CS747 : Foundations of Intelligent & Learning Agents

Programming Assignment 3

REPORT

Vinit P. Doke
190260018[1]

[1] Email: 190260018@iitb.ac.in

November 2, 2021

Content

1 Task 1 : SARSA Tabular

1.1 State Discretization and Representation

The position space varies from $[-1.2, 0.6]$ and is divided into **25 uniform segments**.
The velocity space varies from $[-0.07, 0.07]$ and is divided into **25 uniform segments** too.

The index of the final state is then calculated as the follows:
The bin numbers of the (x, v) observation are $(x^{'}, v^{'})$ and the index is then

$$i = x^{'} + (v^{'} - 1) \times 25$$

The corresponding weight is stored in a 625×3 Matrix : $25 \times 25 \times 3$ $(Actions)$

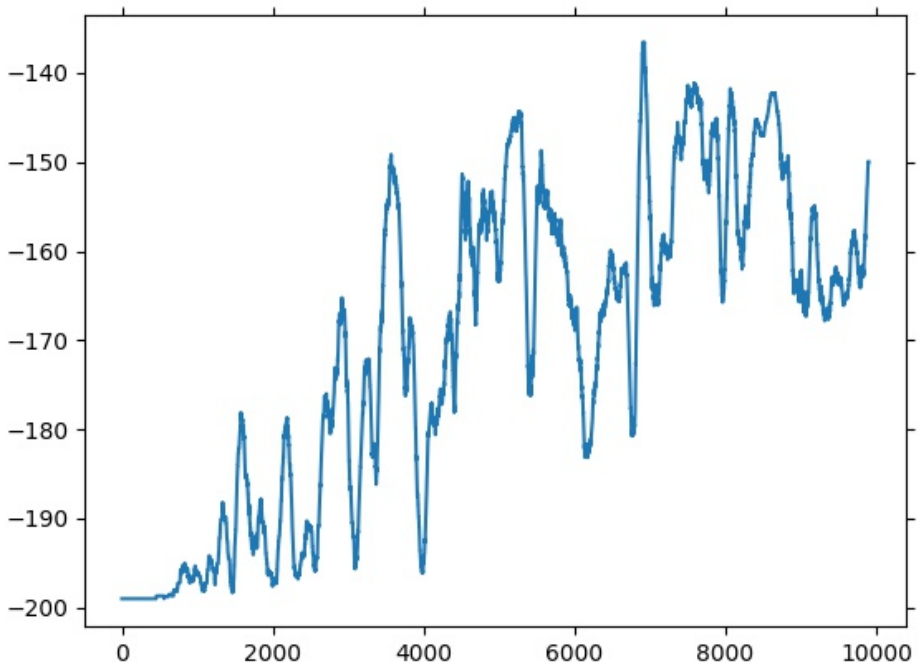
1.2 Other Hyper-parameters

- Learning Rate $\alpha = 0.1$
- Discount Factor $\gamma = 0.9$
- Epsilon $\epsilon = 0.001$

1.3 Plots & Results

Average Reward for 100 Episodes is :

-144.78



(a)

Figure 1: On X - Axis : Number of Episodes
On Y - Axis : Reward

2 Task 2 : SARSA with Linear Function Approximation

2.1 Discretization & Tile Coding

Number of Tilings Used : **12**
Divisions per Tile per Dimension (both position and velocity) : **25**
Position Offset : **0.0065**
Velocity Offset : **0.0005**
3D Weight Matrix of dimensions : $[25 \times 25, 12, 3]$

