The Rescorla-Wagner learning rule

Christian Machens
Group for Neural Theory
Ecole normale supérieure Paris
Classical conditioning a la Pavlov
Classical conditioning

1. food $\rightarrow$ salivation

2. sound $\rightarrow$ no salivation

3. sound $\rightarrow$ food $\rightarrow$ salivation

4. sound $\rightarrow$ salivation
# Extinction

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>sound</td>
<td>food</td>
</tr>
<tr>
<td>2.</td>
<td>sound</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>sound</td>
<td>no food</td>
</tr>
<tr>
<td>4.</td>
<td>sound</td>
<td></td>
</tr>
</tbody>
</table>
What does the dog want?

Assume: The dog wants to be able to predict the reward!

\[ u_i \text{ stimulus in trial } i: \ u_i = 0 \ \text{ or } \ u_i = 1 \]
\[ r_i \text{ reward in trial } i: \ r_i = 0 \ \text{ or } \ r_i = 1 \]
\[ \nu_i \text{ reward that the dog expects in trial } i \]
What does the dog learn?

Assume: The dog wants to be able to predict the reward!

\[ u_i \quad \text{stimulus in trial } i: \quad u_i = 0 \quad \text{or} \quad u_i = 1 \]

\[ r_i \quad \text{reward in trial } i: \quad r_i = 0 \quad \text{or} \quad r_i = 1 \]

\[ v_i \quad \text{reward that the dog expects in trial } i \]

Assume: The dog learns to minimize a “loss” function:

\[ L = \sum_{i=1}^{N} (r_i - v_i)^2 \]
What does the dog learn?

Assume: The dog wants to be able to predict the reward!

- $u_i$: stimulus in trial $i$: $u_i = 0$ or $u_i = 1$
- $r_i$: reward in trial $i$: $r_i = 0$ or $r_i = 1$
- $v_i$: reward that the dog expects in trial $i$

Assume: The dog learns to minimize a “loss” function:

$$L = \sum_{i=1}^{N} (r_i - v_i)^2$$

Assume: 

dog’s model of the world

$$v_i = w u_i$$
What does the dog learn?

Assume: The dog wants to be able to predict the reward!

- $u_i$  stimulus in trial $i$: $u_i = 0$ or $u_i = 1$
- $r_i$  reward in trial $i$: $r_i = 0$ or $r_i = 1$
- $v_i$  reward that the dog expects in trial $i$

Assume: The dog learns to minimize a “loss” function:

$$L = \sum_{i=1}^{N} (r_i - v_i)^2$$

Assume: The dog’s model of the world parameter that the dog needs to learn from observations
What does the dog learn?

Assume: The dog wants to be able to predict the reward!

- stimulus in trial $i$: $u_i = 0$ or $u_i = 1$
- reward in trial $i$: $r_i = 0$ or $r_i = 1$
- reward that the dog expects in trial $i$: $v_i$

Assume: The dog learns to minimize a “loss” function:

$$ L = \sum_{i=1}^{N} (r_i - v_i)^2 $$

$v_i = wu_i$

“Loss” in the $i$-th trial:

$$ L_i = (r_i - wu_i)^2 $$
How should the dog adopt its world model?

“Loss” in the i-th trial:

\[ L_i = (r_i - w u_i)^2 \]

Update parameter w to decrease loss!

\[ w \rightarrow w - \epsilon \frac{d}{dw} L_i \]
How should the dog adopt its world model?

$$\frac{d}{dw} L_i = \frac{d}{dw} (r_i - w u_i)^2$$

Update parameter $w$ to decrease loss!

$$w \rightarrow w - \epsilon \frac{d}{dw} L_i$$
How should the dog adopt its world model?

\[
\frac{d}{dw} L_i = \frac{d}{dw} (r_i - w u_i)^2 = -2 u_i (r_i - w u_i)
\]

Update parameter \( w \) to decrease loss!

\[
w \to w - \epsilon \frac{d}{dw} L_i
\]
How should the dog adopt its world model?

\[
\frac{d}{dw} L_i = \frac{d}{dw} (r_i - wu_i)^2
\]

\[
= -2u_i (r_i - wu_i)
\]

\[
= -2u_i \delta_i
\]

\[
\delta_i = r_i - uw_i = r_i - v_i
\]

“prediction error”

Update parameter \( w \) to decrease loss!

\[
w \rightarrow w - \epsilon \frac{d}{dw} L_i
\]
Minimizing the loss function

Minimize loss function (≈ maximize ability to predict reward)

\[
\frac{d}{dw} L_i = \frac{d}{dw} (r_i - wu_i)^2
\]

\[
= -2u_i(r_i - wu_i)
\]

\[
= -2u_i \delta_i
\]

“Rescorla-Wagner”-rule

\[
w \rightarrow w + \epsilon \delta_i u_i
\]

\[
\delta_i = r_i - uw_i = r_i - v_i
\]

“prediction error”
Rescorla-Wagner rule: Conditioning and extinction

stimulus + reward (CS + US)

\( u_i = 1 \quad r_i = 1 \)

stimulus (CS)

\( u_i = 1 \quad r_i = 0 \)

\( v_i = w u_i \)

\( v_i = w u_i \)
Rabbit eye blinking: Conditioning and extinction

Schneiderman et al, Science, 1962
### Classical conditioning: blocking

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>training</strong></td>
<td>sound</td>
<td>food</td>
<td>salivation</td>
<td>no salivation</td>
</tr>
<tr>
<td><strong>clapping</strong></td>
<td>sound</td>
<td>food</td>
<td>salivation</td>
<td>no salivation</td>
</tr>
<tr>
<td><strong>sound</strong></td>
<td>food</td>
<td>salivation</td>
<td>salivation</td>
<td>no salivation</td>
</tr>
</tbody>
</table>
Reward prediction with multiple stimuli: vectorized Rescorla-Wagner rule

Simple linear model (i-th trial)

\[ v_i = w \cdot u_i \]

Expected reward
Weight vector
Stimulus vector

“Rescorla-Wagner”-rule

\[ w \rightarrow w + \epsilon \delta u_i \]

\[ \delta = r_i - v_i \]

“prediction error”
Classical conditioning: blocking

\[ v_i = w \cdot u_i \]

\[ w \rightarrow w + \epsilon \delta u_i \]
<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>sound → food → salivation</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>clapping → food → salivation</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>light, sound → salivation disappears</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>clapping → salivation</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>light, clapping → reduced salivation</td>
<td></td>
</tr>
</tbody>
</table>
Classical conditioning:
inhibitory conditioning
Secondary conditioning

1. sound $\rightarrow$ food $\rightarrow$ salivation
2. clapping $\rightarrow$ sound $\rightarrow$ food $\rightarrow$ salivation
3. sound $\rightarrow$ salivation
4. clapping $\rightarrow$ salivation

$\rightarrow$ cannot be explained by Rescorla-Wagner ...