

The Final Exam of Deep Learning and Practice (2017-Spring)

June 15th, 2017

(Deep Reinforcement Learning)

1. How can POMDP be viewed as MDP?
2. Anything related to the lab for 2048.
3. Convergence of Bellman Optimal Equation.
 - Contraction Mapping View? Convergence.
 - Policy Improvement?
 - Policy Iteration vs. Value Iteration
4. Policy Iteration With Monte-Carlo Evaluation
 - Why use action-value function, instead of value function?
 - Why use epsilon-greedy policy improvement, instead of greedy policy improvement?
Prove the convergence of epsilon-greedy policy improvement.
5. How does Q-learning work? And why use Q-Learning?
6. What is Score Function Gradient Estimator?
7. What is the Actor-Critic Algorithm?