The Enhanced Rescorla-Wagner Model

Suyamoon Pathak
241110091
suyamoonp24@cse.iitk.ac.in

Abstract

Rescorla-Wagner model is a great model that helps us to understand Association (Rescorla and Wagner, 1972). The way they have modeled how the brain treats two stimuli (conditioned, and unconditioned), relates them, and comes up with a totally different representation in the form of 'learning' is truly amazing. But, it has many limitations. Rescorla-Wagner model was a fairly simple model which didn't account for many real-world conditions (Miller et al., 1995). In this paper, we will be discussing what assumptions Robert A. Rescorla, and Allan R. Wagner used in their model, and how we can tweak them a little to come up with a better model. In this paper, we will first be discussing the important assumptions, the cases where the assumptions fail, and the changes we can create for those cases, and finally propose a combined enhanced model. We will also test each of our proposed technique on a simulated environment, whose plots have been put under their respective headings. The link for colab notebook used for conducting the experiments can be found in the References.

1 Introduction

First, let us briefly understand about the equation that the Rescorla-Wagner model proposed:

$$\Delta V = \alpha\beta(\lambda - V_{tot}) \tag{1}$$

where:

- ΔV is the change in associative strength.
- α is the salience of the conditioned stimulus.
- β is the learning rate.
- λ is the maximum associative strength.
- V_{tot} is the total sum of associative strengths. (Rescorla and Wagner, 1972)

2 Assumptions and Proposed Extensions

1. **Learning is driven by prediction error:** Rescorla and Wagner assume that learning occurs only when the agent experiences a "surprise" or a "shock" (Hollis, 2019). They imply that when the agent's predictions perfectly match the outcome ($V_{tot} = \lambda$), no further learning takes place. However, this assumption fails to account for phenomena like latent learning.

   So, it can very well happen that the prediction error is zero, but it still continues to learn.

   To allow latent learning, we introduce a constant background learning rate (γ) that ensures a baseline level of learning. So, the updated equation becomes:

$$\Delta V = \alpha\beta(\lambda - V_{tot}) + \gamma \tag{2}$$

   where γ is the baseline learning rate independent of the prediction error.

   To check how the updated model performs against the original Rescorla-Wagner model, we designed a simulation where an agent explores an environment without immediate reinforcement. Figure 1 shows how the comparison looks with a constant background learning rate γ.

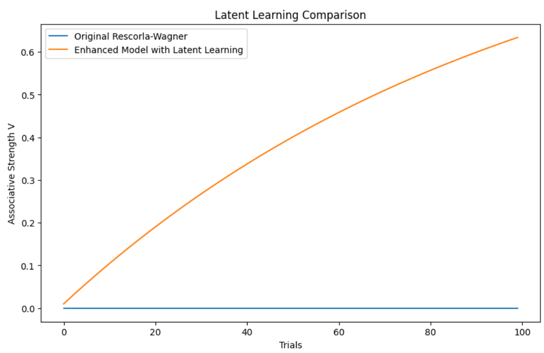

Figure 1: Comparison of associative strength V over trials in the original Rescorla-Wagner model and the enhanced model with latent learning. In the simulation, the agent was exposed to a single conditioned stimulus A over 100 trials without reinforcement ($\lambda = 0$). Parameters used were $\alpha = 0.1$, $\beta = 0.1$, and $\gamma = 0.01$ (for the enhanced model) (Pathak, 2024).

2. **Blocking:**
The Rescorla-Wagner model assumes when multiple conditioned stimuli are present, the combined associative strength (V_{tot}) determines the prediction error (Yau and McNally, 2023). And, if one stimulus already explains the unconditioned stimulus, others won't gain any strength. This doesn't necessarily have to be true in real-life situations. The brain can also treat the combinations of stimuli as unique patterns (configural learning). So, to model this type of behavior, instead of just summing up the associative strengths, we can use a non-linear function $f(V)$, which can change according to the trial, and its stimuli. So, the updated model becomes:

$$\Delta V = \alpha\beta(\lambda - f(V_{tot)}) \tag{3}$$

To evaluate how the enhanced model addresses the phenomenon of blocking and configural learning, we designed a simulation with two phases. In Phase 1, stimulus A was repeatedly paired with reinforcement, allowing it to acquire associative strength. In Phase 2, stimuli A and B were presented together with reinforcement. Check Figure 2 for the plots.

