Machine Learning and Neuroscience

Michel Besserve

MPI for Intelligent Systems - MPI for Biological Cybernetics,

Tübingen, Germany.

Introduction

Machine learning (ML)

- Deep learning (DL) a source a major successes,
- New challenges require smaller datasets, robustness guaranties, interpretability,
- Why DL works and how?

Neuroscience

- Major improvement in experimental tools,
- Difficult to make sense of such high dimensional data,
- Major questions remain unsolved.

- Some DL principles have been inspired by brain research (and still are),
- Some DL-based approaches reach or go beyond human performance,
- Humans typically do not use a large amount of labeled data, can easily adapt/generalize/be creative.

Questions:

- What are the similarities and differences between computations in brains and state of the art ANNs?
- Can Neuroscience and ML help each other facing new challenges?

Overview

- Biological vs artificial neural networks,
  - Single neurons, circuits.
  - Plasticity Backprop and biological credit assignment.
- Major brain functions and their neural correlates, applications of causality.
  - Perception,
  - Memory.
- Big questions and theories,
  - The generative approach to Perception,
  - Learning, memory and sleep.

1943: year of the beginning

Let us define a *temporal propositional expression* (a *TPE*), designating a *temporal propositional function* (*TPF*), by the following recursion:

1. A ${}^1p^1[z_1]$ is a *TPE*, where p_1 is a predicate-variable.
2. If S_1 and S_2 are *TPE* containing the same free individual variable, so are SS_1, $S_1 \mathrm{v} S_2$, $S_1 . S_2$ and $S_i . \infty S_2$.
3. Nothing else is a *TPE*.

THEOREM I.

Every net of order 0 can be solved in terms of temporal propositional expressions.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

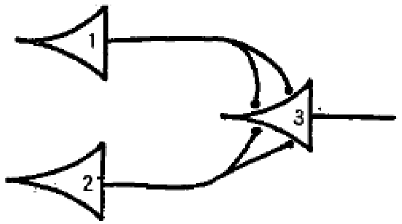
A LOGICAL CALCULUS OF THE
IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,
AND THE UNIVERSITY OF CHICAGO

a

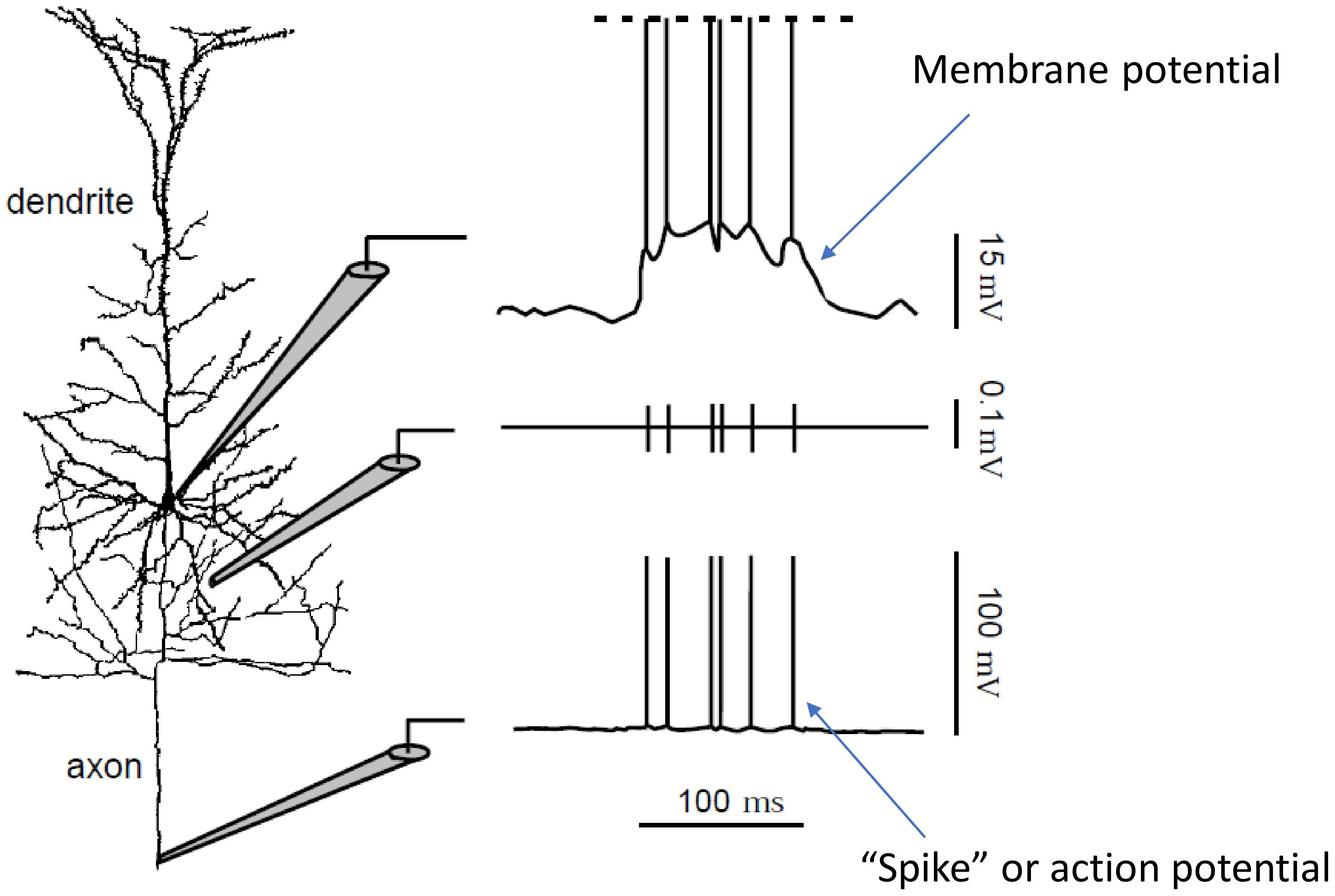
b



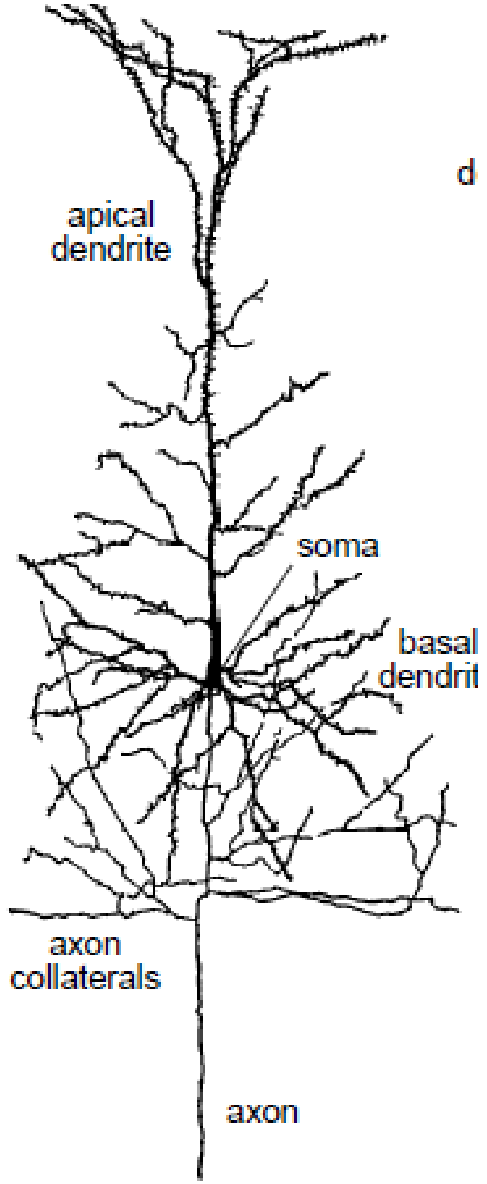
A “Physiological” Turing machine.

