Sharif University of Technology

Advance Neuroscience

Dr.Ali Ghazizadeh

Motivation and Classical conditioning Assignment

By:Mohammad Khorshidi

The aim of this assignment is modeling the learning in classical conditioning and see whether a model can correctly follow the values or not.

1 Rescula Wagner Model

A well know model in classical conditioning and reinforcement learning is the Rescola-Wagner. This model predicts that violations of our expected reward for each stimuli or combination of stimuli causes incremental changes in our belief about their values. RW Model:

$$v = wu$$

Where, w is the weight, v is the expected reward and u is a binary variable that represents the presence or absence of the stimulus. To update w with RW Model:

$$w \to w + \epsilon\delta u \quad with \quad \delta = r - v$$

Where ϵ is the learning rate and r is immediate reward at the current stage.

Below are some of the well-known classical conditioning paradigms and the behavioral results of each paradigm as shown by experiments:

Paradigm	Pre-Train	Train	Result
Pavlovian		$s \to r$	$s \to$ 'r'
Extinction	$s \to r$	$s \to \cdot$	$s \to$ '$\cdot$'
Partial		$s \to r \quad s \to \cdot$	$s \to \alpha$ 'r'
Blocking	$s_1 \to r$	$s_1 + s_2 \to r$	$s_1 \to$ 'r' $\quad s_2 \to$ '$\cdot$'
Inhibitory		$s_1 + s_2 \to \cdot \quad s_1 \to r$	$s_1 \to$ 'r' $\quad s_2 \to -$'r'
Overshadow		$s_1 + s_2 \to r$	$s_1 \to \alpha_1$'r' $\quad s_2 \to \alpha_2$'r'
Secondary	$s_1 \to r$	$s_2 \to s_1$	$s_2 \to$ 'r'

1. To see whether the RW model can learn the Paradigms and final values, I simulate the paradigms using numerical calculation. As we can see in Figure 1, the RW model can completely follow the real results and can be used in this conditioning. In following simulations in this section, the learning rate, ϵ, was 0.05.

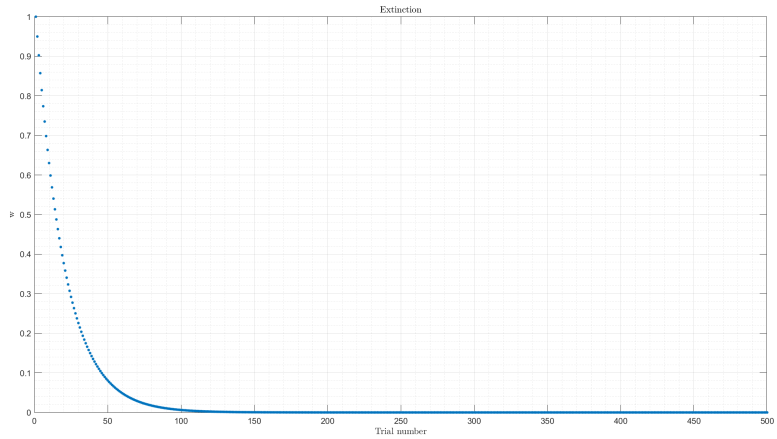


Figure 1: The Extinction Paradigm Simulation Using RW Model.

For Partial Conditioning for different α, as we can see in Figure 2, the RW model can completely learn the α value and can trace it.

