

Chapter 4

Test & Results

4.1 Simulation Results

As we got through our tests, we discovered that Q-Learning has weak results compared to our desired ones. This is mostly because the algorithm works with discrete state space and discrete action space. This was the main problem we had to solve. Robot movement is not a discrete function. It updates its current state continuously and the actions is between the numbers 0 and π also in a continuous spectrum. Although, since we wanted to make an approach which is time and complexity efficient moving our approach to a continuous state and action space, we wouldn't also have a desired result. So, the answer was using Deep Reinforcement Learning and especially Deep Q-Network. DQN as we saw before uses a Neural Network for approximate Q-values instead of calculating them using the Bellman equation. This is more stable as a learning process especially when you are using a large state space. In our case the state space is ideally continuous having too many variables so that makes this approach far more suitable for our problem. Even though, DQN has also the characteristic (as in Q-Learning) that we use only discrete action space with a fine representation and discretization of a large set of values we can make our approach more holistic.

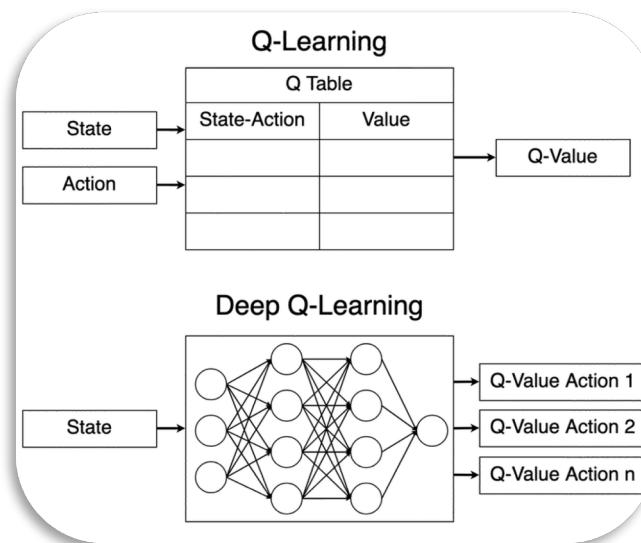


Figure 17: Table that shows the difference between DQ-learning and Q-learning

4.1.1 Q-Learning Results

Letter "C"

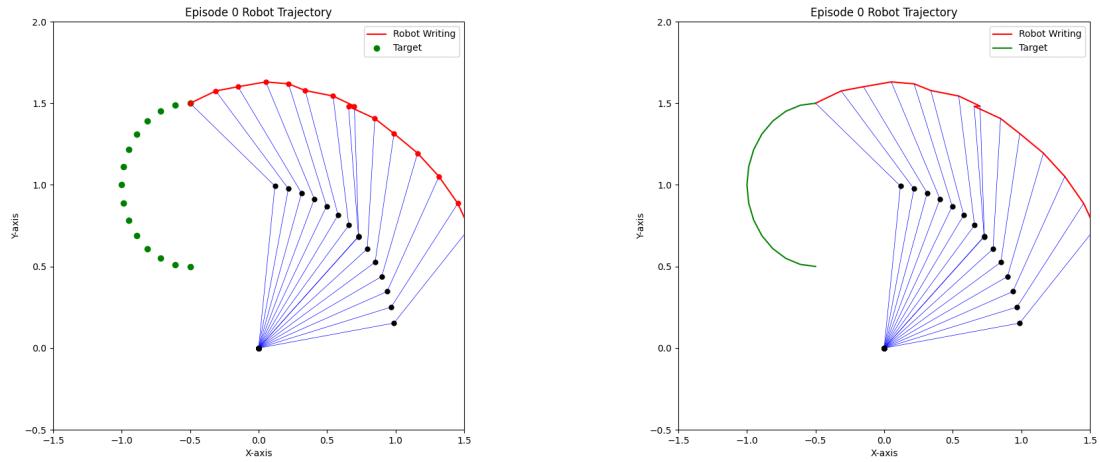


Figure 18: Q-Learning, Letter "C", Episode 0

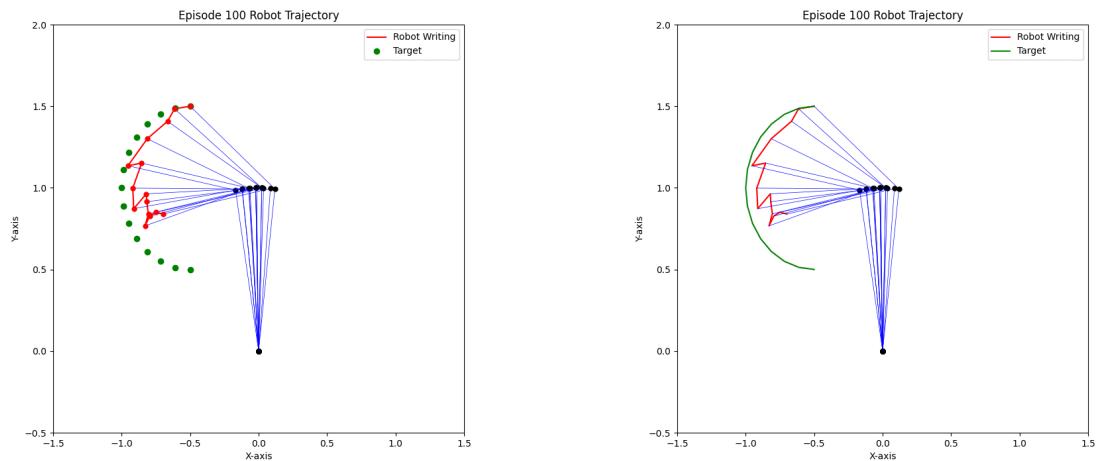


Figure 19: Q-Learning, Letter "C", Episode 100

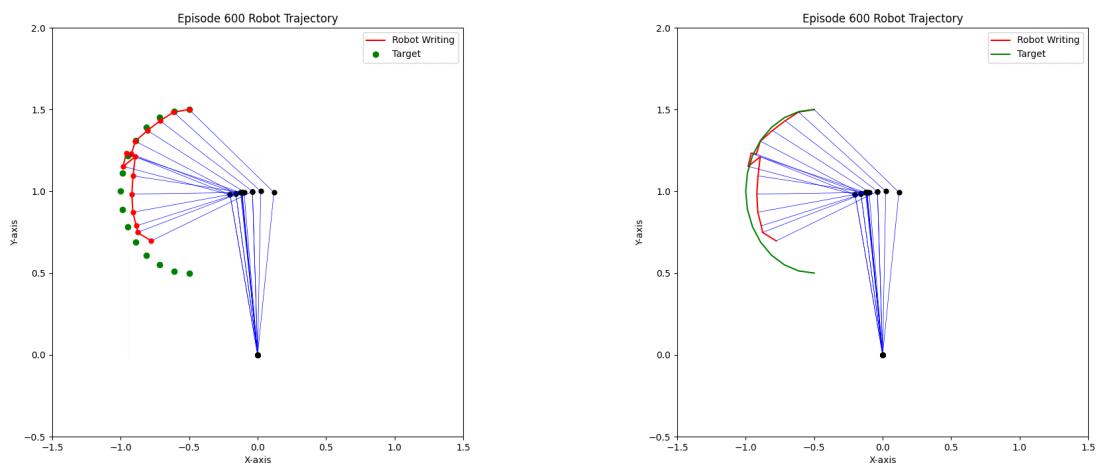


Figure 20: Q-Learning, Letter "C", Episode 600

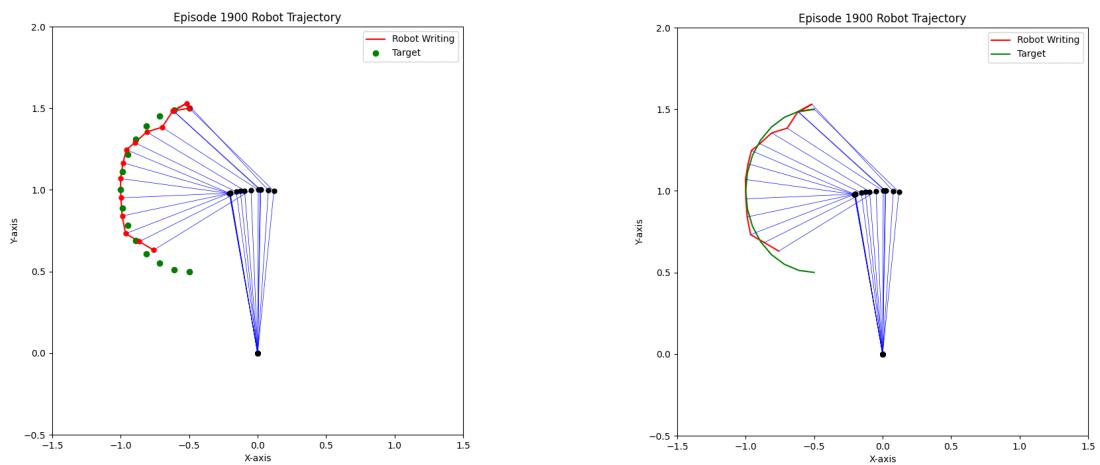


Figure 21: Q-Learning, Letter "C", Episode 1900

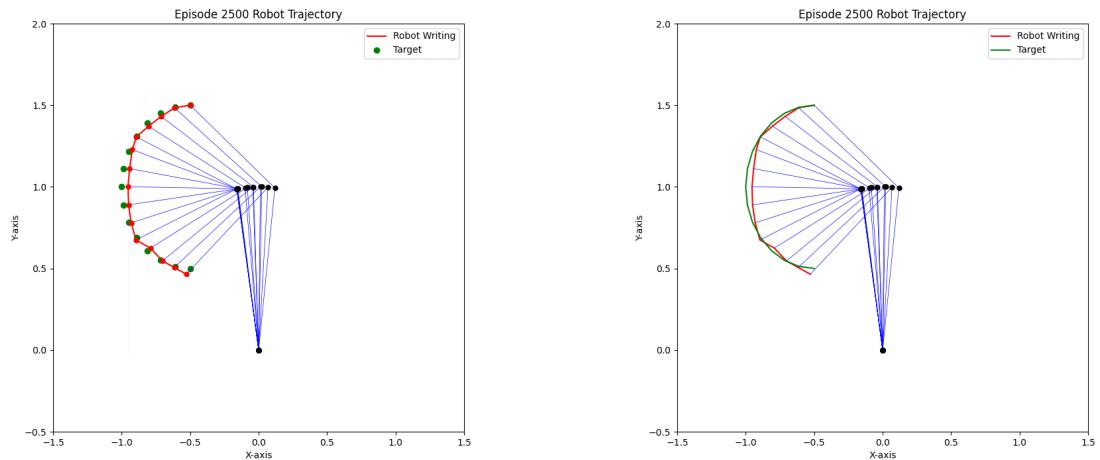


Figure 22: Q-Learning, Letter "C", Episode 2500

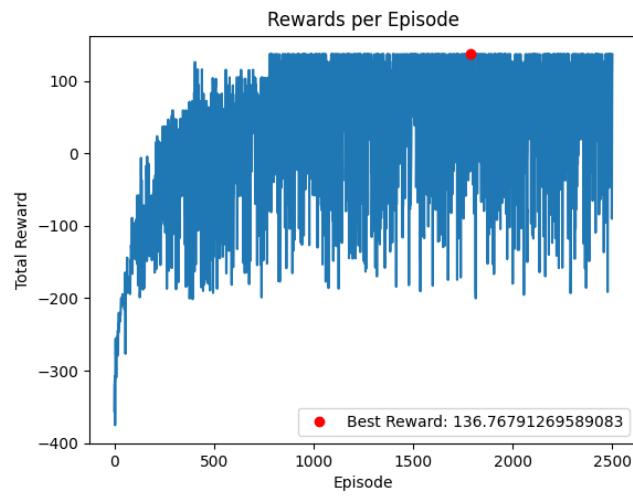


Figure 23: Q-Learning, Letter "C", Rewards per Episodes

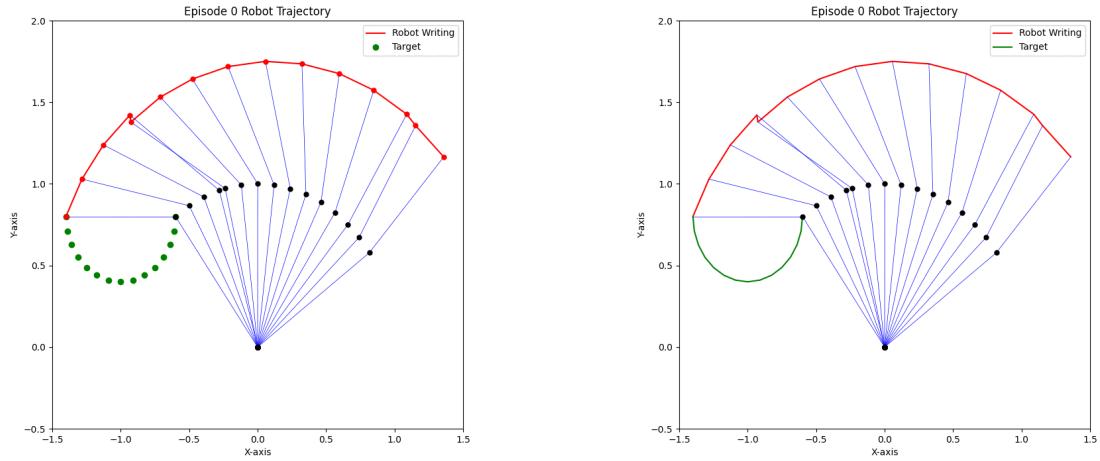
Letter "U"

