Reinforcement Learning (RBE595)

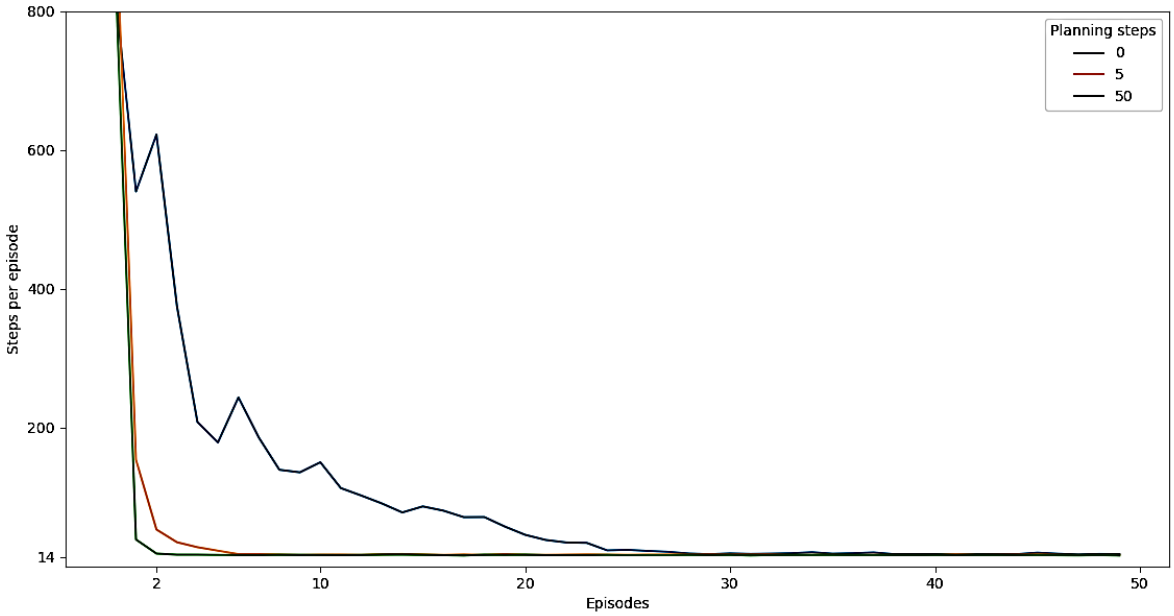
Programming Exercise 5

Model-based RL Programming exercise

Assignment By:
Bhaavin Jogeshwar (WPI ID: 888602054)
Chinmayee Prabhakar (WPI ID: 101177306)
Swapneel Dhananjay Wagholikar (WPI ID: 257598983)

In the DynaMaze Problem, an agent navigates a maze environment to reach a goal state while avoiding obstacles. The agent can take actions such as moving up, down, left, or right to navigate the maze. The task involves a trade-off between exploration and exploitation, as the agent must explore the environment to learn about the optimal policy while also exploiting its current knowledge to maximize rewards.

Output:



The figure shows average learning curves from an experiment in which Dyna-Q agents were applied to the maze task. The curves for each n display the average number of steps the agent took over 30 iterations of the experiment to complete the goal in each episode. The original seed for the random number generator was kept consistent across algorithms during each iteration. Because of this, the first episode was exactly the same for all values of n. After the first episode, performance improved for all values of n, but much more rapidly for larger values.