Biological vs. Artificial networks

Biological neurons

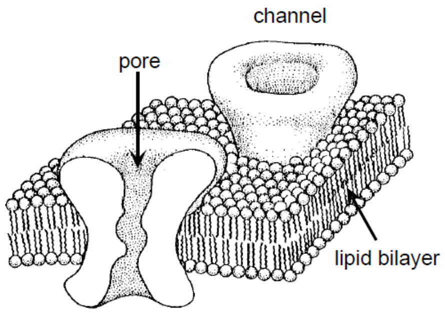


Pyramidal cell

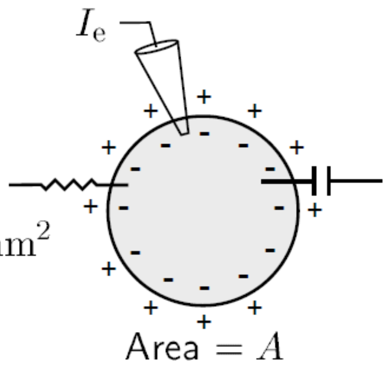


Modified from Dayan & Abbott, 2001

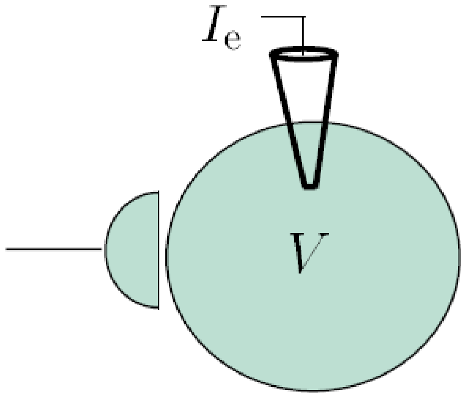
Cell membrane and ion channels

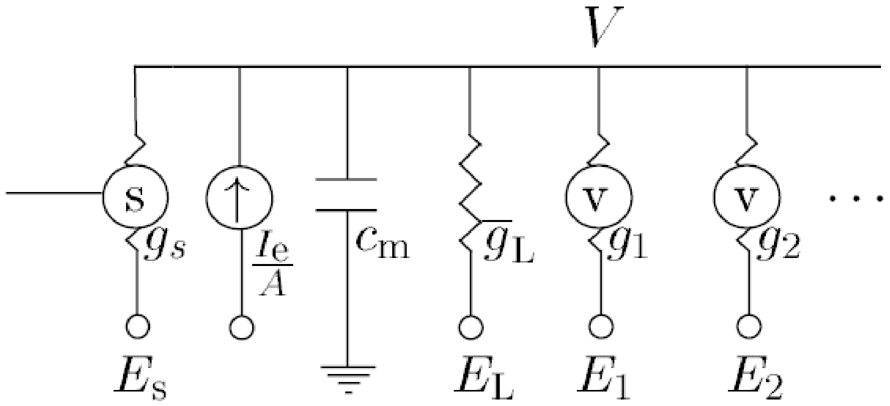


$$\Delta V = I_e R_m$$
$$R_m = r_m / A$$
$$r_m \approx 1\mathrm{M}\Omega \cdot \mathrm{mm}^2$$

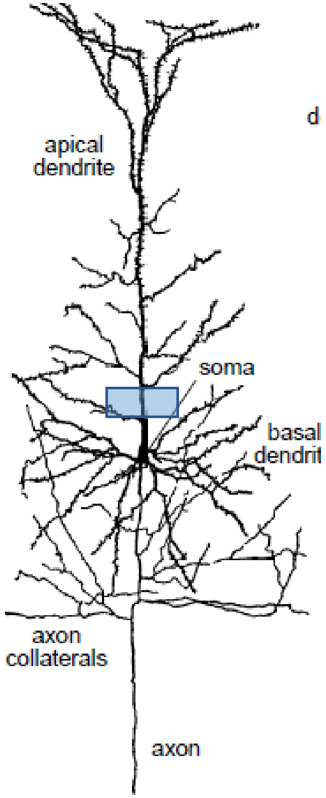


$$Q = C_m V$$
$$C_m = c_m A$$
$$c_m \approx 10\mathrm{nF}/\mathrm{mm}^2$$



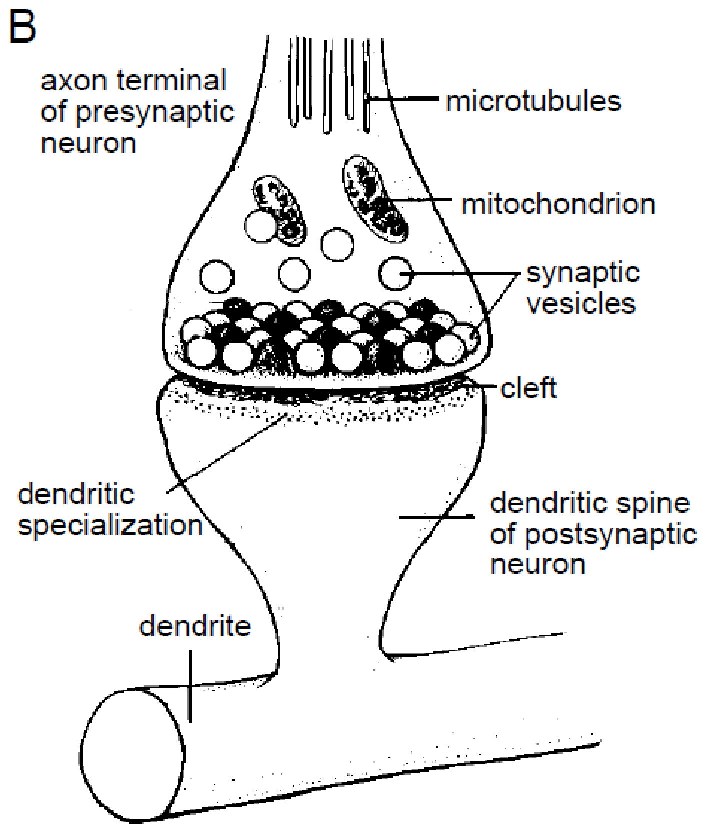


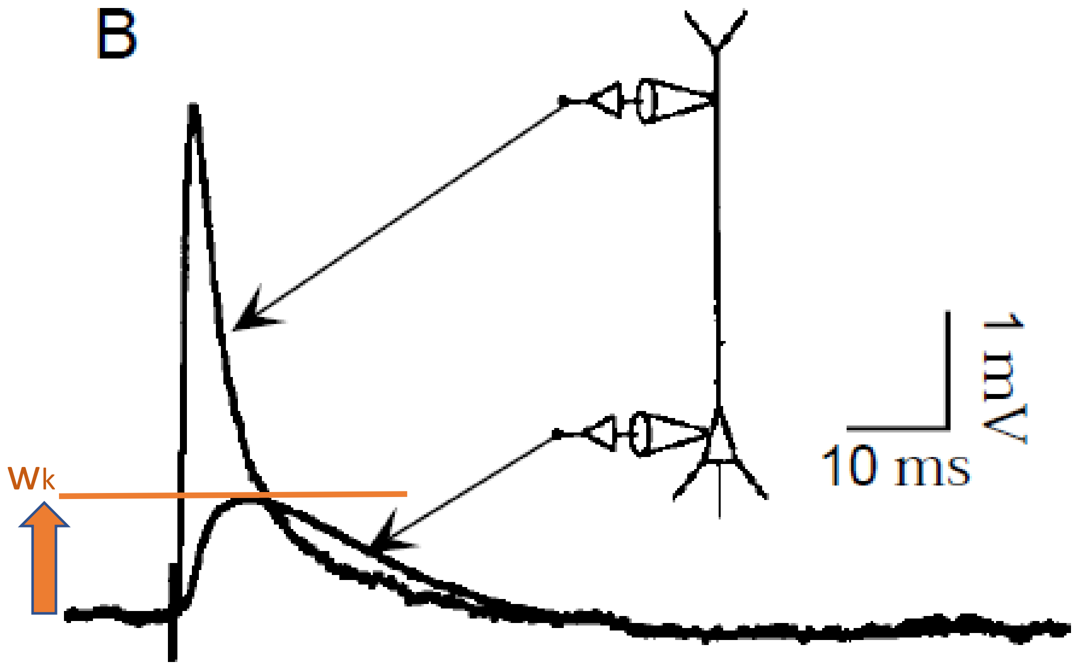
Pyramidal cell

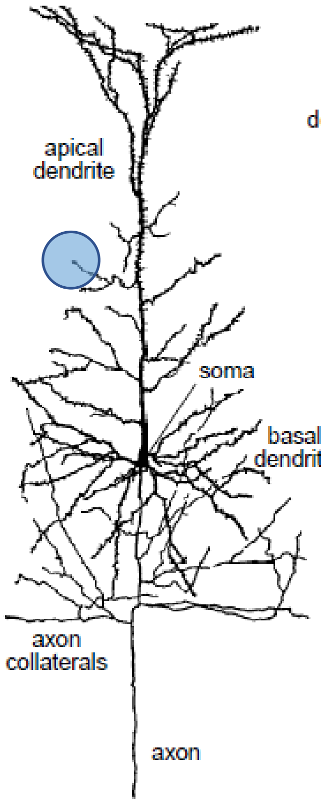


Modified from Dayan & Abbott, 2001

Synapses and post-synaptic potentials (PSPs)







