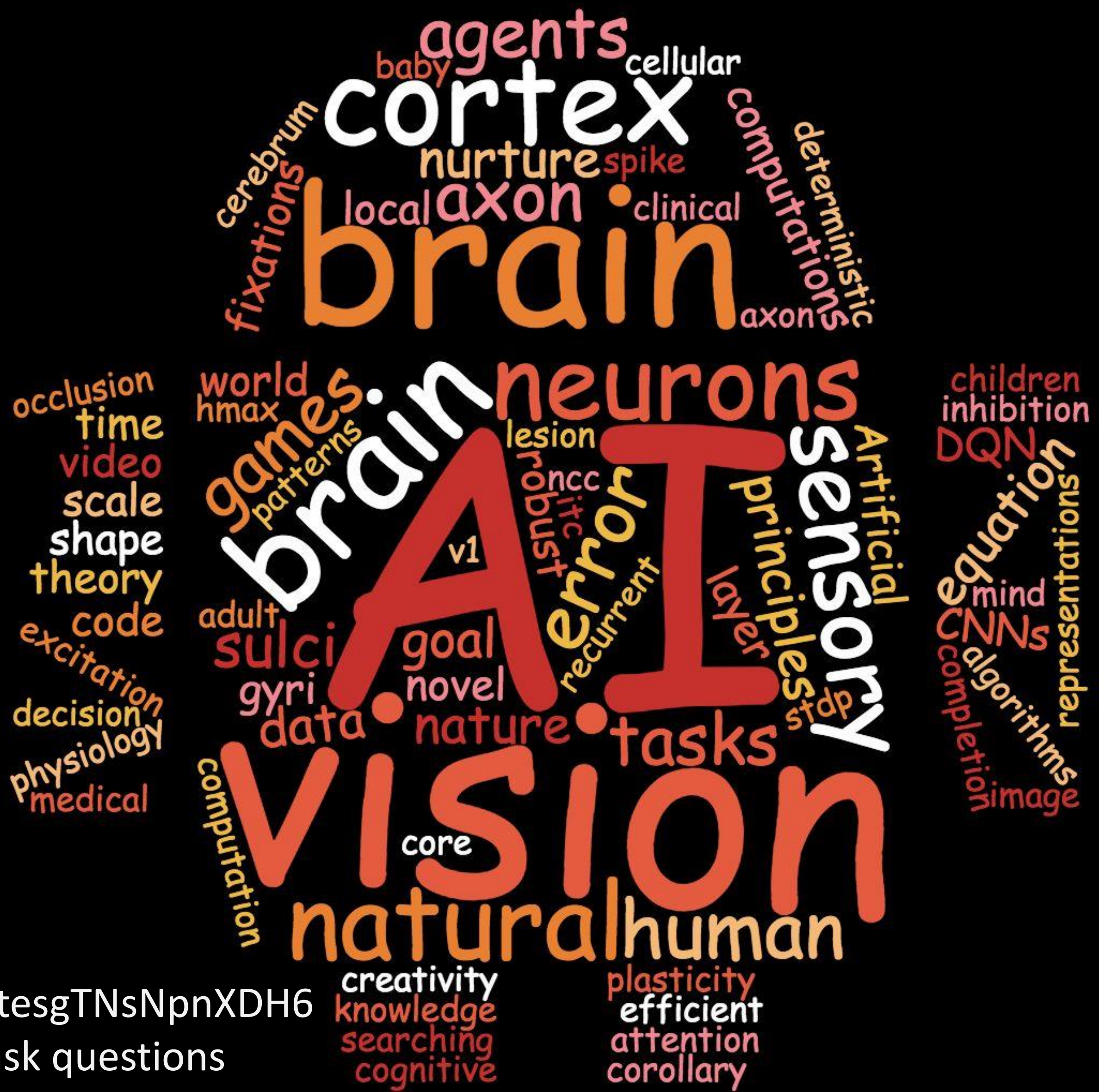
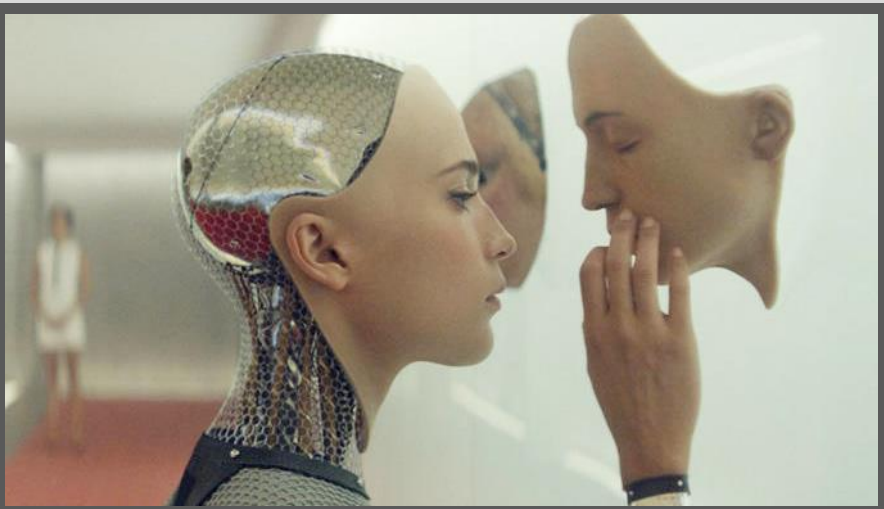
Neuro 140: Biological and Artificial Intelligence



1. https://forms.gle/adFtesgTNsNpnXDH6
2. Please interrupt and ask questions

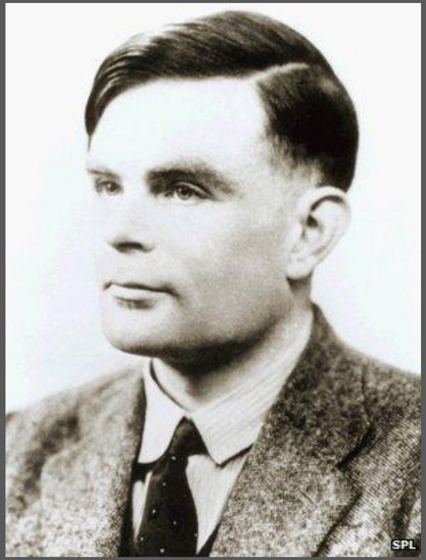
The last machine and the Turing test



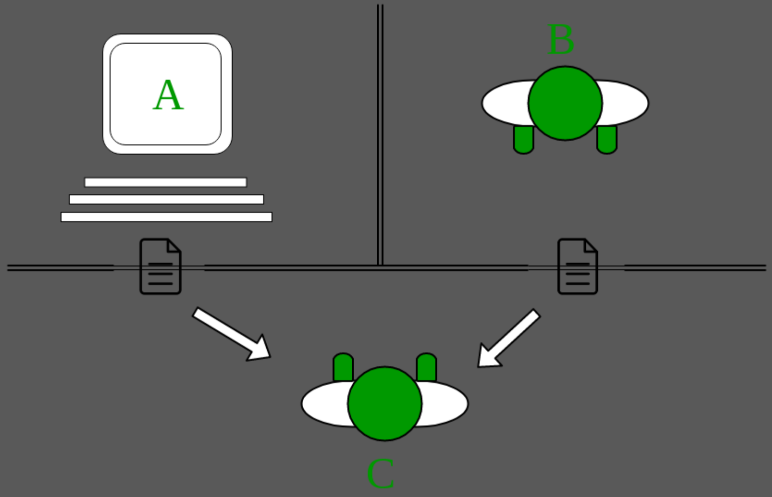
Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. **Thus the first ultraintelligent machine is the last invention that man need ever make . .**

I.J.Good, in Speculations regarding the first ultra-intelligent machine (1965)

Alan Turing



The Turing Test



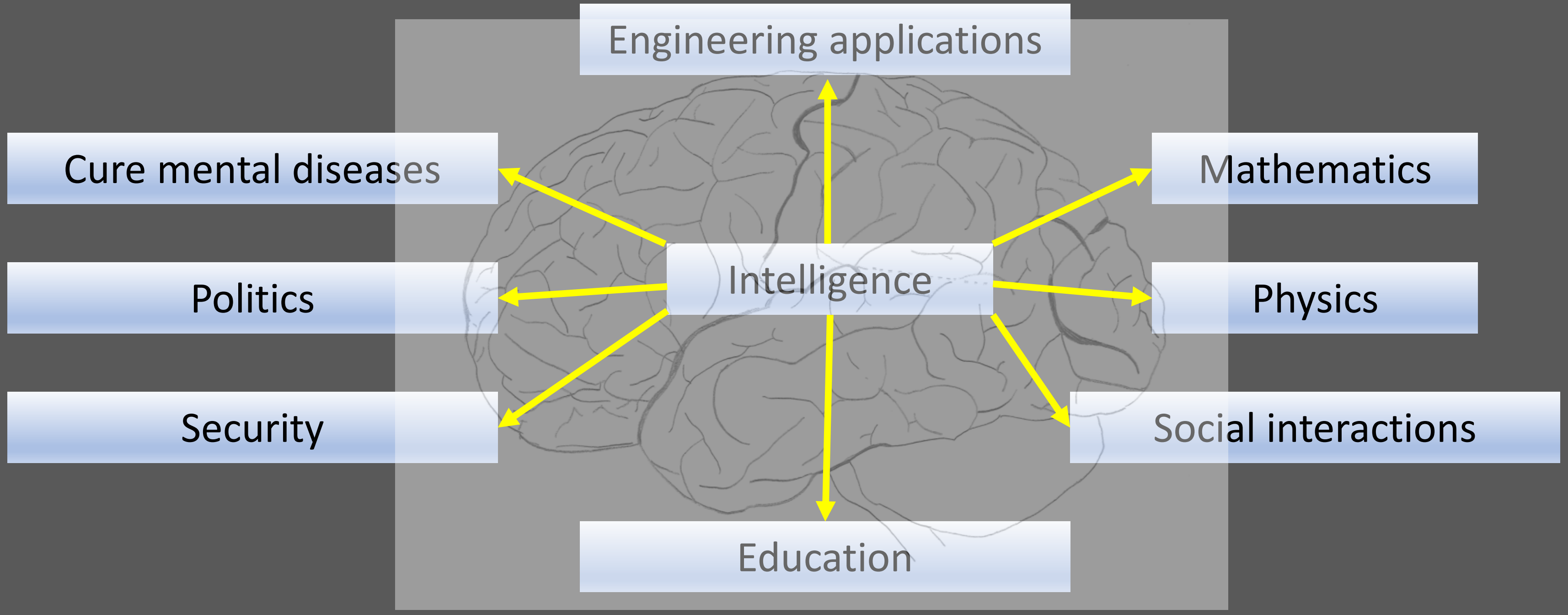
https://forms.gle/adFtesgTNsNpnXDH6

Turing, A. (1950). "Computing Machinery and Intelligence." Mind **LIX**(236): 433-460.

