Gaming Using Reinforcement Learning

Capstone Project Review I

K Jivitesh Narayan
18BCE1290
Presenter and Project Lead/Author

Dr. NITHYA DARISINI P S
Associate Professor
Project guide

PPT approval - K Jivitesh Narayan 18BCE1290 - UPDATED 2

Inbox ×

K JIVITESH NARAYAN 18BCE1290

Gaming using Reinforcement Learning: Review 1 v3

Nithya Darisini P S

to me

Approved for Review-1 Presentation

Thanks and Regards,

Nithya Darisini P.S.

VIT – Recognised as Institution of Eminence (IoE) by Government of India

Agenda

01

About Project

A Description of the Project

02

Project Plan

Timeline, Schedule, etc

03

Status Report

Current Progress, Work done, etc

04

Next Review Goals

Projected Progress, Project Work Done, etc

01

About Project

A Description of the Project

Project Motivation

- Learning about RL and its applications in Gaming
- Building an AI to be better than humans in a particular task
- Contributing to the RL community

About the Project

Problem Statement

- There is no AI for the Super Auto Pets
- This is a potential area for exploration of application of RL in Gaming

Research Objectives

- Applying RL Algorithms to a new domain (super auto pets)
- Evaluate the performance and scalability of different RL algorithms to the domain