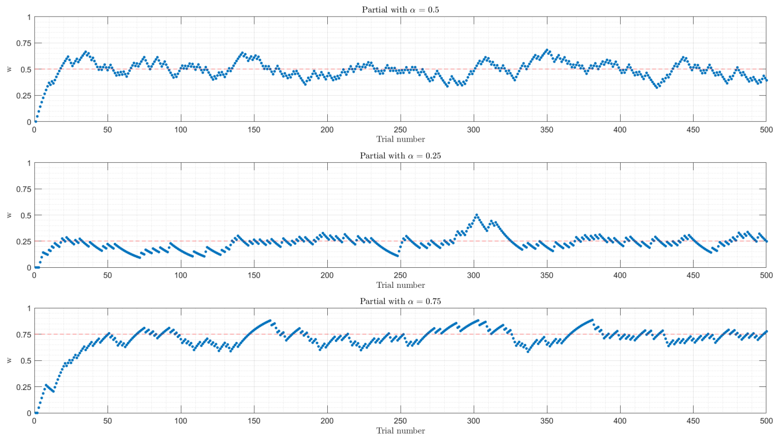
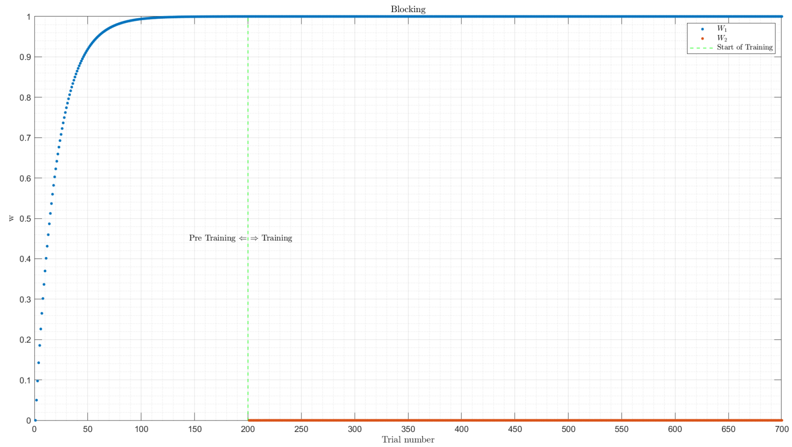


Figure 2: The Partial Paradigm With Different α Simulation Using RW Model.

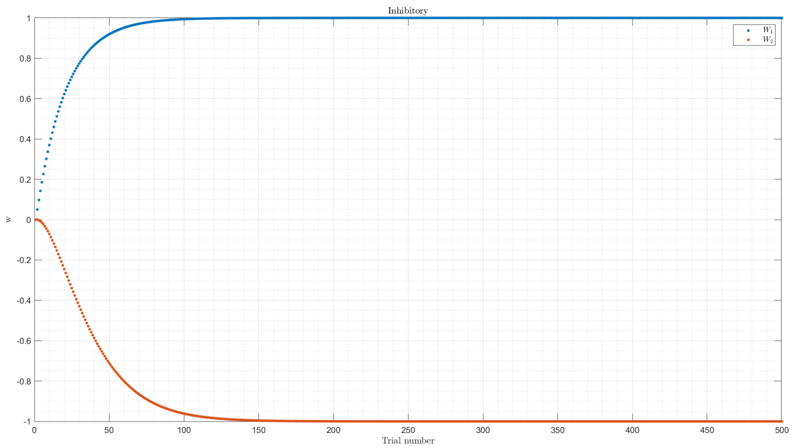
If we have two stimulus, we can re write the main formulation as:

$$\begin{bmatrix} w_1(n+1) \\ w_2(n+1) \end{bmatrix} = \begin{bmatrix} w_1(n+1) \\ w_2(n+1) \end{bmatrix} + \begin{bmatrix} \epsilon s_1 \\ \epsilon s_2 \end{bmatrix} (r - w_1(n)s_1(n) - w_2(n)s_2(n))$$

To simulate the Blocking and Inhibitory Paradigms, I used the above formulation, and as we can see in Figure 3, the model can be used for simulating these two paradigms.



(a) Blocking Paradigm



(b) Inhibitory Paradigm

Figure 3: Two stimulus Paradigms

2. For Overshadow Paradigm, the difference between two stimuli is the learning rate for each one and the final value's differences are determined by this. As we can see in Figures 4,5,6, the model can learn the learning rate ratio in two stimuli and can follow the real results.

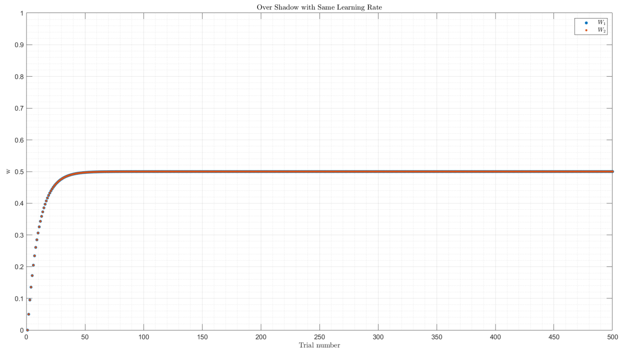


Figure 4: Overshadow Paradigm With Same Learning Rate.

