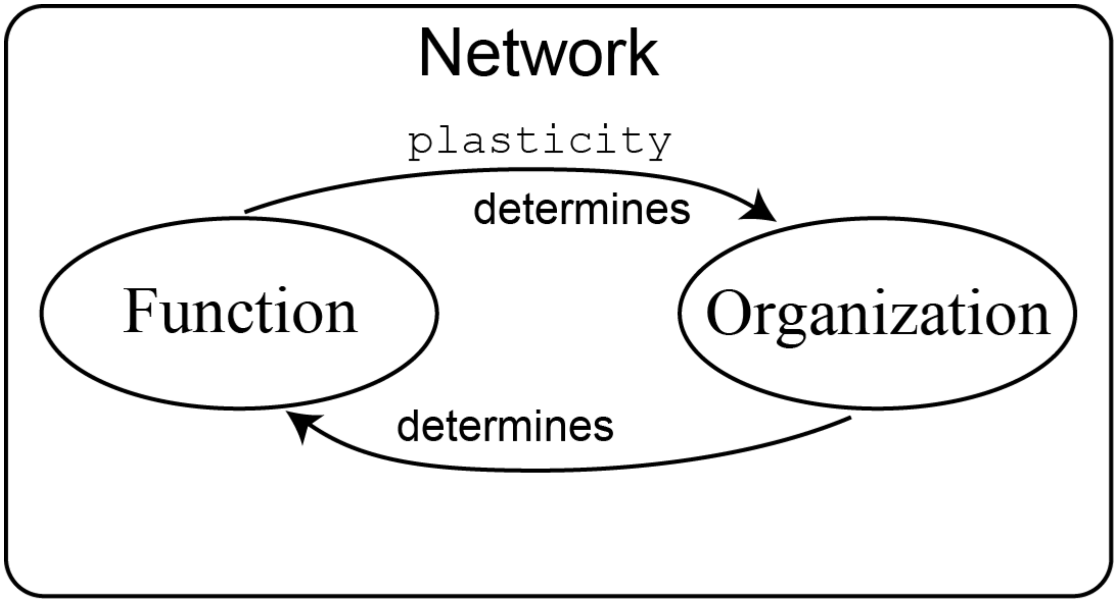
Plasticity in Neural Circuits and Neurobiology of Learning

Jenia Jitsev

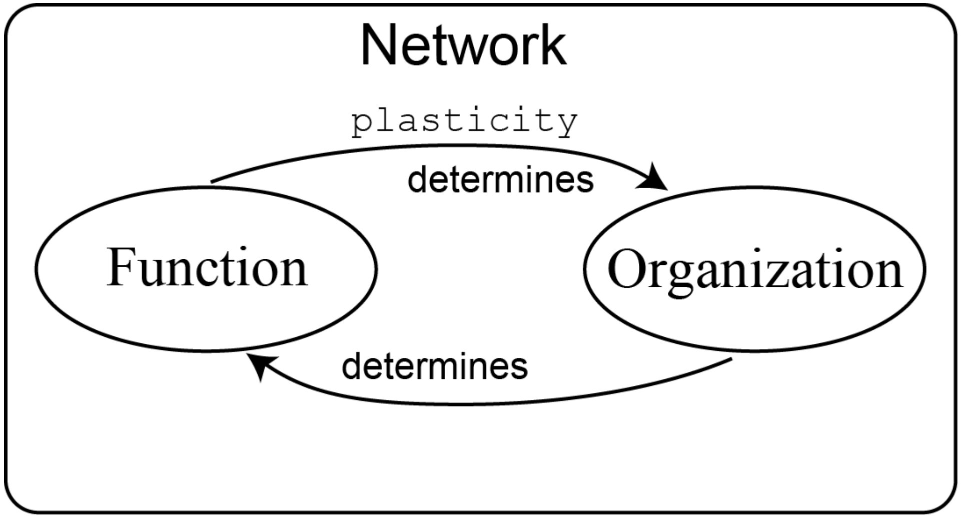
Institute of Neuroscience and Medicine, INM-6 &
Institute for Advanced Simulation, IAS-6

Functional Neural Circuits Lab
Tutors : Phillip Weidel, Susanne Kunkel

Environment, Organism and Plasticity



Environment, Organism and Plasticity



Activity changes organization (structure) changes activity …

Y(t) : activity at time *t*;

X(t) : organization / structure at time *t*;

E : environmental impact at time *t*;

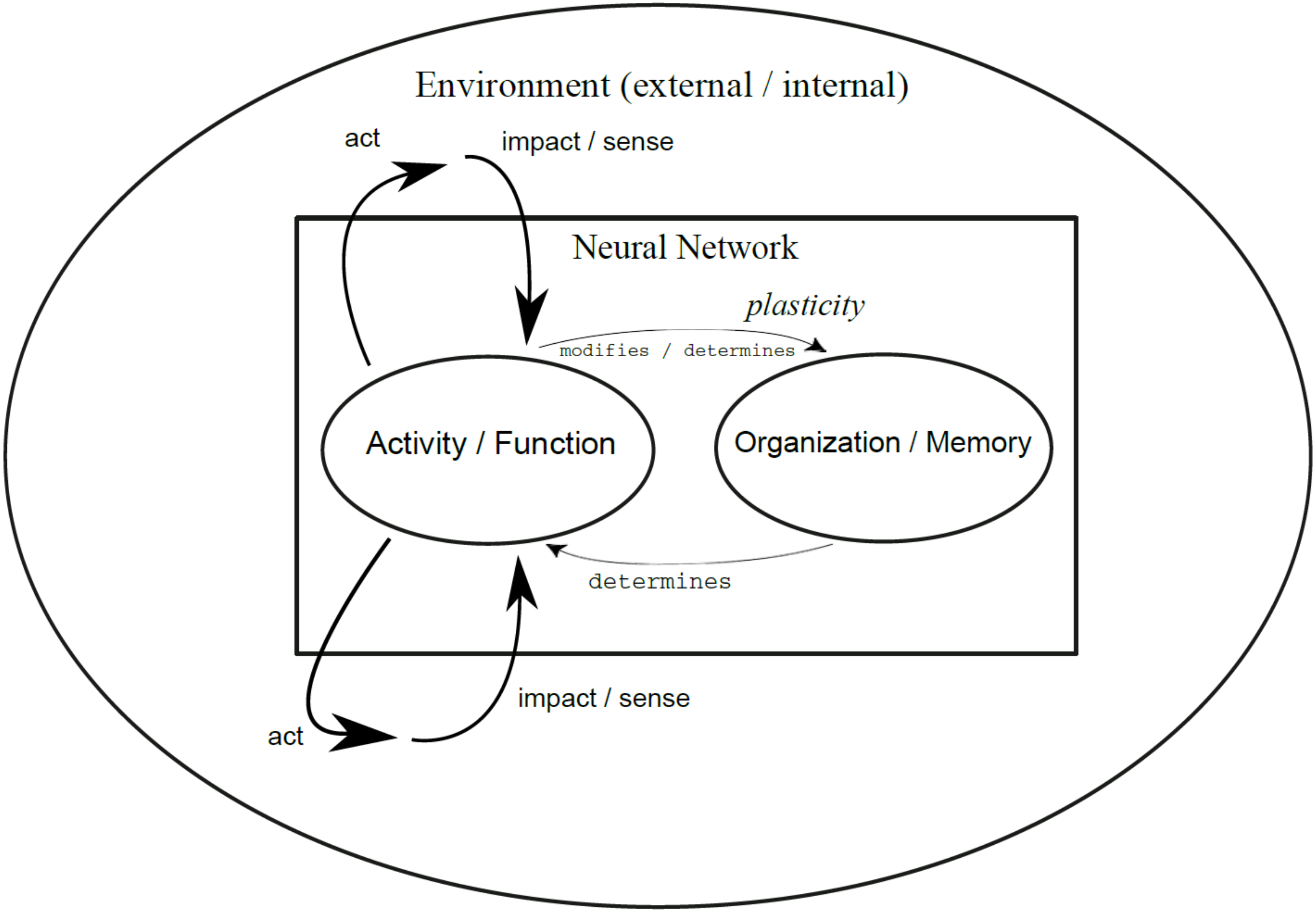
$$Y(t) = f(X(t-\Delta t), Y(t-\Delta t), E(t-\Delta t)), (\text{activity at } t)$$

$$X(t) = g(Y(t-\Delta t), X(t-\Delta t)), (\text{structure at } t)$$

$$Y(t) = f(g(Y(t-\Delta t), X(t-\Delta t)), E),$$

$$Y(t) = f(g(Y(t-\Delta t), g(Y(t-\Delta t_1), g(Y(t-\Delta t_2), ...X(t_0)...), E)$$

Environment, Organism and Plasticity



Adaptation and Learning in Biology

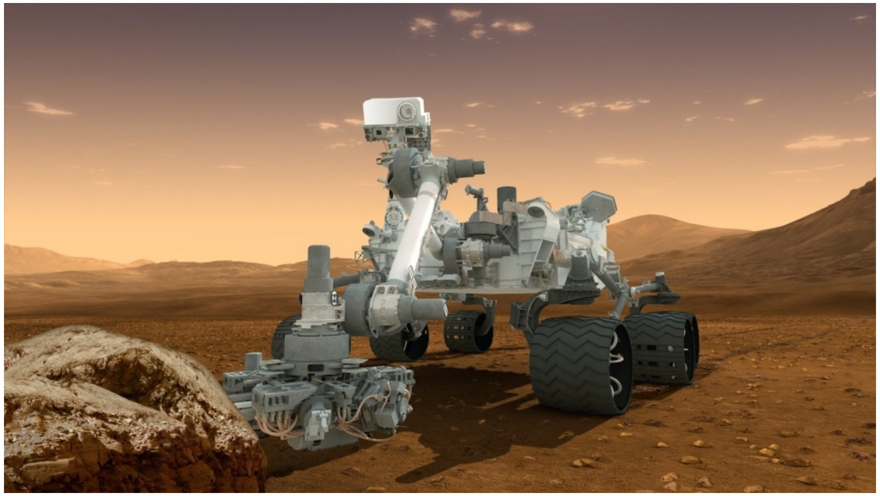
Adaptive, self-sustaining, autonomous

Assisted, engineered, non-autonomous

Adaptation and Learning in Biology

Adaptive, self-sustaining, autonomous

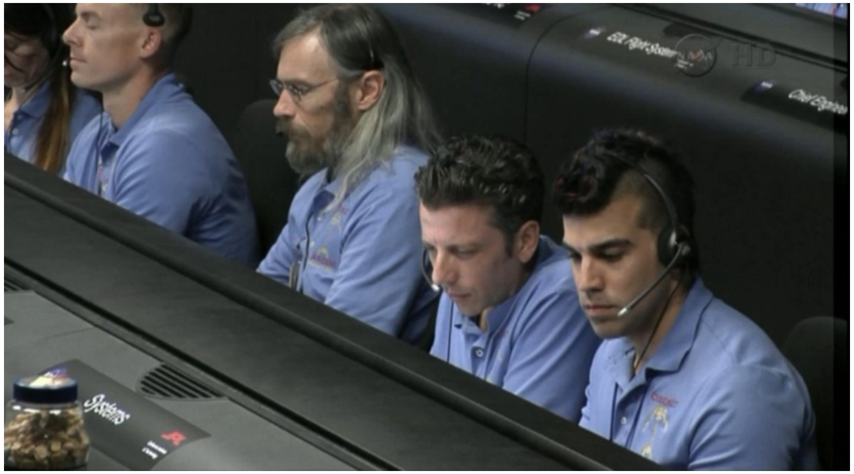
Assisted, engineered, non-autonomous



Adaptation and Learning in Biology

Adaptive, self-sustaining, autonomous

Assisted, engineered, non-autonomous



Adaptation and Learning in Biology

Maintaining long-term system's integrity and well-being

Adaptation and Learning in Biology

<u>Predict</u> the environment, <u>act</u> to gain benefit and avoid harm,
<u>Adapt</u> behavior from experience.

Loop forever and optimize.

Adaptation and Learning in Biology

Prediction, **adaptation**:

- Probabilistic inference of own internal and world‘s model from sensory experience
- Iterative reduction of model prediction error
- → Plasticity, learning, memory (slow processes)

Adaptation and Learning in Biology

<u>**Decision making and action**</u> **:**

- Neural dynamics of (near) optimal decision making (fast)
- Neural dynamics of motor execution (fast)

Adaptation and Learning in Biology

A functional neural architecture :
Prediction, decision making, action and learning

→ Closed-loop : observe, predict, act, observe, adapt, observe, predict, act, observe, adapt, …

What is the computational nature and objective of this loop?
Can we formalize it? Can we disassemble its neural substrate?

Environment, Organism and Plasticity

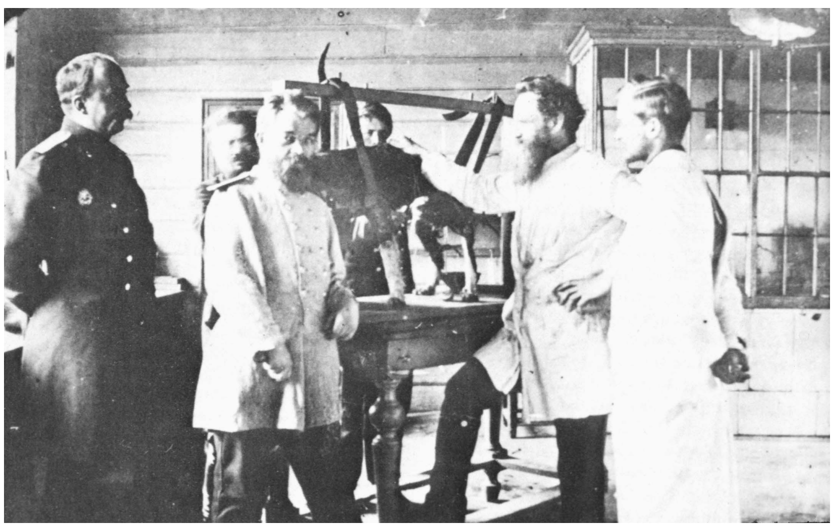
Brains "absorbs" information from the environment to improve organism's fidelity

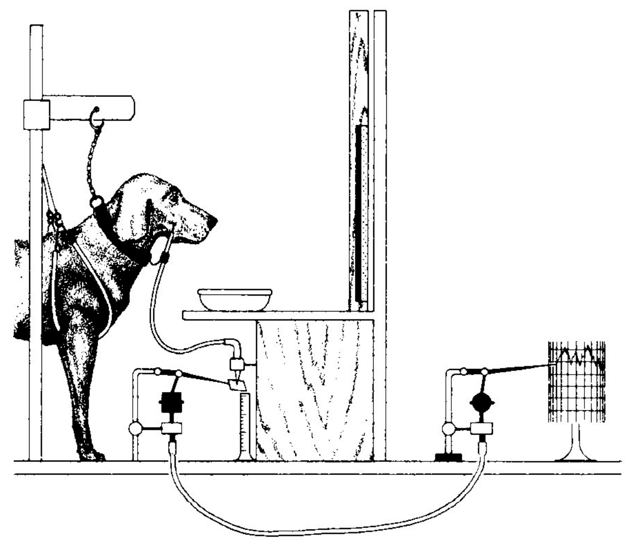
- **Objective that maintains biological persistence** : survive and produce off-spring in a changing environment
- **Prerequisites** : competence about environment
- **Learning** : extracting the information from the environment that is not yet incorporated in the nervous system
- **Plasticity** : ability of the nervous system to change its properties (structure, organization), driven by network's own activity
- **Optimization** : adapting information processing towards maximal benefit for the organism

Different Forms of Biological Learning

Basic forms of learning (classical studies in behavioral psychology) :

- **Habituation** (reduce of response to repetitive stimulus)
- **Sensitization** (increase of response to repetitive stimulus)
- **Classical Conditioning** (passive stimulus-reaction association)
- **Operant Conditioning** (active stimulus-reaction association)



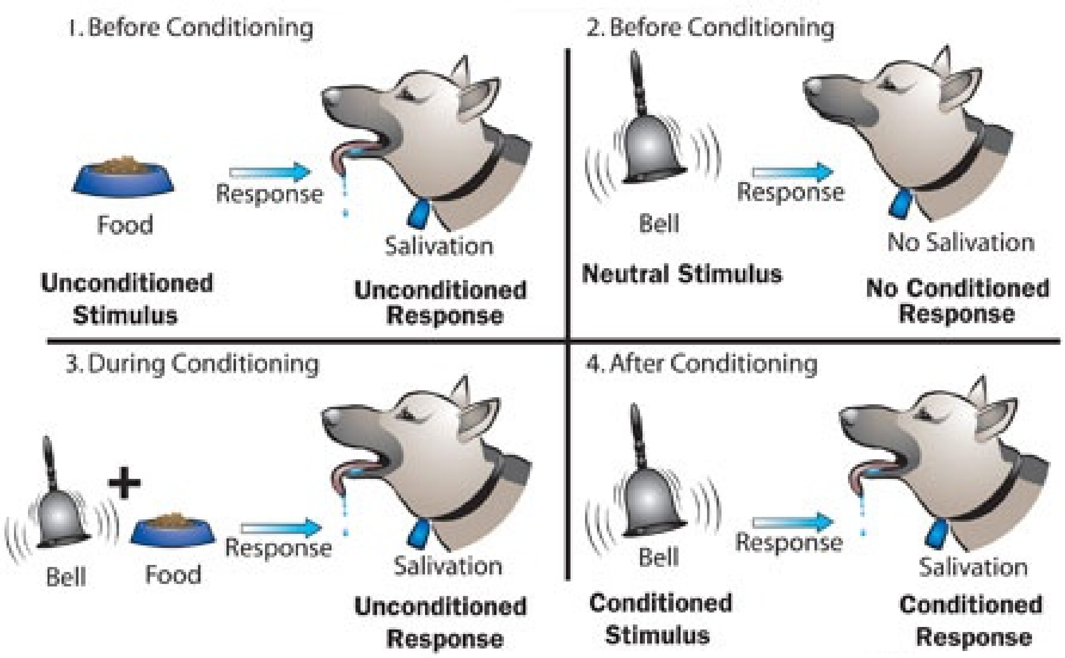


Pavlov Lab, around 1900

Different Forms of Biological Learning

Basic forms of learning (classical studies in behavioral psychology) :

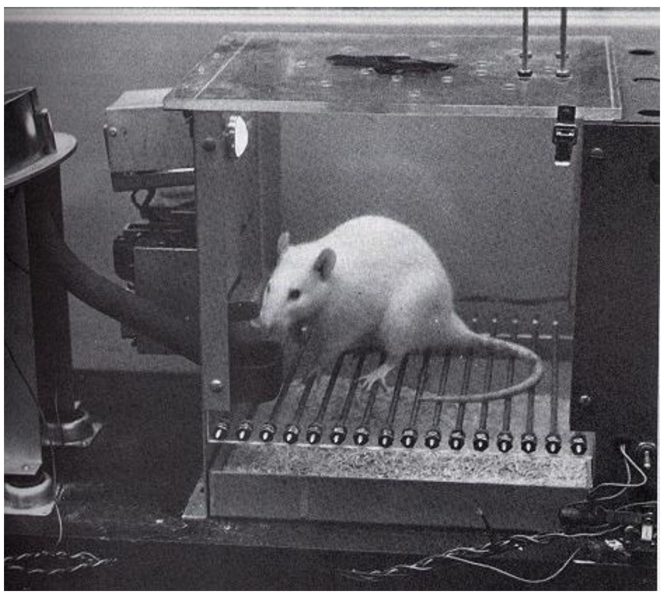
- **Habituation** (reduce of response to repetitive stimulus)
- **Sensitization** (increase of response to repetitive stimulus)
- **Classical Conditioning** (passive stimulus-reaction association)
- **Operant Conditioning** (active stimulus-reaction association)

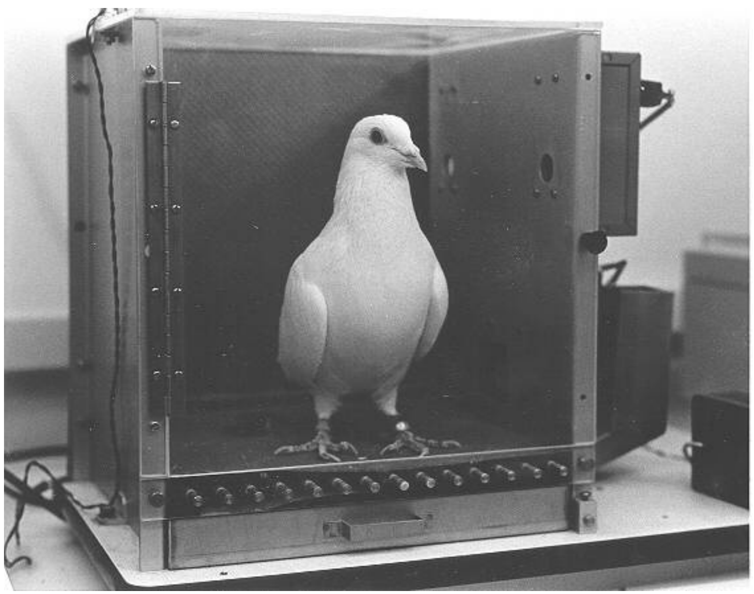


Different Forms of Biological Learning

Classical behavioral studies (Pavlov, Watson, Thorndike, Tolman, Skinner)

- Shaping complex behavior by reward and punishment
- Studying learning phenomena in rather simple setup
- Neural substrates of adaptation were not accessible

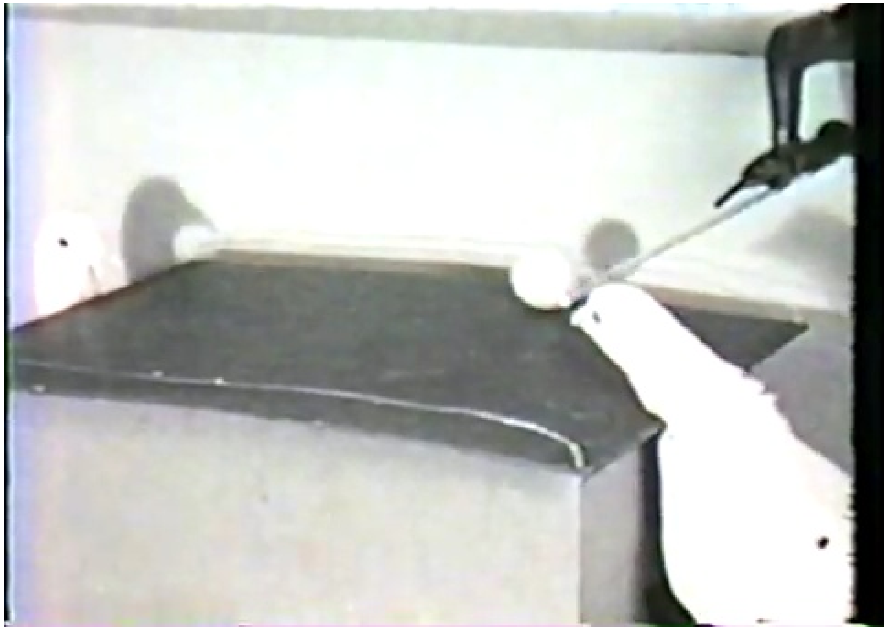




Skinner boxes, around 1940 - 1950

Different Forms of Biological Learning

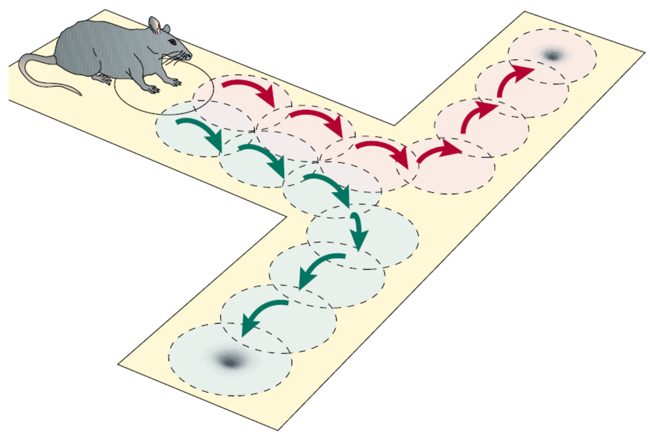
Pigeons learn to play pong by operant conditioning

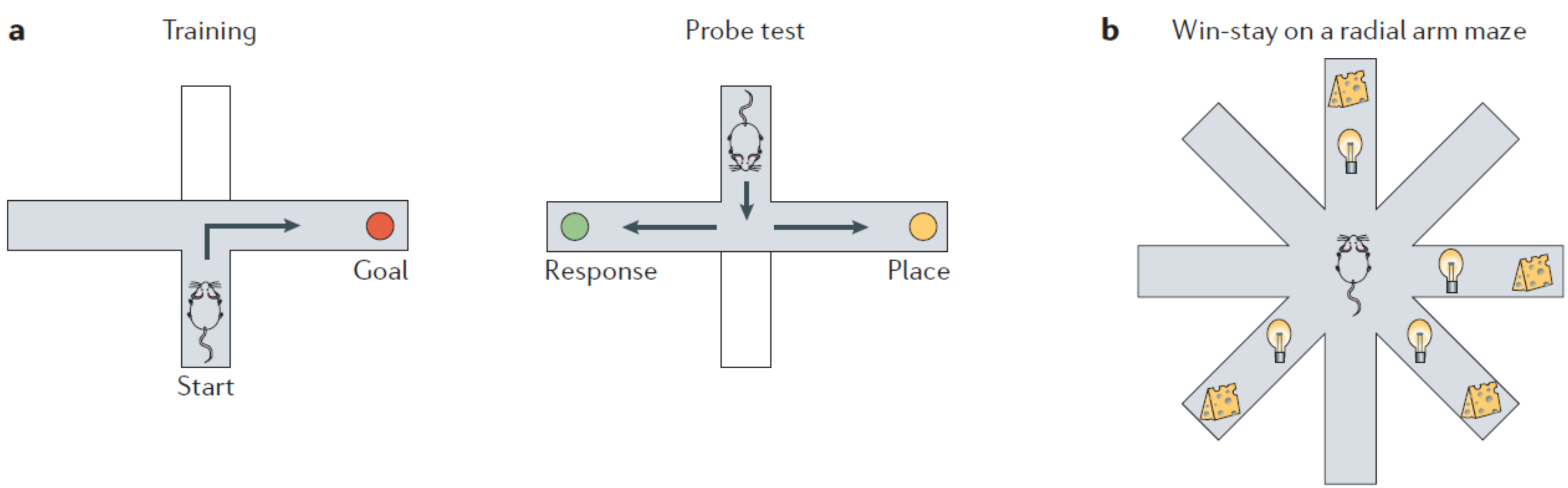


Skinner Lab, around 1950

Different Forms of Biological Learning

Established experimental learning tasks : T-Maze, radial arm maze

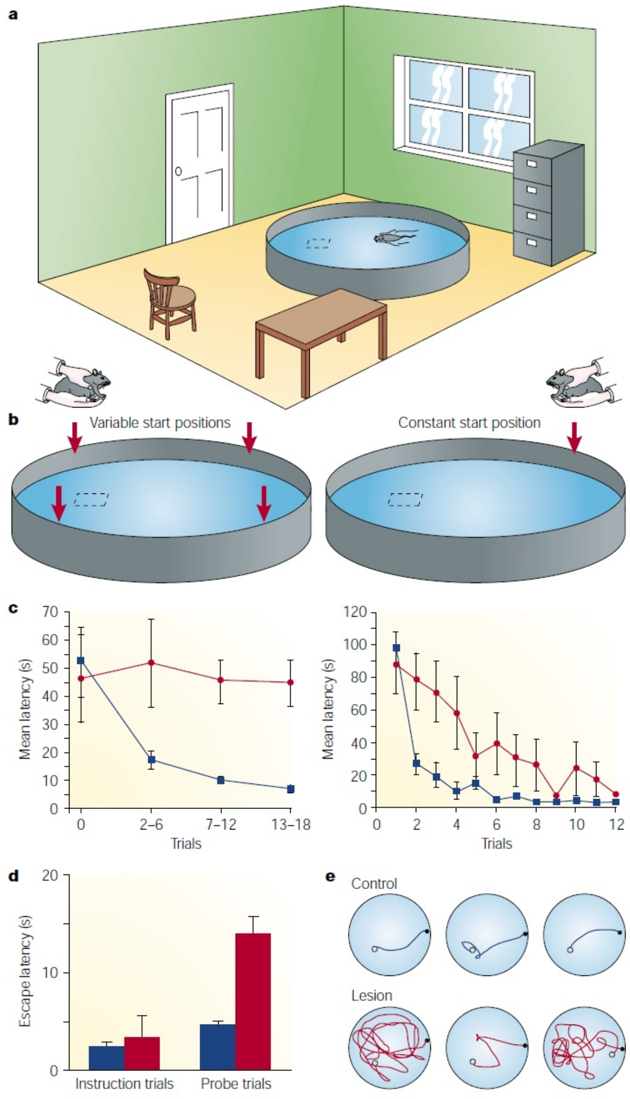




Eichenbaum, 2000; Yin & Knowlton, 2006

Different Forms of Biological Learning

Established experimental learning tasks : Morris water maze

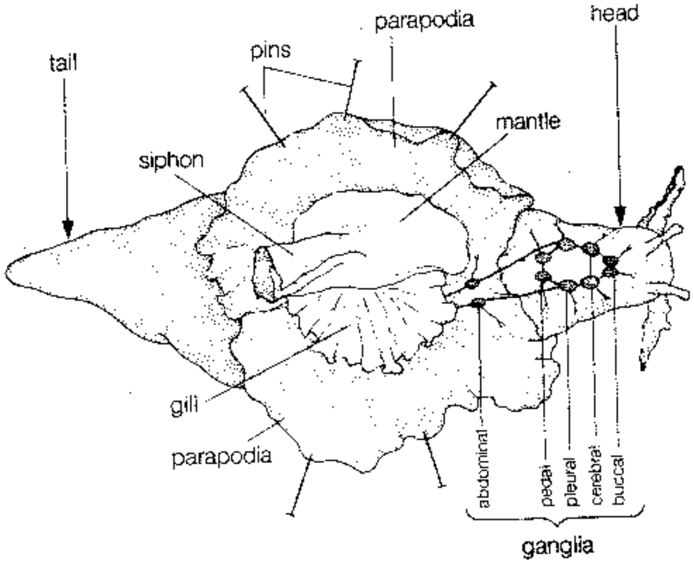


Morris et al., 1982; Eichenbaum, 2000

Neural Substrate of Learning : Plasticity

Basic forms of adaptation and learning, studied in Aplysia (sea slug)

