

Temporal-Difference Learning Study Guide (AIMA 22, RLAI 6)

Artificial Intelligence

1 Reinforcement Learning

1. What is reinforcement learning?
2. What is the *reward hypothesis*?
3. What does “model-free” mean in the context of reinforcement learning algorithms that assume there is an underlying MDP model of the environment?
4. How many time steps does TD(0) learning use in its target to update to the value of a state?