Intelligence is the greatest problem in science

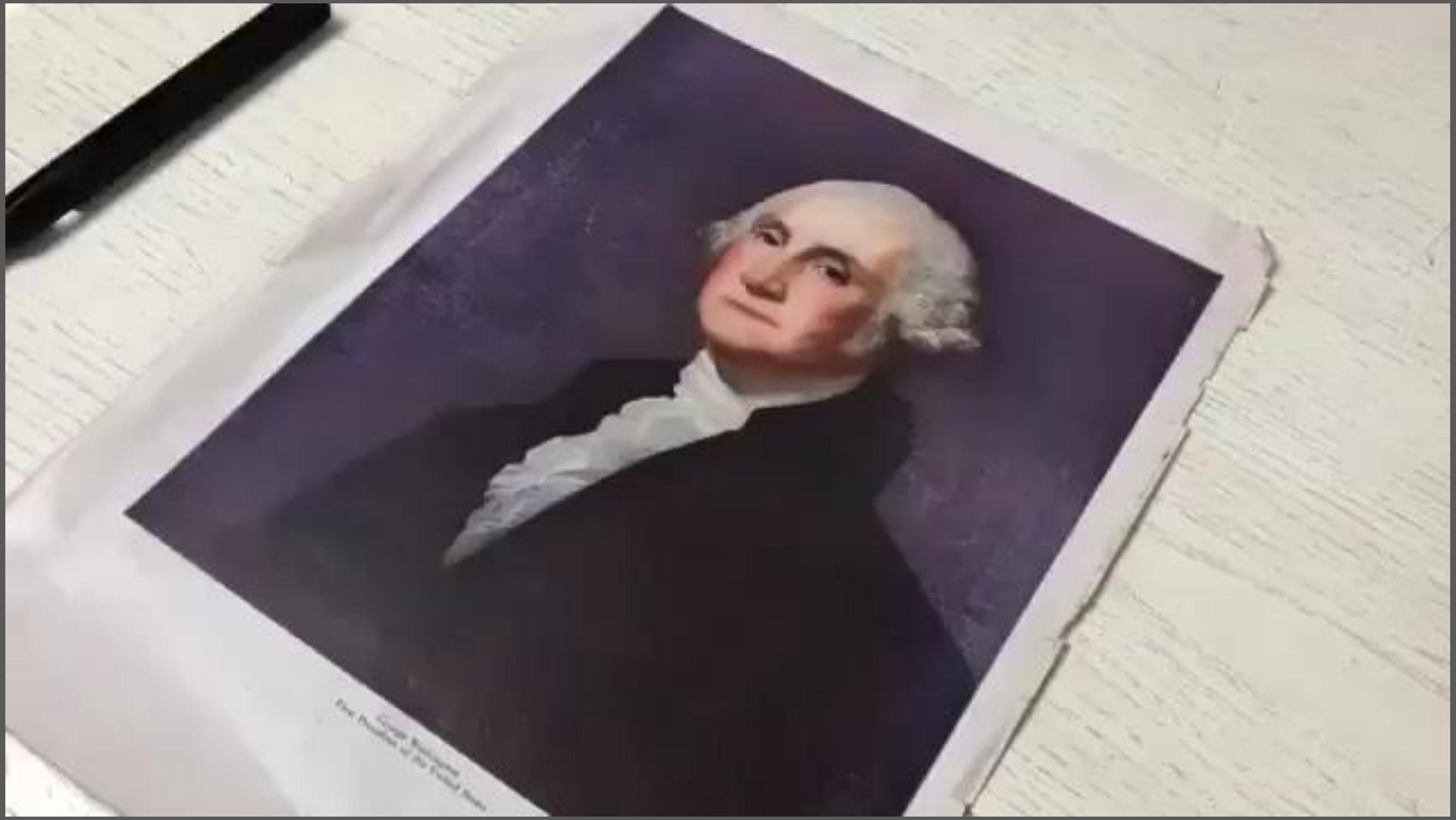
If we understand the brain and we understand intelligence … we could find ways to make us smarter and to build smart machines to help us think

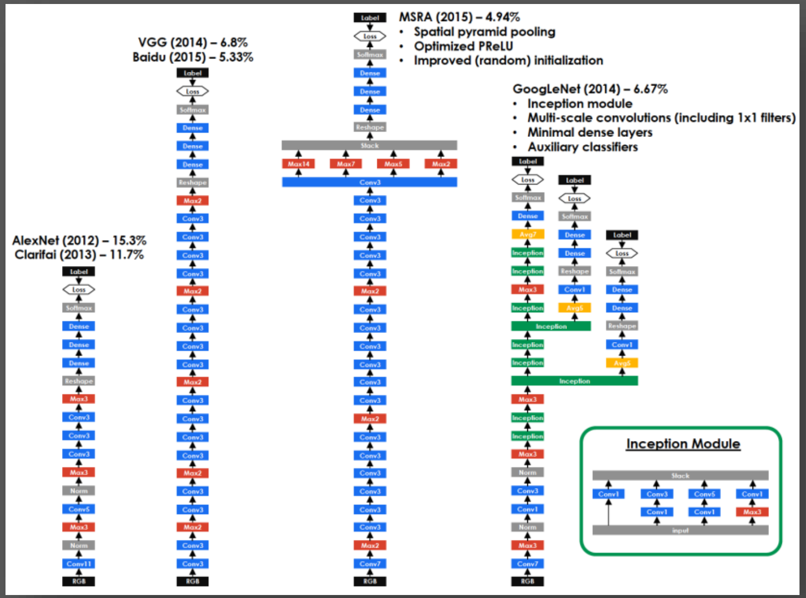
Tomaso Poggio, MIT



Rapid progress in AI

Deep convolutional networks





What can deep convolutional networks do?

0. Handwritten digit recognition

1. Classification of large image datasets

2. Better at face recognition than “superrecognizers” and face forensic experts

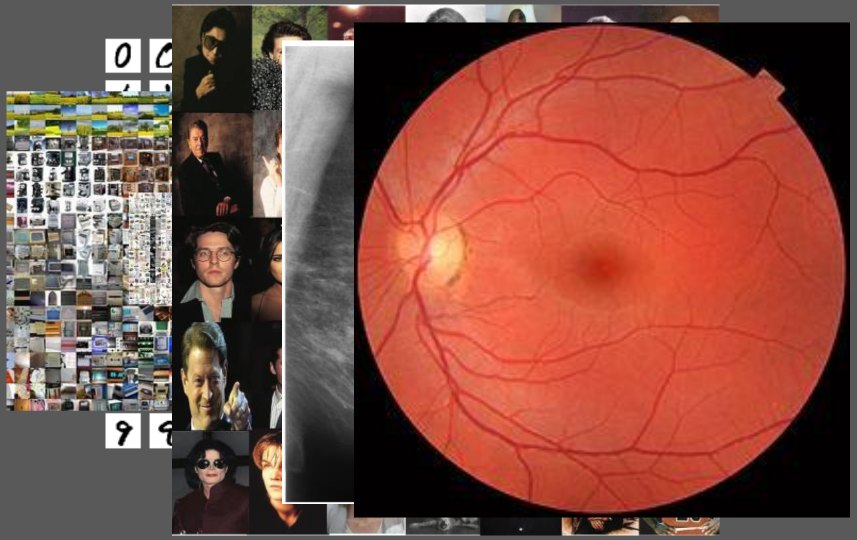
3. Better at diagnosing breast cancer than radiologists

4. Better than ophthalmologists at diagnosing diabetes of retinopathy. Also, can extract other information such as cardiovascular disease from images of the eye!

5. Classification of plants

6. Classification of galaxies

7. Extension to other domains
  - 7a. Speech recognition
  - 7b. Sentiment analysis of short texts
  - 7c. Decision-making in health care
  - 7d. Automatic translation
  - 7e. Predictive advertising
  - 7f. Predicting earthquakes

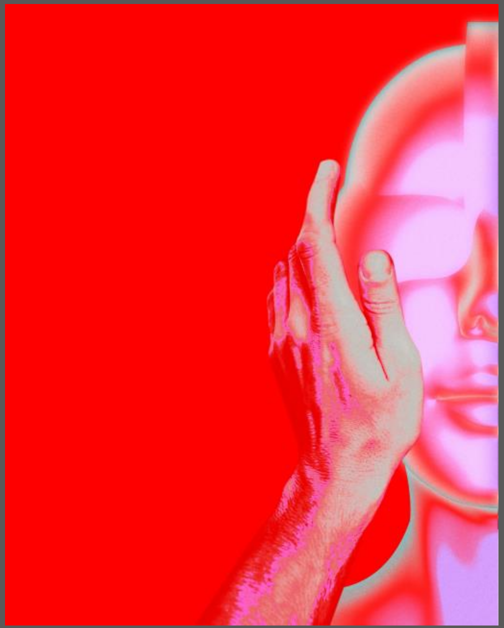


What can’t deep convolutional networks do?

A lot!

Ascribing feelings to machines

The Tamagotchi effect



Do you take this robot …
NY Times 19Jan2019

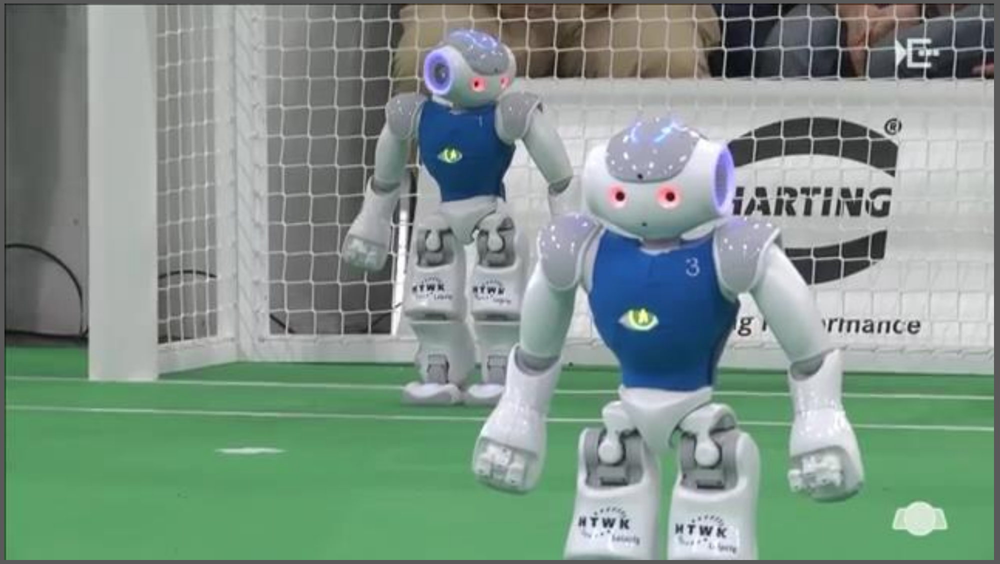
Is it evil to push Atlas?

Perils of AI

1. Redistribution of jobs (akin to but perhaps larger than the Industrial Revolution)
2. Unlikely: Terminator-like scenarios
3. Military applications
4. To err is algorithmic (human too)
5. Biases in training data (note that humans have biases too)
6. Lack of “understanding” (note that we do not necessarily understand how humans make decisions either)
7. Social, mental, and political consequences of rapid changes in labor force
8. Rapid growth, faster than development of regulations

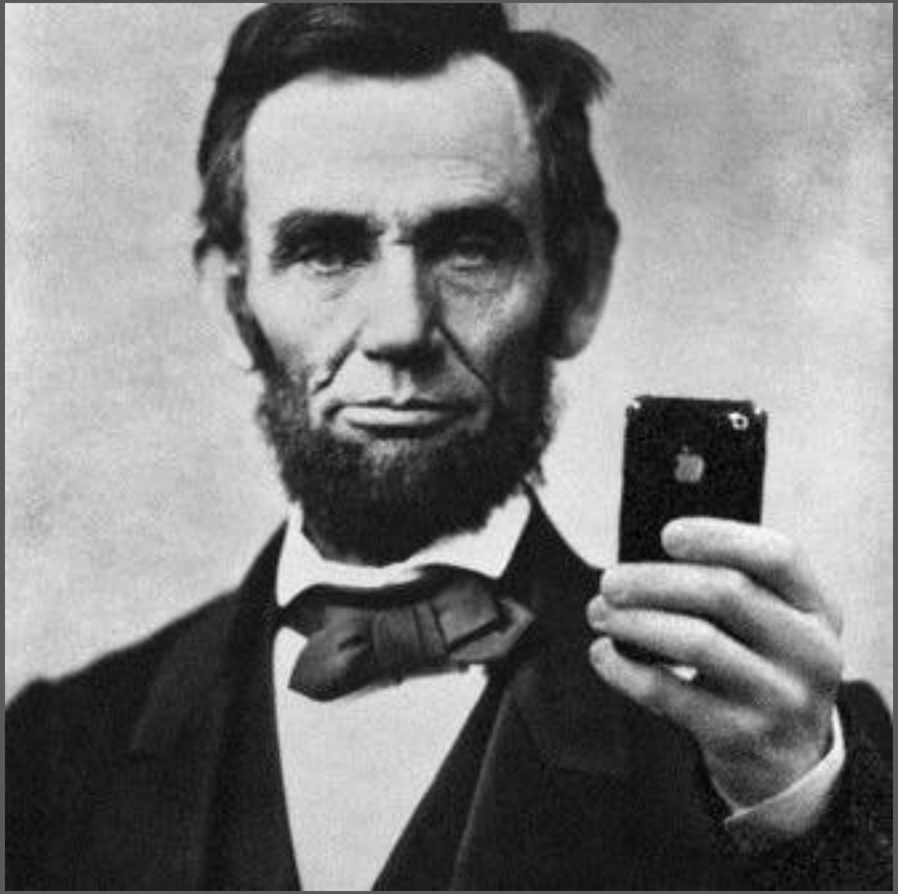
Robots don’t play soccer (yet)

