

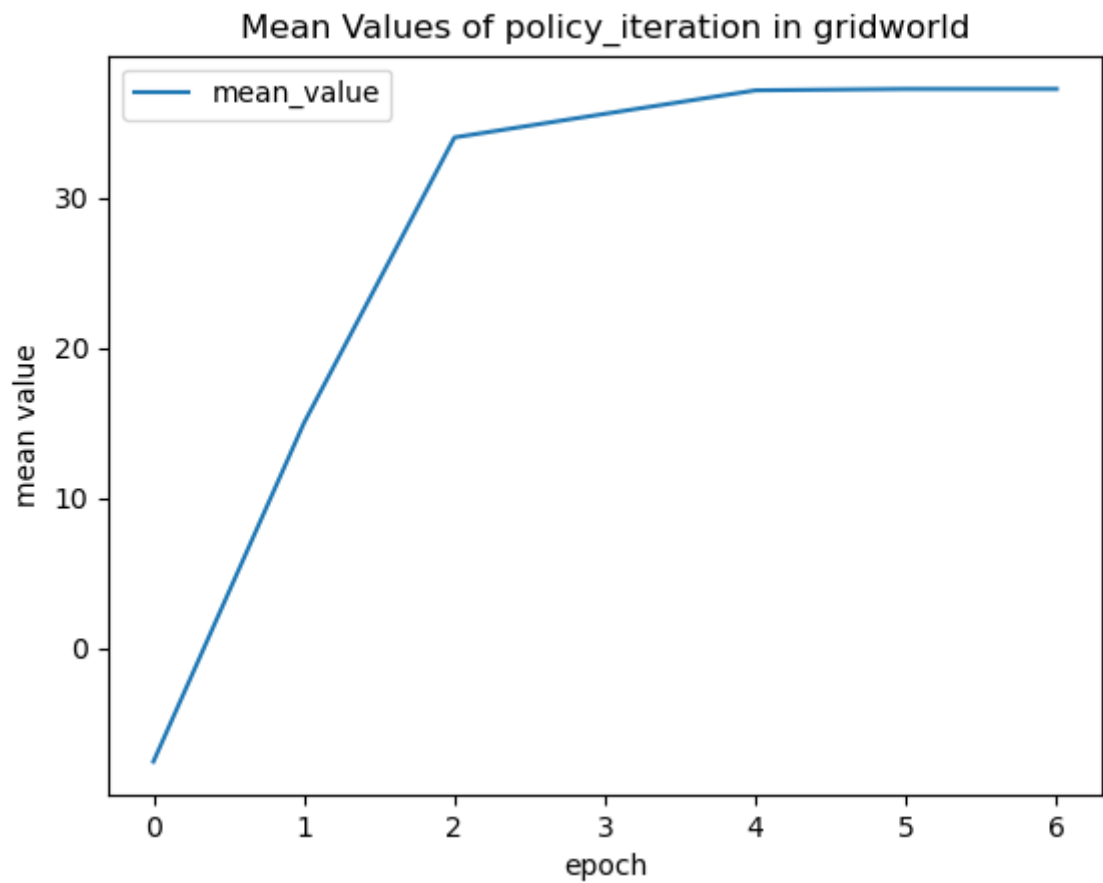
# Result - Mike Liu xl142

---

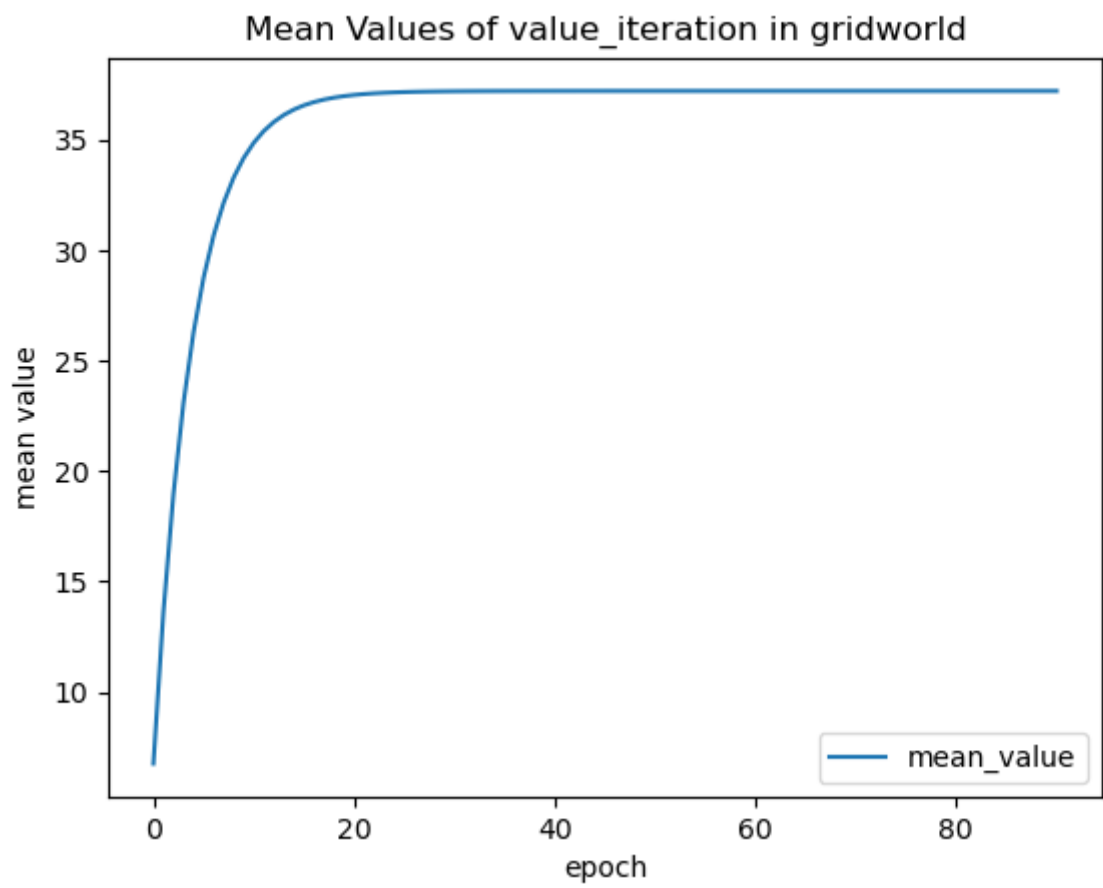
## Learning curves

Gridworld

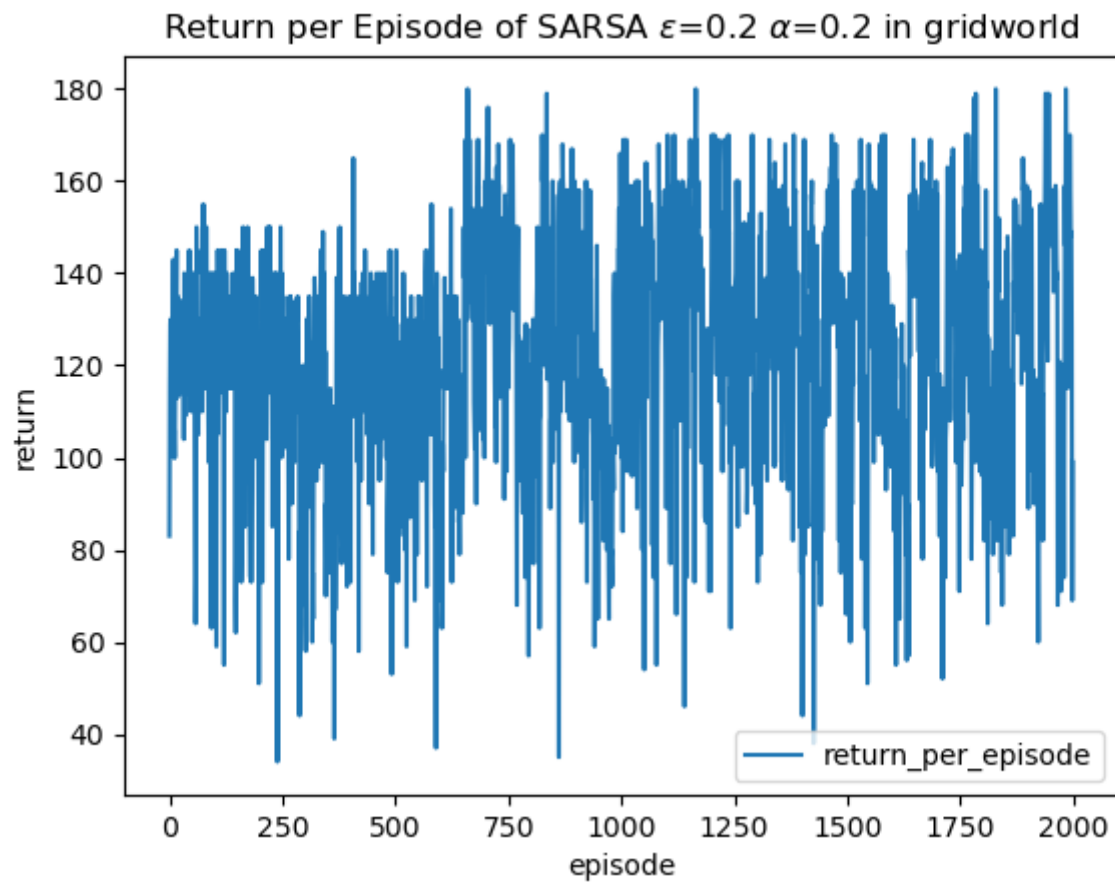
### Policy iteration



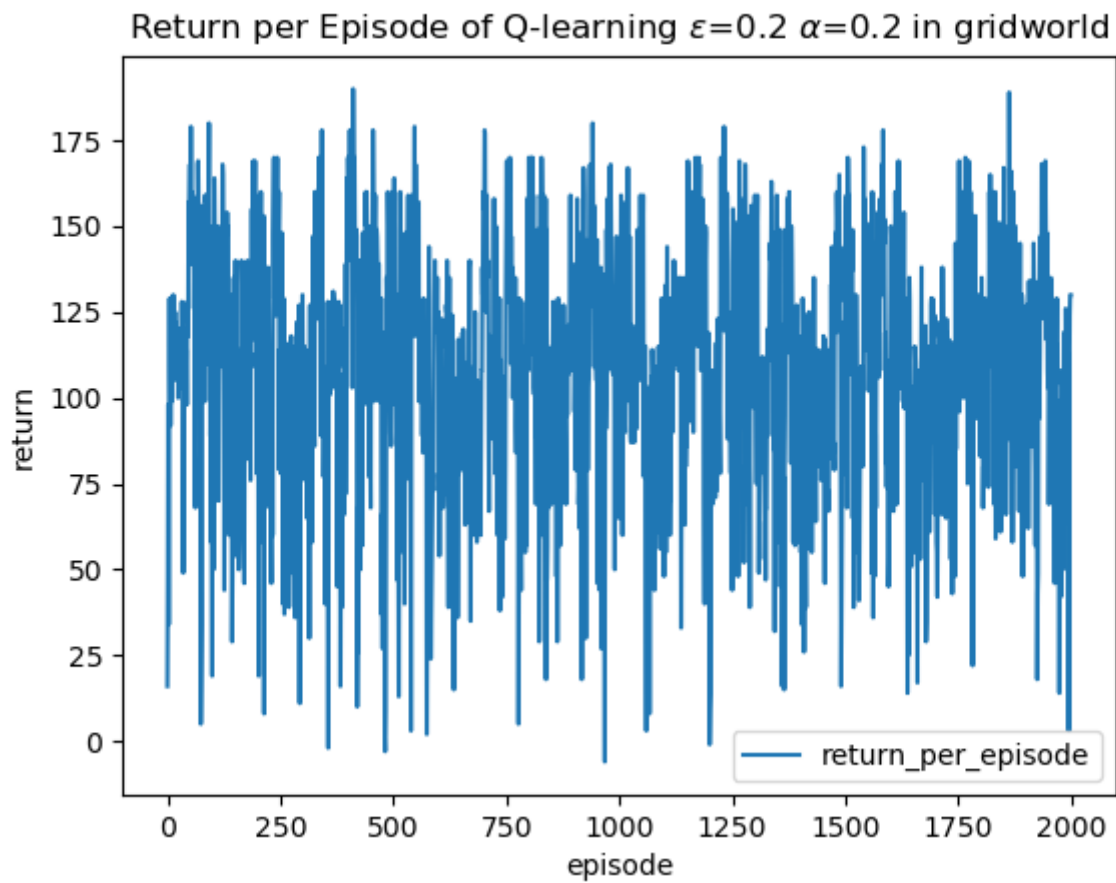
### Value iteration



**SARSA**

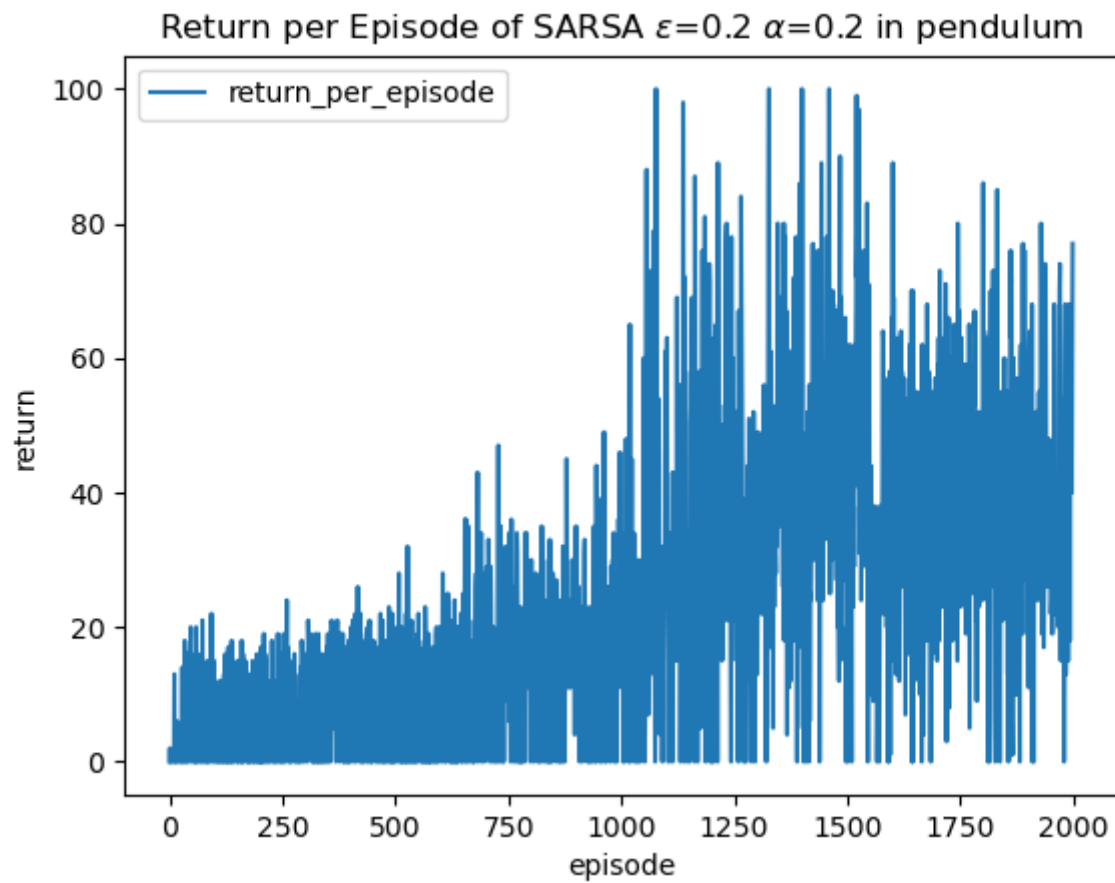


## Q-learning

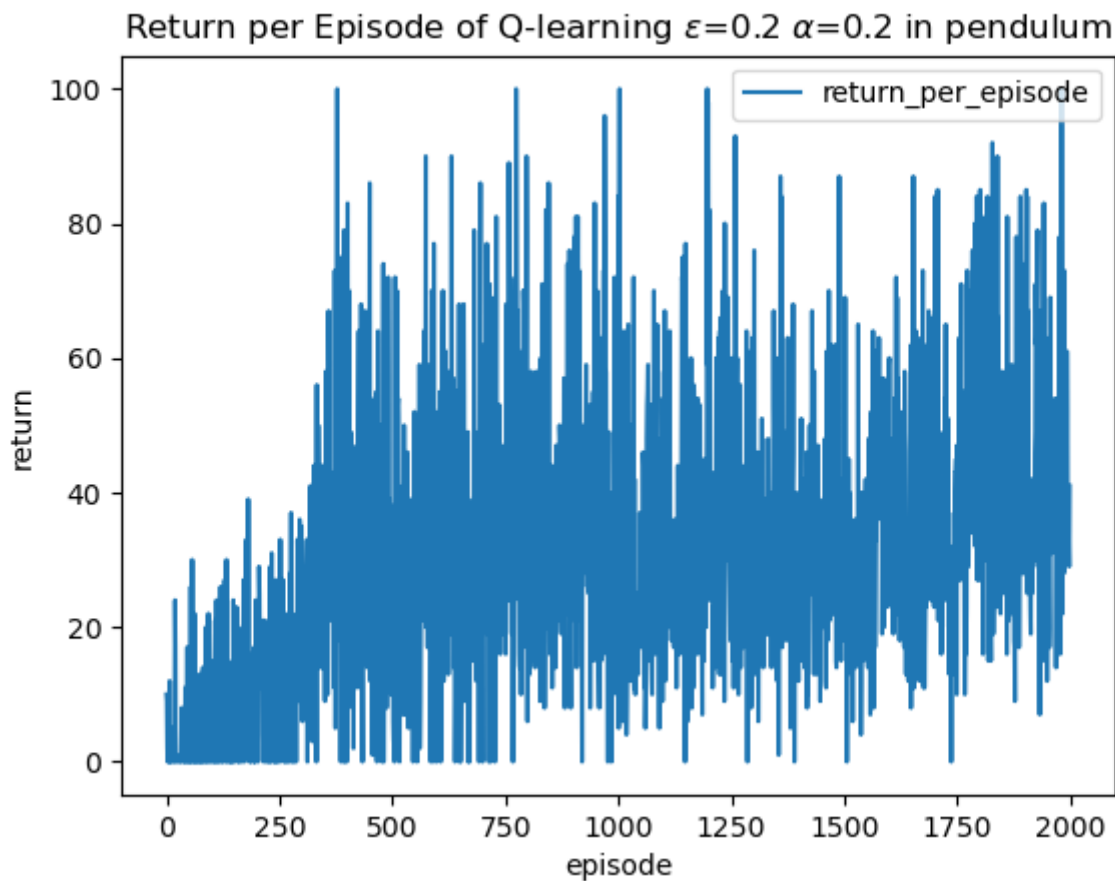


Pendulum

**SARSA**



## Q-learning



## Learning curves for different values of $\epsilon$

### Gridworld

#### SARSA

![sarsa\_eps\_gridworld](./figures/Return