

An optimal control approach to Reinforcement Learning and its algorithms

Maurizio Falcone

University of Roma "La Sapienza"

email: falcone@mat.uniroma1.it

Abstract

We consider an optimal control problem where the exact dynamics is unknown and only a probability distribution π on the vector field is given. The associated optimal control problem tries to minimize an average cost, where the average is made over the ensemble given by π . Our goal is to determine if the value function and the optimal control of this average cost problem converge respectively to the value function and the optimal control of the classical optimal control problem with underlying dynamics f as far as the distribution π gets closer and closer to the Dirac delta centered on the vector field f [2]. Finally, we show how these results can give a better understanding of some Reinforcement Learning algorithms and give some hints on their implementation [1].

Joint works with A. Pacifico, M. Palladino and A. Pesare.

References

1. A. Pacifico, A. Pesare, M. Falcone, *A New Algorithm for the LQR Problem with partially unknown dynamics*, Lecture Notes in Computer Science, 13127 LNCS, 2022, 322-330
2. A. Pesare, M. Palladino, M. Falcone, *Convergence results for an averaged LQR problem with applications to reinforcement learning*, Mathematics of Control, Signals, and Systems, (33) 2021, 379-411