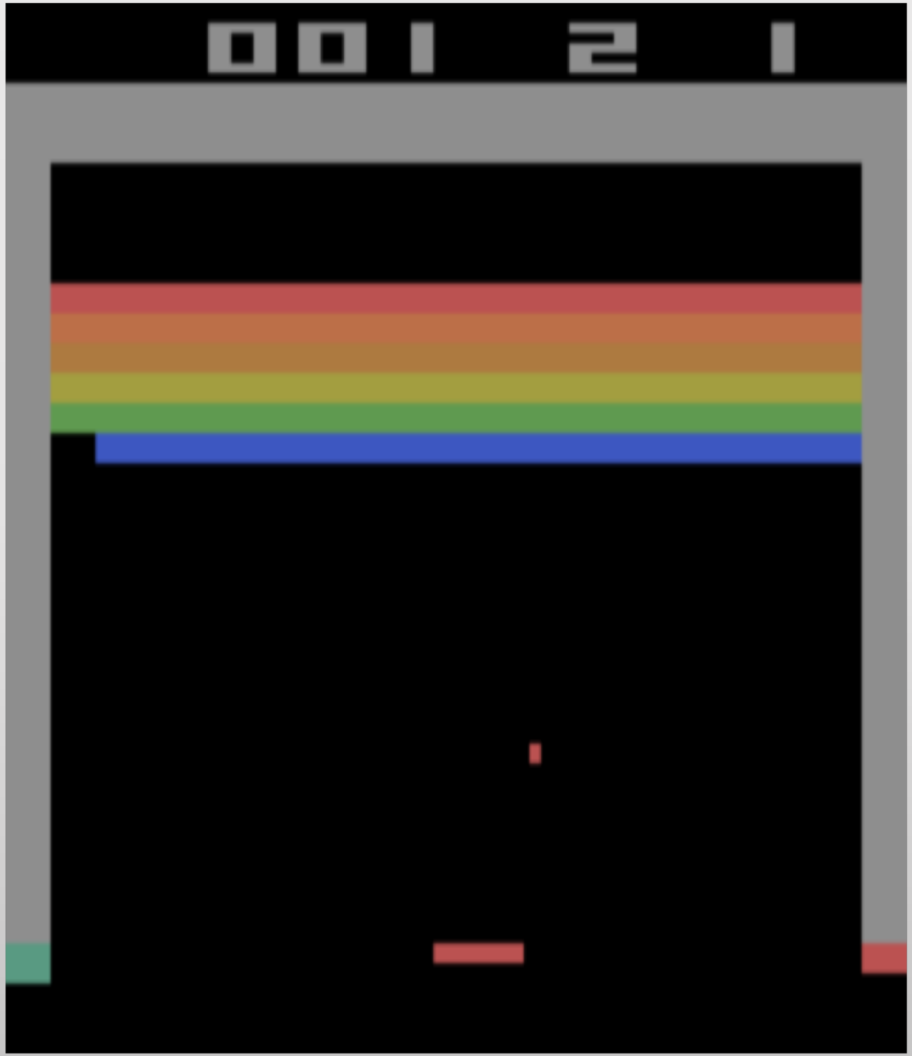


Breakout

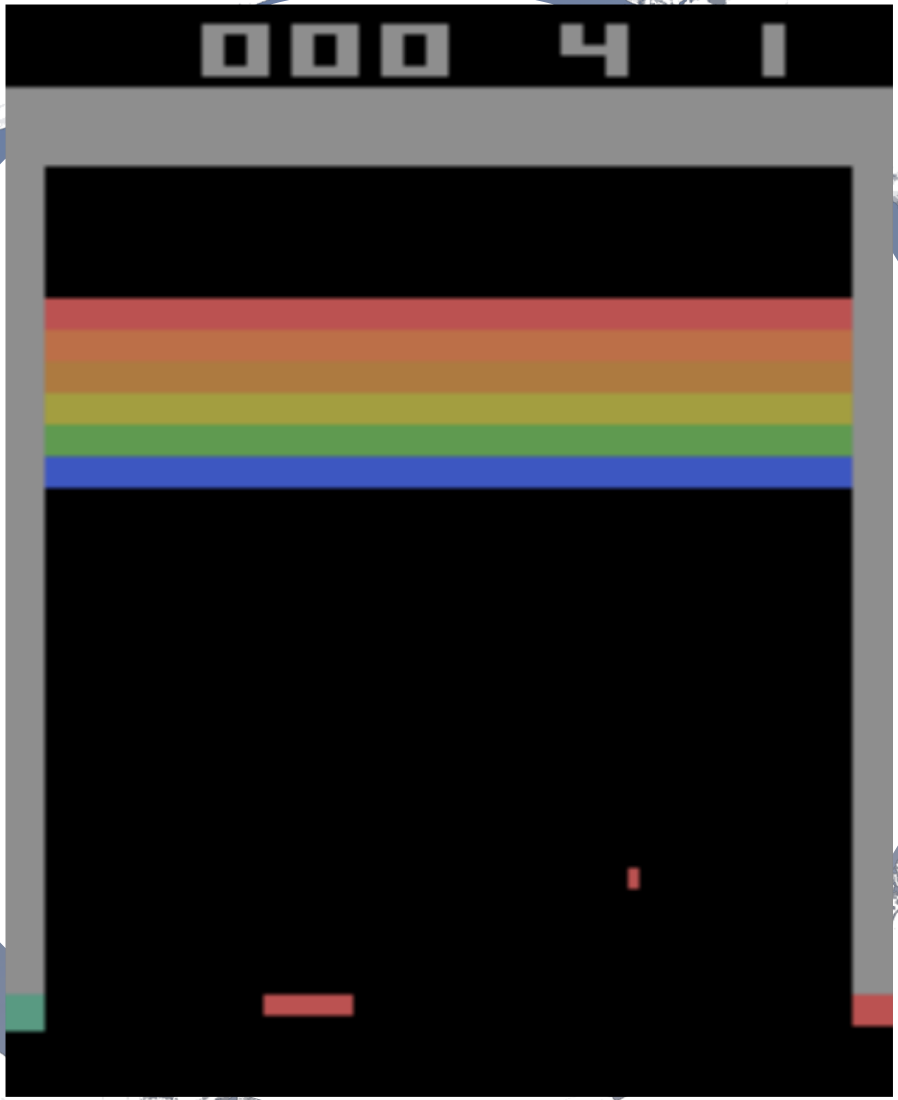
By,

Sai Krishna Theja Bhavaraju

Rohit Vavaldas

Contents

- Project domain
- Experiments
- Learning strategies
- Results
- Discussion
- Future work



Project Domain

Objects: Paddle, Ball and Stack of Bricks

Agent: Paddle

Goal: To score maximum points

Constraints: Only 5 lives available.

End Of Game: When all the bricks are broken, or the agent loses its 5 lives.

Approaches

Complete State space:

All possible combinations of the ball and the paddle.

Initial state space: 5 million

Reduced State space:

Factors considered in forming the reduced state space:

- Paddle tracking
- Ball tracking
- Movement tracking
- Constrained frame area

Reduced state space: 260,000

Frame skipping:

Environment returns 60 frames per second

Consider every 4th frame rather than every frame

Action space: "Noop", "Fire", "Move right", "Move left"

Experiments

- Experiment 1: Approach considering complete state space, each frame as a state
- Experiment 2: Approach considering reduced state space with constrained ball and paddle tracker(relative distance between them, in each frame)
- Experiment 3: Approach considering reduced state space with user defined rewards(relative positions of ball and paddle in two consecutive frames)
- Experiment 4: Approach considering reduced state space with frame skipping and user defined reward strategies(virtual replay memory)

Sample Q-Table:

Actions:

- 0 – NOOP
- 1 – FIRE
- 2 – RIGHT
- 3 – LEFT

	0	1	2	3
002225stationary	0.383017	2.881860	0.000000	0.000000
002326stationary	1.092599	1.626729	0.975501	0.175917
002427stationary	2.199810	2.037096	1.851212	1.946144
002528stationary	2.265779	2.007653	2.596364	2.715576
002629stationary	0.826841	1.305755	1.927002	2.148532
002730stationary	2.361902	1.293474	1.576076	0.986547
002831stationary	0.292200	2.340712	1.081521	0.055555
24101013moveleft	0.012853	0.000000	0.000000	0.000000
24101013moveright	0.480000	0.000000	0.000000	0.000000
24101013stationary	0.019205	0.000000	0.000000	0.000000
24101114moveleft	0.230825	0.000000	0.000000	0.000000
24101114moveright	0.604856	0.000000	0.000000	0.000000
24101114stationary	0.237926	0.000000	0.000000	0.000000
24101215moveleft	0.000000	0.000000	0.000000	0.000000
24101215moveright	0.000000	0.000000	0.000000	0.000000
24101215stationary	0.025833	0.000000	0.000000	0.000000

Learning Methods

Q-Learning

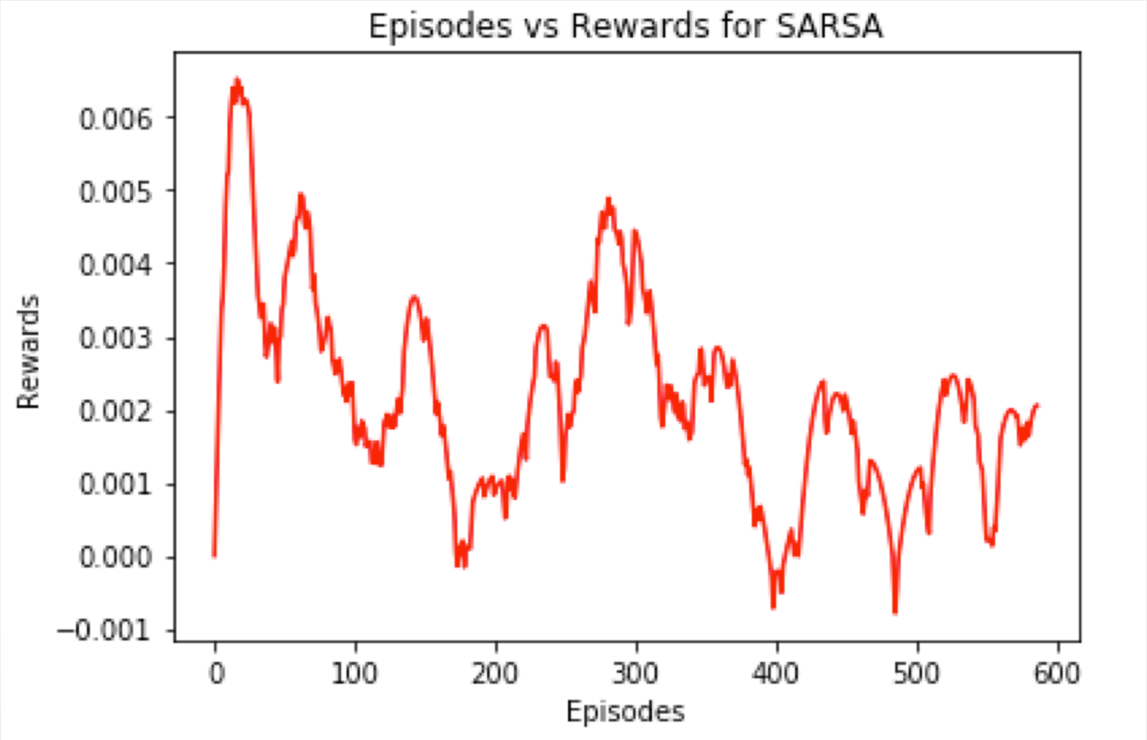
SARSA

Reasons to consider Q learning and SARSA:

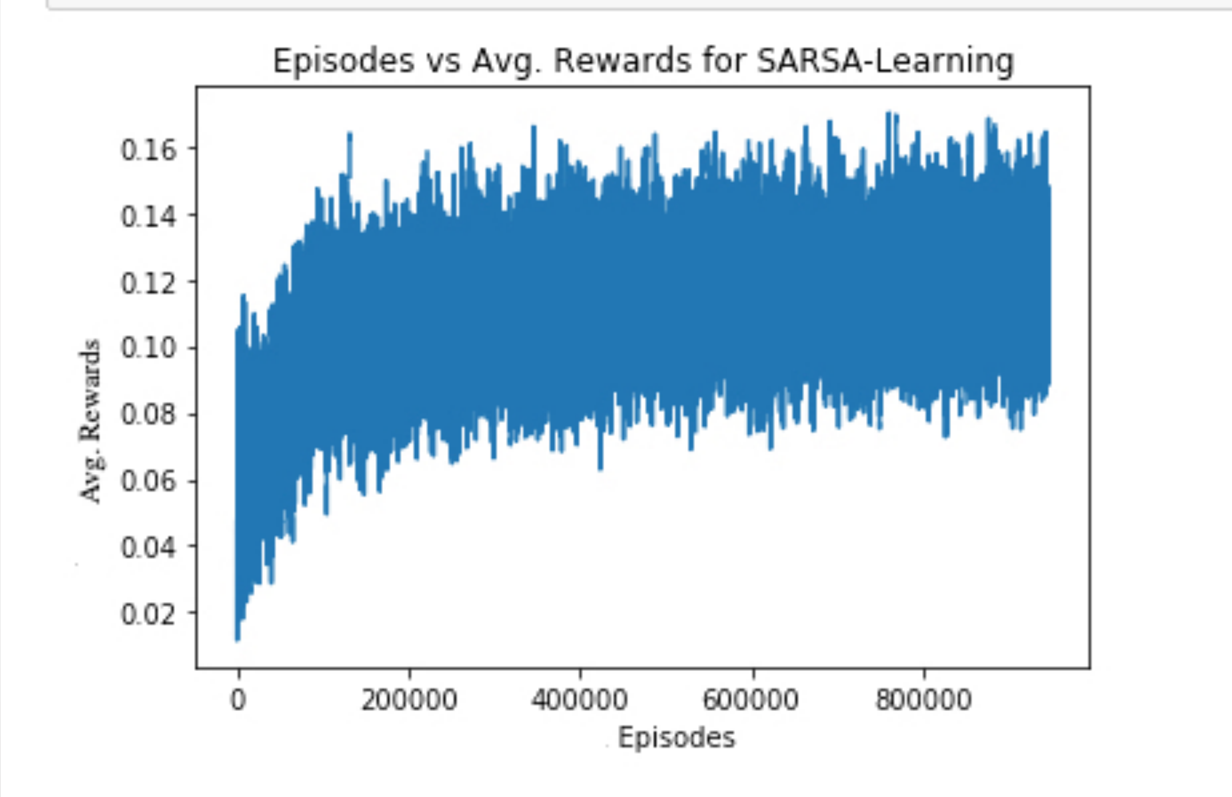
- Model free learning algorithm