Figure 24: Q-Learning, Letter "U", Episode 0

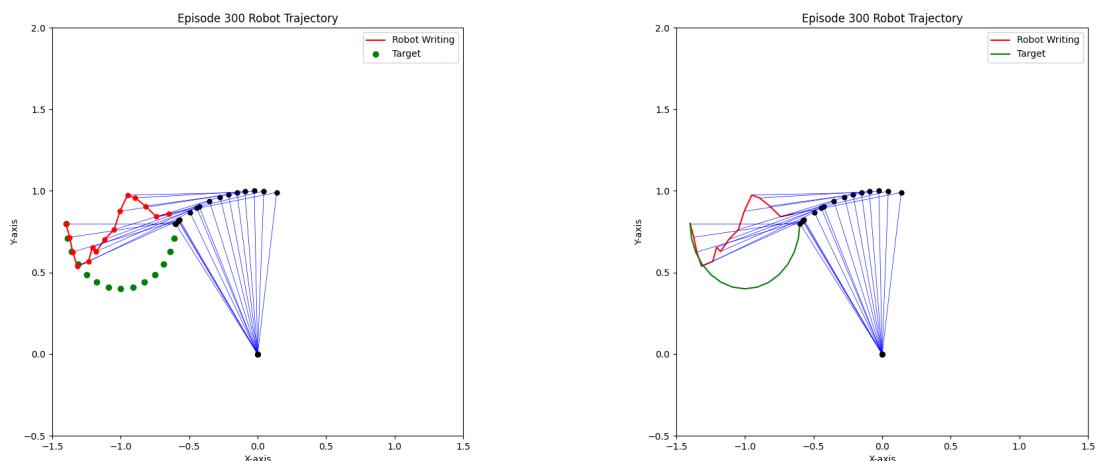


Figure 25: Q-Learning, Letter "U", Episode 300

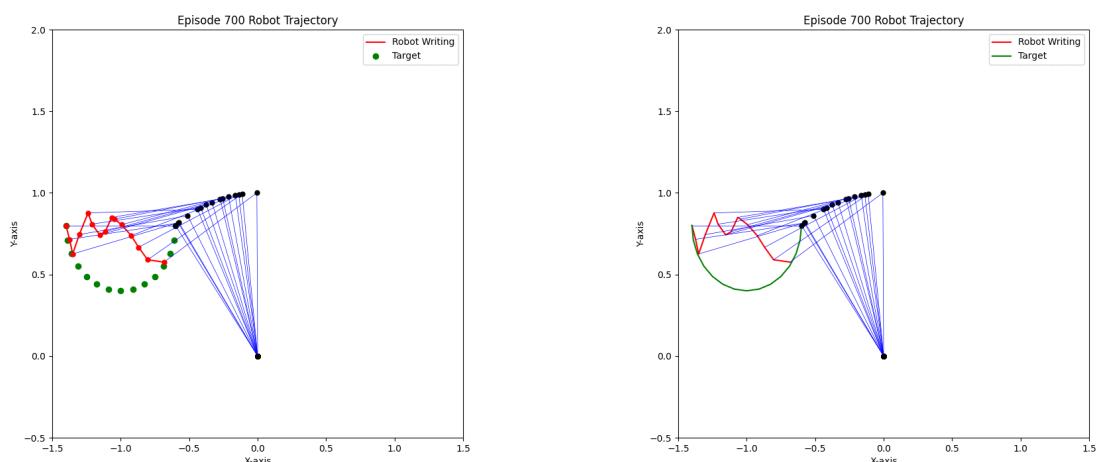


Figure 26: Q-Learning, Letter "U", Episode 700

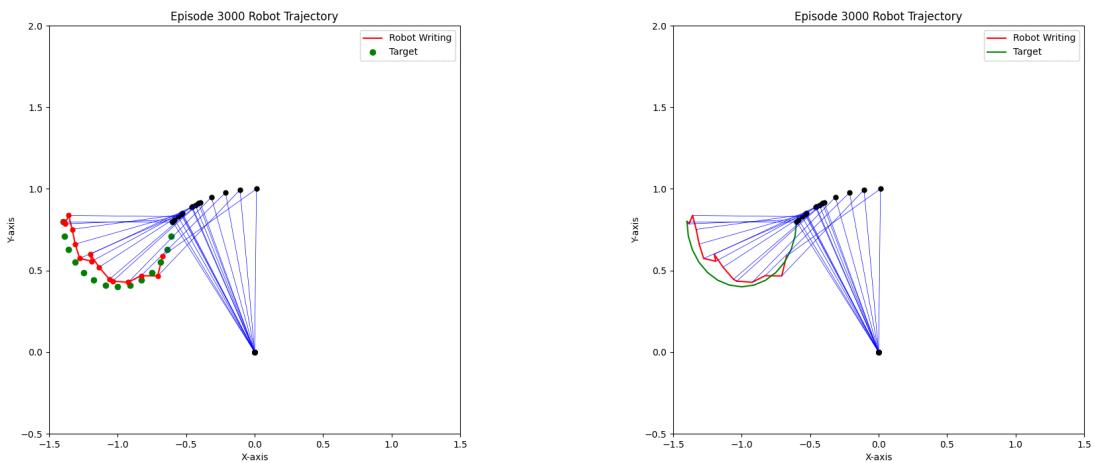


Figure 27: Q-Learning, Letter "U", Episode 3000

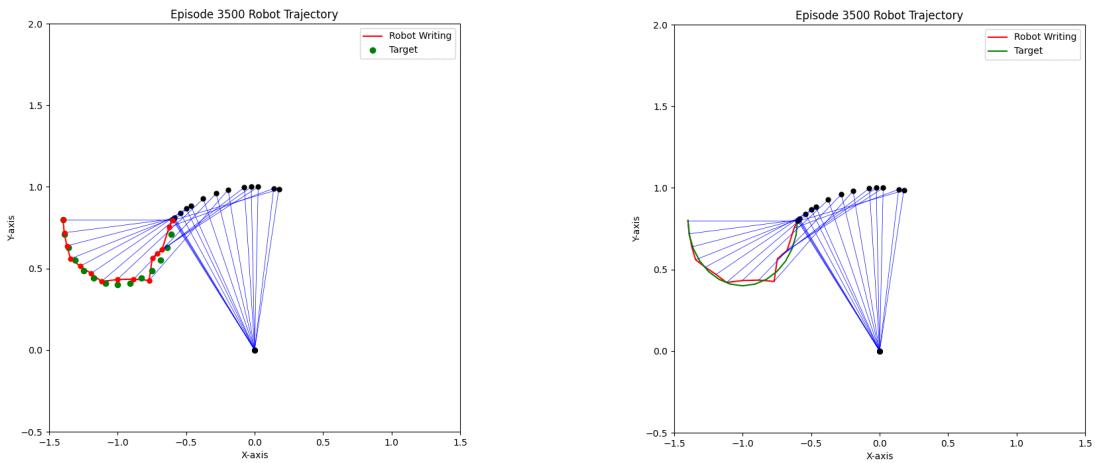


Figure 28: Q-Learning, Letter "U", Episode 3500

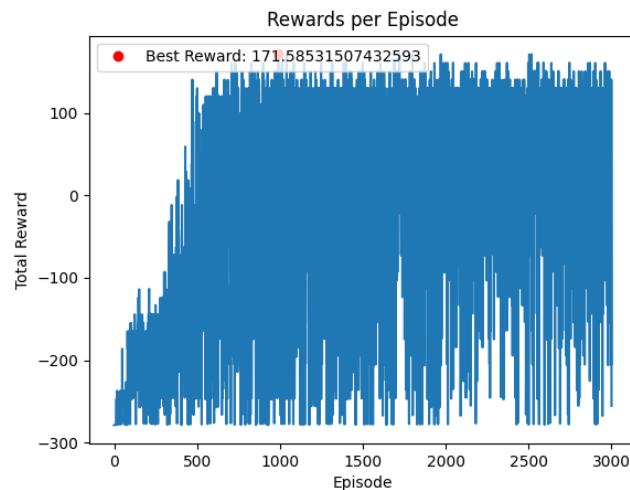


Figure 29: Q-Learning, Letter "U", Rewards per Episodes

Sinewave

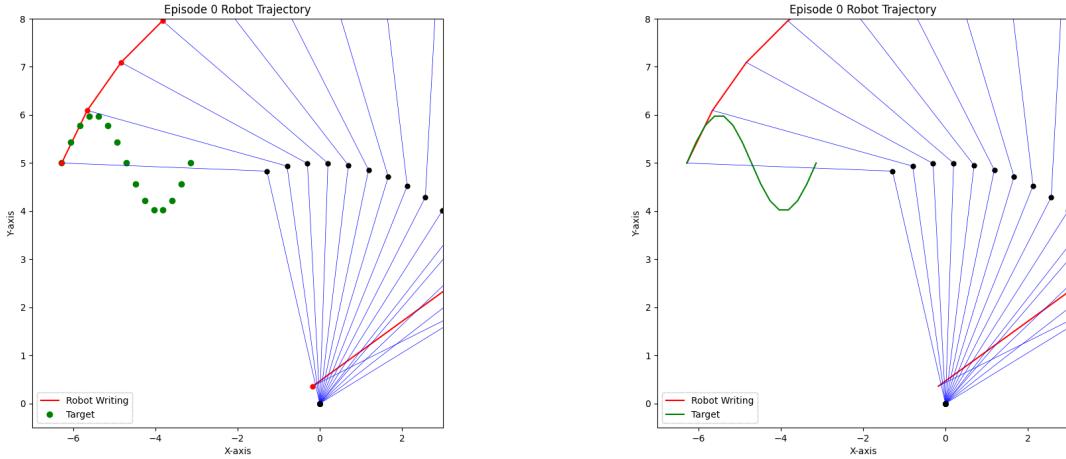


Figure 30: Q-Learning, Sinewave, Episode 0

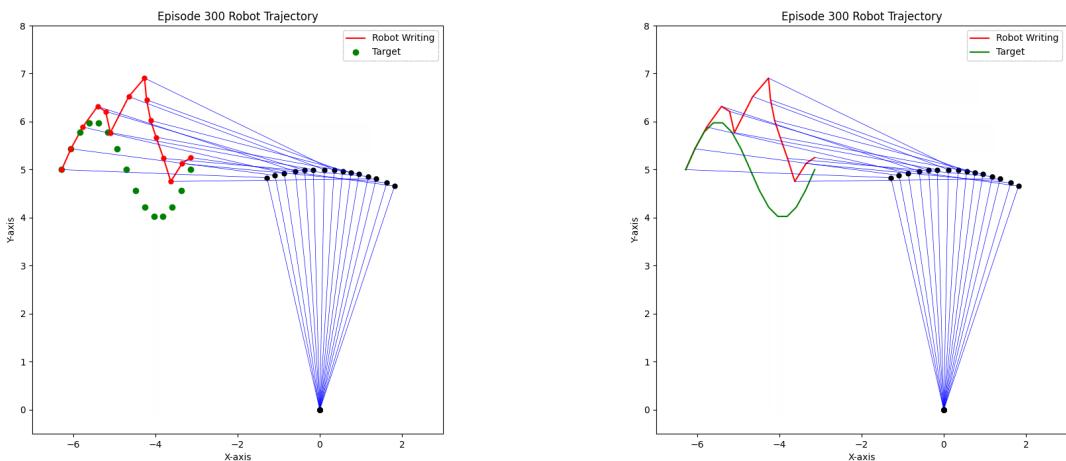


Figure 31: Q-Learning, Sinewave, Episode 300

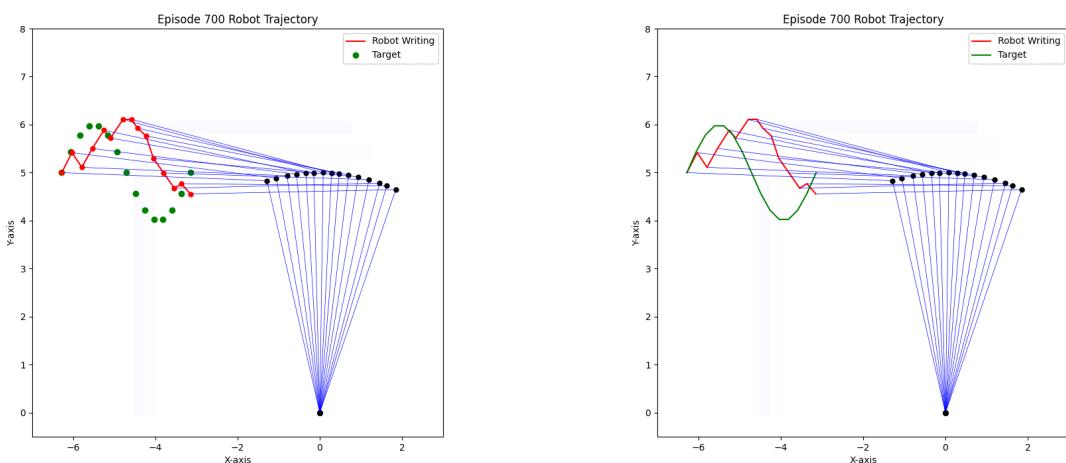


Figure 32: Q-Learning, Sinewave, Episode 700

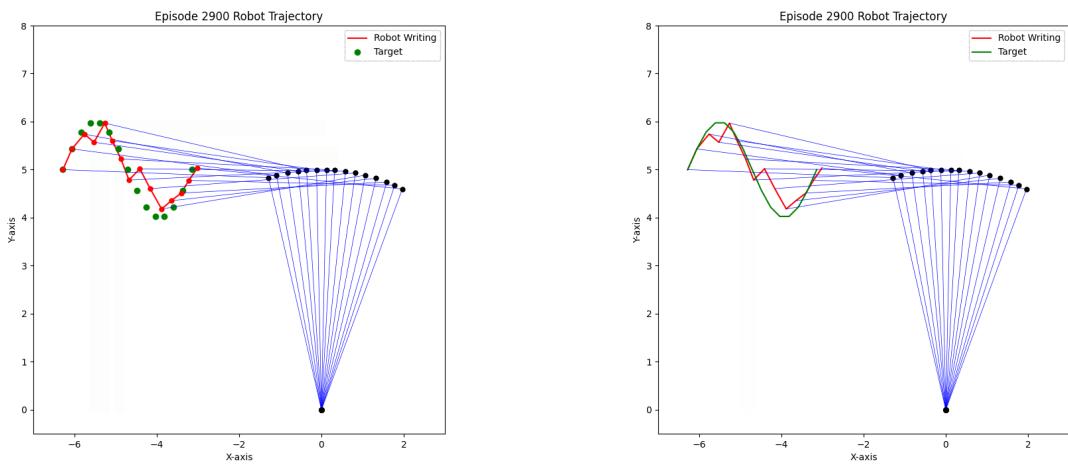


Figure 33: Q-Learning, Sinewave, Episode 2900

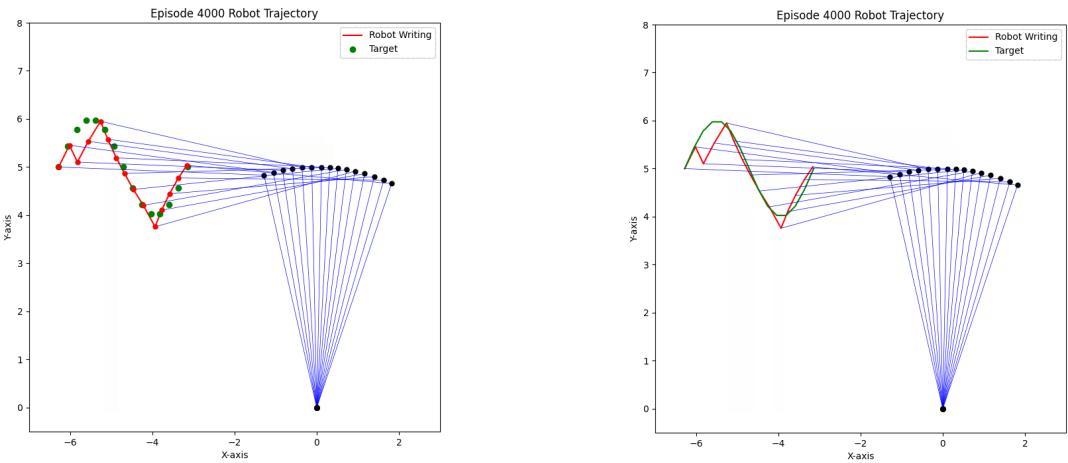


Figure 34: Q-Learning, Sinewave, Episode 4000

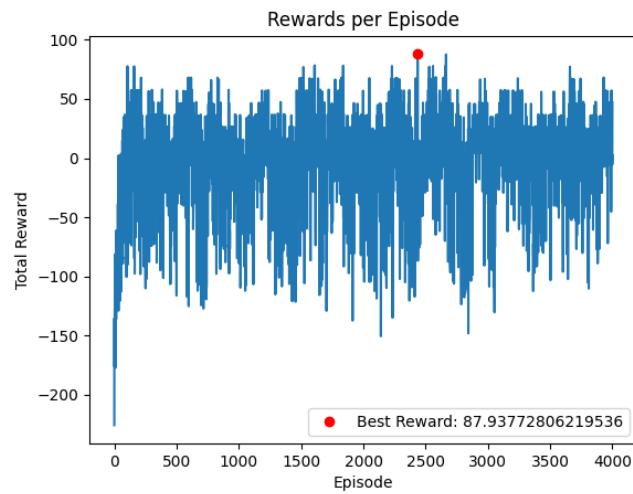


Figure 35: Q-Learning, Sinewave, Rewards per Episodes

Letter "O"