per episode for different epsilon in gridworld of SARSA.png)

#### Q-learning

![qlearning\_gridworld](./figures/Return per episode for different epsilon in gridworld of Q-learning.png)

### Pendulum

#### SARSA

![sarsa\_eps\_pendulum](./figures/Return per episode for different epsilon in pendulum of SARSA.png)

#### Q-learning

![qlearning\_eps\_pendulum](./figures/Return per episode for different epsilon in pendulum of Q-learning.png)

## Learning curves for different values of $\alpha$

### Gridworld

## SARSA

![sarsa\_alpha\_gridworld](./figures/Return per episode for different alpha in gridworld of SARSA.png)

## Q-learning

![qlearning\_alpha\_gridworld](./figures/Return per episode for different alpha in gridworld of Q-learning.png)

## Pendulum

## SARSA

![sarsa\_alpha\_pendulum](./figures/Return per episode for different alpha in pendulum of SARSA.png)

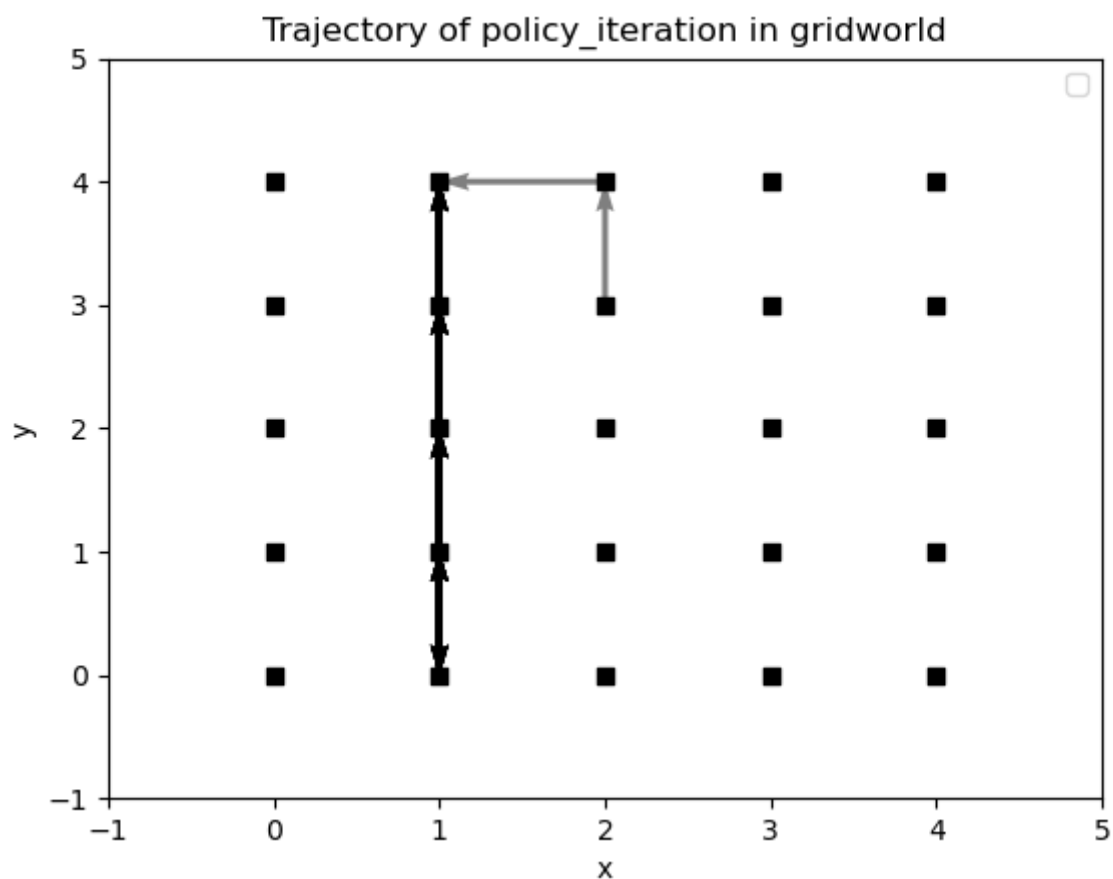
## Q-learning

![qlearning\_alpha\_pendulum](./figures/Return per episode for different alpha in pendulum of Q-learning.png)

