# Final Exam Sample Questions

1. The attack named MobilBye targets which type of sensor?
2. Camera
3. Lidar
4. Radar

ANS: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A

1. Can an ASIL-D system be composed of multiple ASIL-B subsystems?
2. Yes
3. No

ANS: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A

1. In PID control, which term(s) can be used to eliminate steady-state error?
2. P
3. I
4. D

ANS: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

B

1. True or False: twiddle() for PID controller tuning can be viewed as a form of Reinforcement Learning.
2. True
3. False

ANS: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A

1. Why does AlphaGo not use Tabular RL, such as tabular Q Learning or Sarsa?

ANS: the state space is too large.

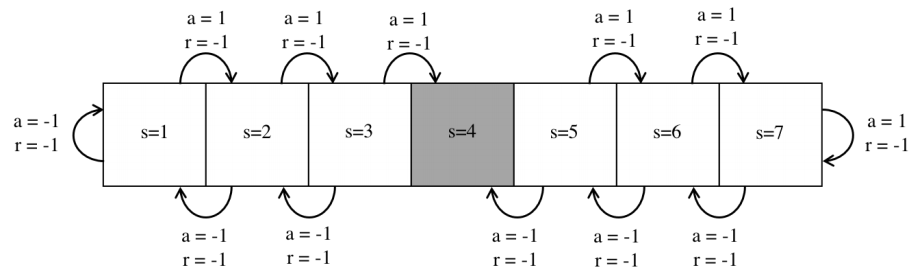
1. What is the probability of selecting the greedy action (assuming there is only one) in state in epsilon-greedy action selection, if the total number of possible actions in state is ?

ANS:

1. What is an Episodic task?
2. A task that has a limited number of actions and then ends.
3. A task with memory

ANS: A

1. Consider the following MDP. Environment is deterministic. In each state, there are two possible actions a ∈{−1, 1}, where −1 corresponds to moving left, and 1 corresponds to moving right. Each movement incurs a reward of r=-1. State s=4 is the goal state: taking any action from s=4 results in reward of r=0 and ends the episode, hence for any action .



Consider this episode in the form of (s,a,r): (3, −1, −1),(2, 1, −1),(3, 1, −1),(4, 1, 0). Assume . All value functions are initialized to 0. Derive the following (only show the changed parts):

1. State value functions after TD learning.
2. Action value functions after Sarsa.
3. Action value functions after Q learning.

ANS:

TD update equation:

Sarsa update equation:

Q learning update equation: