

COLIN MCDONNELL

LEARNING

THE PAPERS

- ▶ Innate attentional biases and associated effects
- ▶ Thalamocortical loops for prediction learning
- ▶ Method of comparing human classification capability to neural networks

how the brain learns
and how it informs machine learning
and vice versa
with no pretensions of completeness

STRUCTURE OF TALK

- ▶ What is learning
- ▶ Neuroscience inspires machine learning
- ▶ Machine learning guides neuroscience

SOME DEFINITIONS

what do I mean neuroscience?

- Affective neuroscience
- Behavioral neuroscience
- Cellular neuroscience
- Clinical neuroscience
- Cognitive neuroscience
- Computational neuroscience
- Cultural neuroscience
- Developmental neuroscience
- Evolutionary neuroscience
- Molecular neuroscience
- Neuroengineering
- Neuroethology
- Neuroheuristics
- Neuroimaging
- Neuroinformatics
- Neurolinguistics
- Neurophysics
- Neurophysiology
- Neuropsychology
- Paleoneurology
- Social neuroscience
- Systems neuroscience

what do I mean ML?

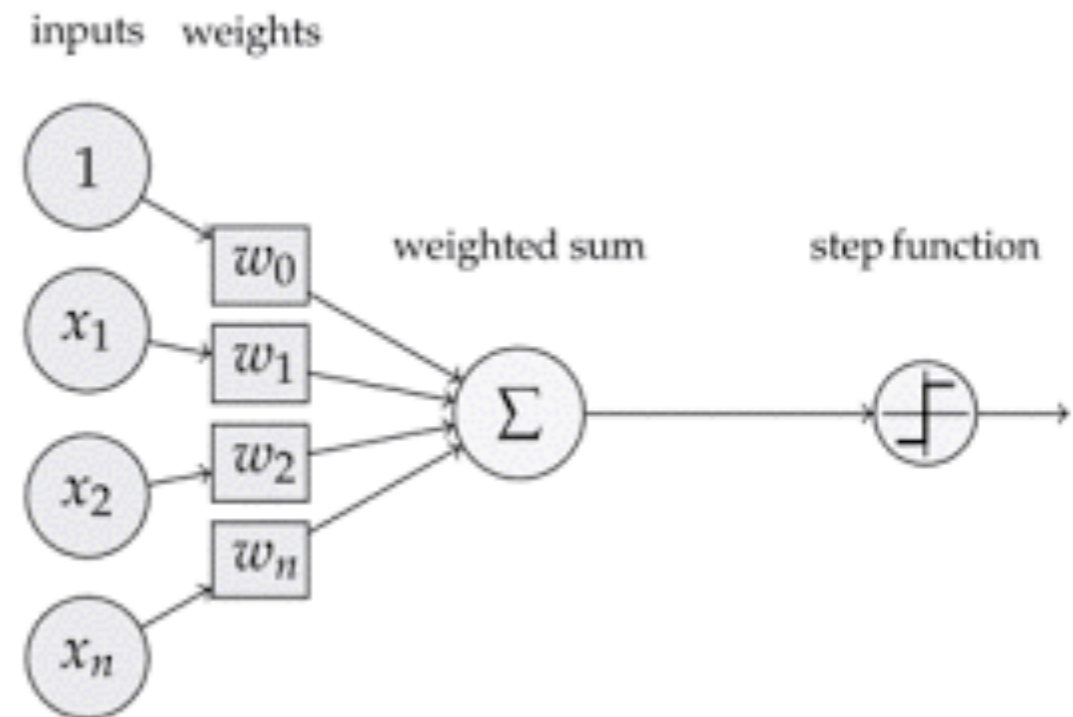
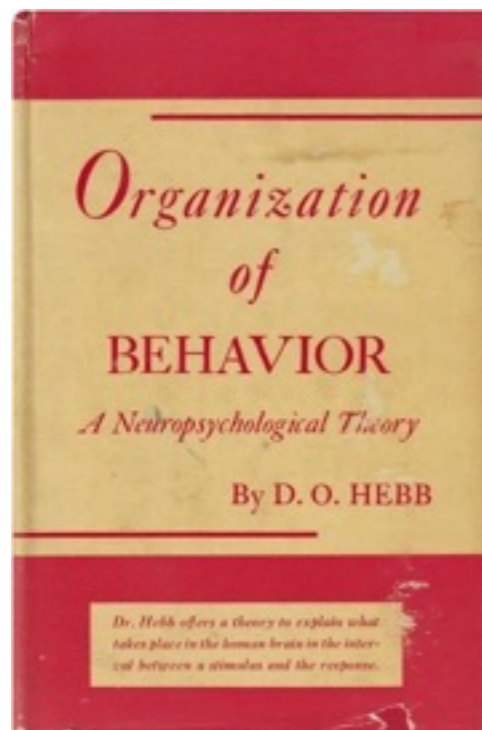
NEURAL NETWORKS