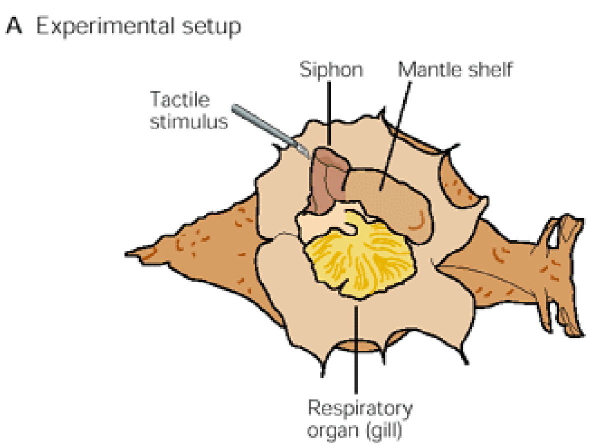
- Eric Kandel Lab (70's-00's, Nobel Prize in Medicine, 2000)
- All basic forms of learning found in more complex organisms exist also here
- Circa 20.000 Neurons
- Cellular and molecular mechanisms revealed

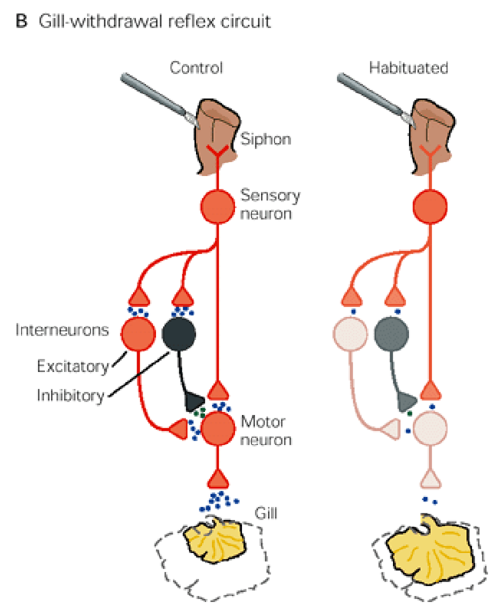


Neural Substrate of Learning : Plasticity

Basic forms of adaptation and learning, studied in Aplysia (sea slug)

- Eric Kandel Lab (70's-00's, Nobel Prize in Medicine, 2000)
- All basic forms of learning found in more complex organisms exist also here
- Circa 20.000 Neurons
- Cellular and molecular mechanisms revealed



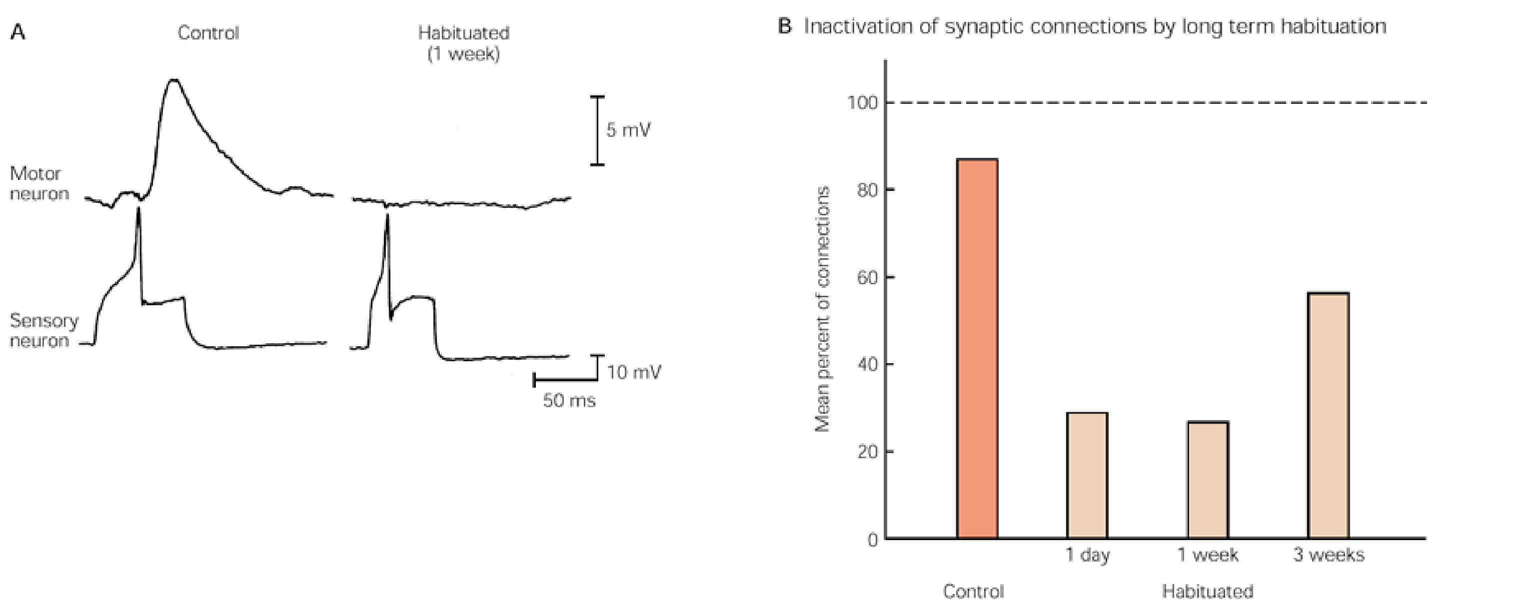


Castellucci et al., 1978

Neural Substrate of Learning : Plasticity

Basic forms of adaptation and learning, studied in Aplysia (sea slug)

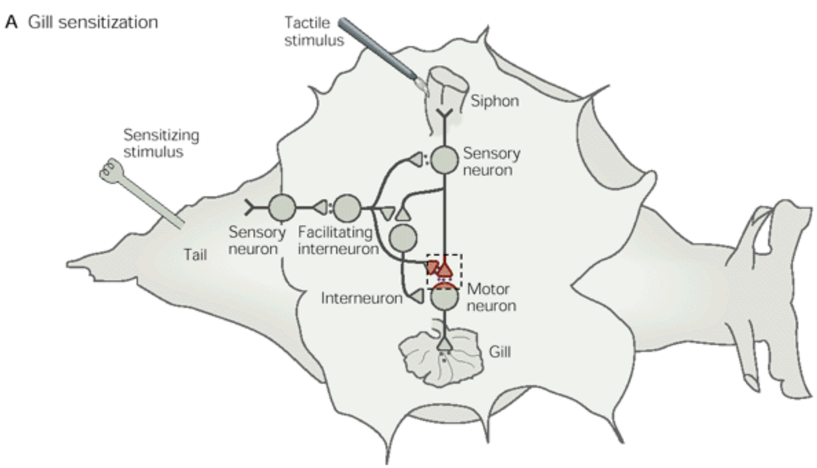
- Eric Kandel Lab (70's-00's, Nobel Prize in Medicine, 2000)
- All basic forms of learning found in more complex organisms exist also here
- Circa 20.000 Neurons
- Cellular and molecular mechanisms revealed

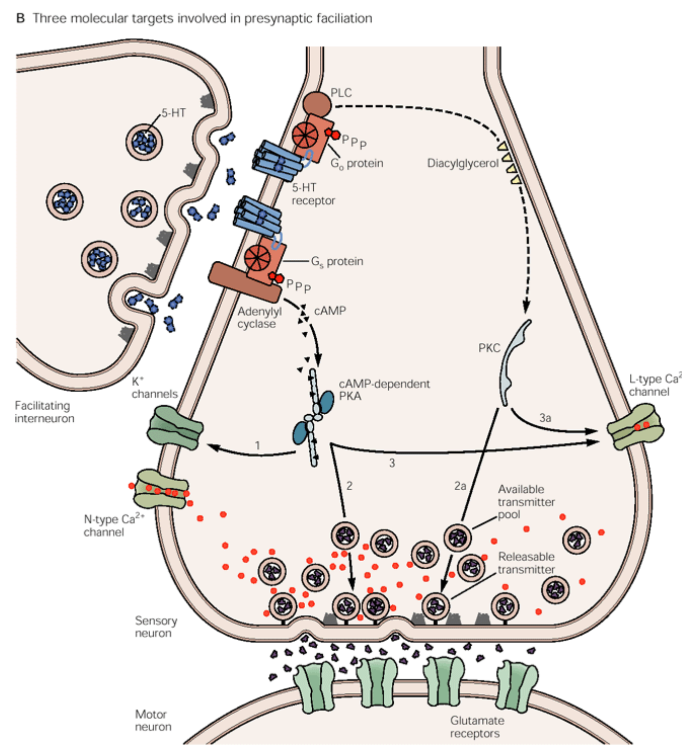


Neural Substrate of Learning : Plasticity

Basic forms of adaptation and learning, studied in Aplysia (sea slug)

- Eric Kandel Lab (70's-00's, Nobel Prize in Medicine, 2000)
- All basic forms of learning found in more complex organisms exist also here
- Circa 20.000 Neurons
- Cellular and molecular mechanisms revealed

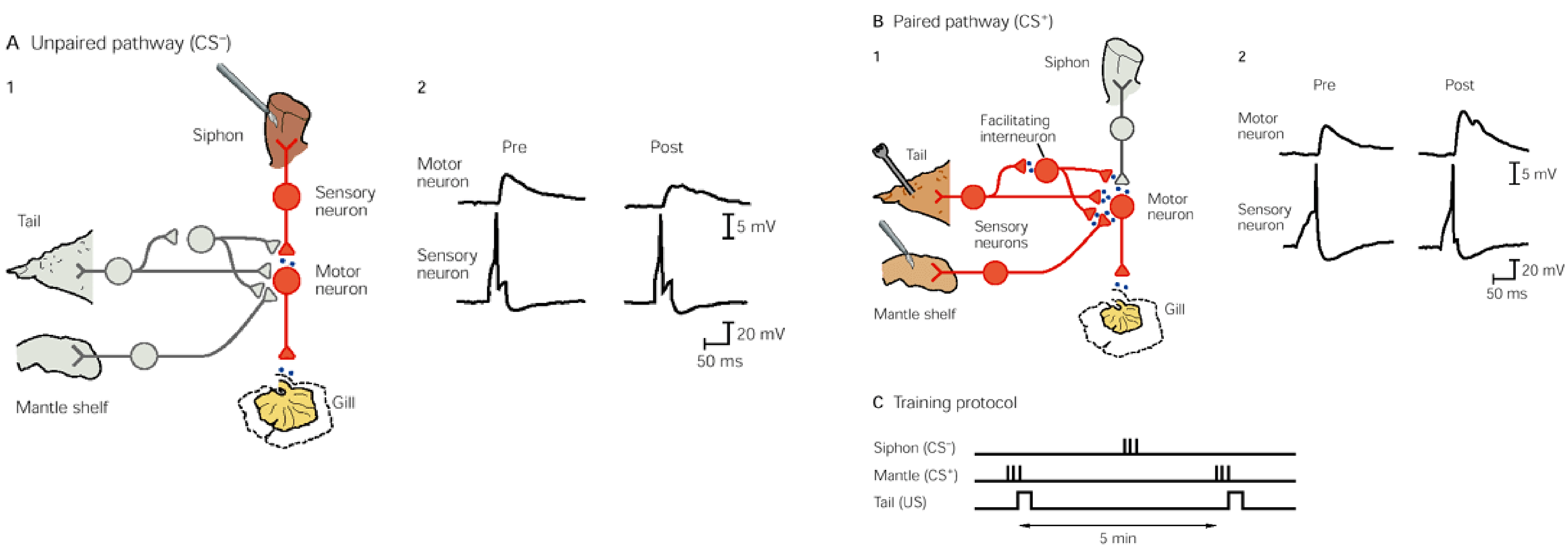




Neural Substrate of Learning : Plasticity

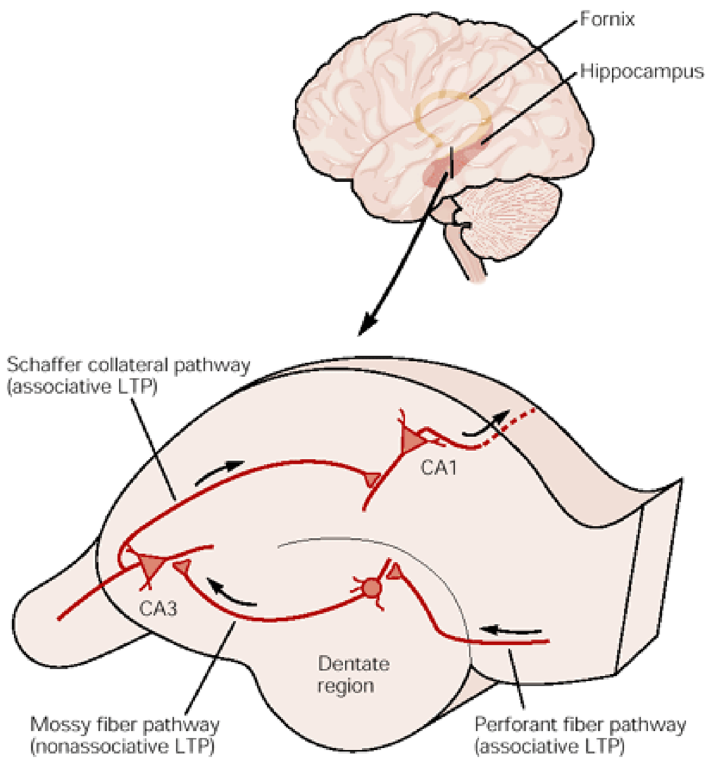
Basic forms of adaptation and learning, studied in Aplysia (sea slug)

- Eric Kandel Lab (70's-00's, Nobel Prize in Medicine, 2000)
- All basic forms of learning found in more complex organisms exist also here
- Circa 20.000 Neurons
- Cellular and molecular mechanisms revealed



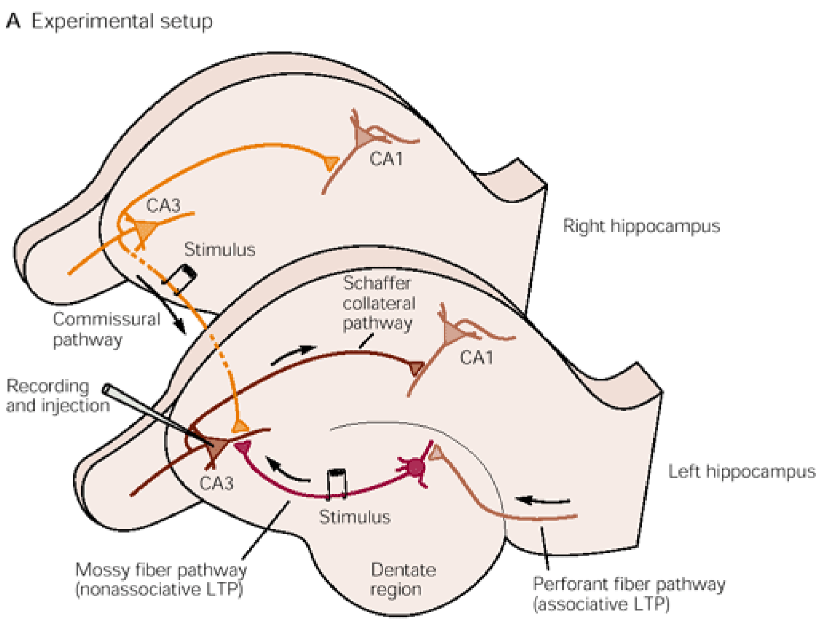
Neural Substrate of Learning : Plasticity

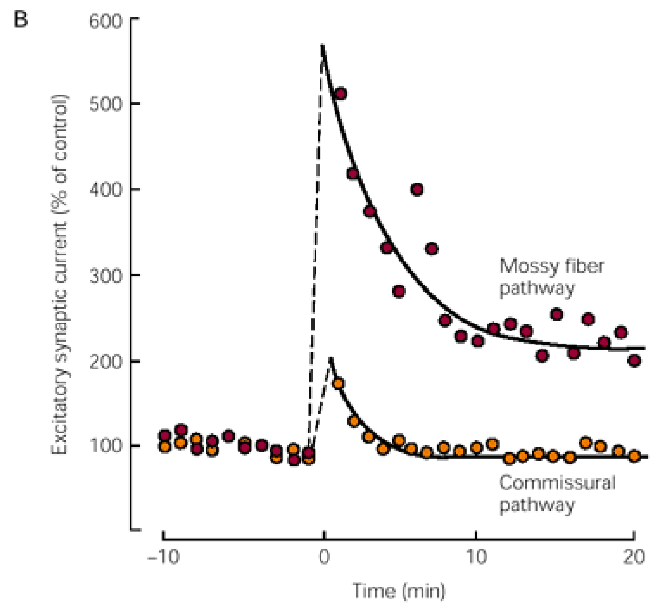
Plasticity has universal cellular and molecular mechanisms throughout the species



Neural Substrate of Learning : Plasticity

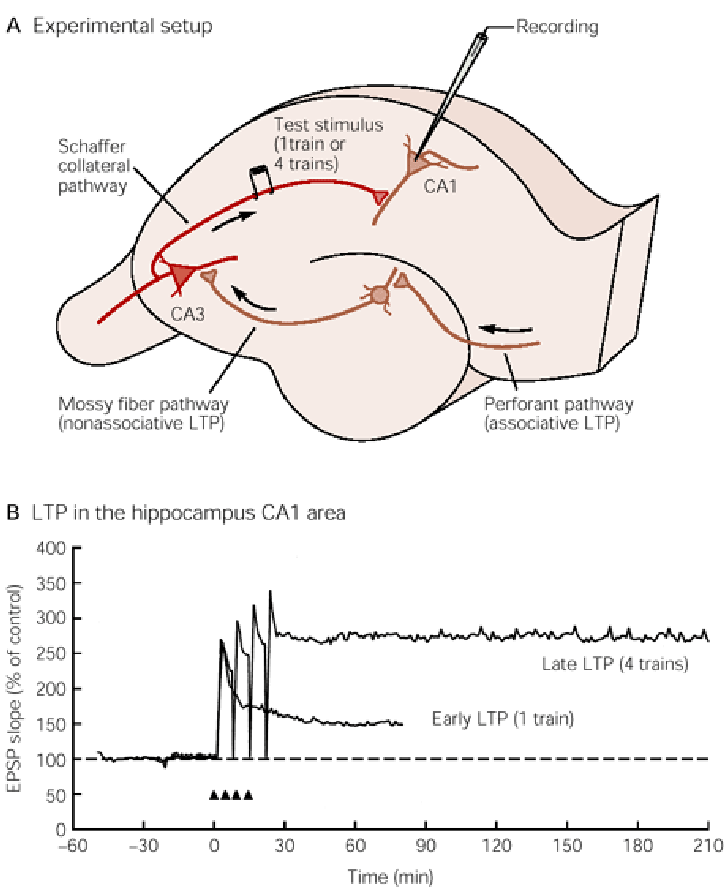
Plasticity has universal cellular and molecular mechanisms throughout the species (Induction of long-term potentiation, LTP)

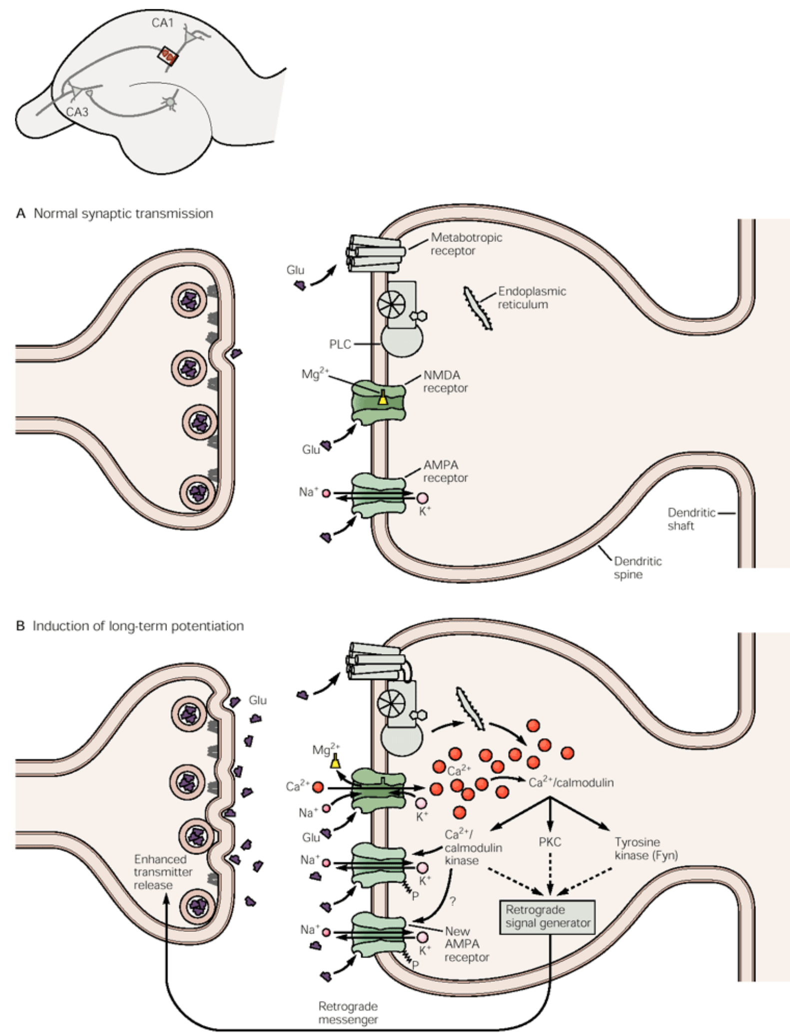




Neural Substrate of Learning : Plasticity

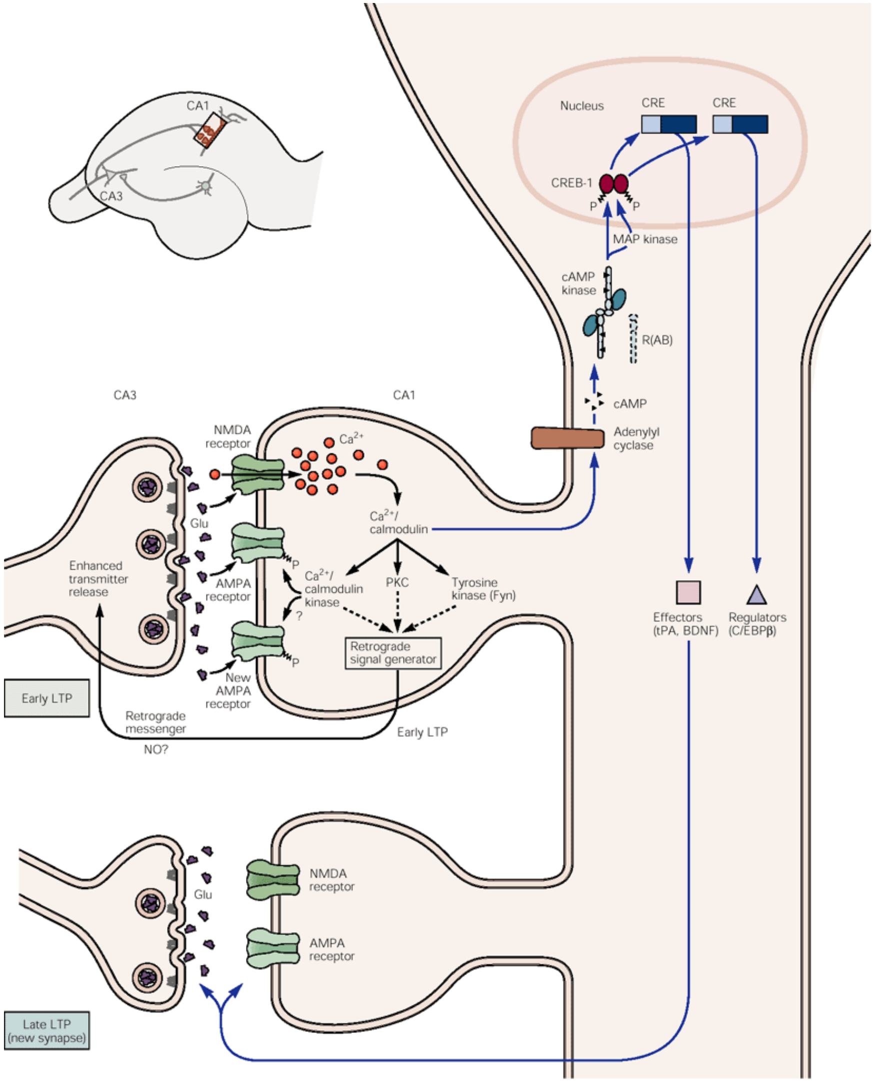
Plasticity has universal cellular and molecular mechanisms throughout the species





Neural Substrate of Learning : Plasticity

Plasticity has universal cellular and molecular mechanisms throughout the species



Neural Substrate of Learning : Plasticity

Plasticity has universal cellular and molecular mechanisms throughout the brain regions and species

ARTICLES

nature neuroscience

Visualization of NMDA receptor–dependent AMPA receptor synaptic plasticity *in vivo*

Yong Zhang[1,3], Robert H Cudmore[1,3], Da-Ting Lin[1,2], David J Linden[1] & Richard L Huganir[1]

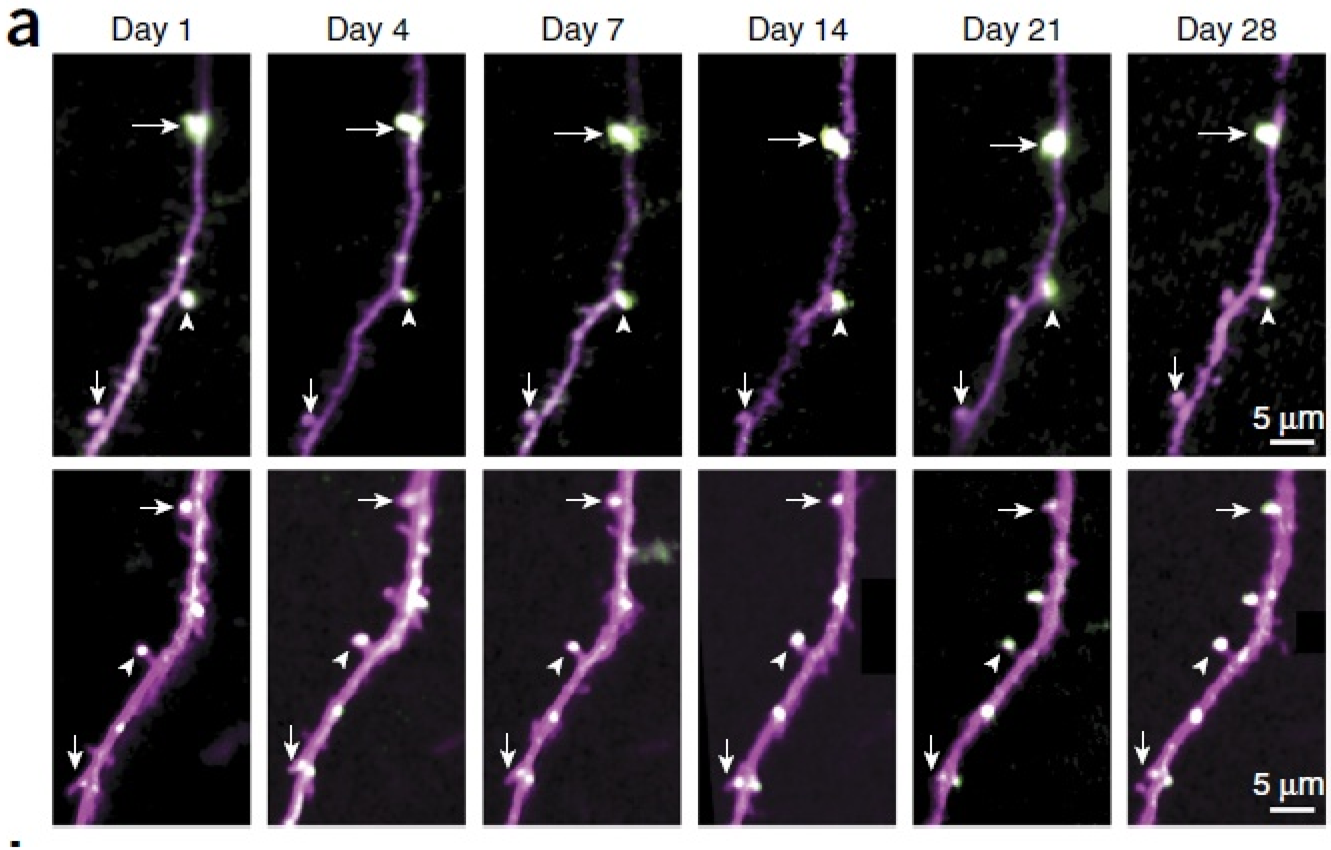
Regulation of AMPA receptor (AMPAR) membrane trafficking is critical for synaptic plasticity, as well as for learning and memory. However, the mechanisms of AMPAR trafficking *in vivo* remain elusive. Using *in vivo* two-photon microscopy in the mouse somatosensory barrel cortex, we found that acute whisker stimulation led to a significant increase in the intensity of surface AMPAR GluA1 subunit (sGluA1) in both spines and dendritic shafts and a small increase in spine size relative to prestimulation values. Interestingly, the initial spine properties biased spine changes following whisker stimulation. Changes in spine sGluA1 intensity were positively correlated with changes in spine size and dendritic shaft sGluA1 intensity following whisker stimulation. The increase in spine sGluA1 intensity evoked by whisker stimulation was NMDA receptor dependent and long lasting, similar to major forms of synaptic plasticity in the brain. In this study we were able to observe experience-dependent AMPAR trafficking in real time and characterize, *in vivo*, a major form of synaptic plasticity in the brain.

[1]Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [2]Present address: National Institute on Drug Abuse, Baltimore, Maryland, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to R.L.H. (rhuganir@jhmi.edu).

