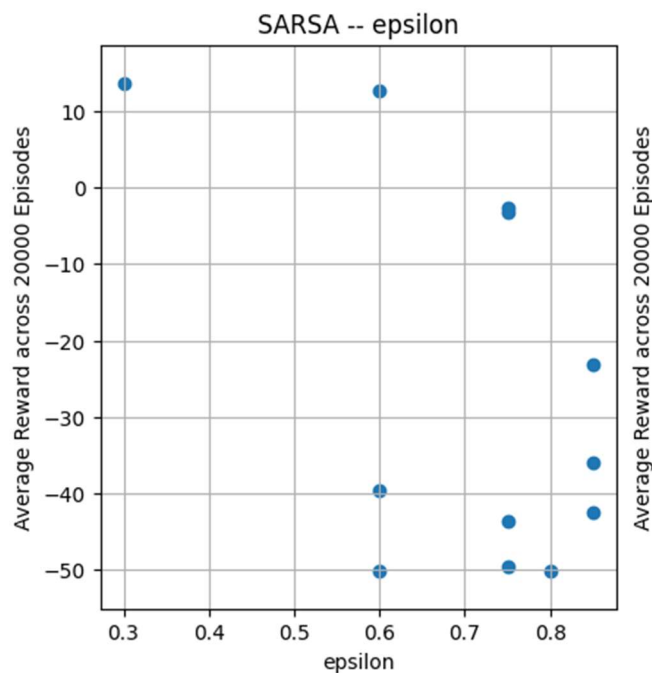


Q3)

- How does the performance of SARSA compare to Q-learning? It is helpful to run each algorithm multiple times and average the observations.
 - SARSA seems to more often achieve the large goal or the small goal, and rarely falls into the volcano, whereas Q-learning either falls into the volcano or achieves the large goal. The result is a better overall policy score for SARSA than Q-learning.
- How does your conclusion change for different exploration probabilities?



- As the chart above shows, epsilon maximizes reward in SARSA in the range of 0.3 and 0.6. When too much exploration is encouraged, the agent tends to fall into the volcano more regularly.
- Rerun your algorithms for both maps (4X3 and 8X6), and analyze the results.

- As you can see from the charts below, I tried several different learning rates on both the “4x3” and the “8x6” maps. In general Q-Learning outperformed SARSA in these experiments. SARSA performed better across a range of learning rates, whereas Q-Learning only performed well when the learning rate was at or below a value of 1.