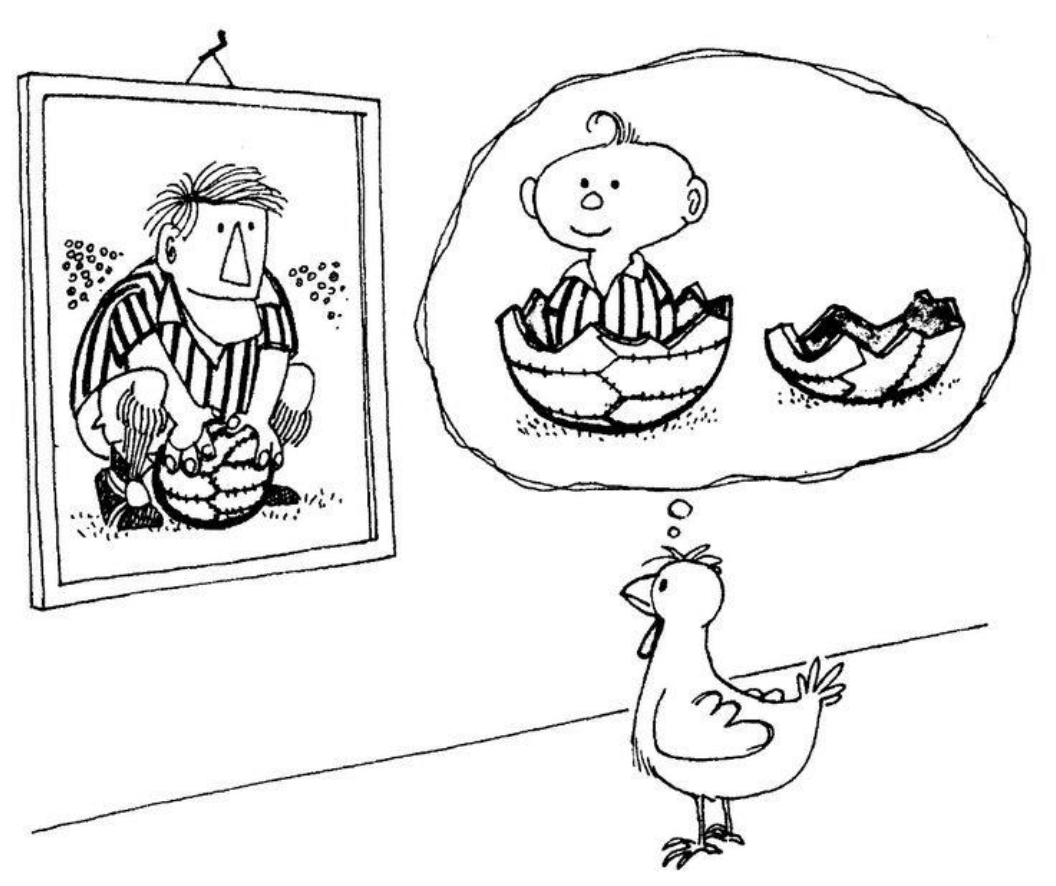
Lionel Messi



GO 2015 Finals: Nao-Team versus B-Human

Example challenge in AI: Understanding humor

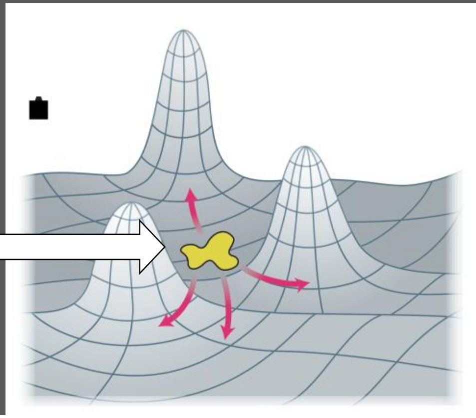




What does biology tell us about intelligence?

Biology finetunes the search for artificial intelligence by showing us what intelligence is most often used for:

Adaptation in Uncertainty




PERSPECTIVES

EVOLUTION

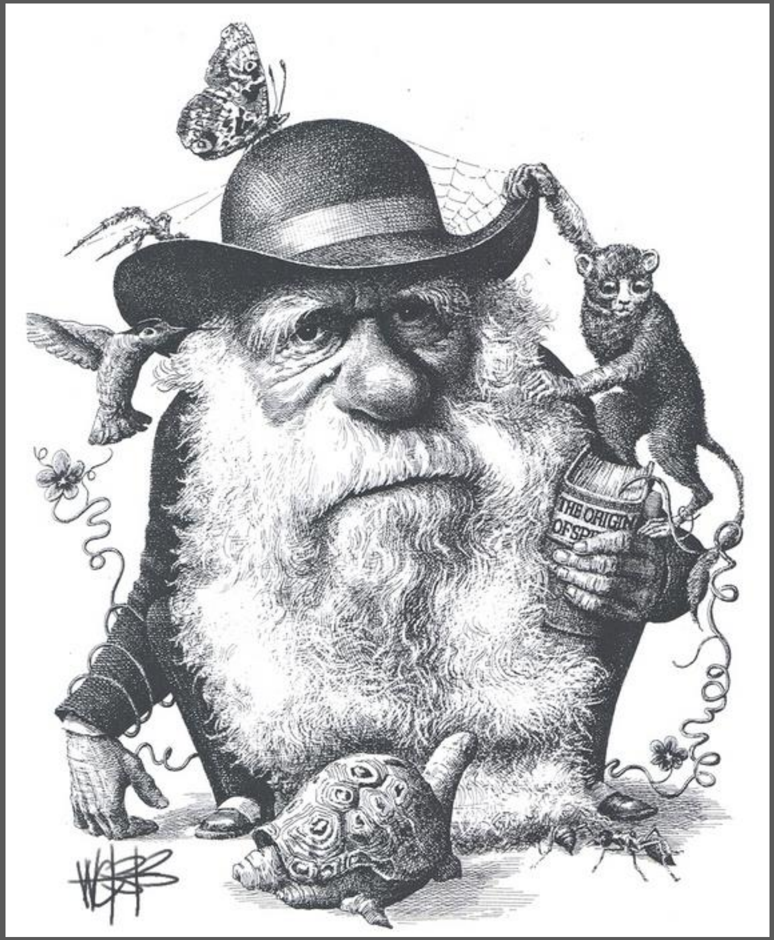
Climb Every Mountain?

Santiago F. Elena and Rafael Sanjuán

What does biology tell us about intelligence?

Two Broad Domains of Adaptation:

Niche Generalism
(General Intelligence)



Niche Specialism
(Specialist Intelligence)

What does it mean to be a generalist?

What material is this made of?

Is that material hard or soft?

If I threw this at you, would it hurt?

If you stubbed your toe on it in the driveway, how would you attend to the toe?

Which of the Bop-It functions can you still perform with just your toe?

If you dropped this in the ocean, would there be a visible displacement in the tide line on the shores?

Does Bop-It have a lunar cycle?

What tool would you need to replace the batteries if the batteries were dead?

What does a battery store?

Can you eat this?

If you could, would you accompany it with a crème brulée or a Godiva chocolate?

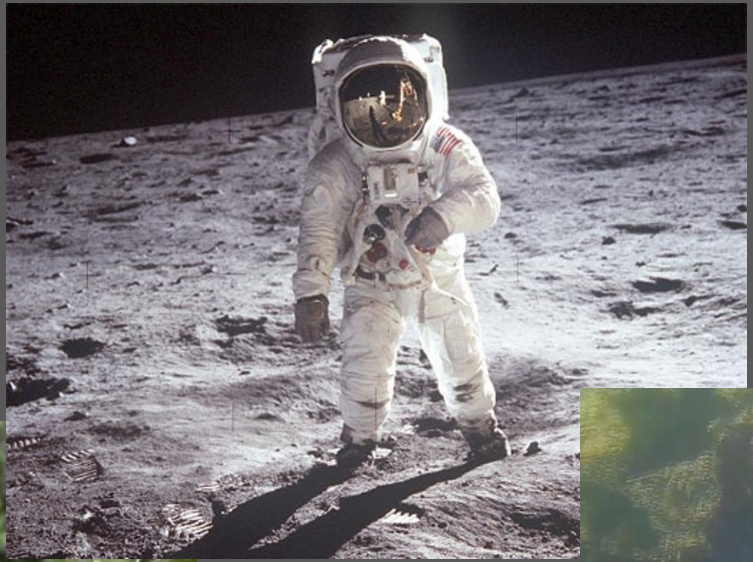
What age of child do you think should not be playing with this toy?

What age of adult do you think should not be playing with this toy?

Can this toy provide me some sense of self-actualization in an otherwise long and tawdry life?

Only Biology Has Created Generalism

Humans are the consummate generalists of intelligent life.



Only Biology Has Created Niche Generalism

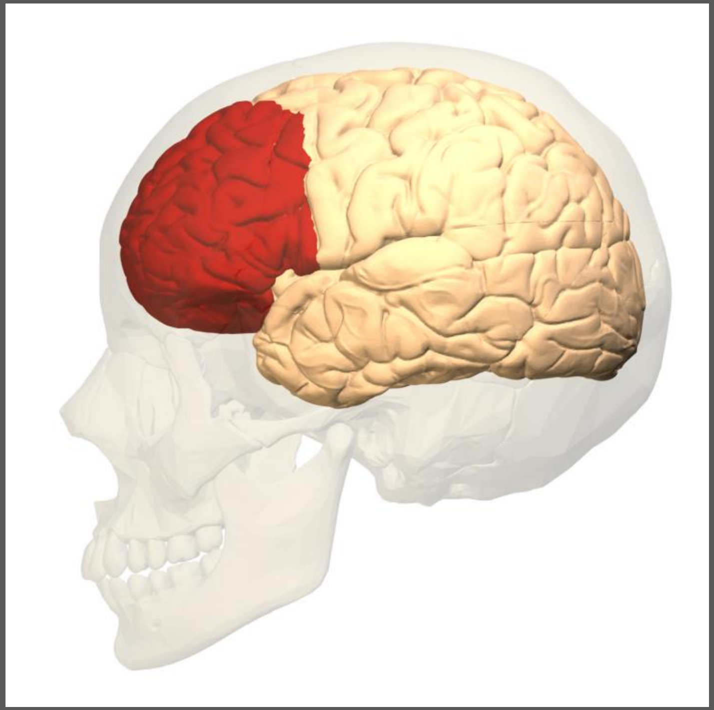
Modern intelligent machines are usually *extreme* specialists.

Shift the paddle 4 pixels and the model fails.

What sets humans apart?

Language?

Culture?