Received 3 October 2014; accepted 7 January 2015; published online 2 February 2015; doi:10.1038/nn.3936

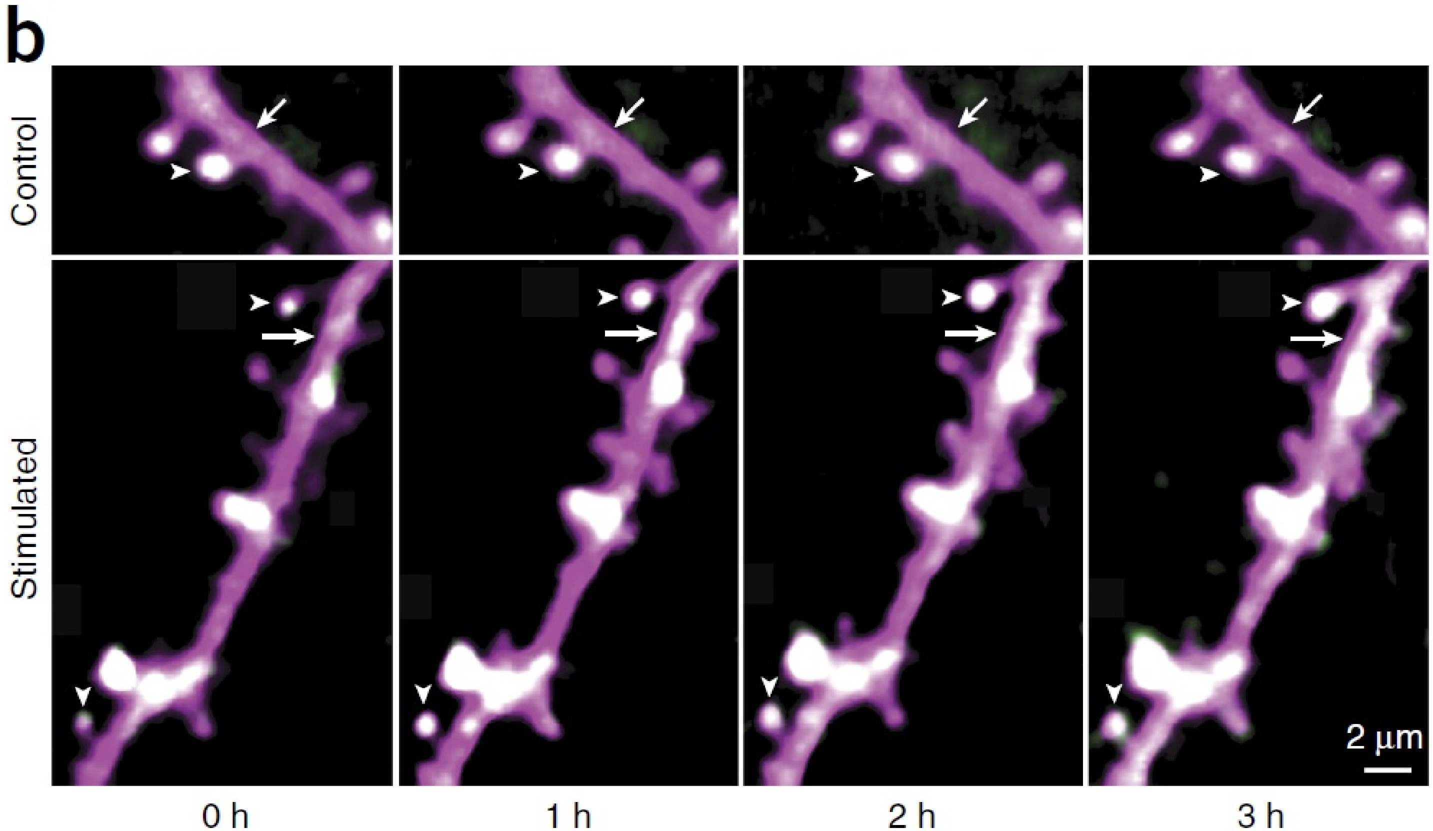
Neural Substrate of Learning : Plasticity

Measuring and visualizing AMPAR expression intensity *in vivo*
(gene transfer into the embryonic mouse brain using in vivo electroporation, GFP labeled AMPAR subunits, two photon microscopy)



Neural Substrate of Learning : Plasticity

AMPAR intensity increases in specific synaptic location after stimulation; this change depends on NMDAR function



Neural Substrate of Learning : Plasticity

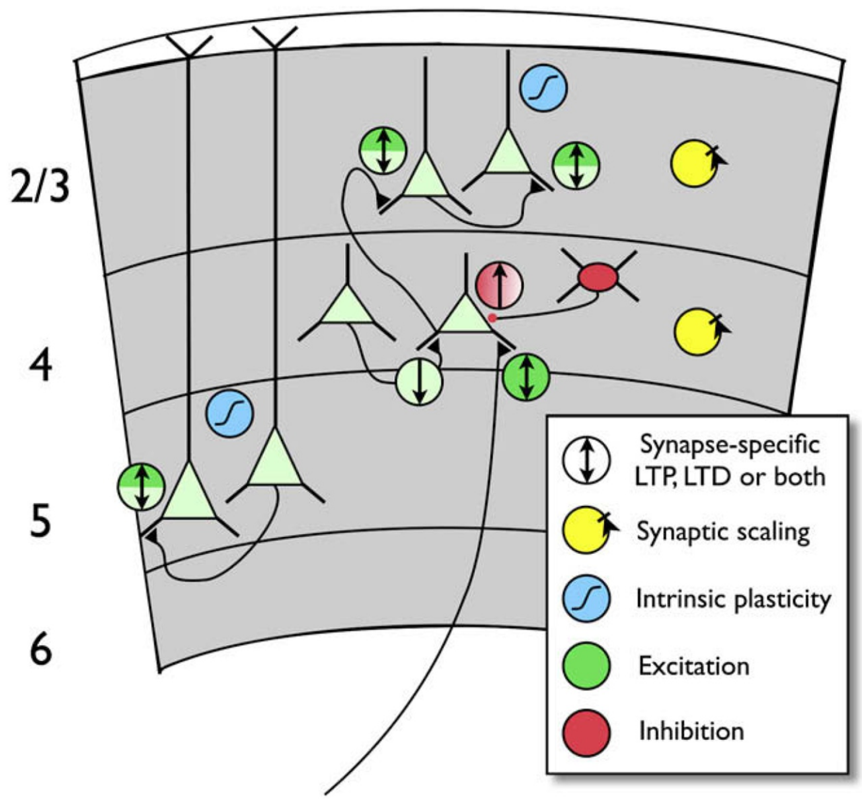
Plasticity can affect different sites in neuronal networks

- Synaptic plasticity (synaptic LTP / LTD, STP / STD)
- Dendritic plasticity (excitability changes, spine modification)
- Somatic plasticity (excitability changes, receptor changes)
- Structural plasticity (adding new synapses, spines or removing them)
- Neurogenesis (adding or removing cells or fibres)

Neural Substrate of Learning : Plasticity

Plasticity can affect different sites in neuronal networks

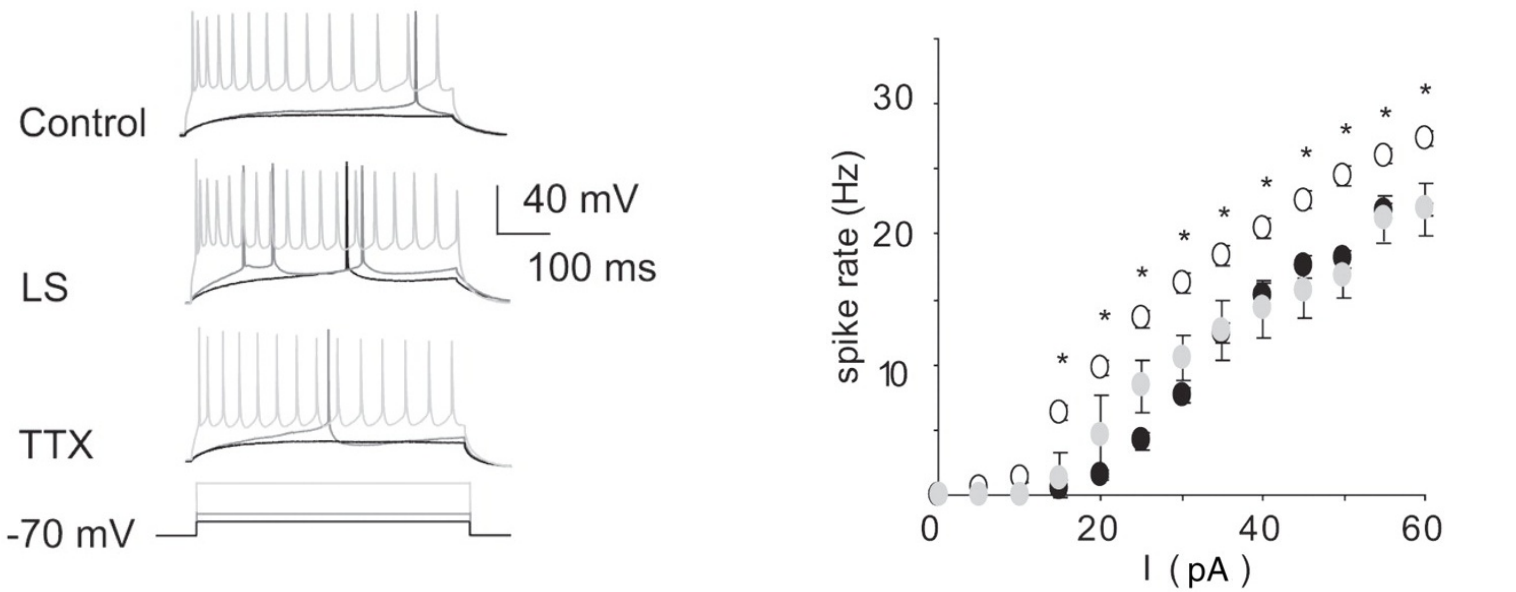
- Synaptic plasticity (synaptic LTP / LTD, STP / STD)
- Dendritic plasticity (excitability changes, spine modification)
- Somatic plasticity (excitability changes, receptor changes)



Maffei, Turrigiano, 2008

Neural Substrate of Learning : Plasticity

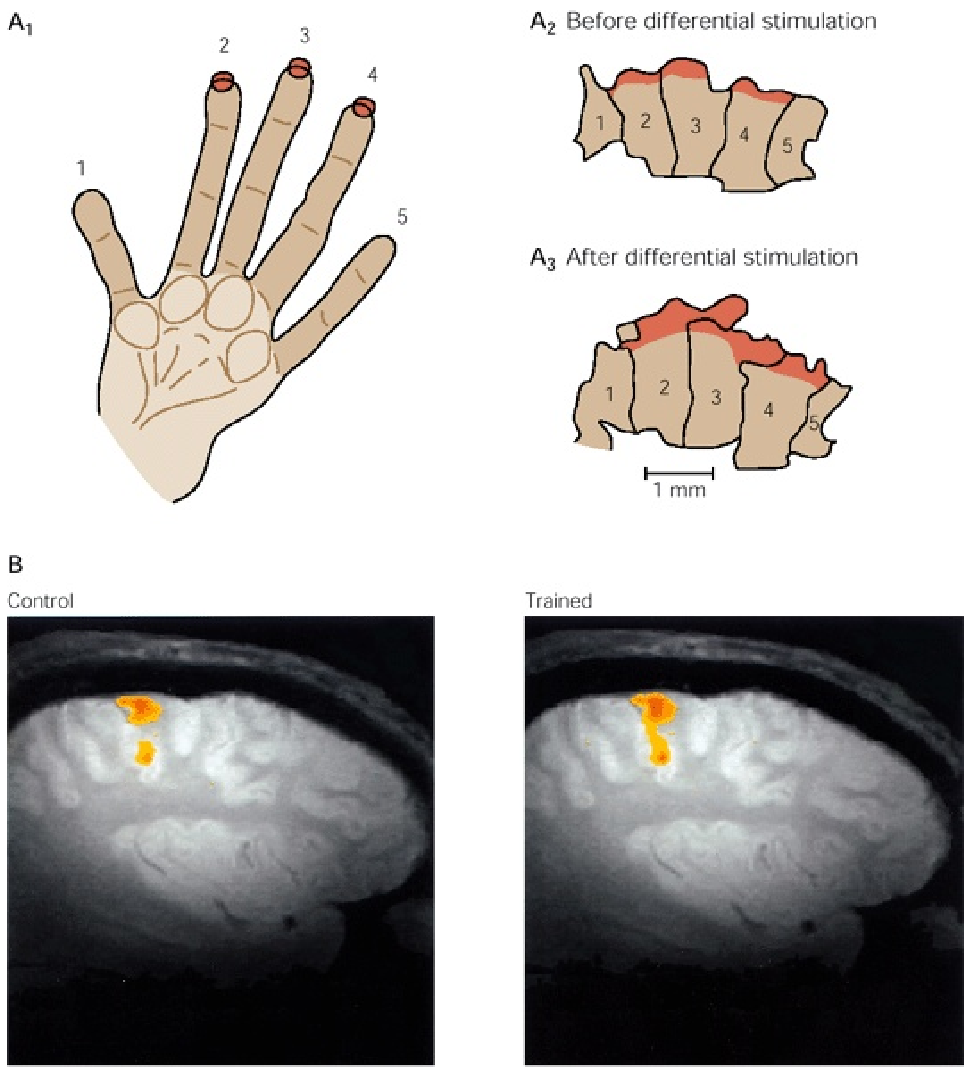
Intrinsic somatic plasticity is synapse-unspecific



Maffei, Turrigiano, 2008

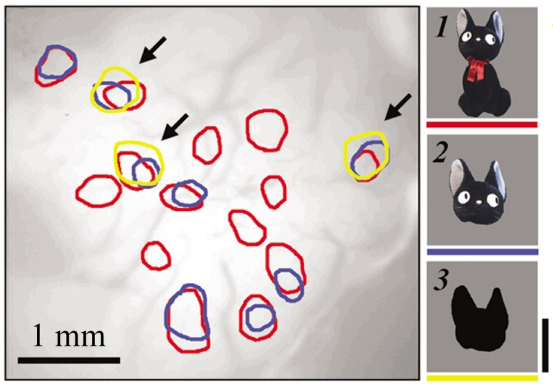
Neural Substrate of Learning : Plasticity

Local plasticity mechanisms and network system-level learning

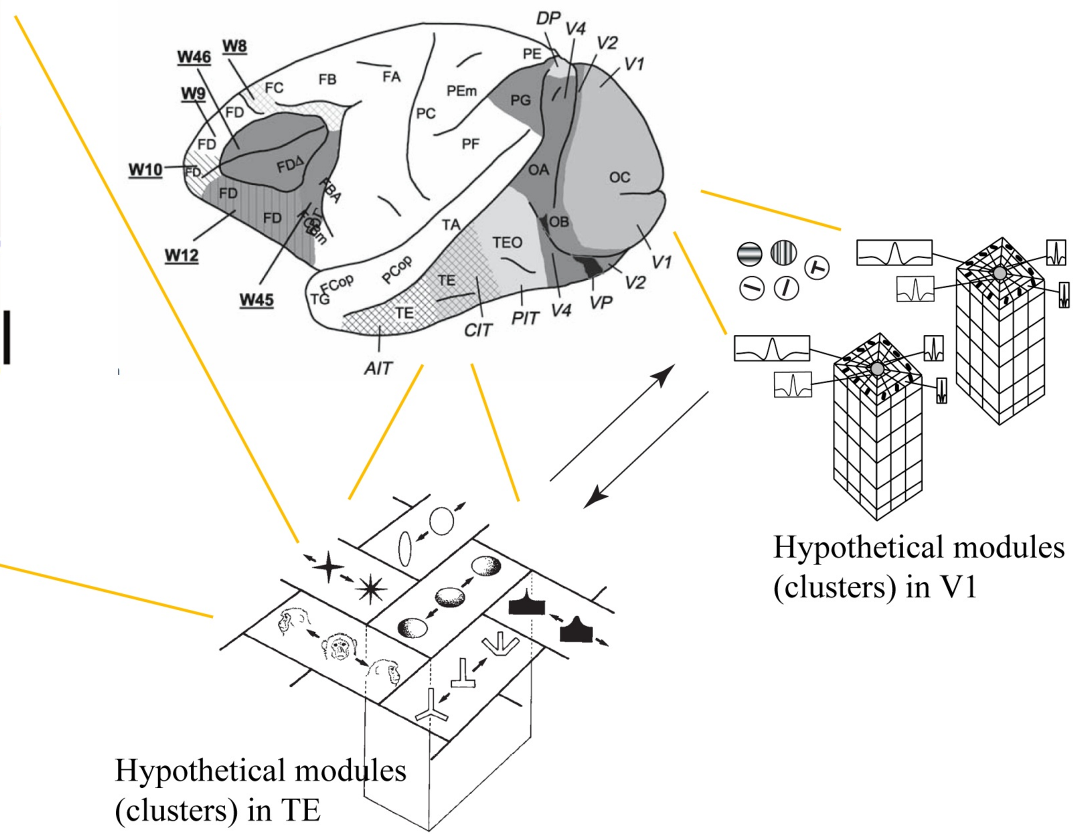


Neural Substrate of Learning : Plasticity

Local plasticity mechanisms and network system-level learning



Cluster responses in dorsal TE, macaque
(via optical instrinsic imaging)

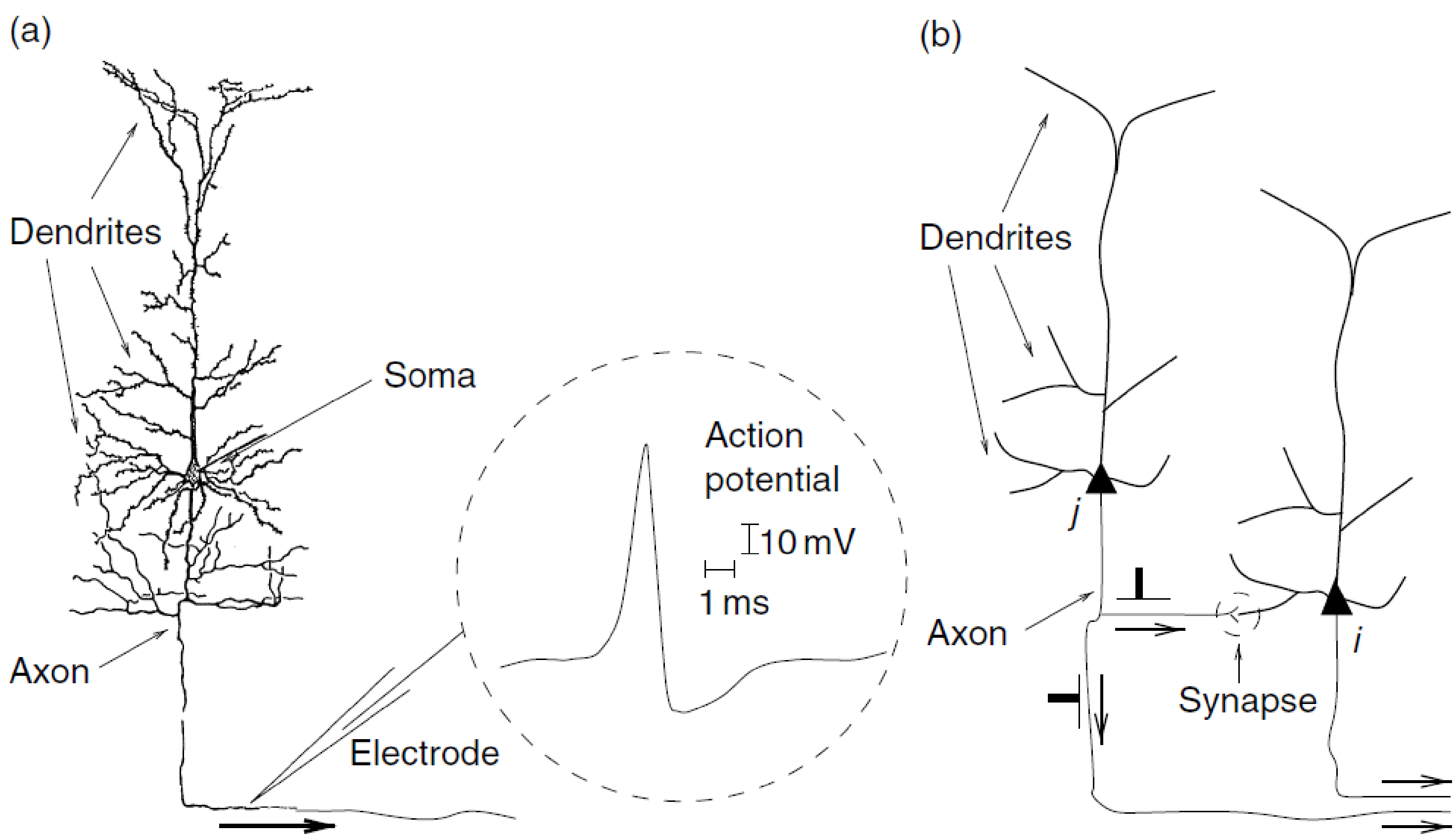


Hypothetical modules
(clusters) in V1

Hypothetical modules
(clusters) in TE

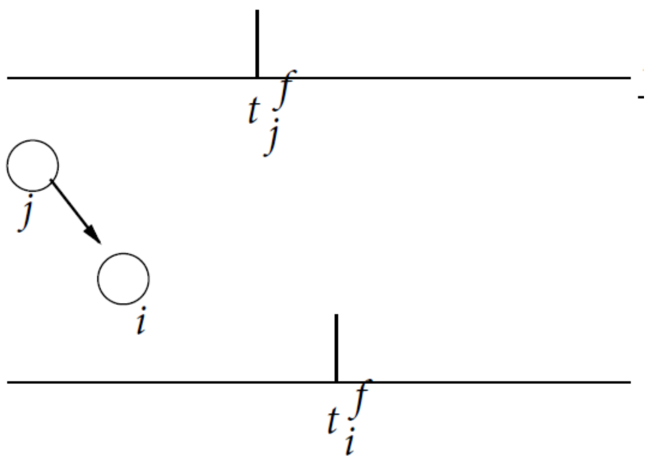
DeValois, 1990; Tanaka, 1996, 2003; Tsunoda et al., 2001

Mathematical models of neural plasticity

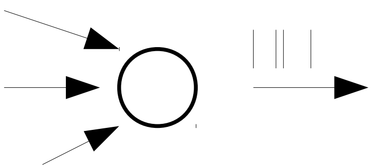


Neural Substrate of Learning : Plasticity

Locality of plasticity mechanism :



Do not forget important ingredient: dynamics of the neuron itself !



$$\frac{d}{dt} v_i = g(\vec{w}, \vec{v_j}, \mathbf{D})$$

Generic form of a local synaptic plasticity rule:

$$\frac{d}{dt} w_{ij} = F(w_{ij}, v_i, v_j, \mathbf{D})$$

where all variables are locally accessible at the synapse

Neural Substrate of Learning : Plasticity

Taylor expansion gives all possible local terms:

$$\begin{aligned}\frac{d}{dt}w_{ij} = c_0(w_{ij}) \ + \ c_1^{pre}(w_{ij})v_j \ + \ c_1^{post}(w_{ij})v_i \ + \\ + \ c_2^{pre}(w_{ij})v_j^2 \ + \ c_2^{post}(w_{ij})v_i^2 \ + \\ + \ c_1^{corr}1(w_{ij})v_iv_j \ + \ \mathcal{O}(v^3)\end{aligned}$$

<u>Any</u> specific synaptic plasticity rule can be derived from this generic form

- Standard Hebb (fire together – wire together): $\frac{d}{dt}w_{ij} \propto v_iv_j$
- Hebb with decay: $\frac{d}{dt}w_{ij} \propto v_iv_j - c_0$
- Hebb with LTP/LTD threshold: $\frac{d}{dt}w_{ij} \propto (v_i - v_\theta)v_j$, $\frac{d}{dt}w_{ij} \propto v_i(v_j - v_\theta)$
- Covariance rule: $\frac{d}{dt}w_{ij} \propto (v_i - \langle v_i \rangle)(v_j - \langle v_j \rangle)$
- Bienenstock-Cooper-Monro (BCM): $\frac{d}{dt}w_{ij} = \eta v_iv_j(v_i - v_\theta(t))$
- *etc*

Neural Substrate of Learning : Plasticity

Usage of pre- and post variables in terms defines the direction of synaptic change (potentiation, LTP, depression, LTD, or none)

Post v_i	Pre v_j	$\mathrm{d}w_{ij}/\mathrm{d}t \propto v_i v_j$	$\mathrm{d}w_{ij}/\mathrm{d}t \propto v_i v_j - c_0$	$\mathrm{d}w_{ij}/\mathrm{d}t \propto (v_i - v_\theta) v_j$	$\mathrm{d}w_{ij}/\mathrm{d}t \propto v_i (v_j - v_\theta)$	$\mathrm{d}w_{ij}/\mathrm{d}t \propto (v_i - \langle v_i \rangle)(v_j - \langle v_j \rangle)$
ON	ON	+	+	+	+	+
ON	OFF	0	–	0	–	–
OFF	ON	0	–	–	0	–
OFF	OFF	0	–	0	0	+

Neural Substrate of Learning : Plasticity

Stability analysis : determining stable synaptic weight configurations given a particular rule

A simple example : Hebb like rule with weight dependance.

$$\frac{d}{dt}w_{ij} = \eta v_i (v_j - w_{ij})$$

Weights converge to presynaptic activities v_j (v_j are fixed, stable points of the synapic modification)

$$f(w_{ij}) = \eta v_i v_j - \eta v_i w_{ij},$$
$$\eta v_i v_j - \eta v_i w_{ij} = 0, (v_i > 0, v_j > 0) \Leftrightarrow w_{ij} = v_j;$$
$$f'(w_{ij}) = -\eta v_i < 0 \Rightarrow \text{stable fixed point}$$

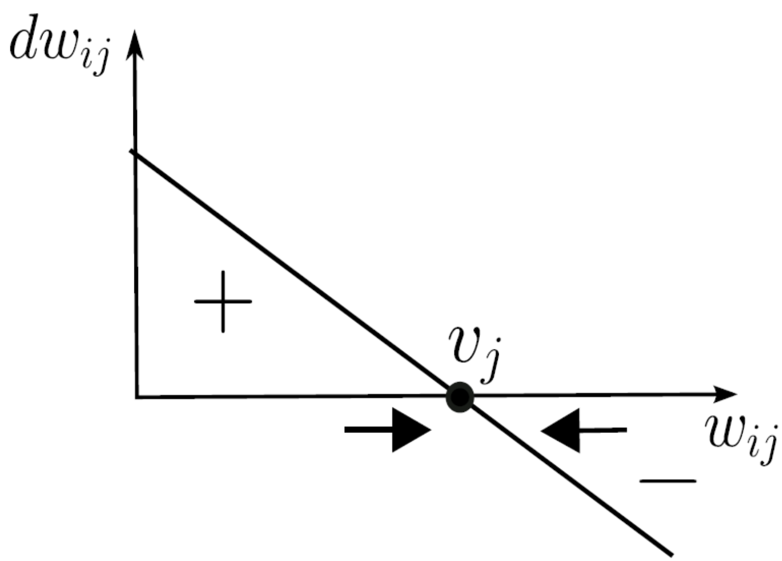
Neural Substrate of Learning : Plasticity

Stability analysis : determining stable synaptic weight configurations given a particular rule

A simple example : Hebb like rule with weight dependence

$$\frac{d}{dt}w_{ij} = \eta v_i (v_j - w_{ij})$$

Weights converge to presynaptic activities v_j (v_j are fixed, stable points of the synapic modification)



Neural Substrate of Learning : Plasticity

Stability analysis : determining stable synaptic weight configurations given a particular rule

A more advanced example : inferring weight distribution from the type of synaptic modification rule

$$c_{11}^{corr}(w_{ij}) = \gamma_2(w^{max} - w_{ij})^{\beta}$$

Depending on form of weight dependence (regulated by β), weight distribution can become

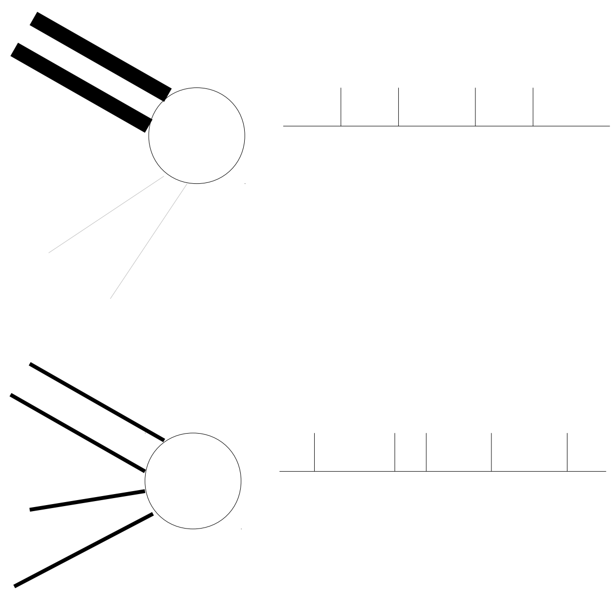
- Bimodal for additive weight dependence (β = 0 : strong competition between the weights, weights either very strong or very weak)
- Unimodal for multiplicative weight dependence (β = 1 : weak competition between the weights, intermediate values are possible)

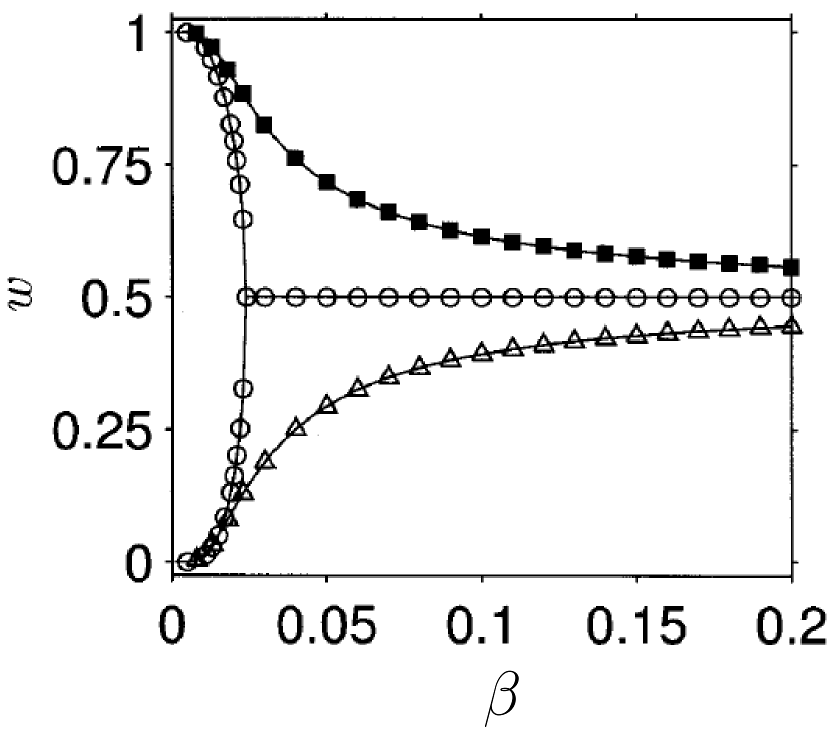
Neural Substrate of Learning : Plasticity

Stability analysis : determining stable synaptic weight configurations given a particular rule

A more advanced example : inferring weight distribution from the type of synaptic modification rule

$$c_{11}^{corr}(w_{ij}) = \gamma_2(w^{max} - w_{ij})^{\beta}$$





Gütig et al, 2003

Neural Substrate of Learning : Plasticity

Phenomenological or functional plasticity rules

Reproduce

- phenomena observed in a neurophysiological experiment), or
- a function / computation that is able to do information processing / learning in a defined task
- or ideally both

