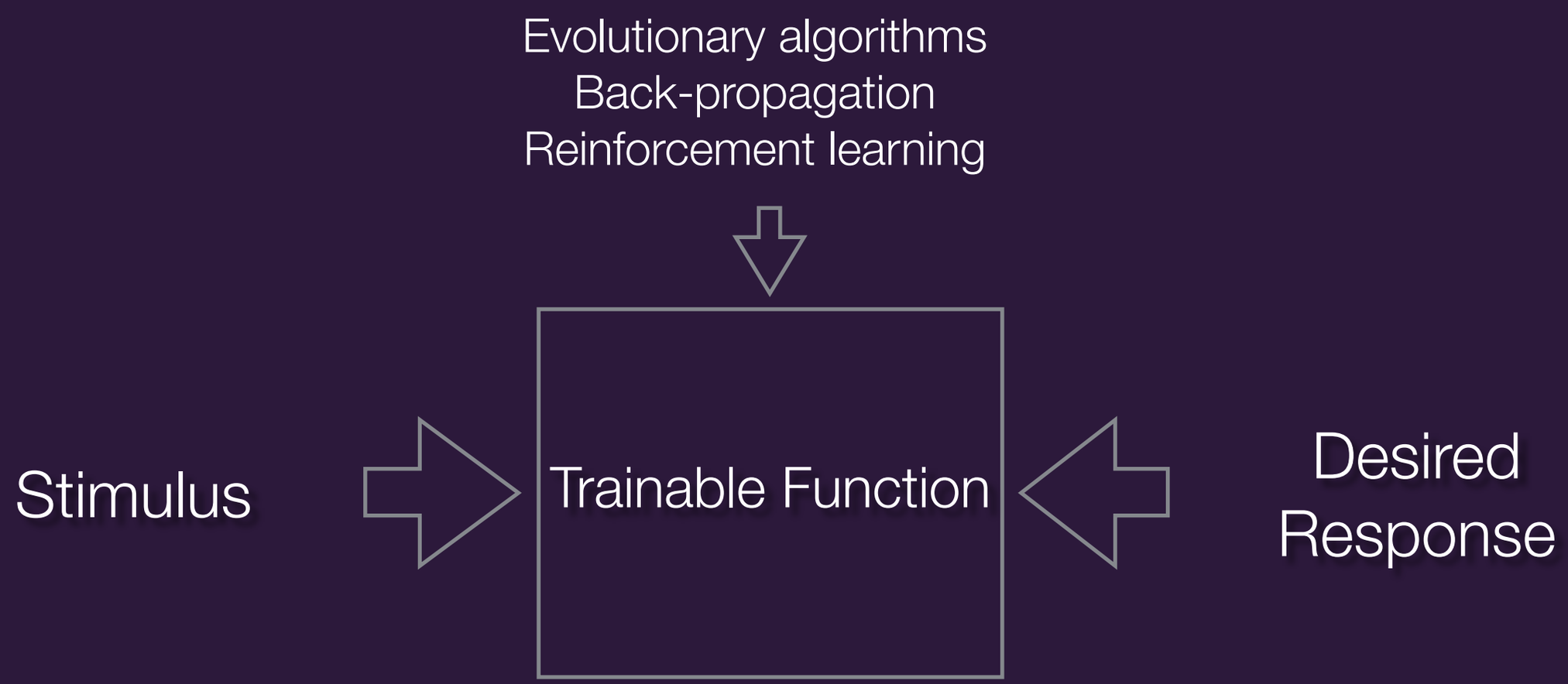
Sep 21st

Dec 31st

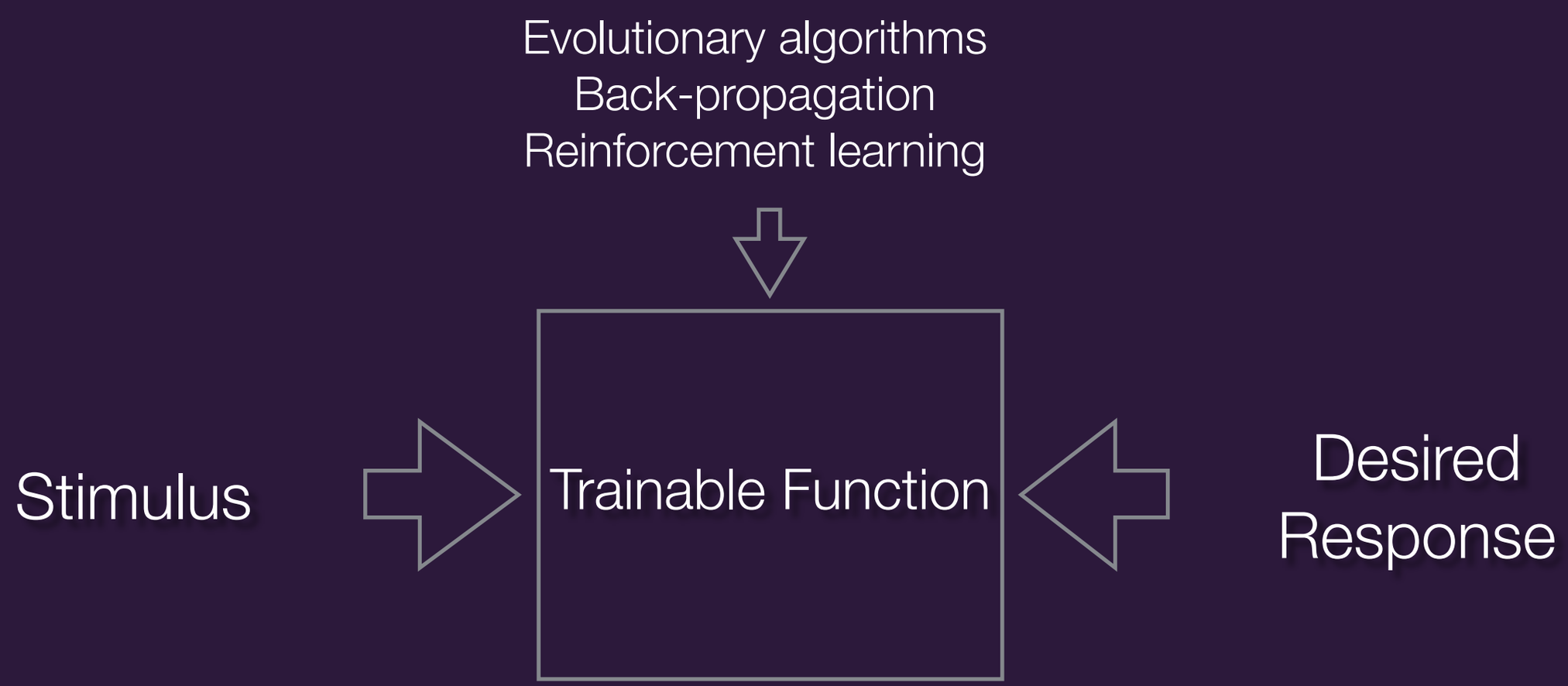
Old brain was very successful. Dinosaurs had walnut-sized brains