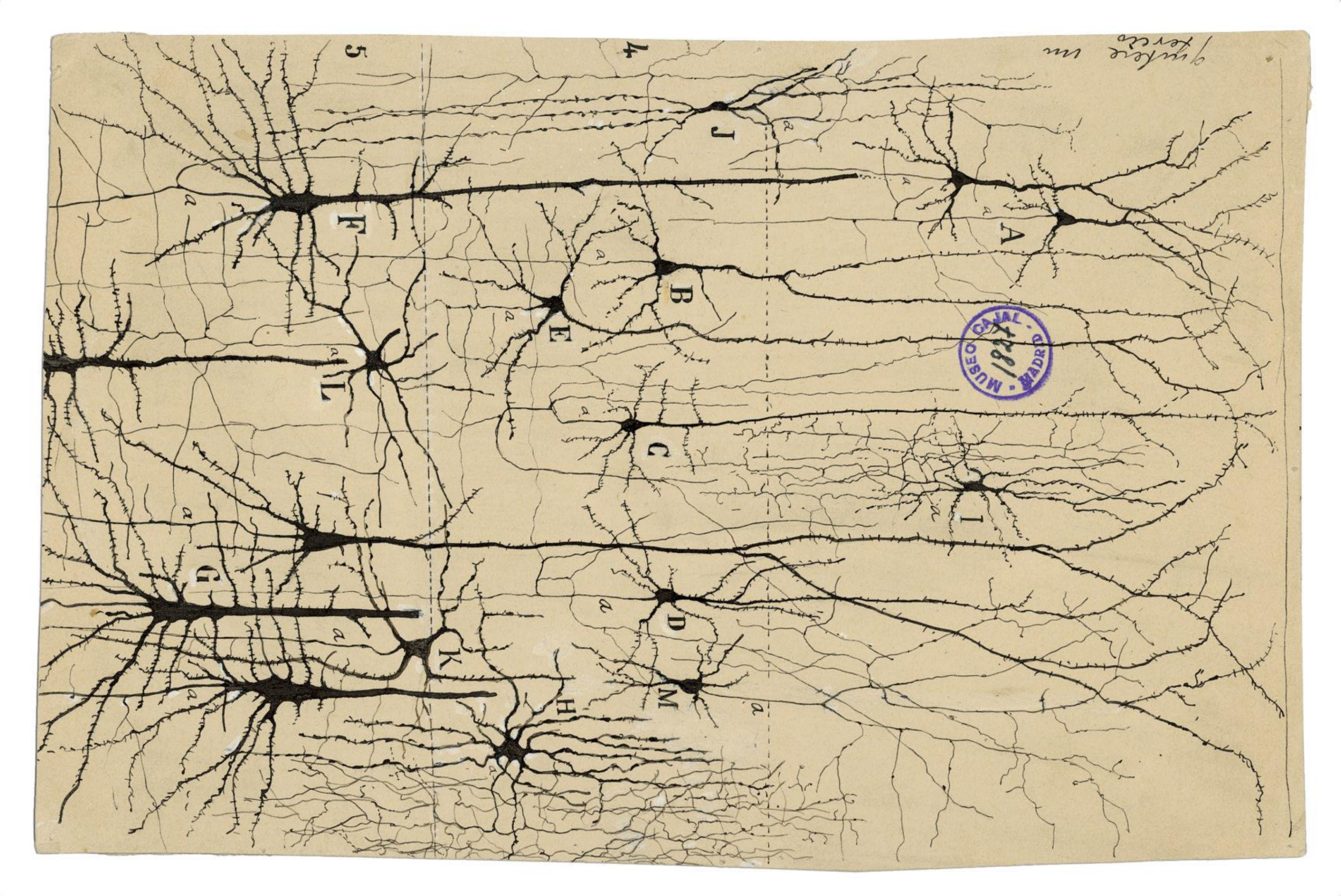
An expanded prefrontal cortex?

Is human intelligence actually exceptional?

- While humans may be the most prolific generalists, many non-human animals aren't far behind.

What is the biology of general intelligence?

- At the base of biological general intelligence are hundreds of millions of small sinewy pocket calculators called neurons.



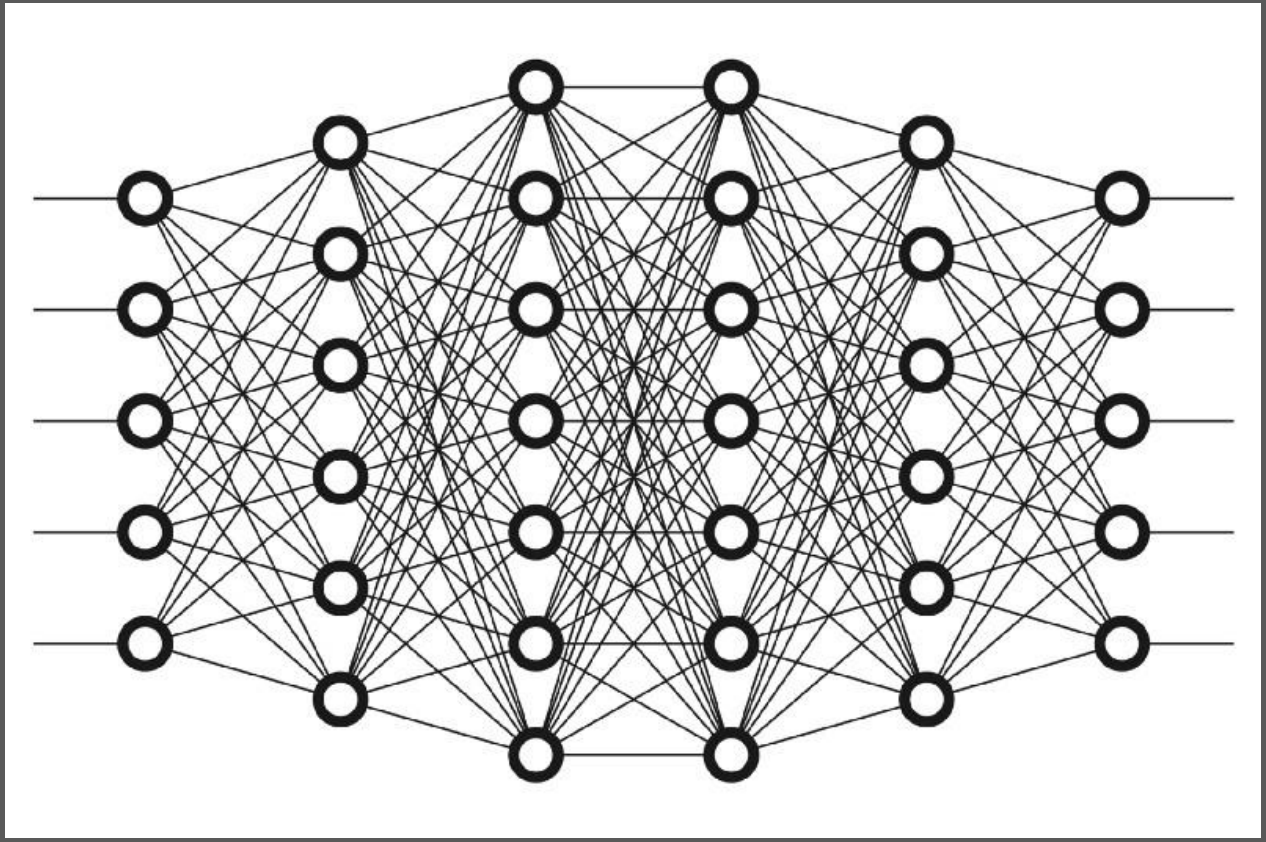
Intelligence is Computational Teamwork

- Neurons Combine Into Neural Circuits ->
- Neural Circuits Combine Into Neural Networks ->
- Neural Networks Combine into Brains ->
- Brains Combine Into Communities…

  - Small, seemingly insignificant computations combined *in parallel, at massive scale,* and *across space* is the recurring motif of biological intelligence.

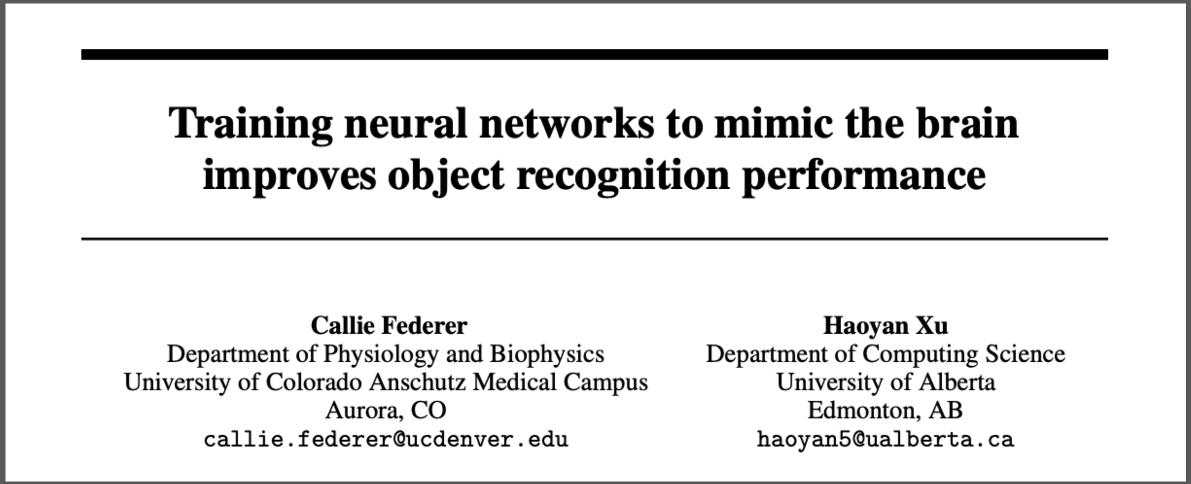
Neural Networks: The Dominant Game in Town

The ”third wave of artificial intelligence” is driven by massively parallel distributed processing in the form of biologically inspired deep neural networks.



The Power of Deep Neural Networks

Extrapolating Trends: More Biology, More Intelligence



Training neural networks to mimic the brain improves object recognition performance

Callie Federer
Department of Physiology and Biophysics
University of Colorado Anschutz Medical Campus
Aurora, CO
callie.federer@ucdenver.edu

Haoyan Xu
Department of Computing Science
University of Alberta
Edmonton, AB
haoyan5@ualberta.ca

Biomimetics

Neuromorphics

Wattage of the Human Brain:

2400 Calories = 2400KCal/24Hours
= 27.8 Cal/Sec = 116.38 J/s = 116 W
Assuming ~90% of energy consumption by the brain, a maximum of 100 watts.

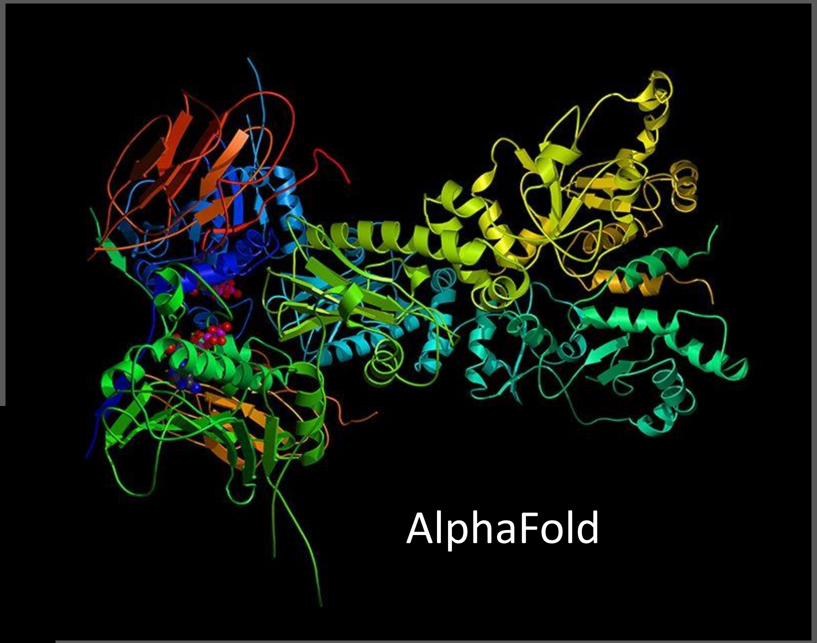
Wattage of Deepmind's AlphaGo:

Approximately the total wattage of all London Metropolitan Consumption in July 2015 over the course of approximately 6 hours

