Motor Control and Reinforcement Learning

Computational Cognitive Neuroscience
Randall O’Reilly

Learning Rules Across the Brain

<table>
<thead>
<tr>
<th>Area</th>
<th>Learning Signal</th>
<th>Dynamics</th>
<th>Attractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primitive</td>
<td>Reward</td>
<td>Error</td>
<td>Self Org</td>
</tr>
<tr>
<td>Basal Ganglia</td>
<td>+++</td>
<td>-</td>
<td>- - -</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>- - -</td>
<td>+++</td>
<td>- - -</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>- - -</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Neocortex</td>
<td>++</td>
<td>-</td>
<td>- - -</td>
</tr>
</tbody>
</table>

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

Learning Signal Dynamics

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

Primitive, Basic Learning..

Cerebellar Error-driven Learning

- Granule cells = high-dimensional encoding (separation)
- Purkinje/Olive = delta-rule error-driven learning

Lookup Table & Pattern Separation

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:
  \[ S(t-100) \to R(t) \wedge \text{Error}(t+100) \]
Basal Ganglia and Action Selection

Basal Ganglia: Action Selection

- Selects motor and "cognitive" actions across frontal areas

Release from Inhibition

Basal Ganglia Reward Learning

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

Reinforcement Learning: Dopamine

Temporal Differences Learning

- $\delta = r(t) + \gamma V(t+1) - V(t)$
- $f = -\gamma V(t+1) \quad \leftarrow \text{this is the future!}$
Network Implementation

Actor - Critic

Biology of Dopamine

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?)

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