# Motor Control and Reinforcement Learning

Computational Cognitive Neuroscience

Randall O'Reilly

## Learning Rules Across the Brain

| Area | Learning Signal: Reward | Error | Self Org | Separator | Dynamics: Integrator | Attractor |
|---|---|---|---|---|---|---|
| *Primitive* Basal Ganglia | +++ | - - - | - - - | ++ | - | - - - |
| Cerebellum | - - - | +++ | - - - | +++ | - - - | - - - |
| *Advanced* Hippocampus | + | + | +++ | +++ | - - - | +++ |
| Neocortex | ++ | +++ | ++ | - - - | +++ | +++ |

\+ = has to some extent … +++ = defining characteristic – definitely has
\- = not likely to have … - - - = definitely does not have

2

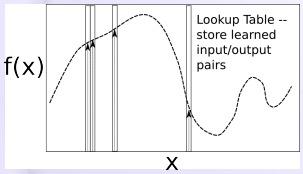
## Primitive, Basic Learning..

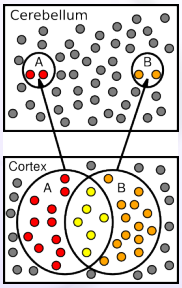
| Area | Learning Signal: Reward | Error | Self Org | Separator | Dynamics: Integrator | Attractor |
|---|---|---|---|---|---|---|
| *Primitive* Basal Ganglia | +++ | - - - | - - - | ++ | - | - - - |
| Cerebellum | - - - | +++ | - - - | +++ | - - - | - - - |

- Reward & Error = most basic learning signals (self organized learning is a luxury..)
- Simplest general solution to any learning problem is a *lookup table* = separator

3

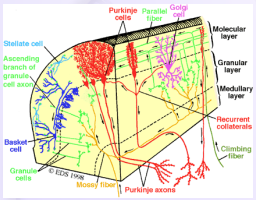
## Lookup Table & Pattern Separation

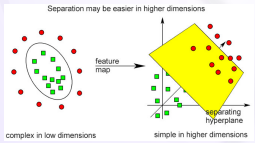



4

## Cerebellar Error-driven Learning




Cerebellum = Support Vector Machine

- Granule cells = high-dimensional encoding (separation)
- Purkinje/Olive = delta-rule error-driven learning
- Classic ideas from Marr (1969) & Albus (1971)

5

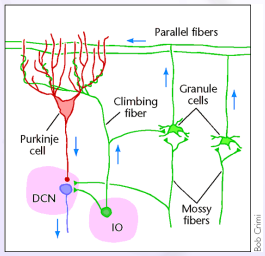
## Cerebellum is Feed Forward

Feedforward circuit:

- Input (PN) -> granules -> Purkinje -> Output (DCN)
- Inhibitory interactions – no attractor dynamics
- Key idea: does delta-rule learning bridging small temporal gap:

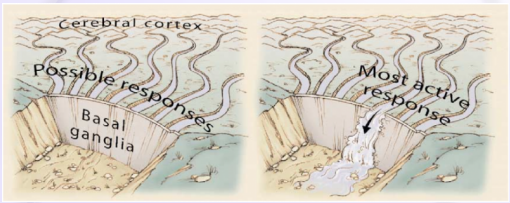
  S(t-100) -> R(t)

  ^ Error(t+100)