Modified from Dayan & Abbott, 2001

The perceptron (Rosenblatt, 1957): a (static) model neuron

Total post-synaptic current

Synaptic weight vector

Afferent/input activities

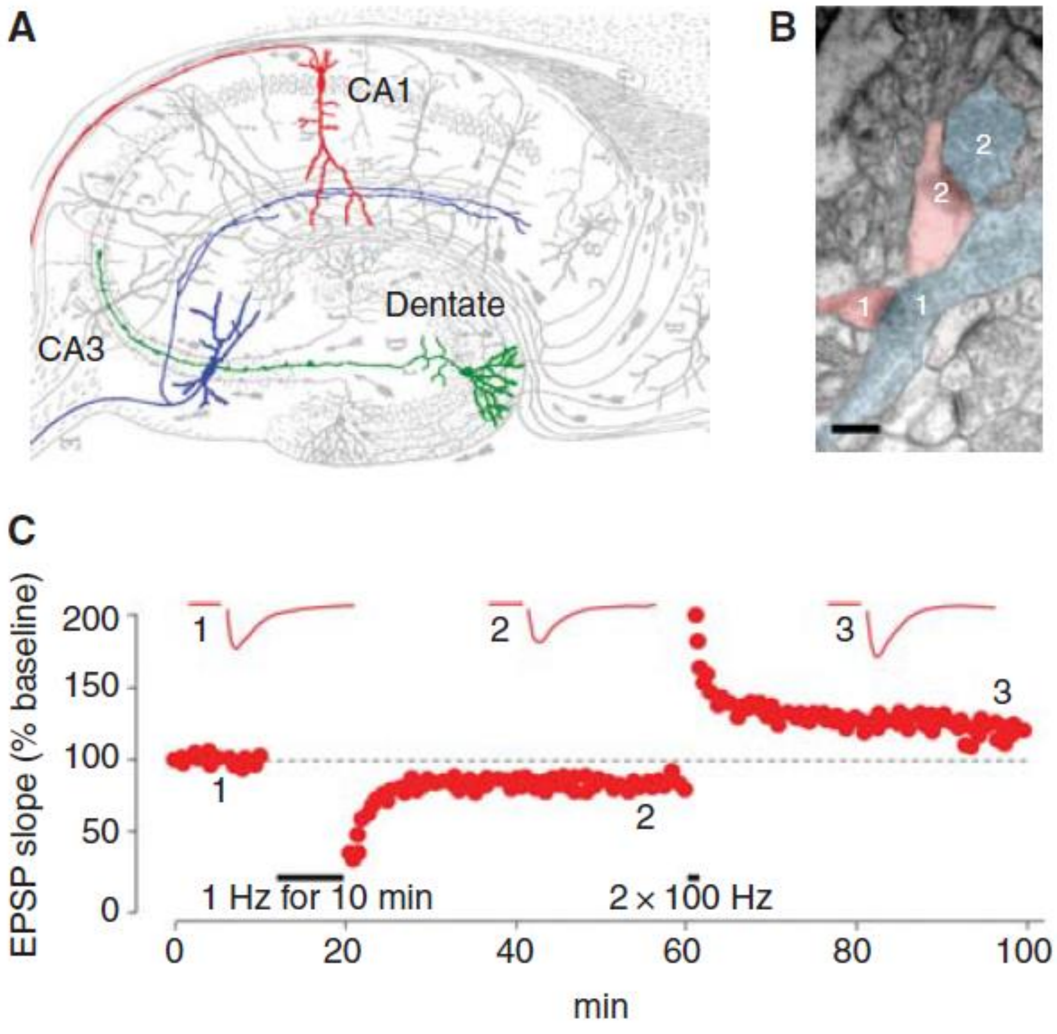
Output Spike

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w}\cdot\mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$

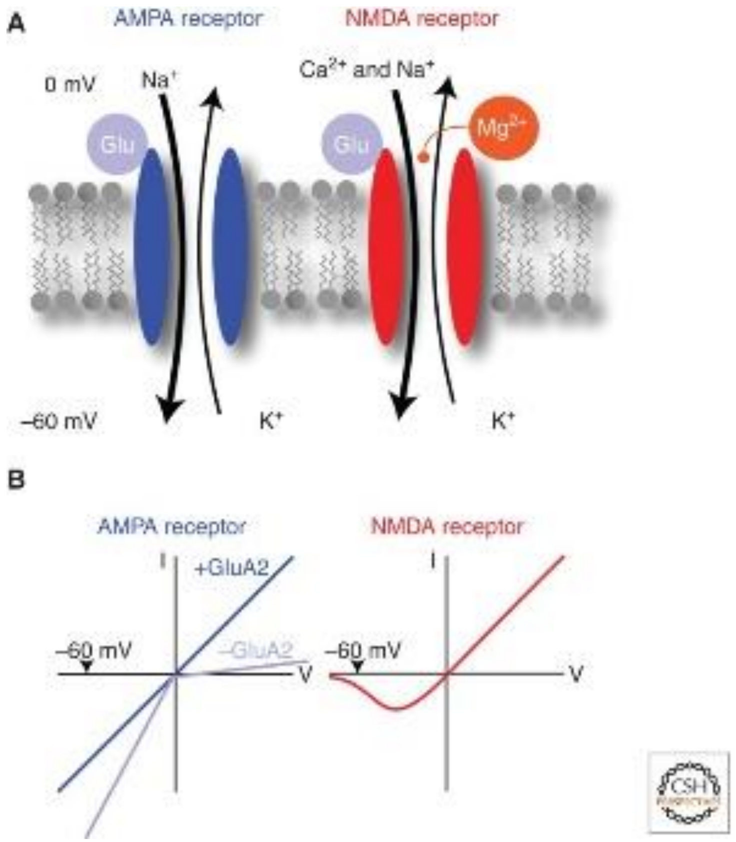
$$\boldsymbol{w}.\boldsymbol{x} = \sum_k w_k x_k$$

Membrane potential

Learning the synaptic weights: Long Term Potentiation/Depression (LTP/LTD)



Lüscher & Malenka, *Cold Spring Harb Perspect Biol.* 2012



The synaptic strength of AMPA receptors is altered by calcium ions flowing inside the cell through NMDA receptors when the membrane is depolarized.

Learning rules

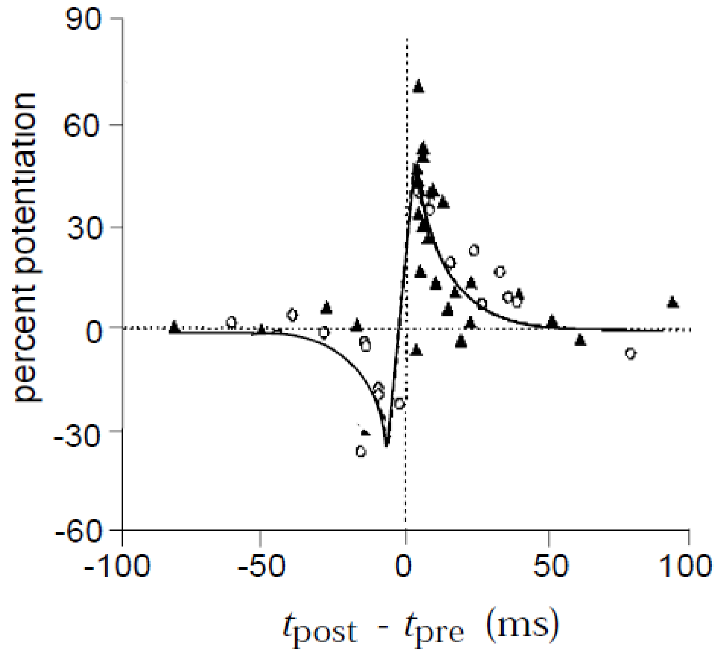
- Hebbian learning (originally rate-based):

$$\tau_w \frac{d\mathbf{w}}{dt} = v\mathbf{u} \,,$$

Post-synaptic activity.

Pre-synaptic activities.

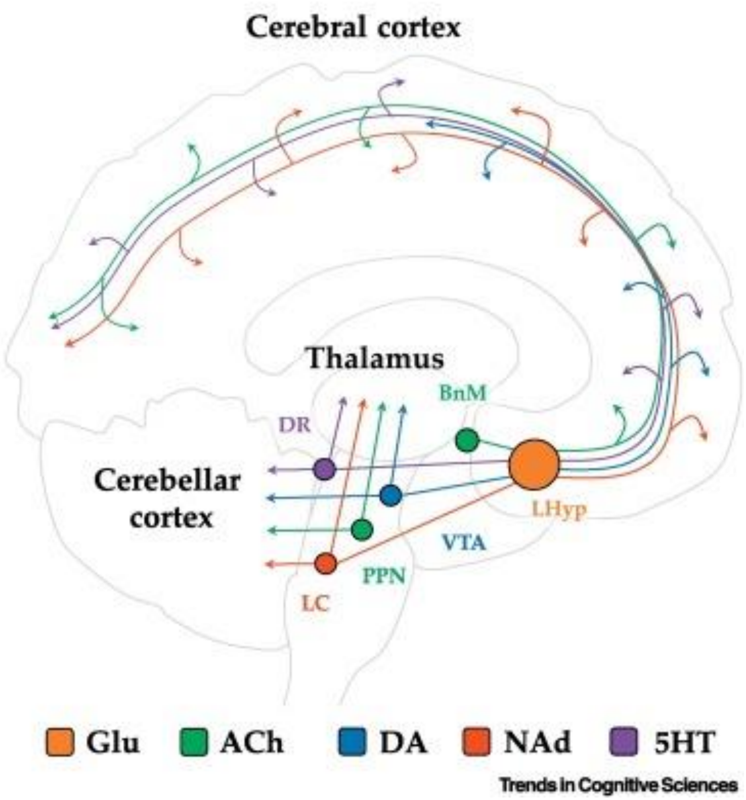
- Spike-Time Dependent Plasticity:
  - Supported empirically, to some extent,
  - But: no good account of LTD



Modified from Dayan & Abbott, 2001

What a neuron can do for its brain?