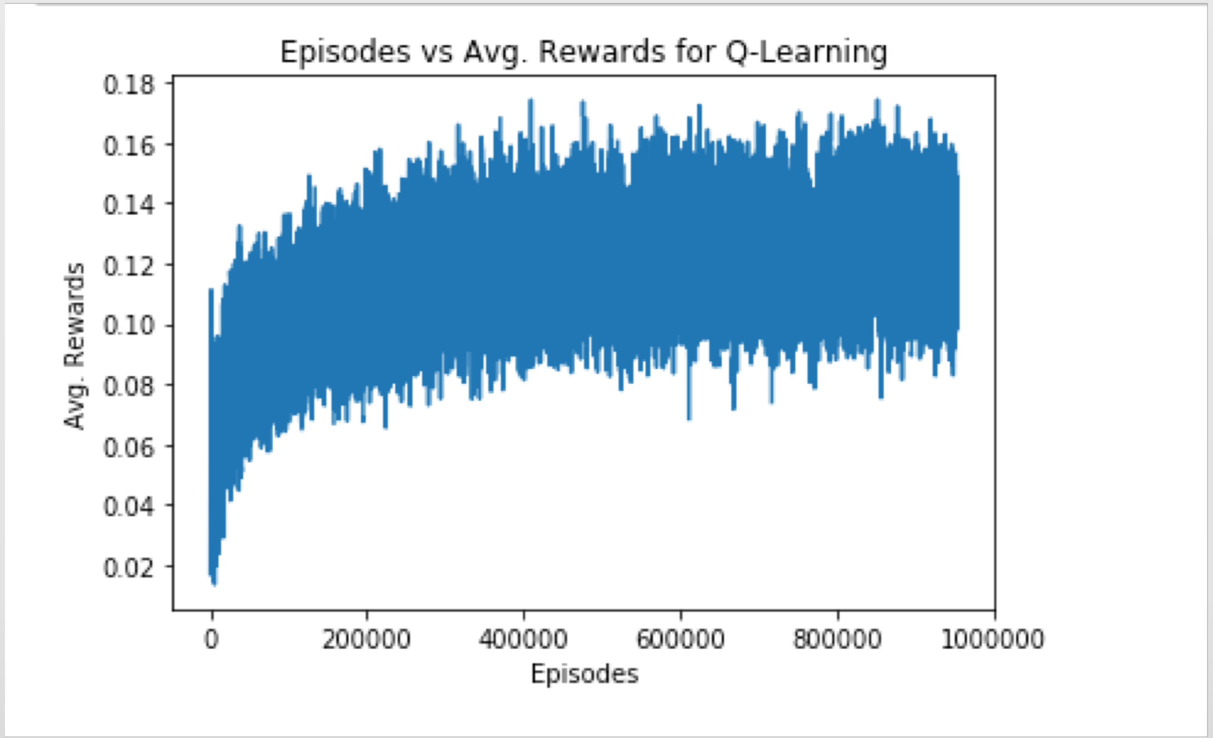
Metrics analyzed:

- Average Rewards per game

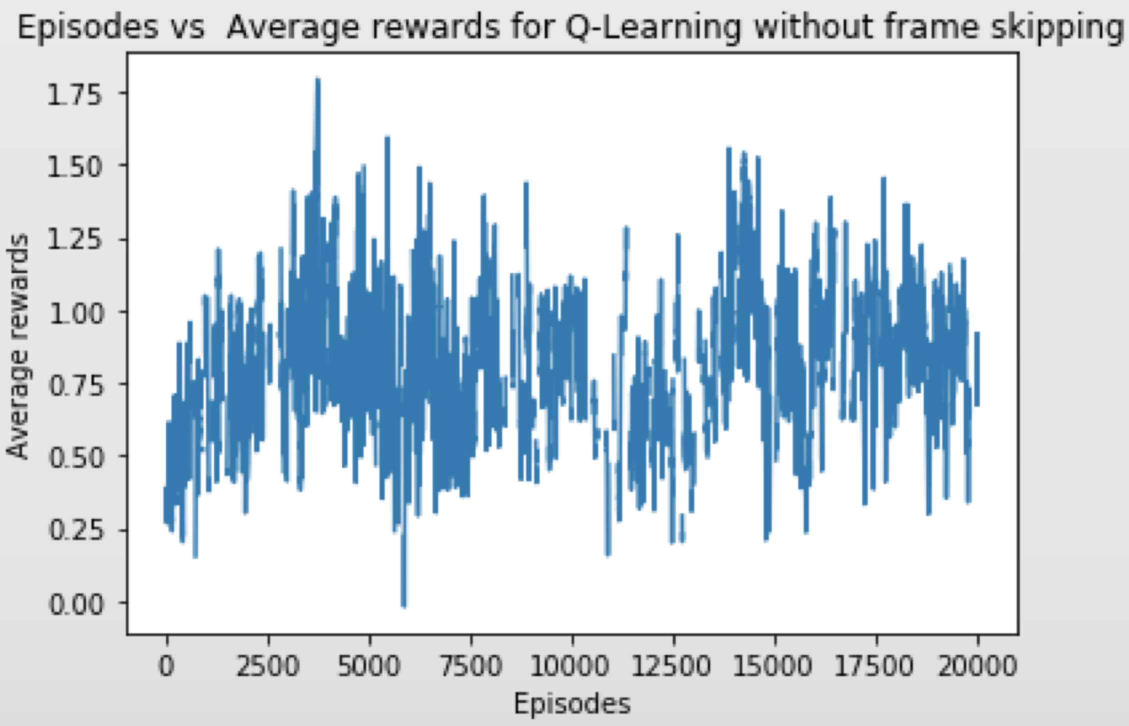


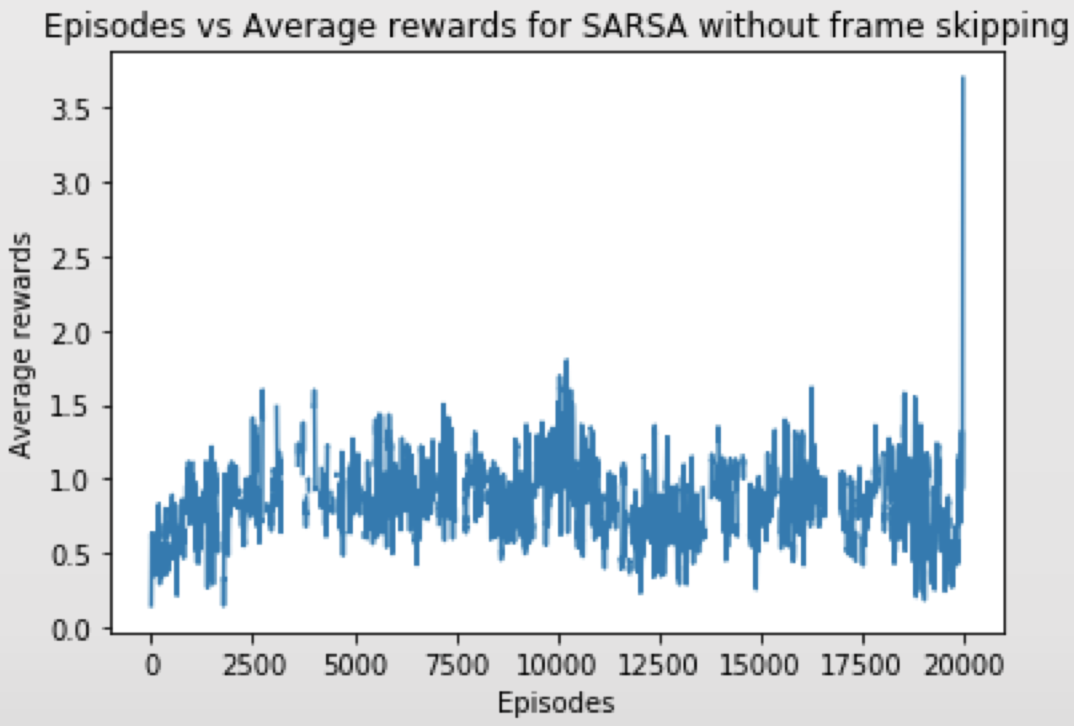
Results for Q-Learning vs SARSA (Experiment 1)