Old-brain style (model-free) learning

Old-brain style (model-free) learning



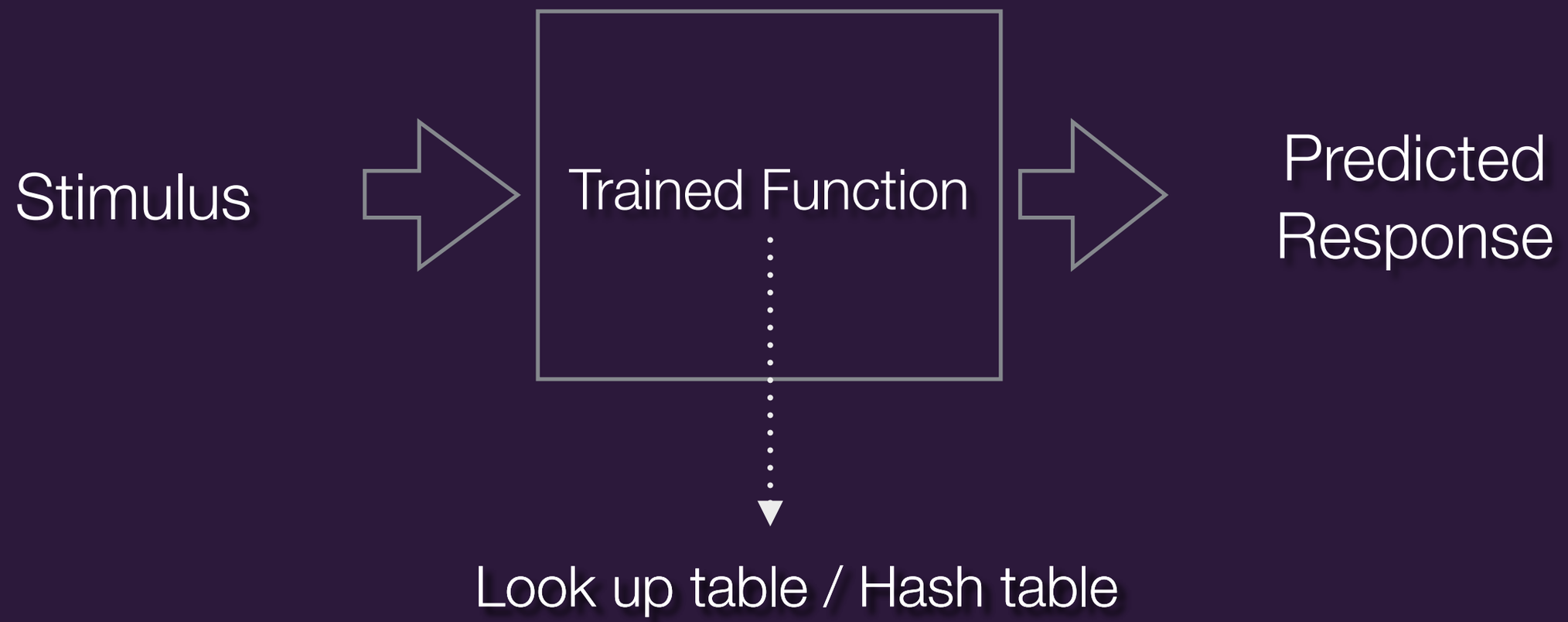
Old-brain style (model-free) learning



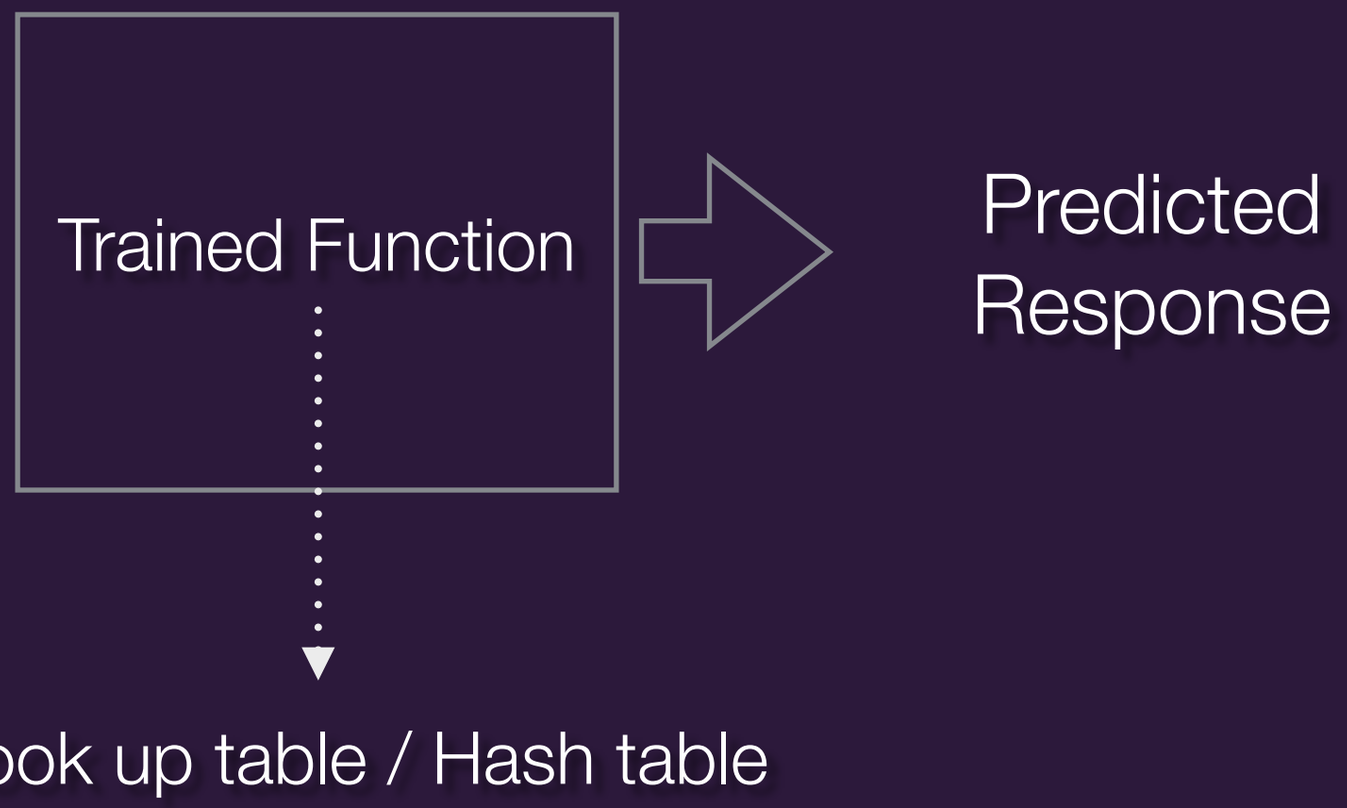
Very large repertoire of training data

Old-brain style (model-free) learning

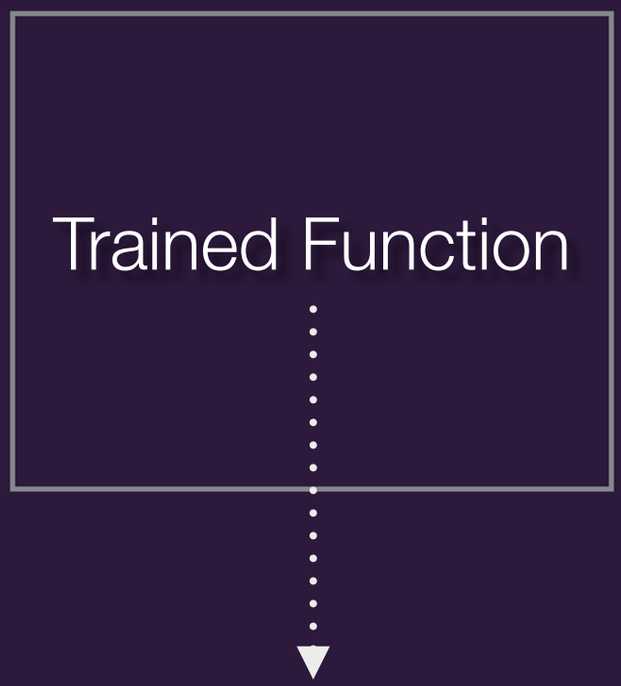
Old-brain style (model-free) learning



Old-brain style (model-free) learning

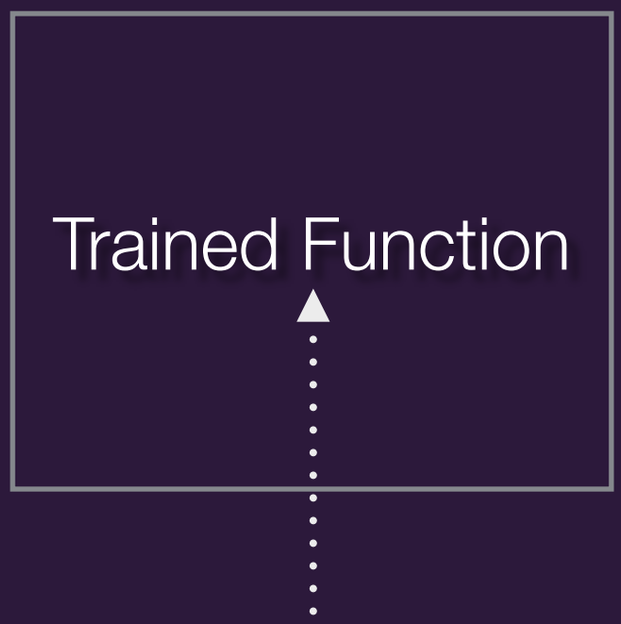


Old-brain style (model-free) learning



Look up table / Hash table

Old-brain style (model-free) learning



Black box with limited capability for thinking, imagination, creativity or planning