Backwards Extrapolation: More Intelligence, More Biology

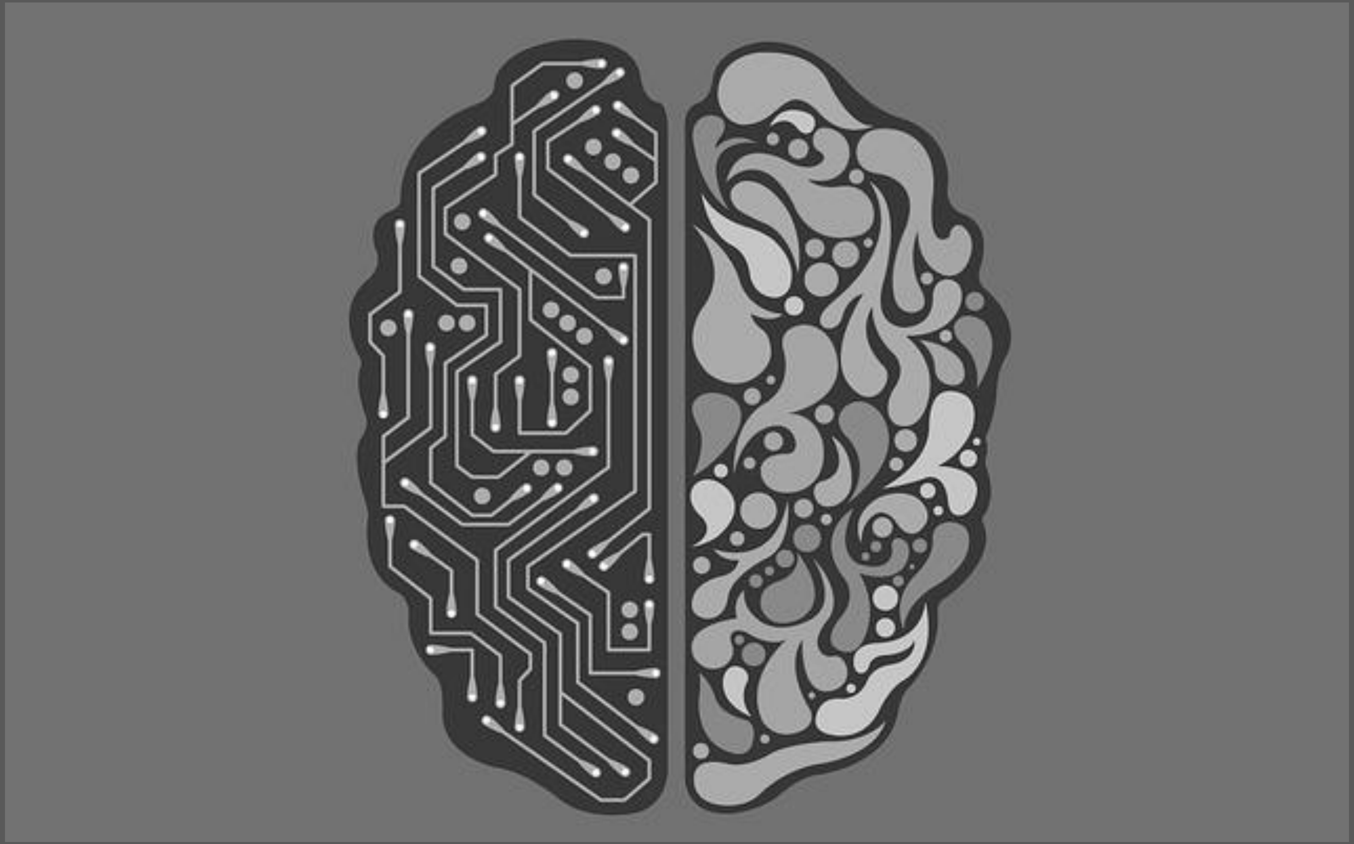
Task-Driven Convolutional Recurrent Models of the Visual System

Aran Nayebi[1,*], Daniel Bear[2,*], Jonas Kubilius[5,7,*], Kohitij Kar[5], Surya Ganguli[4,8], David Sussillo[8], James J. DiCarlo[5,6], and Daniel L. K. Yamins[2,3,9]



An Emerging Framework: Neuro + AI

A virtuous feedback loop between synapse and silicon.



Organizing Principles of Modeling in Neuro + AI

- **Learning:** Whether fully differentiable from end to end, or facilitated by programmatic or architectural biases, models should learn from input.
- **Stimulus Computability**: Models should be trained on similar types of input to what sensory systems receive – images, sounds, chemicals, pressure.
- **(Biological) Cybernetics**: Models should mimic the calibrated feedback loops of communication and control that predominate in biological systems.
- **Computational Efficiency**: Models should be designed with reference to the energetic costs of computation, a reality biological systems cannot escape.
- **Depth / Hierarchy**: Where applicable, models should reflect the multiple levels of information processing that occur in transit from stimulus to response.

Biology + Generalism: Neuro + AI -> AGI?

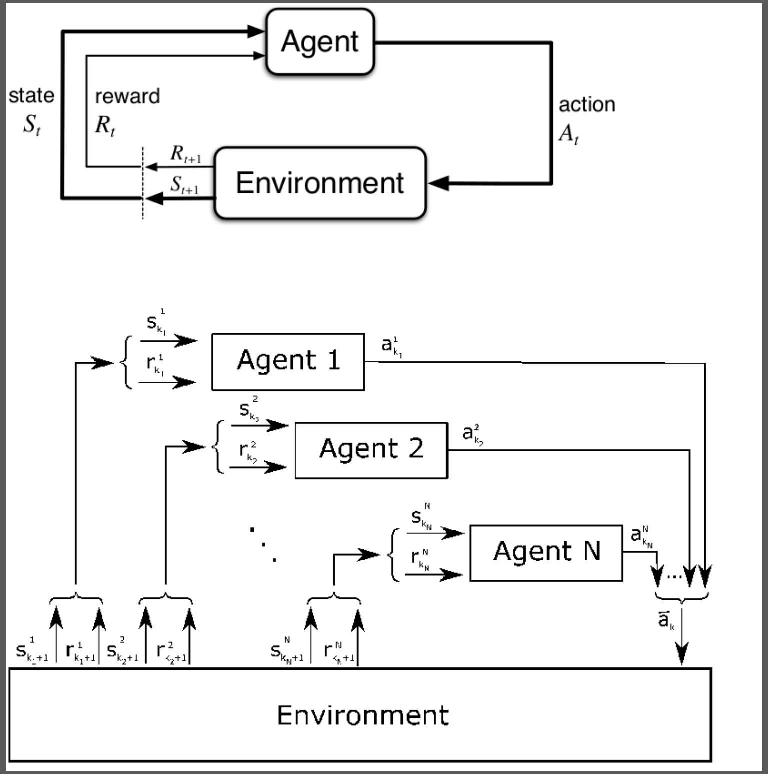
- Neuroscience-inspired artificial intelligence is ascendant… so why haven't we achieved the sort of general intelligence characterized by humans?
- 3 components of biology (and evolution) that AI has mostly ignored:
  - Body (Energetic Consumption + Sensory Interface)
  - Ecology (Social + Physical Environment)
  - Change over Time (Mutation)
- These things combined built niche generalism; how could we mimic them?

Towards Evolutionary Neuro-AI?

DEEP NEUROETHOLOGY OF A VIRTUAL RODENT

Josh Merel[⋆1], **Diego Aldarondo**[⋆2,3], **Jesse Marshall**[⋆3,4], **Yuval Tassa**[1], **Greg Wayne**[1], **Bence Ölveczky**[3,4]
[1]DeepMind, London, UK.
[2]Program in Neuroscience, [3]Center for Brain Science, [4]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.
jsmerel@google.com, diegoaldarondo@g.harvard.edu, jesse_d_marshall@fas.harvard.edu

Virtual Embodied Agents in Plausible Environments



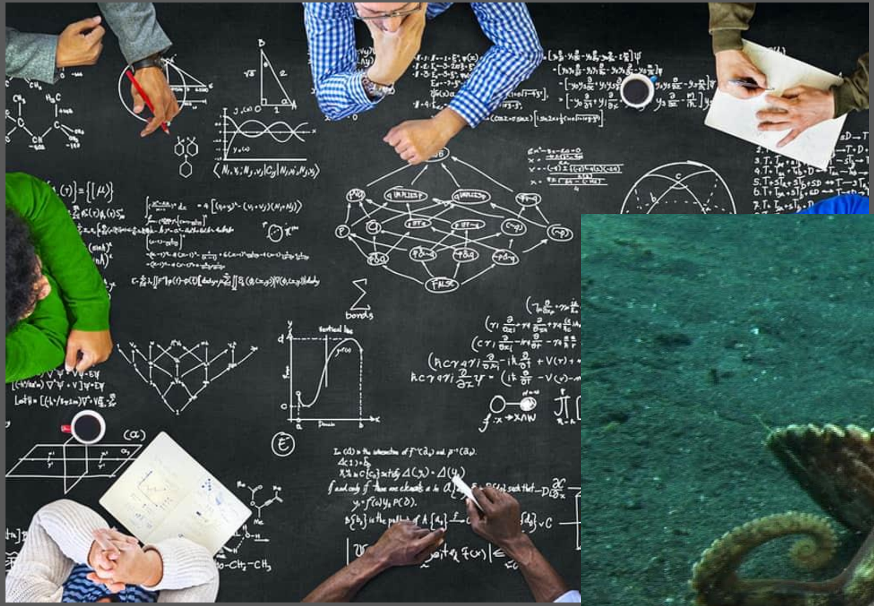
Agent-Based Reinforcement Learning

Towards Evolutionary Neuro-AI?

There exists a deep relationship (effectively an isomorphism) between replicator dynamics and reinforcement learning.

The Future of AIResearch is Diversity of Perspective

- The overwhelming lesson of biology in the search for artificial intelligence is the sheer diversity of forms that intelligence can take.



Neuro 140: Main faculty + Many lecturers

Gabriel Kreiman **– Office Hours**:

- After class (Cambridge)
- Upon request (Longwood): Email klabcoordinator@gmail.com and copy gabriel.kreiman@tch.harvard.edu

A multidisciplinary cadre of experts

Gabriel Kreiman

Computational Neuroscience

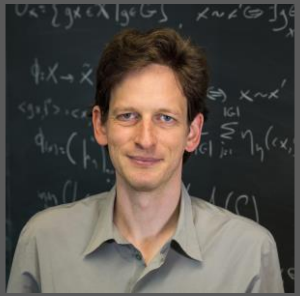
Haim Sompolinsky

Computational Neuroscience

Jan Drugowitsch

Computational Neuroscience

Tomer Ullman



Psychology

Dina Obeid

Computational Neuroscience

Andrei Barbu

Computer Science

Thomas Serre

Computer Science

Sam Gershman

Psychology

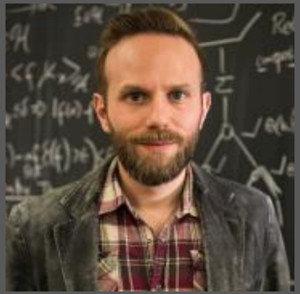
Aravi Samuel

Physics

MacKenzie Mathis

Neuroscience

Xavier Boix



Computer Science

Cengiz Pehlevan

Applied Math

Aude Oliva

Computer Science

Jacob Andreas

Computer Science

Neuro 140: TAs

Colin Conwell
conwell@g.harvard.edu

(Tentative) Office Hours:
Wednesday 2-3pm

WJH 7th Floor (To reach me, ring the red buzzer next to my name, and I'll meet you at the door).

Dániel Barabási
barabasi@fas.harvard.edu

Office Hours:
Wednesdays 4-5pm

Room 636, 1 Brattle Square, Cambridge, MA 02138, USA

Mengmi Zhang
Mengmi.Zhang@childrens.harvard.edu

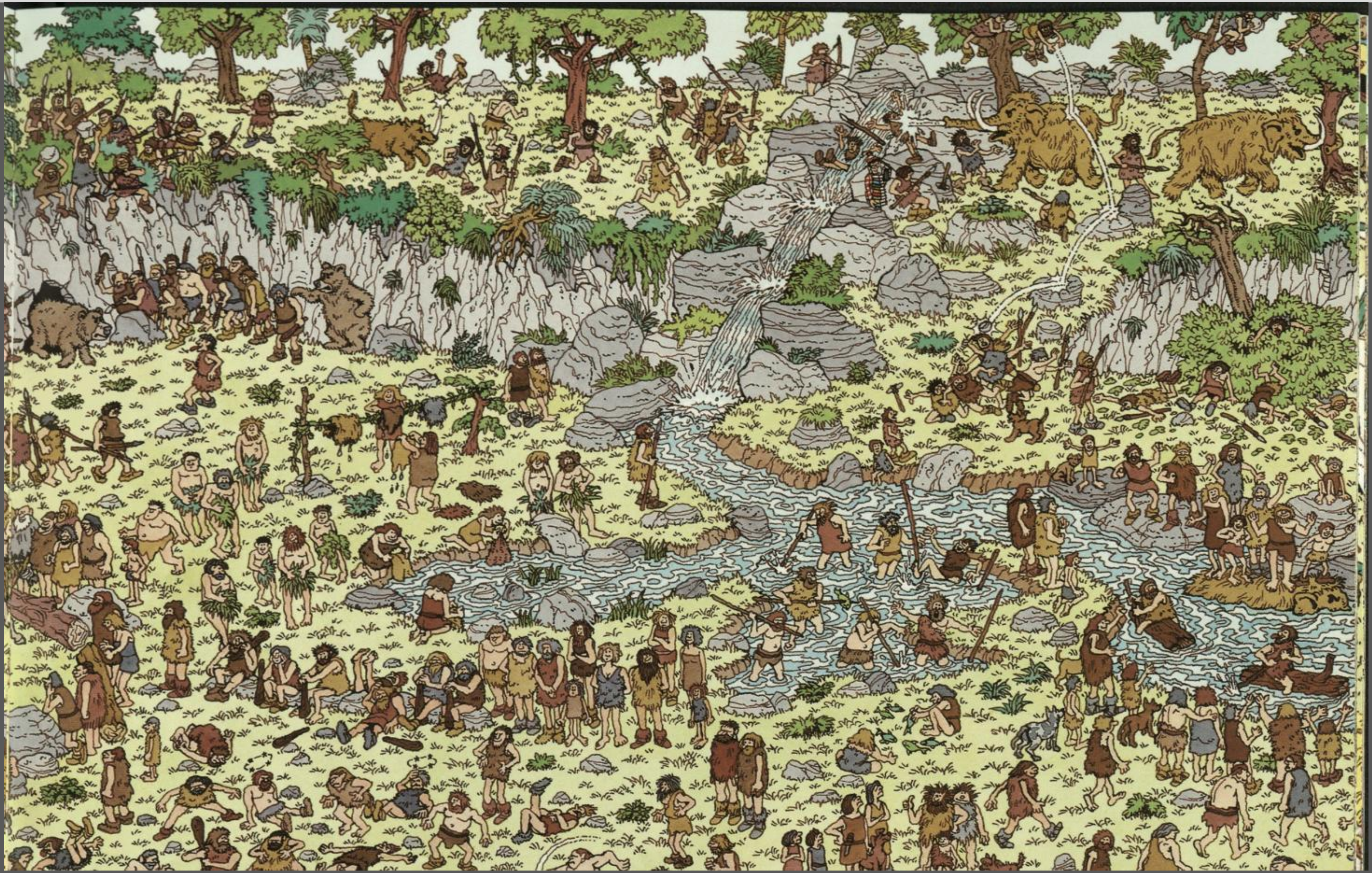
Office Hours:
Wednesdays 2-3pm

Center for Life Science, 13th Floor, 3 Blackfan Circle, Boston, MA 02115

if in Cambridge (by email appointment)

