Question 8 Exercise 6.12

No, even with greedy action selection, they will not be the same.

Sarsa ->

$$Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma Q(S',A') - Q(S,A)]$$

Here the new action is chosen before updating the value of the previous state action pair

Q-Learning

$$Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma \max_a (Q,a) - Q(S,A)]$$

choose next action based on new $Q(S,A)$

Since one chooses the action before updating the state action value (SARSA) and the other selects after, they won't be the same even with a greedy policy.