Many modern methods are still like the old brain

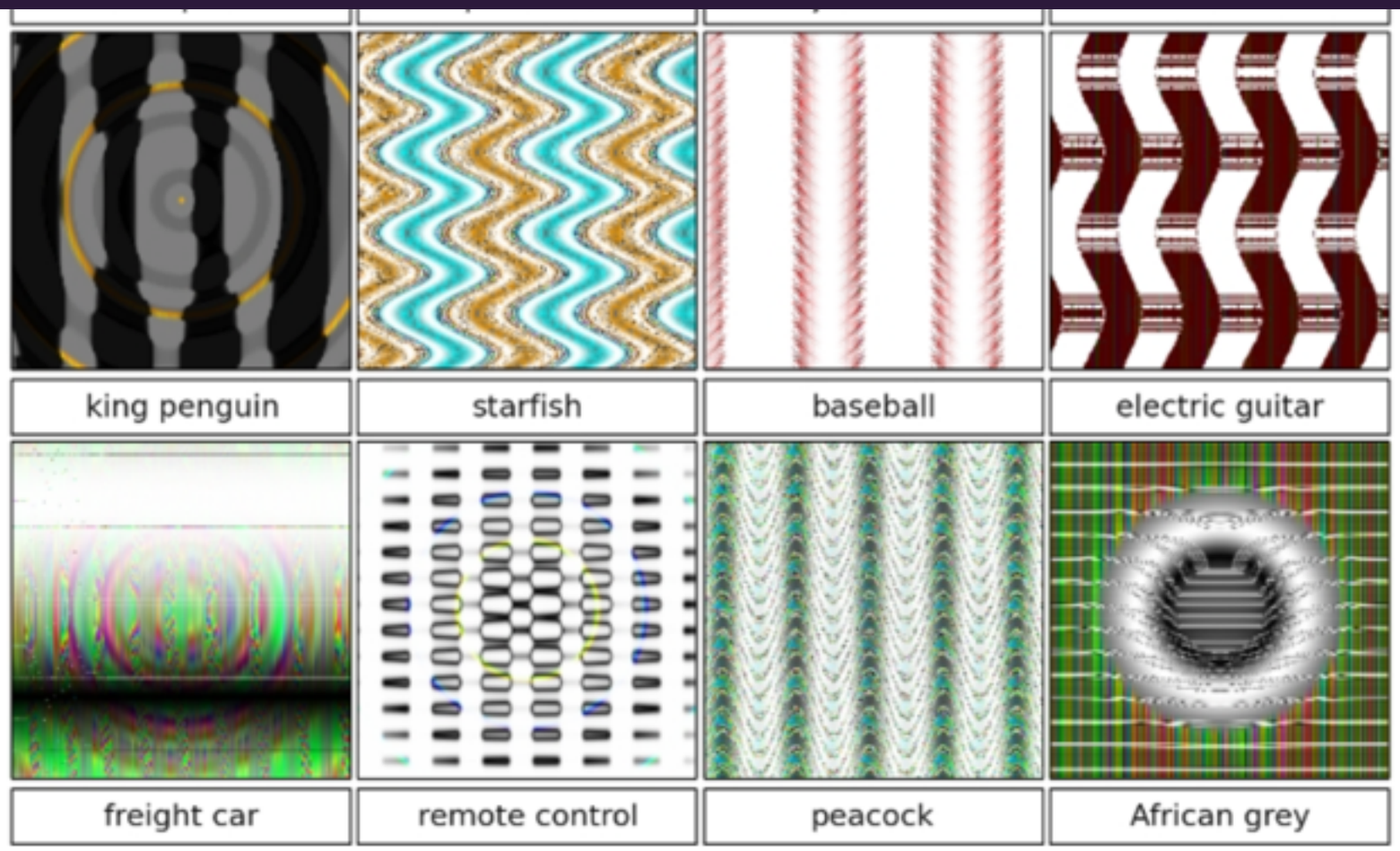
- Extremely large number of training examples
- Inscrutable, black box classifiers
- Not generative. Models lack explanatory power

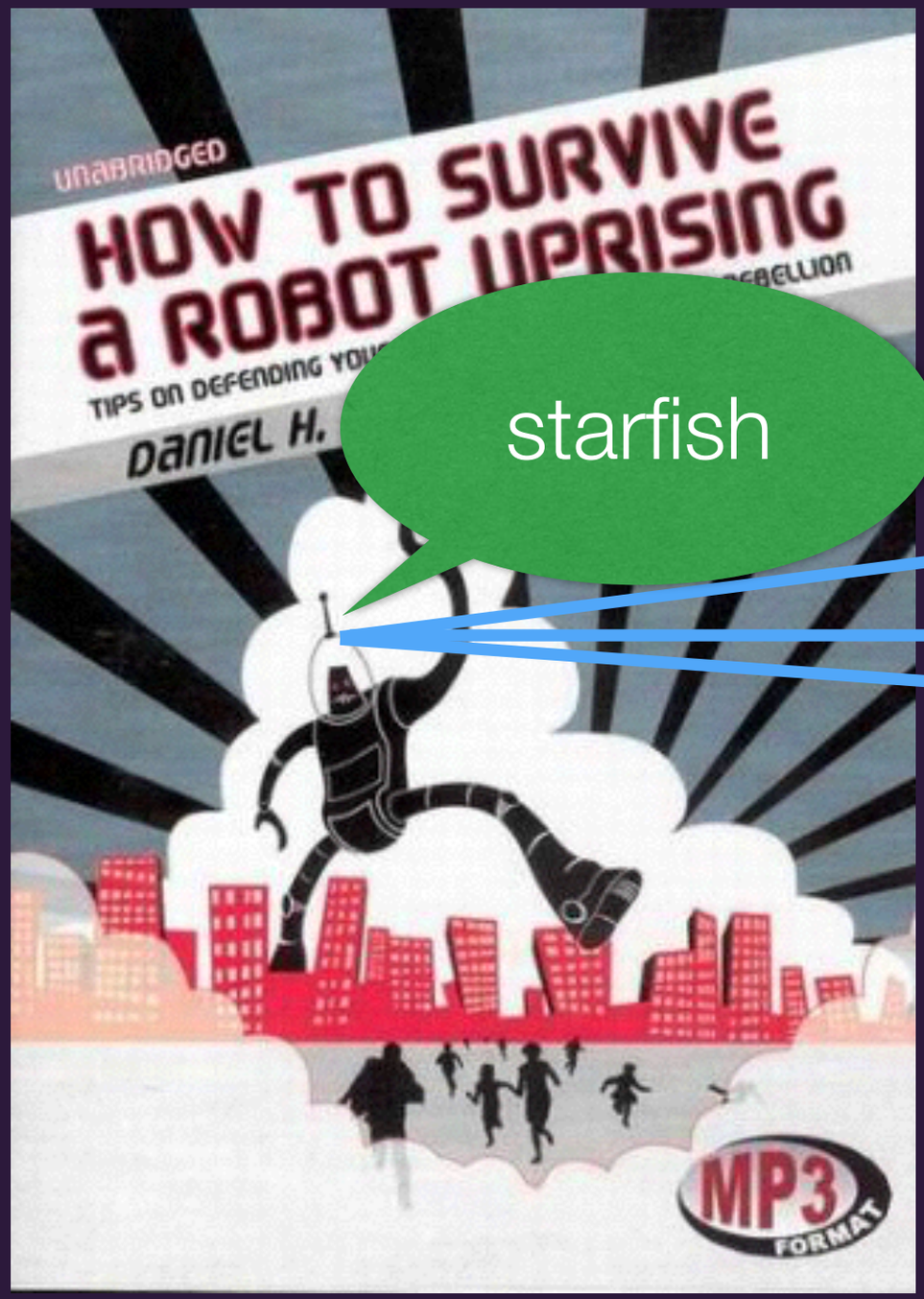
Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images

Anh Nguyen
University of Wyoming
anguyen8@uwyo.edu

Jason Yosinski
Cornell University
yosinski@cs.cornell.edu

Jeff Clune
University of Wyoming
jeffclune@uwyo.edu



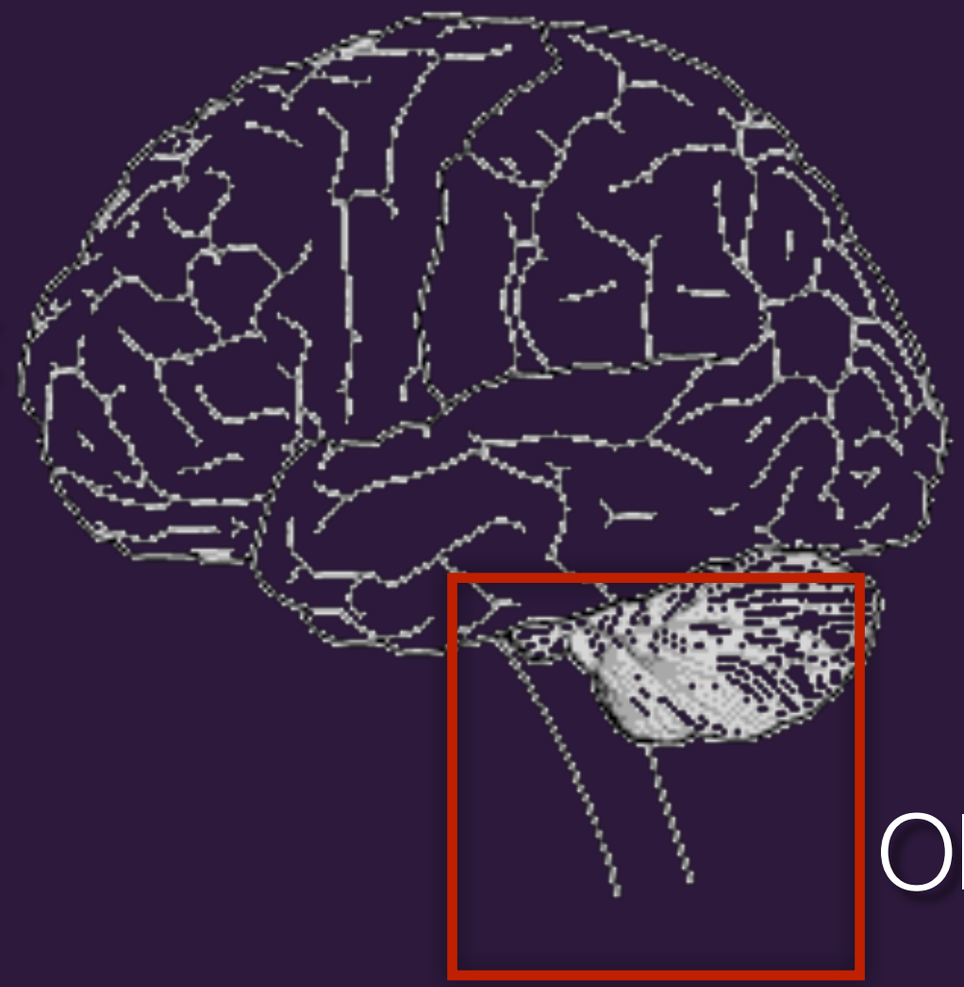

UNABRIDGED
HOW TO SURVIVE
a ROBOT UPRISING
TIPS ON DEFENDING YOU
REBELLION
DANIEL H.
MP3
FORMAT

starfish

Perhaps, the brain can tell us more about how to build truly intelligent machines