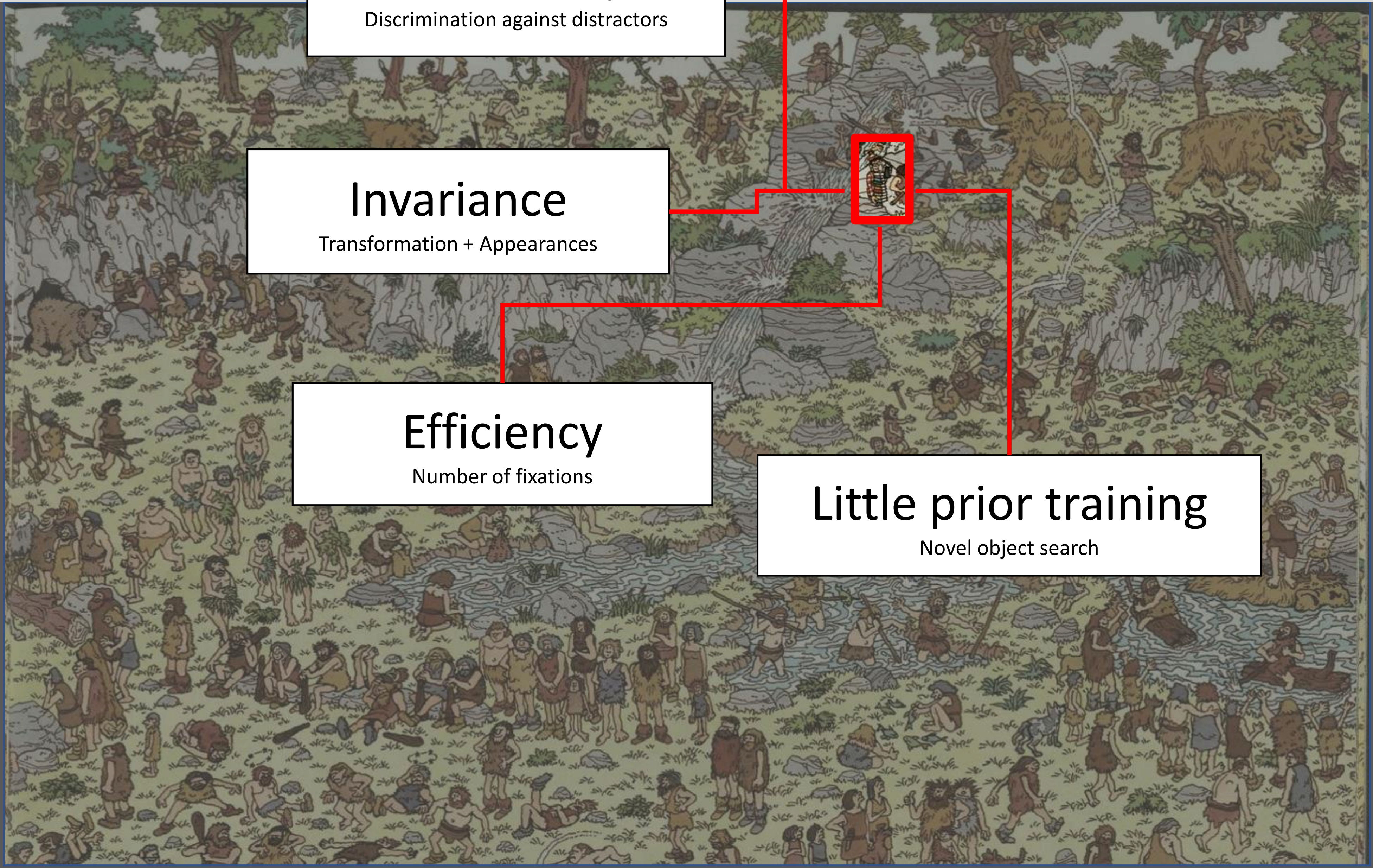
Finding Waldo!

Target



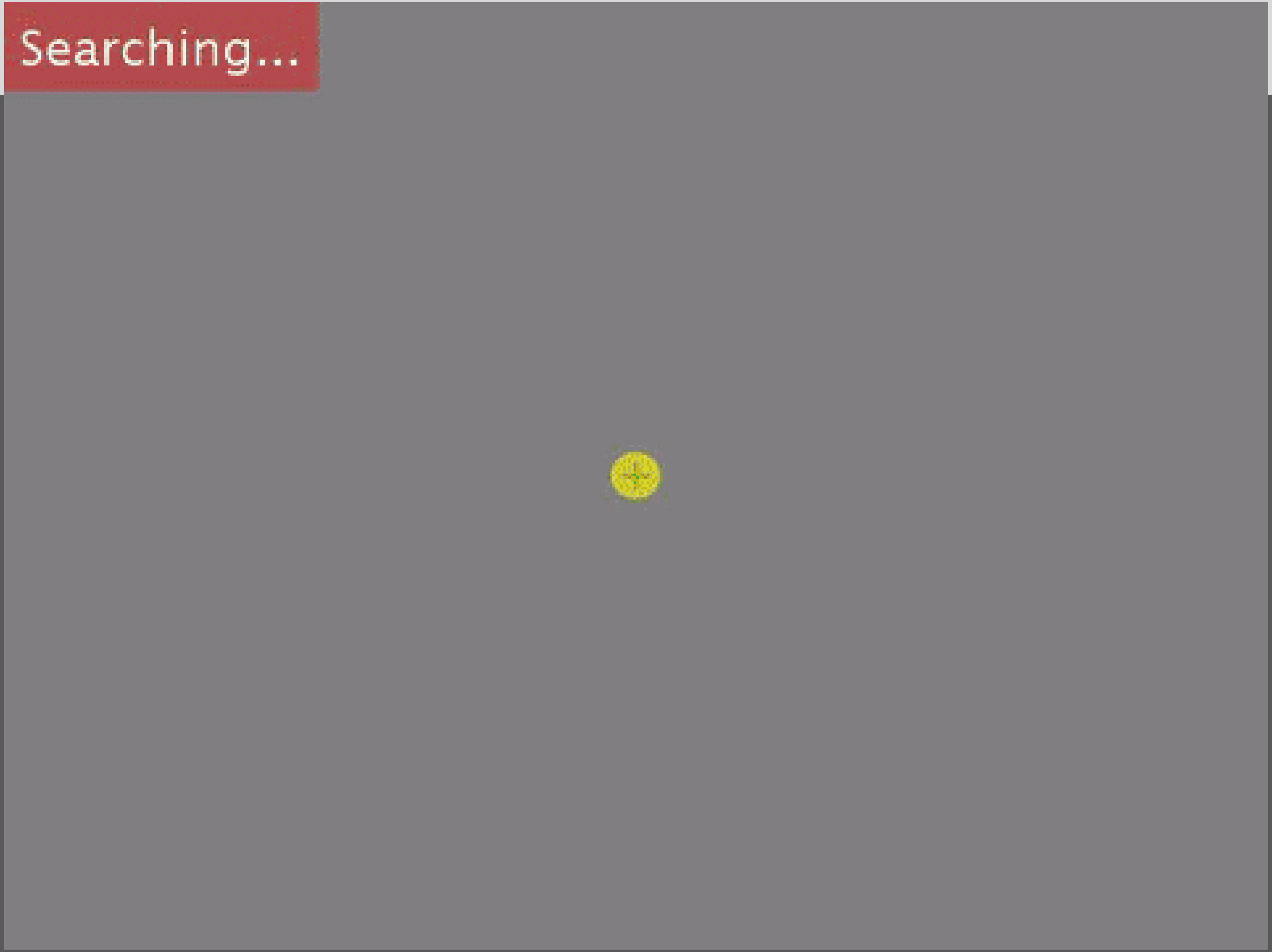
Finding Waldo!

Target



Is this sequence of eye fixations from humans or machines?

Yellow dots
(Eye Fixations)

