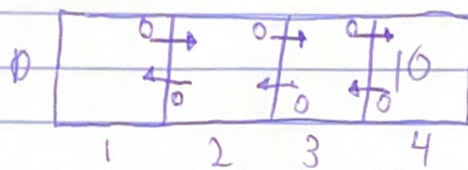


Q-learning HW

Evan Nuss



$$Q((0,1), \rightarrow) = 10 + 0 = 10$$

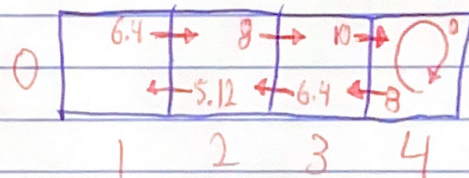
$$Q((0,4), \leftarrow) = 0 + .8(10) = 8$$

$$Q((0,2), \leftrightarrow) = 0 + .8(10) = 8$$

$$Q((0,1), \rightarrow) = 0 + .8(8) = 6.4$$

$$Q((0,2), \leftarrow) = 0 + .8(6.4) = 5.12$$

$$Q((0,3), \leftarrow) = 0 + .8(8) = 6.4$$



Optimal policy