- Neurons have access to little information available about what the whole network is doing: inputs, outputs, neuromodulators.
- This information must be exploited in order to optimize information processing and minimize metabolic costs (e.g. energy consumption).
- One hypothesis is that neuromodulators encode a form of reward for neurons.
- Principle behind learning in biological neural networks can be investigated with learning theory tools and optimization.



The Selectron (Balduzzi & Besserve, *NIPS* 2013)

$$y = f_{\mathbf{w}}(\mathbf{x}) := H\left(\mathbf{w}^\intercal \mathbf{x} - \vartheta\right), \text{ where } H(z) := \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{else} \end{cases}$$

Definition 1. *Define reward function*

$$R(\mathbf{x}, f_{\mathbf{w}}, \nu) = \underbrace{\nu(\mathbf{x})}_{\textit{neuromodulators}} \cdot \underbrace{(\mathbf{w}^\intercal \mathbf{x} - \vartheta)}_{\textit{margin}} \cdot \underbrace{f_{\mathbf{w}}(\mathbf{x})}_{\textit{selectivity}} = \begin{cases} \nu(\mathbf{x}) \cdot (\mathbf{w}^\intercal \mathbf{x} - \vartheta) & \textit{if } y = 1 \\ 0 & \textit{else.} \end{cases}$$

Constrained reward maximization. The selectron solves the following optimization problem:

$$\begin{aligned} &\underset{\mathbf{w}}{\text{maximize:}} \quad \widehat{R}_n := \sum_{i=1}^{n} \nu(\mathbf{x}^{(i)}) \cdot (\mathbf{w}^\intercal \mathbf{x}^{(i)} - \vartheta) \cdot f_{\mathbf{w}}(\mathbf{x}^{(i)}) \\ &\text{subject to:} \quad \|\mathbf{w}\|_1 \leq \omega \quad \text{for some } \omega > 0. \end{aligned}$$

Balduzzi & Besserve, *NIPS* 2013

From reward maximization to plasticity rule

The reward maximization problem cannot be solved analytically in general. However, it is possible to use an iterative approach. Although $f_{\mathbf{w}}(\mathbf{x})$ is not continuous, the reward function is a continuous function of $\mathbf{w}$ and is differentiable everywhere except for the "corner" where $\mathbf{w}^{\intercal}\mathbf{x} - \vartheta = 0$. We therefore apply gradient ascent by computing the derivative of (3) with respect to synaptic weights to obtain online learning rule

$$\Delta \mathrm{w}_j = \alpha \cdot \nu(\mathbf{x}) \cdot \mathrm{x}_j \cdot f_{\mathbf{w}}(\mathbf{x}) = \begin{cases} \alpha \cdot \nu(\mathbf{x}) & \text{if } \mathrm{x}_j = 1 \text{ and } y = 1 \\ 0 & \text{else} \end{cases} \tag{4}$$

where update factor α controls the learning rate.

The learning rule is ***selective***: regardless of the neuromodulatory signal, synapse w_{jk} is updated only if there is both an input $\mathrm{x}_j = 1$ and output spike $y = f_{\mathbf{w}}(\mathbf{x}) = 1$.

Balduzzi & Besserve, *NIPS* 2013

Regularization controls metabolic cost

Define the unsupervised setting *by* $\nu(\mathbf{x}) = 1$ *for all* $\mathbf{x}$*. The reward function reduces to* $R(\mathbf{x}, f_{\mathbf{w}}) = (\mathbf{w}^\intercal\mathbf{x} - \vartheta) \cdot f_{\mathbf{w}}(\mathbf{x})$*. Without the constraint synapses will saturate. Imposing the constraint yields a more interesting solution where the selectron finds a weight vector summing to* ω *which balances (i) frequent spikes and (ii) high margins.*

Theorem 1 (Controlling the frequency of spikes)**.**
Assuming synaptic inputs are i.i.d. Bernoulli variables with $P(spike) = p$*, then*

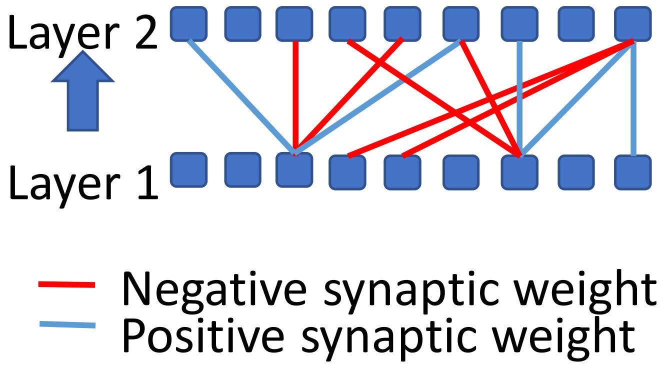
$$P\Big(f_{\mathbf{w}}(\mathbf{x}) = 1\Big) \leq p \cdot \left(\frac{\|\mathbf{w}\|_1}{\vartheta}\right)^2 \leq p \cdot \left(\frac{\omega}{\vartheta}\right)^2.$$

Learning in brains versus ANNs

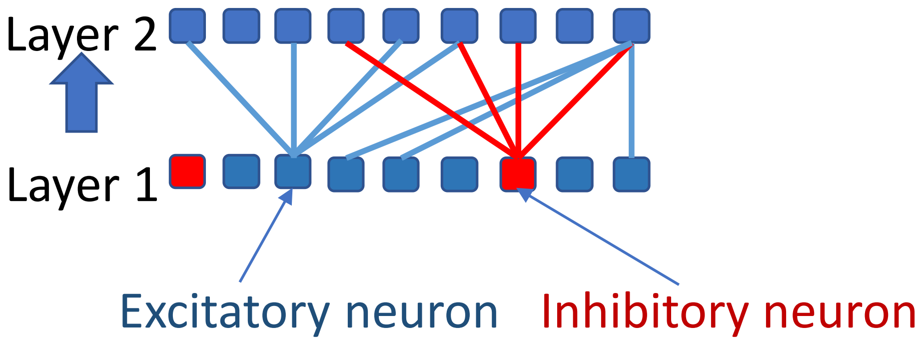
Brains	ANNs
Transmitted information is binary (spikes)	Transmitted information is continuous
Learning and testing take place concurrently (plasticity few hours after learning)	Learning and testing are separated in time
No explicitly defined tasks, domain adaptation “natural”	Well defined tasks, MTL and domain adaptation trough specialized architectures
Synaptic weight update is a function of local activity (+ neuromodulation)	Synaptic weight update is a function of the whole network connectivity (through backpropagation)

Artificial versus Biological neural circuits

Feedforward ANN



Biological network

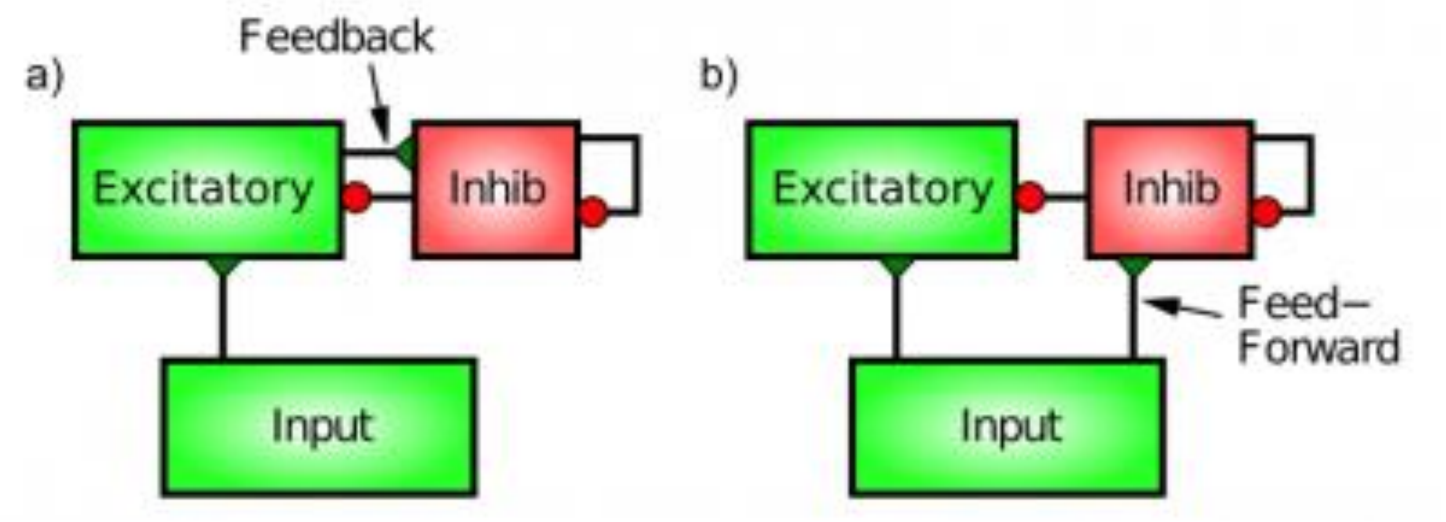


Excitatory/Pyramidal cells: long and short range axons, low firing. ~80%
Inhibitory cells: short range axons, fast firing. ~20%