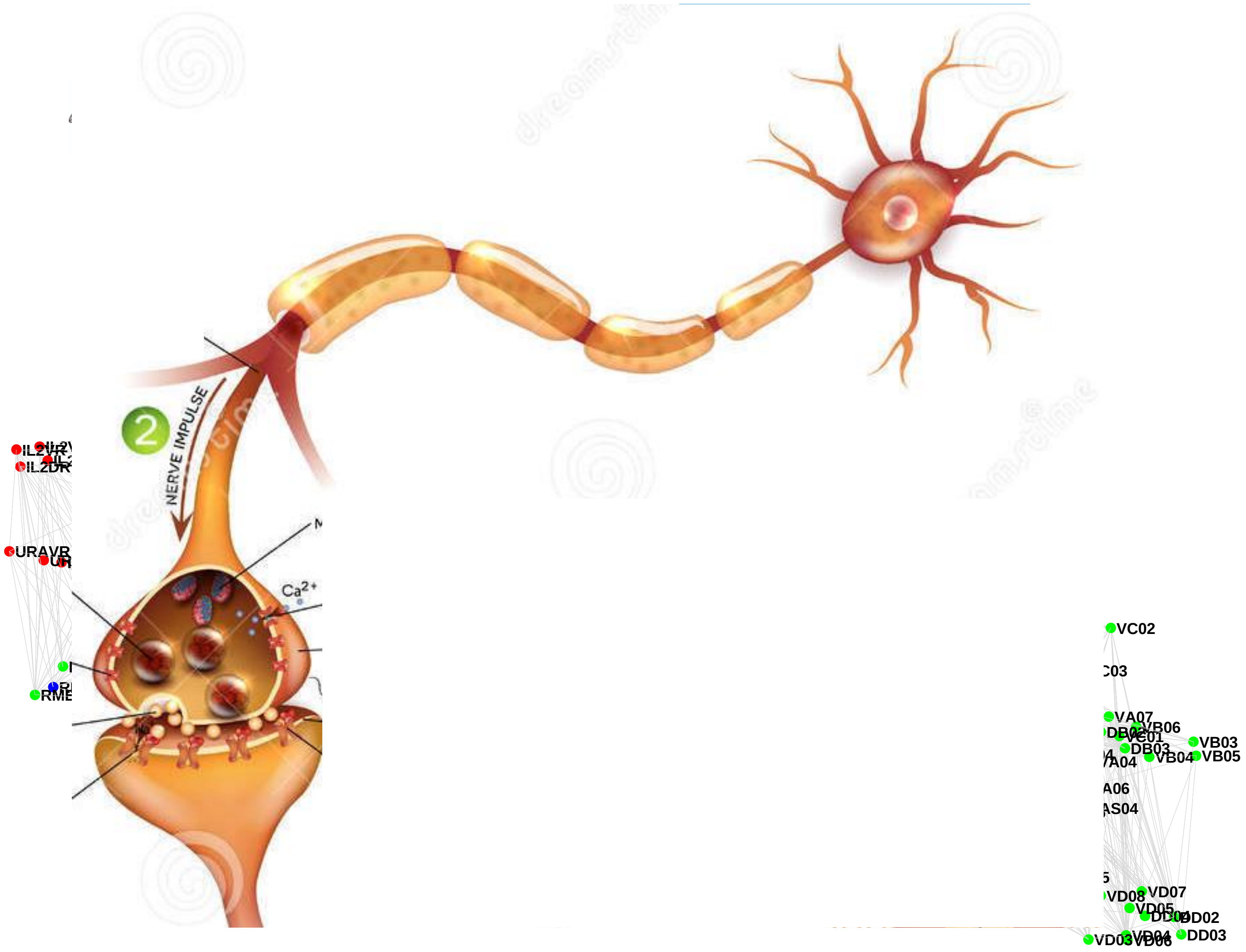


TA Introduction

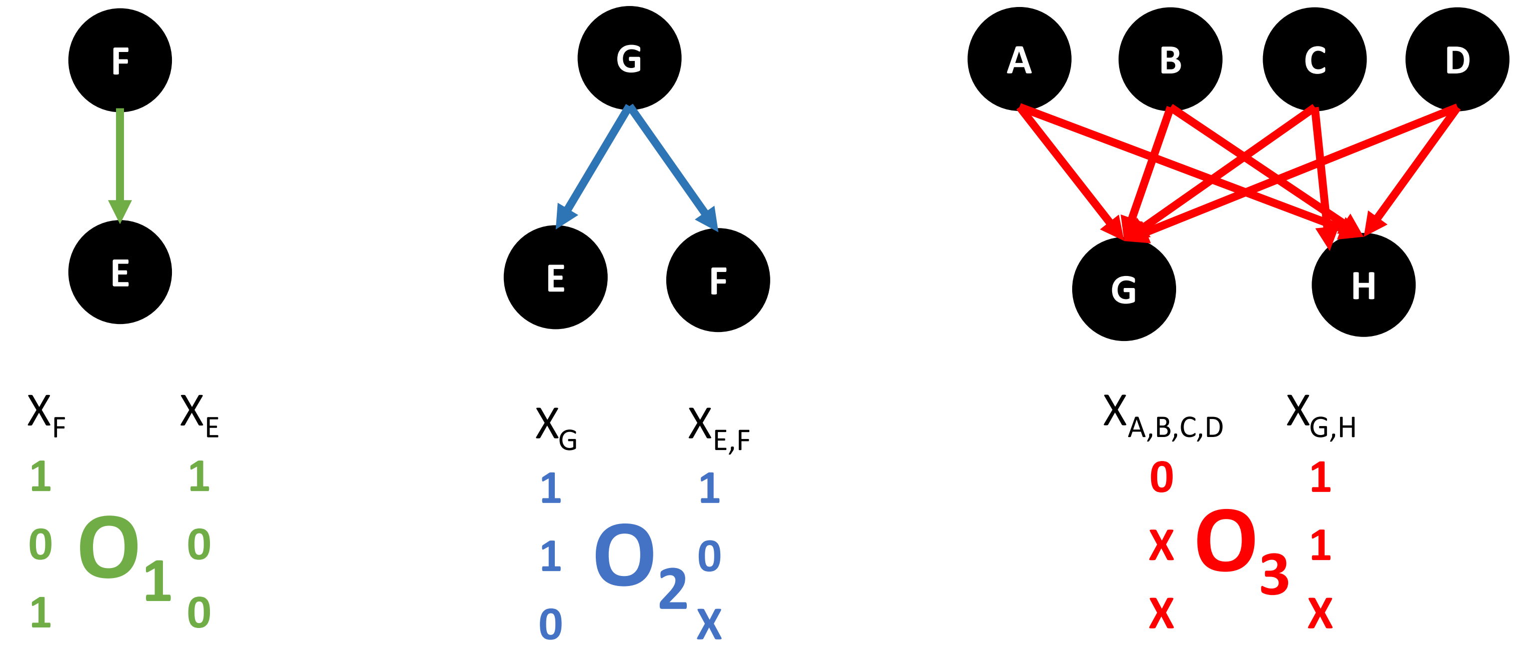
Dániel Barabási
barabasi@fas.harvard.edu

Office Hours:
Wednesdays 4-5pm

Room 636, 1 Brattle Square,
Cambridge, MA 02138, USA

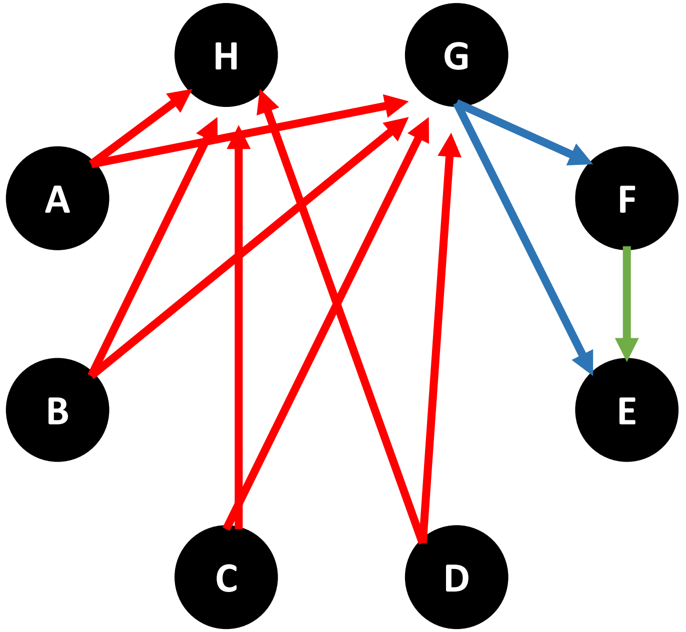


Neuron	X
A	0 0 0
B	0 0 1
C	0 1 0
D	0 1 1
E	1 0 0
F	1 0 1
G	1 1 0
H	1 1 1

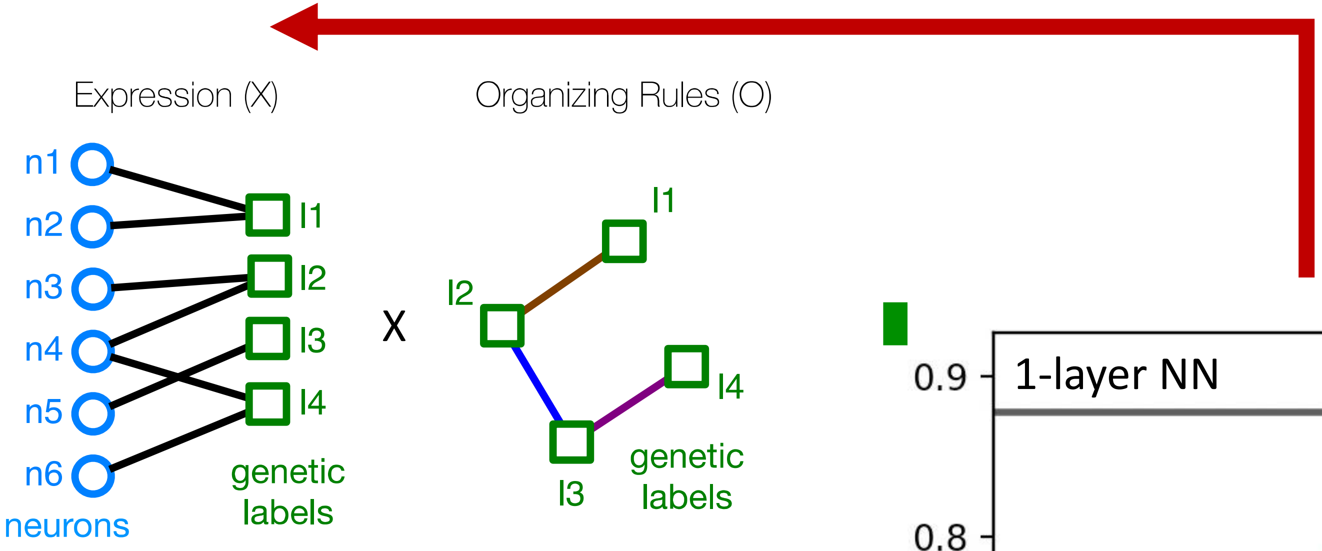


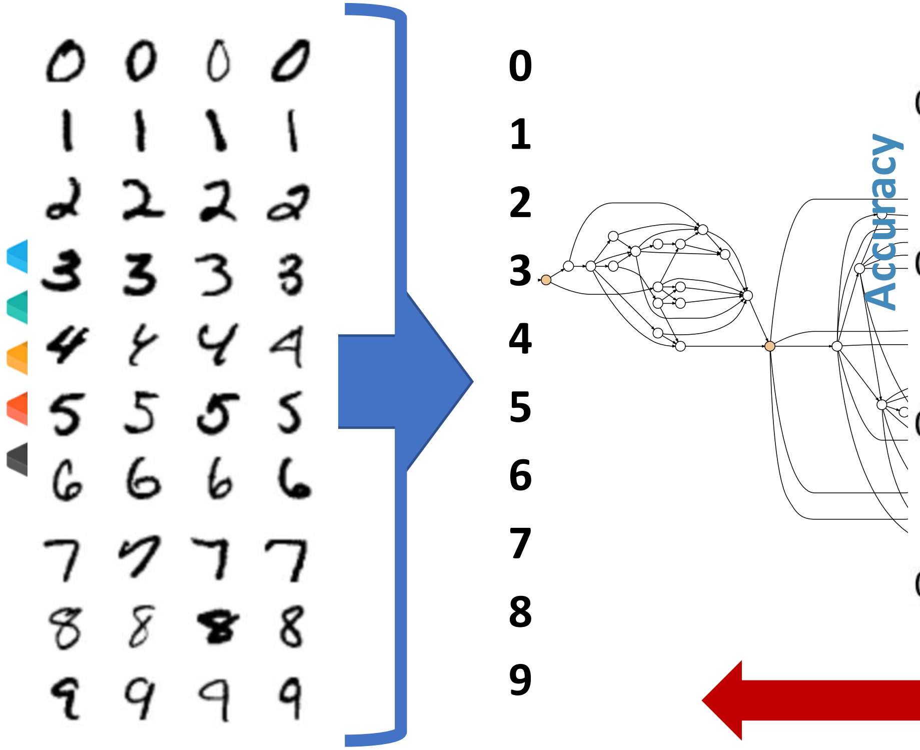
$$B = XOX^T$$

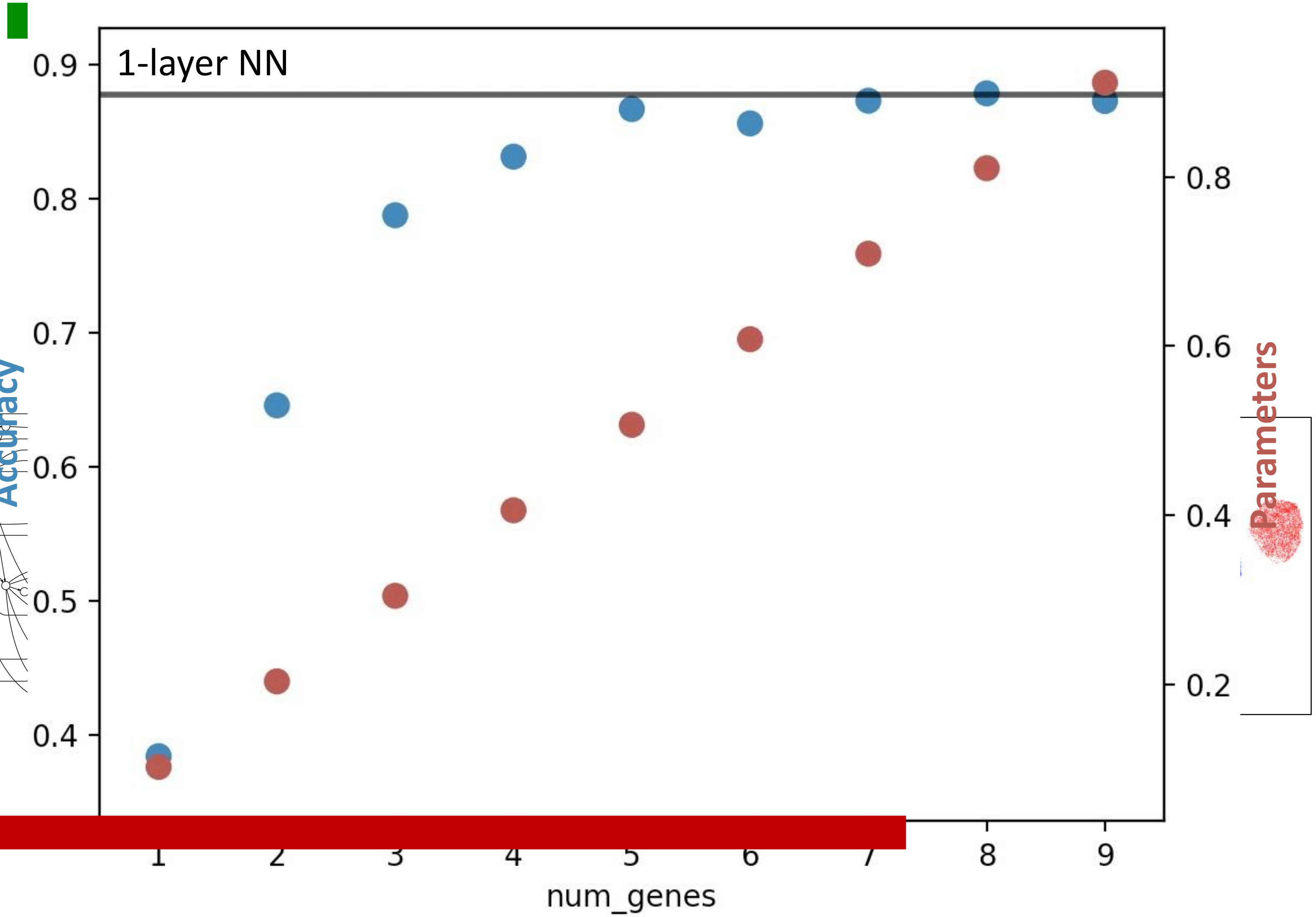
B : Adjacency Matrix
X: Neuronal Identity
O: Operator