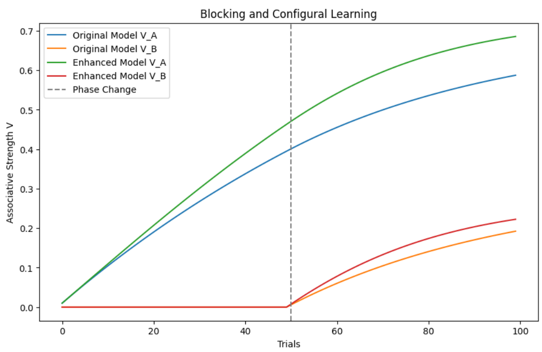

Figure 2: Associative strengths of stimuli A and B over trials in the original Rescorla-Wagner model and the enhanced model with a non-linear function $f(V_{tot}) = V_{tot}^2$. The simulation consisted of two phases of 50 trials each: Phase 1 with stimulus A alone, and Phase 2 with stimuli A and B together. Parameters used were $\alpha = 0.1$, $\beta = 0.1$, and $\lambda = 1.0$. The vertical dashed line indicates the transition between phases (Pathak, 2024).

3. **Extinction is the reverse of learning:**
Extinction basically means when a conditioned stimulus is provided without the unconditioned stimulus. Rescorla and Wagner assume extinction is the reverse of learning. And, it reduces associative strength over time. (Culver et al., 2015)

Mathematically, extinction happens because the maximum associative strength (λ) is zero, and the prediction error (λV_{tot}) is negative, which finally leads to ΔV being negative.

But, in the real-world, extinction is not always the reverse of learning. For example, if you're teaching a dog to sit, when you say "sit!". Firstly you're conditioning it by giving a sweet as a reward every time the dog sits. But, once it has learned it, it will still sit for the rest of its life even if you stop providing the sweet (reward) after some time. The effect can be suppressed, but is not negative.

This can be better explained by phenomena like spontaneous recovery, reinstatement, or renewal. They all suggest that the original association is not lost, but reduced. To account for these things, lets separate the associative strength (V) into two components– V_{active} and $V_{inactive}$.

So, during acquisition:

$$\Delta V_{active} = \alpha\beta(\lambda - \sum V_{active}) \tag{4}$$

During extinction:

$$\Delta V_{active} = -\alpha\beta(\sum V_{active}) \tag{5}$$

$$\Delta V_{inactive} = \eta V_{active} \tag{6}$$

where η is the transfer rate between active and inactive associations.

And, when recovery:

$$V_{inactive} = V_{active} \tag{7}$$

To see how the enhanced model accounts for extinction and spontaneous recovery, we conducted a simulation with three phases: acquisition, extinction, and recovery. We observe the following things in Figure 3:

- In Acquisition phase (Trial 1-50), active association strength increase, but inactive association strength remains at 0.
- In Extinction phase (Trial 51-80), active association strength decreases, but inactive association strength increases. This shows that the association is becoming inactive but not being erased completely.

- In Recovery phase (Trial 81-100), active association strength increases again without any reinforcement (spontaneous recovery). The inactive association strength decrease.

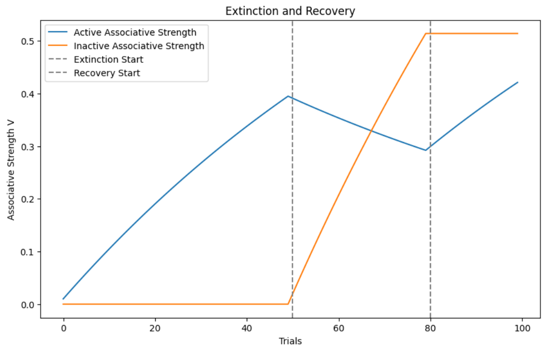

Figure 3: Active and inactive associative strengths of stimulus A over trials in the enhanced model. The simulation consisted of 50 acquisition trials, 30 extinction trials, and 20 recovery trials. Parameters used were $\alpha = 0.1$, $\beta = 0.1$, $\eta = 0.05$, and $\lambda = 1.0$. Vertical dashed lines indicate the transitions between phases (Pathak, 2024).