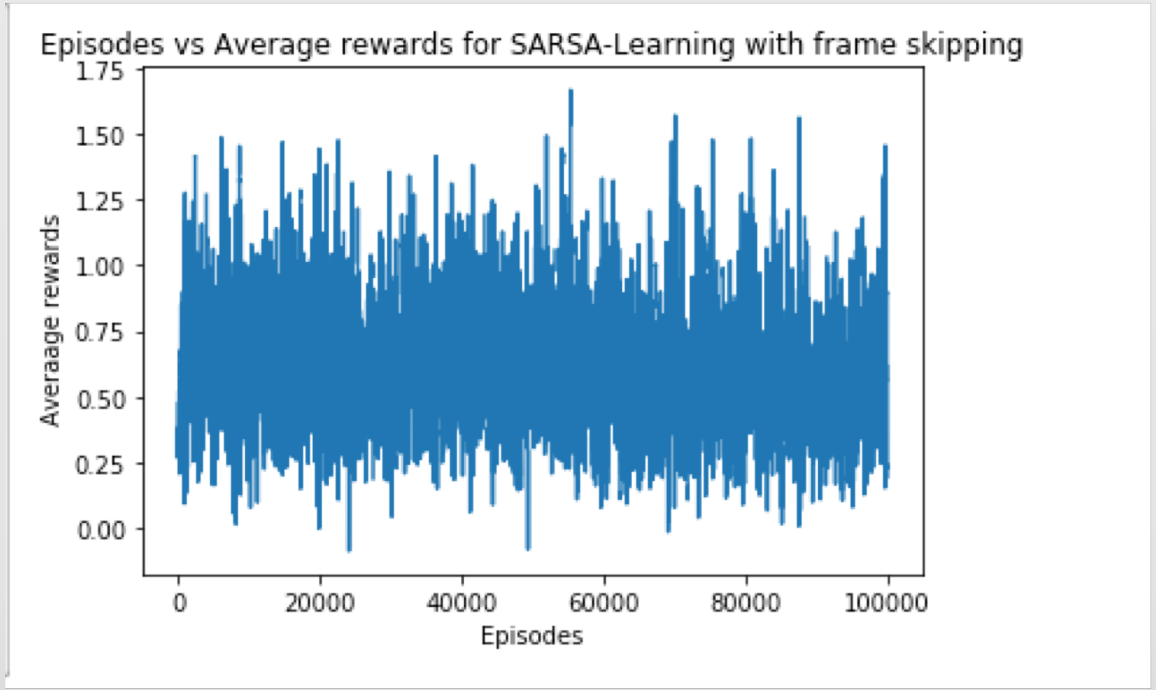


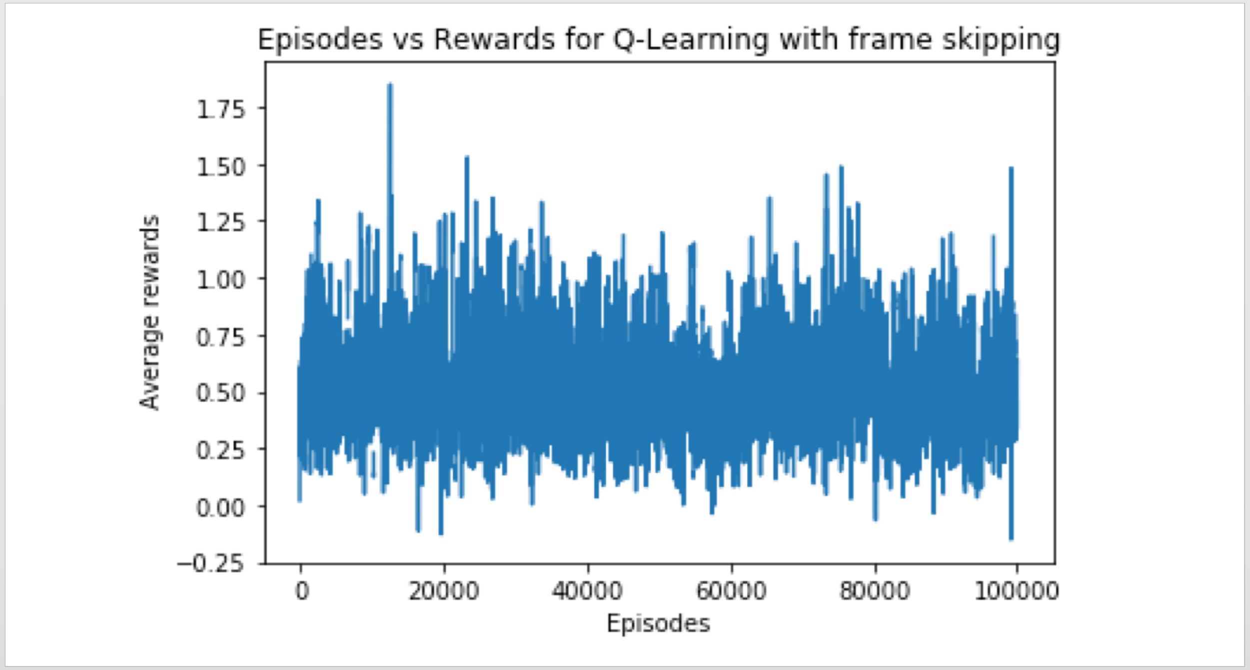
Results for Q-Learning vs SARSA (Experiment 2)





