Idan Hershcovich
CS 481
Assignment 2
4/24/20

Assignment 2

1. Three white and three black balls are distributed in two urns in such a way that each contains three balls. We say that the system is in state i, i=0,1,2,3, if the first urn contains i white balls. At each step, we simultaneously draw one ball from each urn and place the ball drawn from the first urn into the second, and conversely with the ball from the second urn. Let Xn denote the state of the system after the nth step. Explain why {Xn, n=0,1,2,...} is a Markov chain. Calculate its transition probability matrix and stationary probabilities π0, π1, π2, π3 for the states 0,1,2,3.

   It is a Markov Chain, since the state of the system depends on the amount of white balls in the first urn, not in any of the other past states. It doesn't matter what came before the current state, the next state will be determined by the current one.

   Probability Matrix:

Next State→	**0**	**1**	**2**	**3**
Current State↓				
0	**0**	**1**	**0**	**0**
1	**1/9**	**4/9**	**4/9**	**0**
2	**0**	**4/9**	**4/9**	**1/9**
3	**0**	**0**	**1**	**0**

   Stationary Probabilities: $A^t A\pi = A^t b$

$$\begin{bmatrix} \pi_0 \, \pi_1 \, \pi_2 \, \pi_3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} & 0 \\ 0 & \frac{4}{9} & \frac{4}{9} & \frac{1}{9} \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{20} \\ \frac{9}{20} \\ \frac{9}{20} \\ \frac{1}{20} \end{bmatrix}$$

Idan Hershcovich
CS 481
Assignment 2
4/24/20

2. Read Martin Gardner's Mathematical Games column from the March 1962 issue of Scientific American. The 24 states of Hexapawn in the example constitute a complete set of states for which the game's second player has a choice of actions, with the exception of symmetric variants of the opponent's first move. Notice that states arising in which the player has already lost are not included. (Note: because the article has been reproduced in grayscale, the color information for the moves has been lost. It is sufficient to select your own colors; simply make sure each possible move for a given state has a unique color.)
   Follow the instructions for reinforcement learning. How many games did you need to play before before learning was complete? Specify the policy that has been learned by giving an action for each of the 24 states. Indicate which states, if any have multiple options remaining.

   a. **By game 17, the AI had won the last 7 games, but then lost like 3 game by game 23. By game 33 It was winning 100% of the time. I'd say it takes about 15-16 losses for a computer to reach a 100% win rate**
   b. **After the learning has been completed, the policy that has been learned (the policies that result in a win), in the order given by the article:**
      i. **State 3,4,5,6,7,9,11,13**
      ii. **Some of these policies then lead to other states that can result in a win, some are an instantaneous win depending on the option the AI takes.**
   c. **All states in the section 2 (the AIs first move) have multiple options remaining**

3. Programming assignment analysis:
   a. In the assignment, I used a learning rate of 0.1 and did 100 trials of the reinforced learning AI (RL) vs the 3 other AIs and 50 vs the human (me)
   b. Choosing 0.1 for the learning rate made it so it took more runs to get some high values in my learn function V
   c. After the 100 runs were complete, my program mostly learned the correct values. The final values for each state were:
      {11: -0.034710392423216704, 12: -0.8618738417887959, 13:-0.8839755172551941, 22: 0.6354955058971038, 23: 0.001285325524342985, 33: 0.999718039563302}
   d. It successfully learned that the lower states like 12,13 were not the best, but it had issues with 11 and 23. State 23 could be because of how ties are handled, since they could've been handled better, but since the HW didn't specify anything about it, I chose not to dwell too much on them. State 11 I'm not too sure about,

maybe because players almost always discard 1, or lack of Exploration. Even so, it definitely learned that 33 is the best hand possible, and 22 not too far behind