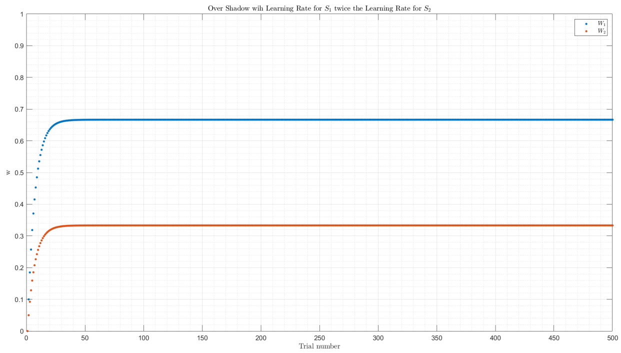


Figure 5: Overshadow Paradigm With Learning Rate Ratio = 2.

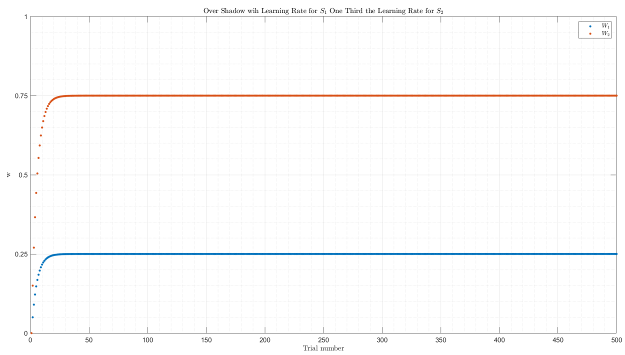
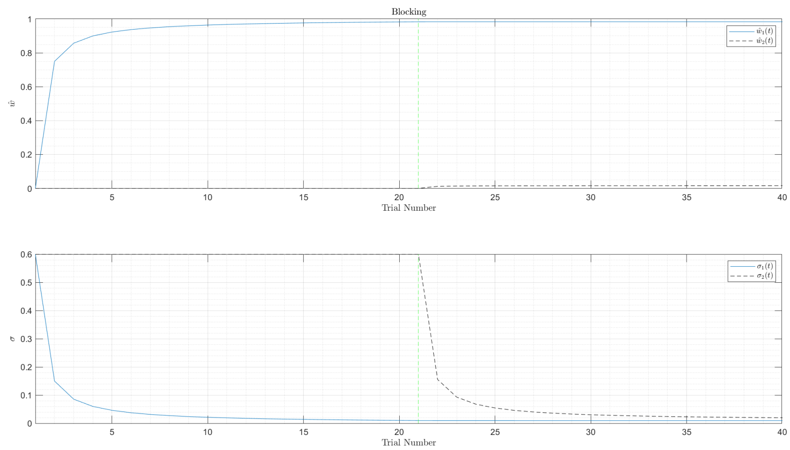


Figure 6: Overshadow Paradigm With Learning Rate Ratio = $\frac{1}{3}$.

