Data Visualization on Reinforcement Learning

Yuetong Liu

February 2021

1 Hyper-parameter Tuning

To find the optimal model architecture should be for a given model, we'd like to be able to explore a range of possibilities. Parameters which define the model architecture are referred to as hyperparameters and thus this process of searching for the ideal model architecture is referred to as hyperparameter tuning.

There are four hyperparameters in Policy Iteration:

- Exploration Rate: eps = [0,1];
- Discount Factor: gamma=[0,1];
- Maximum Iterations in policy evaluation: k = positive integer;
- Synchronous/Asynchronous: synchronous = Boolean;

There are two hyperparameters in Monte Carlo:

- Exploration Rate: eps = [0,1];
- Maximum Iterations : k = positive integer;

In this example, we use king as the agent to evaluate minimal number of steps it takes from (0,0) to (5, 7), given that the true minimal step is 8. Figure 1 and Figure 2 are line charts of estimated minimal steps vs hyperparameters value.

For Policy iteration, there is no obvious optimal hyperparameter.

For Monte Carlo, we can find the optimal hyperparameter which Exploration Rate = 0.1 and maximum iterations = 500.

2 Density of Steps

Figure 3 and 4 are density plot for number of step among 1000 episodes.

According the the plot, Monte Carlo has a heavy tail, compared with Policy Iteration