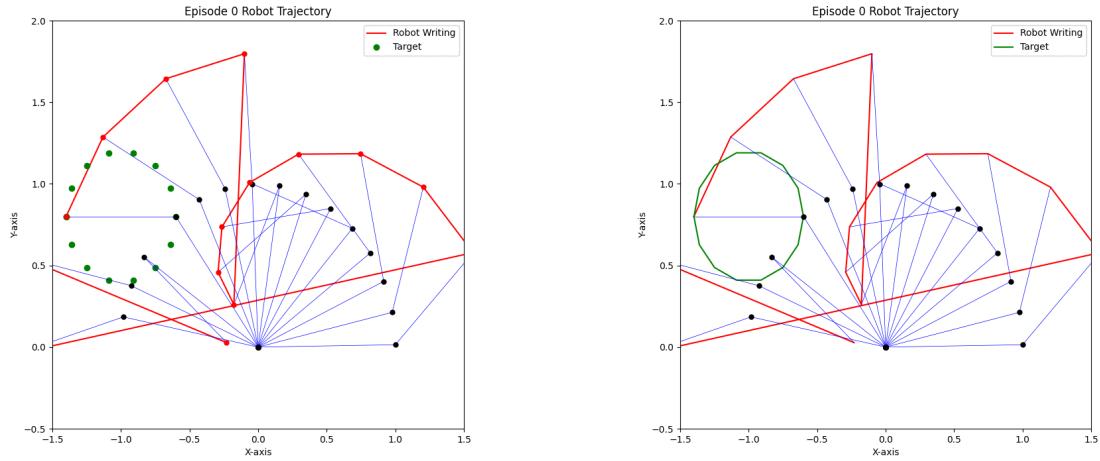


Figure 36: Q-Learning, Letter "O", Episode 0

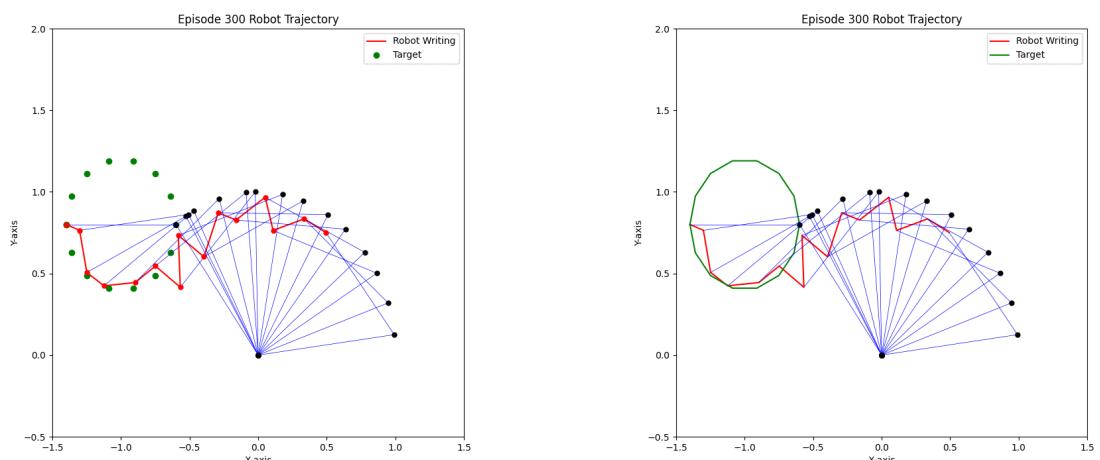


Figure 37: Q-Learning, Letter "O", Episode 300

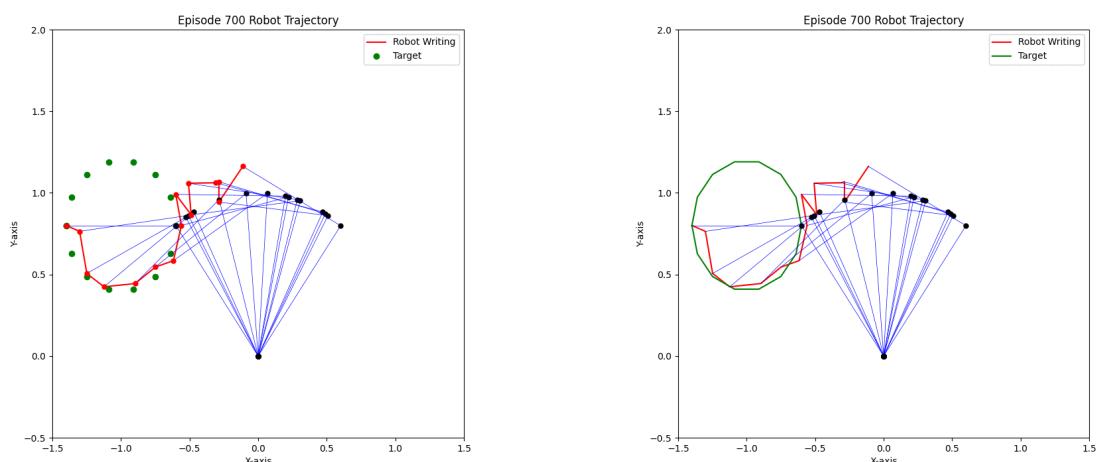


Figure 38: Q-Learning, Letter "O", Episode 700

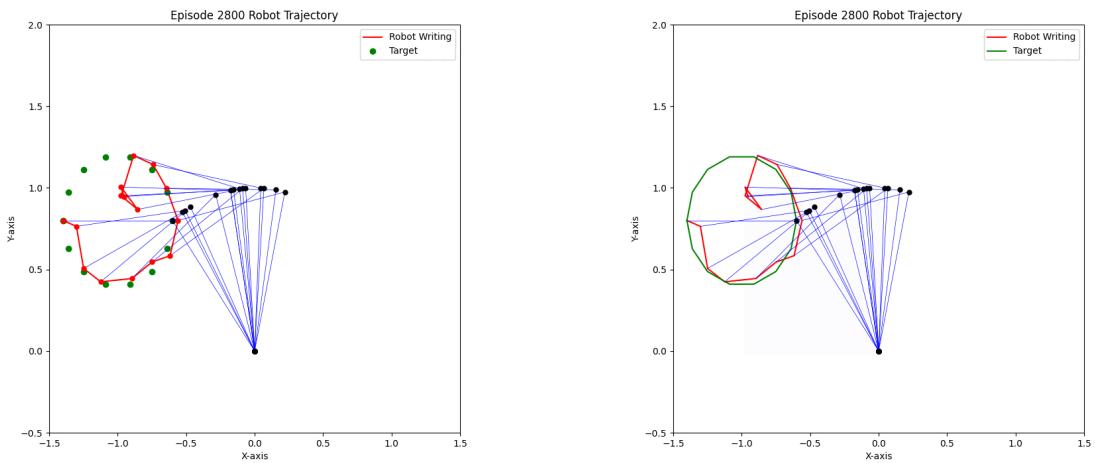


Figure 39: Q-Learning, Letter "O", Episode 2800

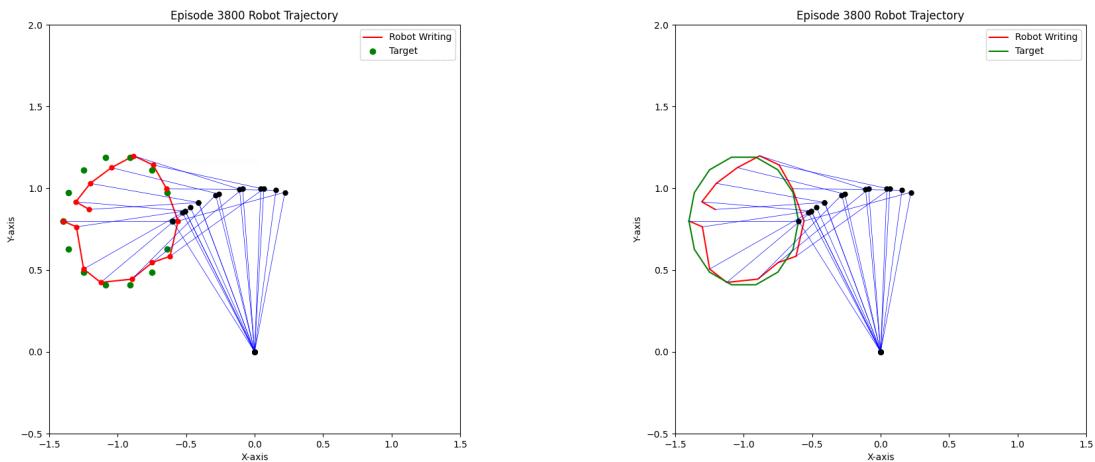


Figure 40: Q-Learning, Letter "O", Episode 3800

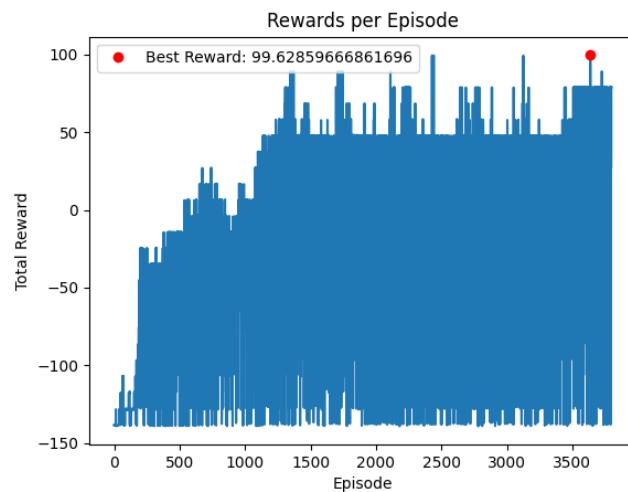


Figure 41: Q-Learning, Letter "O", Rewards per Episodes

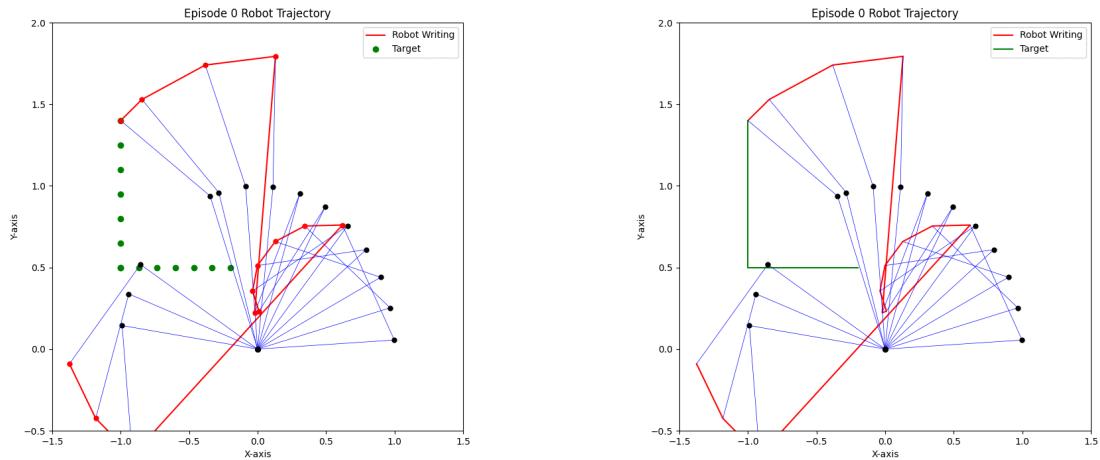
