

Exercise 7

Task 1 - Linear function approximation

a) The feature vector $x(s) = (x_1(s), \dots, x_n(s))^T$ contains information about the state space. This information can be for example configurations of pieces in chess or the distance of a robot from the walls. Each entry represents one information. We can construct a feature vector for tabular methods, as well. This vector would then be $x'(s) = ([s = s_1], \dots, [s = s_n])^T$, where $[s = s_i]$ are indicator function which are 1 if we are in this state and 0 otherwise. Therefore, this vector would contain exactly one 1 at the state we are in and lots of 0s. If we multiply this by the weight vector, this would simply pick the weight corresponding to this state. This shows, that tabular methods are a special case of linear function approximation.

b) • tabular:

$$\begin{aligned}\delta_{t+1}(Q) &= R_{t+1} + \gamma Q(Y_{t+1}, A'_{t+1}) - Q(X_t, A_t) \\ Q_{t+1}(x, a) &= Q_t(x, a) + \alpha_t \delta_{t+1}(Q_t) \mathbb{I}_{x=X_t, a=A_t}\end{aligned}$$

• linear function approximation:

$$\begin{aligned}e_t &= \gamma \lambda e_{t-1} + \phi(s, a) \\ \theta_{t+1} &= \theta_t + \alpha_t e_t (r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)) \phi(s_t, a_t)\end{aligned}$$

Task 2 - Mountain Car

a)

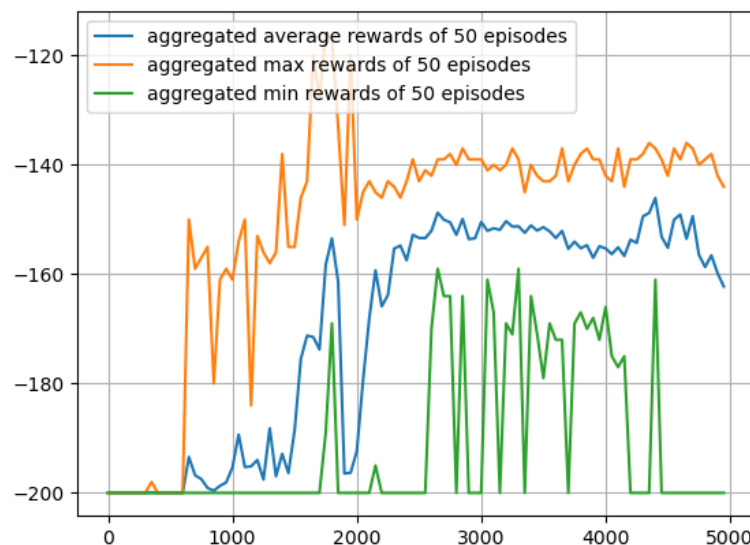


Figure 1: average rewards

b)

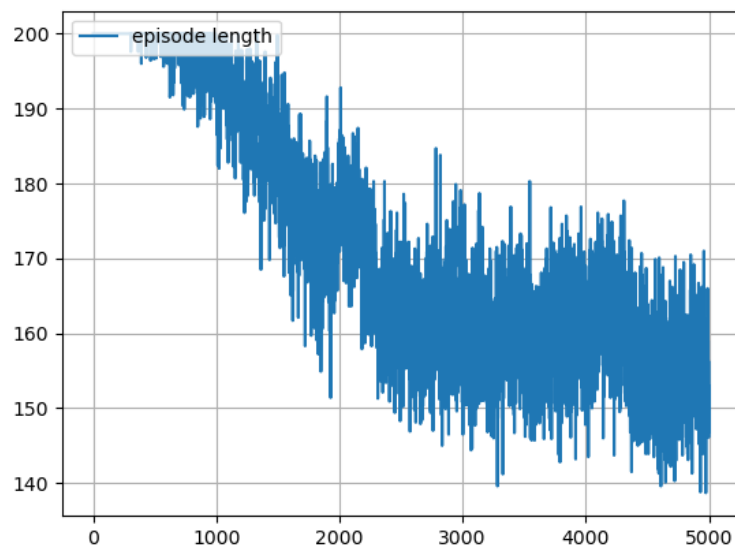


Figure 2: length

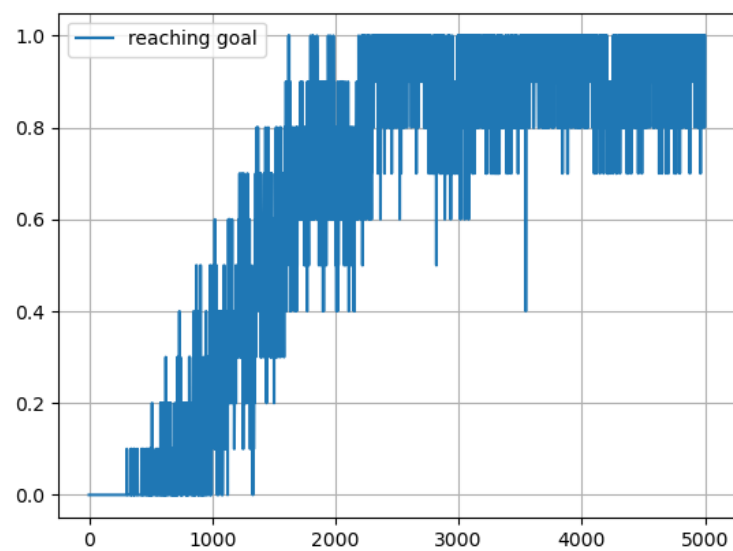


Figure 3: goal