

Reinforcement Learning exam program

1. RL problem statement
 - a. State, Action, Reward, Environment, Action
2. Crossentropy method
3. Value function
4. Value Iteration, Policy Iteration
5. Q-function
6. Q-learning
7. Approximate Q-learning: DQN,
8. DQN enhancements: Experience replay, Double DQN,
9. Autocorrelation problem in approximate Q-learning
10. Policy gradient and REINFORCE algorithm
11. Baselines idea in Policy gradient
12. Advantage Actor Critic (A2C)
13. TRPO
14. DDPG
15. Monte-Carlo tree search (MCTS)
16. Policy gradient applications in other domains (outside RL). How Self-Critical Sequence Training is performed? What is used as a baseline?