1. Which of the following algorithms are appropriate in a control setting in which updates will be made at every time step? [Select all that apply] **1 point**

- [x] SARSA
- [x] Expected SARSA
- [x] Q-learning

2. Which of the following algorithms are appropriate in a prediction setting in which updates will be made at the end of each episode? [Select all that apply] **1 point**

- [ ] Exploring Starts Monte-Carlo
- [x] Monte-Carlo Prediction
- [x] Off-Policy Monte-Carlo

3. Which of the following algorithms are appropriate in a tabular setting in which we will be learning a model and using it for planning? [Select all that apply] **1 point**

- [x] Dyna-Q+
- [x] Dyna-Q
- [ ] Expected SARSA

4. Which of the following algorithms are appropriate in a control setting in which we are given access to a model? [Select all that apply] **1 point**

- [x] Dyna-Q

- [ ] Iterative Policy Evaluation
- [x] Value Iteration
- [x] Policy Iteration

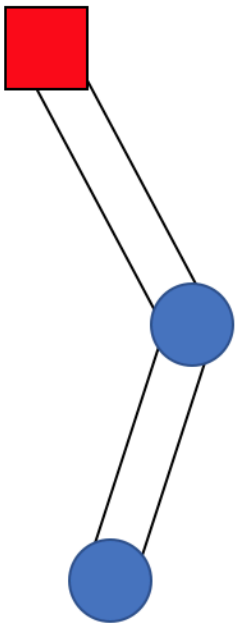
5. Which of the following algorithms are appropriate in a continuing control setting with a discrete action space and function approximation? [Select all that apply] **1 point**

- [ ] Gaussian Actor-Critic
- [x] Differential Softmax Actor-Critic
- [x] Differential Semi-Gradient SARSA

6. Which of the following algorithms are appropriate in an online prediction setting with linear function approximation? [Select all that apply] **1 point**

- [x] Semi-Gradient TD
- [ ] Gradient Monte Carlo
- [ ] SARSA

7. In the **continuing** acrobot system (shown below), a double pendulum is fixed to the red square. The goal is to swing the double pendulum such that the height of the mass on the end of the lower pendulum exceeds the height of the black line. When the goal is reached, a reward of one is given and the double pendulum transitions to a vertical position. Otherwise, the reward is zero. Which of the following algorithms are appropriate for control in this context? [Select all that apply] **1 point**



- [ ] Expected SARSA
- [ ] Q-learning
- [x] Average Reward Actor-Critic

8. Which of the following algorithms are appropriate for control in the lunar lander MDP, as it is described in the lecture “Initial Project Meeting with Martha: Formalizing the Problem”? [Select all that apply]

- [x] Expected SARSA

- [ ] Average Reward Actor-Critic
- [x] Q-learning