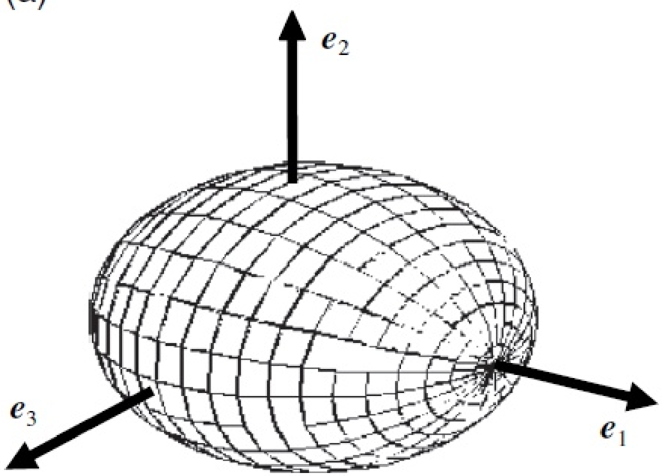
Neural Substrate of Learning : Plasticity

Oja rule (Oja, 1982) : LTP or LTD depending on pre, post and the current synaptic weight

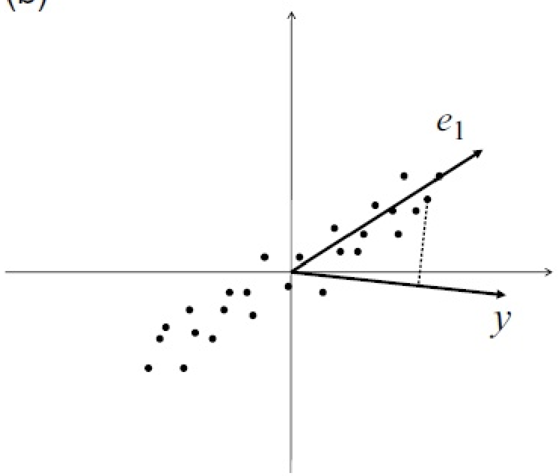
- No experimental support
- Self-stabilizing rule : synaptic weights are self-normalized and cannot grow without bounds
- Unsupervised learning : formally equal to Principal Component Analysis (PCA → provides a mean to perform dimensionality reduction)

$$\frac{d}{dt} w_{ij} = \gamma (v_i v_j - w_{ij} v_i^2)$$

(a)



(b)

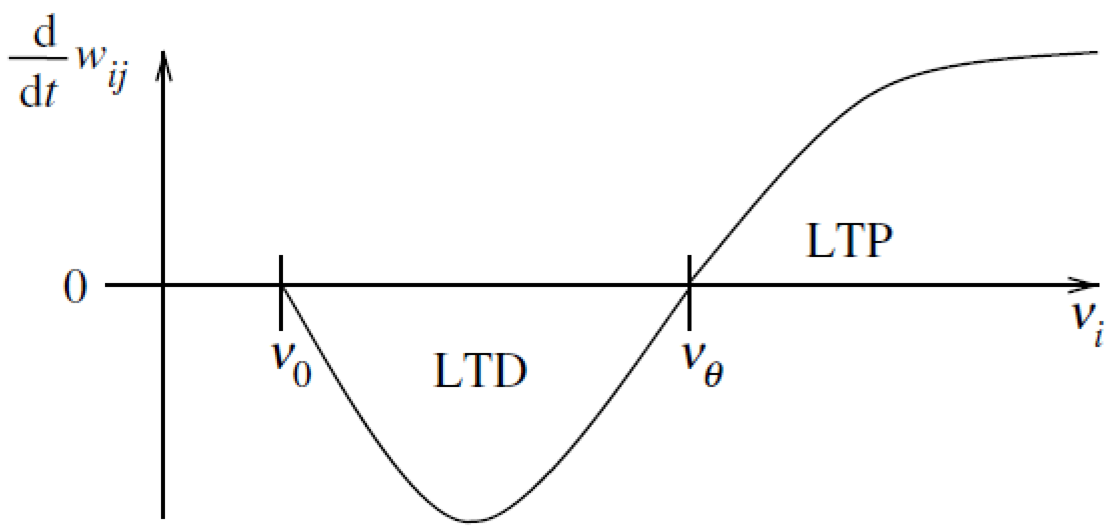


Neural Substrate of Learning : Plasticity

BCM (Bienenstock-Cooper-Monroe, 1982) : sliding threshold (modified by previous history of activation) separating LTP / LTD zones

- Experimentally supported (visual cortex plasticity during early development, monocular / binocular deprivation experiments)
- Unsupervised learning : development of input selectivity (receptive field formation)
- Relation to projection pursuit, function / density approximation

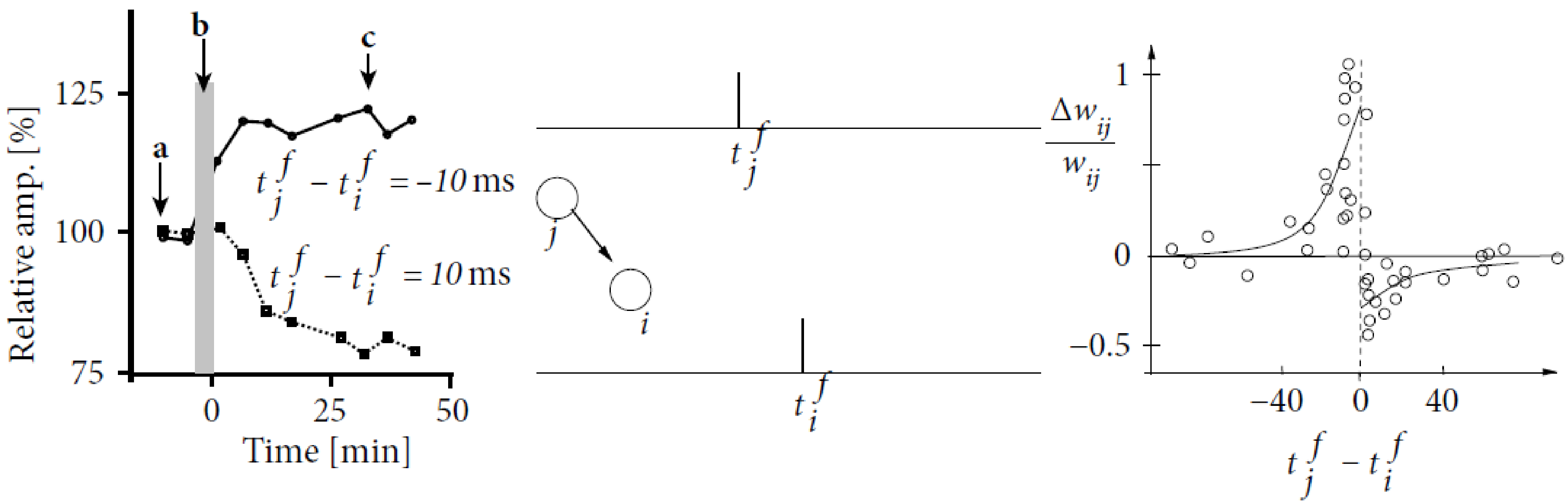
$$\frac{d}{dt}w_{ij} = \eta v_i v_j (v_i - v_\theta(t))$$



Neural Substrate of Learning : Plasticity

Spike timing dependent plasticity (STDP) family of rules : take into account explicitly the time of spikes t_j, t_i on pre and post synaptic sides

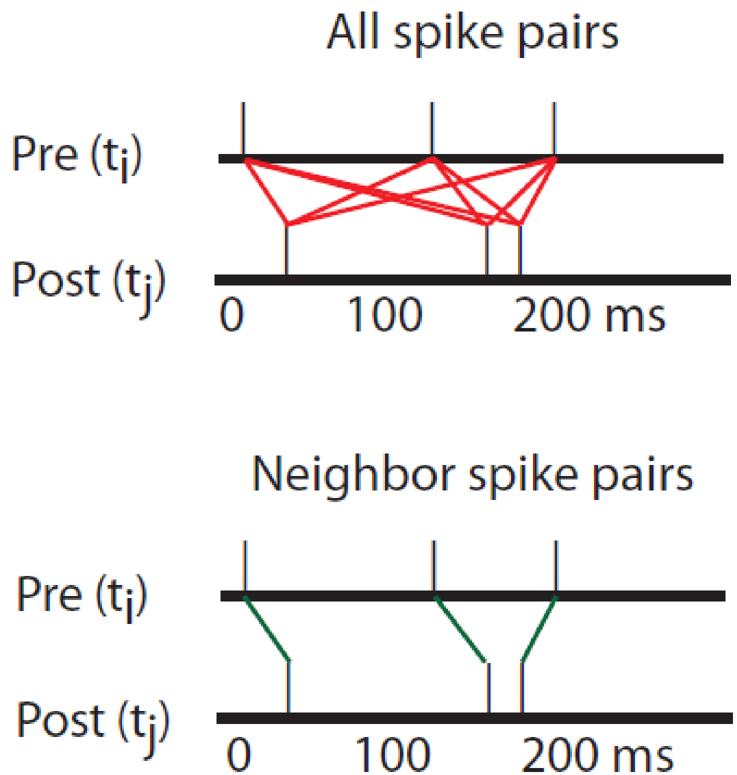
- Experimental support that timing often matters, but no experimental support for universal synaptic plasticity based on timing only
- Unsupervised learning : development of input selectivity, receptive field formation, Independent Component Analysis (ICA), etc



Markram et al, 1997

Neural Substrate of Learning : Plasticity

Spike timing dependent plasticity (STDP) family of rules : pair-based form, taking only spike pairs into account

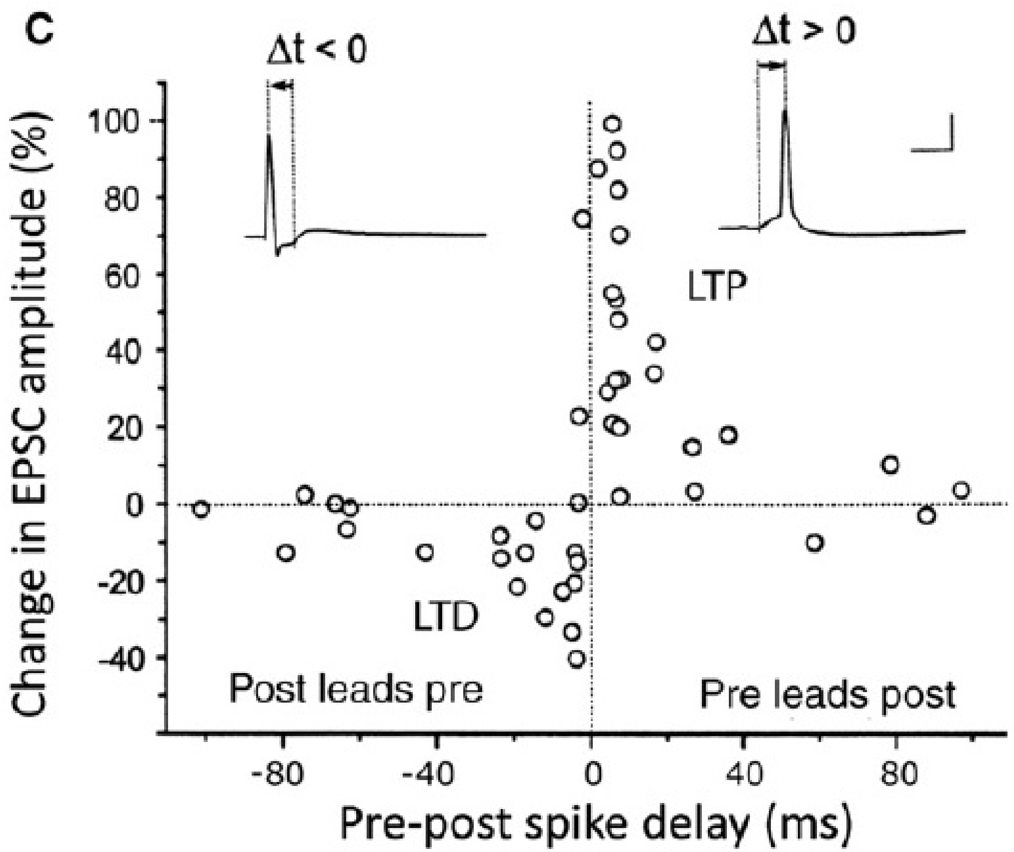


$$\Delta w_+ = A_+(w) \cdot \exp(-|\Delta t|/\tau_+) \text{ at } t_{\text{post}} \quad \text{for } t_{\text{pre}} < t_{\text{post}},$$

$$\Delta w_- = A_-(w) \cdot \exp(-|\Delta t|/\tau_-) \text{ at } t_{\text{pre}} \quad \text{for } t_{\text{pre}} < t_{\text{post}},$$

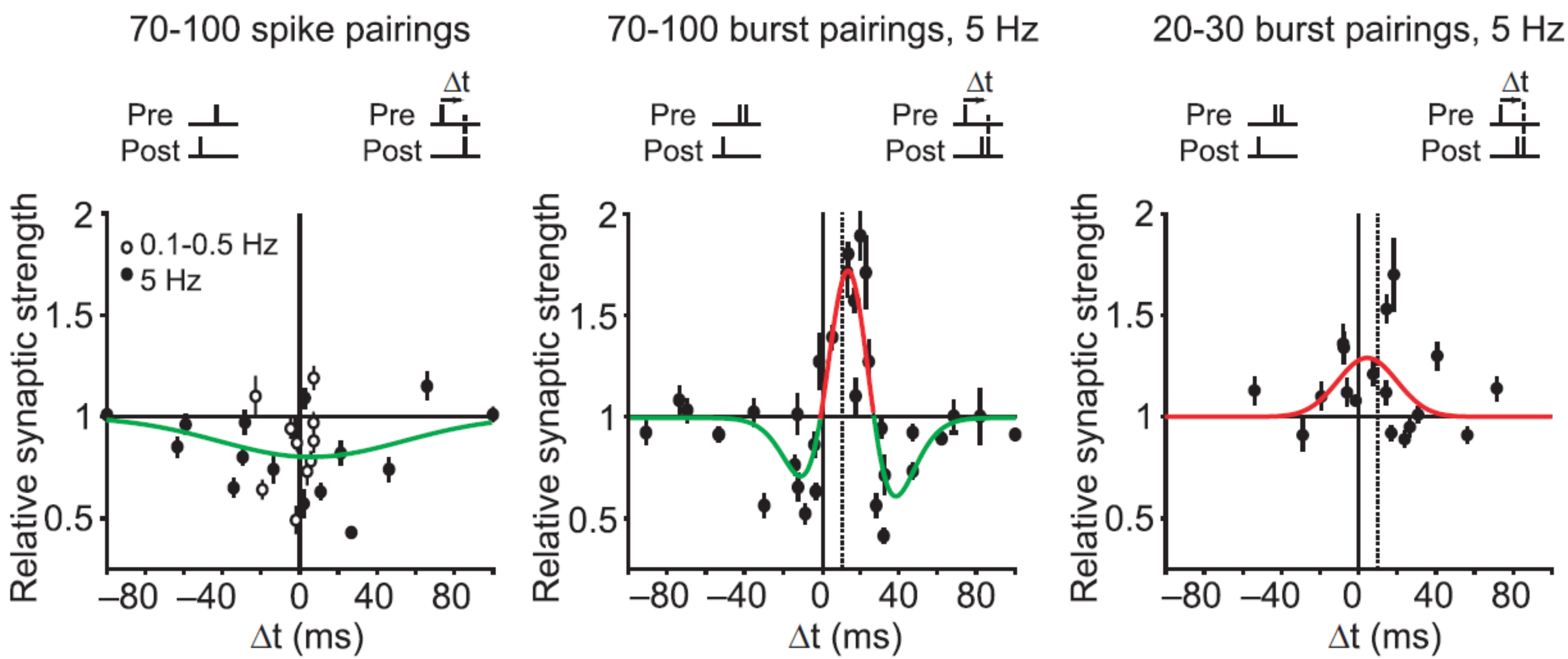
Neural Substrate of Learning : Plasticity

Spike timing dependent plasticity (STDP) family of rules :
however, in experiments you can generate almost arbitrary forms of modification that do not fit any pair-based spike timing only scheme



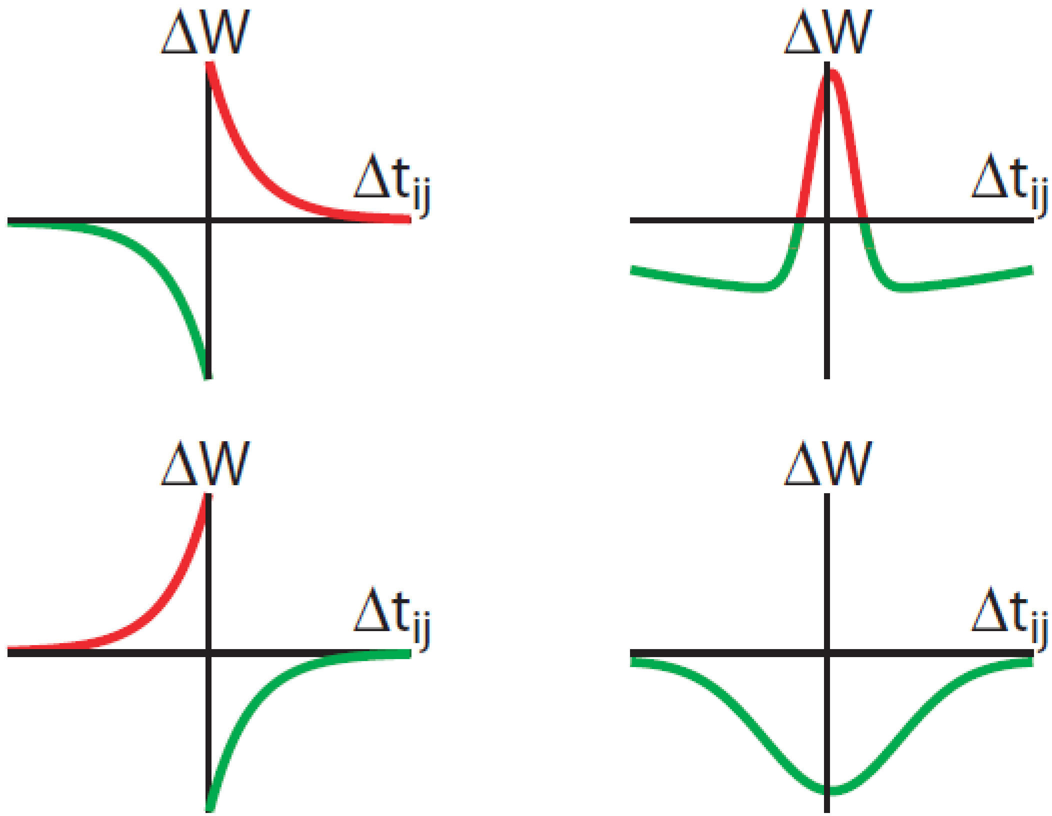
Neural Substrate of Learning : Plasticity

Spike timing dependent plasticity (STDP) family of rules :
however, in experiments you can generate almost arbitrary forms of modification that do not fit any pair-based spike timing only scheme



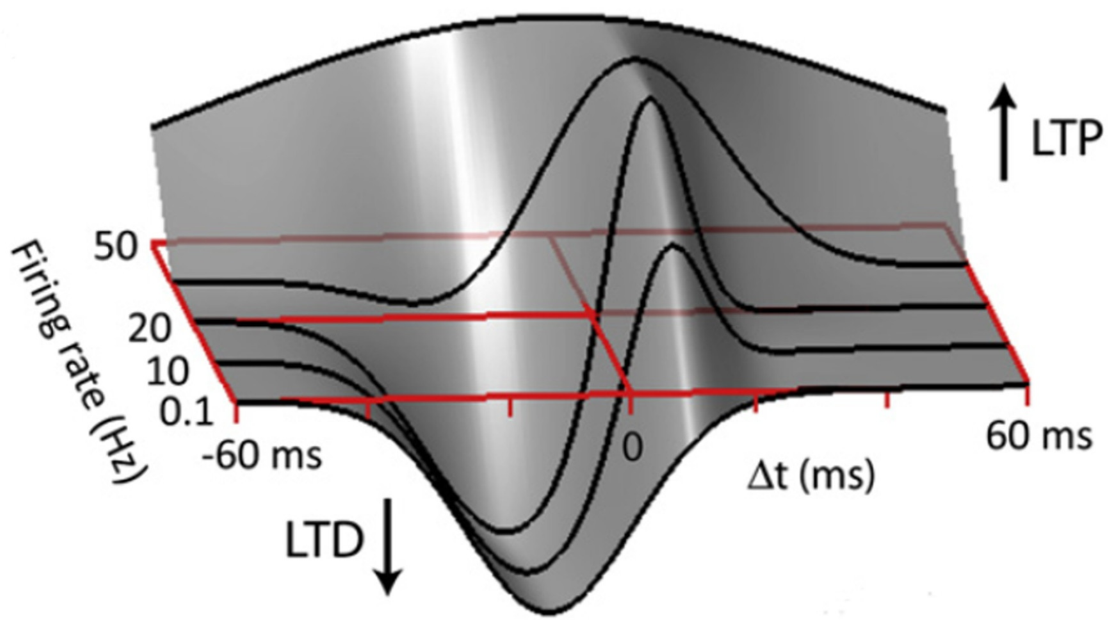
Neural Substrate of Learning : Plasticity

Spike timing dependent plasticity (STDP) family of rules :
however, in experiments you can generate almost arbitrary forms of modification that do not fit any pair-based spike timing only scheme



Neural Substrate of Learning : Plasticity

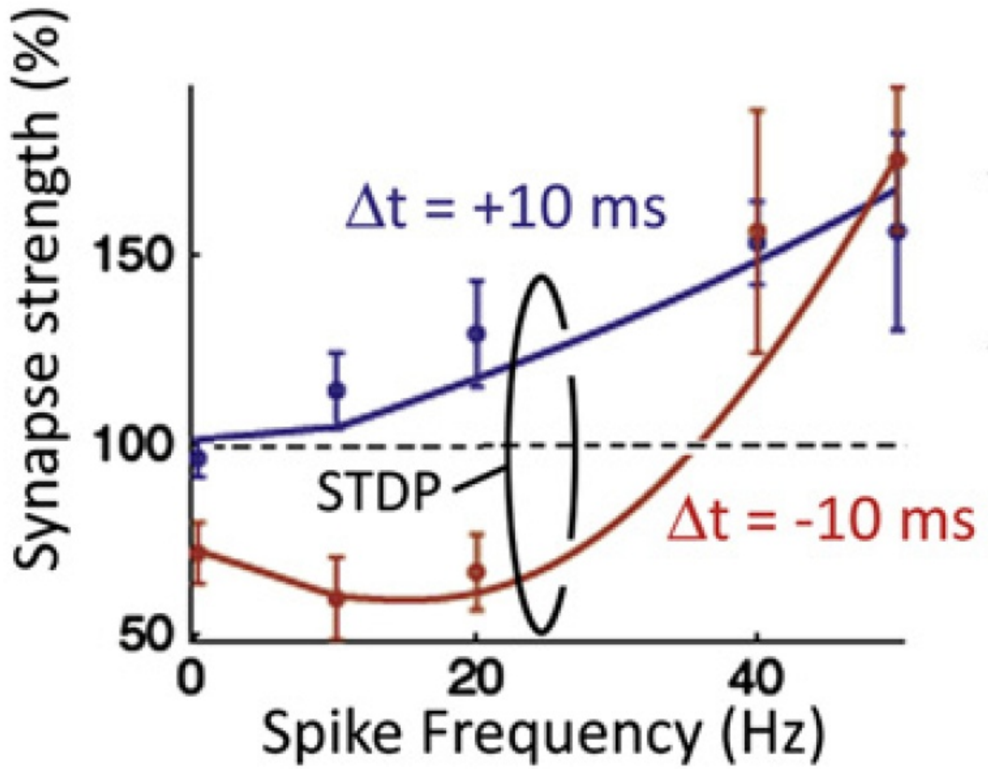
Spike timing dependent plasticity (STDP) family of rules :
however, in experiments you can generate almost arbitrary forms of modification that do not fit any pair-based spike timing only scheme



Neural Substrate of Learning : Plasticity

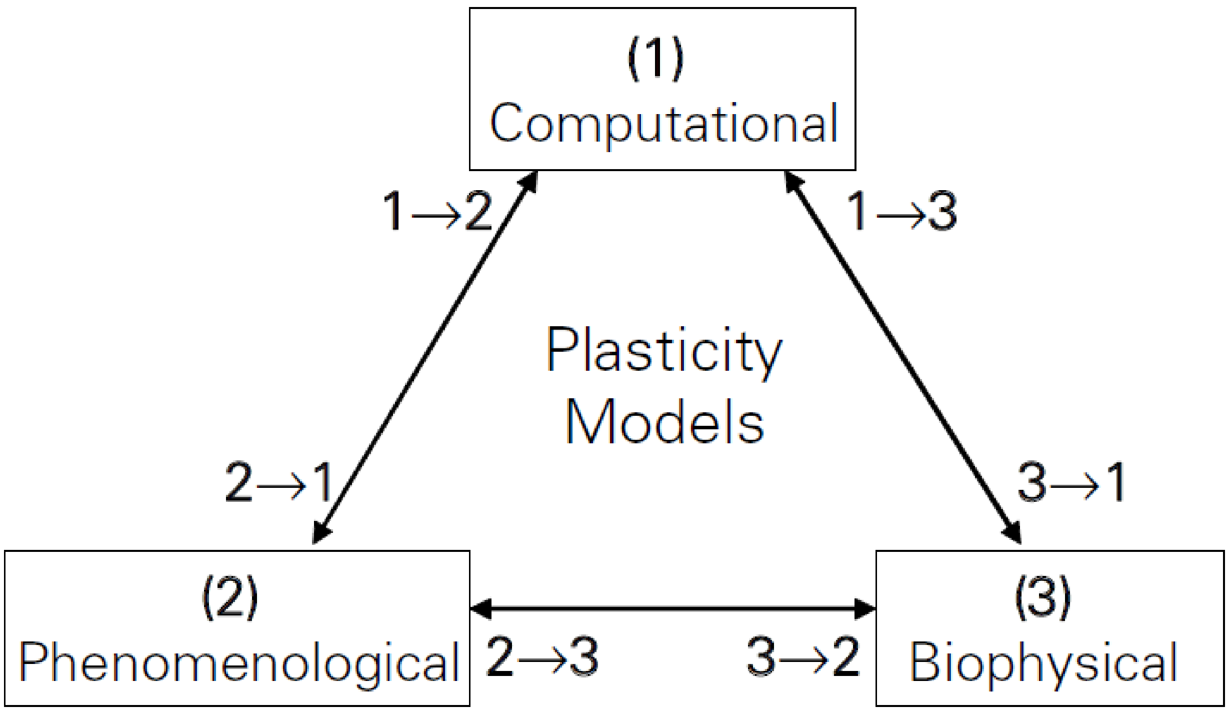
Spike timing dependent plasticity (STDP) : timing is not everything, other factors are at least as important
→ too simple synaptic plasticity rules may be just wishful thinking

Voltage-dependent STDP, Clopath et al, 2010 : example of an attempt to correct pure STDP by taking into account further variables



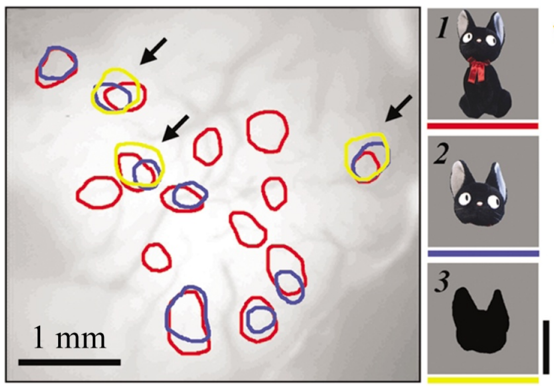
Neural Substrate of Learning : Plasticity

Understanding plasticity requires a good guess of the objective of learning : synaptic changes occur under very diverse conditions, without an idea what information processing may stay behind them, one gets lost in this diversity.

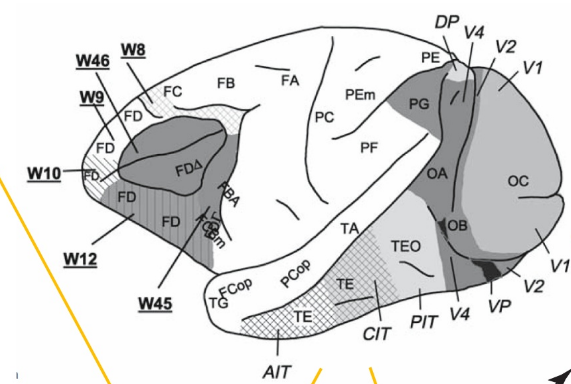


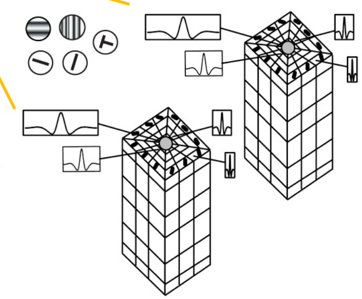
Neural Substrate of Learning : Plasticity

System level learning : what does the neural network with changing synapses trying to do?