## Example Trajectories

### Gridworld

### Policy iteration

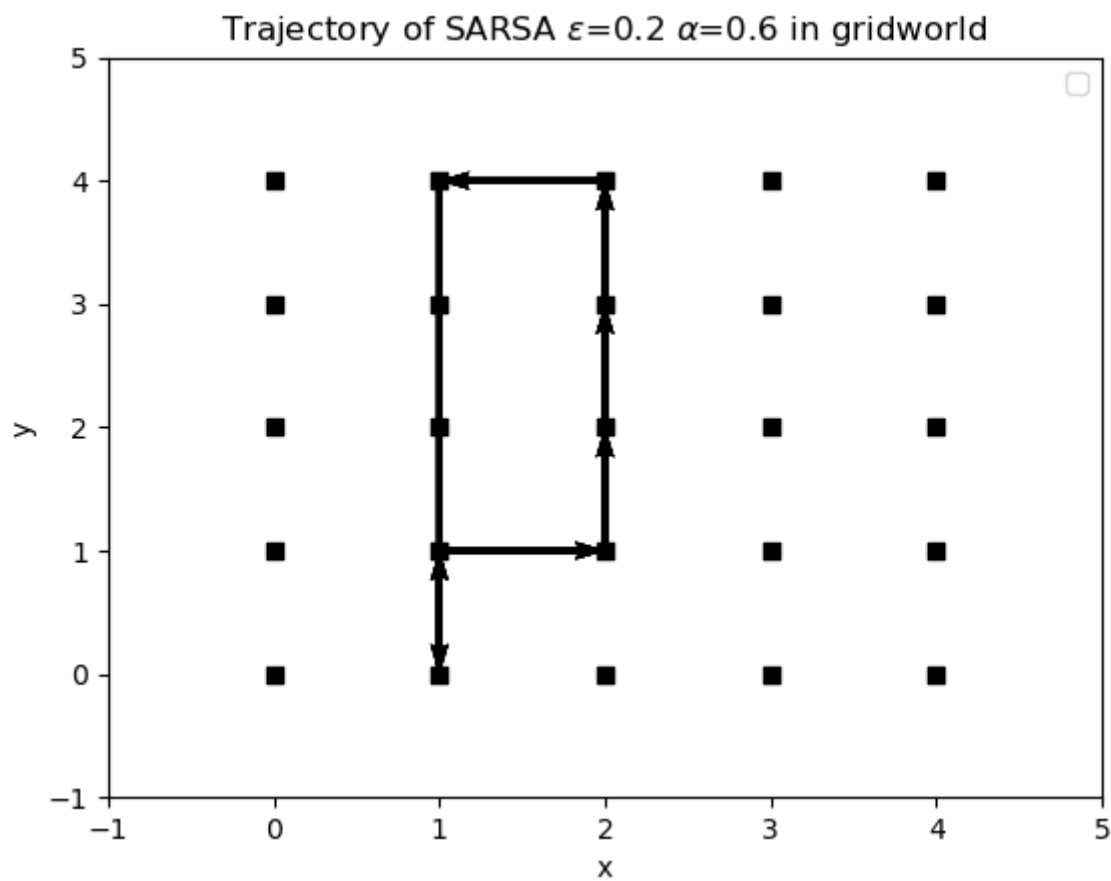


Value iteration



SARSA



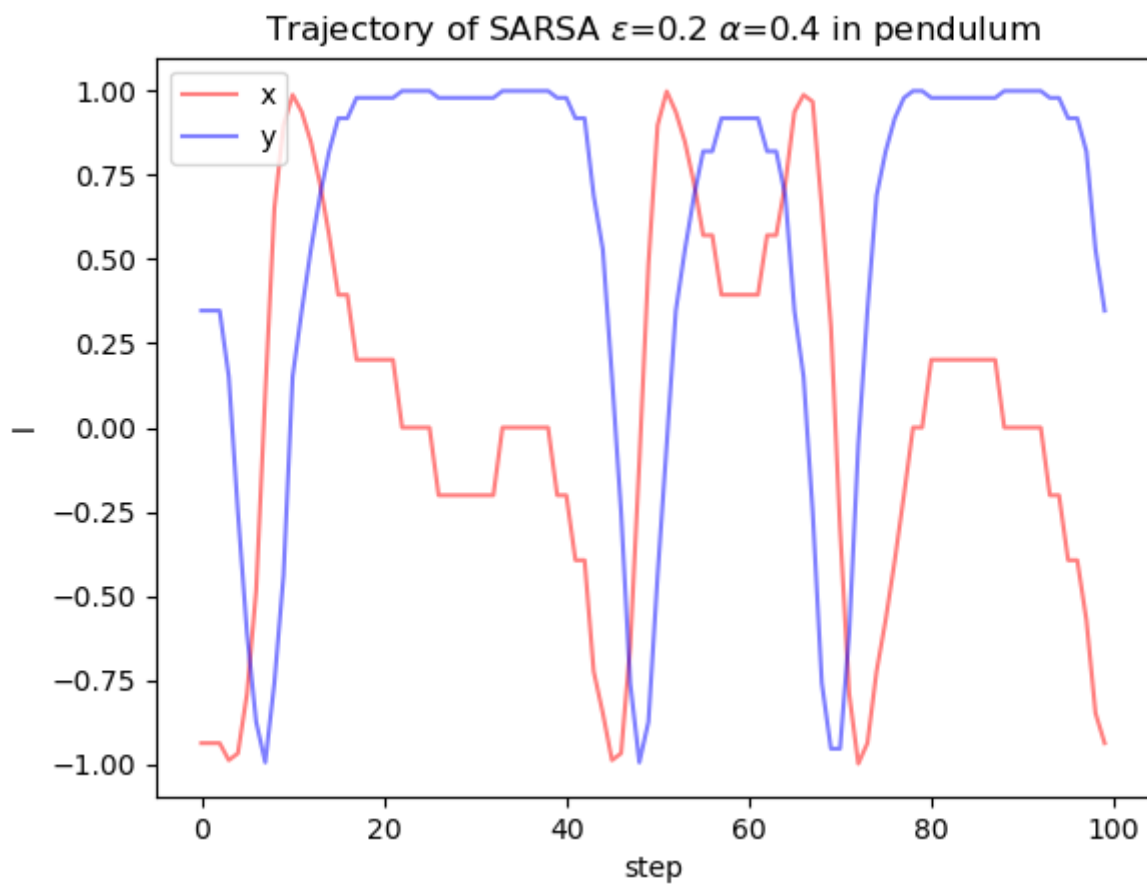


