

Optimal Control of An Invasive Plant in n-D Action Space

1 Main Ideas

The biggest advantage of using Q-learning in adaptive signal control is that it's not supervised learning. Rather than dealing with a large number of training data set, in Q-learning the controller learns how to react to the system dynamics in order to optimize a reward function.

The state space can comprise of the queue length, the elapsed phase time, various combination of upstream and downstream queue length and the offset of signal changes controlling upstream and downstream. At the end, it's a tread-off between the information you really need and care about, and the computational resources you have available to solve the problem. The reward, here, is straight forward; the time a car spends in a queue, which is in the form of penalty.

Convergence of Q-learning is highly dependent on the quality of exploration to ensure that all regions in our state space have been visited enough time or not. In order to do that, an ϵ -greedy policy has been used.