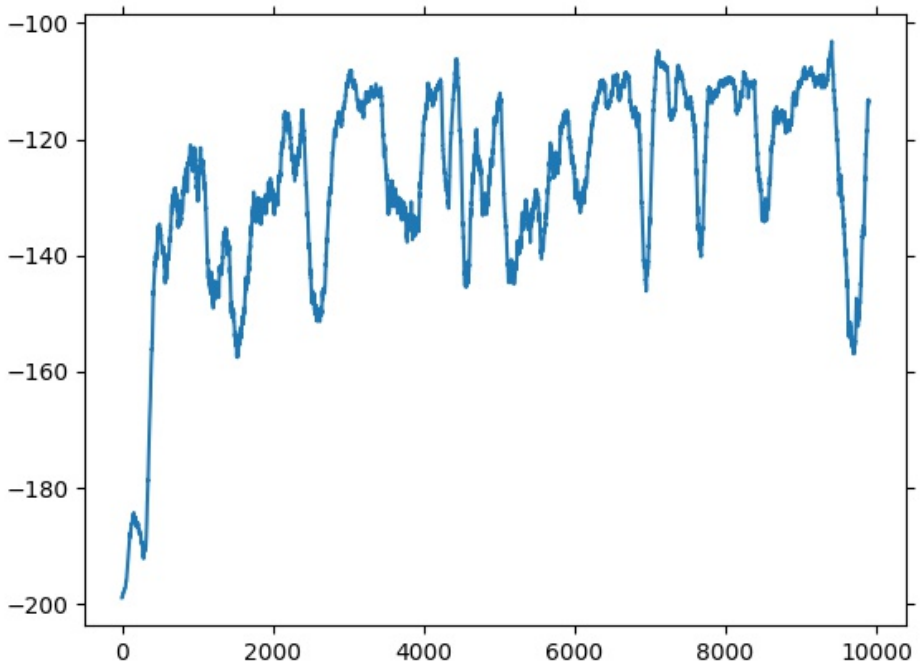
2.2 Other Hyper-parameters

- Learning Rate $\alpha = 0.1$
- Discount Factor $\gamma = 0.9$
- Epsilon $\epsilon = 0.001$

2.3 Plots & Results

Average Reward for 100 Episodes is :

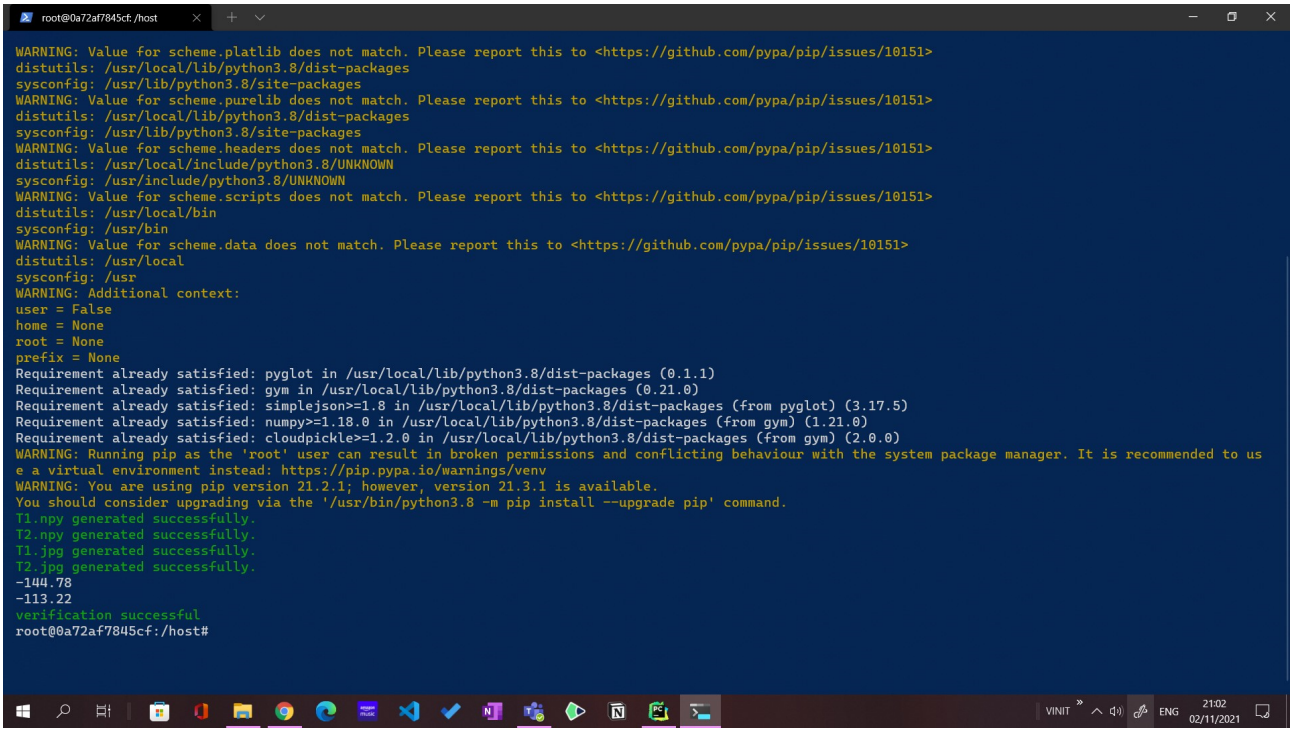
-113.22



(a)

Figure 2: Task 2 : SARSA with Tile Coding
On X - Axis : Number of Episodes
On Y - Axis : Reward

3 Screenshot of verifyOutput.sh



(a)

References

1. Tile Coding
https://towardsdatascience.com/reinforcement-learning-tile-coding-implementation-7974b600762b

2. SARSA with LFA
https://artint.info/2e/html/ArtInt2e.Ch12.S9.SS1.html

3. Book : Sutton & Barto

4. OpenAI Gym
https://gym.openai.com/docs/