Lecture 19: Reinforcement Learning in Humans

04/04/23

Lecturer: Prof. Subrahmanya Swamy Peruru *Scribe: Rohan Baijal, Harsh Garg*

1 RL in Human Cognition, Emotions and Behavior

The setting in RL where an agent interacts with an environment is very similar to how humans act. Understanding how humans learn is a step towards building intelligent agents that can learn efficiently by interacting with their environments.

Studies suggest that the basal ganglia in our brain computes the rewards associated with an action. This effect is studied through the dopamine levels in the brain.

The following section demonstrates how the settings of RL make sense by understanding dopamine levels during interactions.

2 Dopamine

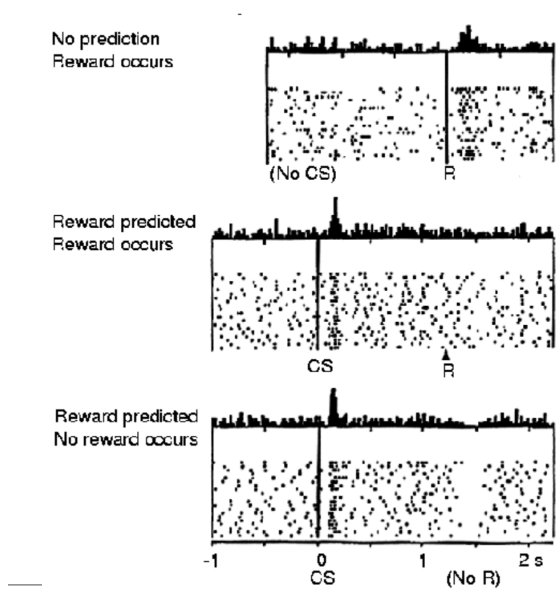

Figure 1: Dopamine Neuron response vs prediction of reward

Dopamine is the brain's "pleasure chemical". Dopamine release can be triggered by various stimulants such as food, psychomotor stimulants, opiates, etc. Apart from that, Beautiful faces, images of lovers, monetary rewards can also trigger the dopamine reward system.

The experiment was performed on monkeys where they were presented with small quantity of fruit juice as a reward and then their dopamine neuron response was studied. In the first Figure, we can see that before learning about the reward, the dopamine neuron is activated after the unpredicted occurrence of the reward. Figure 2 suggests that after learning, monkey's conditioned stimulus predicts reward and there is a spike in dopamine even before the actual reward is presented. In Figure 3, the conditioned stimulus predicts a reward, but the reward fails to occur because of a mistake in the behavioural response of the monkey. Due to the absence of reward, we can observe a depression in dopamine levels which occurs more than 1 s after the conditioned stimulus which reveals an internal representation of the time of the predicted reward in monkeys.

2.1 Dopamine and Food

Studies related to the connection between food and dopamine have also shown interesting results. According to a study by Nestler (2001), presentation of palatable food induced dopamine release.

Appetite is a very good example of delayed rewards in the real world. We eat now (action) and our nutrients rise hours later (reward).

2.2 Dopamine and Effort

While making an optimal choice, our brain takes actions which not only depends on the reward/punishment obtained after the response is made, but also in the cost of the effort incurred while performing the action. For example, in experiments where a rat is given food rewards for pressing a lever multiple times, the quantity that determines the animal's effort is not just food reward, but some sort of an effective reward relating the value (#food pellets) and effort (#lever presses), and the effective reward can be written as:

Effective Reward = (#food_pellets) - effort_factor* (#lever_presses)

where effort_factor denotes a suitable factor that converts cost of effort into equivalent quantity of food pellets. For this experiment the Fixed Ratio Schedule is the number of times the rat has to press lever for 1 unit of food. As we can see in the above image, with an increase in Fixed Ratio Requirement, the number of times the lever is pulled first increases, but then eventually starts decreasing. Hence with dopamine depletion, the effort_factor increases. Which means that the animal becomes more sensitive to cost(number of lever presses in this case).

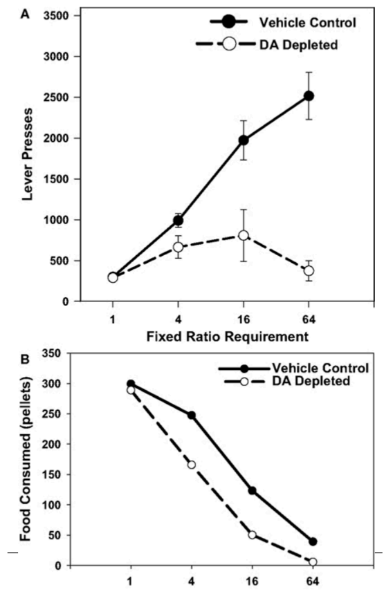

Figure 2: FRR vs #lever_presses and #food_pellets

References

1. Lectures notes by "Prof. Pragathi"