Q-learning

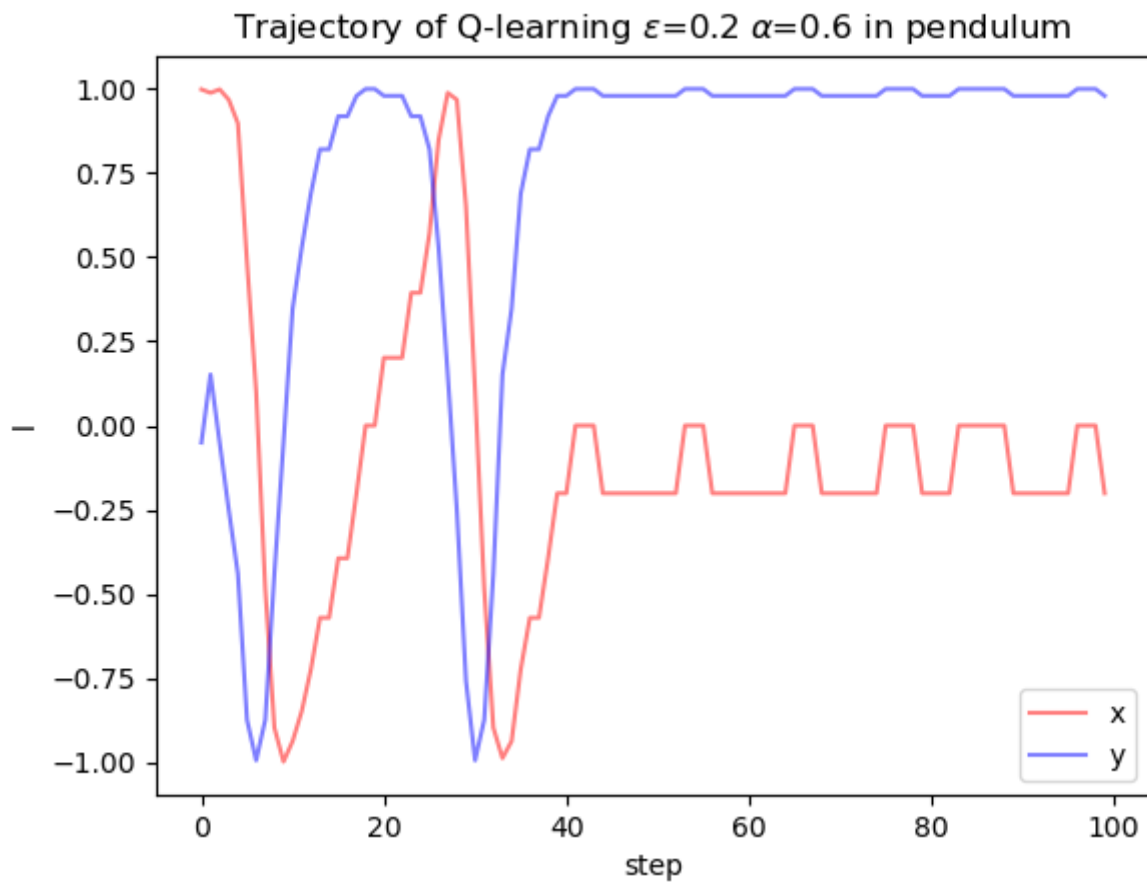


Pendulum

**SARSA**



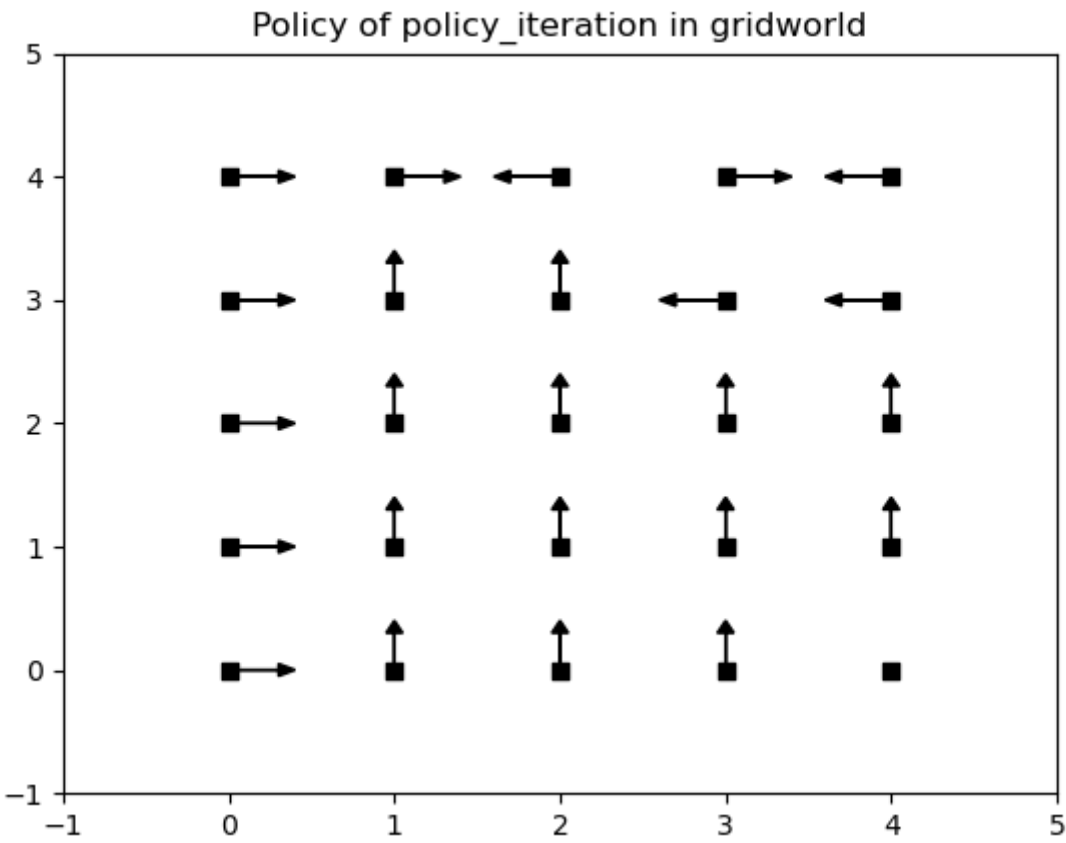
## Q-learning



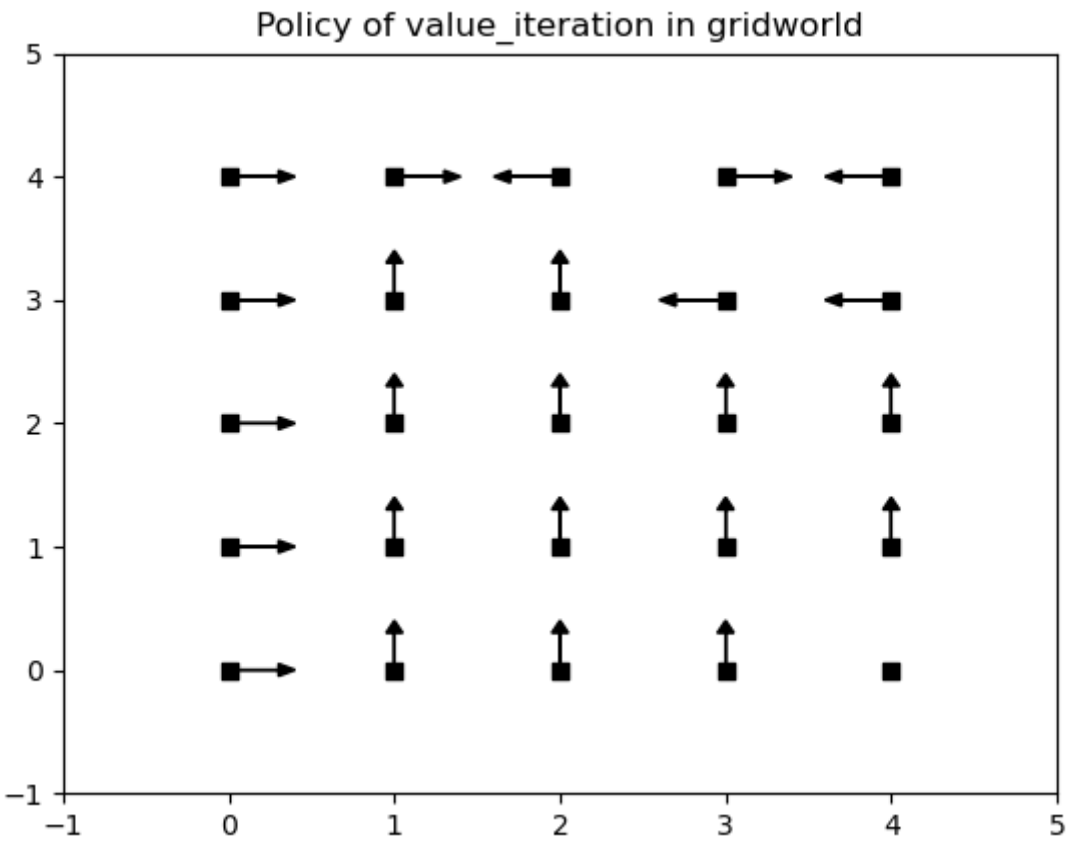