WHAT IS LEARNING

- ▶ pattern recognition
- ▶ finding an efficient encoding
- ▶ discovering beneficial behavior based on past observations
- ▶ abstraction
- ▶ Australian or European?

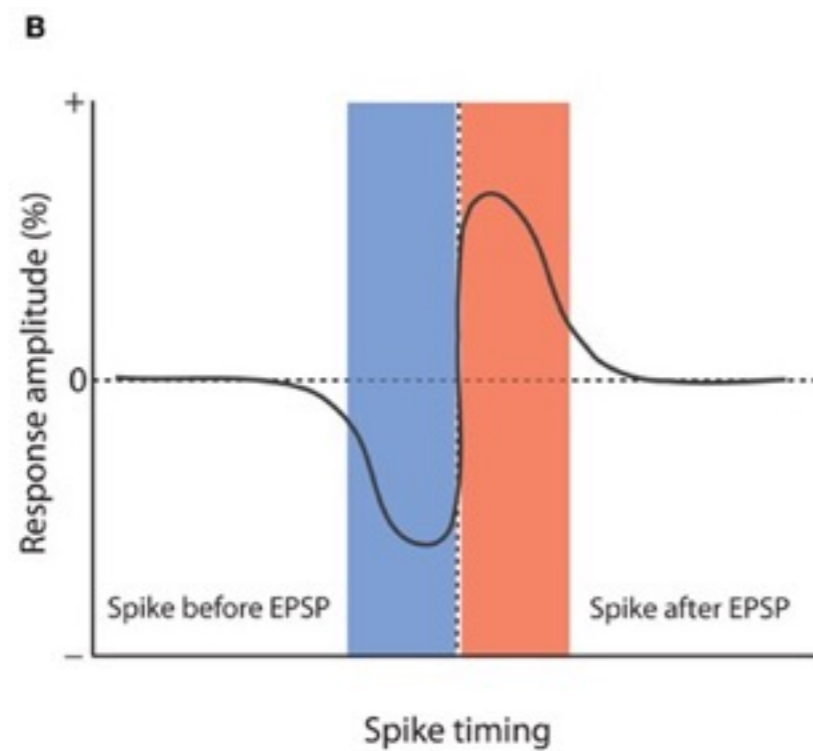
NEUROSCIENCE INSPIRES MACHINE LEARNING

Hebbian learning



spike timing dependent plasticity

404



synaptic re-wiring

sort of a 404, but...

sparsity

winner-take-all learning

ReLU?

many excitatory synapses
(for temporal learning?)

neurons that remember
window of past
inputs/firings

**Why Neurons Have Thousands of Synapses, A
Theory of Sequence Memory in Neocortex**

[Jeff Hawkins](#), [Subutai Ahmad](#)

(Submitted on 31 Oct 2015 (v1), last revised 1 Dec 2015 (this version, v2))

dendritic computation

layers collapse to neurons
mathematically equivalent

reflection of action
potential into input
dendrites

gradient learning

going smaller...

