Logan Griffin

EN.665.645.8VL.SU25

Final Project Pick and Place Agent Analysis

8/17/2025

<div align="center">Analysis</div>

The following plots can be used to visualize learning episodes and improvement over time. The plot directly below shows the average number of steps the agent took to complete the task of finding the object, picking it up, and moving it to the correct container. The step limit was 100 during the training episodes. Each dot represents the average of 25 episodes. During the early episodes, the average stayed near 75-80, as the agent was learning its purpose. This meant that many episodes were reaching the maximum number of steps without completing the task. As the agent learned and updated its Q-table, the averages started to decrease. Around the 1000th episode, the agent appeared to have learned to complete the task efficiently. The variation in steps after the 1000th episode can be explained by the random placement of the object within the 5x5 grid. Since a single container was located in each of the 4 corners, it might take more steps during one episode to complete the task if the object appeared far away from its appropriate container. The second thousand episodes hovered around an average of 10 steps.
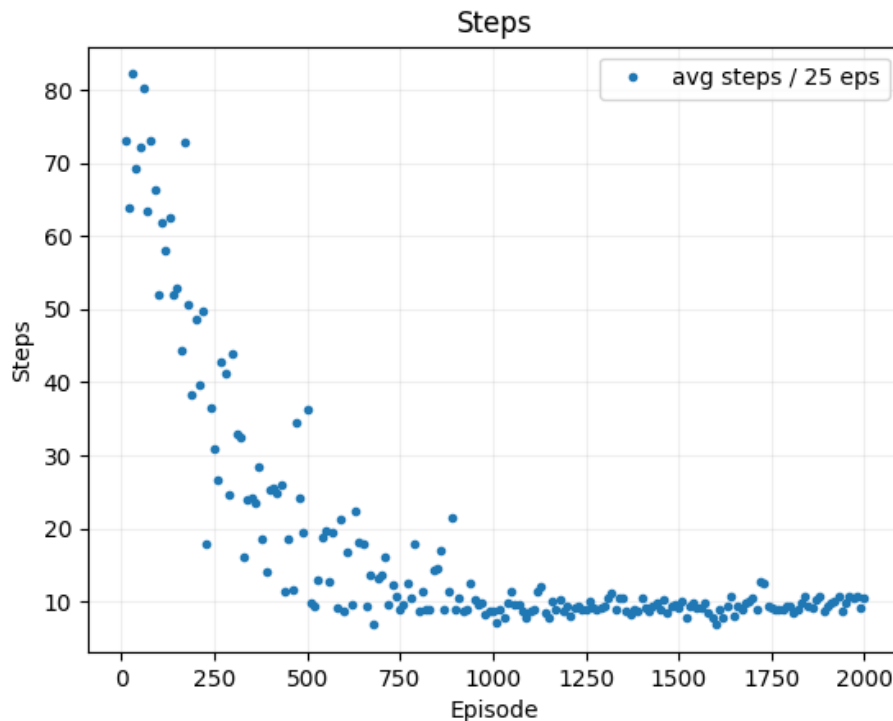
The following plot shows the average reward gained by the agent over the training period. The reward system was +1 for picking up the object, +1 for dropping the object in the correct container, 0 for a neutral move around the grid, and -1 for dropping the object in the correct container. Each dot on the plot represents the average over 25 episodes. At the beginning of the training period, the agent was mostly wandering around, not knowing its purpose. Around 150 episodes in, it figures out that its supposed to pick up the object. Between then and around the 750 episode mark, it figures out that its supposed to put the object in the correct container. After that, the plot shows that the agent knows its task and can achieve the maximum reward of 2. This shows an improvement over time in the success rate of the agent. An additional testing period was conducted after training the agent. Over 100 episodes, the agent succeeded in picking and placing the object I the correct container 100/100 times.