PART I:

Aim:

Modelling learning in classical conditioning paradigms

Description:

We often need to rapidly learn about the value of new stimuli that we encounter or be ready for changes to familiar stimulus values that we were familiar without prior notice. Classical conditioning includes a large group of paradigms where different combination of reward histories with stimuli can form our judgement about their current values.

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.

Instructions:

Use the 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.

And update w with the Rescola-Wagner (RW) rule:

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

Where ϵ is the learning rate

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) Using RW rule, simulate and plot the outcome of following paradigms: extinction, partial, blocking, inhibitory and overshadow. Assume a fixed learning rate and number trials in each

phase to be such that learning almost saturates at the each of each phase. Which one the predictions of RW rule match the above table?

2) For overshadow condition, how can one have different amount of learned value for each stimuli? The ambiguity is a form of a concept known as 'credit assignment' in reinforcement learning literature

According to the paper, 'Uncertainty and Learning', implement Kalman filter method to explain blocking and unblocking in conditioning.

1) Simulate the results shown in figures 1-2 in the Dayan and Yu paper.

2) How does result change by changing process noise vs measurement noise?

3) What factors determine the value of Kalman gain at steady state? Can you derive an approximate relationship between steady state Kalman gain and the model parameters?

4) Does the change in uncertainty depend on the errors made in each trial or changes in the learning context?

5) Using this paradigm if we learn S_1-> 'r'. first and then S_1->'-r'. how does learning the second context compares to the first context? What happens to the learning rate?

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 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:

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

1) Simulate the results shown in Figure 3 of the paper.