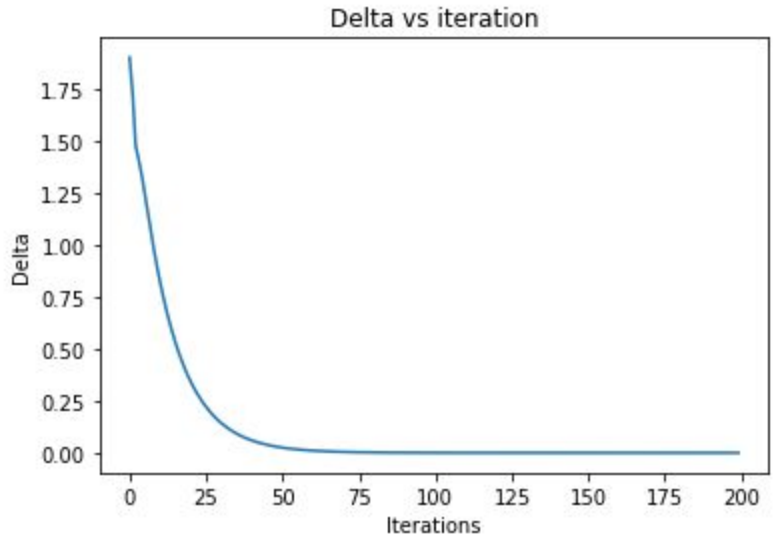
RL Assignment 2

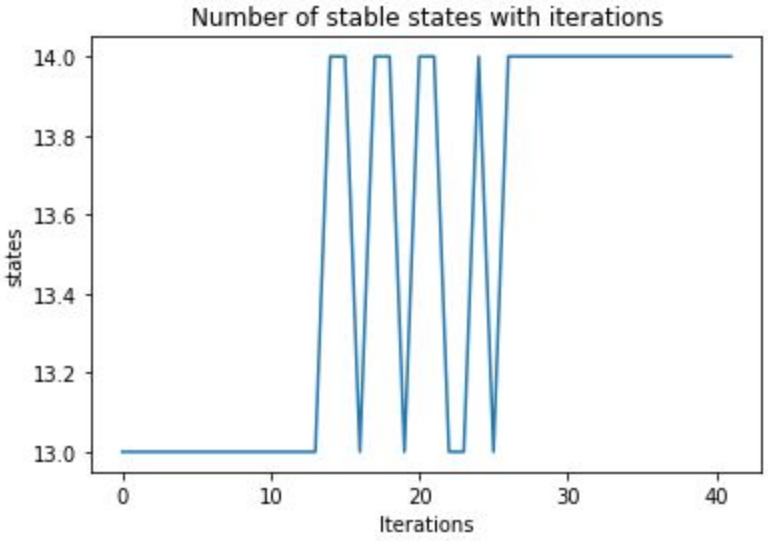
Shaurya Goel
2016194

Ans-6

Policy Iteration

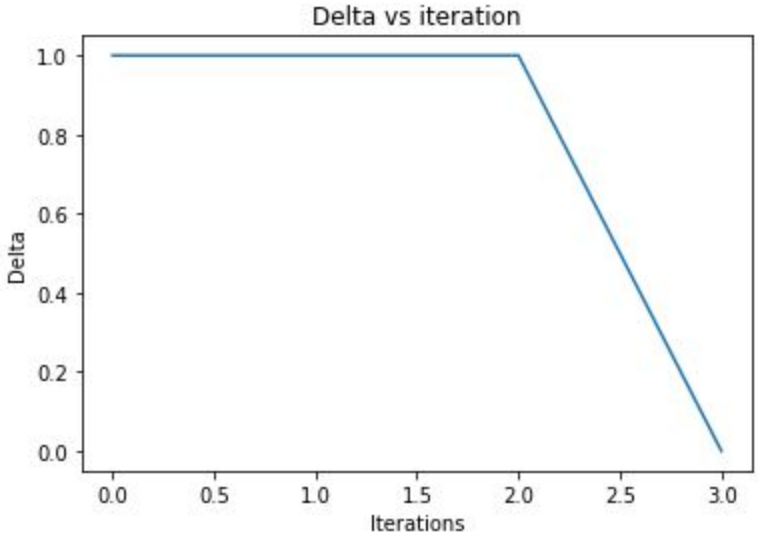


This graph shows that our value function is converging to a stable value.



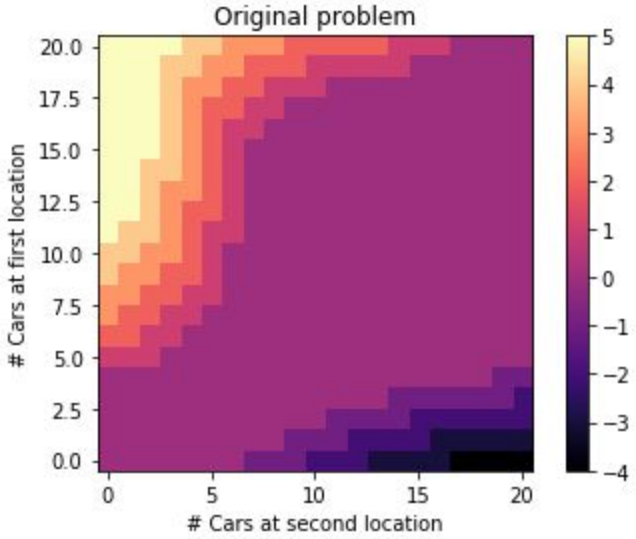
This show that our policy is becoming stable with time, hence ensuring convergence.

Value Iteration

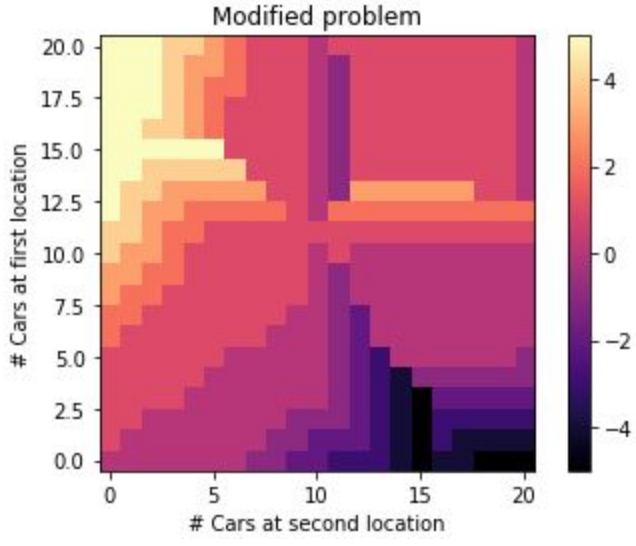


This shows that our value function is becoming stable.

ANS-7



Heatmap for Jack's car rental problem. This shows that if we have large number of cars in any location, then move these cars to the other location. Prefer moving from location 1 to location 2 as location 2 has higher rental rate and low return rate.



This is heatmap of modified version of Jack's car problem.
Overall same conclusion as before. In addition to that, if we have 11-12 cars in any location, then, it is preferable to transfer it to the other location, as >10 cars occurs a cost of 4. If number of cars >12, then there is no benefit to transfer as cost of transfer is 2 per car.