Jun 15 - Jun 22 Seminar Review

Jiaxin Hu

June 16, 2020

1 MDFS 6.16

Title: Representation, Modeling, and Gradient Based Optimization in Reinforcement Learning
Representer: Sham Kakade

Regrarding representation

How can we address the large/infinite table in Reinforcement Learning? This seminar discusses this problem from three aspects: representation, modeling, and optimization. ~~In representation,~~ the log-linear policy method with 100 parameters is shown to perform no worse than a full neural network with 1000 parameters. This implies that fewer features ~~also support good learning~~ achieves good learning performance. A hard threshold to the goodness of the features is also provided. ~~In~~ Regarding modeling, Sham discusses the sampling efficiency and provides a minimax optimal sampling criterion to achieve a good performance. ~~In~~ Regarding optimization, training from different configurations is suggested to strengthen the robustness of the model. The Natural Policy Gradient (NPG) is proposed to solve the non-convex optimization with a global convergence which is independent of dimensions. The theoretical performance of NPG for the linear policy method is also ~~guaranteed.~~ provided The theoretical performance guarantee for NPG