## Policy

Gridworld

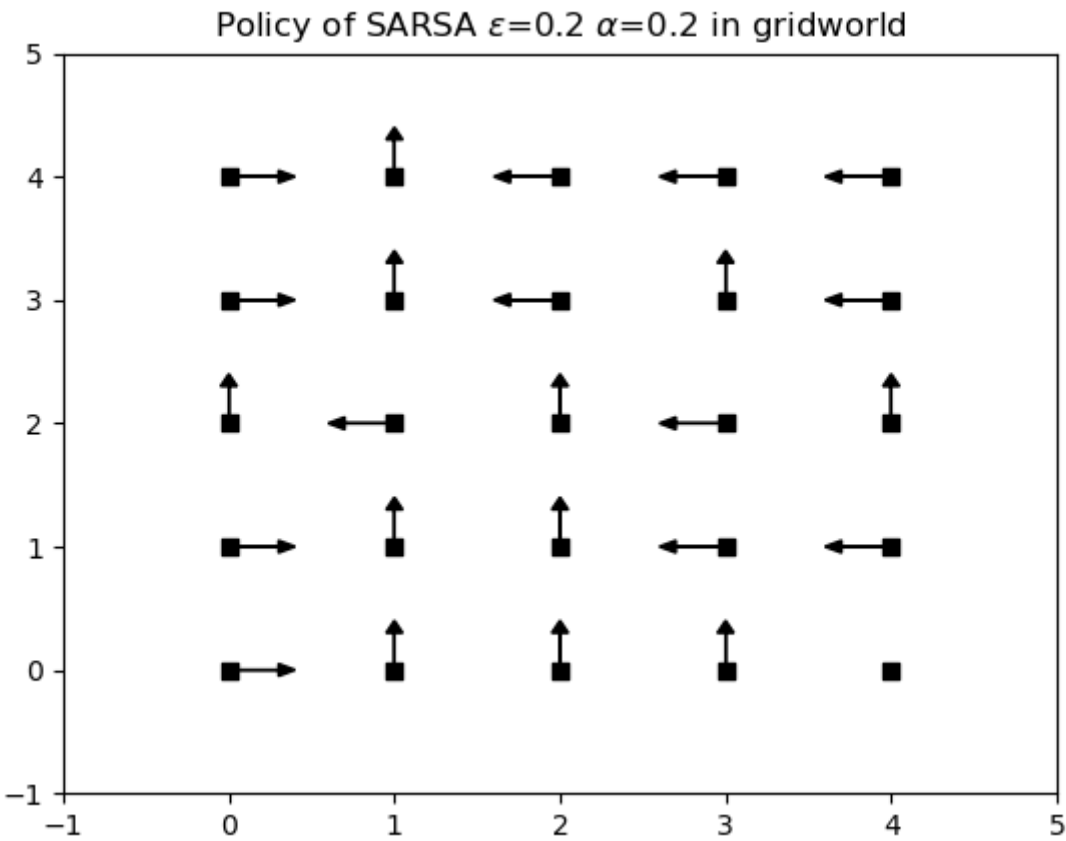
**Policy iteration**



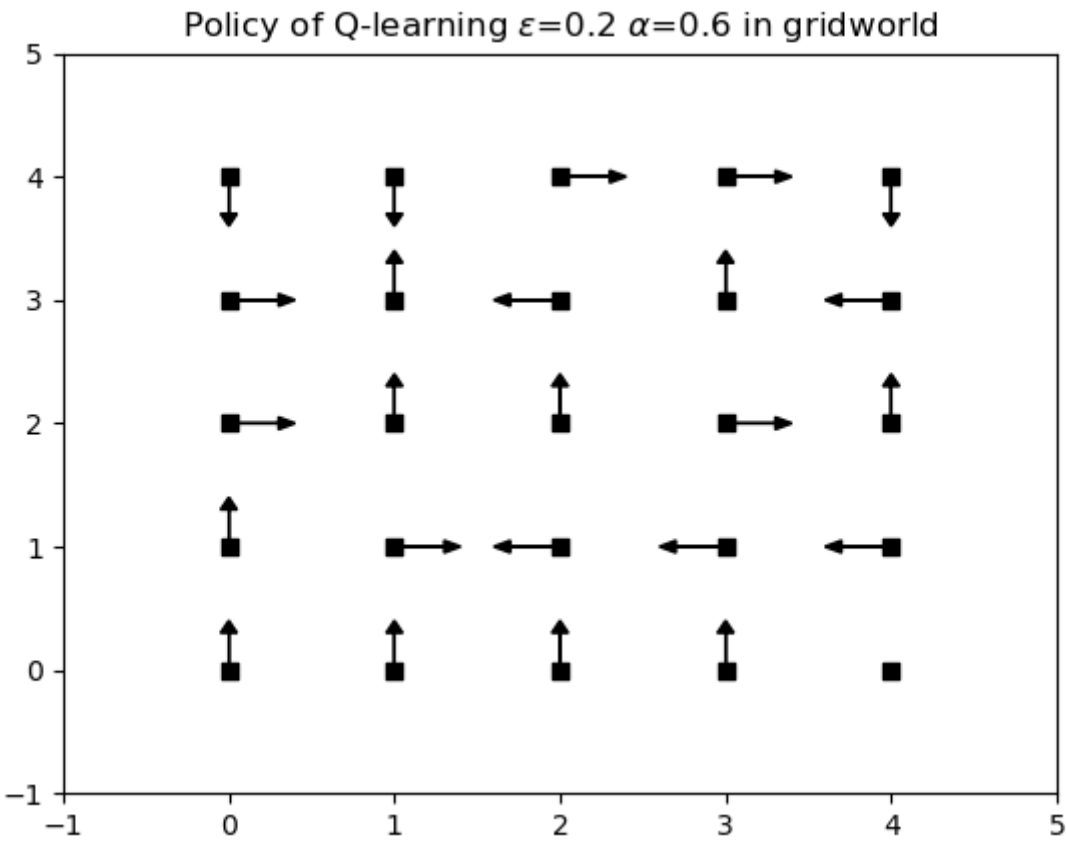
Value iteration