4. **Linear summation of associative strength:** According to the Rescorla-Wagner model, the total expectation (V_{tot}) is the sum of associative strengths of all the stimuli (Blaisdell et al., 2009). This implies every stimulus contributes independently to the overall prediction. But, in real-world, it is not always the case. Real-world can also have non-linear interactions, and it might be naive to assume the independent assumption. To accomodate this, we can generalize the model to use a weighted interaction matrix (W). So, the updated model equation becomes:

$$\Delta V_i = \alpha_i \beta \left(\lambda - \sum_j W_{ij} V_j \right) \tag{8}$$

Where W is a matrix that scales the contribution of each stimulus j to the total prediction for stimulus i.

Again, to check how this updated model performed with respect to the original model, we conducted a simulation, which suggests the updated model performs better. Check Figure 4 for the plots.

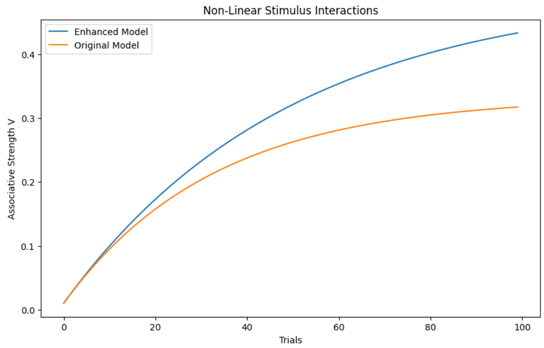

Figure 4: Associative strengths of stimulus over 100 trials in the original Rescorla-Wagner model and the enhanced model with weighted interactions. Parameters used were $\alpha = 0.1$, $\beta = 0.1$, $\lambda = 1.0$. The weighted interaction matrix W for the enhanced model was $W = [1.0\ 0.5]$ (Pathak, 2024).

3 Final Combined model

By integrating all the above-mentioned techniques, we can define the final updated model by this equation below:

$$\Delta V_i = \alpha_i \beta \left(\lambda - f\left(\sum_j W_{ij} V_j \right) \right) + \gamma - \eta V_{active} \tag{9}$$

where:

- γ is the baseline learning rate (latent learning).
- f is the non-linear function (configural learning).
- W_{ij} is the weighted interaction matrix (non-linear stimulus interactions).
- V_{active} and $V_{inactive}$ are the split associative strengths (extinction and recovery).

4 Conclusion

In a nutshell, this paper tries to make the Rescorla-Wagner model more robust to the real-world learning environment. We accounted for the latent learning by introducing a baseline learning rate, introduced a non-linear function mapping to avoid the blocking that happens in case of multiple stimuli, we proposed not erasing the original learning, rather suppressing it to account for spontaneous recovery, and finally introduced a weighted interaction matrix to provide different weights to different stimuli. These all techniques were brought together to propose the final combined model.

References

A. P. Blaisdell, K. J. Leising, and I. P. L. McLaren. 2009. Overexpectation and trial massing. *Behavioral Neuroscience*, 123(5):992–1000.

Najwa C. Culver, Bram Vervliet, and Michelle G. Craske. 2015. Compound extinction: Using the rescorla–wagner model to maximize exposure therapy effects for anxiety disorders. *Clinical Psychological Science*, 3(3):335–348.

Geoff Hollis. 2019. Learning about things that never happened: A critique and refinement of the rescorla-wagner update rule when many outcomes are possible. *Memory & cognition*, 47(7):1415–1430.

Ralph R Miller, Robert C Barnet, and Nicholas J Grahame. 1995. Assessment of the rescorla-wagner model. *Psychological bulletin*, 117(3):363.

Suyamoon Pathak. 2024. Code for running experiments. Google Colab Notebook.

RA Rescorla and Allan Wagner. 1972. *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement*, volume Vol. 2.

Joanna Oi-Yue Yau and Gavan P. McNally. 2023. The rescorla-wagner model, prediction error, and fear learning. *Neurobiology of Learning and Memory*, 203:107799.