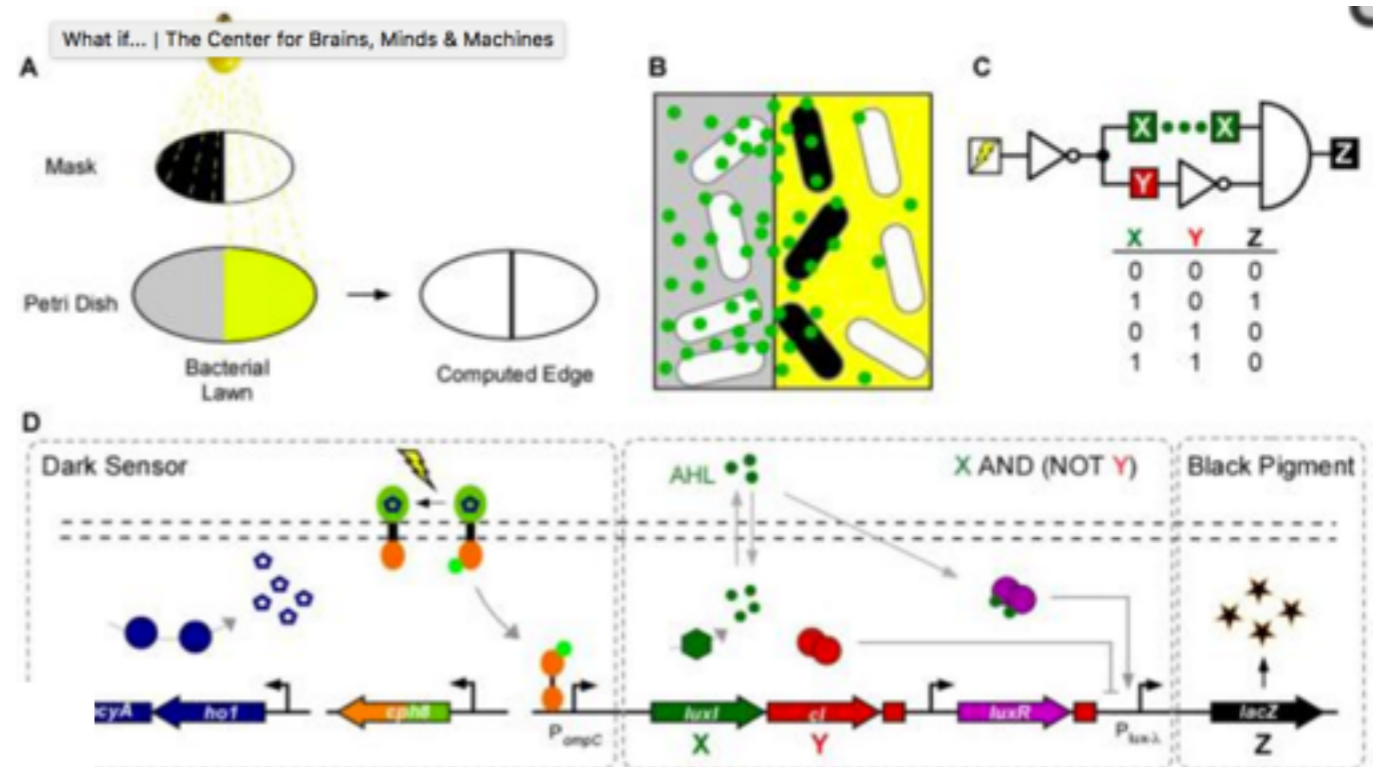
molecular diffusion

neuromodulators

neuropeptides

nitrous oxide

message-passing networks



retrograde chemical
signaling

reverse connections

going bigger...

attractor circuits

should arise organically...

(later: LeabraTI)

404

communicating brain
regions

eventually, interacting
networks:

top-down feedback [Ullman 2015]