**SARSA**



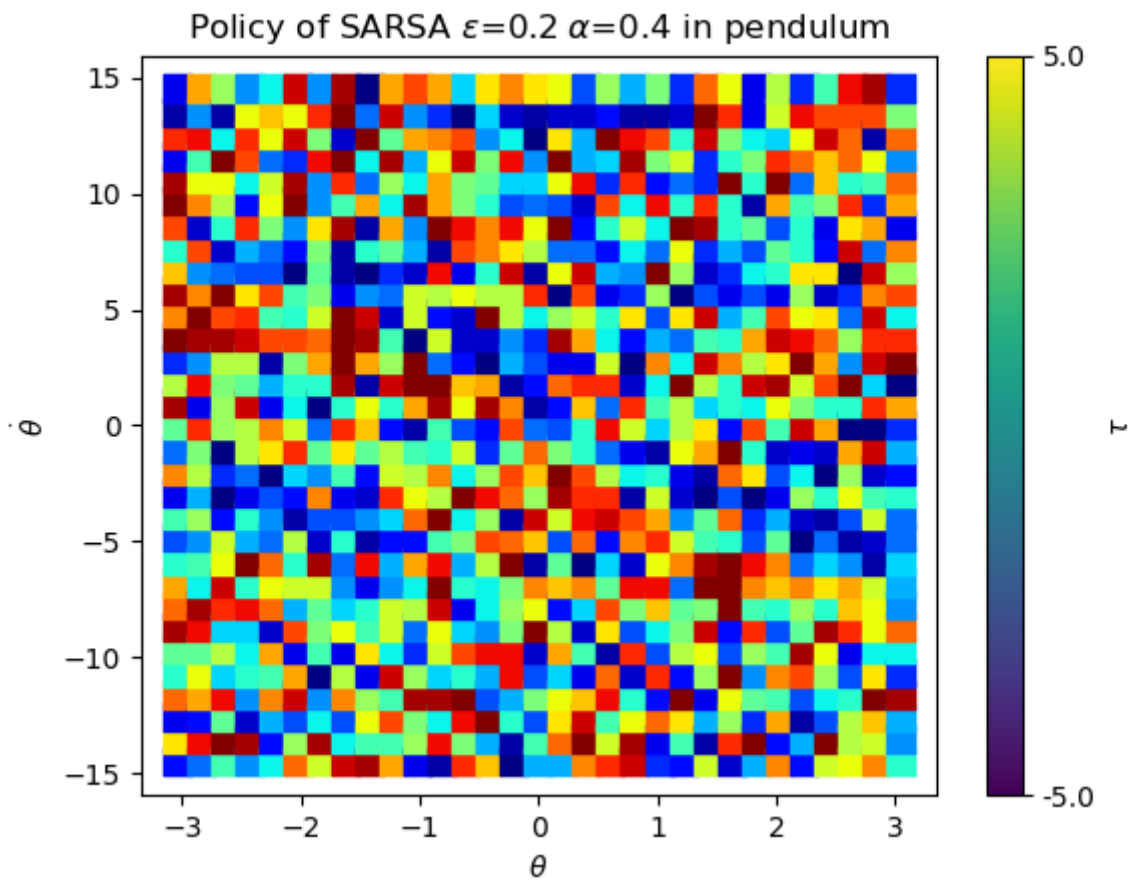
Q-learning



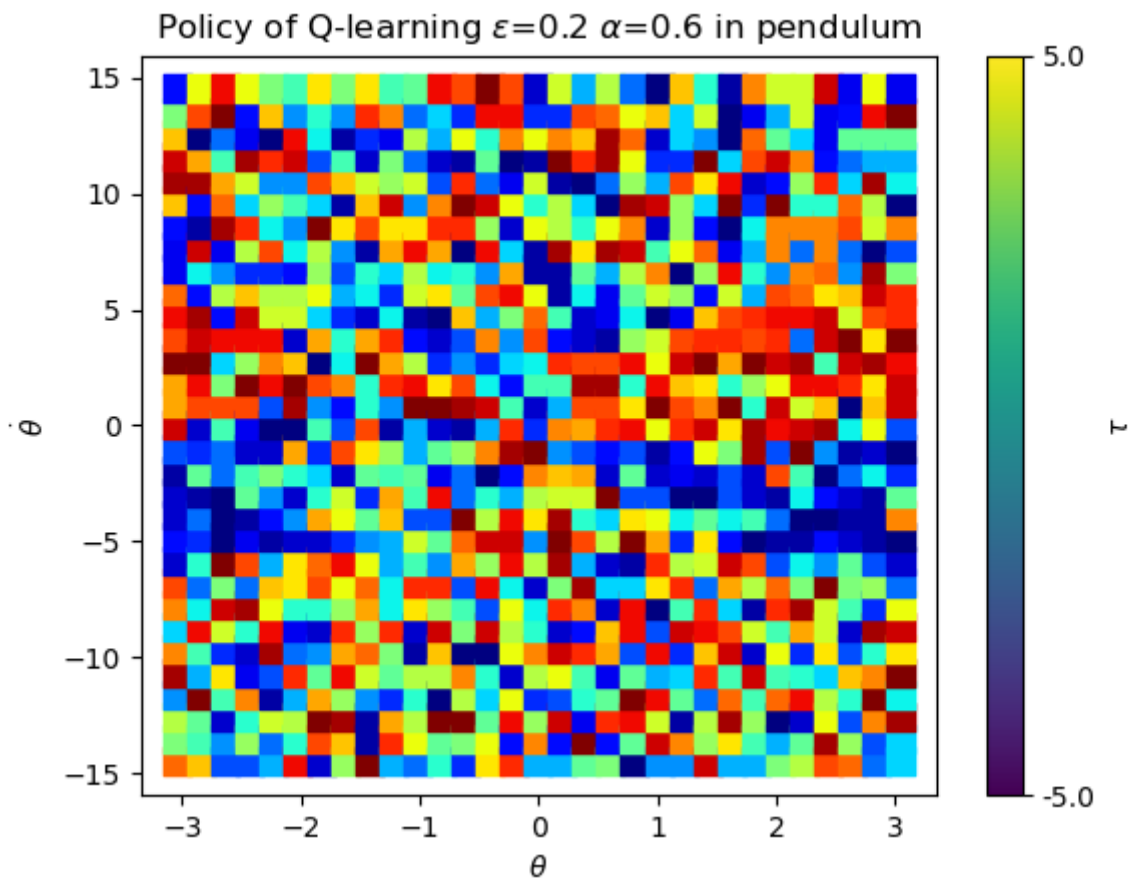
Pendulum

**SARSA**





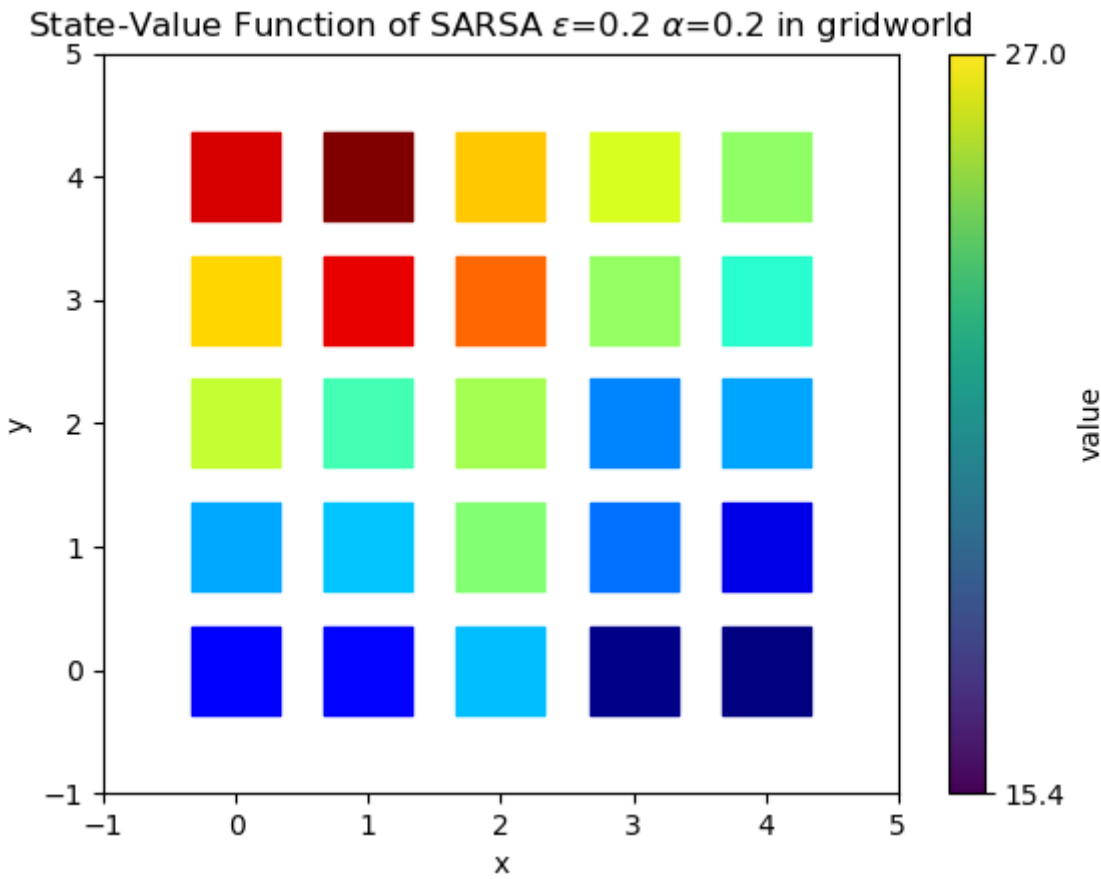