Cortical microcircuits

Feedback inhibition

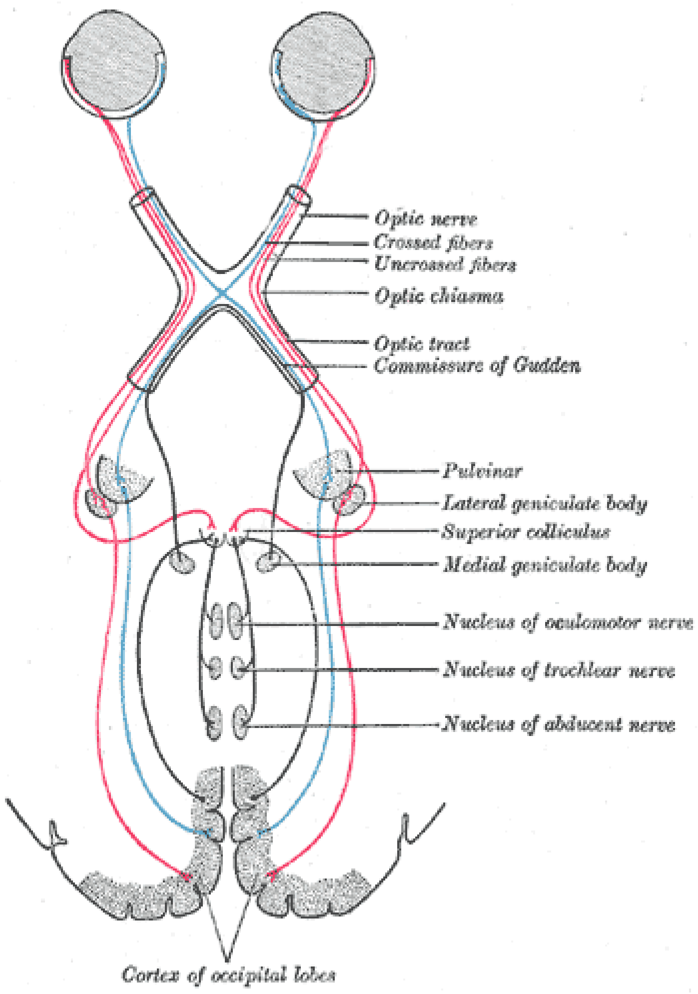
Feedforward inhibition



O'Reilly and Munakata, 2012

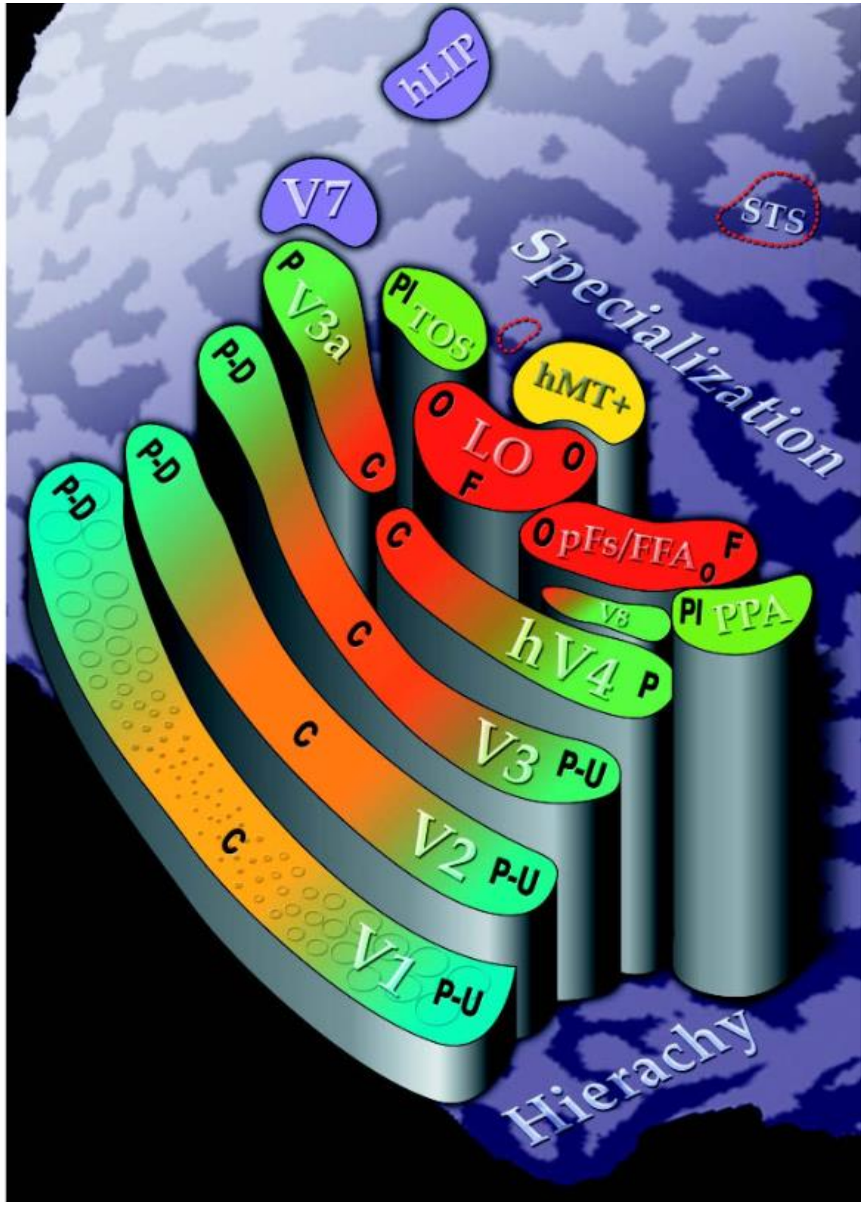
Visual perception

Visual perception



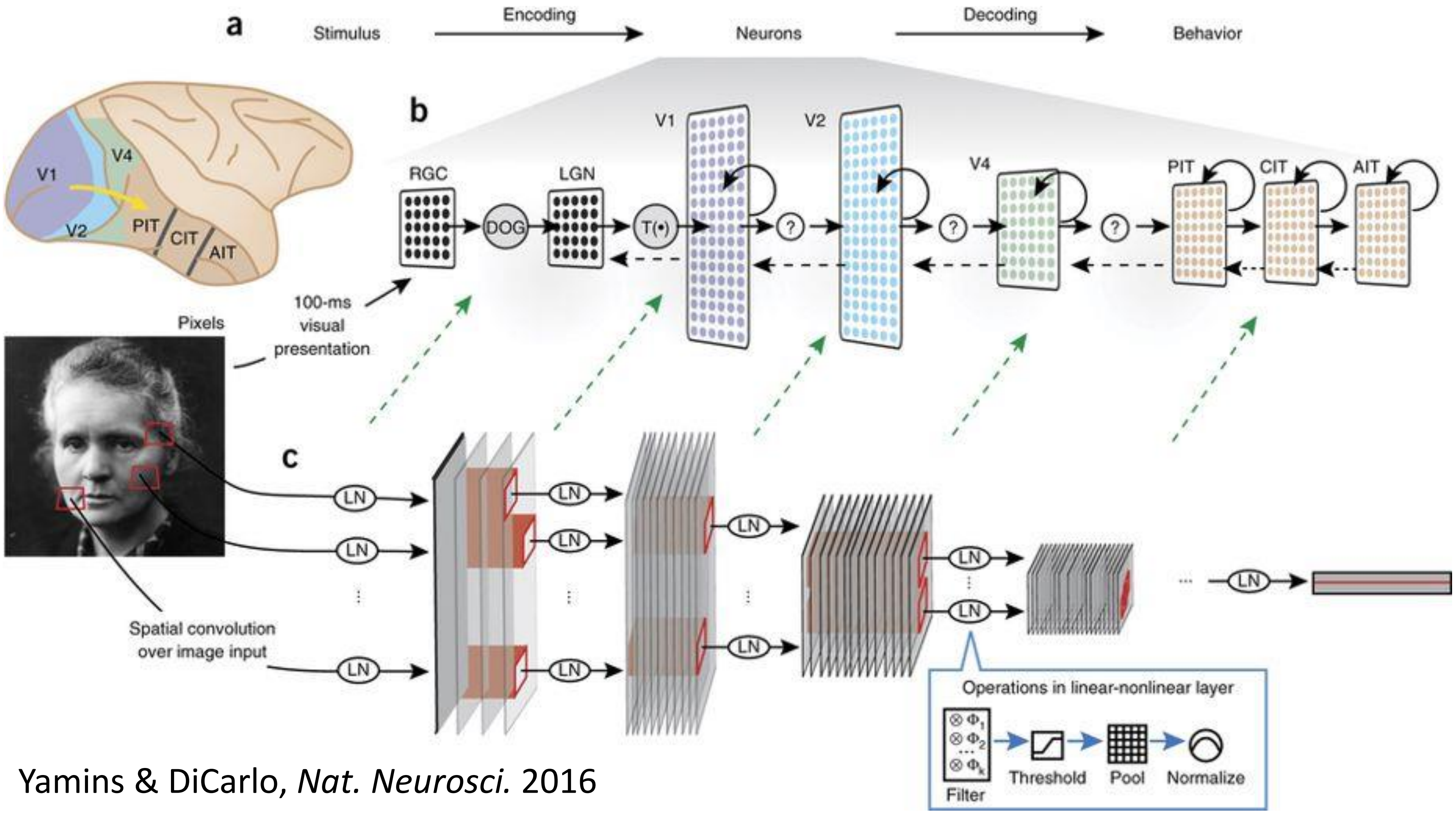
Henry Gray: Gray's Anatomy (20th edition)