deep reinforcement learning

basal ganglia for decisions

variable binding

multi-sensory learning

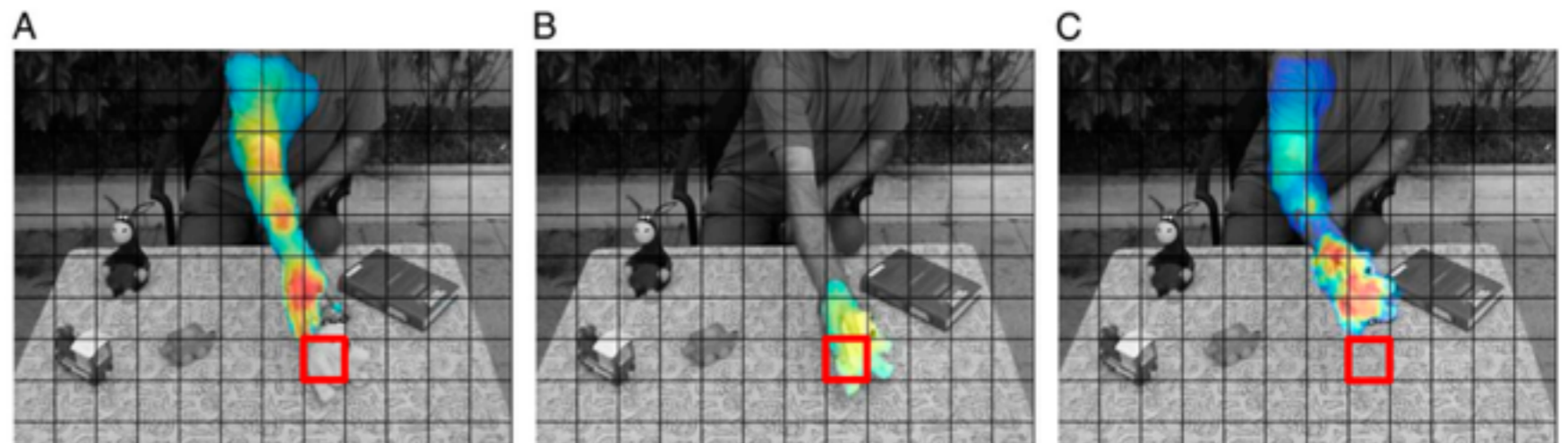
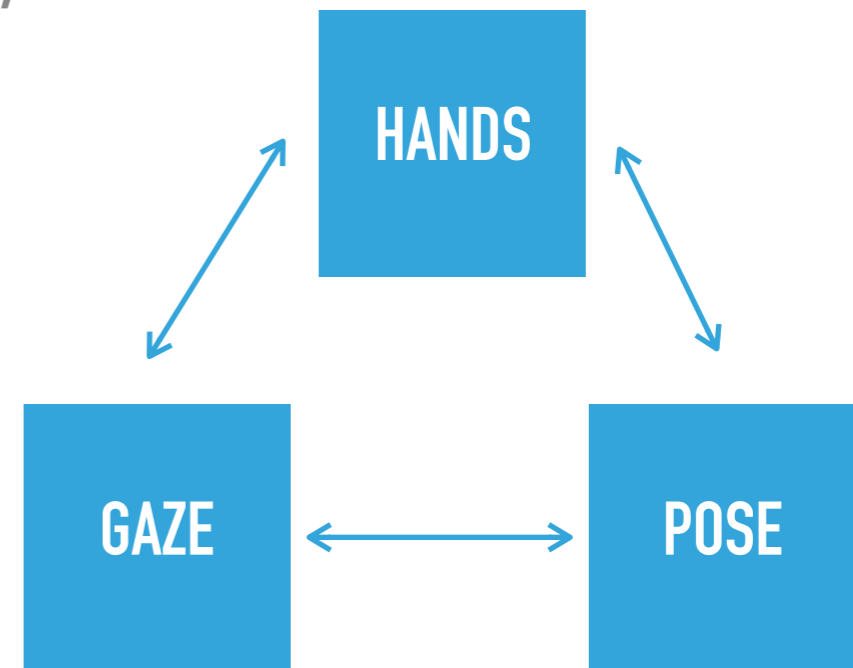
information routing

innate biases

attention mechanisms

"feature weights"

- ▶ bias network towards “mover events”
- ▶ first pass:
 - ▶ high precision, low recall
- ▶ second pass:
 - ▶ tracking/continuity
 - ▶ body context
- ▶ third pass:
 - ▶ gaze
- ▶ fourth pass
 - ▶ mirroring
 - ▶ own-hands
- ▶ comparison
- ▶ but was it necessary?