Letter "L"

Figure 42: Q-Learning, Letter "L", Episode 0

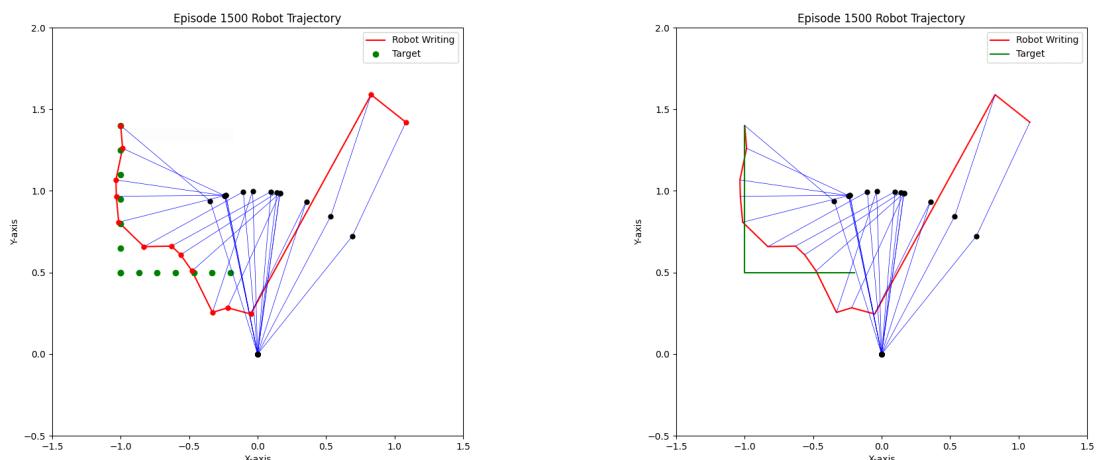


Figure 43: Q-Learning, Letter "L", Episode 1500

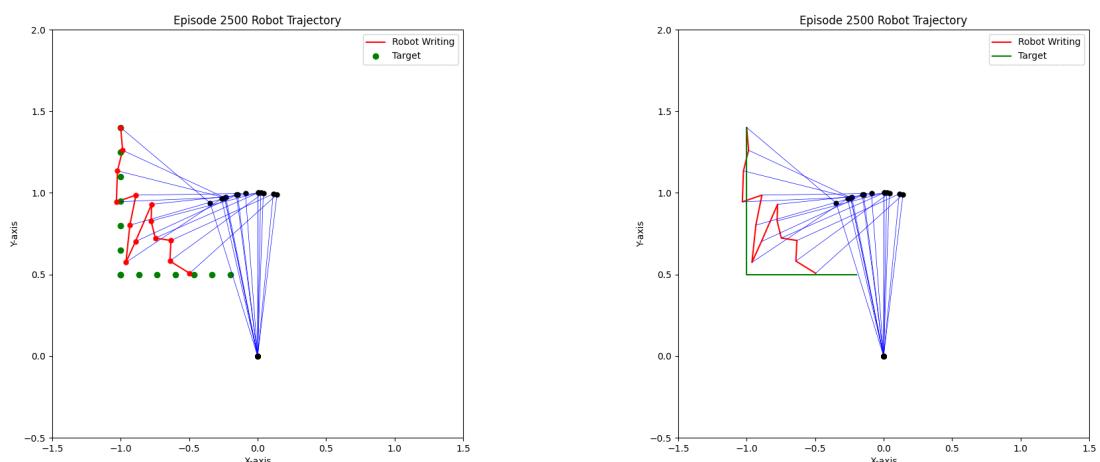


Figure 44: Q-Learning, Letter "L", Episode 2500

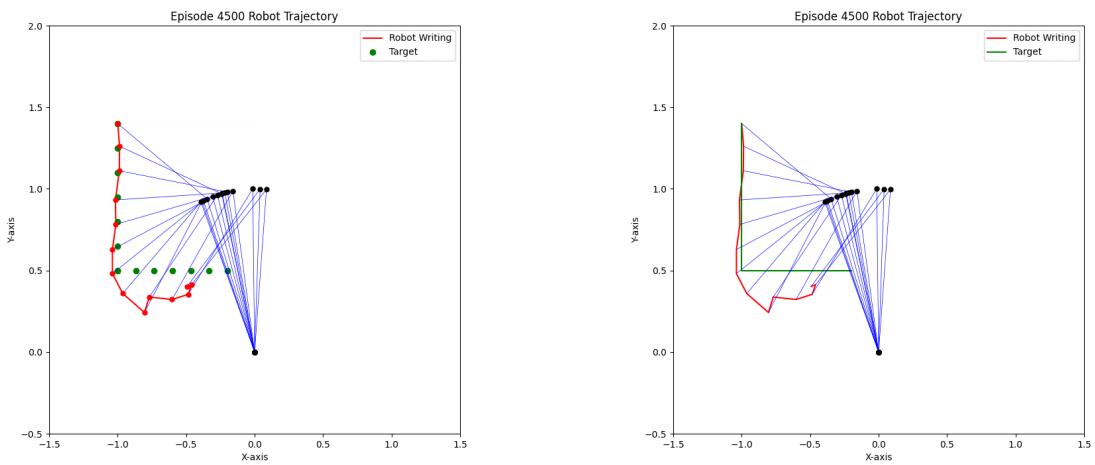


Figure 45: Q-Learning, Letter "L", Episode 4500

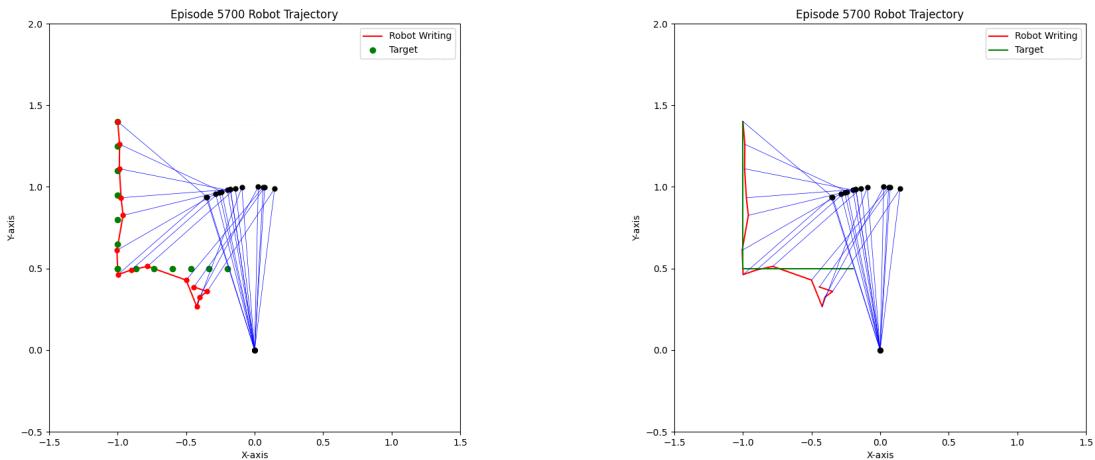


Figure 46: Q-Learning, Letter "L", Episode 5700

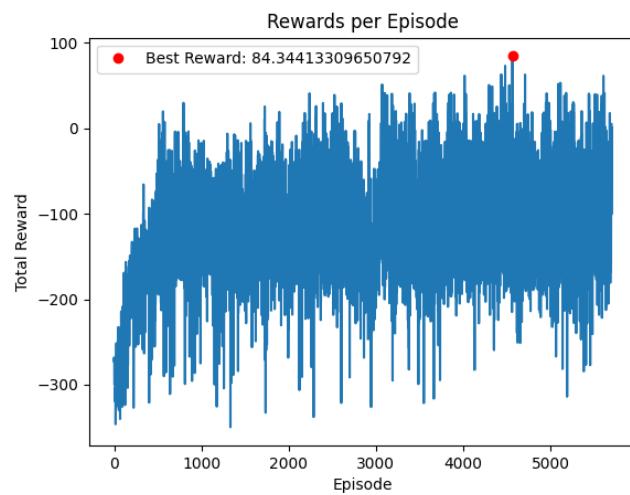


Figure 47: Q-Learning, Letter "L", Rewards per Episodes

4.1.2 Deep Q-Learning Results

Letter "C"