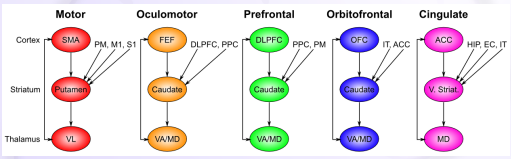

6

## Basal Ganglia and Action Selection




7

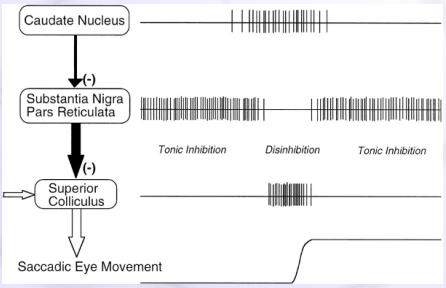
## Basal Ganglia: Action Selection




- Selects motor and "cognitive" actions across frontal areas

8

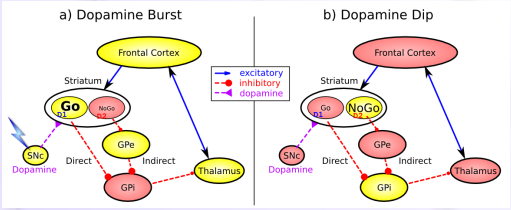
## Release from Inhibition




9

## Basal Ganglia Reward Learning

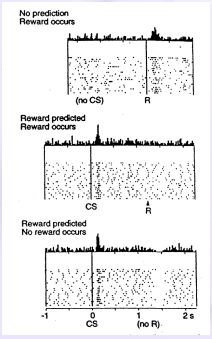
(Frank, 2005…; O'Reilly & Frank 2006)




- Feedforward, modulatory (disinhibition) on cortex/motor (same as cerebellum)
- Co-opted for higher level cognitive control -> PFC

10

## Reinforcement Learning: Dopamine




Rescorla-Wagner / Delta Rule:

- $\delta = r - \hat{r}$
- $\delta = r - \sum xw$

But no CS-onset firing – need to Anticipate the future!

- $\delta = (r + f) - \hat{r}$

CS-onset = future reward = f

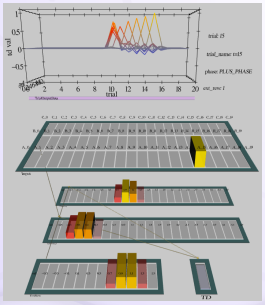
11

## Temporal Differences Learning

- $V(t) = r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2)...$
- $\hat{V}(t) = r(t) + \gamma \hat{V}(t+1)$
- $0 = \left(r(t) + \hat{V}(t+1)\right) - \hat{V}(t)$
- $\delta = \left(r(t) + \hat{V}(t+1)\right) - \hat{V}(t)$
- $f = \gamma \hat{V}(t+1)$ <- this is the future!

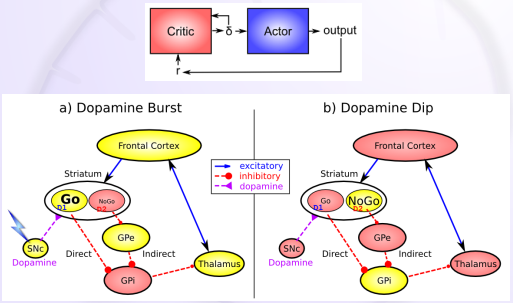
12

## Network Implementation



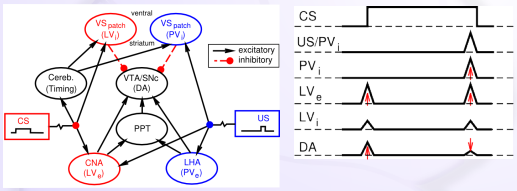
13

## Actor - Critic




14

## Biology of Dopamine




15

## BG + Cerebellum Capacities

- Learn what satisfies basic needs, and what to avoid (BG reward learning)
  - And what information to maintain in working memory (PFC) to support successful behavior
- Learn basic Sensory -> Motor mappings accurately (Cerebellum error-driven learning)
  - Sensory -> Sensory mappings? (what is going to happen next..)

16

## BG + Cerebellum Incapacities

- Generalize knowledge to novel situations
  - Lookup tables don't generalize well..
- Learn abstract semantics
  - Statistical regularities, higher-order categories, etc
- Encode episodic memories (specific events)
  - Useful for instance-based reasoning
- Plan, anticipate, simulate, etc..
  - Requires robust working memory

17