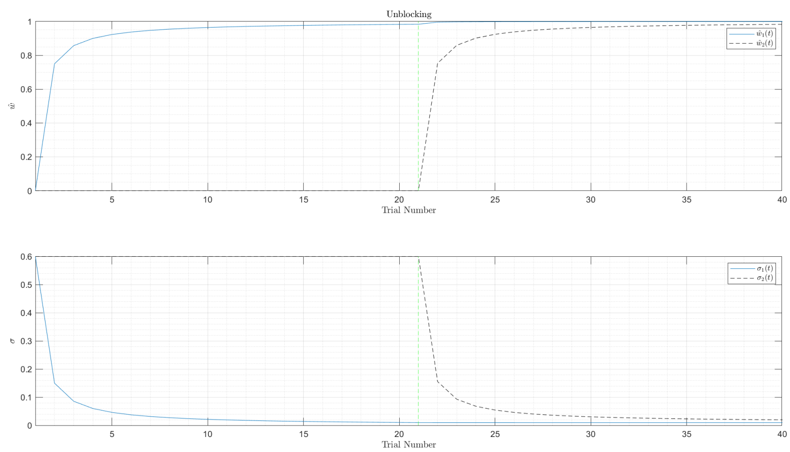
2 Kalman Filter

Delta rule cannot explain the backward blocking. Delta rule predicts $w_1 = r$ and $w_2 = \alpha r$. So we have to change the model. Kalman filter as Dayan et al. [**1**] mentioned, can be used here. The Kalman filter with the help of uncertainty tries to model the way humans or animals learn new values. In this method, it assumes that in a space where uncertainty is high, the learning also is very high and when we learn all stages or stimuli's values, the uncertainty (learning rate) will reduce and there is nothing to learn.

1. To see how the Kalman filter works, I regenerate Figure 1 and 2 of [**1**], which is the blocking and unblocking paradigms. As we can see in Figure 7, the Kalman filter learns faster than the RW model. Another thing that brings attention is that the variances are modeling the uncertainty from the world, and we can see that when the variances are near zero, there is no learning.



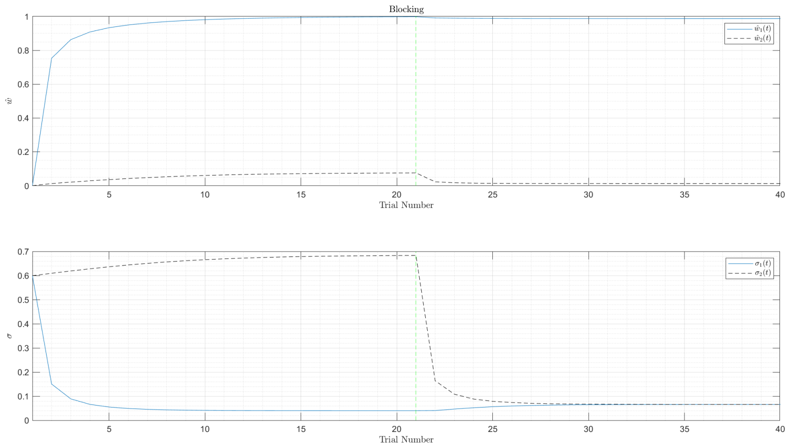
(a) Blocking Paradigm



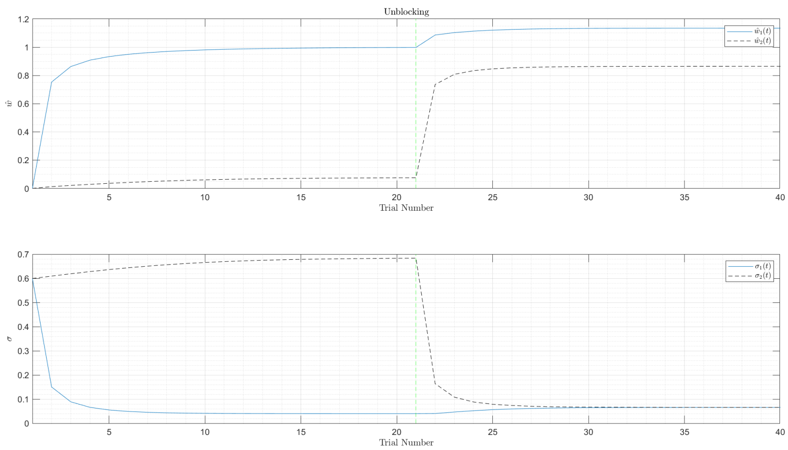
(b) Unblocking Paradigm

Figure 7: Kalman Filter Simulation

2. To see the effect of Process Noise and Measurement Noise, I regenerate the last section's Figures using new values for these noises. As we can see in Figure 8, it causes the variances to be more flexible and even in pre-training when they are near zero, in training have a higher value so for example when σ_1 becomes higher in training, the final value of $\hat{w}_i$. It is compatible with the definition "Process Noise" because we have an inner uncertainty which brings us to here to have flexible variances.



