Assignment 10

The due date for submitting this assignment has passed.
You may not resubmit this assignment.

1) Which of the following is true about Markov and Semi-Markov Options? 1 point
   - A Markov Option's policy depends only on the current state
   - A Semi-Markov Option's policy depends only on the current state
   - A Semi-Markov Option's policy depends only on the current state and actions
   - A Semi-Markov Option is always a Markov Option (but not vice versa)

   No, the answer is incorrect.
   Accepted Answers:
   - A Semi-Markov Option is always a Markov Option (but not vice versa)

2) Which type of policies will you get for your policy to find the optimal solution consistent with the hierarchical structure of the problem? 1 point
   - Markov optimal solution
   - Hierarchically optimal solution
   - Hierarchical optimal solution
   - Partial optimal solution

   No, the answer is incorrect.
   Accepted Answer:
   - Partial optimal solution

3) Which of the following is a correct Bellman equation for an SMPD? 1 point
   \[ V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \]
   \[ V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \]
   \[ V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \]
   \[ V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \]
   \[ V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \]

   No, the answer is incorrect.
   Accepted Answer:
   - Partial optimal solution

4) In the above Abstract Machine the core MDPs all change only when we visit a choice point 1 point
   - True
   - False

   No, the answer is incorrect.
   Accepted Answer:
   - False

5) Consider a SMPD in which the next state and the reward only depend on the previous state and action. Is \( V(s, a, n, t) = \max_{\pi} \left\{ R(s, a, n, t) + \gamma E[V(s', a', n', t') | s', a', n', t', \pi] \right\} \)

   Yes, because it only changes anything and we end up with a unique state and action sequence.
   No, because it depends on the action and discounting may have an effect on the final policies.
   Yes, because the next state will depend on \( s \).
   Yes, because the Bellman equation is same for both methods in this case.

   No, the answer is incorrect.
   Accepted Answer:
   - No, because it depends on the action and discounting may have an effect on the final policies

6) In the above SMPD if \( f \) is fixed for all actions, action values will always get the same policy for conventional Q learning and SMPD Q learning? For \( f \) defined by \( f(\pi) = \pi(x_1, x_2, x_3, x_4, x_5) \) 1 point
   - Yes, yes
   - Yes, no
   - No, yes
   - No, no

   No, the answer is incorrect.
   Accepted Answer:
   - Yes, yes

7) Does the optimal solution give the solution with highest expected return? 1 point
   - Yes
   - No

   No, the answer is incorrect.
   Accepted Answer:
   - No

8) Suppose we have a deterministic Markov option \( \pi \) being executed, and that the current state is \( s \). The action taken in \( s \) is \( \pi(s) \); which is of course 1 point
   - \( s \)
   - \( \pi(s) \)
   - \( s' \)
   - \( \pi(s') \)

   Yes, the reward is the immediate reward received from taking action \( \pi(s) \) in state \( s \) while executing option \( \pi \) is obtained.

   No, the answer is incorrect.
   Accepted Answer:
   - No

9) Which of the following statements are true? 1 point
   - If we have two deterministic Markov options, \( \pi_1 \) and \( \pi_2 \), then \( \pi_1 \) takes action \( \pi_1(s) \) when \( s \) is in \( \pi_1 \) and \( \pi_2 \) takes action \( \pi_2(s) \) when \( s \) is in \( \pi_2 \).
   - The reward received by the second option is the immediate reward received from taking action \( \pi_2(s) \) in state \( s \).

   No, the answer is incorrect.
   Accepted Answer:
   - No

10) Which of the following statements are true? 1 point
    - If we have two deterministic Markov options, \( \pi_1 \) and \( \pi_2 \), then \( \pi_1 \) takes action \( \pi_1(s) \) when \( s \) is in \( \pi_1 \) and \( \pi_2 \) takes action \( \pi_2(s) \) when \( s \) is in \( \pi_2 \).
    - The reward received by the second option is the immediate reward received from taking action \( \pi_2(s) \) in state \( s \).

    No, the answer is incorrect.
    Accepted Answer:
    - No