EN 520.650 Homework 4 Spring 2022

Due on 04/13/2022

1. Consider an undiscounted MDP having three states, (1, 2, 3), with rewards −1, −2, 0, respectively. State 3 is a terminal state. In states 1 and 2 there are two possible actions: a and b. The transition model is as follows:  
   - In state 1, action a moves the agent to state 2 with probability 0.6 and makes the agent stay put with probability 0.4.  
   - In state 2, action a moves the agent to state 1 with probability 0.6 and makes the agent stay put with probability 0.24  
   - In either state 1 or state 2, action b moves the agent to state 3 with probability 0.2 and makes the agent stay put with probability 0.8.  
   Answer the following questions:  
   1. What can be determined *qualitatively* about the optimal policy in states 1 and 2?  
   2. Apply policy iteration, showing each step in full, to determine the optimal policy and the values of states 1 and 2. Assume that the initial policy has action b in both states.  
   3. What happens to policy iteration if the initial policy has action a in both states? Does discounting help? Does the optimal policy depend on the discount factor?
2. Compare and contrast the various CNN-based object detectors using the following metrics: number of parameters, training and test times, performance on ImageNet dataset.
3. Discuss methods for dropout and searching for the best neural architecture.
4. For the neural network shown below with initial weights indicated in the figure, calculate the next values of w7 and w3 assuming a single input i1=0.05, i2=0.10 with corresponding outputs o1=0.01 and o2=0.99. Hint: Follow the calculations done in class on 10/06/2021!