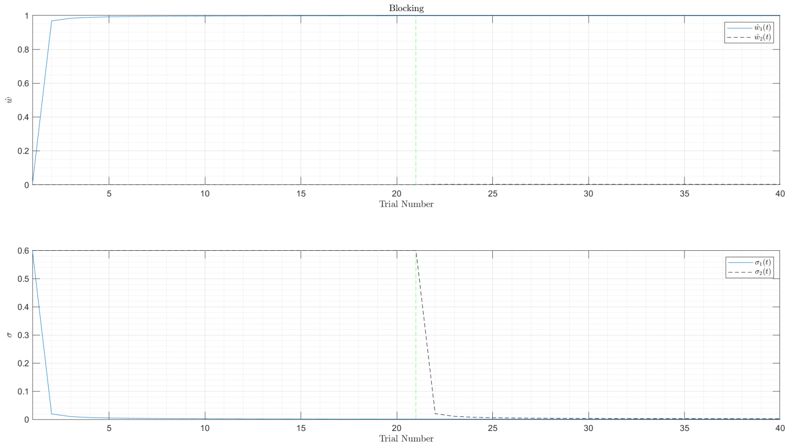
(a) Blocking Paradigm



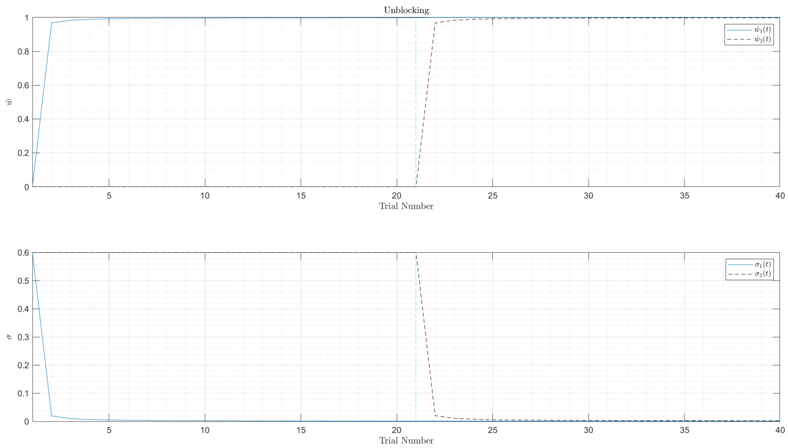
(b) Unblocking Paradigm

Figure 8: Kalman Filter Simulation with Higher Process Value Noise Value

In Figure 9, we can see the effect of Measurement Noise. We can see that with higher value for measurement noise, the variances become near zero much faster and the learning also is faster. Also, the flexibility of learning become lower.



(a) Blocking Paradigm



(b) Unblocking Paradigm

Figure 9: Kalman Filter Simulation with Higher Measurement Value Noise Value

3. As we said earlier, the RW model cannot follow the backward blocking paradigm.

Paradigm	Pre-Training	Training	Result
Backward Blocking	$s_1 + s_2 \to r$	$s_1 \to r$	$s1 \to r \quad s_2 \to' 0'$

But as we can see in Figure 12, the Kalman Filter can completely follow this Paradigm.

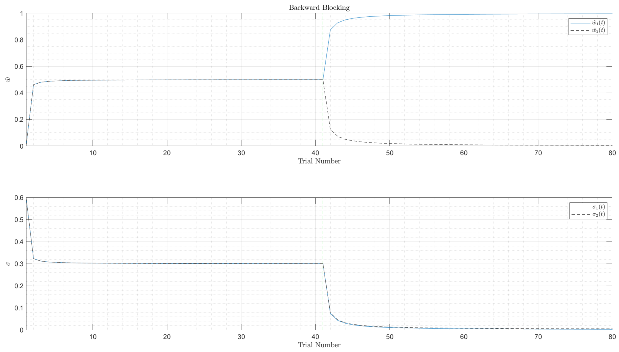
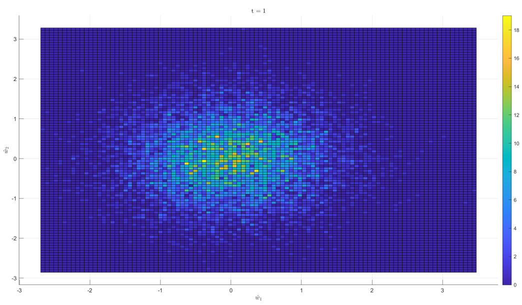
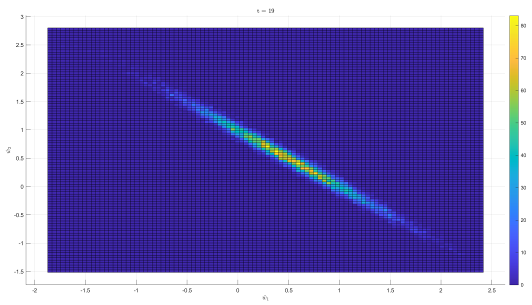