Results for Q-Learning vs SARSA (Experiment 3)





Results for Q-Learning vs SARSA (Experiment 4)

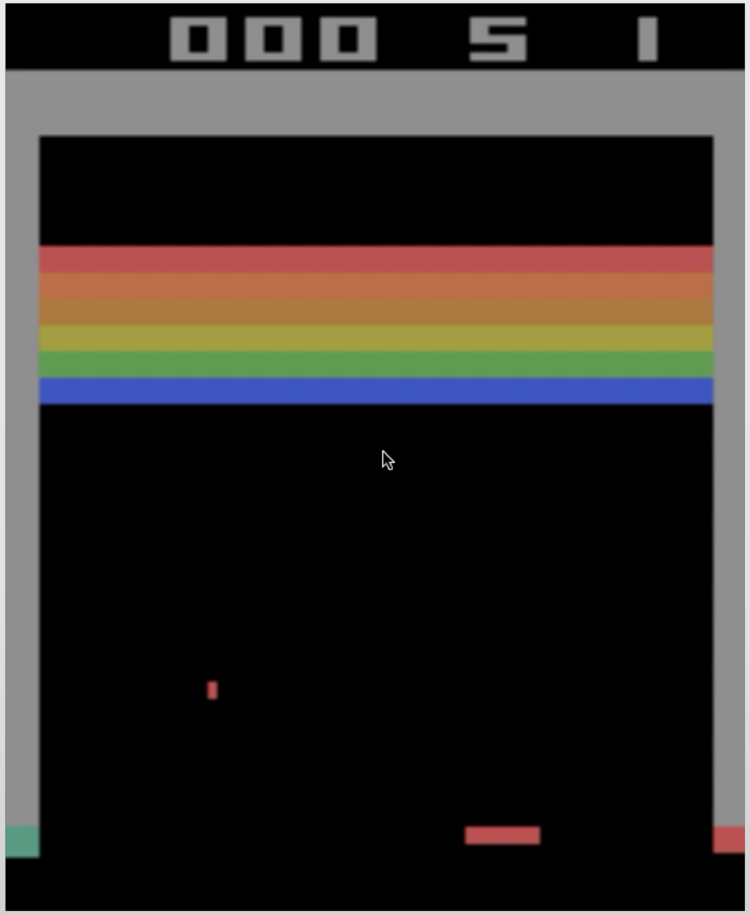
Demo:

Challenges faced:

- Observation space returned by the environment
- No features pertaining to objects are returned by the environment
- Huge state space

Reason for failure:

- Lack of ball history for the agent



Existing work

[1] Researchers at Stanford have implemented their own environment for the game Breakout that returns a set of features such as game state, ball position, paddle position, number of remaining bricks, ball velocity etc. Using these features, they have trained the agent by implementing variants of Q learning and SARSA.

[2] Deep Minds from Google have developed a neural network based reinforcement learning model called Deep Q Network to play Atari games.

Future Work

- Function approximation
- Deep Q Network

References:

[1] Mnih, Volodymyr, Kavukcuoglu, Koray, Silver, David, Graves, Alex, Antonoglou, Ioannis, Wierstra, Daan, and Riedmiller, Martin. Playing atari with deep reinforcement learning. *In NIPS Deep Learning Workshop*. 2013.

[2] Arun Nair, Praveen Srinivasan, Sam Blackwell, Cagdas Alcicek, Rory Fearon, Alessandro De Maria, Vedavyas Panneershelvam, Mustafa Suleyman, Charles Beattie, Stig Petersen, Shane Legg, Volodymyr Mnih, Koray Kavukcuoglu, David Silver. Massively Parallel Methods for Deep Reinforcement Learning. *International Conference on Machine Learning, Lille, France*, 2015.

Questions?