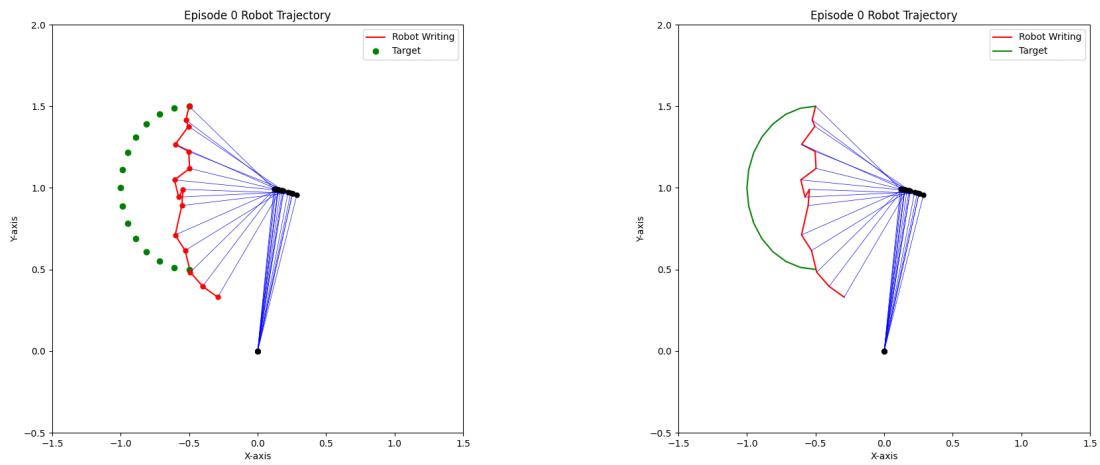


Figure 48: Deep Q-Learning, Letter "C", Episode 0

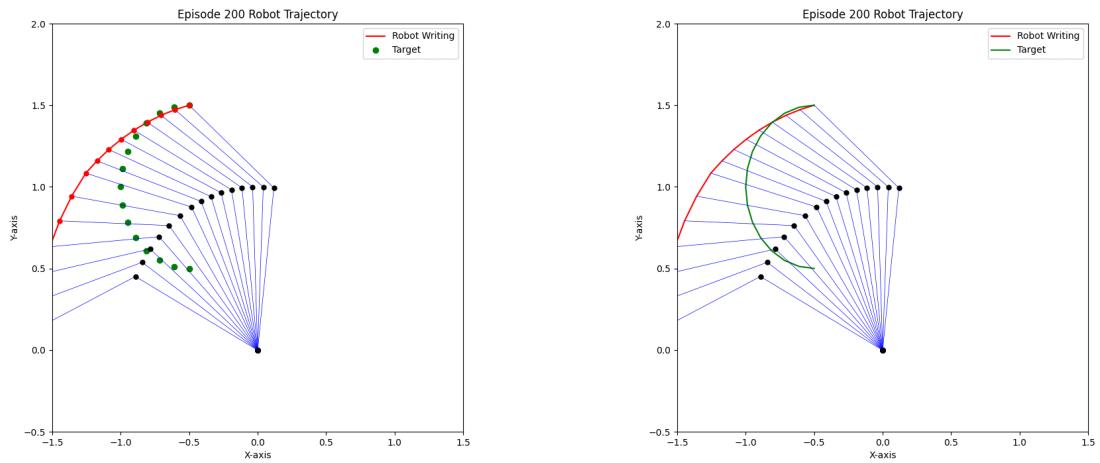


Figure 49: Deep Q-Learning, Letter "C", Episode 200

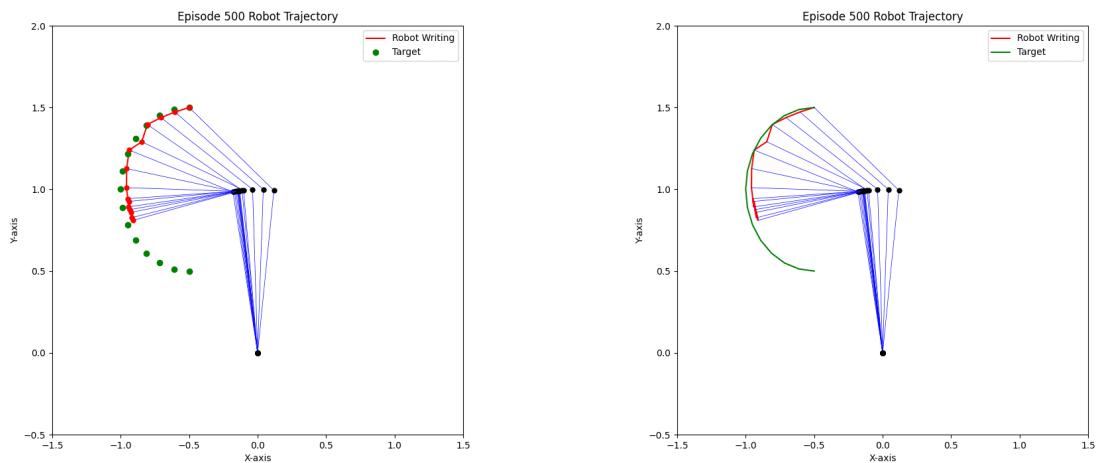


Figure 50: Deep Q-Learning, Letter "C", Episode 500

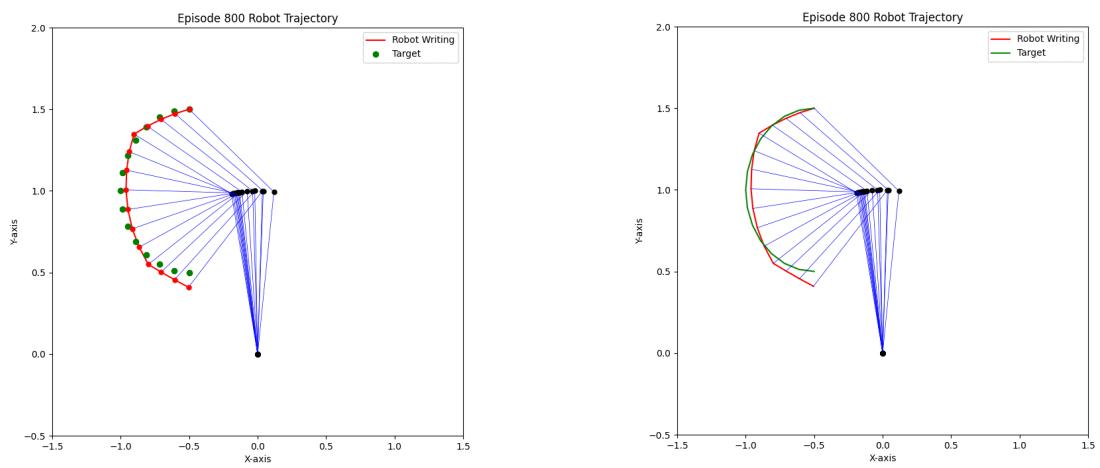


Figure 51: Deep Q-Learning, Letter "C", Episode 800

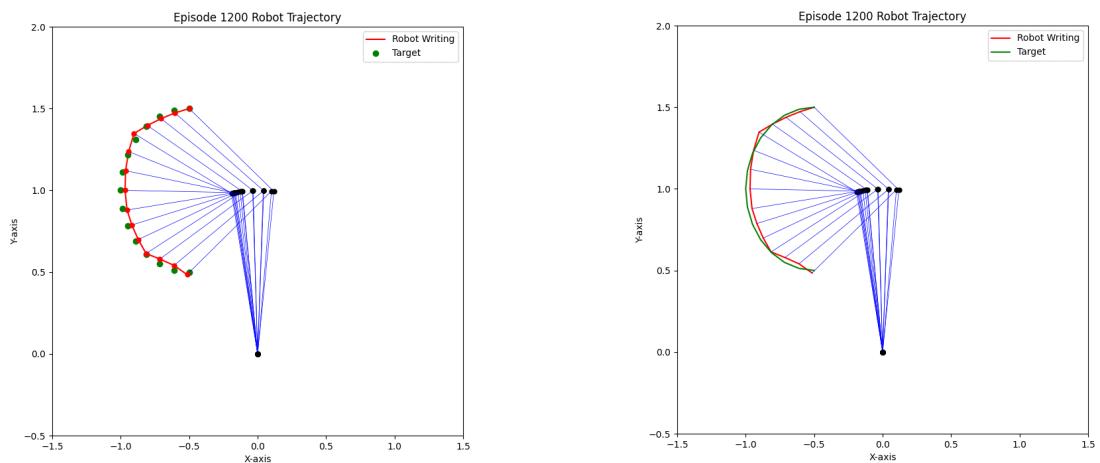


Figure 52: Deep Q-Learning, Letter "C", Episode 1200

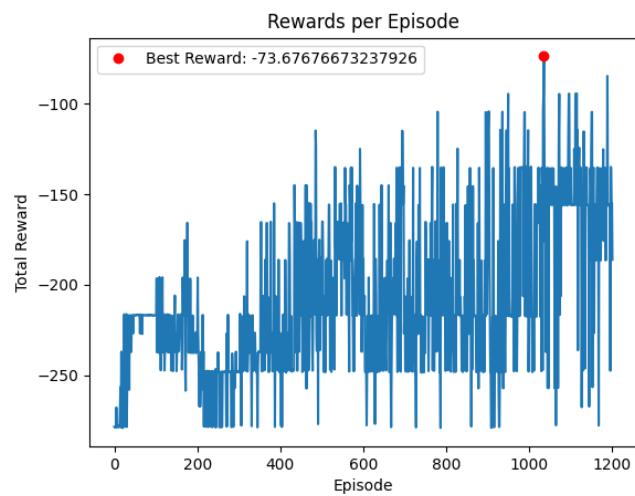


Figure 53: Deep Q-Learning, Letter "C", Rewards per Episodes

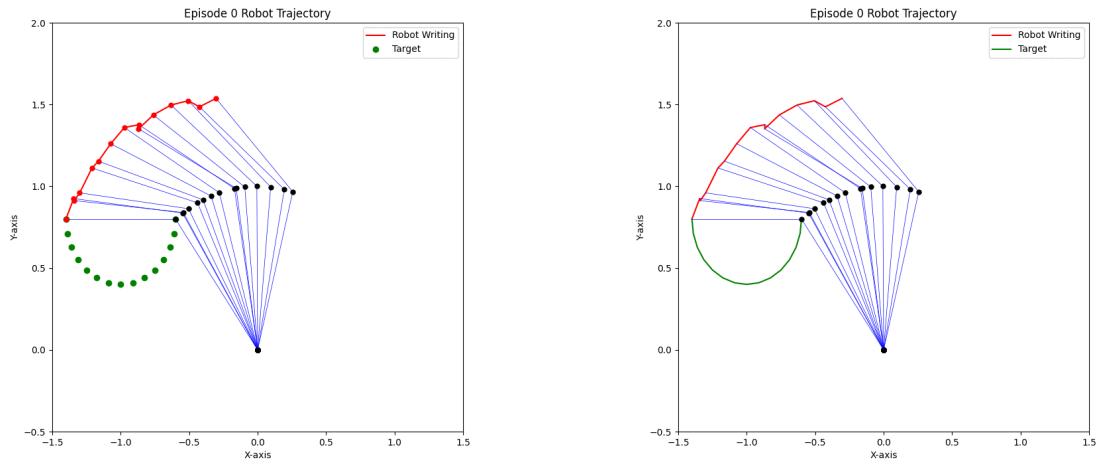
Letter "U"

