The Rescorla-Wagner learning rule

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

1. Food → Salivation

2. Sound → No Salivation

3. Sound → Food → Salivation

4. Sound → Salivation
# Extinction

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><img src="image1" alt="Sound" /> <img src="image2" alt="Food" /> → <img src="image3" alt="Salivation" /></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td><img src="image1" alt="Sound" /> → <img src="image3" alt="Salivation" /></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td><img src="image1" alt="Sound" /> → <img src="image4" alt="No Food" /> → <img src="image5" alt="Salivation Disappears" /></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td><img src="image1" alt="Sound" /> → <img src="image6" alt="No Salivation" /></td>
<td></td>
</tr>
</tbody>
</table>
What does the dog want?

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
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:

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

\[ u_i \text{ stimulus in trial i: } u_i = 0 \text{ or } u_i = 1 \]
\[ r_i \text{ reward in trial i: } r_i = 0 \text{ or } r_i = 1 \]
\[ v_i \text{ reward that the dog expects in trial i} \]

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

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

Assume: dog’s model of the world

\[ v_i = w u_i \]

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!

- \( u_i \) 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:

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

\( v_i = w u_i \)

“Loss” in the i-th trial:

\[
L_i = (r_i - w u_i)^2
\]
How should the dog adopt its world model?

“Loss” in the i-th trial:

\[ L_i = (r_i - wu_i)^2 \]

Update parameter w to decrease loss!

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

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

Update parameter \( w \) to decrease loss!

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

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

Update parameter \( w \) to decrease loss!

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

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

Update parameter \( w \) to decrease loss!

\[ w \rightarrow w - \epsilon \frac{d}{dw} L_i \]

\( \delta_i = r_i - uw_i = r_i - v_i \) “prediction error”
Minimizing the loss function

Minimize loss function (= maximize ability to predict reward)

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

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

\[
= -2u_i \delta_i
\]

\[
\delta_i = r_i - uw_i = r_i - v_i \quad \text{“prediction error”}
\]

“Rescorla-Wagner”-rule

\[
w \rightarrow w + \epsilon \delta_i u_i
\]
Rescorla-Wagner rule: Conditioning and extinction

stimulus + reward
(CS + US)

$$u_i = 1 \quad r_i = 1$$

stimulus
(CS)

$$u_i = 1 \quad r_i = 0$$

$$v_i = w u_i$$
Rabbit eye blinking: Conditioning and extinction

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

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

Simple linear model (i-th trial)

\[ v_i = w \cdot u_i \]

"Rescorla-Wagner"-rule

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

\[ \delta = r_i - v_i \]

"prediction error"
Classical conditioning: blocking

\[ v_i = w \cdot u_i \quad \text{and} \quad w \rightarrow w + \epsilon \delta u_i \]
# Inhibitory conditioning

<table>
<thead>
<tr>
<th>Step</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>sound → food → salivation</td>
<td>sound → food → salivation</td>
</tr>
<tr>
<td>2.</td>
<td>clapping → food → salivation</td>
<td>clapping → salivation</td>
</tr>
<tr>
<td>3.</td>
<td>light, sound → salivation disappears</td>
<td>light, clapping → reduced salivation</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
cannot be explained by Rescorla-Wagner ...