Robert Roche

Hw2 written solutions

1.)

State 0 → State 1 = 100%

State 1 → State 0 = 1/9

State 1 → state 1 = 4/9

State 1 → state 2 = 4/9

State 2 → State 1 = 4/9

State 2 → State 2 = 4/9

State 2 → State 3 = 1/9

State 3 → State 2 = 100%

Transition Matrix:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/9 & 4/9 & 4/9 & 0 \\ 0 & 4/9 & 4/9 & 1/9 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Explanation:

We're in state i if the first urn contains i white balls. A ball is drawn from each urn and placed in the other urn. This explains the reflection of probabilities. Our chances of selecting a white ball from each urn is 2/9 same for black thus we have a 4/9 chance of remaining in the first state. You can use this logic to calculate the probabilities for the transition matrix.

$$\pi_0 = \frac{1}{9}\pi_1$$

$$\pi_1 = \pi_0 + \frac{4}{9}\pi_1 + \frac{4}{9}\pi_2$$

$$\pi_2 = \frac{4}{9}\pi_1 + \frac{4}{9}\pi_2 + \pi_3$$

$$\pi_3 = \frac{1}{9}\pi_2$$

$$\pi_3 = \frac{1}{9}\pi_1$$

$$\pi_1 = \frac{5}{9}\pi_1 + \frac{4}{9}\pi_2$$

$$\frac{4}{9}\pi_2 = \frac{4}{9}\pi_1 \qquad \pi_2 = \pi_1$$

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$$

$$\frac{1}{9}\pi_1 + \pi_1 + \pi_1 + \frac{1}{9}\pi_1 = 1$$

$$\frac{20}{9}\pi_1 = 1$$

$$\pi_1 = \frac{9}{20} \simeq 0.45$$

$$\pi_0 = \frac{1}{9}\pi_1 = \frac{0.45}{9} = 0.05$$

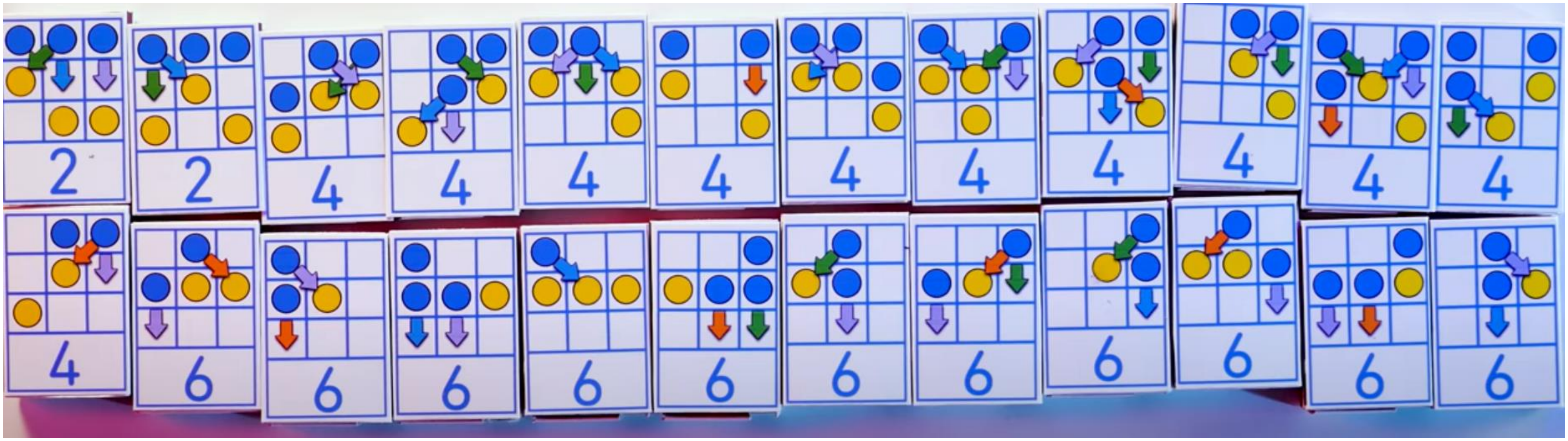
$$\pi_2 = \pi_1 = 0.45$$

$$\pi_3 = \frac{1}{9}\pi_1 = 0.05$$

$$\pi = [0.05, 0.45, 0.45, 0.05]$$

2.)

Below is the image representation of each state I used to conduct my research. The number on the bottom denotes the round the computer is moving on. It took me 27 games for the computer to learn an optimal policy, however the optimal policy means avoiding certain states before I even had a chance to weed them out, thus to make a more robust policy I went back in and created the best move for each state even if the policy I found originally does not allow the said state to occur. The source of the image is the following youtube video discussing how hexapawn and the learning machine works from the channel Vsauce. This video helped me gain a better understanding of how the game and the machine are intended to work. I then used colored pencils and paper to create my own simulation



Below is my policy table. Each move has a brief description and then the color of the arrow corresponding to the move is noted to bring further clarification, it is helpful to look at the states and the policy side by side.

State	Policy	State	Policy
1	Top middle piece captures diagonally (green)	13	Both moves are equally valid (red and purple)
2	Top left piece captures diagonally (blue)	14	Left center piece moves forward (purple)
3	Both diagonal captures lead to a losing state (green and purple)	15	Left center piece moves forward (red)
4	Center piece moves forward or captures diagonally (blue and purple)	16	Center or left center piece moves forward (blue or purple)
5	Top center piece captures diagonally (blue)	17	Next state is a lose state no matter what
6	Top right piece moves forward (red)	18	Center or right center piece moves forward (red or green)
7	Next state is the losing state no matter what	19	Center piece moves forward (purple)
8	Top right piece moves forward (purple)	20	Left center piece moves forward (purple)
9	Center piece captures diagonally (red) or moves forward (blue)	21	Right center piece moves forward (blue)
10	Top right piece moves forward (green)	22	Right center piece moves forward (purple)
11	Left center piece moves forward (red)	23	Center or left center piece moves forward (purple and red)
12	Both moves are equally valid (green and blue)	24	Center piece moves forward (blue)

This policy contains the optimal moves for victory. It should be noted there are a number of moves that could result in a victory or defeat, that are not optimal. Because of their inherent risk as opposed to other moves they were not included in the policy.