Figure 54: Deep Q-Learning, Letter "U", Episode 0

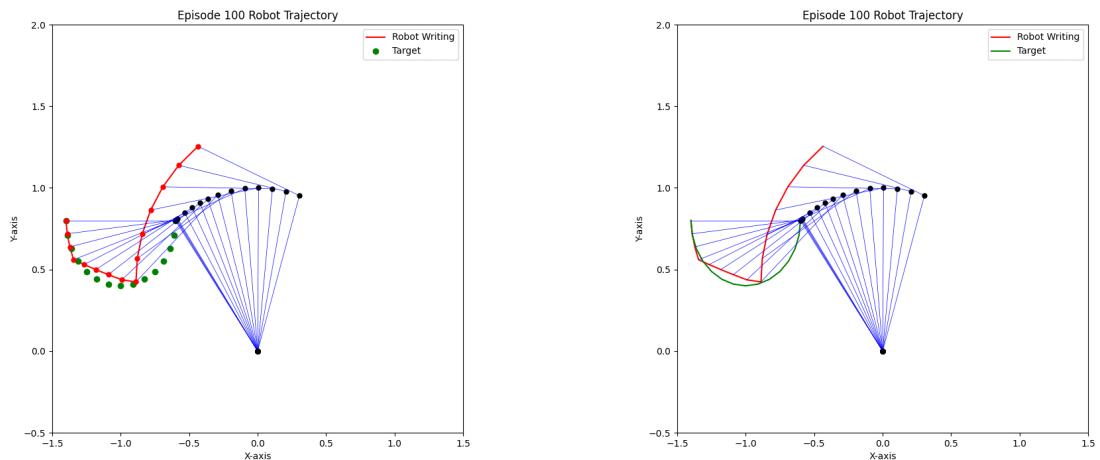


Figure 55: Deep Q-Learning, Letter "U", Episode 100

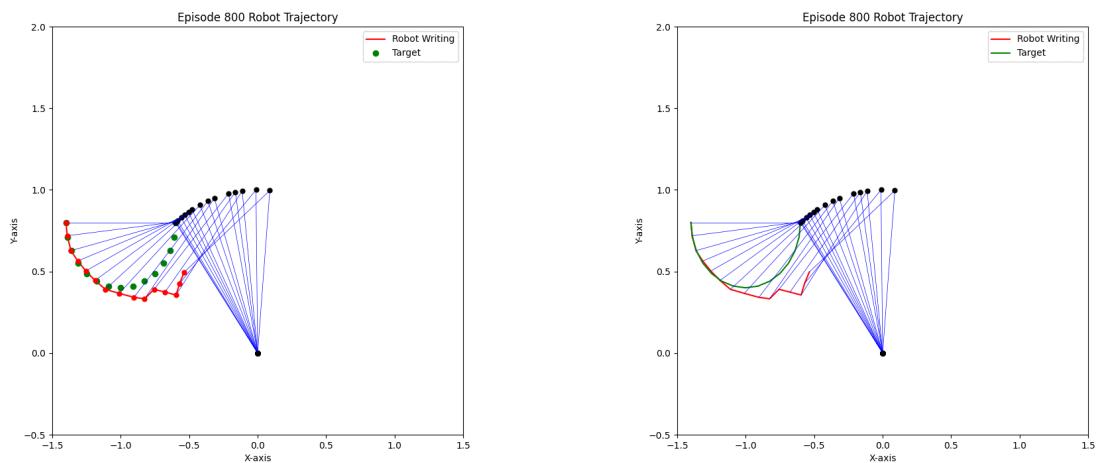


Figure 56: Deep Q-Learning, Letter "U", Episode 800

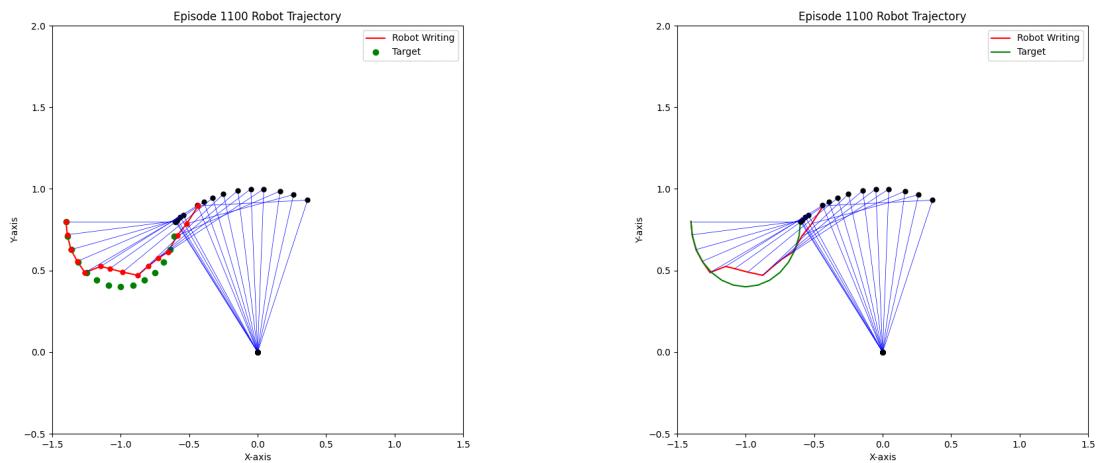


Figure 57: Deep Q-Learning, Letter "U", Episode 1100

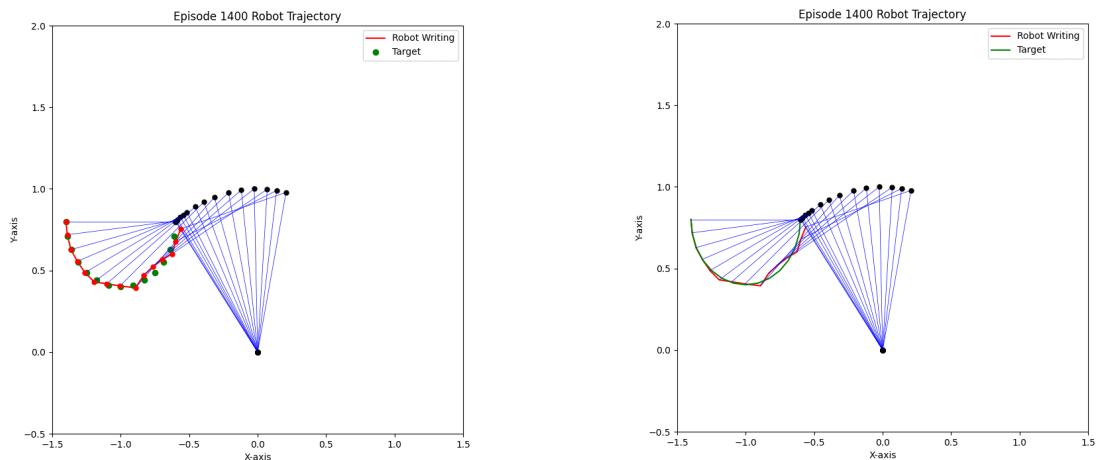


Figure 58: Deep Q-Learning, Letter "U", Episode 1400

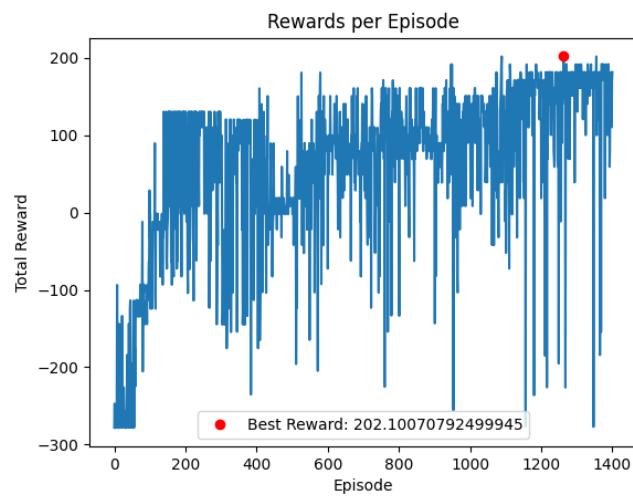


Figure 59: Deep Q-Learning, Letter "U", Rewards per Episodes

Sinewave

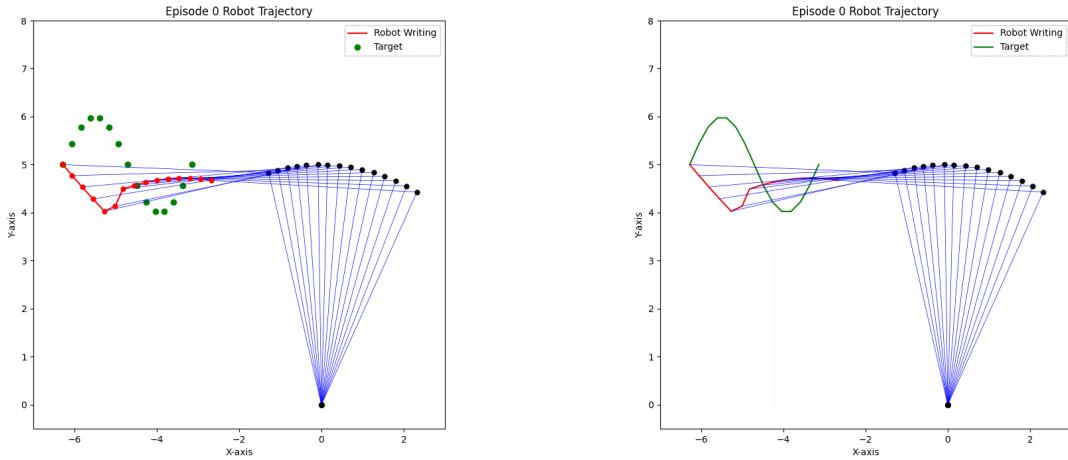


Figure 60: Deep Q-Learning, Sinewave, Episode 0

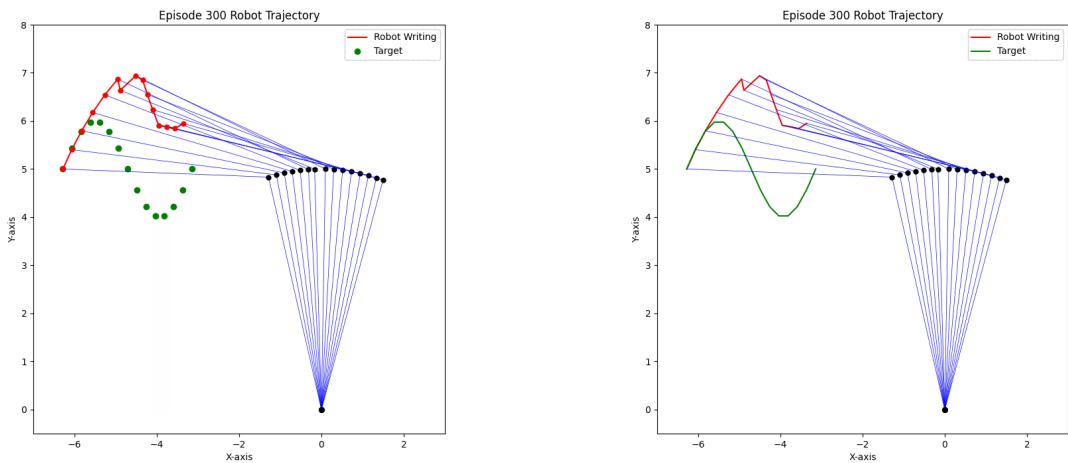


Figure 61: Deep Q-Learning, Sinewave, Episode 300

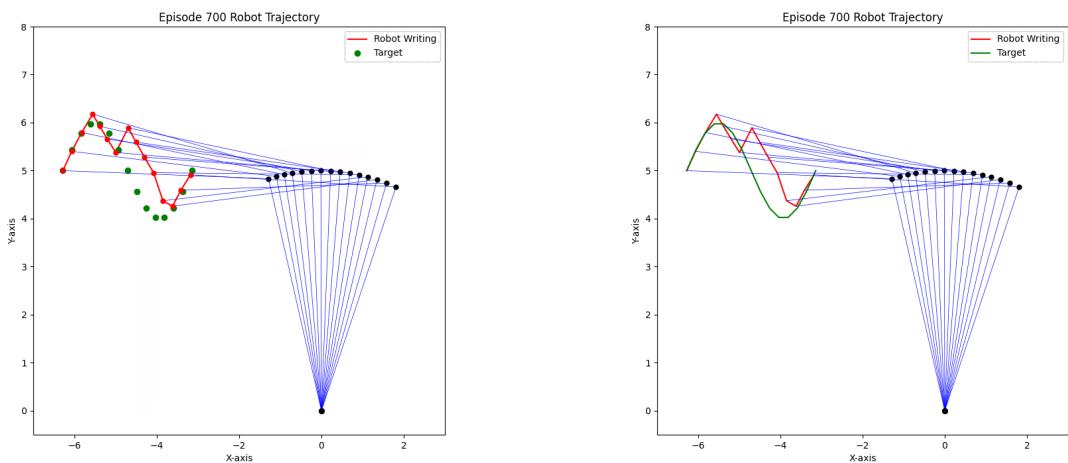


Figure 62: Deep Q-Learning, Sinewave, Episode 700

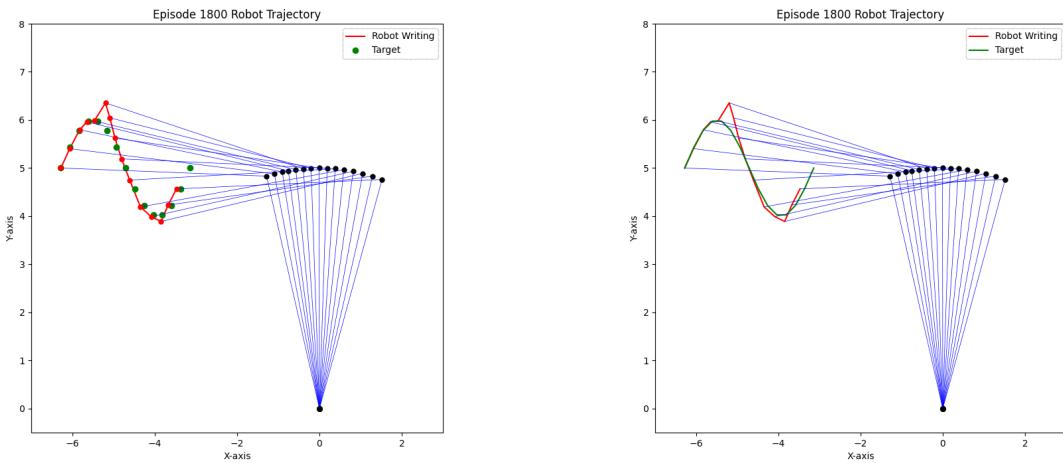


Figure 63: Deep Q-Learning, Sinewave, Episode 1800

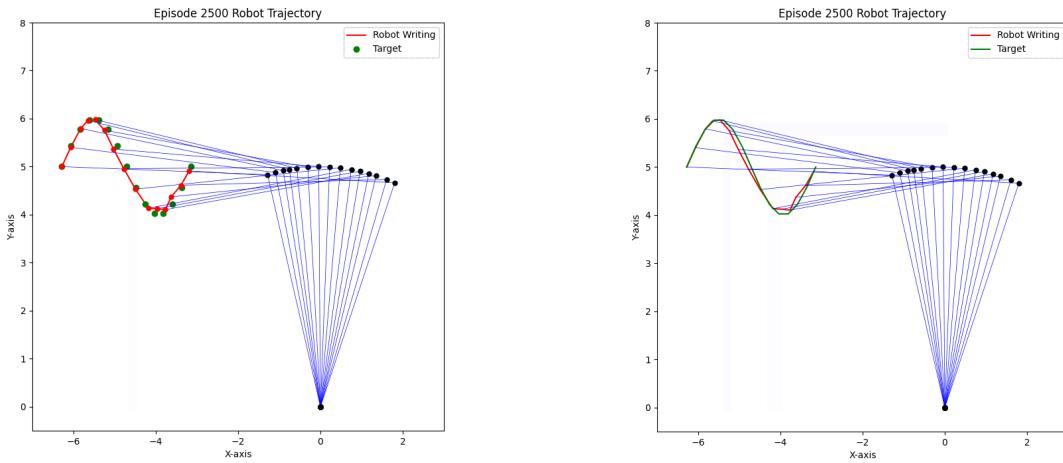


Figure 64: Deep Q-Learning, Sinewave, Episode 2500

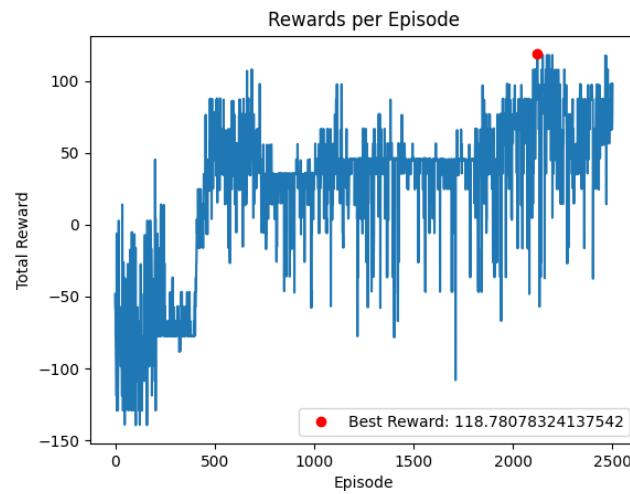


Figure 65: Deep Q-Learning, Sinewave, Rewards per Episodes

Letter "O"

