University of New Haven

TAGLIATELA COLLEGE OF ENGINEERING

POWER ON

Frogger Reinforcment Learning

DSCI-6612 | Thomas Powell

Project Objectives

- Primary Objective
  - Develop and train a reinforcement learning (RL) agent to successfully navigate and score in the Atari game Frogger.
- Sub-Objectives:
  - Implement an agent capable of learning optimal strategies using ALE's Frogger environment and OpenAI Gymnasium's reinforcement learning API.
  - Deploy a state-of-the-art RL algorithm, such as Q-Learning.
  - Analyze and optimize agent performance through iterative training.

aws
©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Environment

- OpenAI Gymnasium's Reinforcement Learning API
- Arcade Learning Environment for Atari Frogger.
- ALE's Python Interface

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

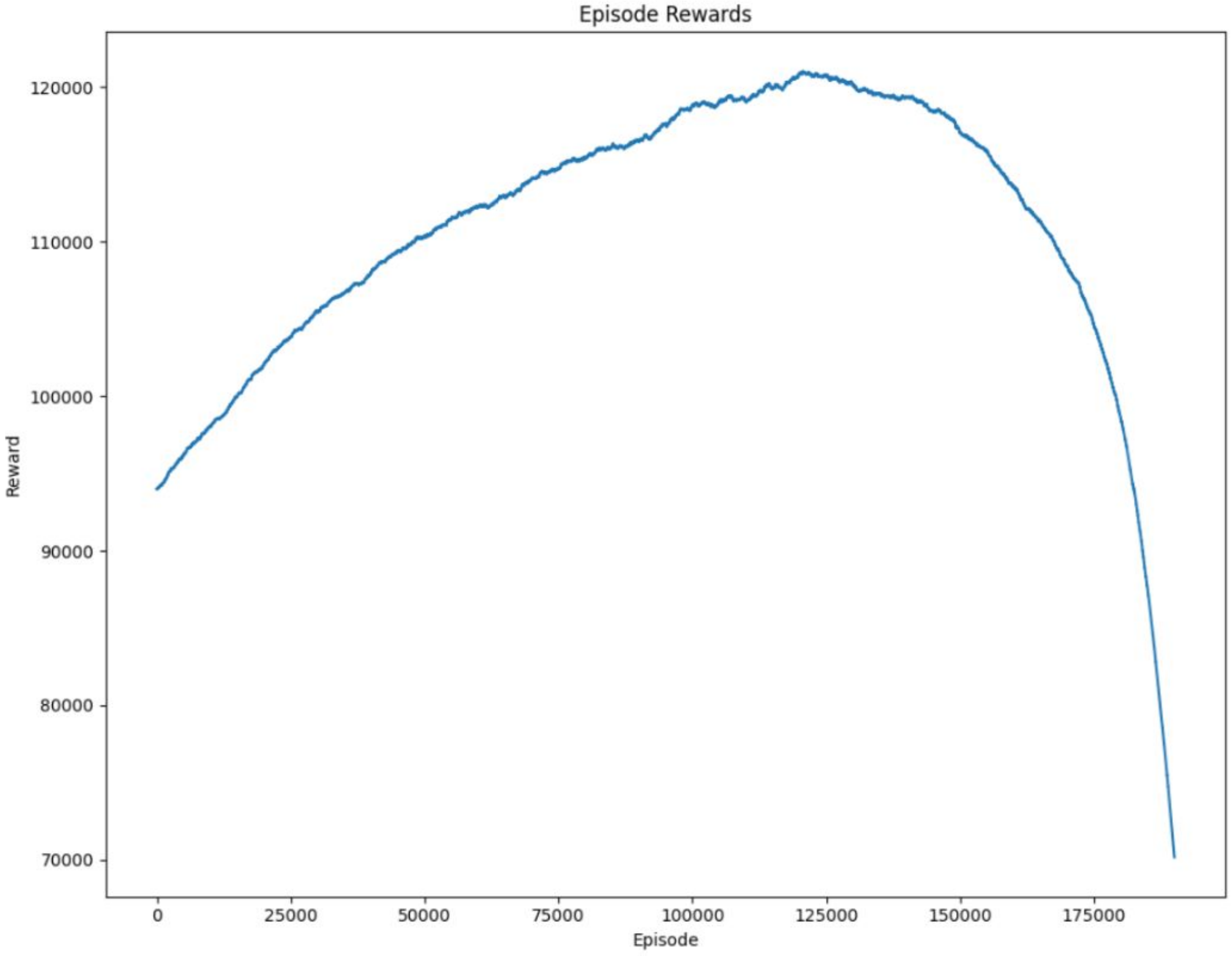
Tabular Q-Learning

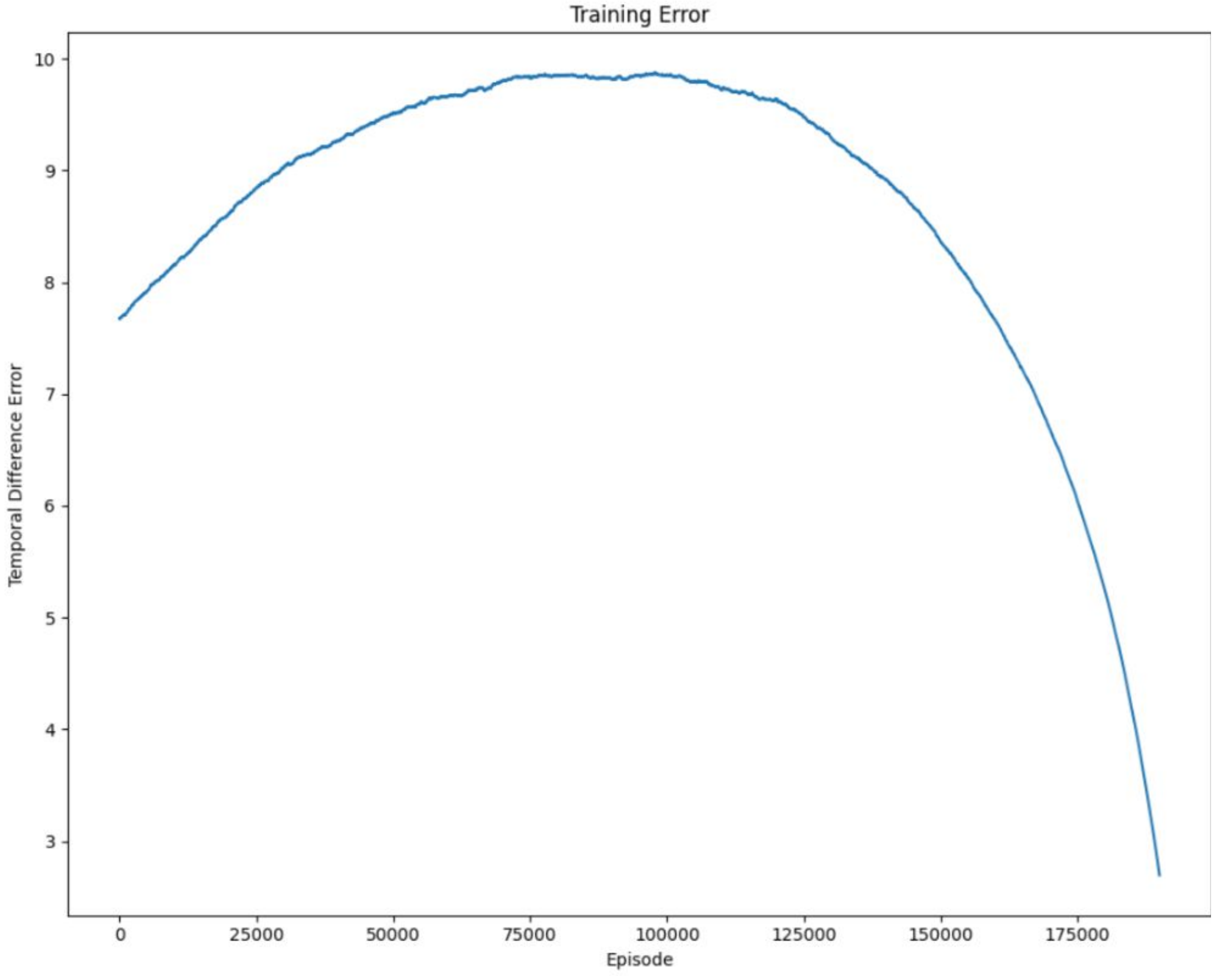
- Uses a Q-value table to store all state-value pairs at the pixel level.
- Updates iteratively using a Bellman Equation.
- Parameters
  - Discount Rate/Gamma: 0.95
  - Learning Rate: 0.01
  - Episilon Decay: 1.0 / # of episodes
  - Q-table is hash-based for lower memory usage
  - States preprocessed for lower memory usage
    - Greyscale
    - Resized
    - Normalized and Flattened
  - Episodes: 200,000

aws

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Results





©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Limitations and Issues

- Not scalable to large or continuous state spaces.
- Memory usage increases exponentially with state and action space sizes.
- Fails with high-dimensional inputs (e.g., raw pixel data).

aws
©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Deep Q-Learning

- Replaces the Q-table with a neural network that approximates the Q-values.
  - Neural Network consists of 3 2DConvolution layers.
- Uses experience replay to stabilize learning
  - Stores past experiences (s,a,r,s', done) in a replay buffer.