From simple innate biases to complex visual concepts

Shimon Ullman^{1,2}, Daniel Harari¹, and Nimrod Dorfman¹

Department of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 76100, Israel

implicit labeling/
proto-concepts
[Ullman 2012]

semi-supervised labeling

basins of attraction as reward
signal generator

MACHINE LEARNING GUIDES NEUROSCIENCE

supervised learning
unsupervised learning

supervised learning

supervised learning

404?

problems

error-driven
backpropagation

- ▶ requires transmission of exact derivatives over distance
- ▶ neurons communicate through (stochastic?) spikes
- ▶ temporal synchronicity
- ▶ what are the targets?

supervised learning

derivative backpropagation

hyper polarization of
dendrites
+
reverse action potential

supervised learning

training epochs

hippocampus-mediated
replay

supervised learning

momentum gradient descent

plasticity modulation

supervised learning

dropout

stochastic models of
neural spiking

supervised learning

layer-by-layer training

developmental drop-off in
cortical plasticity

unsupervised learning

unsupervised learning

404

auto-encoders

layer-size bottleneck

lateral inhibition

unsupervised learning

predictive learning

babies understand
spatiotemporal coherence

thalamocortical loops!
[O'Reilly 2014]

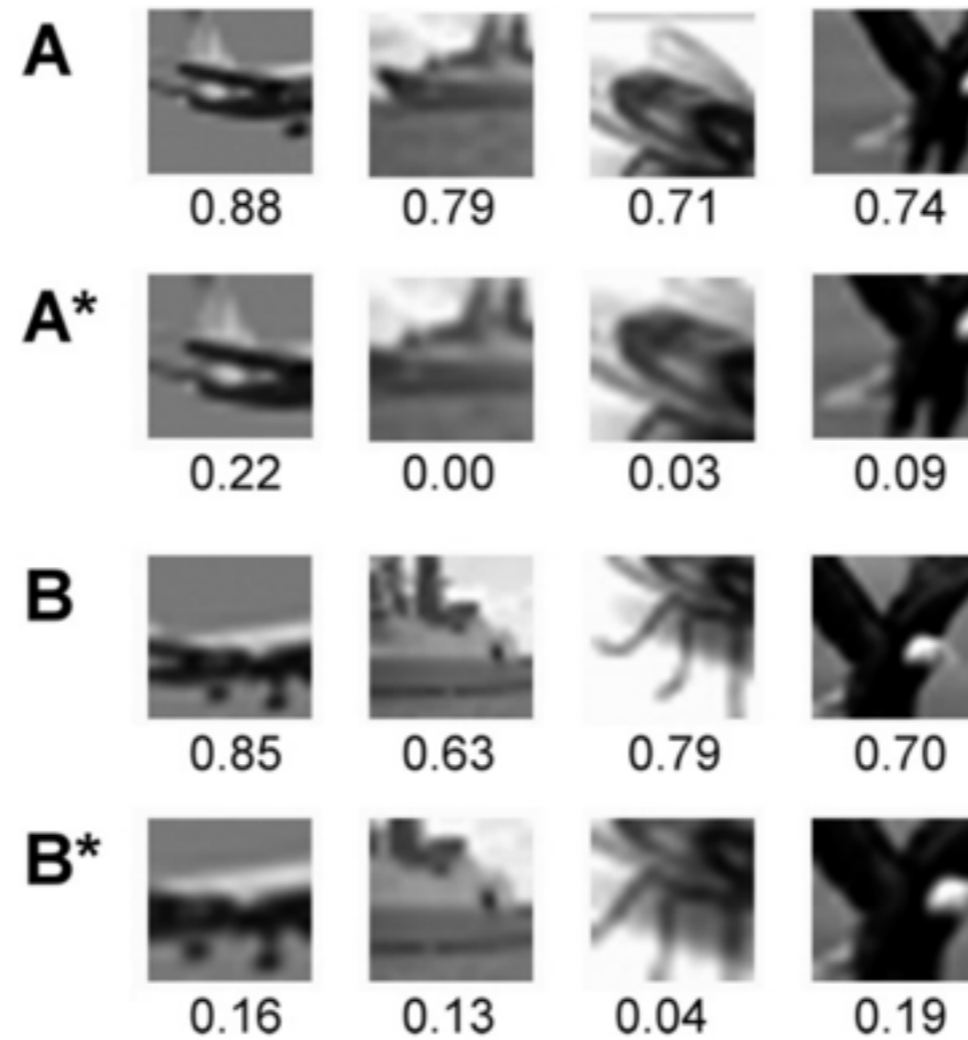
unsupervised learning

prediction learning

babies understand
spatiotemporal coherence

thalamocortical loops!
[O'Reilly 2014]

