Q-values estimation using Dyna agent after 500 episodes 0.068 0.043 0.018 Velocity -0.001 -