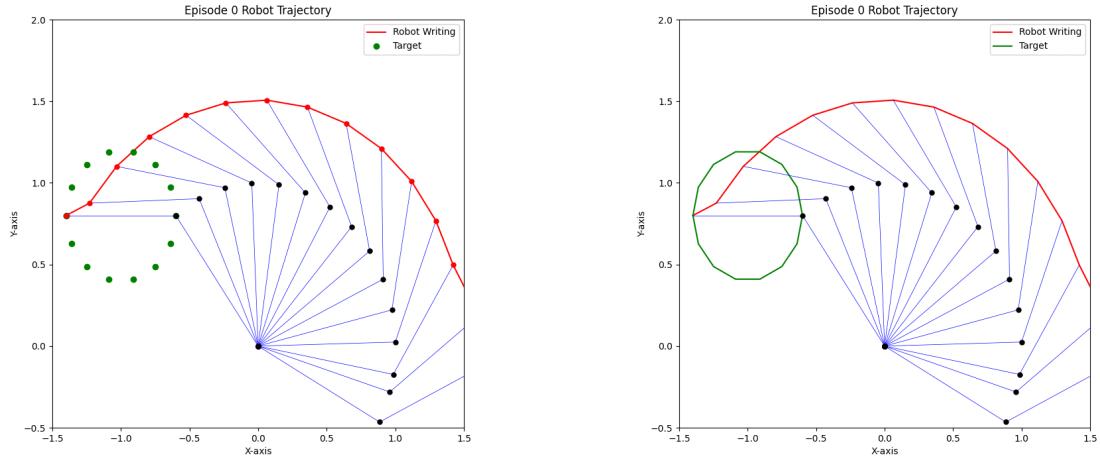


Figure 66: Deep Q-Learning, Letter "O", Episode 0

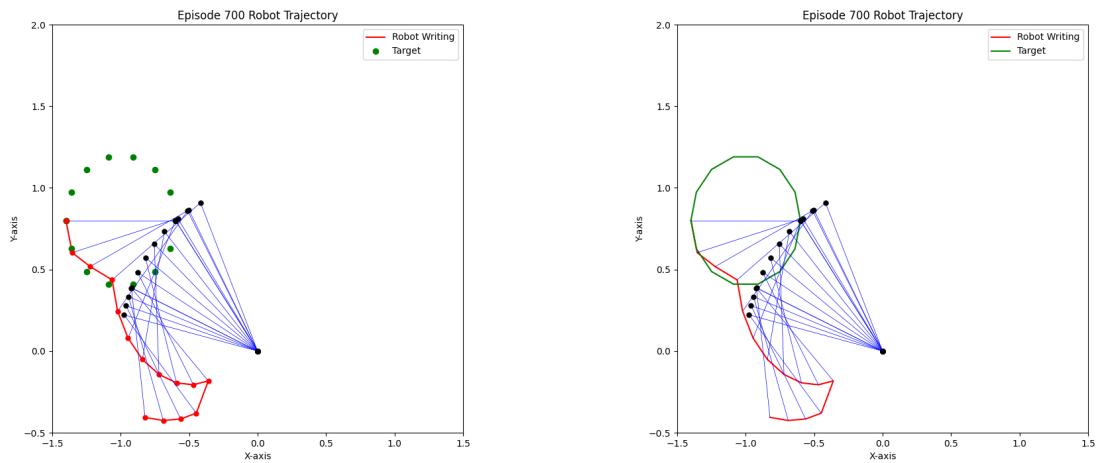


Figure 67: Deep Q-Learning, Letter "O", Episode 700

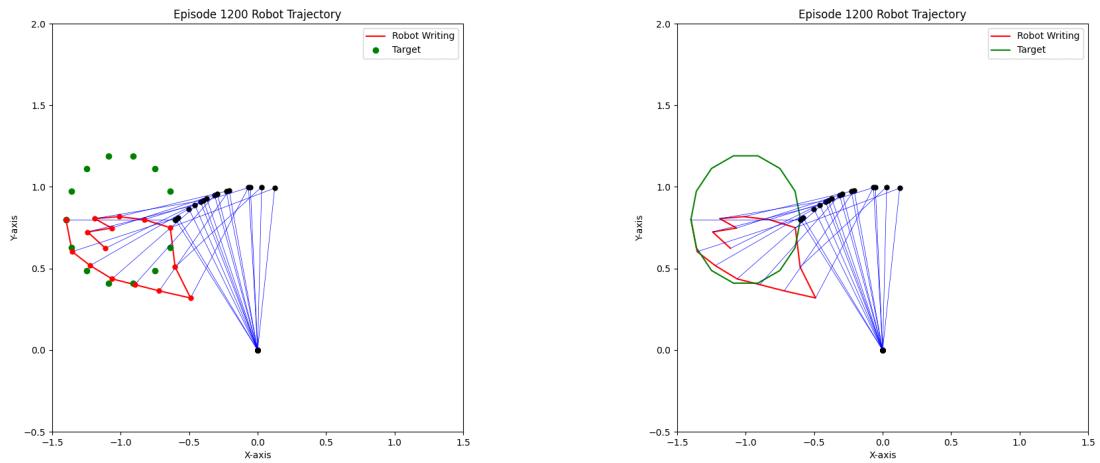


Figure 68: Deep Q-Learning, Letter "O", Episode 1200

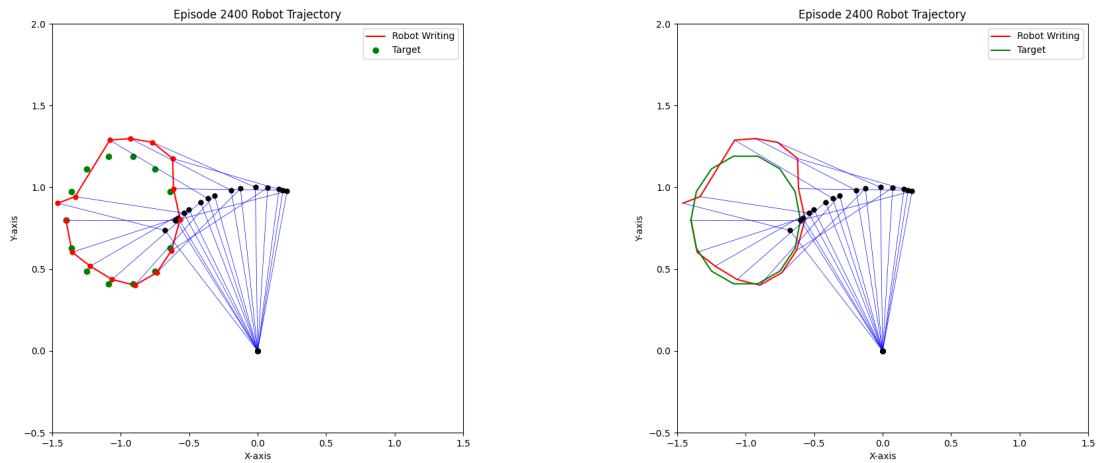


Figure 69: Deep Q-Learning, Letter "O", Episode 2400

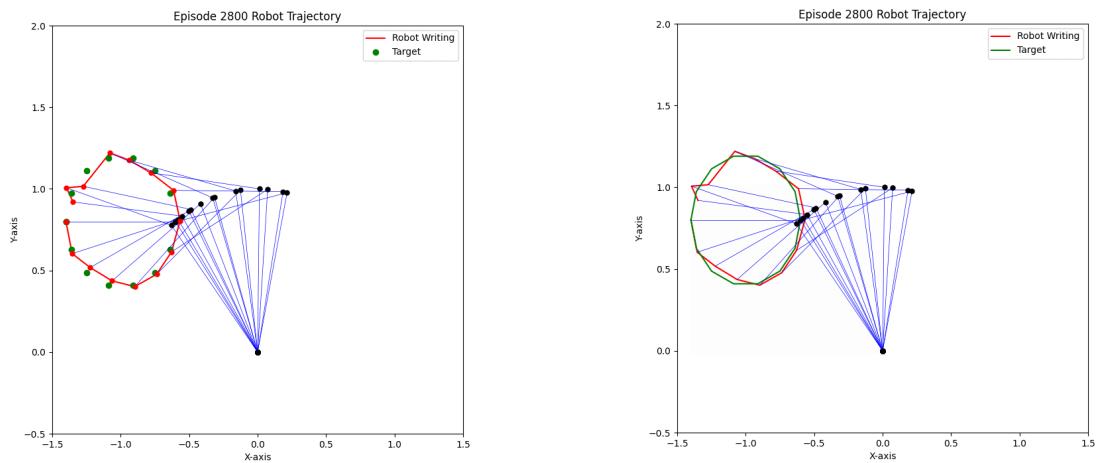


Figure 70: Deep Q-Learning, Letter "O", Episode 2800

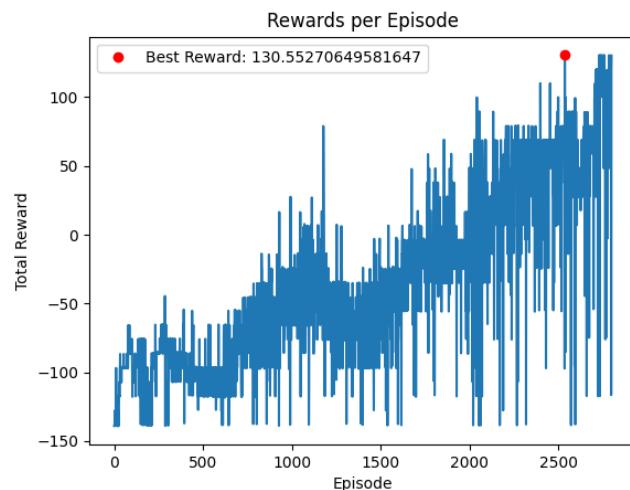


Figure 71: Deep Q-Learning, Letter "O", Rewards per Episodes

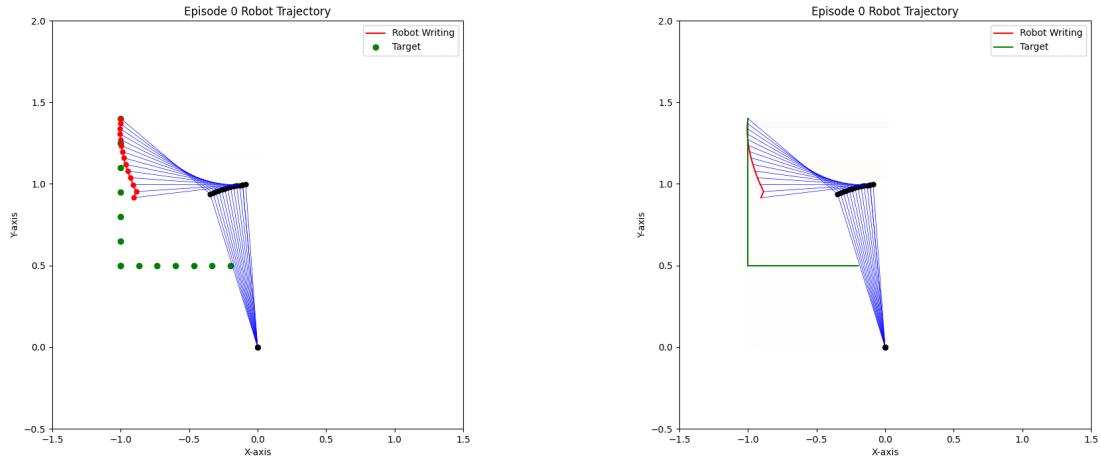
Letter "L"

