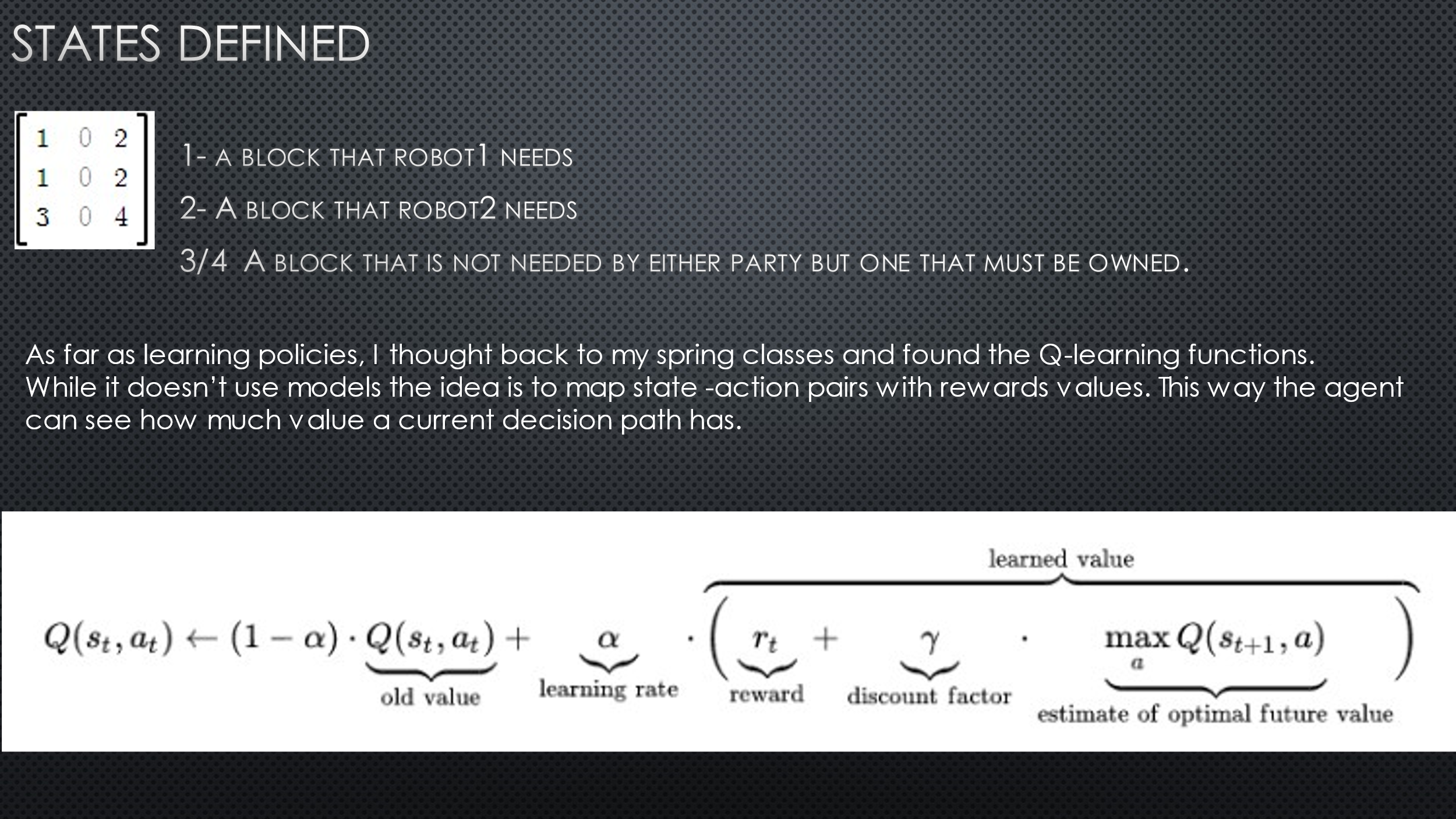
* **Agents**
  + Two agents named “robot1” and “robot2” respectively
  + **Agent-Actions(A)**
    - move a block or ask the other agent to move a block
    - more boiled down
      * move one of the current agents blocks to the middle
      * take one of the blocks from the middle to the current agent’s column
      * ask the other agent to move a block to the middle column
  + **Discount factor(Y)**
    - This is gamma Y, as this will determine the amount of immediate reward vs future possible rewards for a given state or possible sequence of states
* **Environment (reward heuristic)(E)**
  + This will take in the current board state for the current agent and the agents selected action. The environment will then give back to the agent it’s reward and newly created state based off of its action
  + **Reward(R)**
    - The value of the chosen action at the given state
  + **State(S)**
    - This will be the NxN grid with blocks located on them



<https://www.desmos.com/matrix>

<https://wiki.pathmind.com/neural-network>

<https://wiki.pathmind.com/deep-reinforcement-learning>

<https://towardsdatascience.com/reinforcement-learning-implement-tictactoe-189582bea542>