vicarious

Neocortex



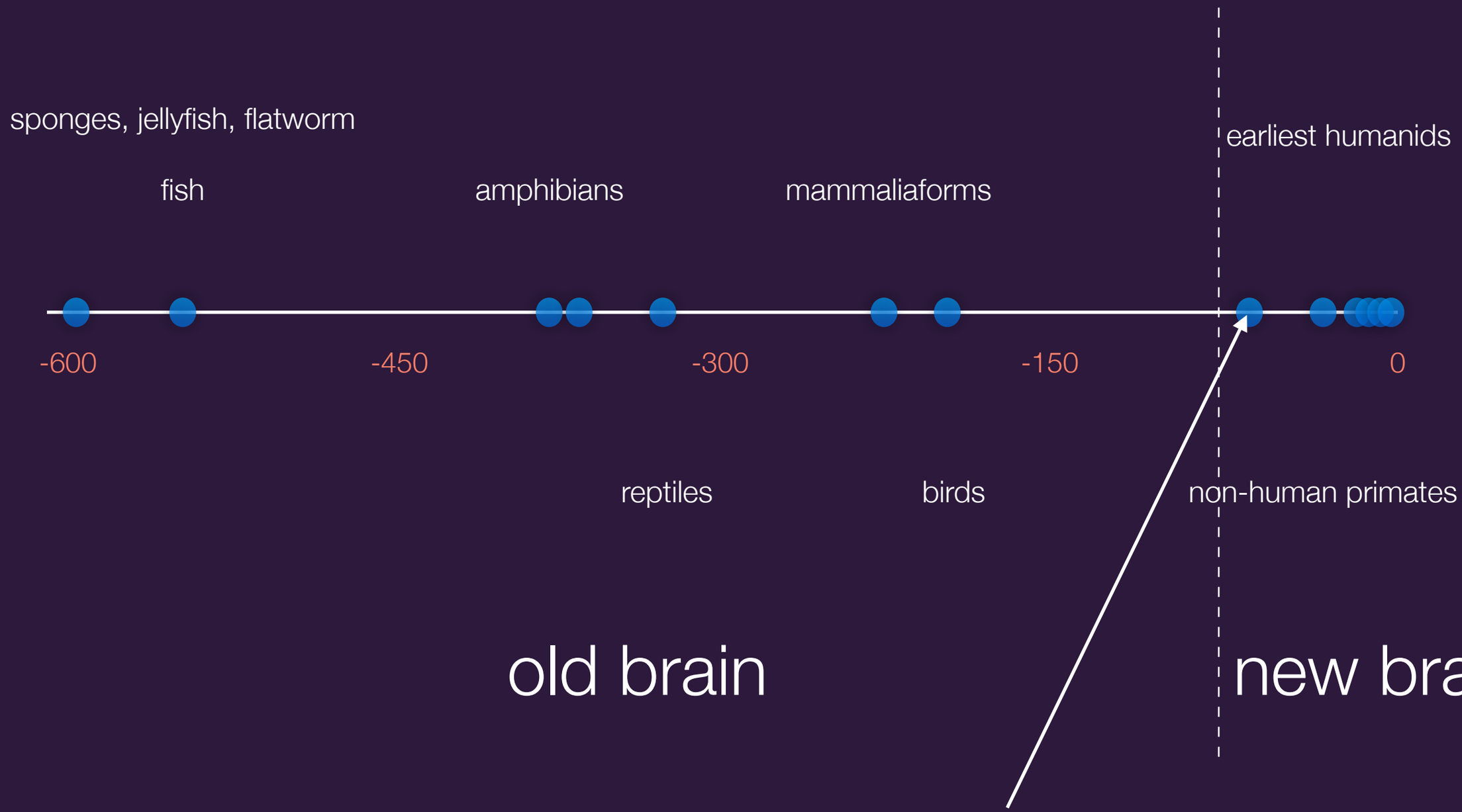
Old brain

vicarious

Assumptions

Learning efficiency

Generality



By discovering neocortex, nature stumbled upon the ‘magic middle’ of learning architectures

This ‘magical architecture’ is :

General enough to solve multiple problems like vision, audition, somatosensory etc.

But specific enough to learn efficiently.

(This also implies that the same set of architectural assumptions work well for multiple domains)

Your brain is not very good at recognizing QR codes

Neural nets

Flat

RBMs

Conv-nets

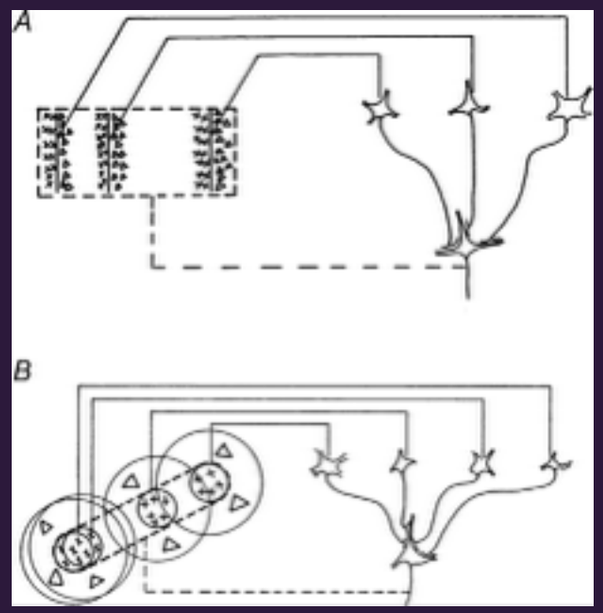
Neo-Cortex

Domain specific architectures

Less structure

More structure

Hubel & Wiesel



Complex cells (Feature pooling)

Simple Cells (Feature detection)