Figure 10: Backward Blocking Paradigm .

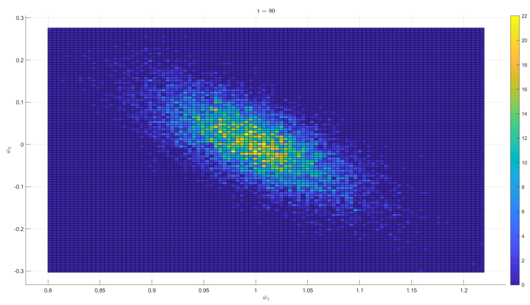
To see how the $\hat{w_1}$ and $\hat{w_2}$ changes over time, we can see at Figure 11. As we can see at t = 0, the weights don't have any correlations. As we go further, the weights structure change into anti-correlated because the sum of two weights must be 1 and if one of them goes up, another must come down. At the end of the experiment, we can see that the mean of weights is at the point where it should be.



(a) t = 0



(b) t = 19



(c) t = 80

Figure 11: Backward Blocking Paradigm

4. to check the following paradigm, $s_1 \to 1$ first and then $s_1 \to -1$, we can see the Figure ??. As we can see, in the beginning, when the variance is high enough, the learning is completed. After that, because variance is so low, the learning rate is near zero so it will take a lot of steps to learn the real value.

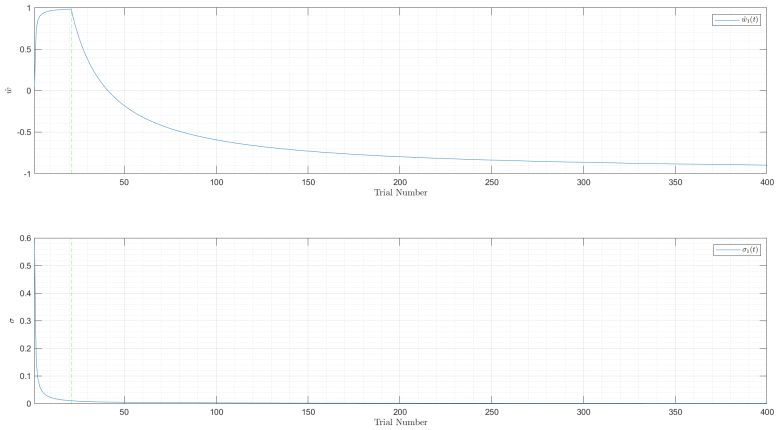


Figure 12: Low Variance Problem.

3 Unknown Uncertainty

The uncertainty modeled by Kalman filter is referred to as 'known uncertainty'. This is the uncertainty about the value of the stimulus for which the agent has some estimate. However there are times when we don't even know how much we do not know about value of a stimulus. This is referred to as 'ambiguity' or 'unknown uncertainty'.

If we do not account for the ambiguity, then our estimate of uncertainty gets smaller over time and we cannot learn about new changes in the environment. To account for this Dayan and Yu [**1**] made their model sensitive to the error magnitude. Large errors served to reset the uncertainty about the values to promote learning according to thresholding this value:

$$\beta(t) = \frac{(r(t) - x(t).\hat{w}(t)))^2}{x(t)^T \Sigma(t) + \tau^2}$$

1. To solve the problem, first we have to determine the threshold. For this, I do a numerical simulation to see which γ has the lowest Mean Squared Error. As we can see in Figure 13, the optimal $\gamma = 1.9$.