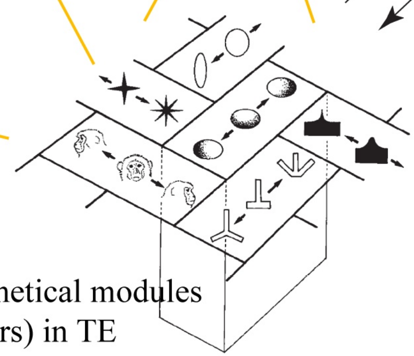


Cluster responses in dorsal TE, macaque
(via optical instrinsic imaging)





Hypothetical modules
(clusters) in V1

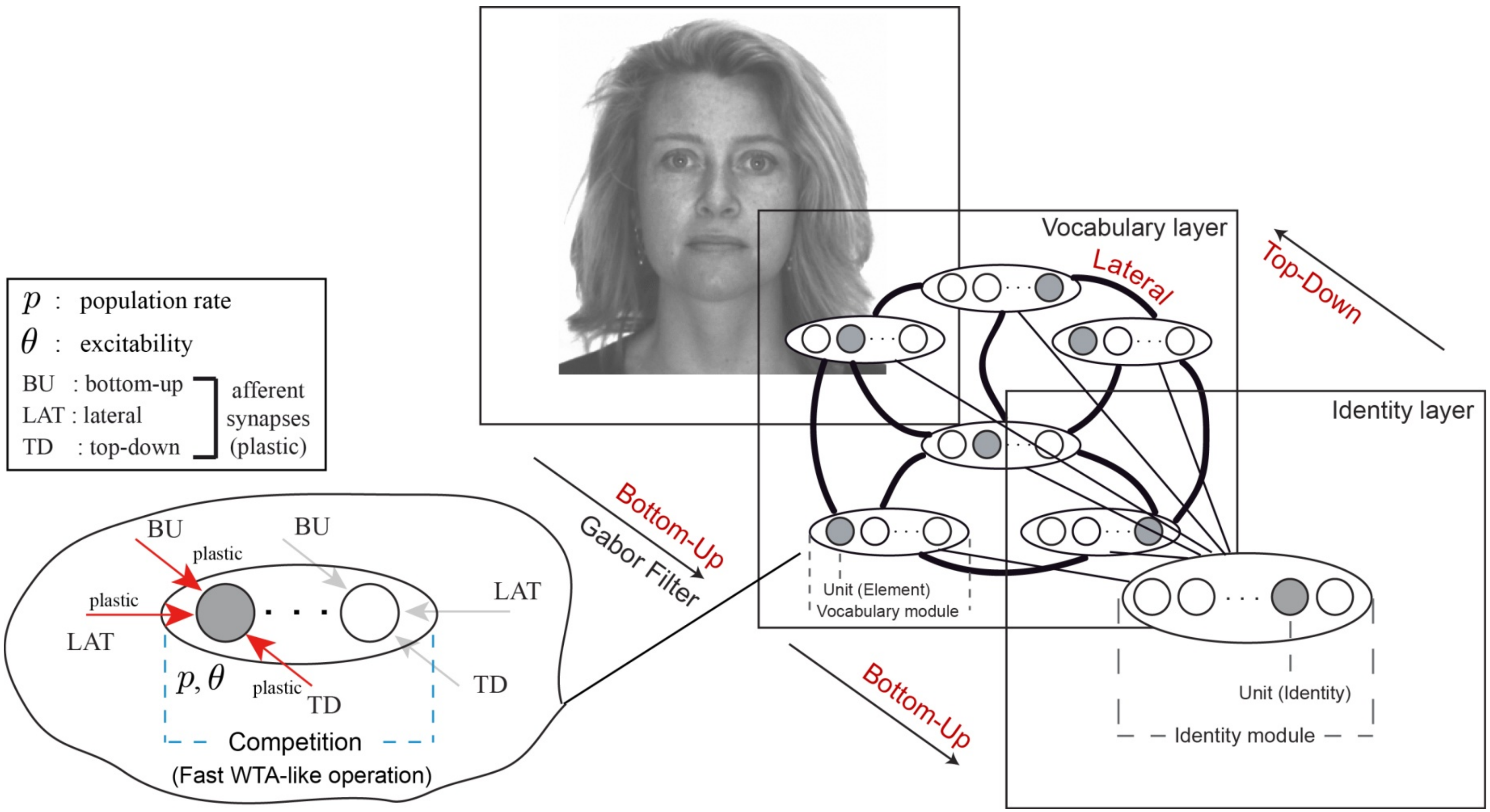


Hypothetical modules
(clusters) in TE

DeValois, 1990; Tanaka, 1996, 2003; Tsunoda et al., 2001

Neural Substrate of Learning : Plasticity

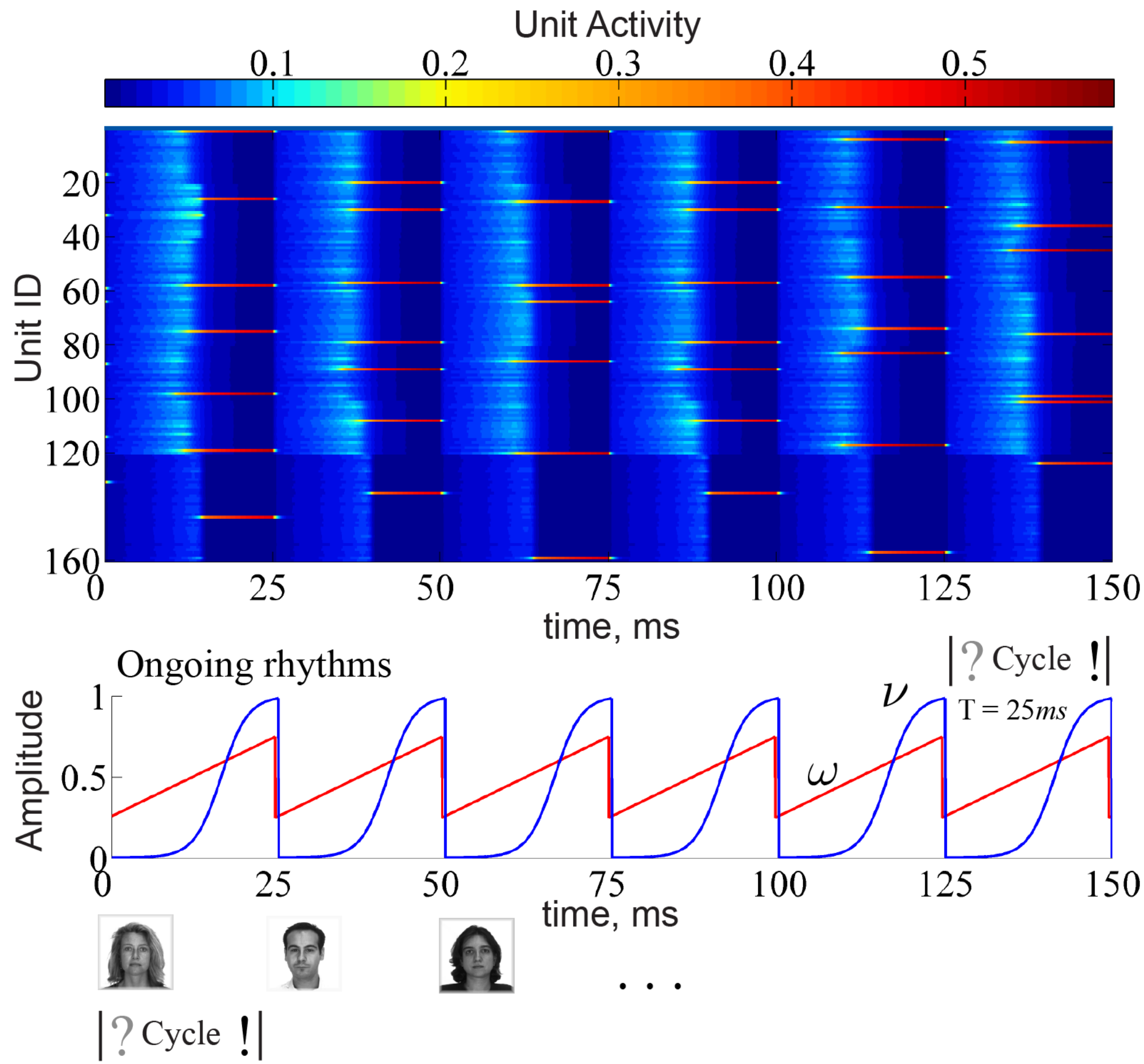
Unsupervised learning in network of winner-take-all (WTA) circuits



Jitsev and von der Malsburg, 2009, 2010; Jitsev, 2014

Neural Substrate of Learning : Plasticity

Unsupervised learning in network of winner-take-all (WTA) circuits



Jitsev and von der Malsburg, 2009, 2010; Jitsev, 2014

Neural Substrate of Learning : Plasticity

System level learning : formalize objective of the learning as an optimization that seeks to maximize organisms well-being
→ maximize “reward”, minimize “punishment”
→ To do that, it is necessary to develop ability to predict upcoming events in the environment
→ To do that, it is necessary to build a model what organism's actions may cause in the environment

Neural Substrate of Learning : Plasticity

"What is painful is avoided and what is pleasant is pursued."
(Aristotle, De Motu Animalium, 400 B.C.)

Neural Substrate of Learning : Plasticity

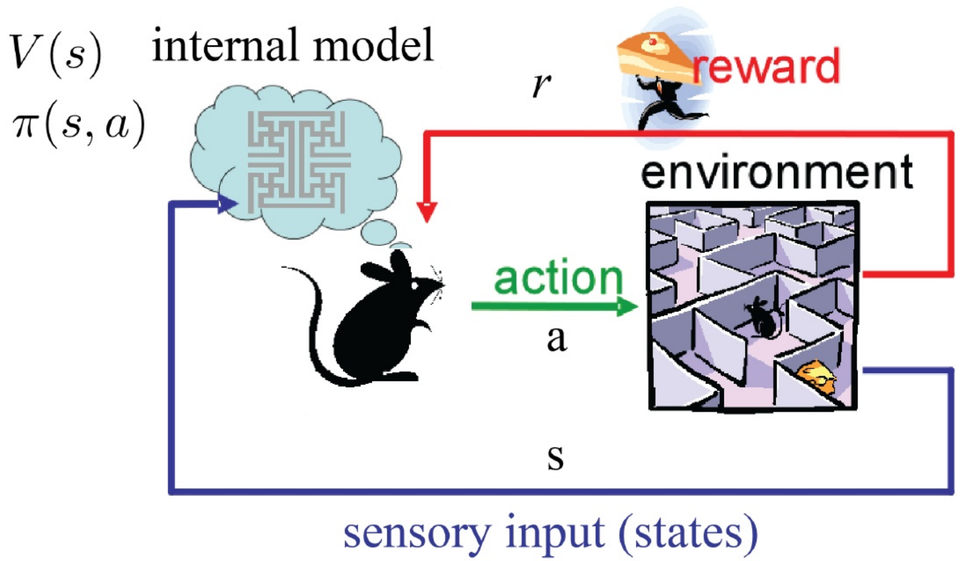
Predict the environment and act to gain benefit and avoid harm

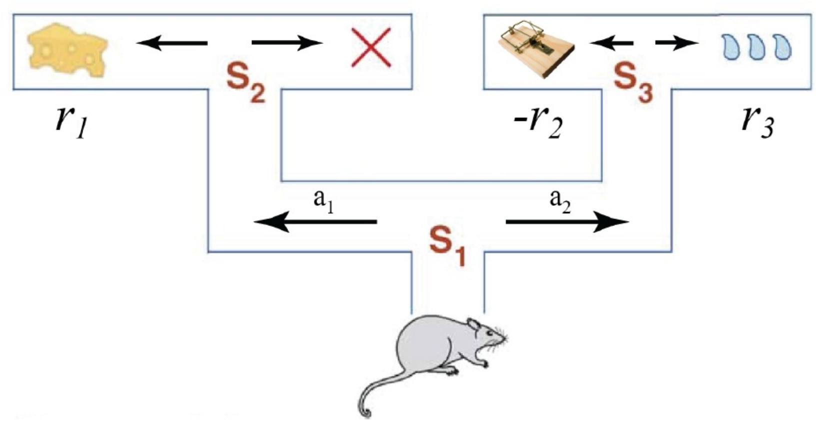
Neural Substrate of Learning : Plasticity

<u>A functional neural architecture</u>:

- **Sense and Predict : internal model and outcome expectations**
- **Act : (near) optimal decision making (based on expectations)**
- **Learn : improve to get more reward and less punishment**

A functional model of reinforcement learning





Barto et al, 1983; Niv et al, 2006

Bellman equations

Bellman equations : mathematical foundation for optimizing decision making and behavior

- **A value function *V(s)* describes expected outcome (given a policy for action selection)**

$$V^\pi(s_t) = E(r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots | \pi) =$$
$$= E(\sum_{k=0} \gamma^k r_{t+k+1} | \pi)$$

- **Optimal value function provides best possible outcome**

$$V^*(s_t) = \max_{a \in \mathcal{A}(s_t)} E\left[r_{t+1} + \gamma V^*(s_{t+1}) | s_t, a\right] =$$
$$= \max_{a \in \mathcal{A}(s_t)} \left[R(s_t, a, s_{t+1}) + \gamma \sum_{s_{t+1}} P(s_{t+1} | s_t, a) V^*(s_{t+1})\right]$$

Barto et al, 1983; Niv et al, 2006

Temporal Difference (TD) learning

Learning driven by the prediction error signal

- **Computing prediction error (difference between internal expectation and the following observation)**

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$$

$\delta_t > 0$: better than expected

$\delta_t < 0$: worse than expected

- **Updating expectations and action preferences**

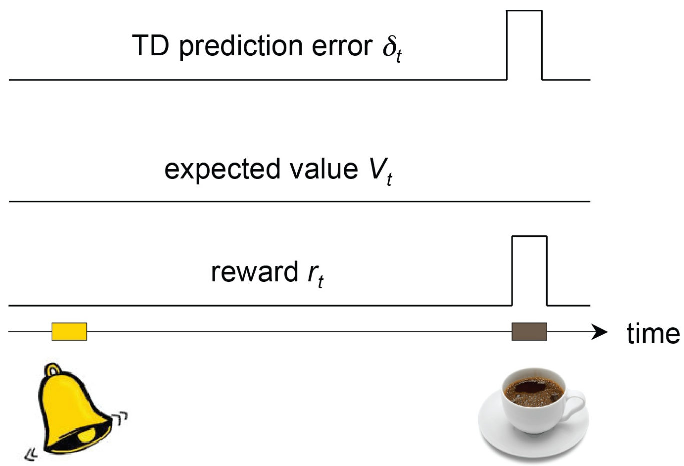
$$V(s_t) \leftarrow V(s_t) + \alpha \delta_t$$

$$\pi(s,a) \leftarrow \pi(s,a) + \varepsilon \delta_t$$

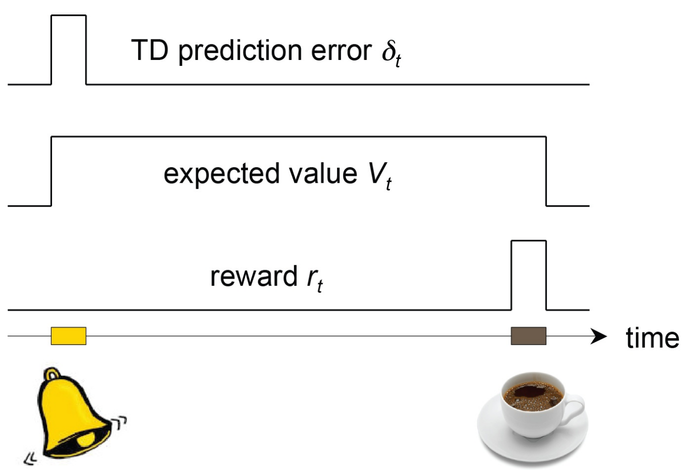
Barto et al, 1983; Niv et al, 2006

Learning and prediction error

Before learning : unpredicted outcome



After learning : predicted outcome



Temporal Difference (TD) learning

Learning driven by the <u>prediction error</u> signal

- **Updating expectations and action preferences**

$$V(s_t) \leftarrow V(s_t) + \alpha\delta_t$$
$$\pi(s,a) \leftarrow \pi(s,a) + \varepsilon\delta_t$$

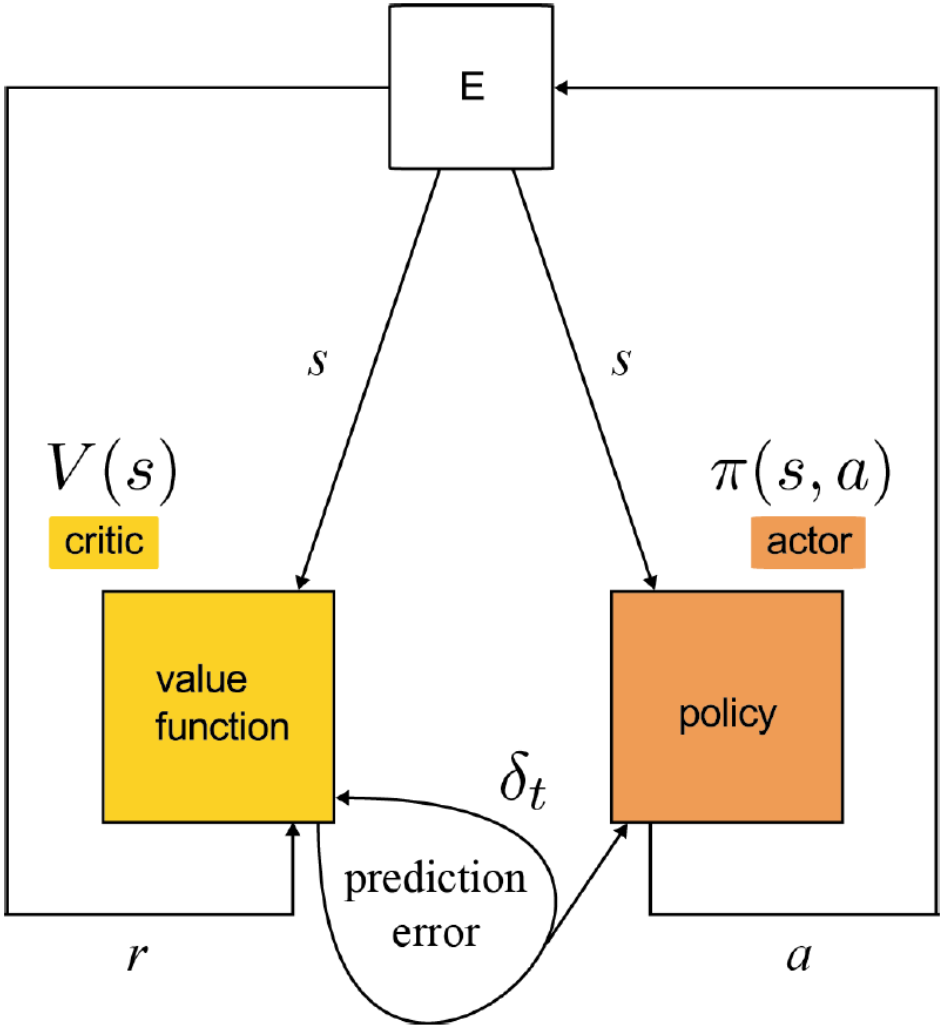
- **Decision making and action selection based on learned preferences (Bolzmann distribution, soft max)**

$$p(a|s) = \frac{e^{-\beta\pi(s,a)}}{\sum_{a'\in\mathcal{A}(s)} e^{-\beta\pi(s,a')}}$$

Barto et al, 1983; Niv et al, 2006

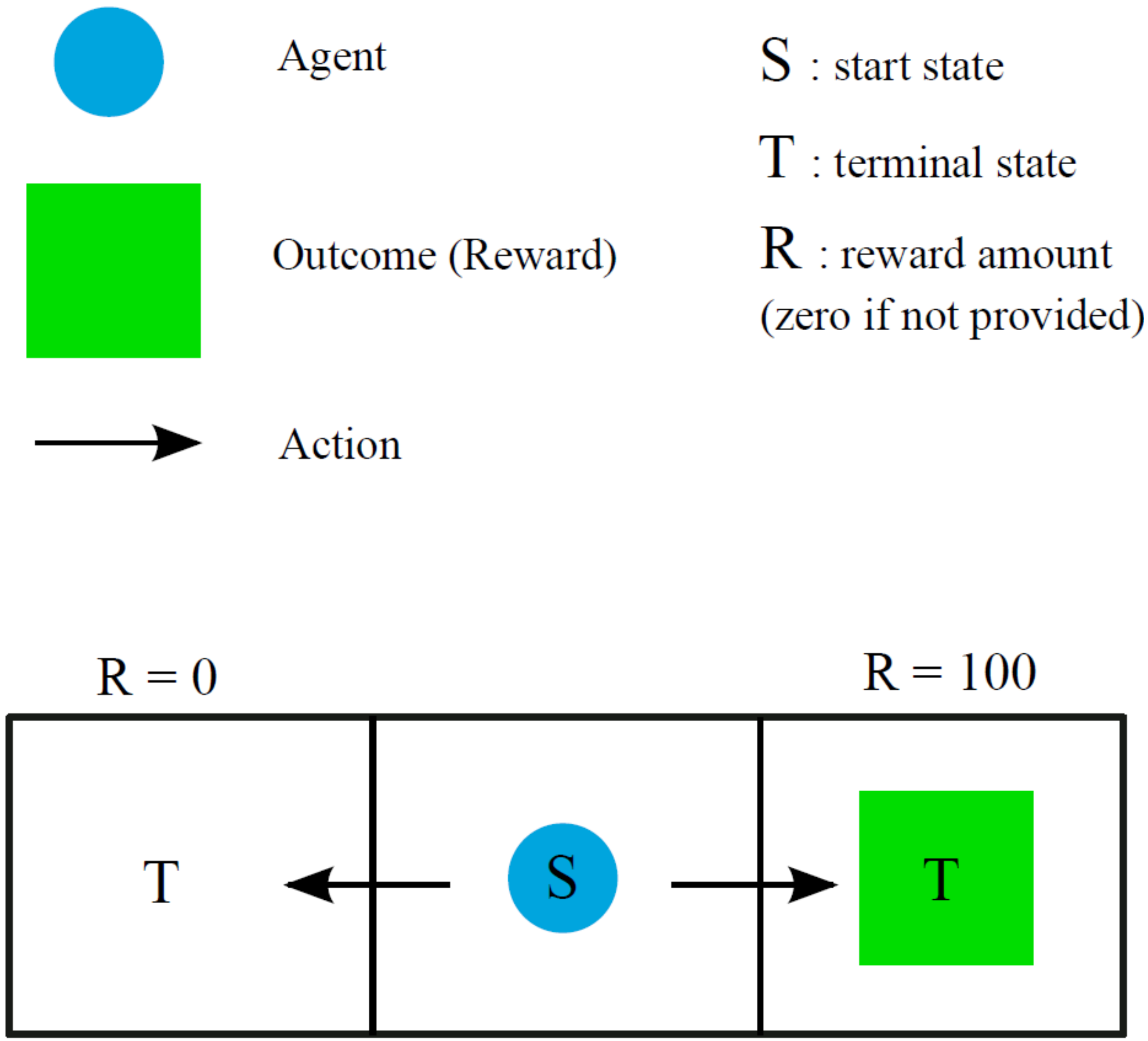
A functional model of reinforcement learning

Instantiation of temporal-difference (TD) learning : Actor-critic architecture

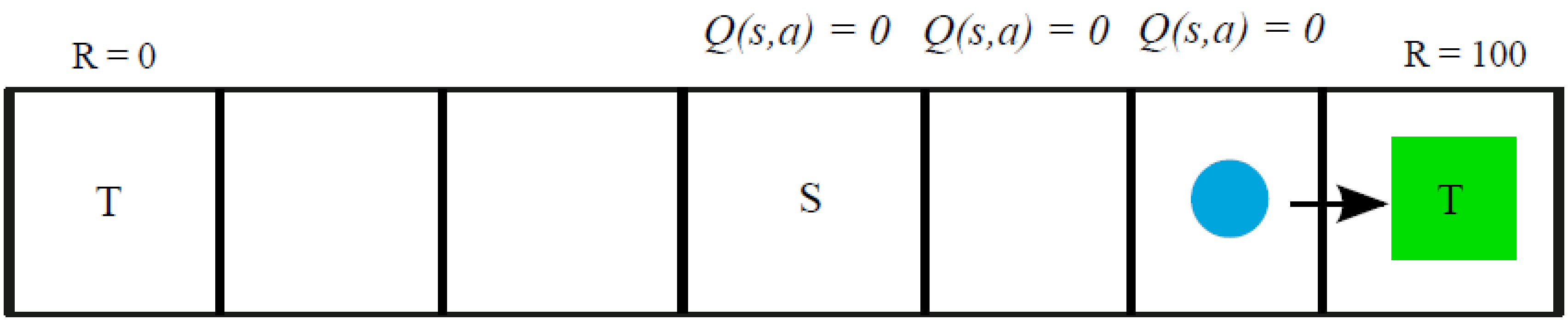


Barto et al, 1983; Sutton and Barto, 1998

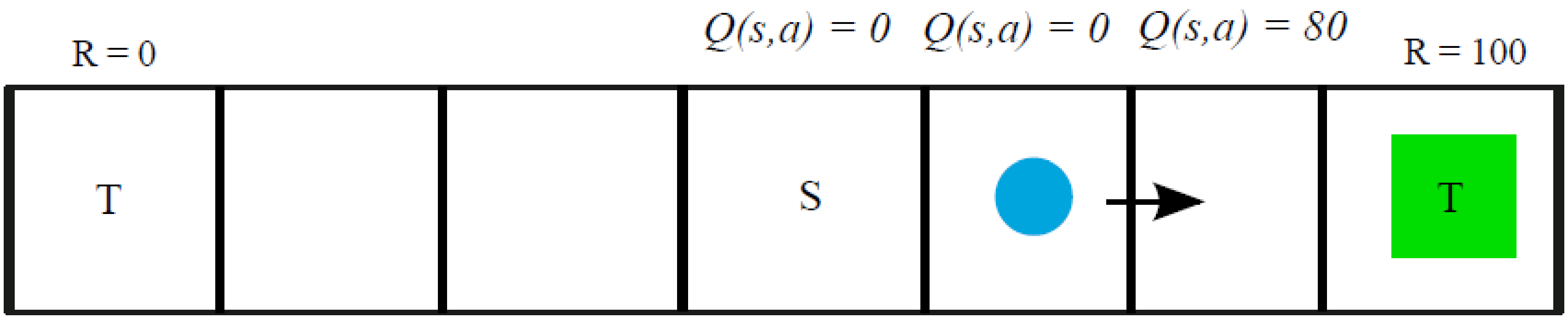
A functional model of reinforcement learning



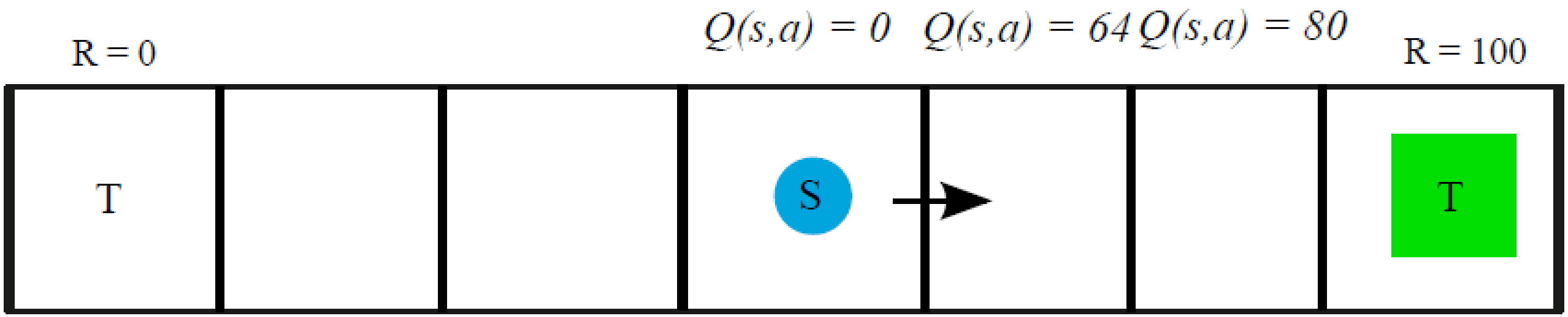
A functional model of reinforcement learning



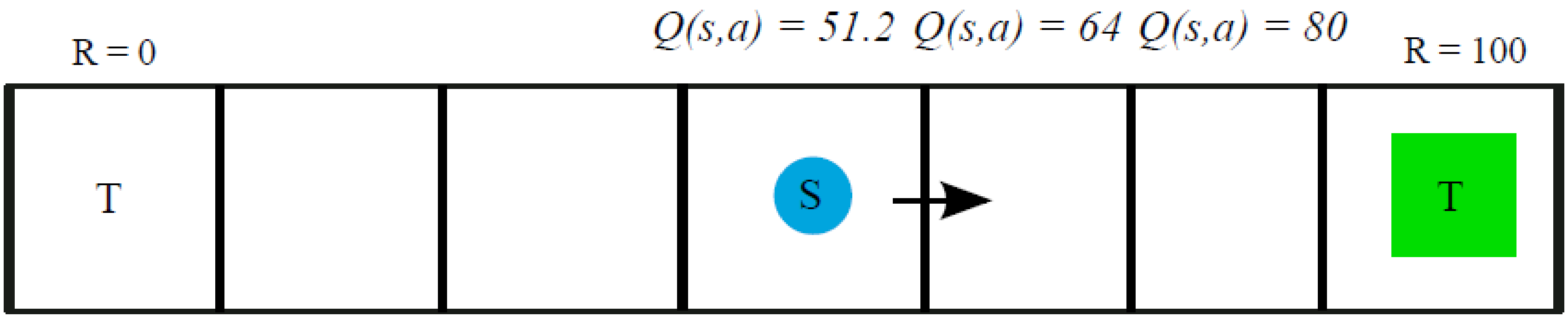
A functional model of reinforcement learning



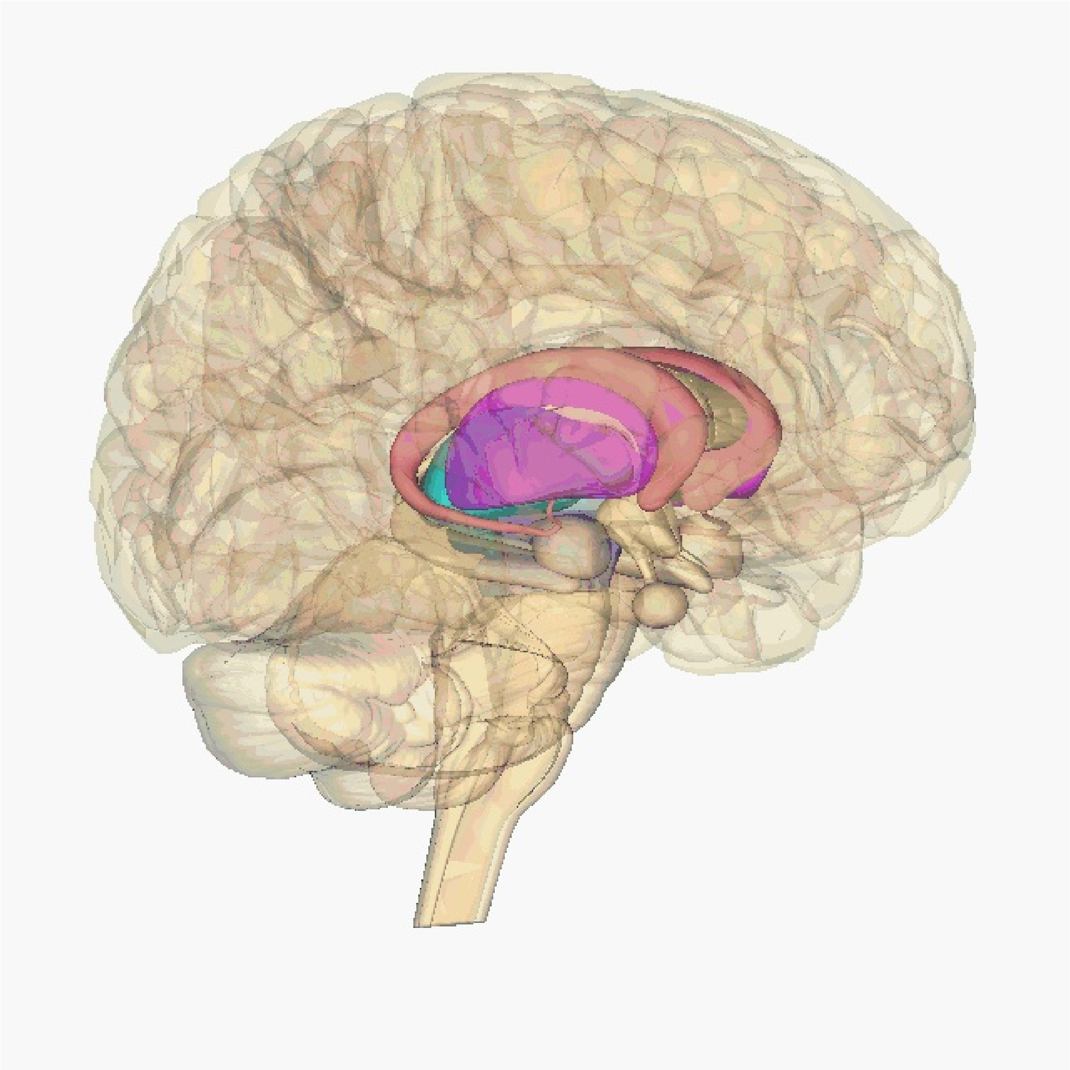
A functional model of reinforcement learning