RECEPTIVE FIELDS, BINOCULAR INTERACTION AND FUNCTIONAL ARCHITECTURE IN THE CAT'S VISUAL CORTEX

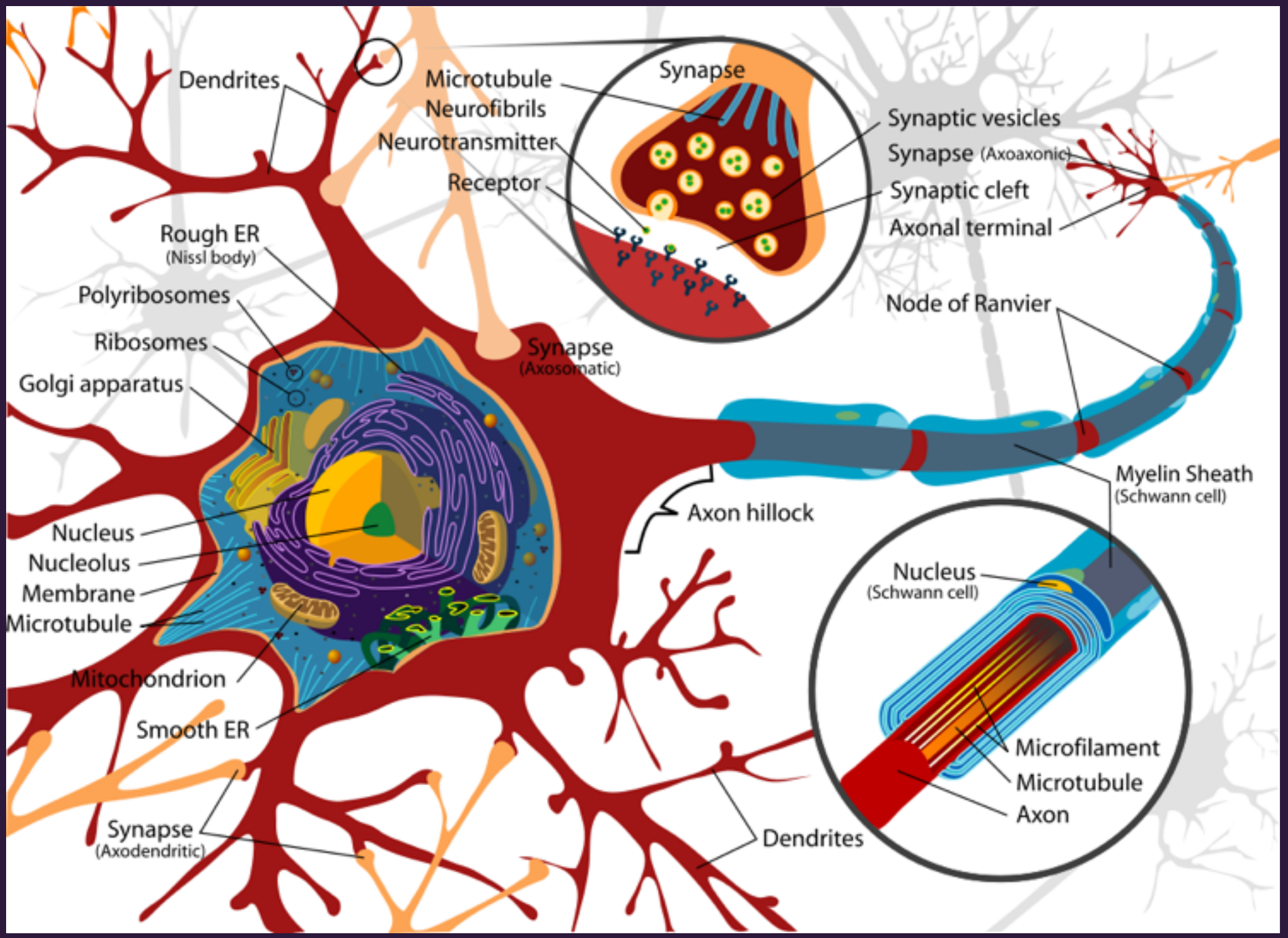
By D. H. HUBEL and T. N. WIESEL

From the Neurophysiology Laboratory, Department of Pharmacology Harvard Medical School, Boston, Massachusetts, U.S.A.

(*Received* 31 *July* 1961)

Neuroscience has had many more advances since 1961.
Can't we use that?

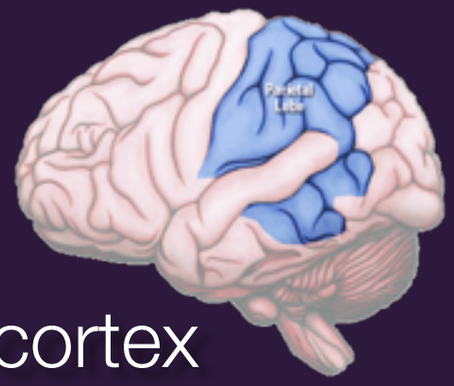
How?


Dendrites
Microtubule
Neurofibrils
Neurotransmitter
Receptor
Synapse
Synaptic vesicles
Synapse (Axoaxonic)
Synaptic cleft
Axonal terminal
Rough ER
(Nissl body)
Polyribosomes
Ribosomes
Golgi apparatus
Synapse
(Axosomatic)
Node of Ranvier
Myelin Sheath
(Schwann cell)
Axon hillock
Nucleus
Nucleolus
Membrane
Microtubule
Mitochondrion
Smooth ER
Synapse
(Axodendritic)
Dendrites
Nucleus
(Schwann cell)
Microfilament
Microtubule
Axon

vicarious

How can we decipher information-processing principles from the brain?

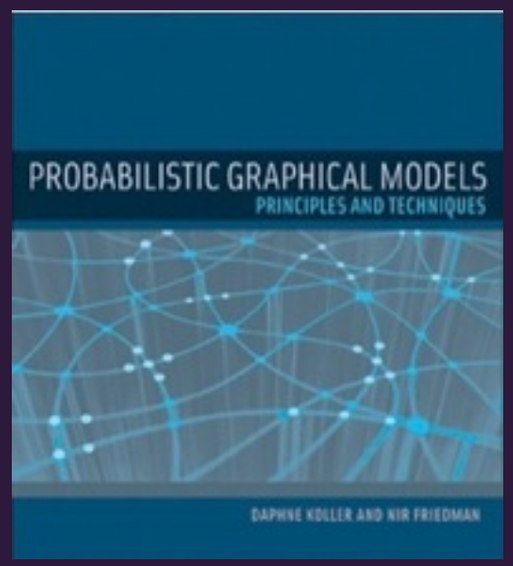
vicarious



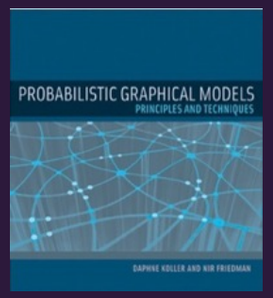
Neocortex

Computational Framework

Physics of World's Data

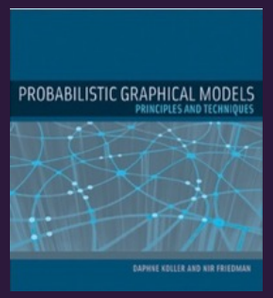


vicarious


PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES
DAPHNE KOLLER AND NIR FRIEDMAN

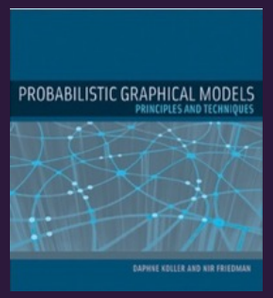
vicarious

Neocortex
Source of assumptions/ constraints



vicarious

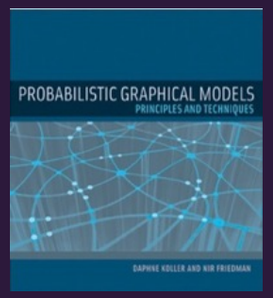
Neocortex
Source of assumptions/ constraints



Physics of World's Data
To find correspondence with neocortex properties

vicarious

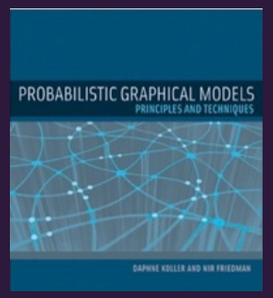
Neocortex
Source of assumptions/ constraints



Physics of World's Data
To find correspondence with neocortex properties

Computational Framework
Understand why neocortex does what it does to design algorithms

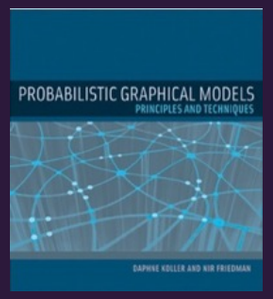
vicarious


PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES
DAPHNE KOLLER AND NIR FRIEDMAN

vicarious

Observed hierarchy in the cortex

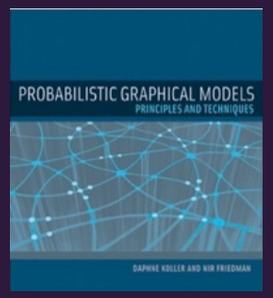
Hierarchical structure of data



Efficiency and re-use.

vicarious

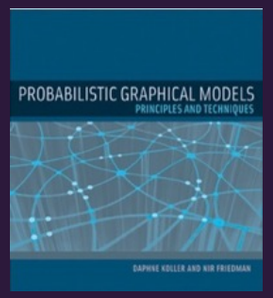
Spikes



?

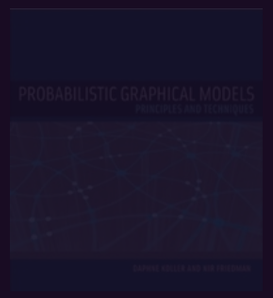
?