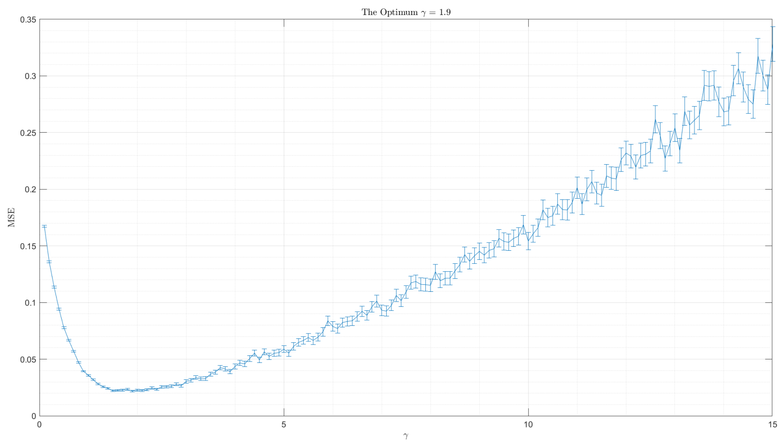


Figure 13: Different Thresholds MSE.

To see if the problem is solved, we can see at Figure 14. As it can be seen, when the β value reaches the threshold γ, the σ value become so large so it can chase the unexpected event. As it can be seen in Figure 15, the $\hat{w}$ values follows the immediate reward very fast. Also, we can see the slow and fast drift in w.

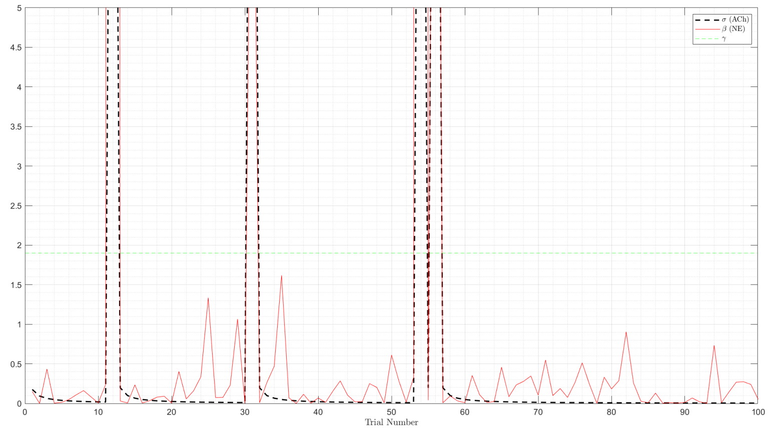


Figure 14: β and σ Changes Over Time.

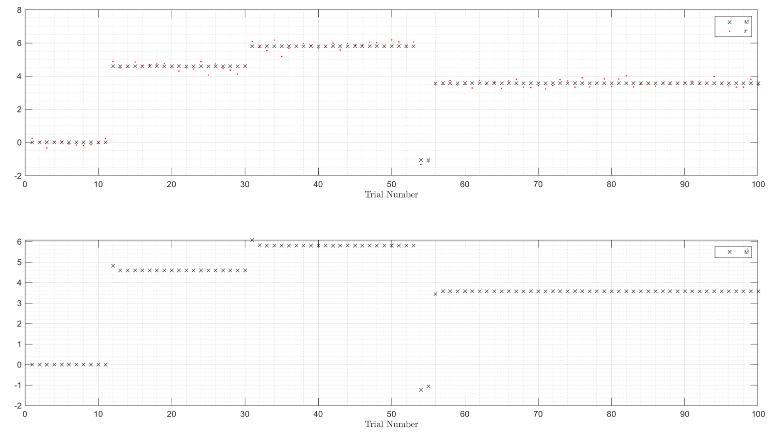


Figure 15: W and r and $\hat{w}$ Changes Over Time.

References

[1] Peter Dayan Angela Jyu (2003) Uncertainty and Learning, IETE Journal of Research, 49:2-3, 171-181, DOI: 10.1080/03772063.2003.11416335