

✓ **Correct**

Correct! The update in Q-learning only learns about the greedy action. As demonstrated in Cliff World, it ignores the outcomes of exploratory actions.

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

1 / 1 point

- ☒ True
☐ False

✓ **Correct**

Correct! The target in this case only depends on the reward.

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

1 / 1 point

- ☐ True
☒ False

✓ **Correct**

Correct! Unlike Monte Carlo methods, Sarsa performs its updates at every time-step using the reward and the next action-value estimate.