vicarious

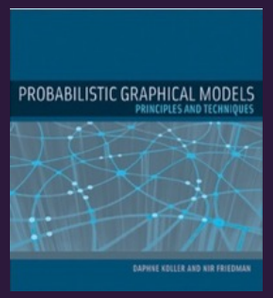

PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES
DAPHNE KOLLER AND NIR FRIEDMAN

vicarious

Blue brain project



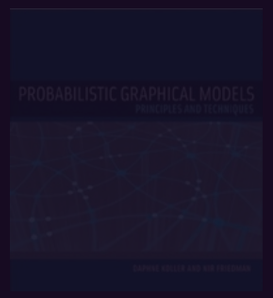
vicarious



Machine learning, statistics

vicarious

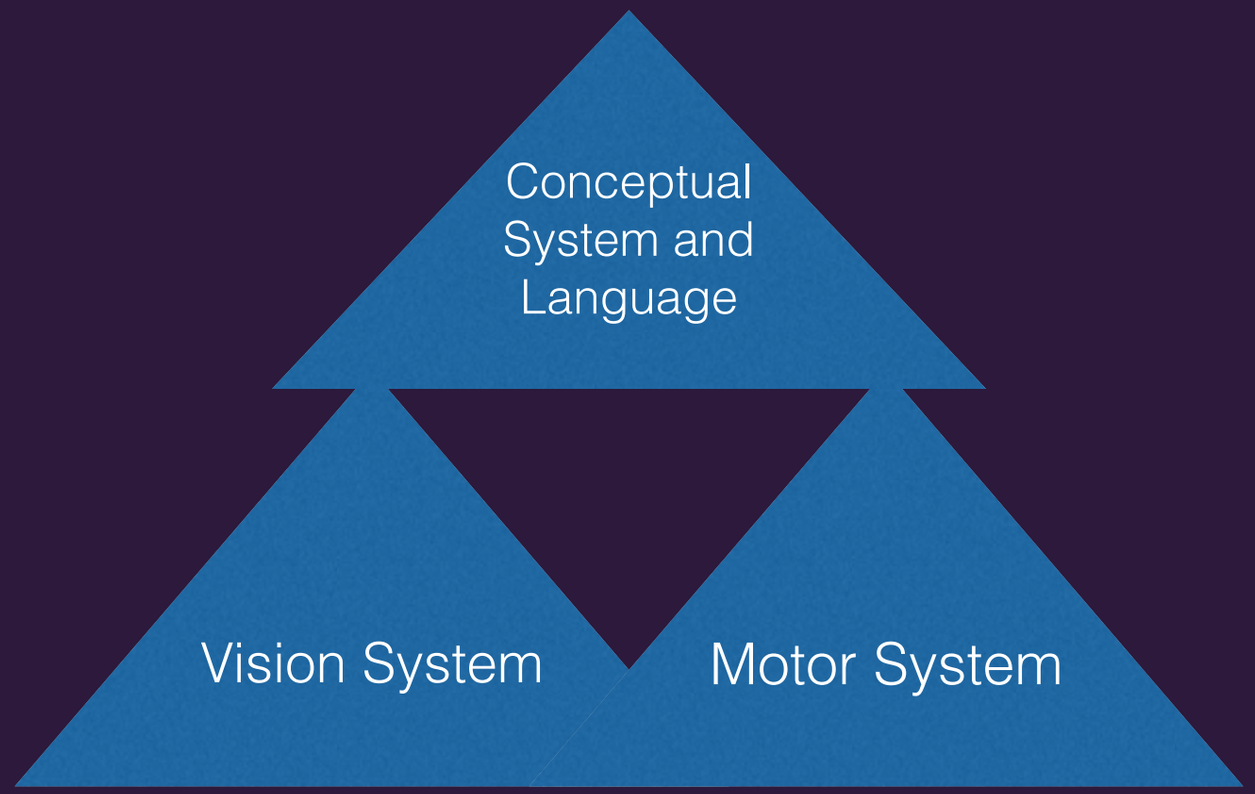
Brain corporation



vicarious

New-brain research questions

- Hierarchy + Feedback + Temporal Learning & Inference
- Scene understanding
- Sensori-motor integration
- Concept learning
- Language understanding


Conceptual System and Language
Vision System
Motor System

How we solved CAPTCHAs

We stuck to new-brain methods:

- Small amounts of training data
  - Trained on clean examples
  - Trained only on positive examples
- Unsupervised training
- ....

Few training examples

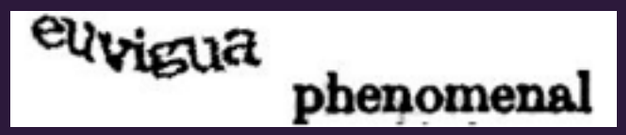
vicarious

Interpolation of millions of training images

VS

Extrapolation from few examples

Solves All Variations



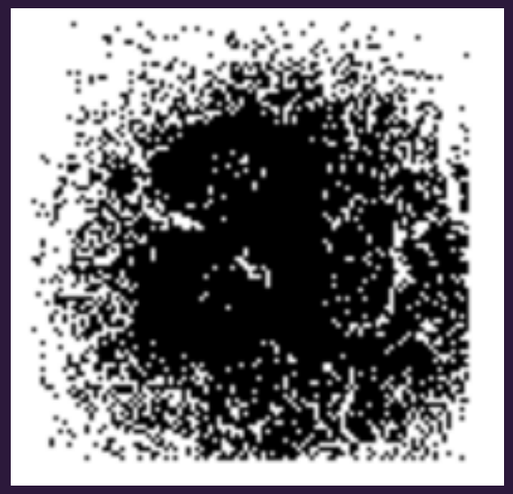
Total training examples < 1000 !

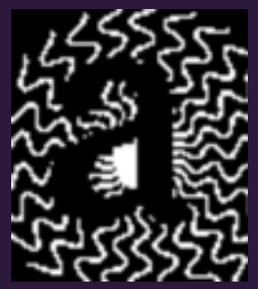
vicarious

Our algorithm produces very detailed segmentation

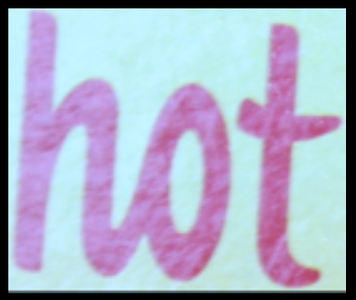
even when the contours are occluded

a





ICDAR

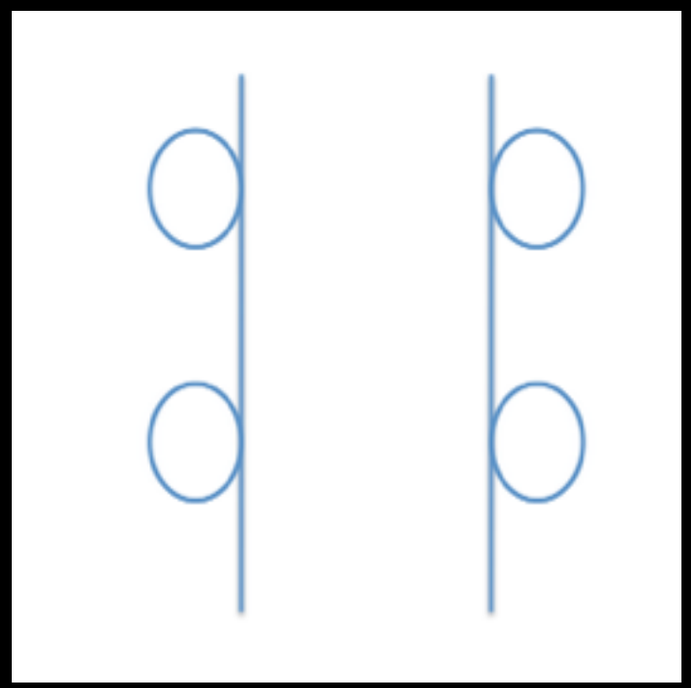


Scene parsing

input	1st (closest)	2nd	3rd	

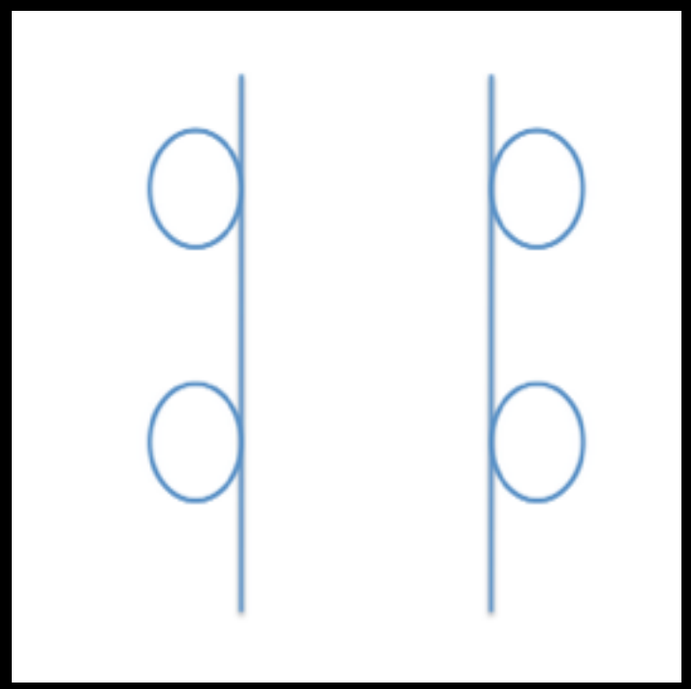
Systems with imagination..

vicarious



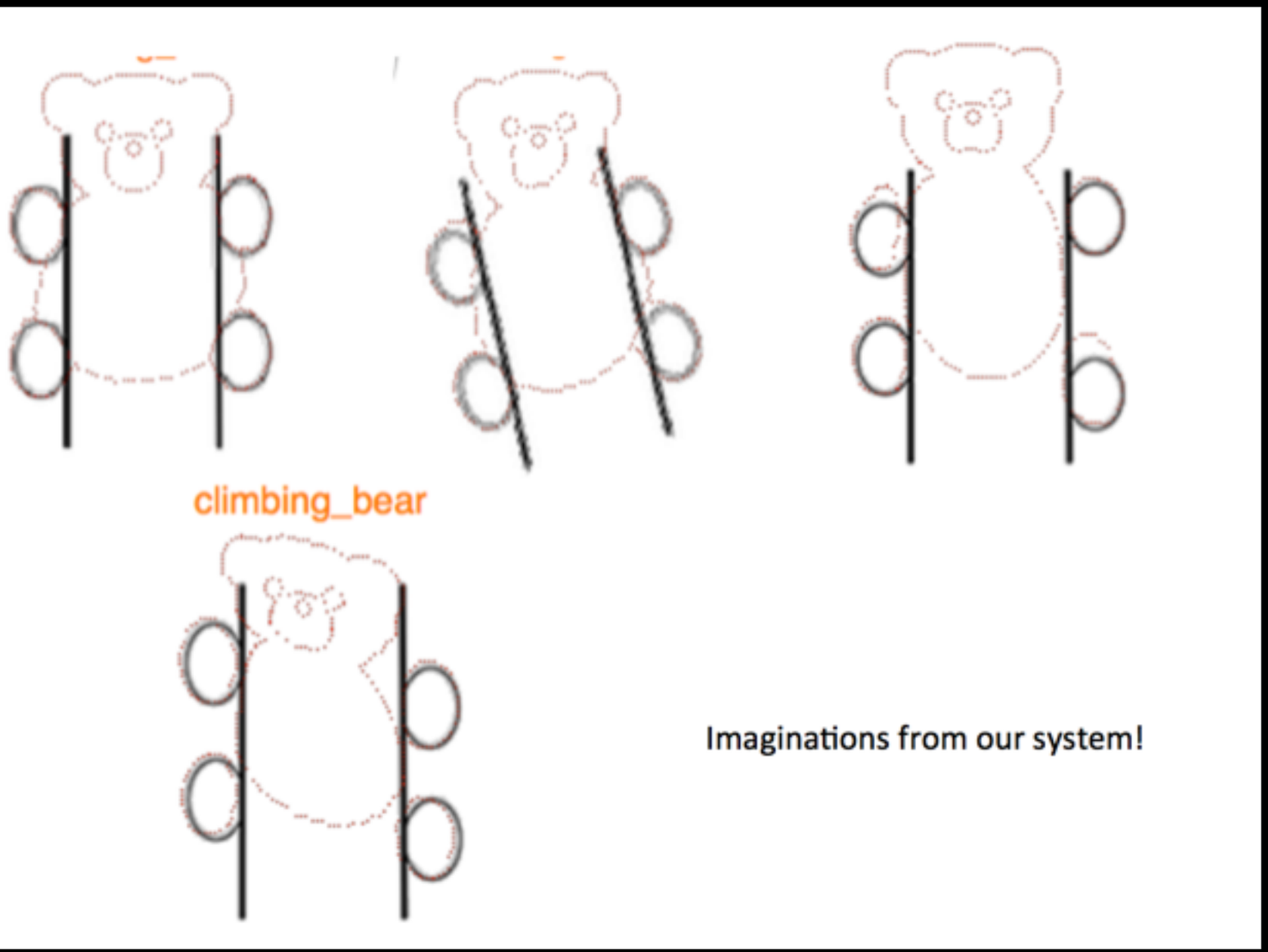
What is this picture?

vicarious



It is the picture of a bear climbing a tree. Can you see the bear?

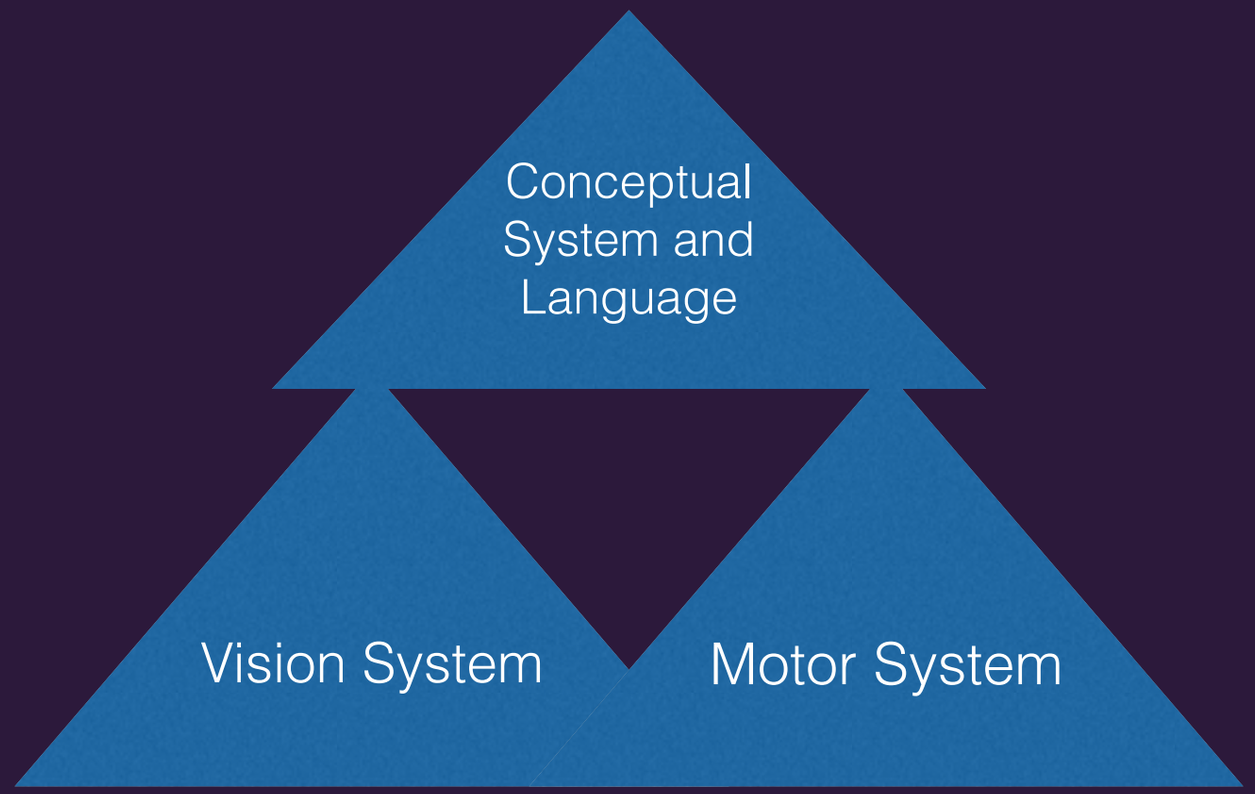
vicarious

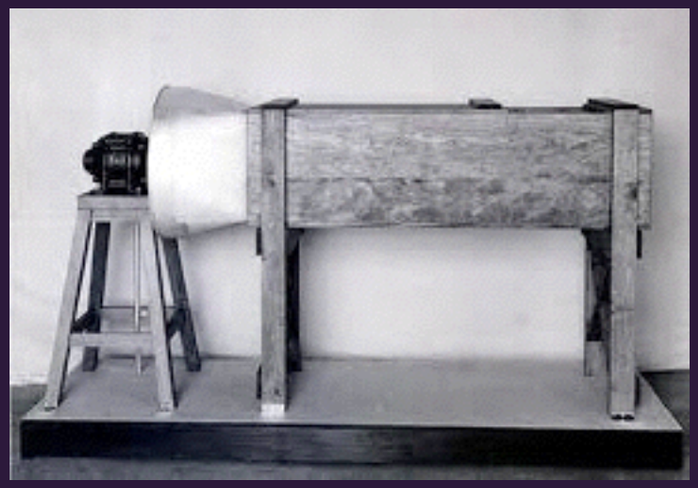


Imaginations from our system!

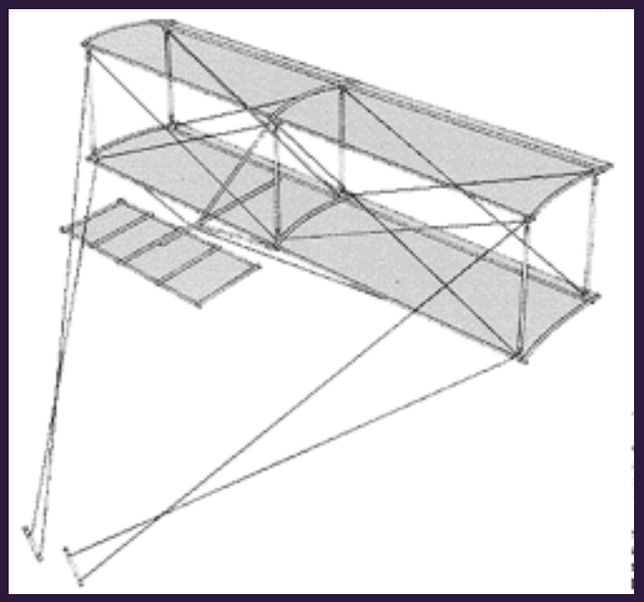
vicarious

Why is imagination important for AGI?