C – center
P – periphery
O – objects
F – faces
Pl – places



Grill-Spector & Malach, *Annu. Rev. Neurosci.* 2004

Clear analogies with deep learning

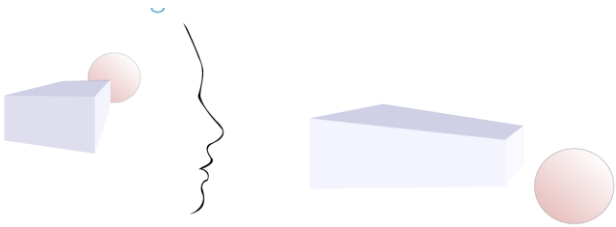


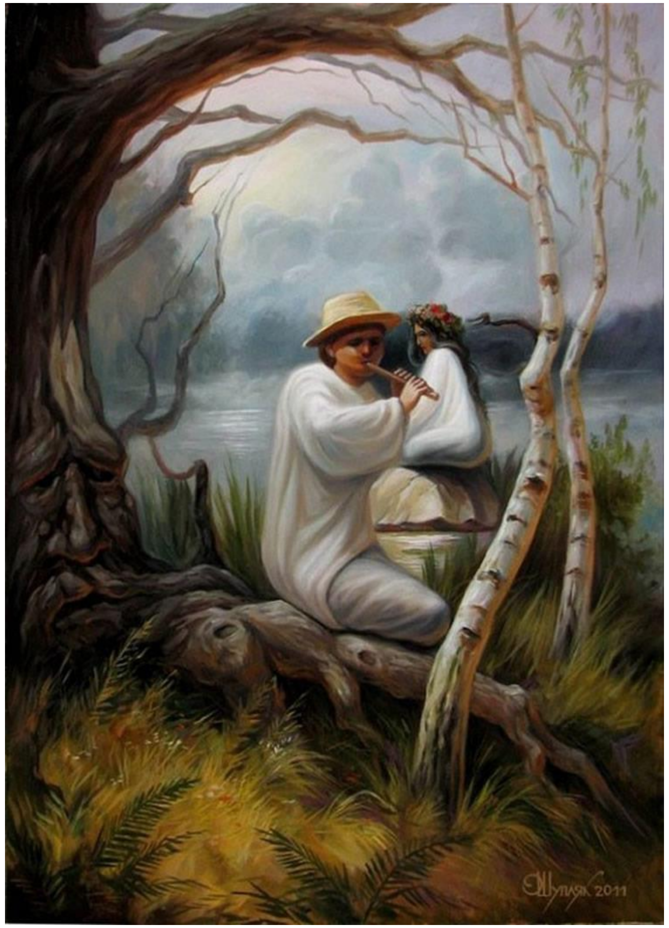
Yamins & DiCarlo, *Nat. Neurosci.* 2016

A sophisticated system… that can get confused!

German postcard (1888)
See also. My wife and my mother in law. *Wikipedia*.

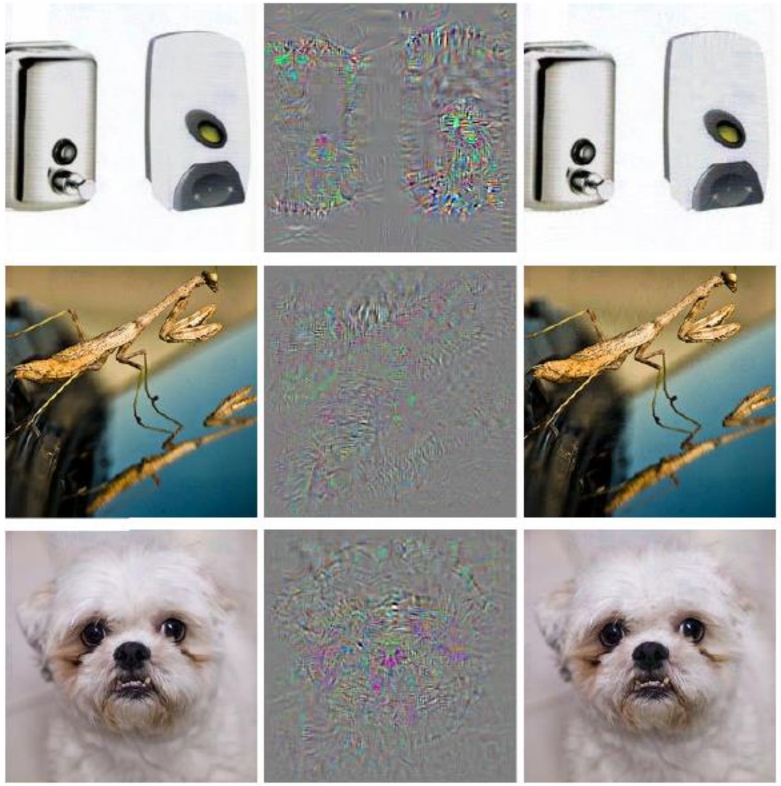
Vision solves an ill-posed problem:
2D image -> 3D scene



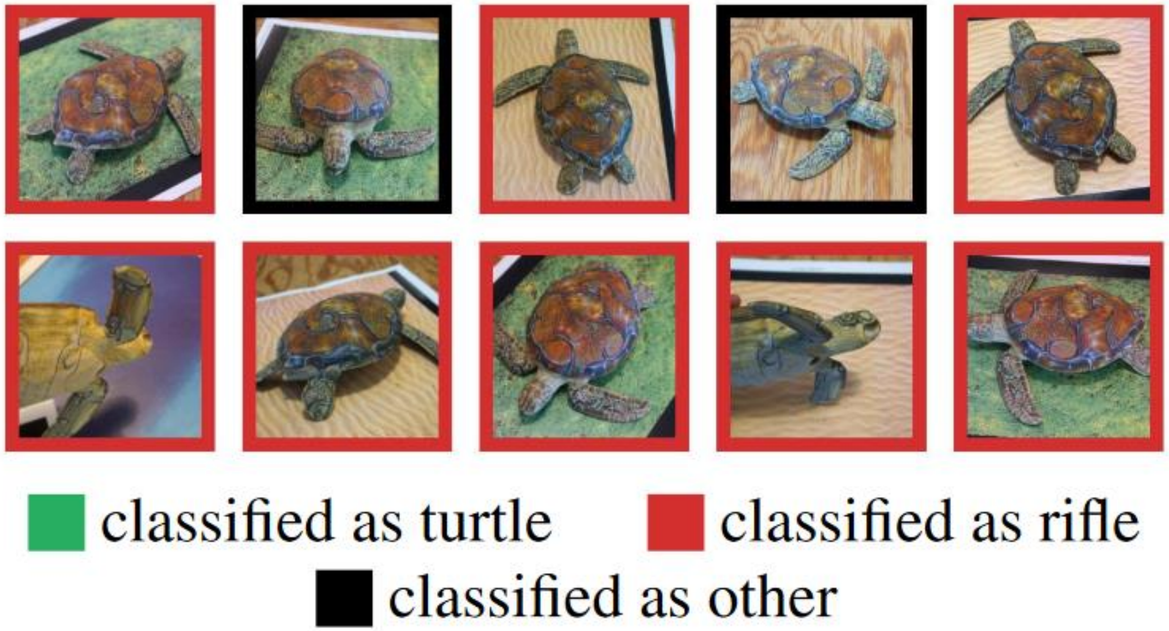


Oleg Shuplyak

ANNs can also get confused: adversarial attacks.



