RL_CW

s1916575

March 2025

1 Logarithmic Decay

In Q3 we explored two different epsilon decay strategies – linear decay and exponential decay. The goal of exploration is for the agent to initially explore until it finds a good strategy and then focus on exploiting that strategy. This suggests we would want to keep the value of epsilon high initially and then decay it more quickly later in training. A log decay strategy would be suited to this, where epsilon is kept high during the initial phase of training, before beginning to decay more rapidly later in training. I implemented a logarithmic decay strategy and tested it on the DQN in Q3. Figure 1 below shows how epsilon decays over training, setting t_max to 350000 (half the total number of training episodes). This hyperparameter controls the point at which epsilon reaches epsilon_min.

I ran this decay strategy on the MountainCar environment for a total of 5 seeds, and achieved a mean return of -134.56 with standard deviation 4.028. This was only using one value of t_max. As with the linear and exponential decay strategies, other hyperparameter choices may give better results.
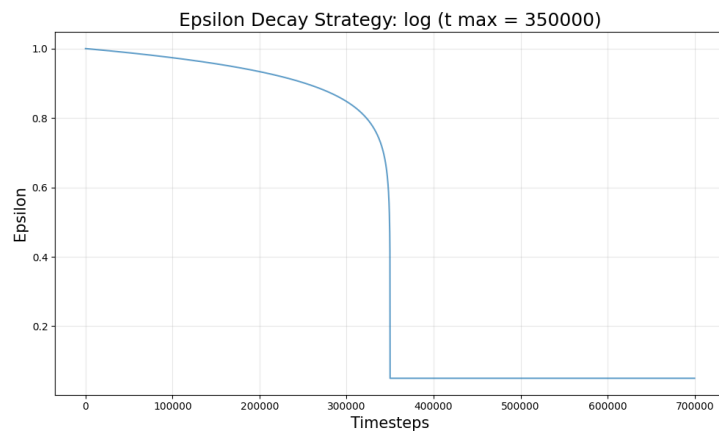
Figure 1: Logarithmic decay graph for t_max = 350000