Dynamic Programming

Reinforcement Learning Homework Assignment – Move 37

Collected by Sudha Mantri (Dean, Leesburg School of AI, VA, USA)

V 1.0

Dated 9/12/2018

Table of Contents

Disclaimer

I thank Siraj, all the internet sources, experts, and School of AI team members. This information is not meant to violate any copyrights. Kindly know that I am an eternal student (always learning) and making notes (literally). However, I wish to add that this information is meant to be used by School of AI Deans/Wizards to help members and aspirants learn AI Free of charge. Please let me know of any copyright violations and I shall duly oblige.

- *School of AI Deans please feel free to use this information and/or create your own sessions/training as you deem fit

Below is the School of AI Guidelines for all the deans. Please follow them and share information if/as need be.

https://gallery.mailchimp.com/21a24cf96b7ec98b2144fca5c/files/c9a14833-10b2-42c0-a024-f5310cb395bc/Copy_of_School_of_AI_Meetup_Guidelines.pdf

School_of_AI_Meetup_Guidelines.pdf

Version History

V 1.0	9/7/2018	Initial Version (Draft) meant for School of AI Deans as reference – Use to teach others for FREE

Dynamic Programming (DP) and Reinforcement Learning (RL) – Reading Assignment

- Many reinforcement learning (RL) algorithms utilize dynamic programming methods. Dynamic Programming methods have been researched for years to solve MDPs (Markov Decision Processes) assuming that the agent has prefect knowledge of functions that define an environment. In machine learning, the environment is usually formulated as an MDP. The main difference between the classical dynamic programming methods and RL algorithms is that RL algorithms do not assume knowledge of an exact mathematical model of the MDP. RL algorithms target large MDPs where exact methods are not possible. RL problems such as Q-learning fits a problem setting under Markov decision processes because the environment is observable and an agent can iteratively and progressively learn to improve. Agent works towards obtaining optimal discounted reward from the environment.

I suggest the following as reading assignment. foundational material for DP in RL with examples.

URL: https://github.com/dennybritz/reinforcement-learning/tree/master/DP/

Additional Reads:

1) Good Beginner/Midlevel Summary
   a) https://towardsdatascience.com/reinforcement-learning-demystified-solving-mdps-with-dynamic-programming-b52c8093c919
2) Comprehensive Reading:
   a) http://incompleteideas.net/book/bookdraft2018jan1.pdf
   b) http://www.wildml.com/2016/10/learning-reinforcement-learning/
3) Midlevel/Advanced Summary:
   a) https://www.cs.cmu.edu/~katef/DeepRLControlCourse/lectures/lecture3_mdp_planning.pdf
4) Advanced
   a) https://djrusso.github.io/RLCourse/index