A functional model of reinforcement learning

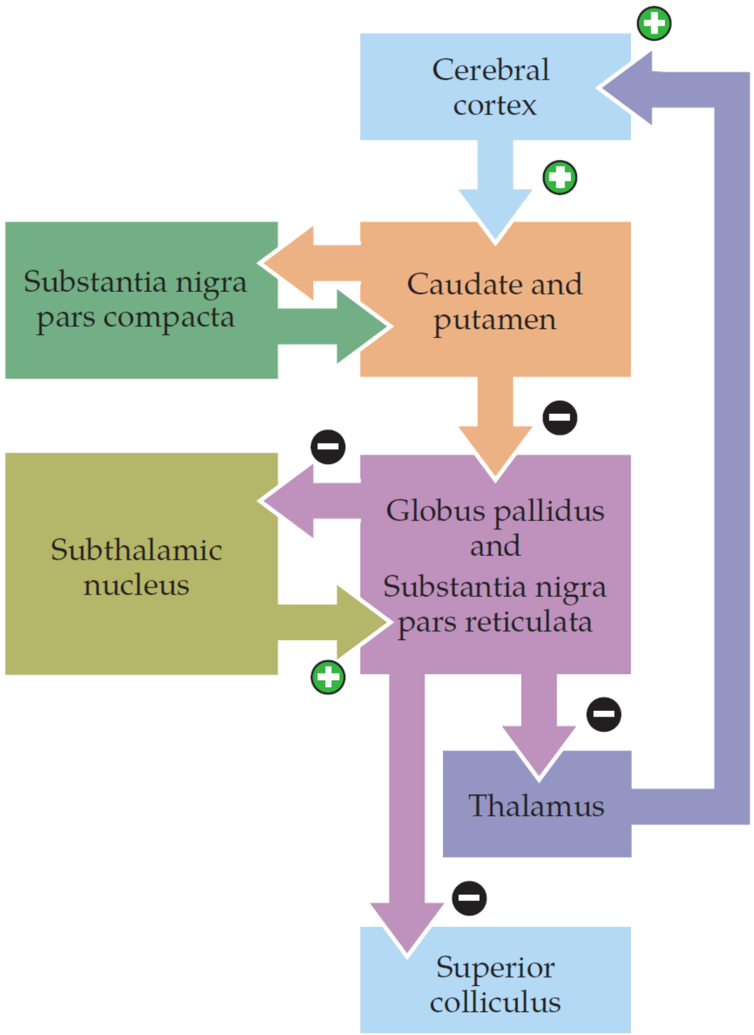


Basal Ganglia and Reward-Based Learning



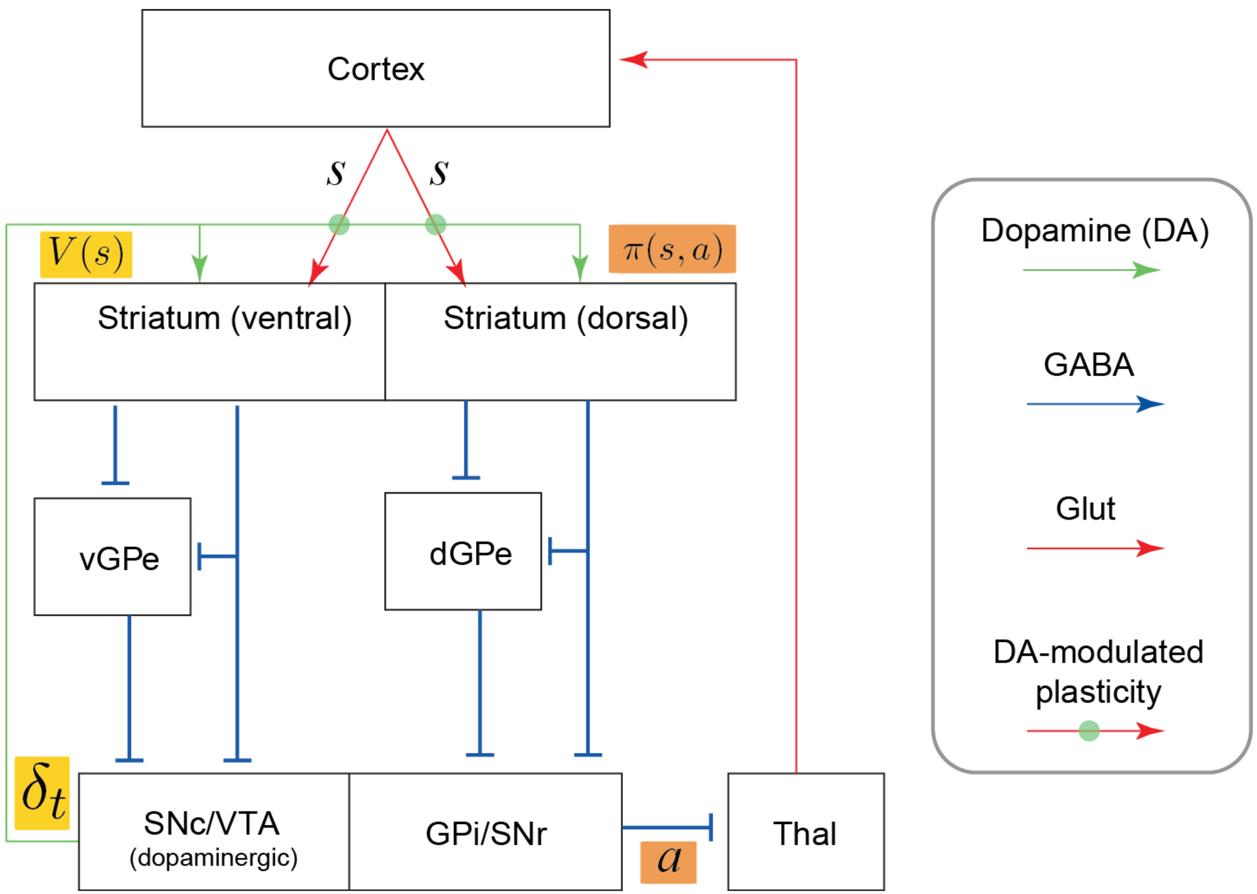
Made by BodyParts3D, Integrated Database Project, MEXT, Japan

Basal Ganglia and Reward-Based Learning

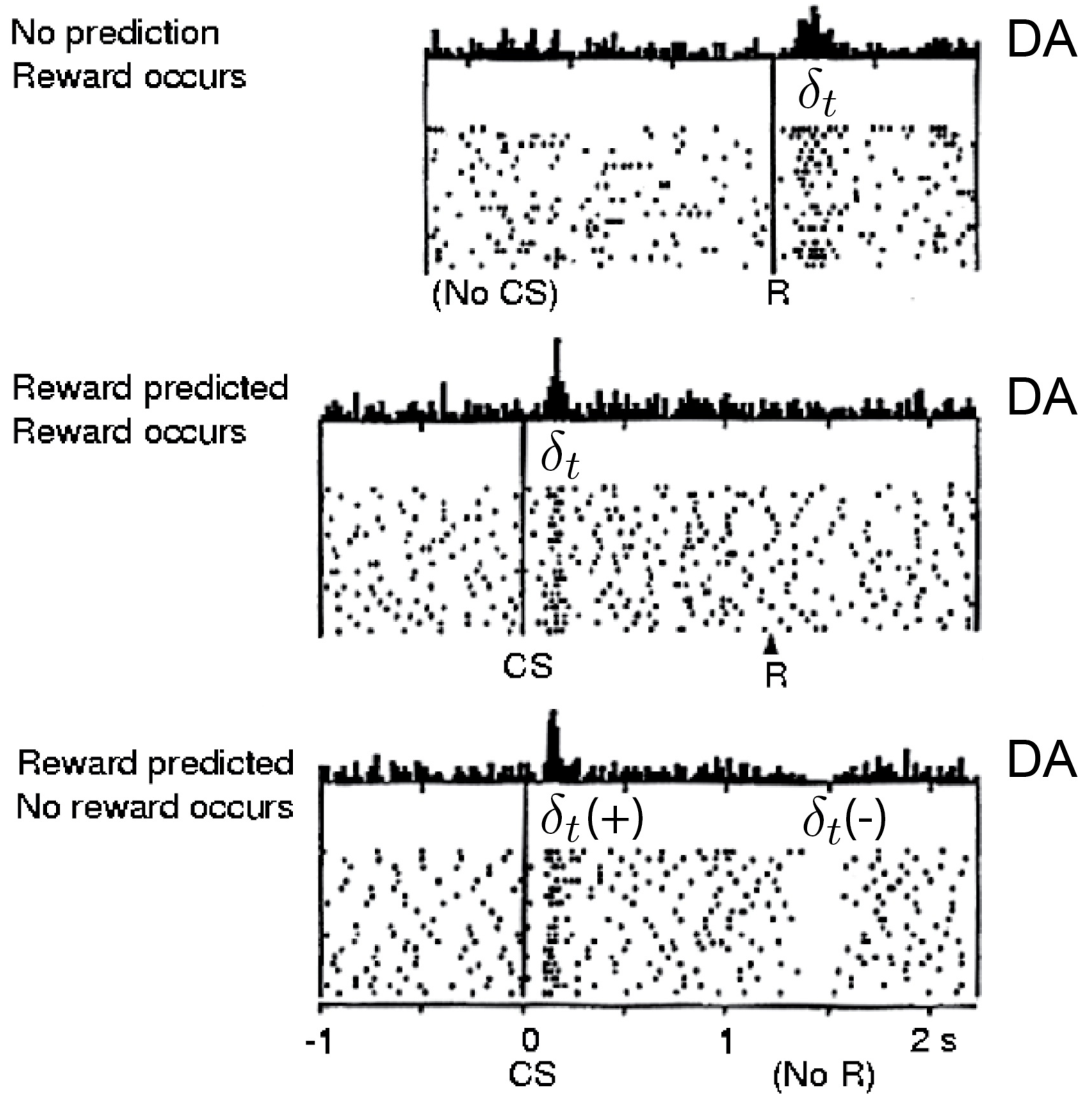


Purves et al, Neuroscience, 3ed, 2003

Basal Ganglia and Reward-Based Learning

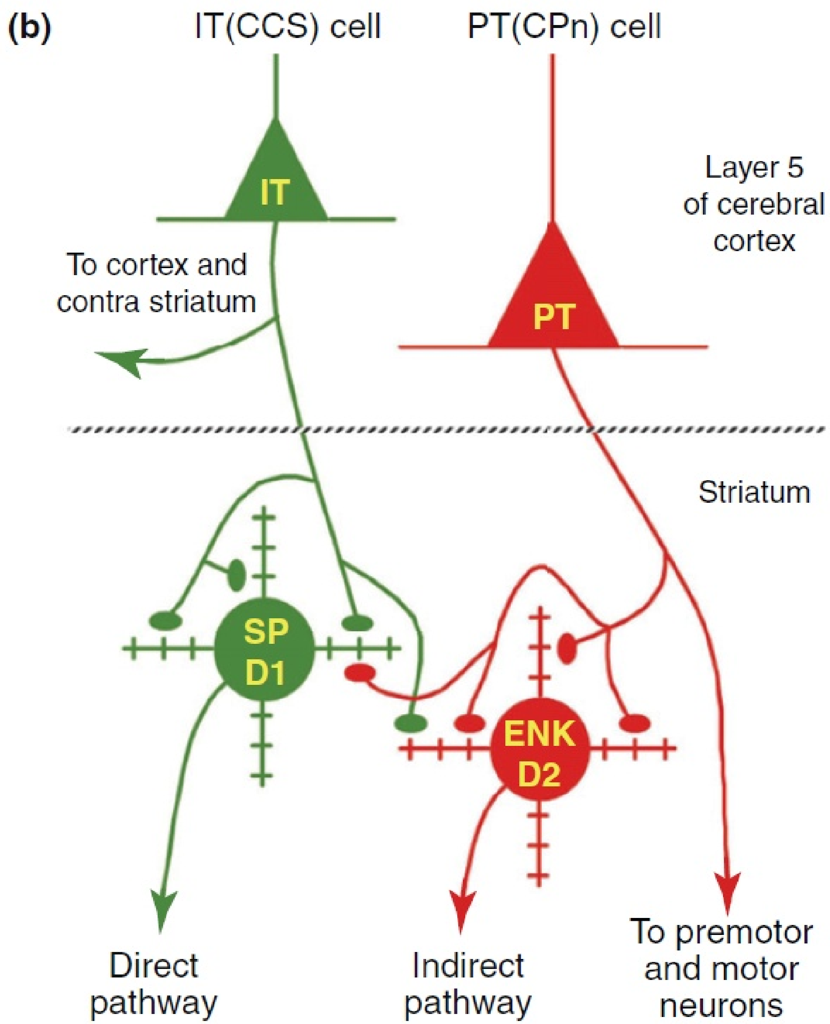


Dopamine and Reward-Based Learning



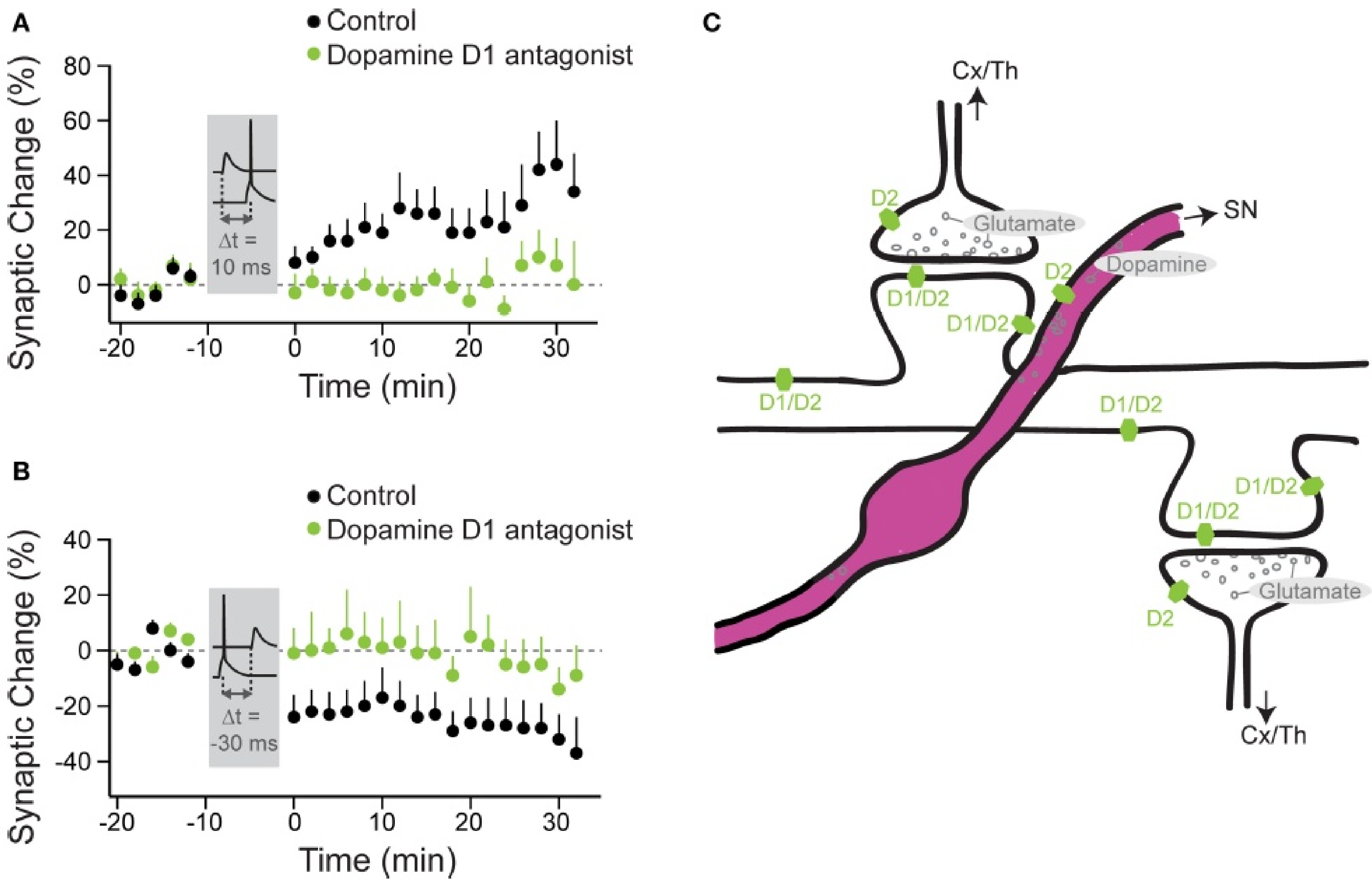
Schulz et al, 1997

Basal Ganglia and Reward-Based Learning



Lei et al., 2004; Morishima & Kawaguchi, 2006; Reiner et al., 2010

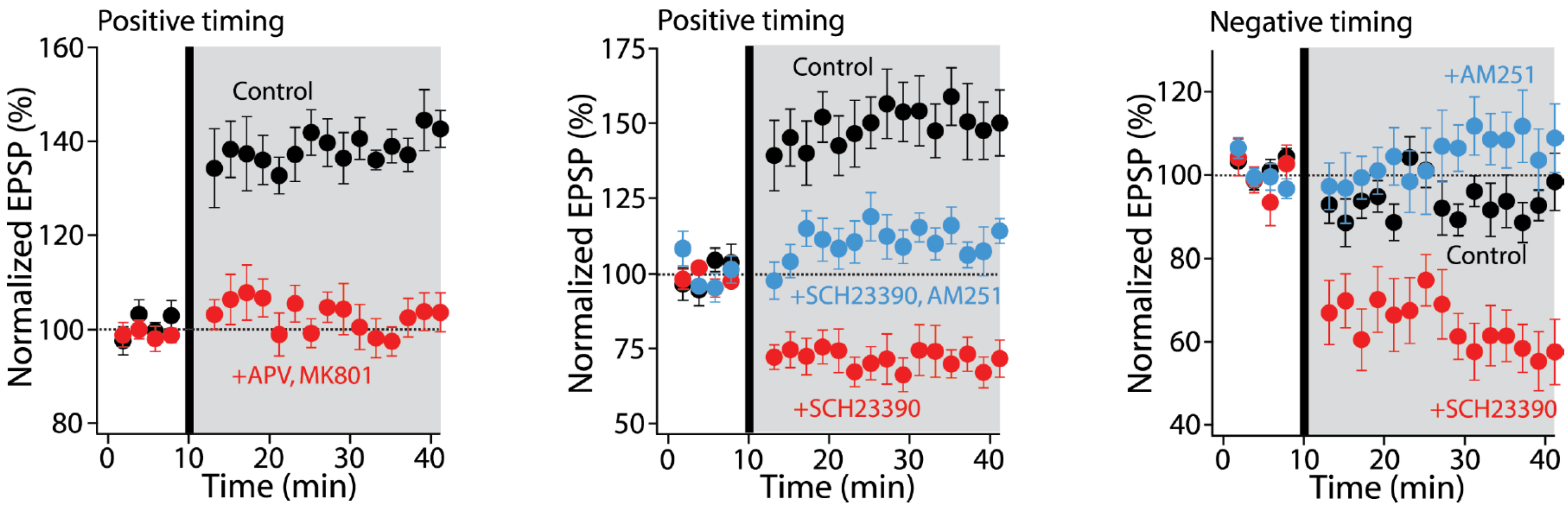
Basal Ganglia and Reward-Based Learning



Pawlak et al,, 2010

Basal Ganglia and Reward-Based Learning

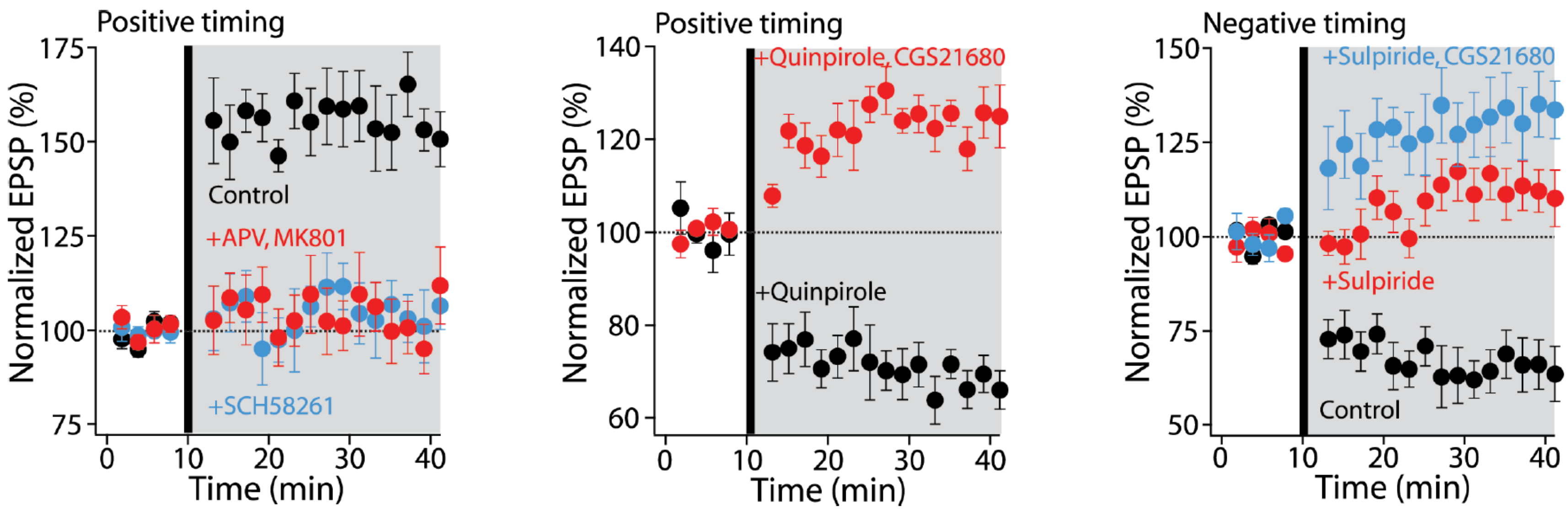
Bidirectional plasticity on D1 MSN synapses.
(DA low : LTD; DA high : LTP)



Shen et al,, Science, 2008

Basal Ganglia and Reward-Based Learning

Bidirectional plasticity on D2 MSN synapses.
(DA low : LTP; DA high : LTD)



Shen et al,, Science, 2008

Neural Substrate of Learning : Plasticity

Learning driven by the prediction error signal
→ Hypothesis : Dopamine represents prediction error and modulates plasticity

- **Computing prediction error (difference between internal expectation and the following observation)**

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$$

$\delta_t > 0$: better than expected

$\delta_t < 0$: worse than expected

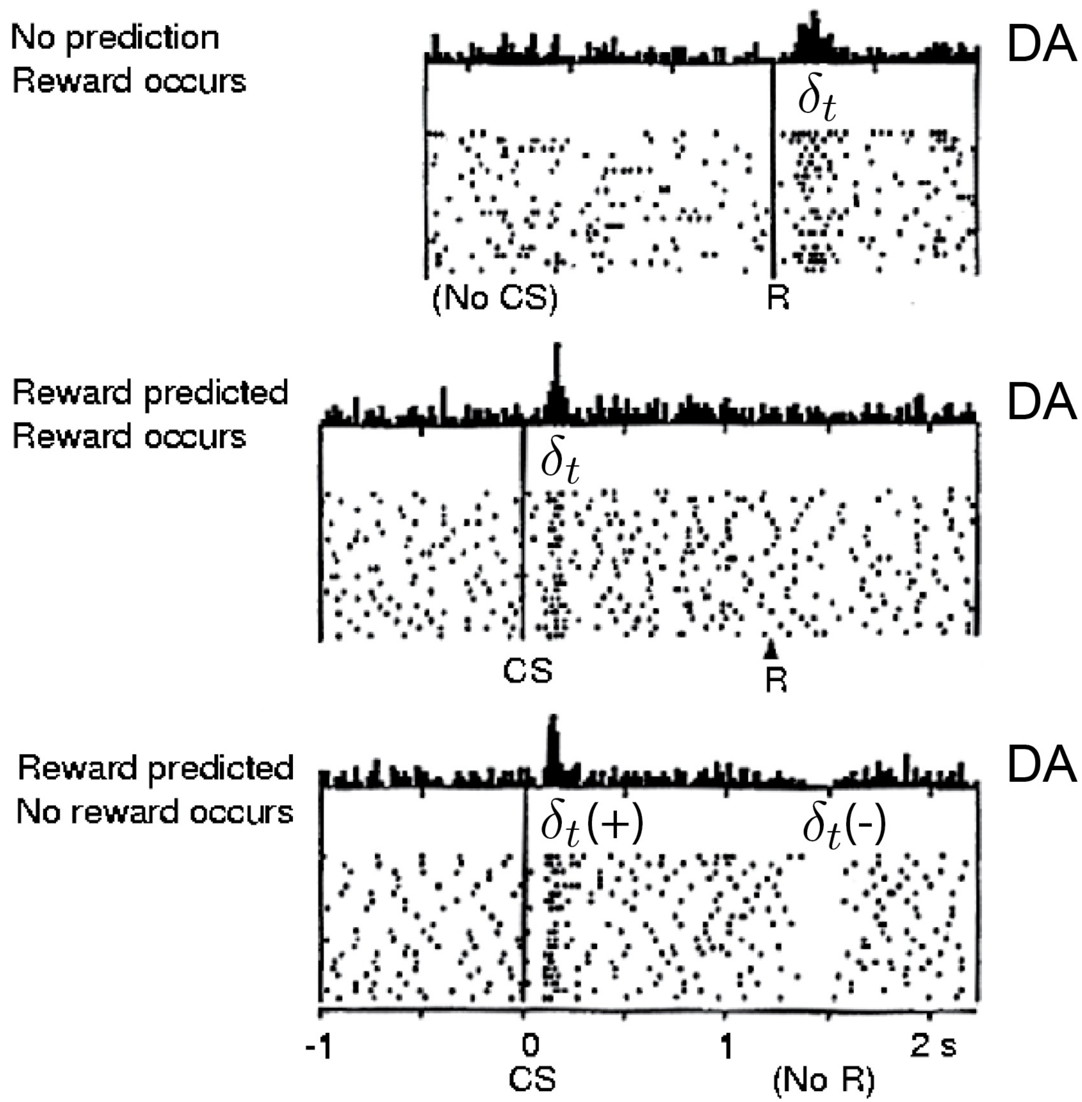
- **Updating expectations and action preferences**