  - Trains the network by sampling mini-batches from the buffer.
  - Updates weights using the loss function
    - $1/N * N\Sigma i=1(y_i - Q(s_i,a_i;\Theta))^2$
-
- Parameters
  - Discount Rate/Gamma: 0.95
  - Learning Rate: 0.01
  - Episilon Decay: 1.0 / # of episodes
  - Q-table is hash-based for lower memory usage
  - States preprocessed for lower memory usage
    - Greyscale
    - Resized
    - Normalized and Flattened
  - Episodes: 50,000

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Results

```
Episode 0/50000, Total Reward: 15.0, Epsilon: 1.00
/usr/local/lib/python3.10/dist-packages/numpy/core/from
  return _methods._mean(a, axis=axis, dtype=dtype,
/usr/local/lib/python3.10/dist-packages/numpy/core/_met
  ret = ret.dtype.type(ret / rcount)
Episode 2500/50000, Total Reward: 10.0, Epsilon: 0.95
Episode 5000/50000, Total Reward: 9.0, Epsilon: 0.90
Episode 7500/50000, Total Reward: 8.0, Epsilon: 0.85
Episode 10000/50000, Total Reward: 9.0, Epsilon: 0.80
Episode 12500/50000, Total Reward: 8.0, Epsilon: 0.75
Episode 15000/50000, Total Reward: 11.0, Epsilon: 0.70
Episode 17500/50000, Total Reward: 23.0, Epsilon: 0.65
```

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Limitations and Issues

- Requires significant computational resources (e.g., GPUs).
- Sensitive to hyperparameter tuning.
- Slower convergence compared to tabular Q-learning in simple environments.

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Improvements

- Better reward shaping for faster convergence.
- More computational power and time.
- Different optimization techniques.
  - Mixed precision training
  - Parallel Environments

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Conclusions

- Both reinforcement learning models showed learning ability and improvement.
- Deep Q-Learning would perform better on the high-dimensional state space
- Lacking time and computational resources for full convergence.

©2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.