Figure 72: Deep Q-Learning, Letter "L", Episode 0

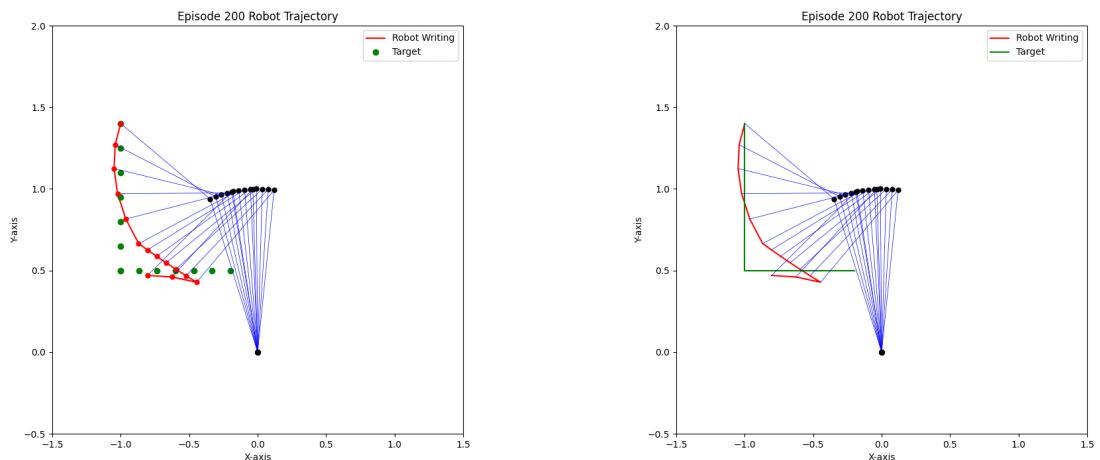


Figure 73: Deep Q-Learning, Letter "L", Episode 200

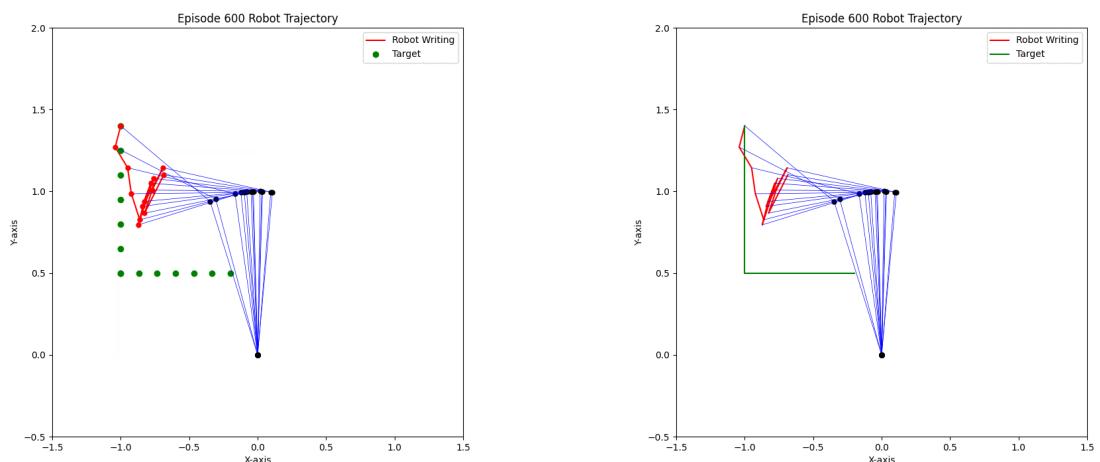


Figure 74: Deep Q-Learning, Letter "L", Episode 600

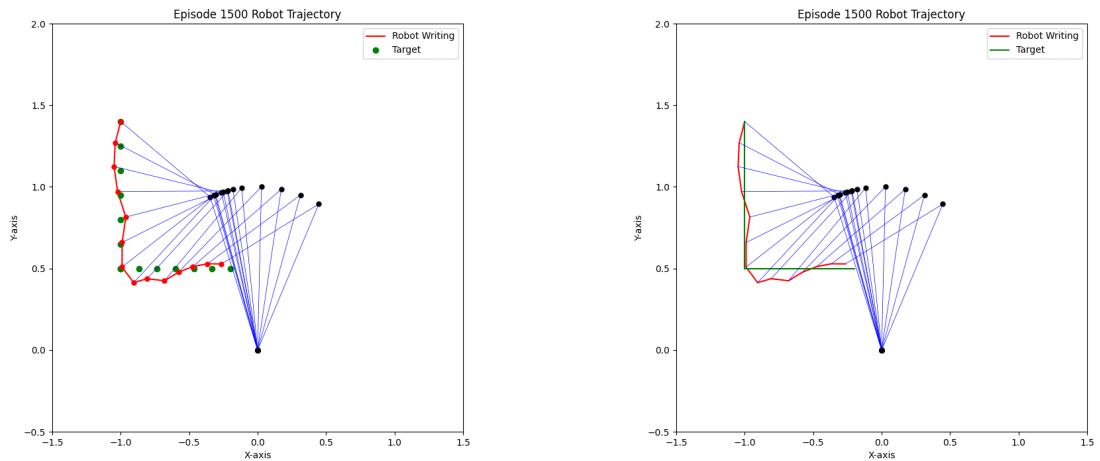


Figure 75: Deep Q-Learning, Letter "L", Episode 1500

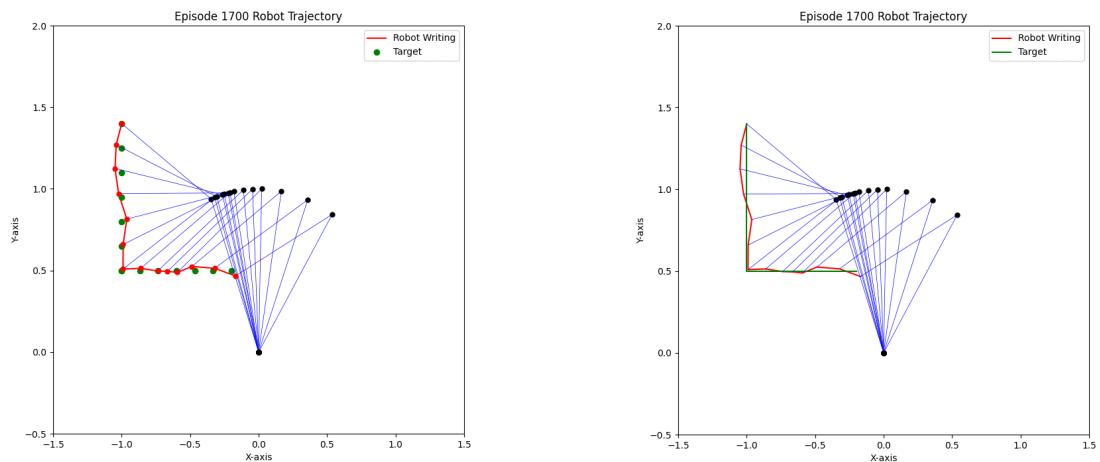


Figure 76: Deep Q-Learning, Letter "L", Episode 1700

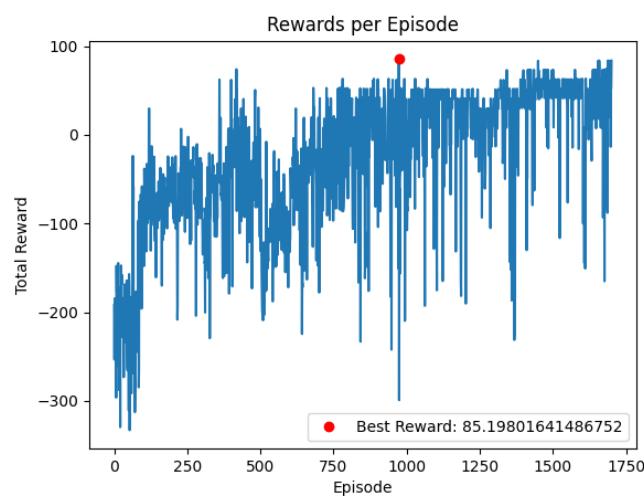
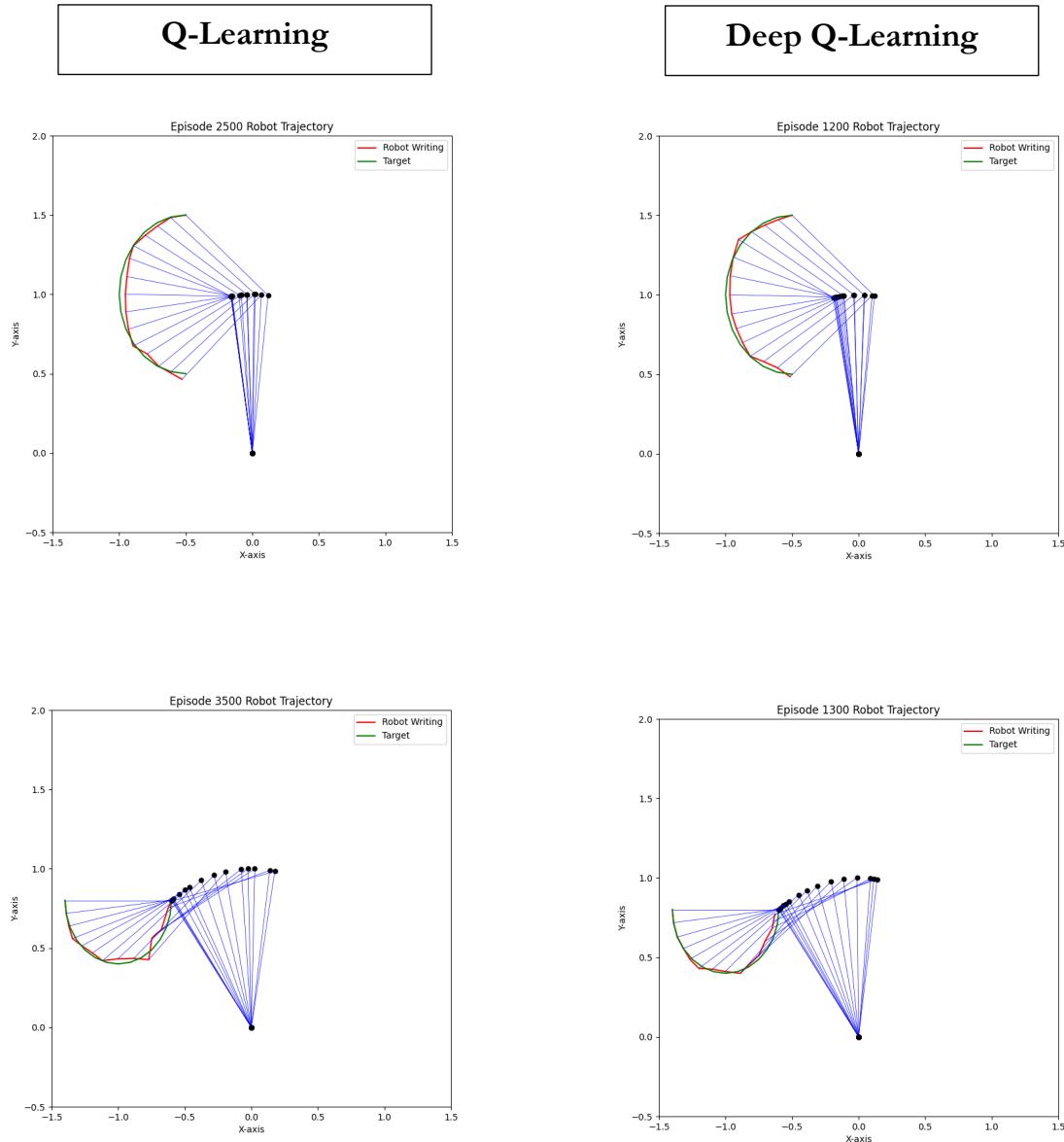


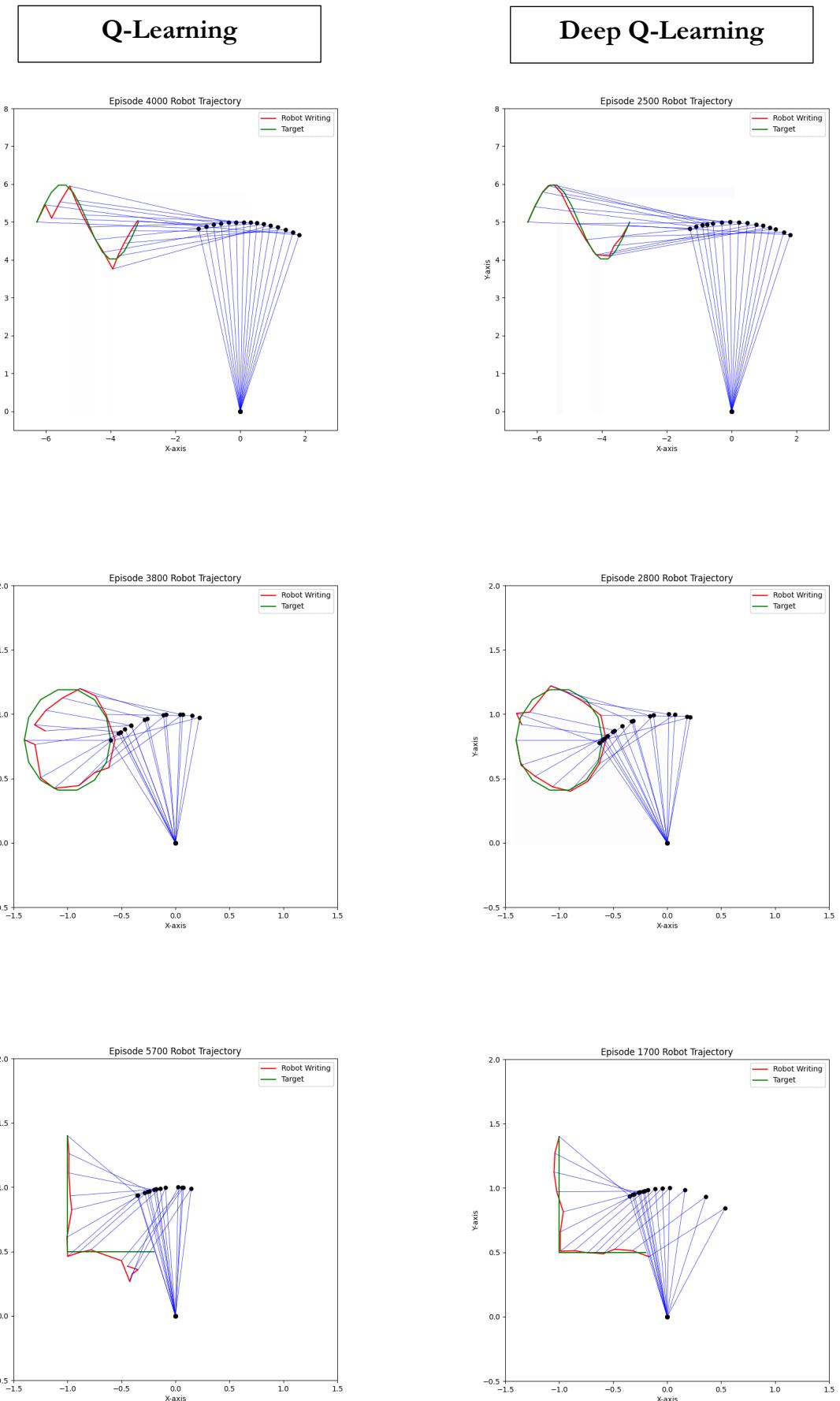
Figure 77: Deep Q-Learning, Letter "L", Rewards per Episodes

Observations & Comments

4.2 Comparing Q-Learning and Deep Q-Learning

As we can see with the above results Q-Learning algorithm is time efficient with less complexity, but as the shapes were more complex and curved it loses in accuracy. The fundamental change in finding Q values that Neural Network propose, the results were noticeably better compared to the Q-Learning ones as you can see in Letter “**O**” and “**Sinewave**”. DQN had efficient also in number of episodes. As you can see the Letter “**U**” in Q-Learning uses about 3000 and “**C**” about 2500 episodes to approximate the desired trajectory while with Deep Q Network only 1400 and 1200 needed respectively.





4.3 Challenges of Value Based Approaches

The challenging part of using Value-Based Reinforcement Learning such as Q-Learning and Deep Q-Learning compared to other approaches is the discretization of the action and state space. The simple Q-Learning method works with discrete actions and discrete states. So, using this method for a 2DOF robot who has continuous action sets so as states, is quite challenging. Although using DQN we managed to use continuous state space, the actions must remain discrete. Trying to achieve the desired results using discrete action and state space made our approach either too time consuming or with pre-known moves that robot should do to make the action set more accurate which means less holistic approach in terms of learning process. That means that using Value-Based approaches doesn't always go with time efficiency and less complexity, but also with a sensitive tuning in discretization and also in hyperparameters (learning rate, exploration etc).