$$V(s_t) \leftarrow V(s_t) + \alpha\delta_t$$

$$\pi(s,a) \leftarrow \pi(s,a) + \varepsilon\delta_t$$

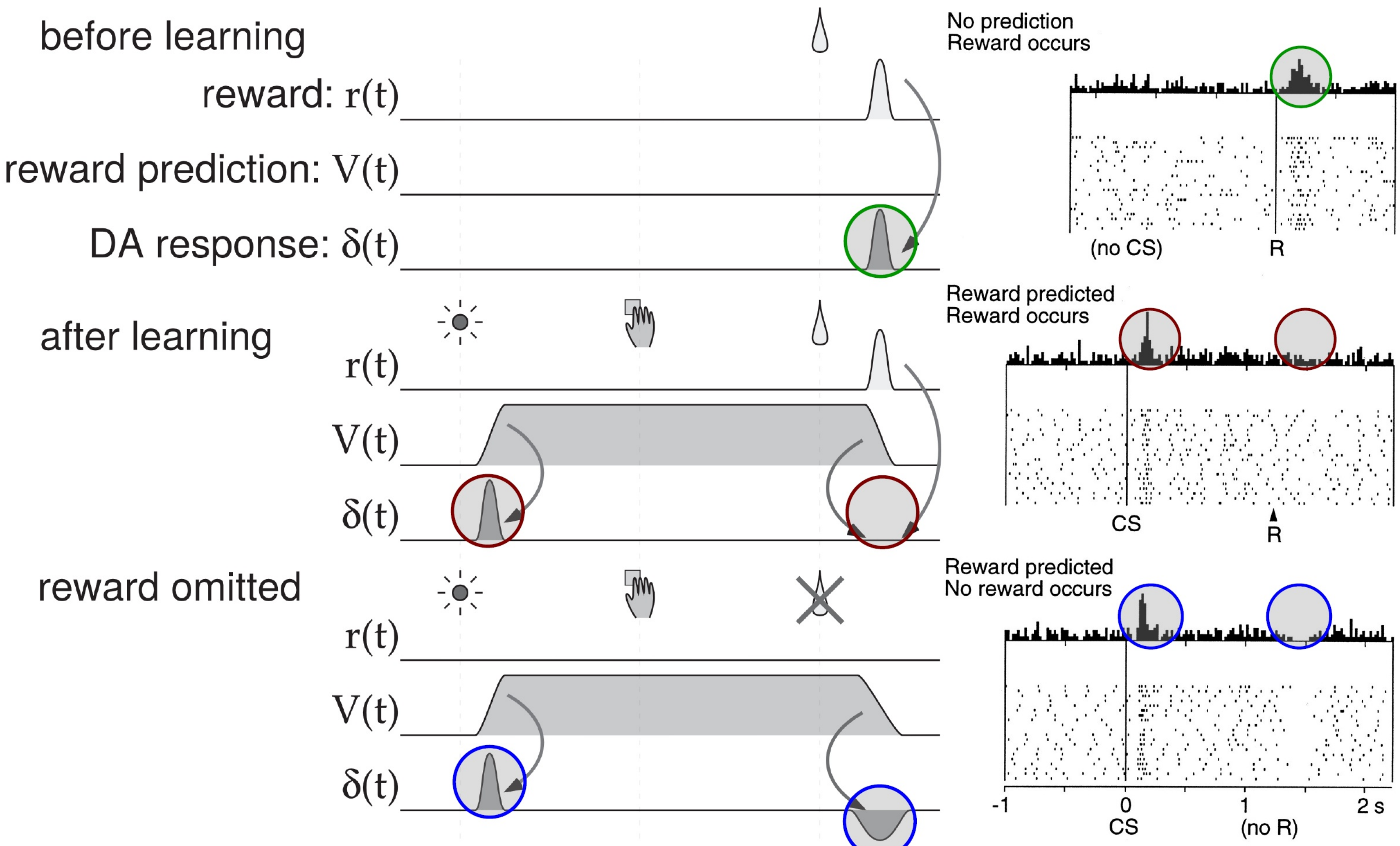
Barto et al, 1983; Niv et al, 2006

Dopamine and Reward-Based Learning



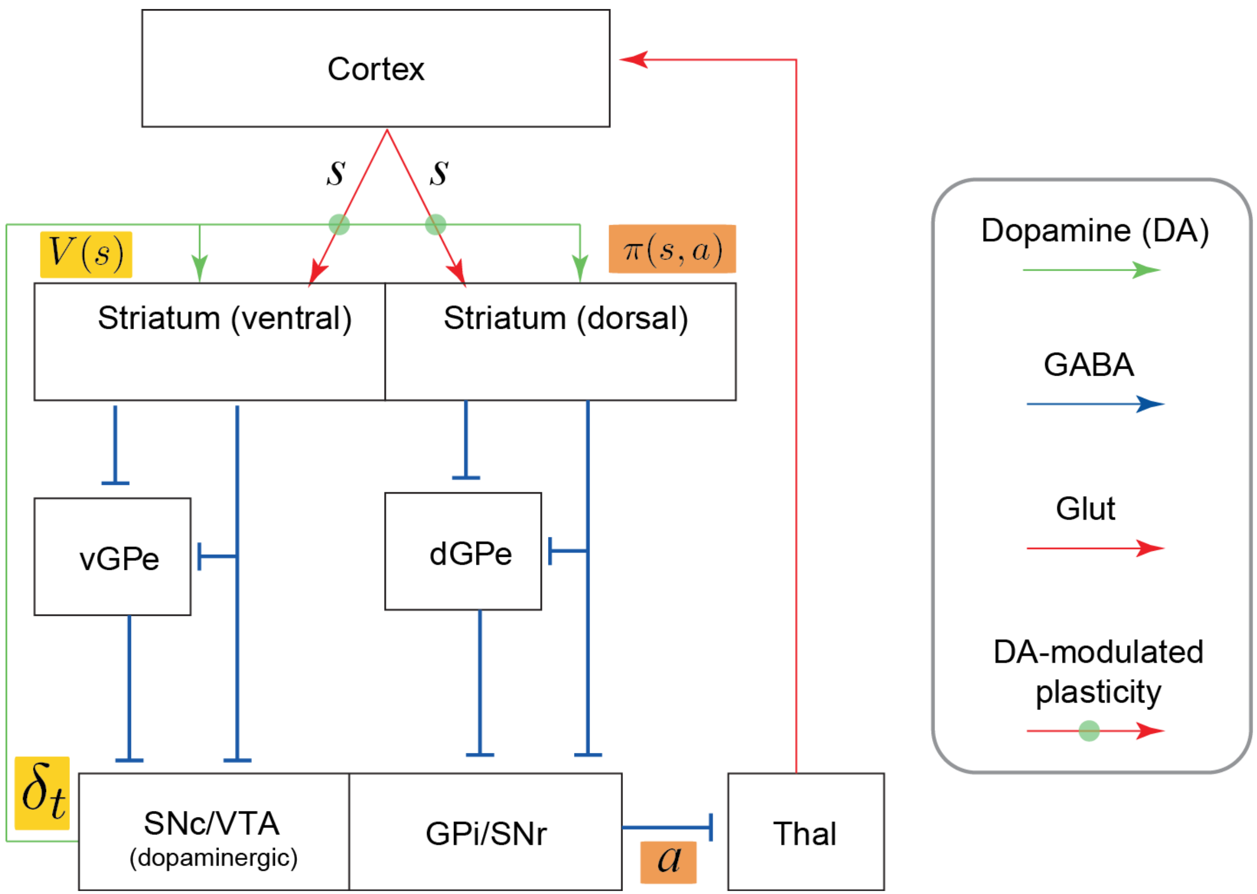
Schulz et al, 1997

Basal Ganglia and Reward-Based Learning

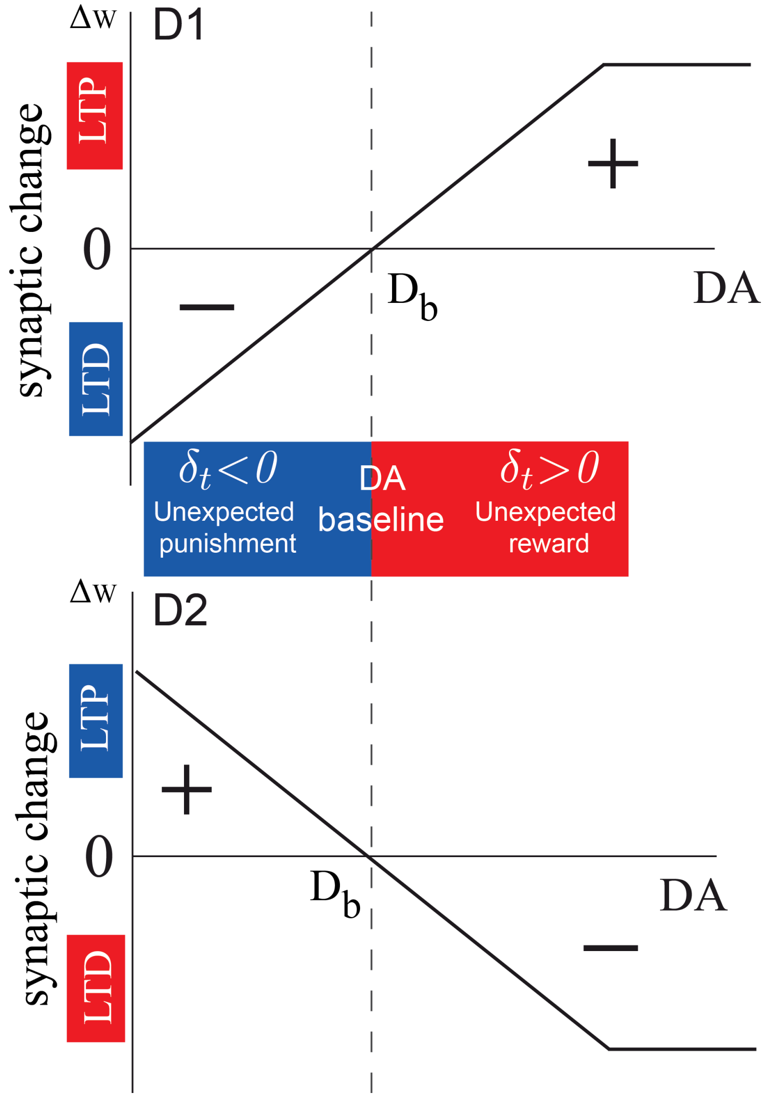


Doya, 2002

Basal Ganglia and Reward-Based Learning

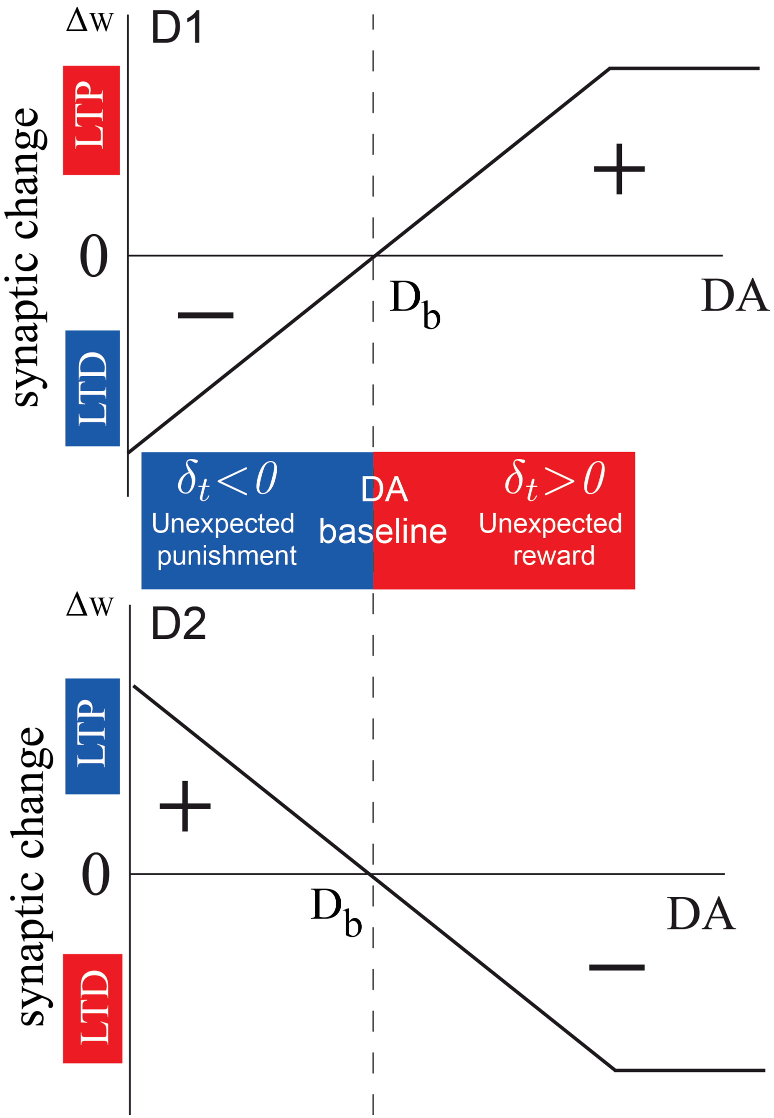


Basal Ganglia and Reward-Based Learning



Jitsev et al,, 2012

Basal Ganglia and Reward-Based Learning

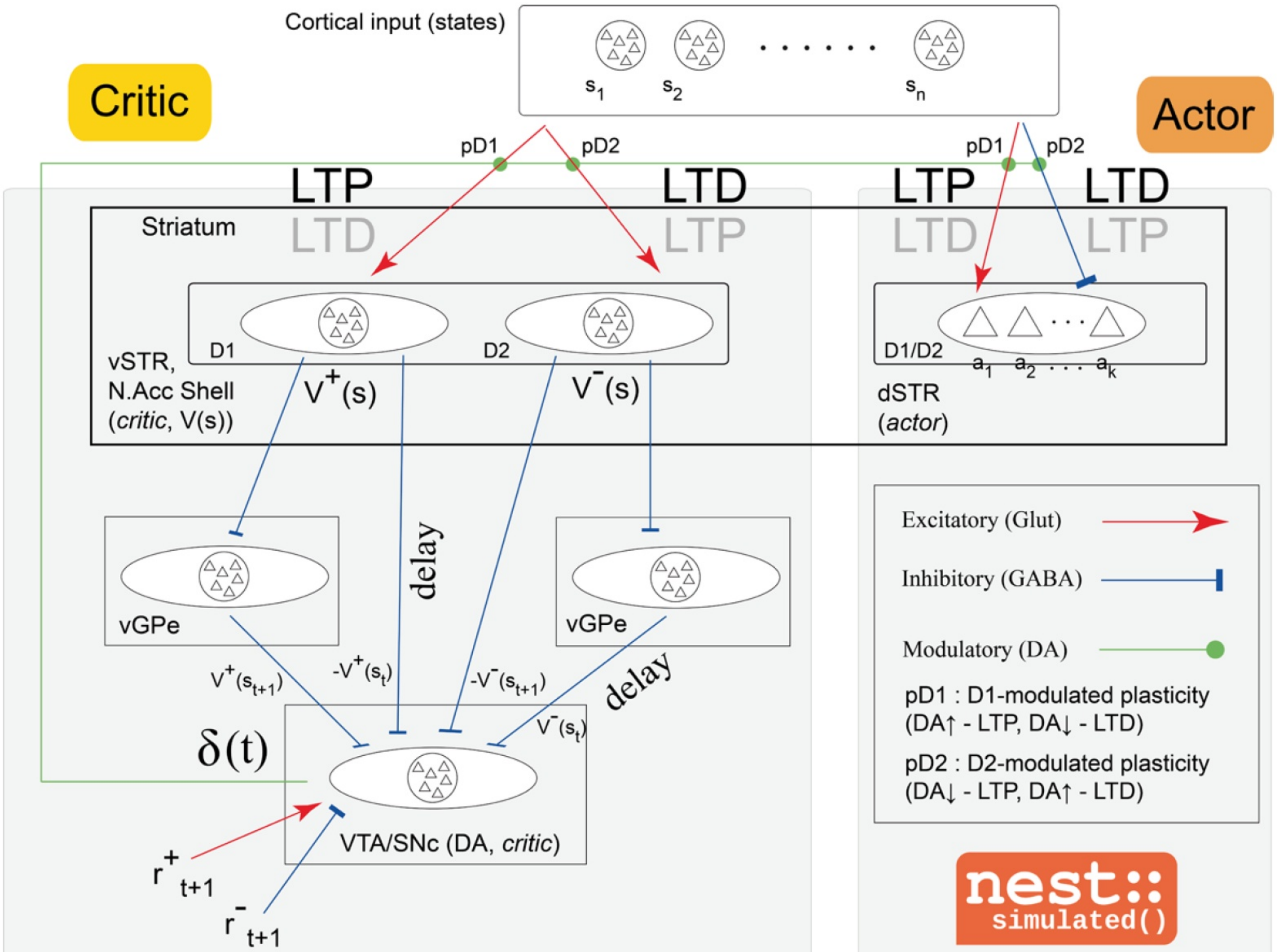


$$\dot{w}_{ij}^{D1} = A\,K_i^-(t)\,\varepsilon_i(t)\,K_j^+(t)\,(D(t) - D_b)$$

$$\dot{w}_{ij}^{D2} = A\,K_i^-(t)\,\varepsilon_i(t)\,K_j^+(t)\,(D(t) - D_b)$$

Jitsev et al,, 2012

A functional model of reinforcement learning



Jitsev et al,, 2012

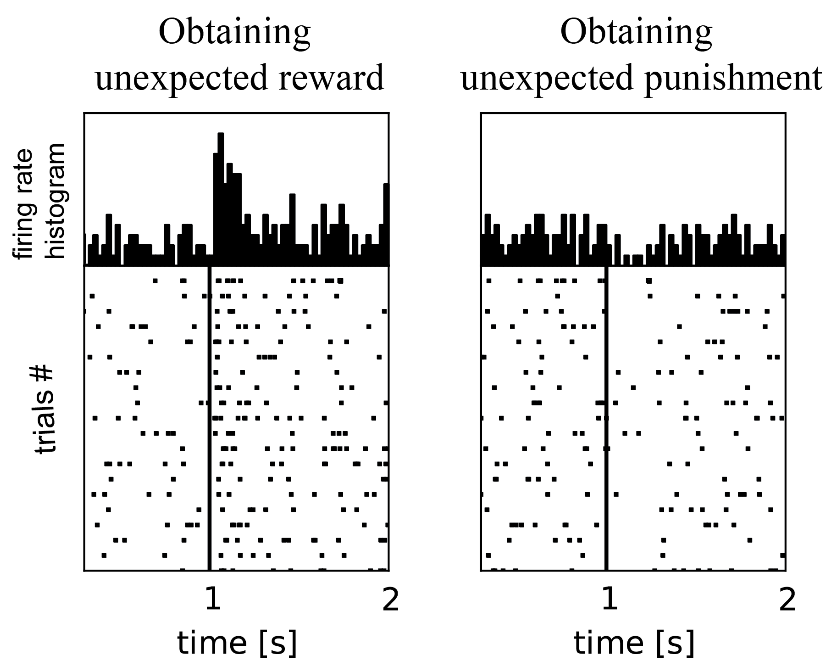
A functional model of reinforcement learning

Spiking neural network circuitry model

$$\dot{\mathbf{V}}_{\mathbf{m}} = -\frac{1}{\tau_m}\mathbf{V}_{\mathbf{m}} + \frac{1}{C_m}I(t),$$

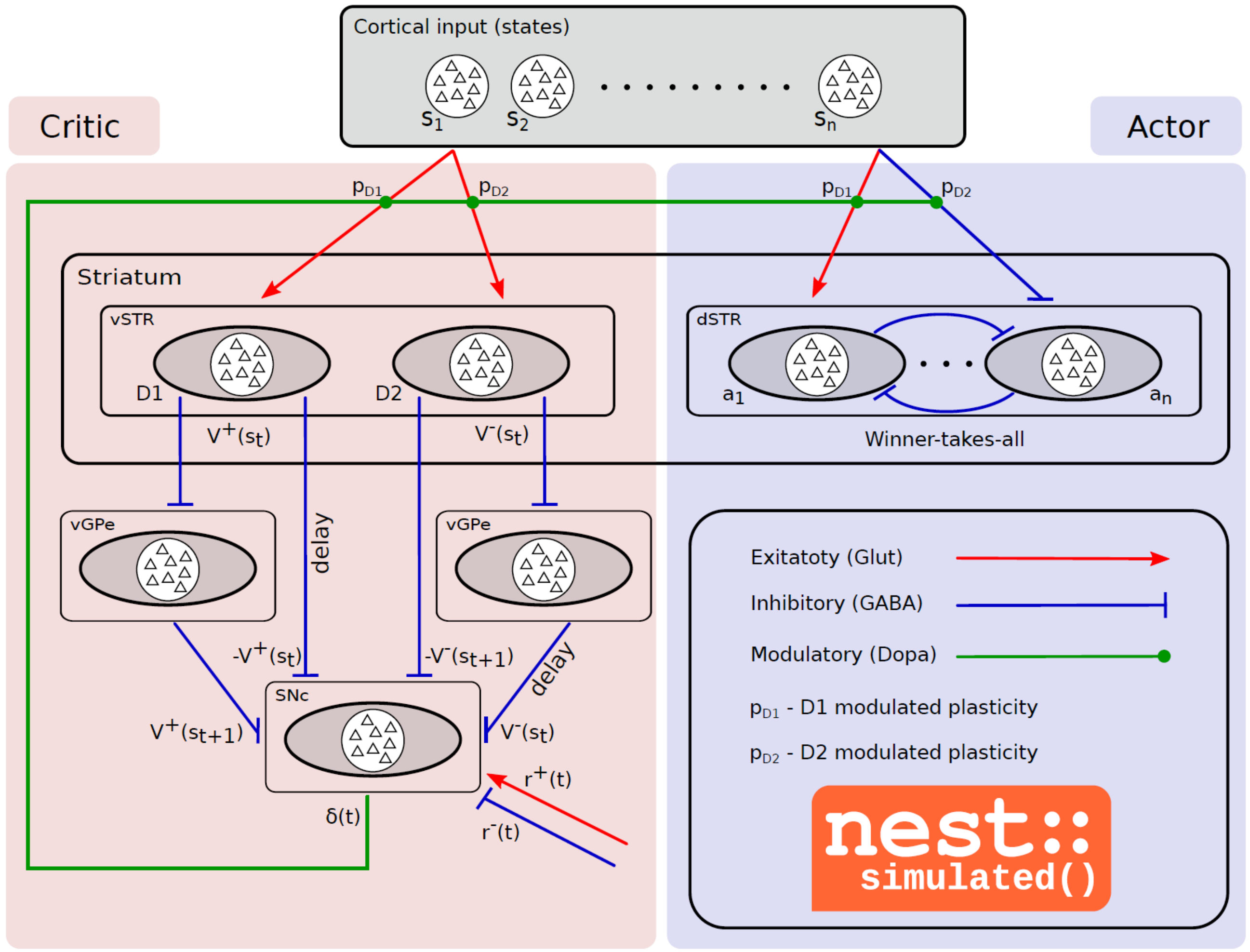
$$I_{syn}(t) = w\frac{e}{\tau_{syn}}te^{-t/\tau_{syn}}$$

DA activity (SNc/VTA, critic)
as prediction error signal



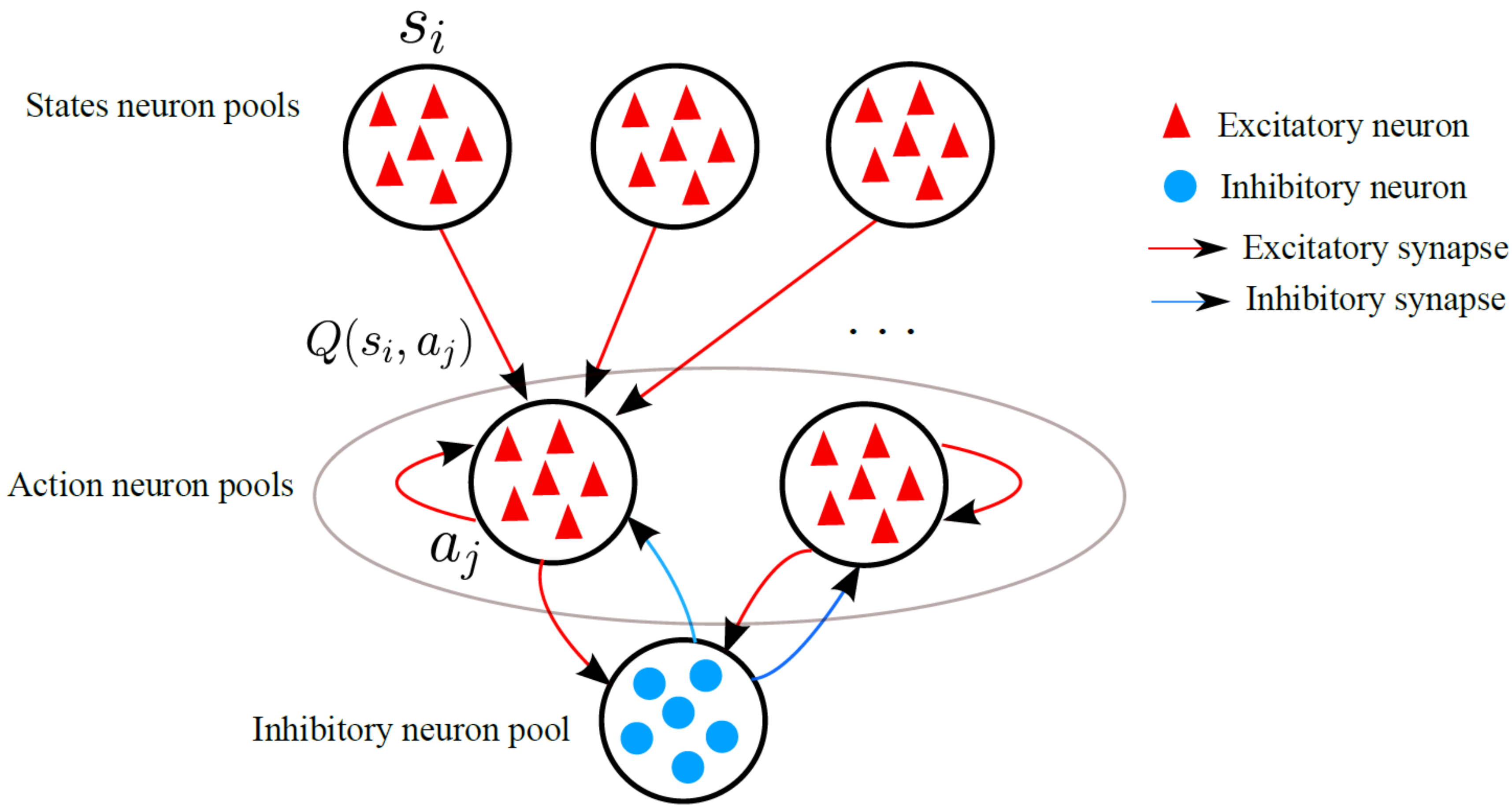
A functional model of reinforcement learning