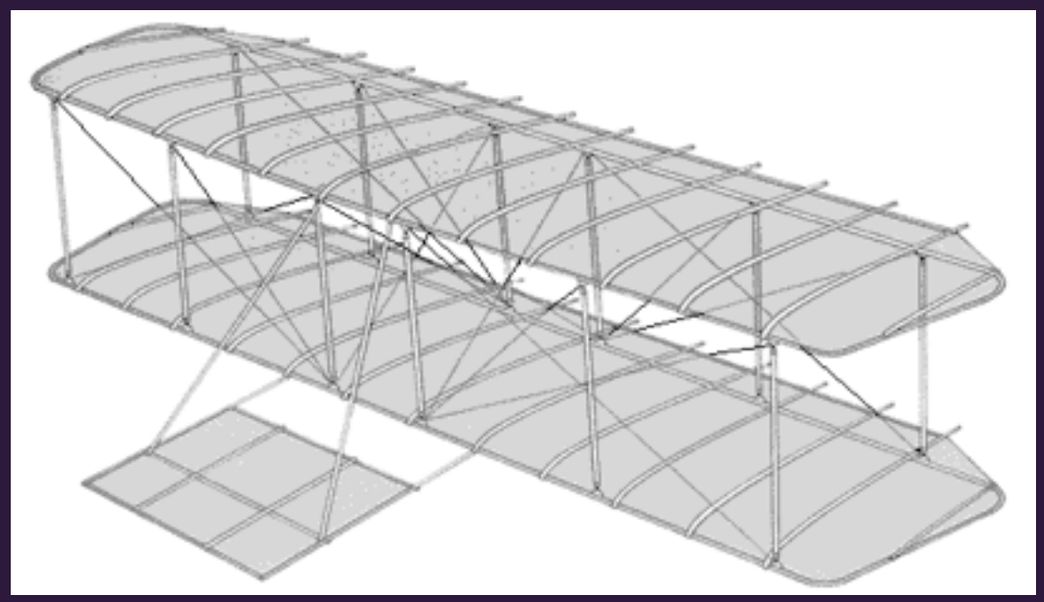

Conceptual System and Language
Vision System
Motor System



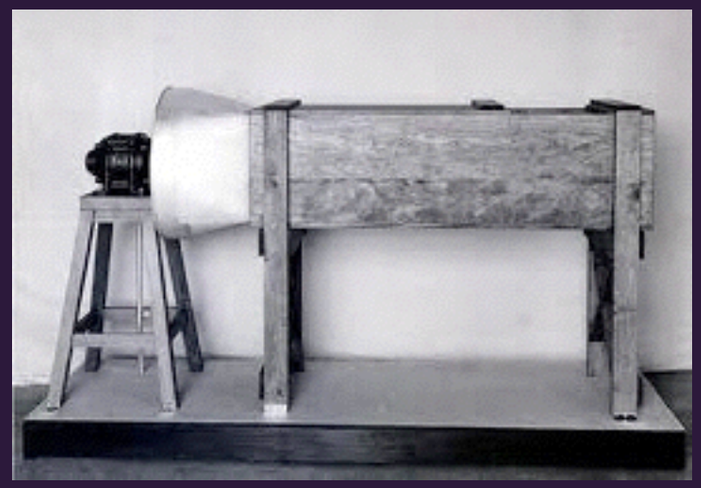
Wind Tunnel



Kite



Glider



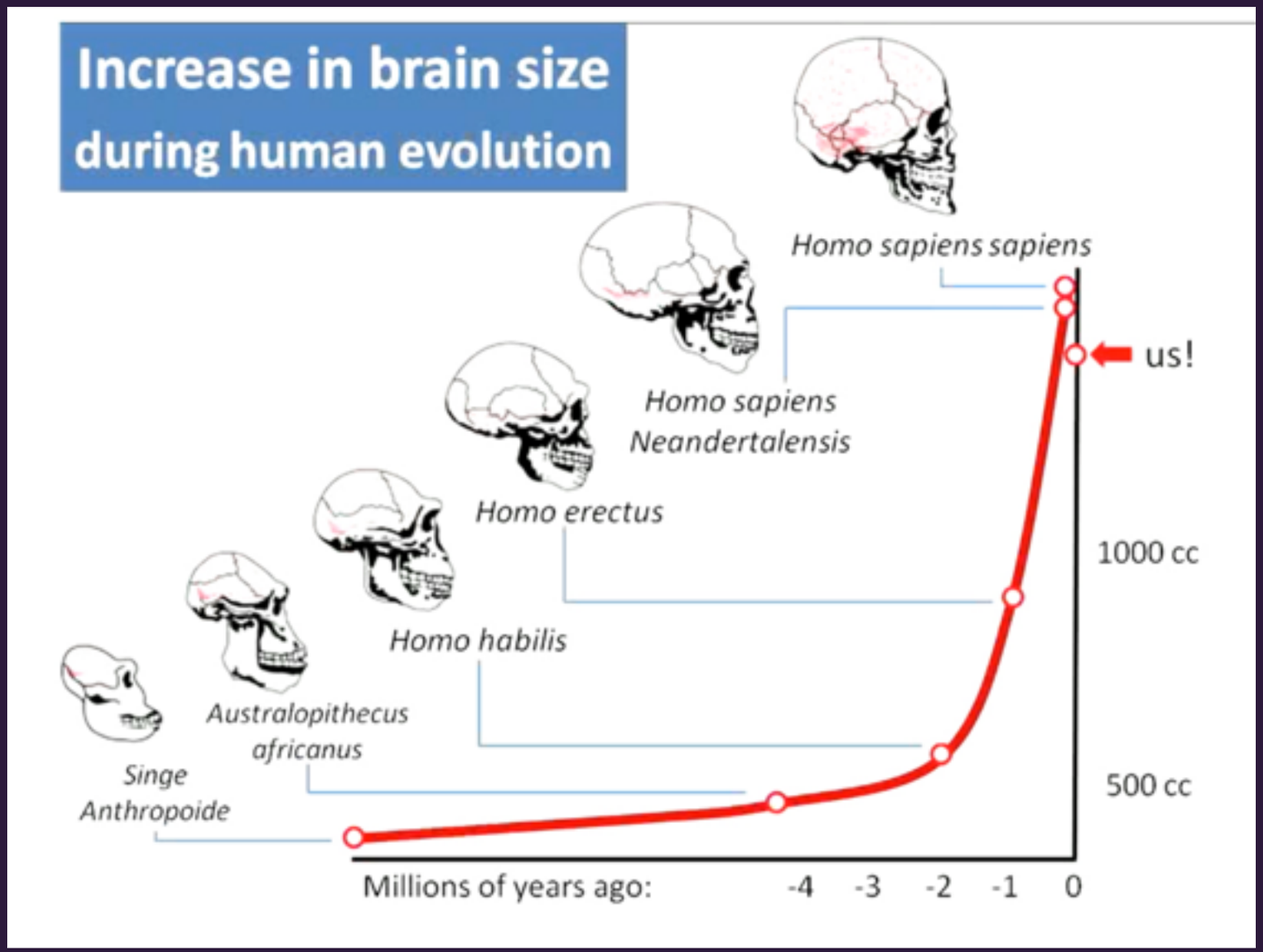
Wind Tunnel

Major breakthroughs in A.I will need some wind-tunnel style research

When?

Will we solve the fundamental research problems in N years?

N <= 5	No way
5 < N <= 10	Small possibility
10 < N <= 20	> 50%


Increase in brain size
during human evolution
Homo sapiens sapiens
us!
Homo sapiens
Neandertalensis
Homo erectus
1000 cc
Homo habilis
Australopithecus
africanus
Singe
Anthropoide
500 cc
Millions of years ago:
-4
-3
-2
-1
0

About intelligence explosion…

Self-limiting forces on A.I that will prevent an uncontrolled growth

vicarious

- Data limit
- Intervention limit
- . . .

vicarious

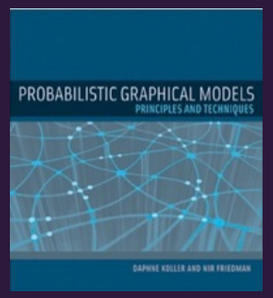
- Data limit
  - Super-human A.I will encounter limits of available data
  - Limit on recursive improvement
    - Going over the same data again and again does not increase the information in the data. (Data Processing Inequality)
- Natural dynamics of data generation

vicarious

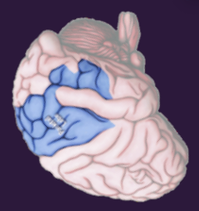
Suppose super-human A.I was created before we had any knowledge about how gravity worked and how earth moved around the sun.

How long will it take for the A.I to discover that it takes the earth 365 days to go around the sun?

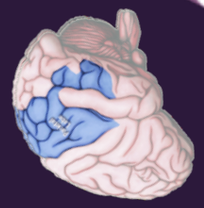
The dynamics of the world imposes limits on how fast data can be acquired.


PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES
DAPHNE KOLLER AND NIR FRIEDMAN

PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES



PROBABILISTIC GRAPHICAL MODELS
PRINCIPLES AND TECHNIQUES



vicarious