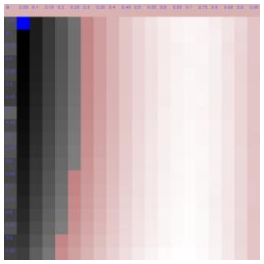
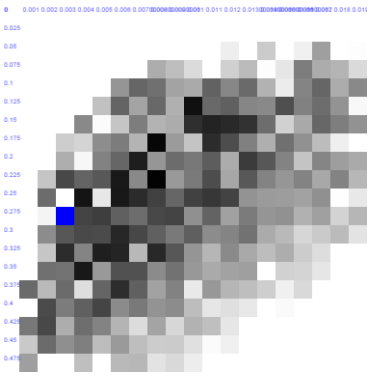
Joel Einbinder
CS 440
einbndr2

1. WorldWithoutThief
    a. With ε set to 0.0, the agent quickly learned how to achieve a score of 0, and never explored further. Average: -0.0095, Maximum: 0, Minimum: -9.5
    b. With ε set to 0.1, the agent started worse, getting an average of about -10 for the first 10 episodes. After about 20 episodes however, the agent quickly improved. Average: 132.9, Maximum: 249, Minimum: -24.5.
    c. With a higher value of ε = 0.5, the agent was able to sometimes deliver the package. However it didn't retain this knowledge. Overall the scores were much worse. Average: -28.65, Maximum: 20.5, Minimum: -79.
2. WorldWithThief
    a. Without knowledge of the thief, the agent never figures out how to achieve a positive score. Average: -13.05, Maximum: -3, Minimum: -38.5
    b. With knowledge of the thief, the agent is very successful at learning the world. After the seventh episode, the agent always gets a positive score. After about 25 episodes, its score never drops below 200.Average: 271.8, Maximum: 304.5, Minimum: -68.
    I graphed the average values over epsilon and the learning rate, zooming in on interesting areas. I settled on an epsilon of 0.003 and a learning rate of 0.275. This gives an average total score of about 315.

    I felt like pasting these, but they aren't going to be much good for reading.
    Starting graph:

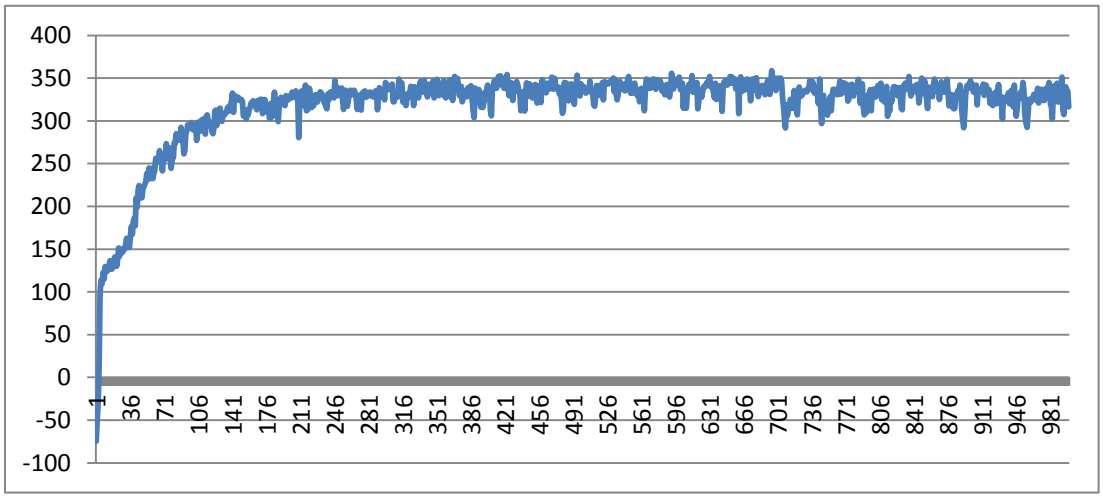


    <Many iterations in between>
    Final graph:



Joel Einbinder
CS 440
einbndr2

3. First iteration: The agent learns how to get about 300 points and levels out. It fluctuates between about 300 and 350 points.



After 10 iterations: The graph looks similar, but much smoother. A resting value of about 330 is reached.