



Practice Quiz

TOTAL POINTS 9

1. What is the target policy in Q-learning?

1 point

- ☐ ϵ -greedy with respect to the current action-value estimates
- ☐ Greedy with respect to the current action-value estimates

2. Which Bellman equation is the basis for the Q-learning update?

1 point

- ☐ Bellman equation for state values
- ☐ Bellman equation for action values
- ☐ Bellman optimality equation for state values
- ☐ Bellman optimality equation for action values

3. Which Bellman equation is the basis for the Sarsa update?

1 point

- ☐ Bellman equation for state values
- ☐ Bellman equation for action values
- ☐ Bellman optimality equation for state values
- ☐ Bellman optimality equation for action values

4. Which Bellman equation is the basis for the Expected Sarsa update?

1 point

- ☐ Bellman equation for state values
- ☐ Bellman equation for action values
- ☐ Bellman optimality equation for state values
- ☐ Bellman optimality equation for action values

5. Which algorithm's update requires more computation per step?

1 point

- ☐ Expected Sarsa
- ☐ Sarsa

6. Which algorithm has a higher variance target?

1 point

- ☐ Expected Sarsa
- ☐ Sarsa

7. Q-learning does not learn about the outcomes of exploratory actions.

1 point

- ☐ True
- ☐ False

8. Sarsa, Q-learning, and Expected Sarsa have similar targets on a transition to a terminal state.

1 point

- ☐ True
- ☐ False

9. Sarsa needs to wait until the end of an episode before performing its update.

1 point

- ☐ True
- ☐ False

☐ I, **Dhawal Gupta**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.



[Learn more about Coursera's Honor Code](#)

Save

Submit