Synaptic theory of gradient learning with empiric inputs

Ila Fiete
Kavli Institute for Theoretical Physics

Types of learning

• Associations among events (and reinforcements).
  e.g., classical conditioning
• How one’s actions affect events (and reinforcements), how to shape actions.
  e.g., instrumental conditioning
### Outlined

- Learning with reinforcements as optimization
- Synaptic learning rule
- Scale-up problem
- Example: birdsong
- Implications with Sebastian Seung
Instrumental conditioning

![Image of Thorndike's Puzzle Box](image)

FIGURE 4.1 One of Thorndike's Puzzle Boxes. A cat could escape from this box by pulling a string, stepping on the platform, and turning one of the two latches on the front of the door. (From Thorndike, 1898.)

Learning by reinforcement

![Graph of Time to Escape](image)

Thorndike, 1898

Try multiple strategies; modify behavior in way that tends to improve reinforcement.
The law of effect

Of several responses made to the same situation those which are accompanied or closely followed by satisfaction to the animal will become more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections to the situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond.

Thorndike, 1910
Trial, reward, and optimization

reward = performance on task

<table>
<thead>
<tr>
<th>behavioral</th>
<th>neural</th>
</tr>
</thead>
<tbody>
<tr>
<td>classical conditioning</td>
<td>Hebbian learning</td>
</tr>
<tr>
<td>instrumental conditioning</td>
<td>?</td>
</tr>
</tbody>
</table>
Neural basis of learning

Global signal
(Reward, motor error, etc.)

Local signals
(Voltage, calcium, etc.)

Synaptic plasticity

The interaction between global and local signals is largely uncharacterized.

Why difficult?
Learning with empiric inputs

Fiete and Seung
A Synaptic Theory of Gradient Learning with Empiric Inputs

learning rule

\[ \Delta W_{ij} = \eta Re_{ij} \]

\[ R = R[V] \]

\[ e_{ij} = \int_0^T dt \left( \xi_i - \langle \xi_i \rangle \right) s_{ij} \] local eligibility

gradient following

\[ \langle \Delta W_{ij} \rangle = \frac{\partial \langle R \rangle}{\partial W_{ij}} \]

equation: spiking network

\[ \frac{dV}{dt} = -g_{syn}^{\text{ion}} (V_i - V_E) - \sum_a g_{aw}^{\text{ion}} (V_i - V_a) \]

\[ \frac{dx_{ij}}{dt} = -f_i (V_i, s_i, \{u_a, \ldots\}) \]

\[ g_{ij}^{\text{syn}} = \sum_j W_{ij} s_{ij} + \xi_i \]

\[ \Delta W_{ij} = \frac{\partial \langle R \rangle}{\partial W_{ij}} \]

Sensitivity lemma

\[ \frac{\partial \langle R \rangle}{\partial W_{ij}} = \frac{\partial}{\partial b_j} \langle R s_{ij} \rangle \]

\[ \left( g_{ij}^{\text{syn}} = \sum_j W_{ij} s_{ij} + b_i + \xi_i \right) \]
Stochastic gradient ascent

- Reward optimization, but only on average.
- Model-free.
- Local to synapses aside from evaluation.
- Deals with delayed evaluation and dynamic synapses.
- Slow?

Learning time

- Individual correlations small \(\rightarrow\) long averaging?
- Apply to biological example.

Birdsong learning

*with Sebastian Seung & Michale Fee*
Behavior

Social feedback and tutor song not needed.
Auditory feedback of own song crucial.
A Synaptic Theory of Gradient Learning with Empiric Inputs

Song circuit

motor

HVC

RA:

resp syrinx

anteriormotor

t forebrain

DLM

X

LMAN

Lesion:

no song

premature crystallization

Dr. Ila Fiete, KITP (KITP, Understanding the Brain Program 9/21/04)
RA activity

Song representation

HVC generates a sparse sequence later stages learn/perform the motor map
LMAN activity is variable

Hessler & Doupe

Song system schematic

auditory evaluation → motor network ← LMAN

motor commands

vocal apparatus

song
A Synaptic Theory of Gradient Learning with Empiric Inputs

Dr. Ila Fiete, KITP (KITP, Understanding the Brain Program 9/21/04)
Scaling to large networks

Correlation of single neuron with reward = 1/N.
Time to learn ~ N?

factor of 4000 more neurons in bird:

<table>
<thead>
<tr>
<th>model</th>
<th>HVC</th>
<th>RA</th>
<th>outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>bird</td>
<td>20000</td>
<td>8000</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>200</td>
<td>2</td>
</tr>
</tbody>
</table>
Redundancy and rank

Correlation of single output with reward = 1/N₀.

Time to learn ~ N₀.

Model: 2 outputs
Bird: 8 outputs

Neural or synaptic noise. Time: <8000 epochs.
‘Simple’ tasks can be learned

- Learning time is independent of network size in redundant networks.
- Zebra finches practice 80,000 times.
- Motor control is redundant and learning is often slow.

Experimental probe

\[ x_j \]
\[ \xi_i \] experimenter
\[ \Delta W > 0 \quad \Delta W < 0 \]
Why is the brain so noisy?

- Why is spiking irregular?
- Why are synapses unreliable?
- Does the brain use noise for learning?
Similarity to Hebbian learning?

\[ \Delta W_{ij} = R \left( \xi_i - \langle \xi_i \rangle \right) s_j \]

strong noise input and weak regular inputs

\[ \approx R \left( \delta_i - \langle \xi_i \rangle \right) s_j \]

need segregated noise inputs so noise baseline correct

Supervised, reinforced, or unsupervised?

noise as exploration

\[ \text{reward} \]

noise as instruction

\[ \text{increase active weights} \]

post noise spike

\[ \text{reward} \]

no noise spike

\[ \text{decrease active weights} \]

however, ‘instruction’ is stochastic, uncorrelated with error
Evidence of heterosynaptic plasticity

Tsukamoto et al., 2003

also cerebellar CF, PF convergence at Purkinje cells

Practice vs. performance

Hessler and Doupe, 1999
A Synaptic Theory of Gradient Learning with Empiric Inputs

Song circuit

Delayed reward

Dr. Ila Fiete, KITP (KITP, Understanding the Brain Program 9/21/04)
IF dynamics

\[ C_m \frac{dV_j}{dt} = -g_L(V_j - V_L) - g_{E_j}(V_j - V_L) \]

\[ V_j = V_{\text{thres}} \]
\[ V_j := V_{\text{reset}} \]
\[ s_j := s_j + \Delta \]

\[ \tau_s \frac{ds_j}{dt} = -s_j(t) \]
\[ g_{E_i} = \sum_j W_{ij} s_j(t) \]

Gradient following:

\[ \Delta R = \frac{\partial R}{\partial W_{ij}} \Delta W_{ij} \]

\[ \Delta W_{ij} = \frac{\partial R}{\partial W_{ij}} \Rightarrow \Delta R = \left( \frac{\partial R}{\partial W_{ij}} \right)^2 > 0 \]

REINFORCE:

\[ \left\langle \Delta W_{ij} \right\rangle = \frac{\partial \langle R \rangle}{\partial W_{ij}} = \frac{\partial}{\partial W_{ij}} \sum_s P(s) R(s) \]

\[ = \sum_s P(s) \left( R(s) \frac{\partial \ln P(s)}{\partial W_{ij}} \right) \]

\[ = \left\langle R(s) \frac{\partial \ln P(s)}{\partial W_{ij}} \right\rangle \]
\[ \Delta W_{ij} = R(s) \frac{\partial \ln P(s)}{\partial W_{ij}} = Re_i \]
A Synaptic Theory of Gradient Learning with Empiric Inputs