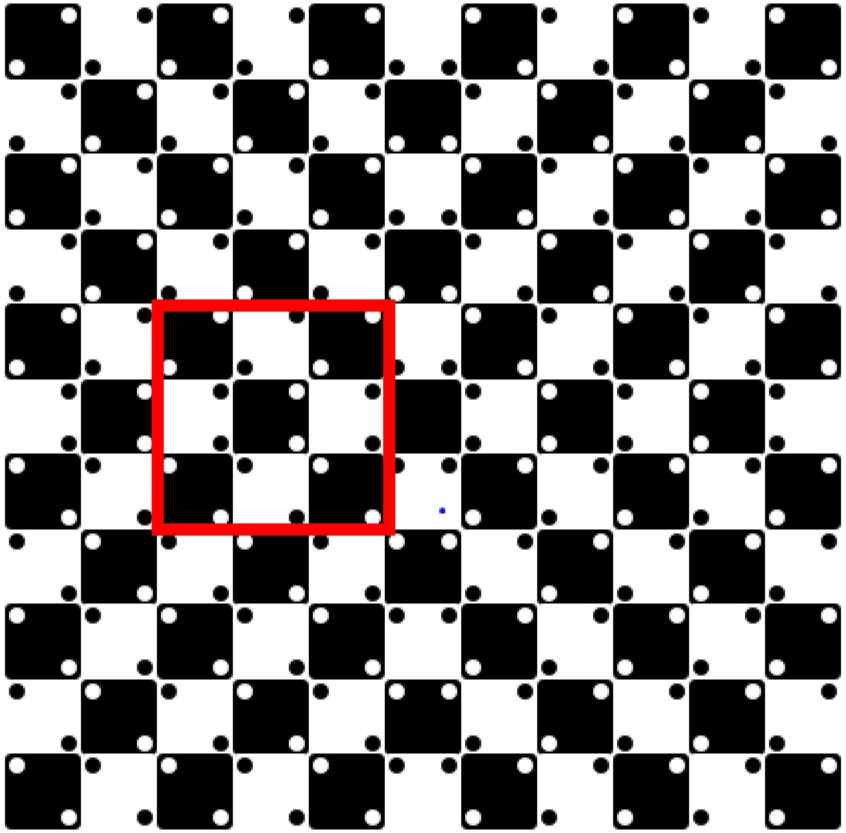
Szegedy et al., ICLR 2014

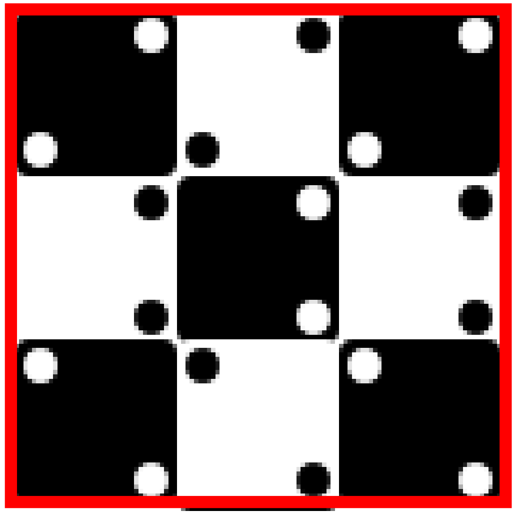


Athalye et al., *ICML* 2018

Video at:
https://www.youtube.com/watch?v=YXy6oX1iNoA

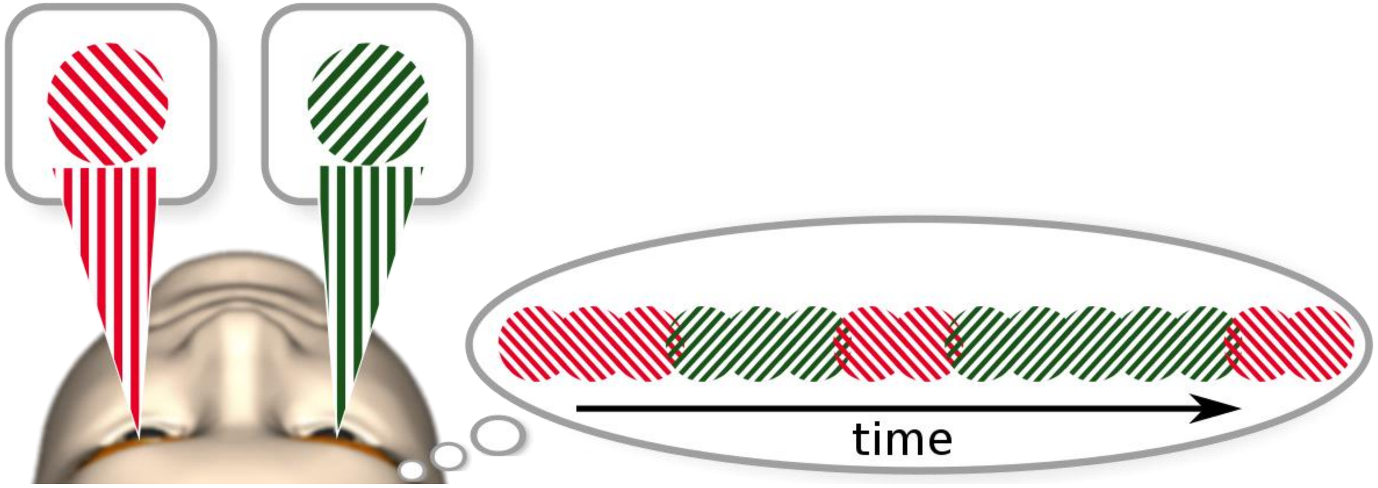
Small perturbation can also induce confusion in humans.