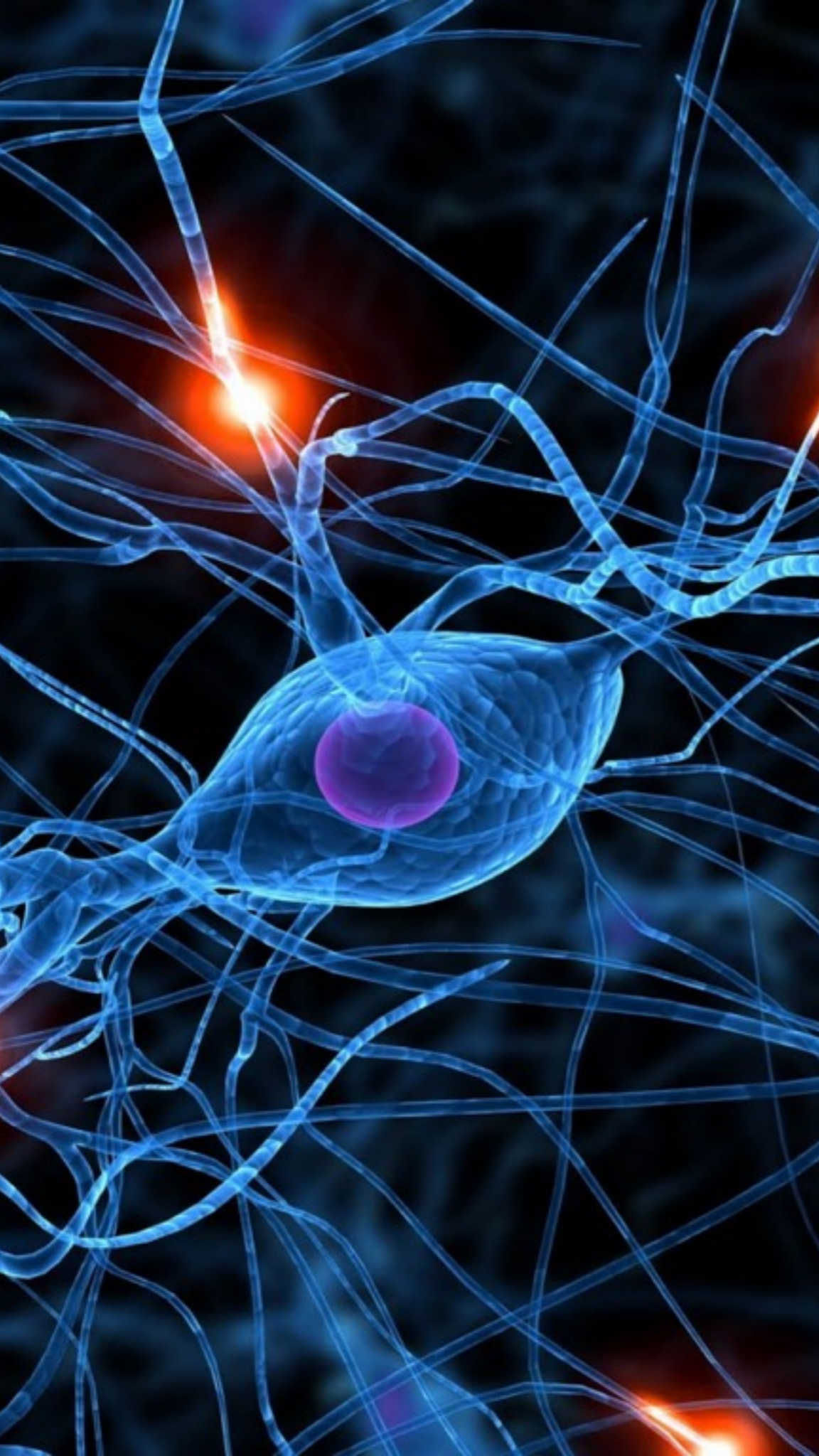
INTER-DISCIPLINE COMPARISONS



[Proc Natl Acad Sci U S A](#). 2016 Feb 16. pii: 201513198. [Epub ahead of print]

Atoms of recognition in human and computer vision.

[Ullman S](#)¹, [Assif L](#)², [Fetaya E](#)², [Harari D](#)³.



QUESTIONS