Winner-take-all (WTA) circuitry used in actor part for action-selection

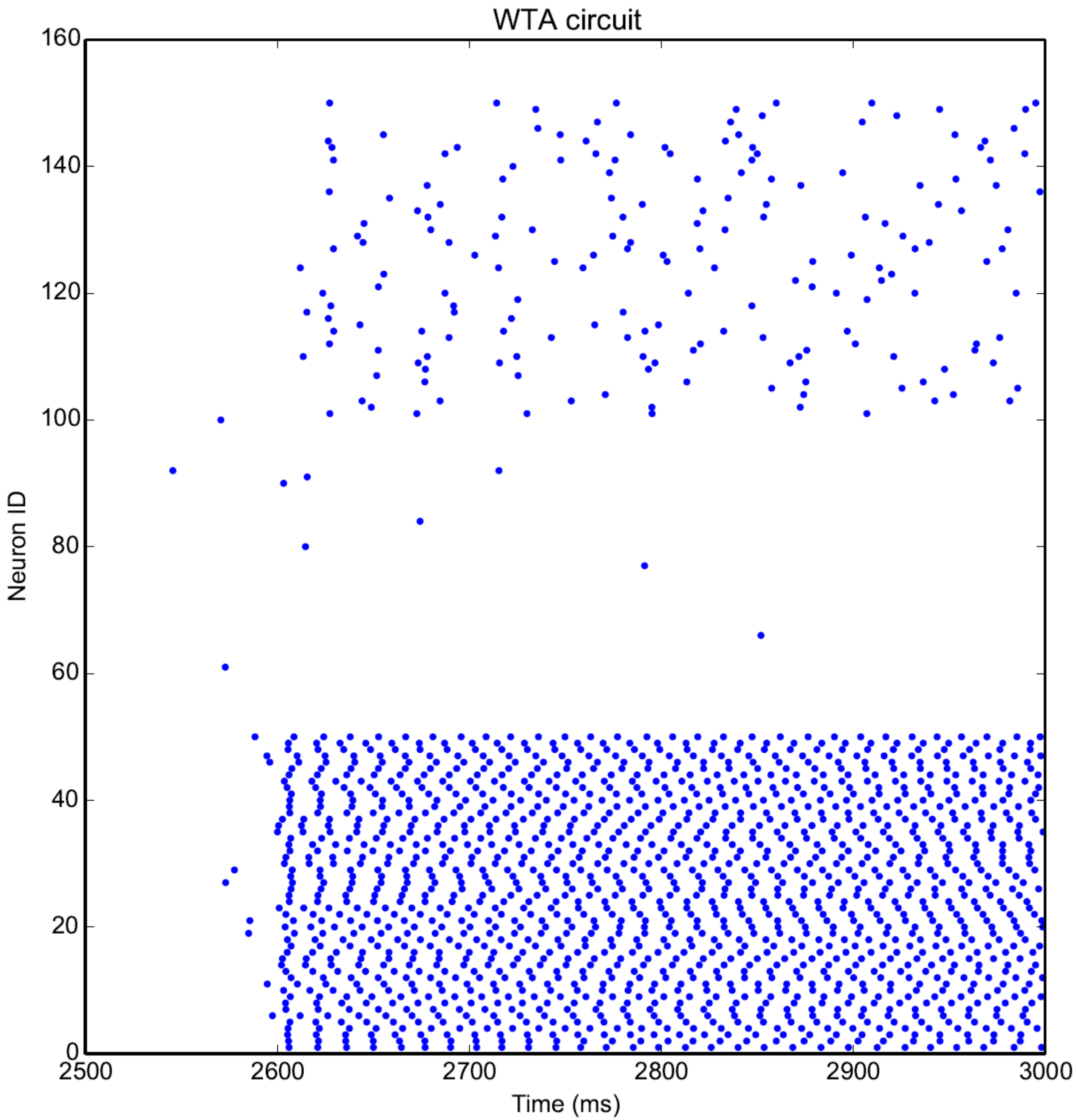


A functional model of reinforcement learning

Winner-take-all (WTA) circuitry used in actor part for action-selection

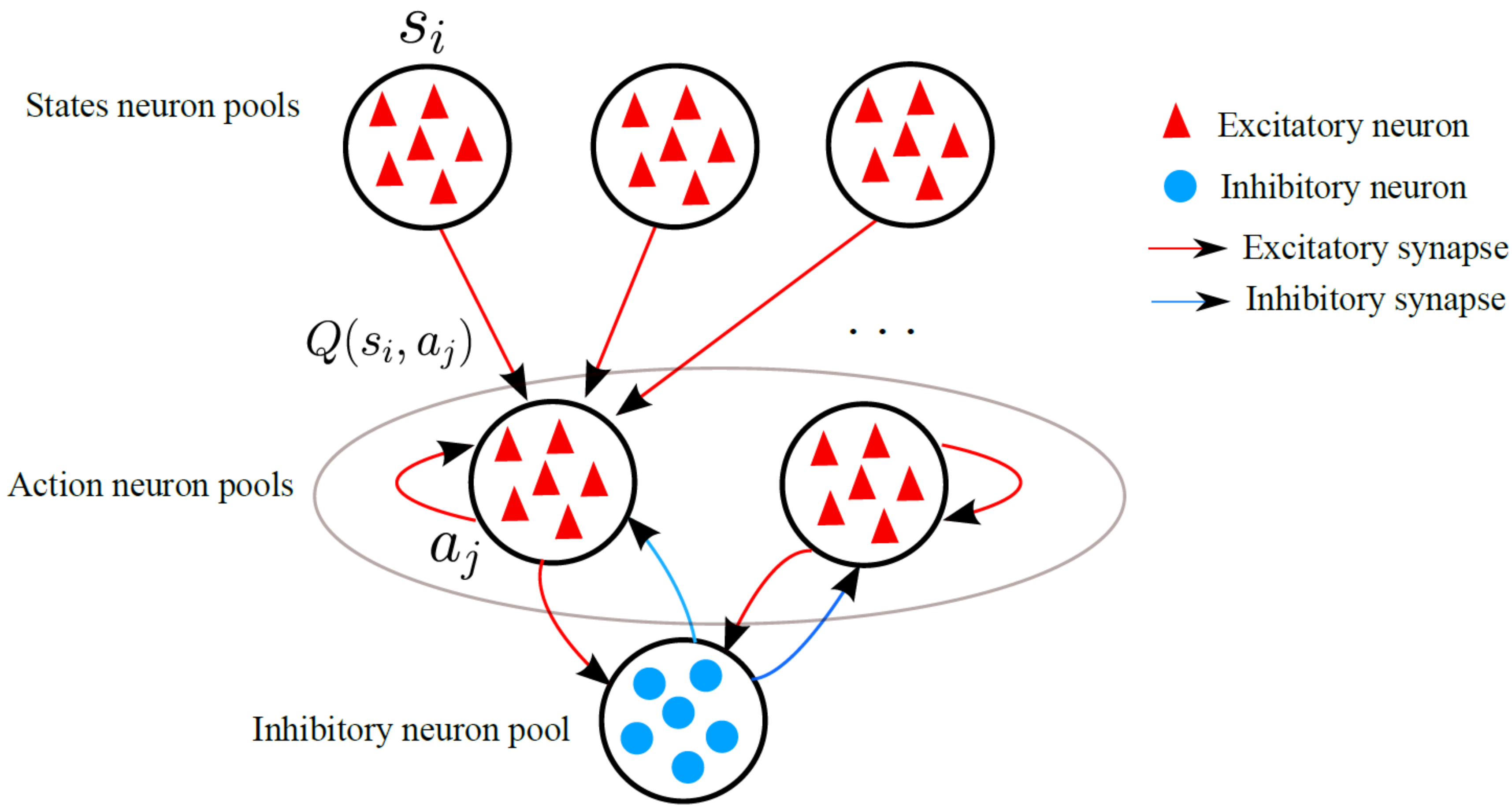


A functional model of reinforcement learning

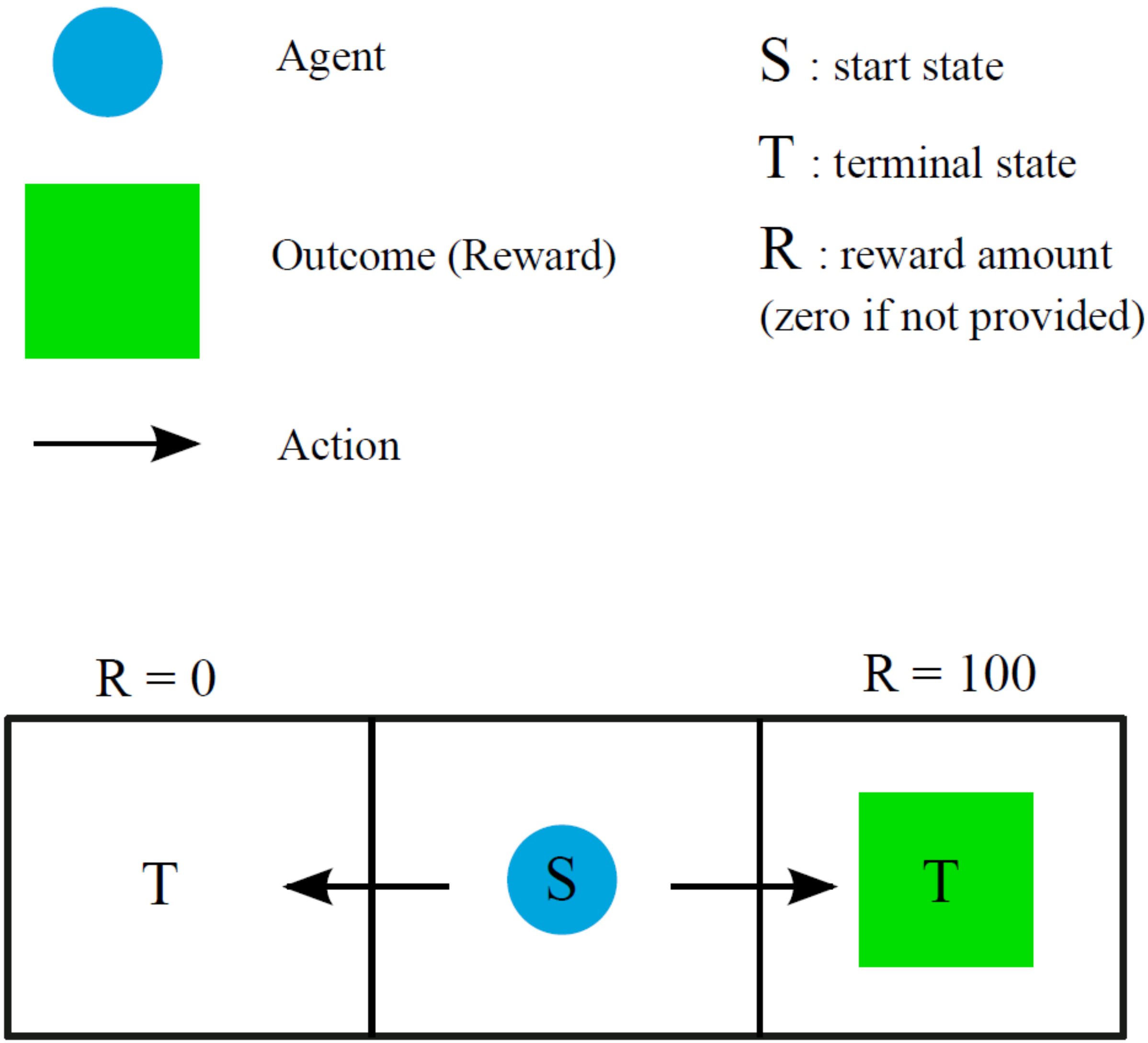


A functional model of reinforcement learning

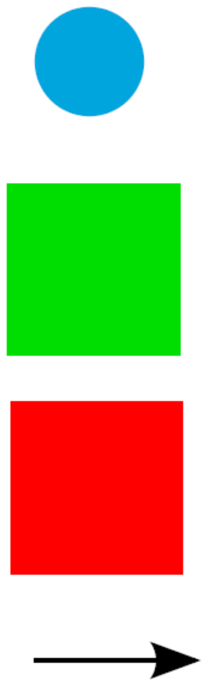
Winner-take-all (WTA) circuitry used in actor part for action-selection



A functional model of reinforcement learning



A functional model of reinforcement learning



Agent

Outcome (Reward)

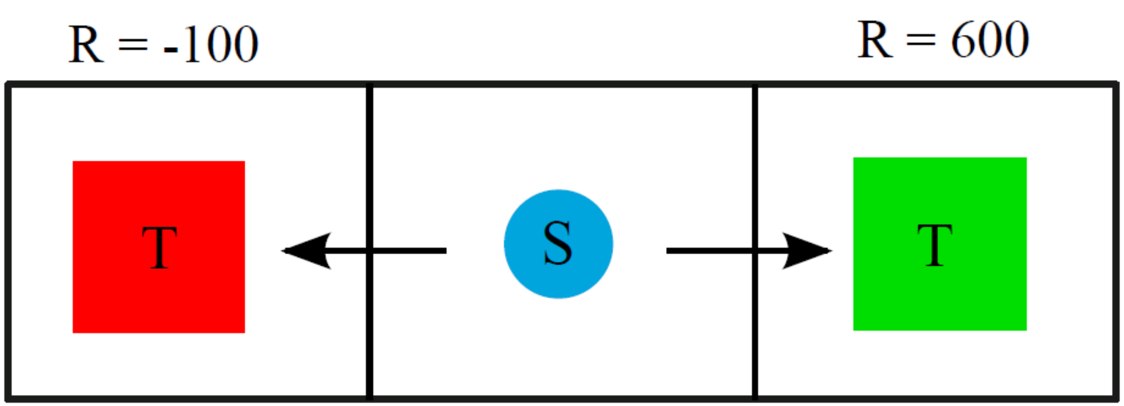
Outcome (Punishment)

Action

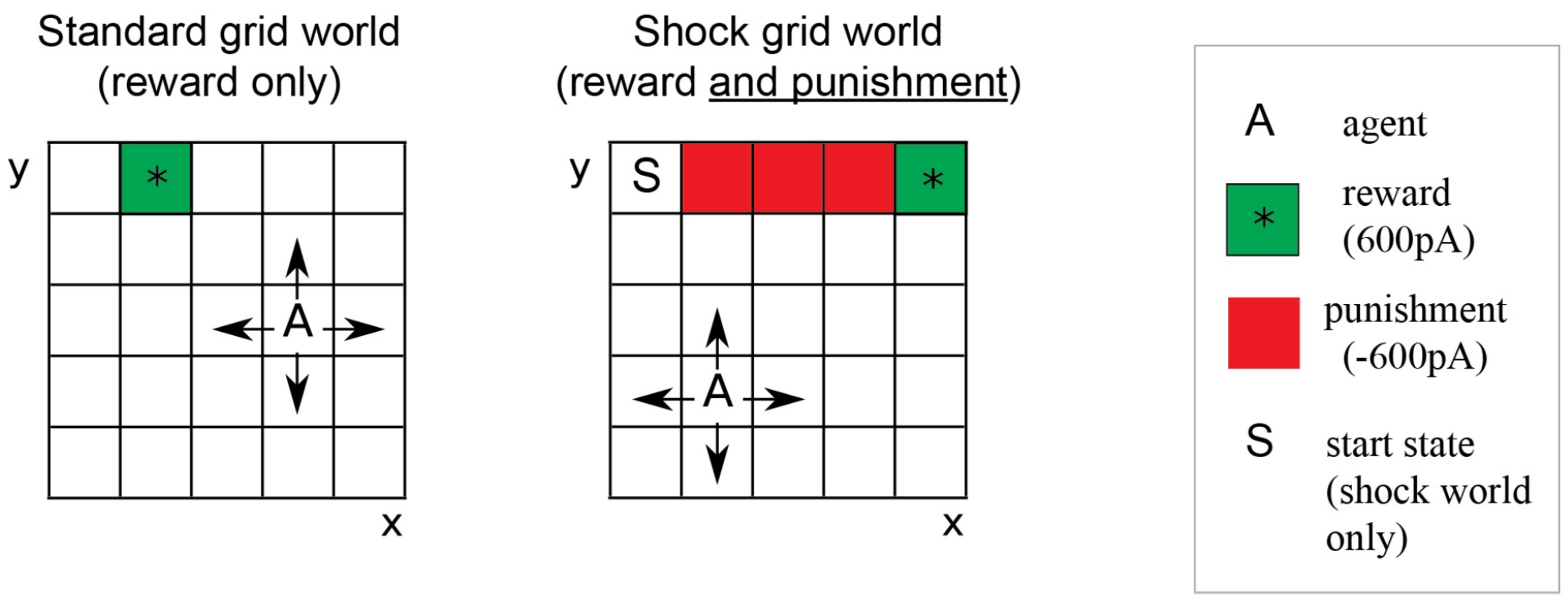
S : start state

T : terminal state

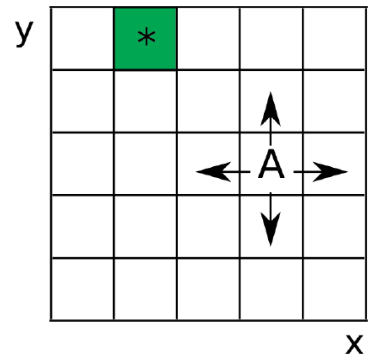
R : reward amount (zero if not provided, rewarding if positive, punishing if negative)

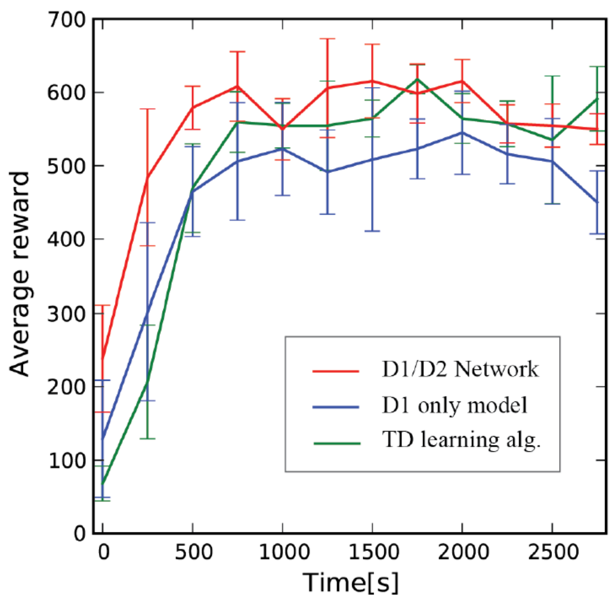


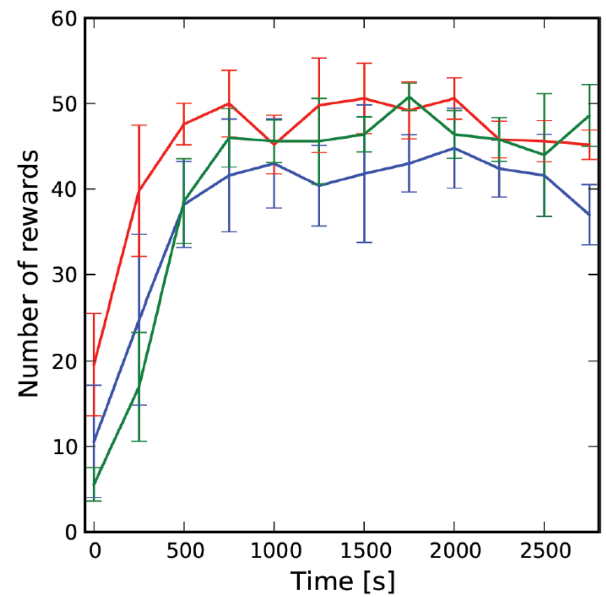
A functional model of reinforcement learning



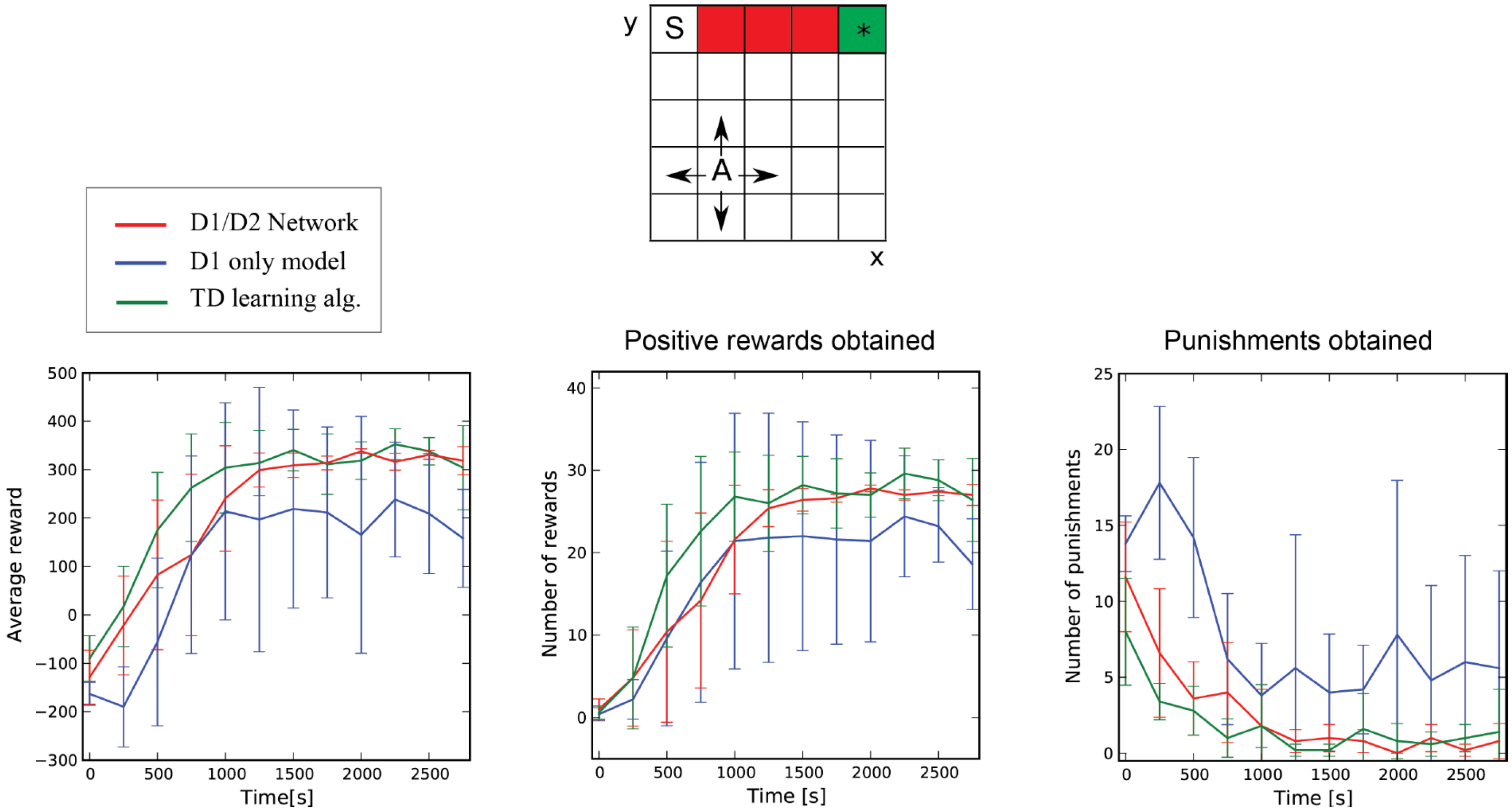
A functional model of reinforcement learning



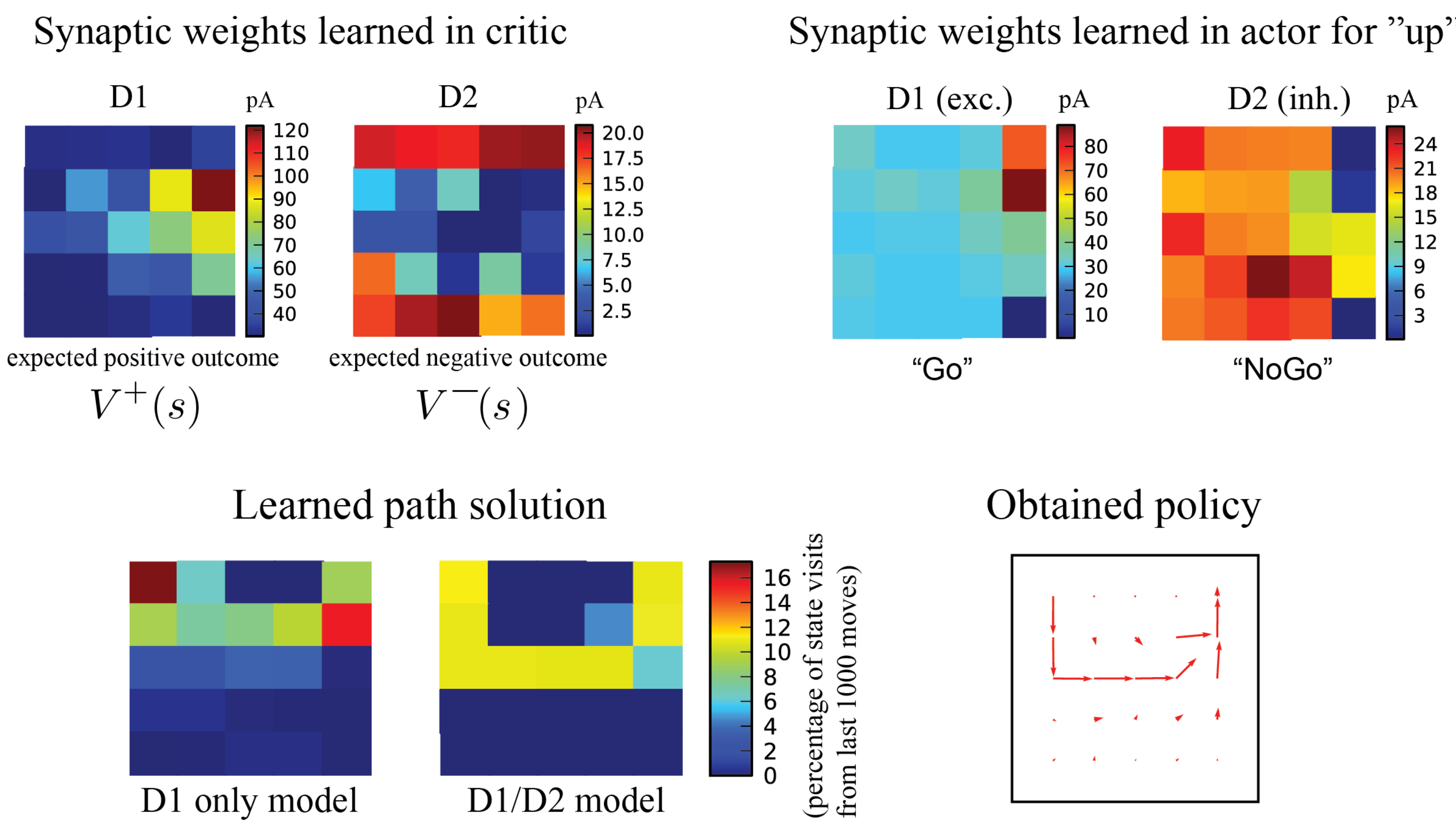




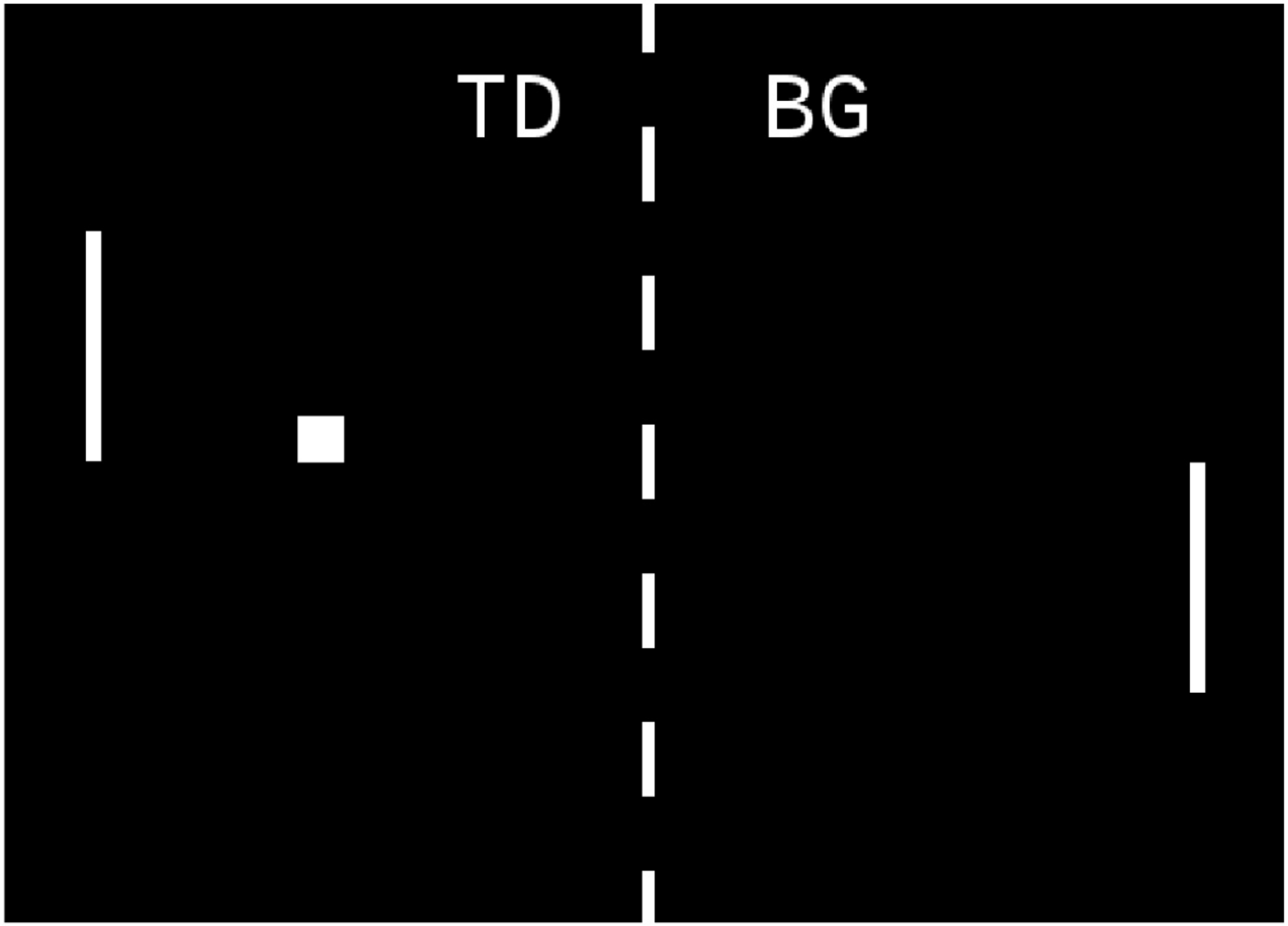
A functional model of reinforcement learning



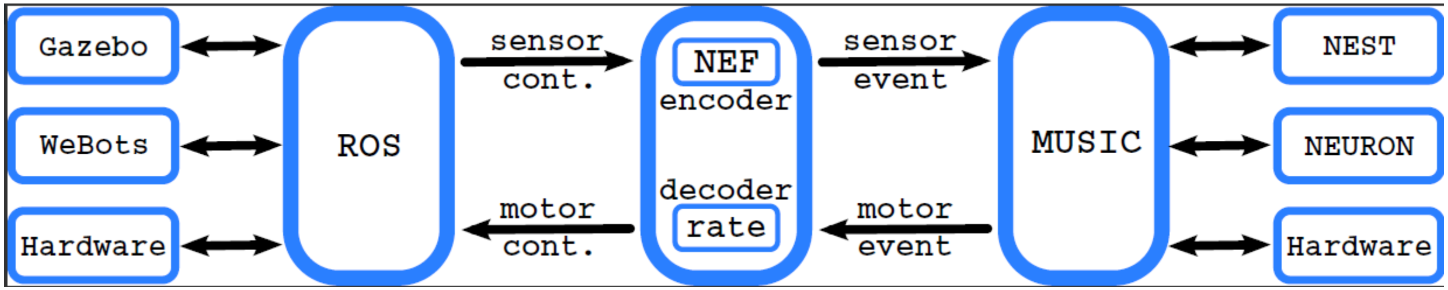
A functional model of reinforcement learning

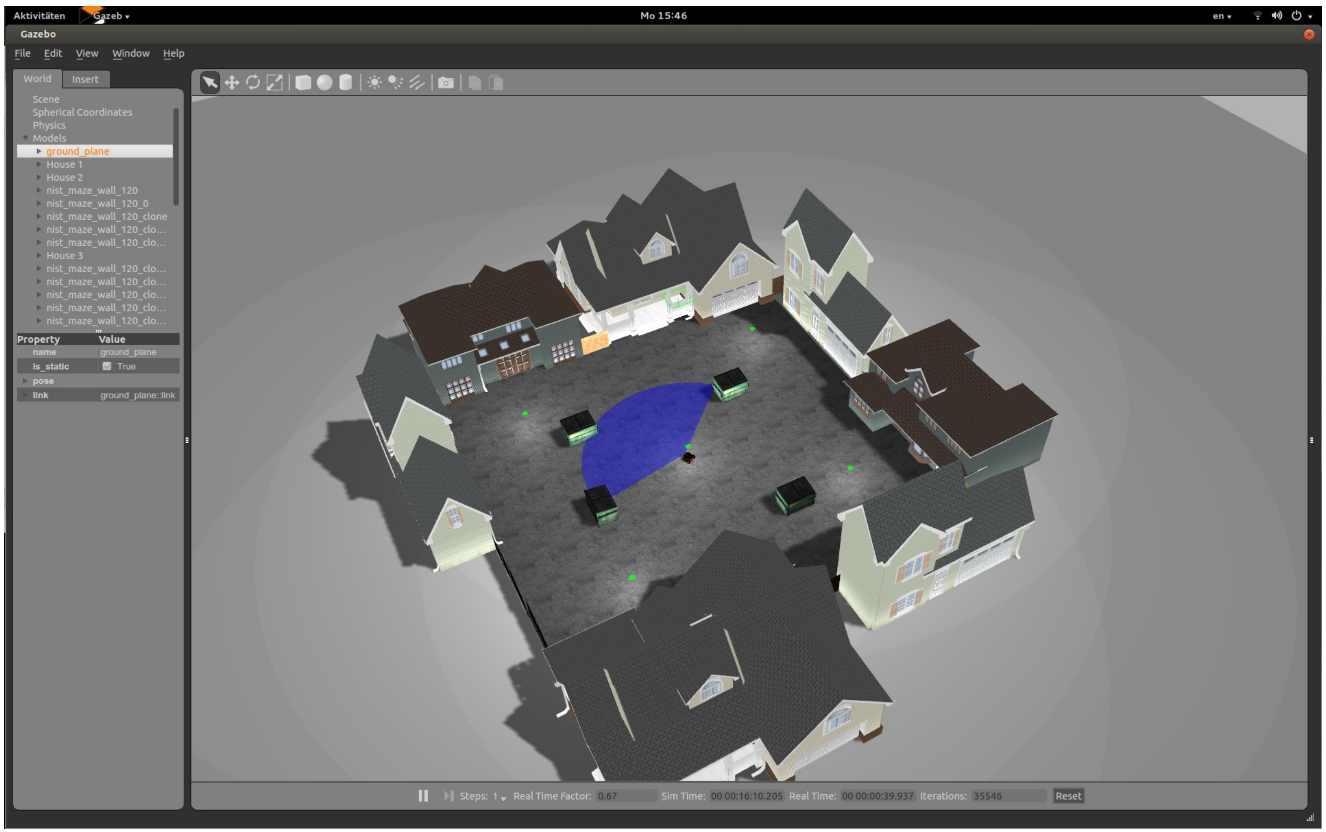


A functional model of reinforcement learning



A functional model of reinforcement learning





Weidel et al., CNS 2015

A functional model of reinforcement learning

People involved

- **Kenji Morita, University of Tokyo (Time and Reward)**
- **Boris Gutkin, ESN & CNRS, Paris (Nature of Reward Signal)**
- **Philipp Weidel, Susanne Kunkel, Abigail Morrison (Plasticity, Learning, Representation, Neural Coding)**

Adaptation and Learning in Biology

- The core feature of adaptivity is updating internal model (expectations, beliefs) in face of experience (consequences) → <u>Learning</u>

- Experience (sensory input) is able to change organization of brain's neural networks → <u>Plasticity</u>

- Plasticity implements learning in yet unknown way

- Figuring it out would mean a breakthrough both in <u>understanding the brain</u> and in <u>creating new technology</u>

Adaptation and Learning in Biology

Just do it.

Thanks
for
your
Attention