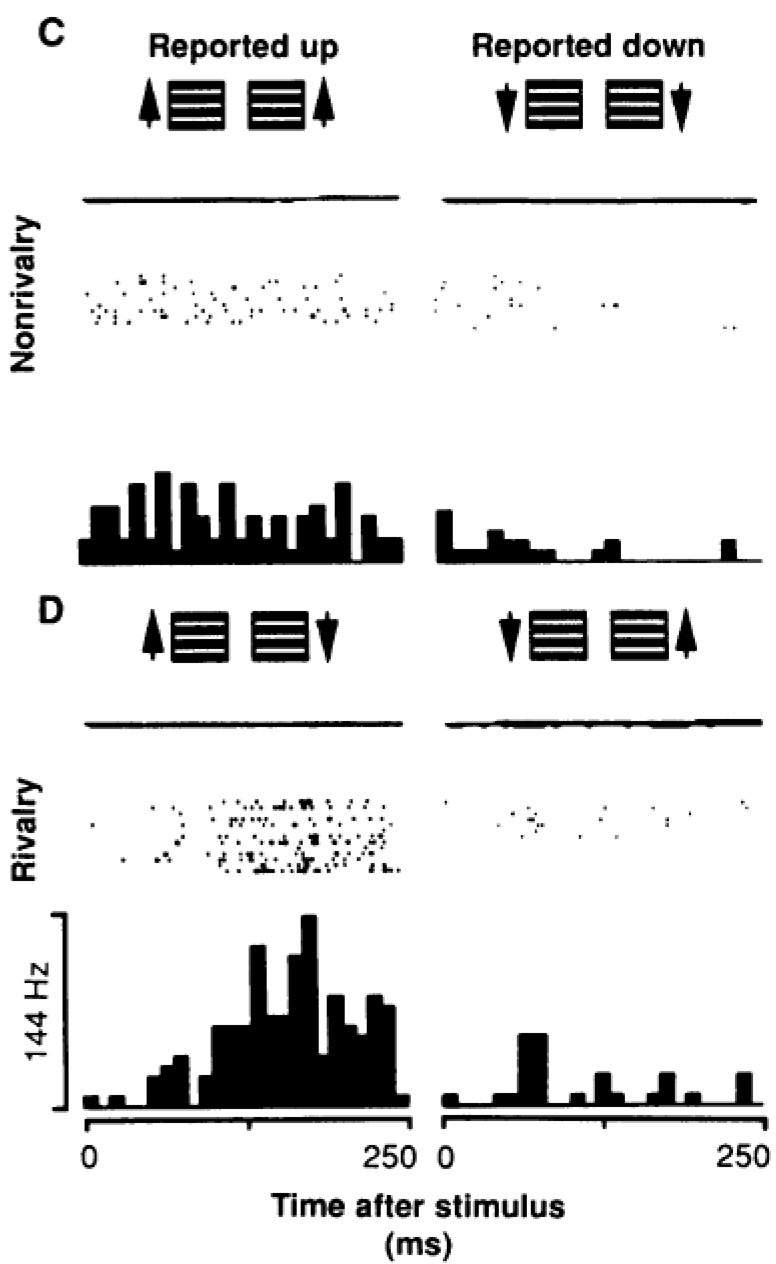


CSS Only Optical illusion by Luca Dimola

The study of perception with binocular rivalry

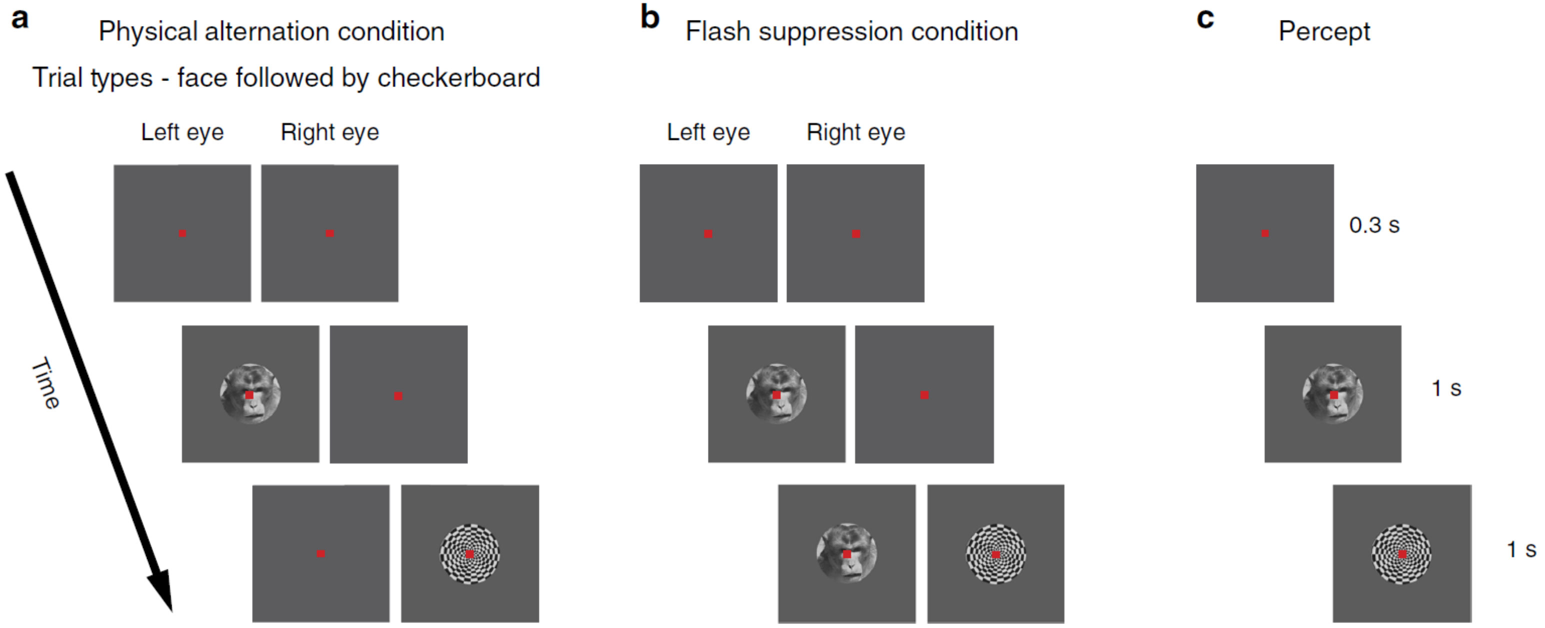


Anatomography by DBCLS licensed under CC-BY-2.1-jp

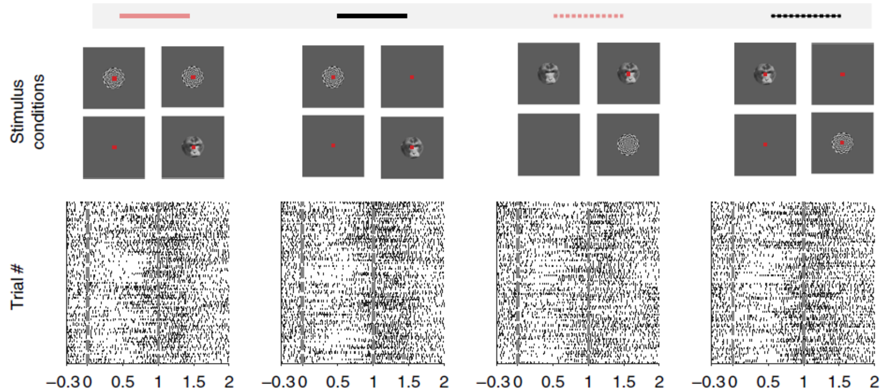


Logothetis and Shall, *Science* 1989

Application of unsupervised learning to perception data analysis



Task-free binocular perception. Awake recordings in the macaque prefrontal cortex (PFC).



Kapoor et al., 2018

Matrix factorization of averaged unit responses

- Previous result: a fraction of PFC units encode the conscious percept (Panagiotaropoulous et al., Neuron 2012)
- Question: what do other neurons do?
- Approach: non negative matrix factorization of unit responses

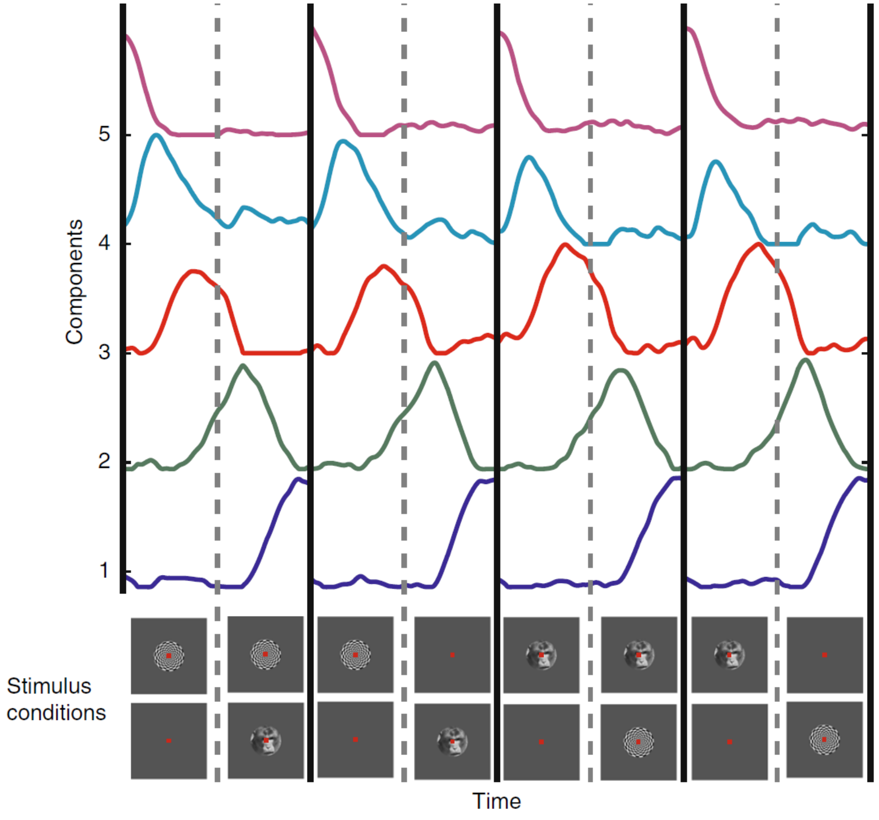
unit

$$\mathbf{R} \approx \mathbf{HW}$$

time

Outcome: most units are "monitoring" the task phase.

Implication: task-free paradigms do not isolate disentangle completely conscious perception.



Kapoor et al., 2018

References

- Athalye et al., Synthesizing Robust Adversarial Examples, *ICML* 2018.
- Dayan & Abbott, *Theoretical Neuroscience* 2001.
- Grill-Spector & Malach, The Human Visual Cortex. *Annu. Rev. Neurosci.* 2004
- Kapoor et al., Parallel and functionally segregated processing of task phase and conscious content in the prefrontal cortex. *Communications Biology* 2018.
- Logothetis and Shall, Neuronal Correlates Of Subjective Visual Perception. *Science* 1989.
- Lüscher & Malenka, The Human Visual Cortex. *Cold Spring Harb. Perspect. Biol.* 2012.
- O'Reilly & Munakata, *Computational Explorations in Cognitive Neuroscience* 2000.
- Panagiotaropoulos et al., Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron* 2012.
- Rosenblatt, The Perceptron--a perceiving and recognizing automaton. Report 85-460-1, *Cornell Aeronautical Laboratory* 1957.
- Szegedy et al., Intriguing properties of neural networks. *ICLR* 2014.