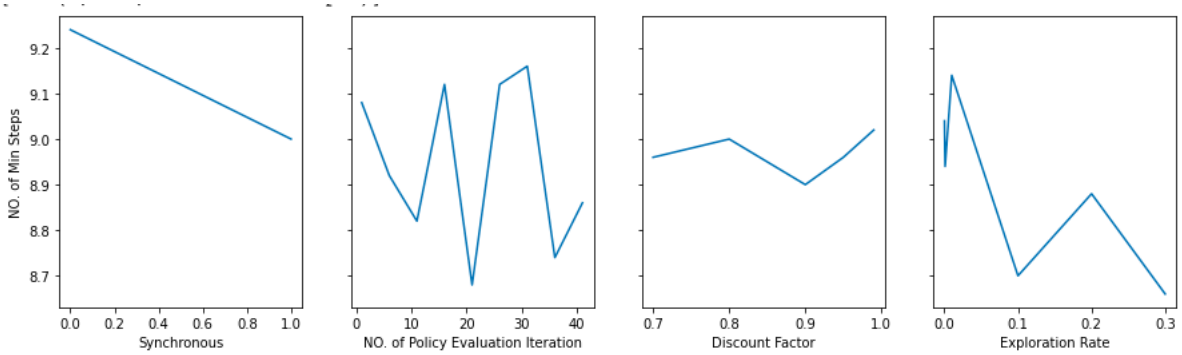

Figure 1: Hyper-parameter of Policy Iteration

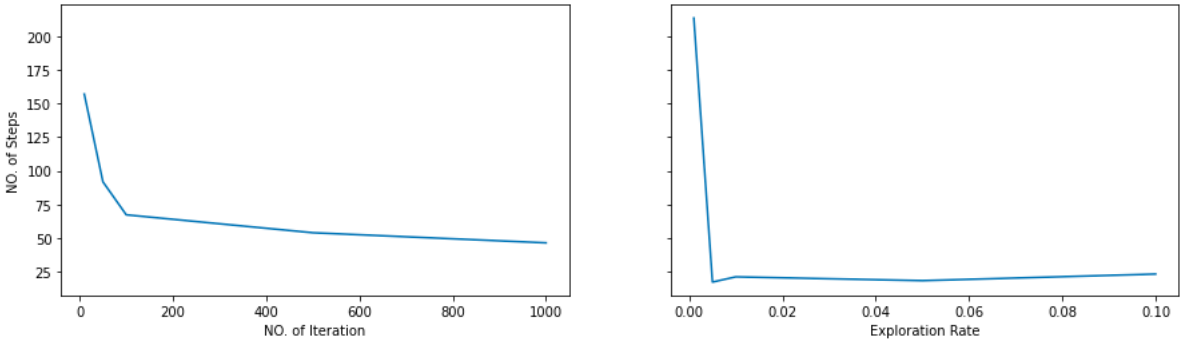

Figure 2: Hyper-parameter of MC

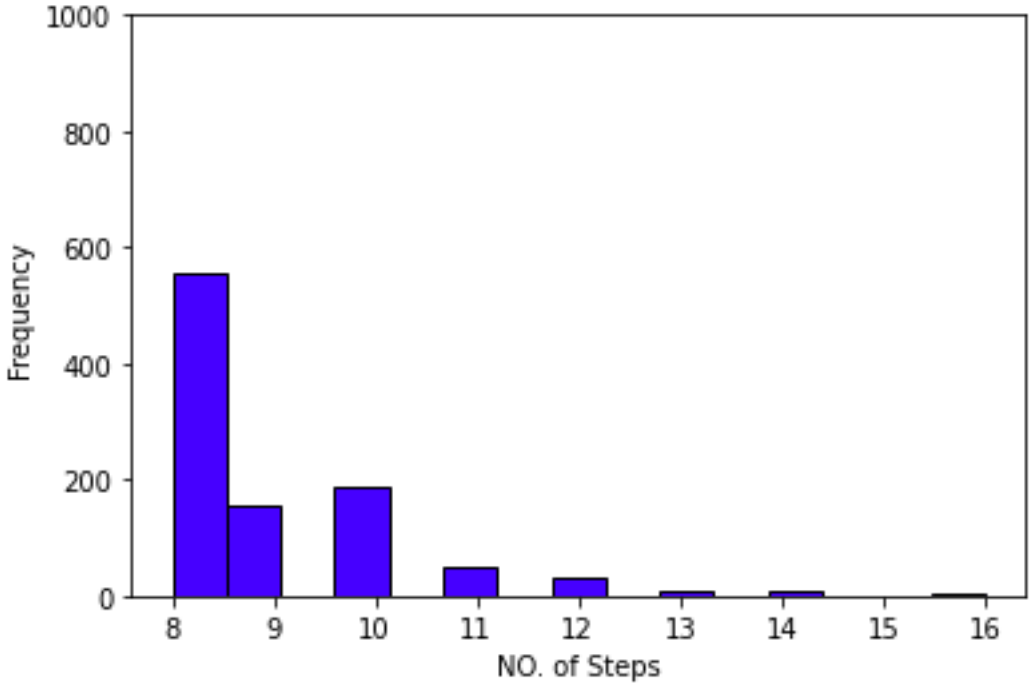

Figure 3: Density of Policy Iteration

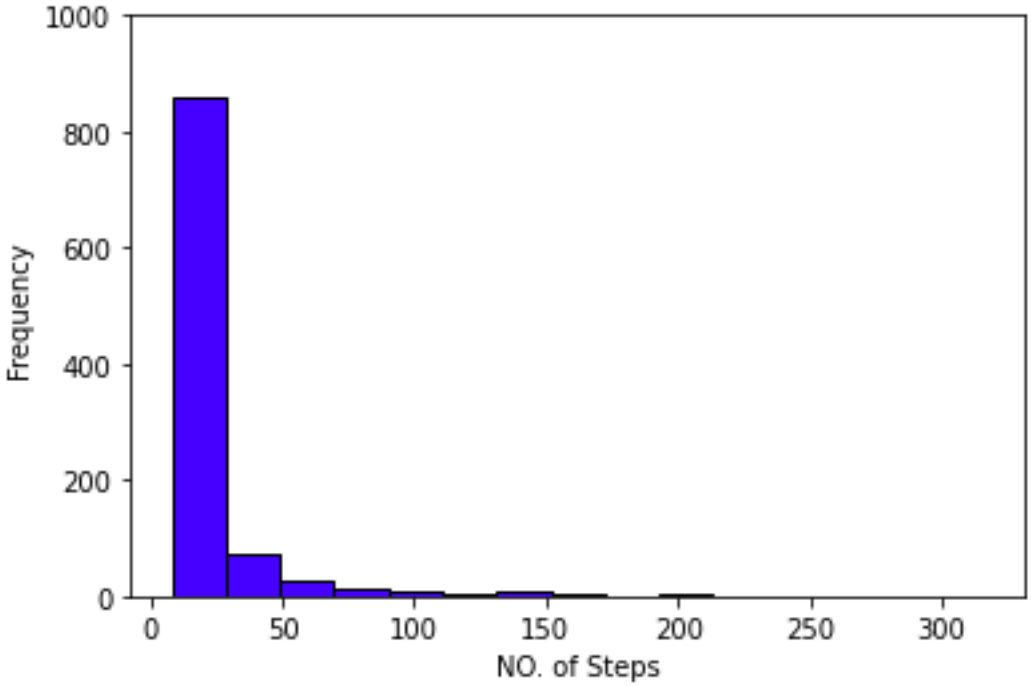

Figure 4: Density of MC