**Q-learning**



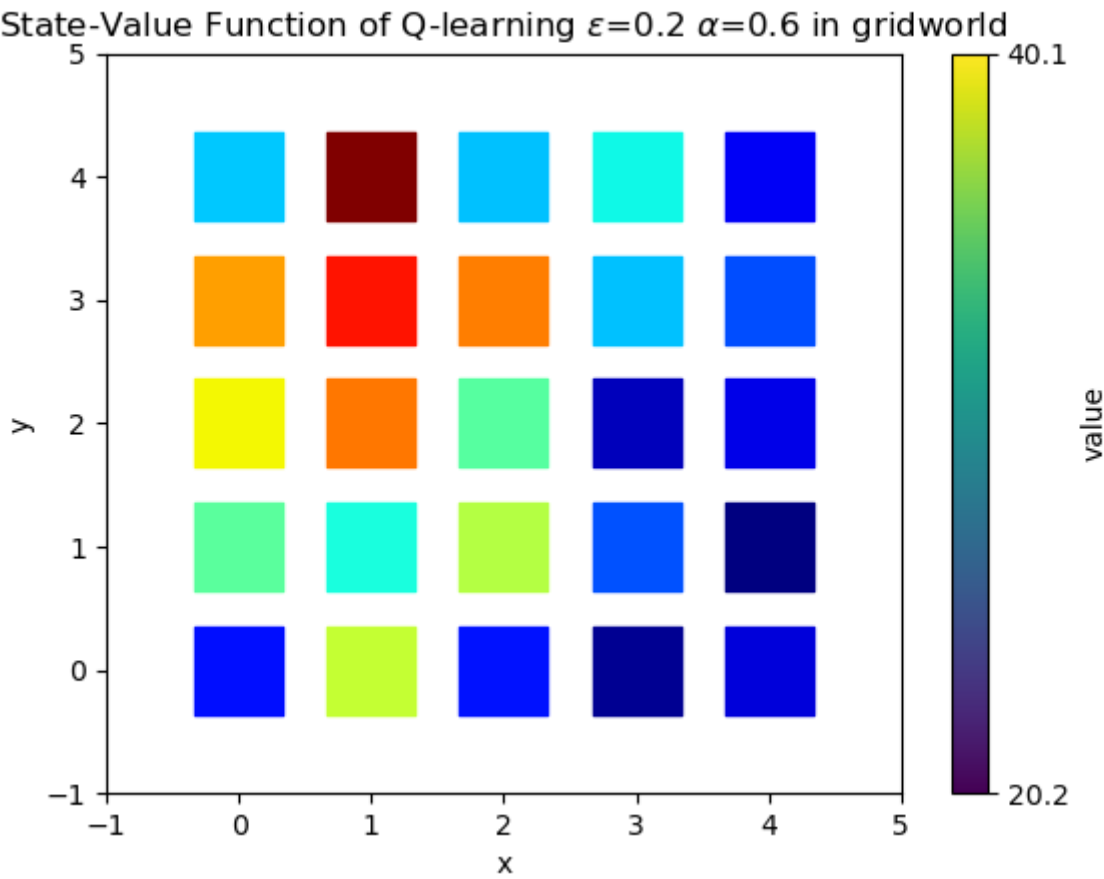
## State-Value Function

Gridworld

**SARSA**

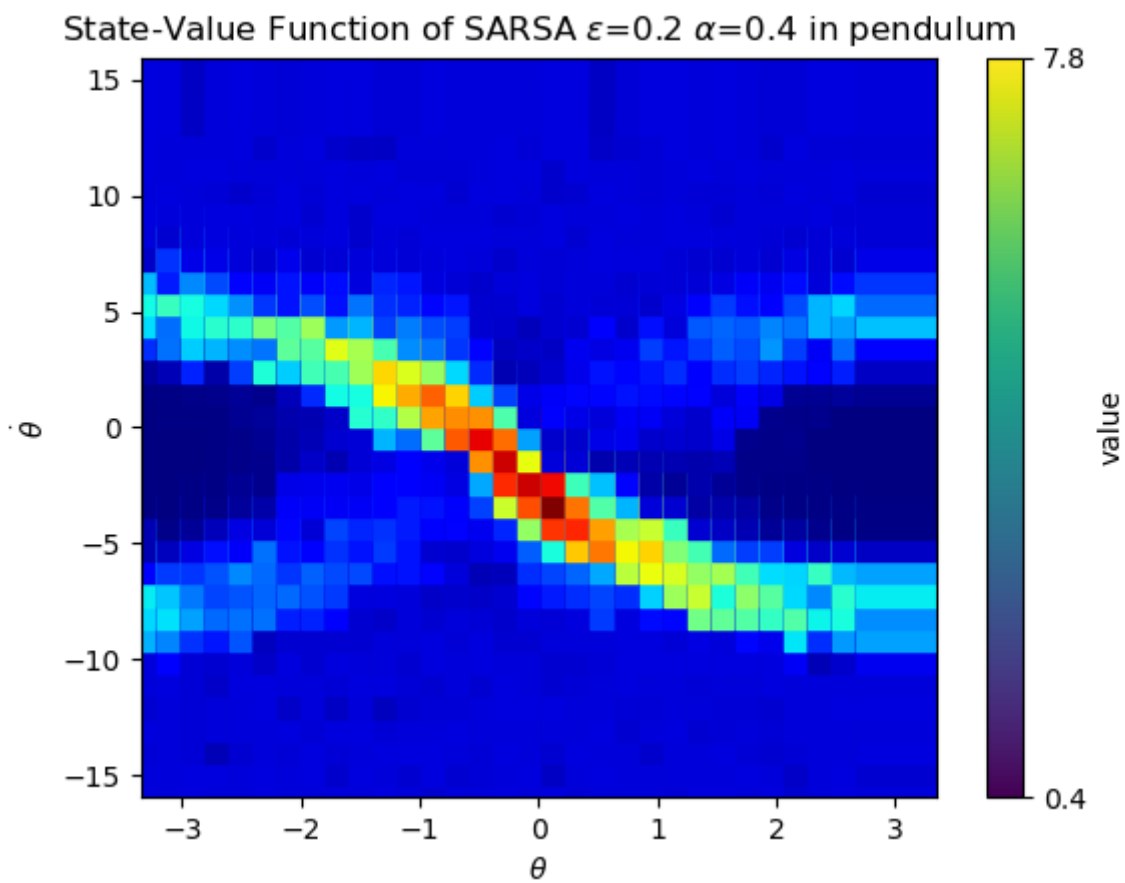


Q-learning



Pendulum

**SARSA**



Q-learning

