

Exam preparation worksheet

Q1. In the context of reinforcement learning what is the exploration exploitation tradeoff?

Q2. In your own words, give an example of a state action space formalism applied to a real-world problem.

Q3. Using an example, explain the Markov property.

Q4. In RL all methods can be classified according to the following taxonomy

- model-free or model-based
- value-based or policy-based
- on-policy or off-policy

Describe in detail each of these

Q5. What is Temporal Difference learning? How does it differ from Dynamic Programming methods?

Q6. Value iteration is considered of limited use when applied to real world problems. List a number of reasons as to why you think this is the case.

Q7. List the steps involved in the Value Iteration Algorithm and describe a method by which the algorithm can “self-terminate” once the difference between successive value updates becomes small.

Q8. List the steps involved in implementing the Q-learning algorithm. Is Q-learning considered model-based or model-free? Explain your answer.

Q9. List the steps involved in implementing the DQN algorithm.

Q10. What is a continuous state variable? Why do tabular methods prove problematic when dealing with large states and action spaces?