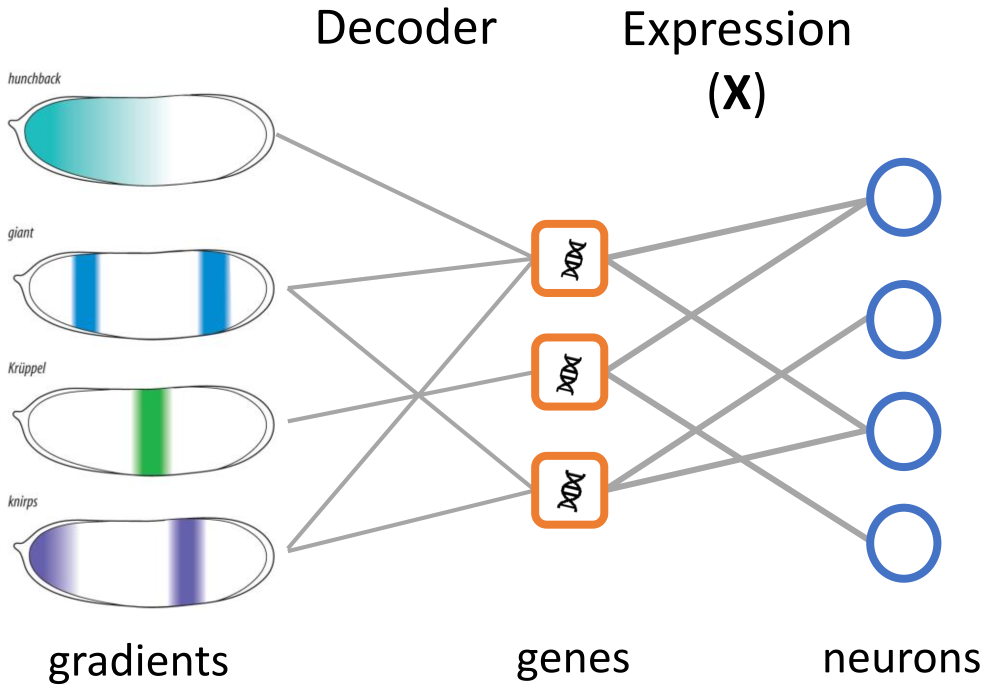
Neuroevolution Model

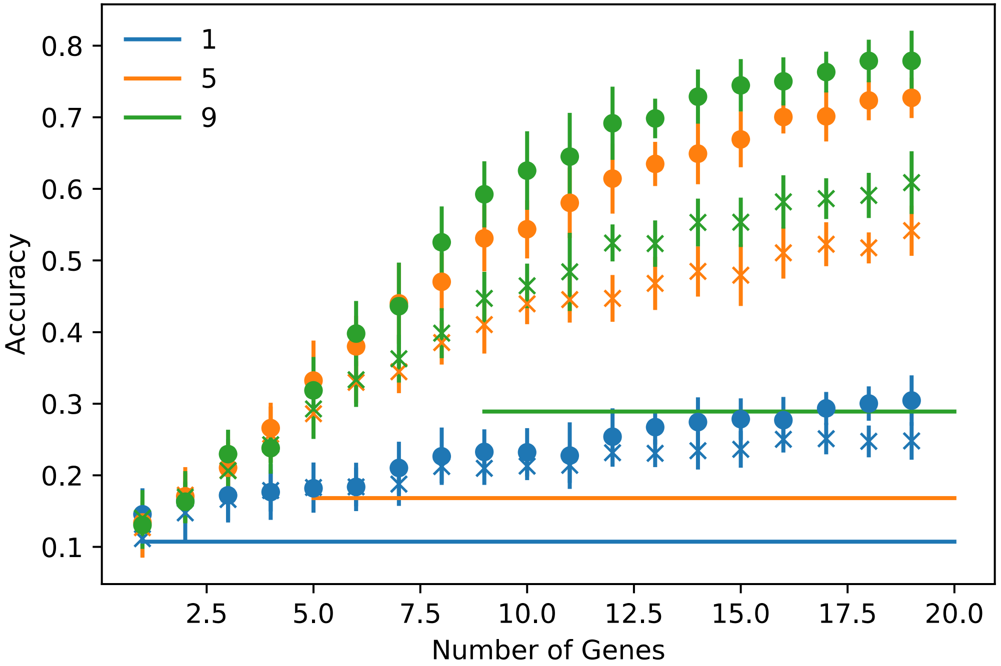


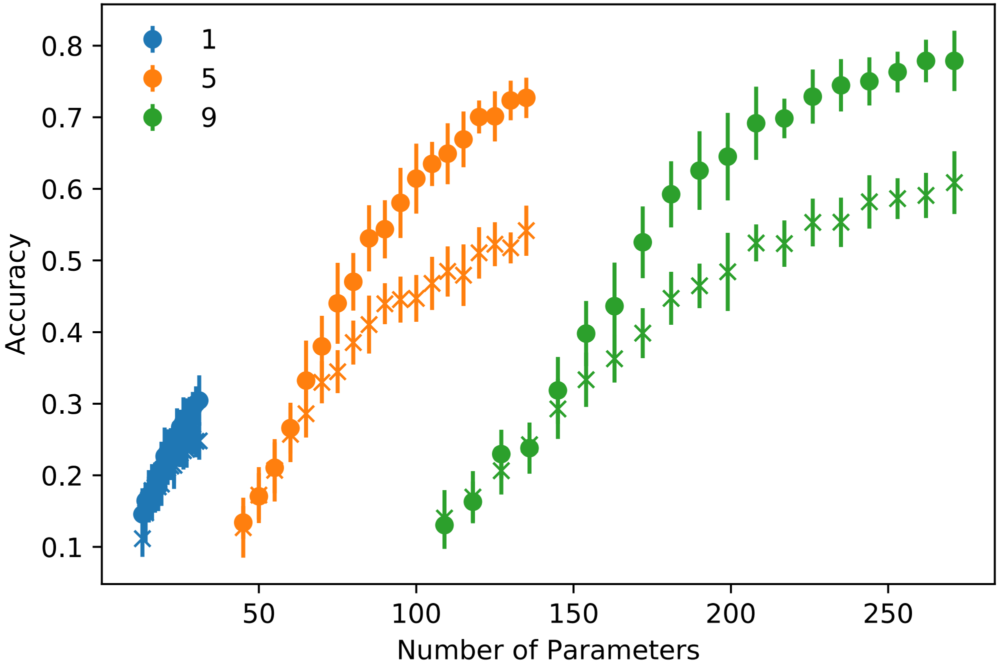




Gradient decoding for Cell ID







Colin Conwell (Coco)

Graduate Student, Vision Sciences Laboratory
Harvard University

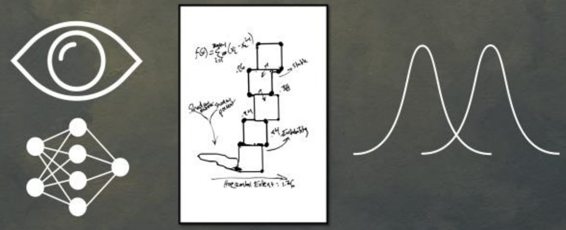
How do we make meaning out of sensory funk?

Interests

Means & Methods

Machine Psychophysics

How can we more systematically compare human and machine behavior?



Can we introduce 'cognitive loss functions' to better guide learning?

Unsupervised Learning

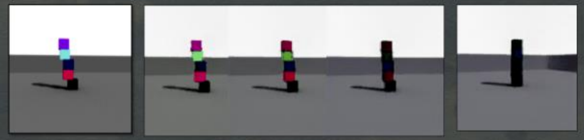
What aspects of human perception and development can we model with unsupervised learning algorithms?

How can we make our algorithms evolve?

Empirical Substrates

Intuitive Physics

Interpolating Between Physical States with a Variational Autoencoder



Input (Designed) Intermediate States (Learned) Output (Learned)

Computational Aesthetics?

BubbleView (Zylinskii & Co., 2015) **Heatmapping**

Can we better decode aesthetics from behavior?
Can we use deep nets to influence aesthetics?

Articulation

How do we translate the subsymbolic fireworks of neurons into the abstract symbols of natural language?

Peabody Essex Museum

Title Image: Alfred Thompson Bricher, 'Time & Tide' (1873)

Neuro 140: Biological and Artificial Intelligence

Logistics:

Class website

https://canvas.harvard.edu/courses/67701 [Login required]

http://bit.ly/BAI2019 [No login required]

Location: Northwest Labs, B101

When: Tue and Th 3-4:15pm

Web will include: Slides, Suggested Reading, Videos (after class), link to other information

Class Logistics: Grades

Midterm project proposal (<=2 pages) and presentation: 15%

One-on-one discussion with (>=2) TAs before midterm presentation is compulsory

Final project presentation: 15%

Final project report: 70%

Learning by doing: Projects

List of projects: https://docs.google.com/document/d/1MsQJPkYiCPbkY8J2LRikfOXAt4bojBBtn4ah8wiMOD4/edit

- Fill in 30-seconds background survey (link also provided on CANVAS): https://forms.gle/SvHfGP393qt8mFMC9
- All projects are computational. Basic programming skills are required: “hello world” and arithmetic operations (+/-/×/÷)
- Students can use open-source codes. MUST properly cite resources. MUST Submit code in the final report.
- Different students can choose the SAME project topic; but each student must work on his/her OWN project
- Choose project early (ample time for preparation)
- $0 <$ Number of projects < 2
- Customize your own project: requires approval by TAs and Prof. Kreiman

Due dates

- Jan 28th: Submit class roster: https://forms.gle/ru8i4uCXskdAdyd26
- Jan 31st: Submit background survey: https://forms.gle/SvHfGP393qt8mFMC9
- Feb 12th: Upload project titles to CANVAS
- March 2nd – 8th: One-on-one discussion with TAs
- March 9th: Upload project proposal to CANVAS
- March 9th: Upload project slides for midterm presentation to CANVAS
- March 10th and 12th: Mid-term project presentations
- Apr 22nd: Upload Final project to CANVAS
- Apr 22nd: Upload Final project slides for final presentation to CANVAS
- Apr 23th and 28th : Final project presentations

Learning by doing: Projects

- Recognition of doodles
- Face recognition with Arduino
- Action recognition in controlled datasets
- The problem of parameters in linear systems
- The role of color in visual recognition
- Size and Position invariance in visual recognition
- Continual Learning for video games
- Create your own object recognition challenge

To be continued…

Have a wonderful semester ahead!

Feel free to interrupt during the class, and interact with faculty members before/during/after classes!