Modules of Project

Computer Vision

Captures the state information from screenshot of the game

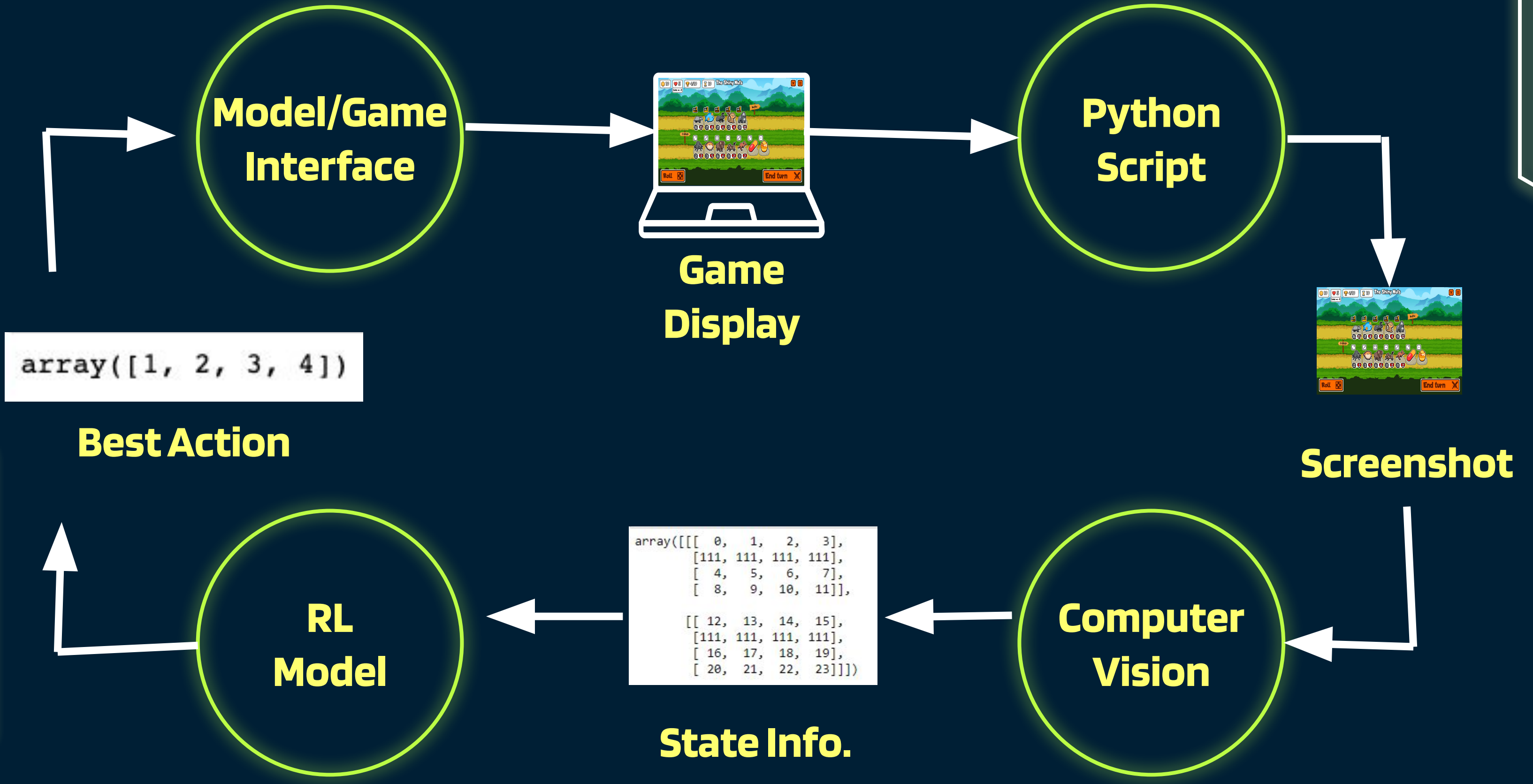
Model/Game Interface

Implements the action given by the model

RL Model

Takes State Info. to give the next best action

System Architecture



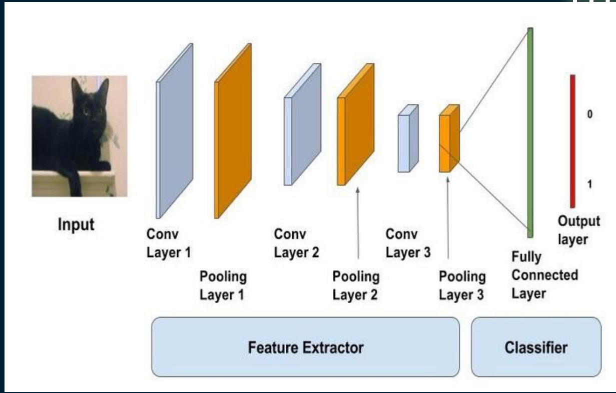
Module Explanation: Computer Vision

Functionality

- Extraction of State Information From ScreenShot

Concepts/Technologies Used

- CNN Multi-Class Classifier - Overfitted with single image for each class
- Image Manipulation Operations - Cropping of Region of Interest



Module Explanation: Model/Game Interface

Functionality

- Realising the action given by the model into the game

Concepts/Technologies Used

- Using Python library pyautogui to control the mouse
- Abstracting the module using classes in python

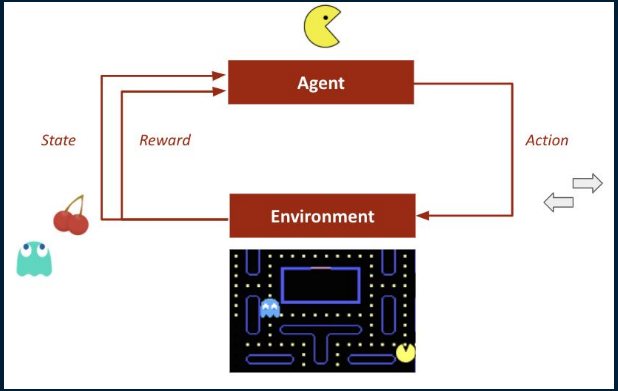
Module Explanation: RL Model

Functionality

- Translates State Information to next best action

Concepts/Technologies Used

- Machine Learning
- Machine Learning frameworks - Tensorflow
- GPU for training
- Reinforcement Learning Concepts



Chosen Domain

Super Auto Pets

- you build a team from a cast of animals who will fight for you
- Animals in store appears according to the probability set by dev.
- They each have unique abilities
- You play against other teams build by players connected through the internet

02

Project Plan

Timeline, Schedule, etc

Project Timeline

01

Phase 1

Proof of Concept

02

Phase 2

CV and Interfacing

03

Phase 3

Environment Construction

04

Phase 4

Algorithms Performance and scalability Evaluation

05

Phase 5

Report Writing

Next Review Schedule

<table>
<tr><th></th><th>Mon</th><th>Tue</th><th>Wed</th><th>Thu</th><th>Fri</th><th>Sat</th></tr>
<tr><td>Week 1</td><td colspan="4">Task 1</td><td></td><td>Task 4</td></tr>
<tr><td>Week 2</td><td colspan="5">Task 2</td><td></td></tr>
<tr><td>Week 3</td><td>Task 4</td><td></td><td colspan="4">Task 3</td></tr>
<tr><td>Week 4</td><td colspan="2">Task 3</td><td></td><td colspan="3">Task 4</td></tr>
</table>

- CV and Interfacing
- Env. Const.
- Algo. Implementation and Eval.
- Report Writing

03

Status Report

Current Progress, Work done, etc

Literature Survey: Papers

Summaries:

https://docs.google.com/document/d/1l9ANdrK7Fzzm9KnKKIsSuA-kXC9LZmzS9VTIPrlGJ0w/edit?usp=sharing

Deep Q Learning

- Deep Recurrent Q-Learning for Partially Observable MDPs
- Deep Reinforcement Learning with Double Q-learning
- Prioritized Experience Replay
- Dueling Network Architectures for Deep Reinforcement Learning
- Rainbow: Combining Improvements in Deep Reinforcement Learning
- Playing Atari with Deep Reinforcement Learning

Deterministic Policy Gradients

- Continuous Control With Deep Reinforcement Learning
- Addressing Function Approximation Error in Actor-Critic Methods
- Deterministic Policy Gradient Algorithms

Literature Survey: Papers

Distributional RL

- A Distributional Perspective on Reinforcement Learning
- Distributional Reinforcement Learning with Quantile Regression
- Implicit Quantile Networks for Distributional Reinforcement Learning
- Dopamine: A Research Framework For Deep Reinforcement Learning

Policy Gradients

- Trust Region Policy Optimization
- High-Dimensional Continuous Control Using Generalized Advantage Estimation
- Asynchronous Methods for Deep Reinforcement Learning
- Sample Efficient Actor-Critic With Experience Replay
- Emergence Of Locomotion Behaviours In Rich Environments
- Proximal Policy Optimization Algorithms
- Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation
- Soft Actor-Critic

Objectives Progress

50%

Interfacing

Need to Configure to Super Auto Pets

30%

Computer Vision

Need to Configure to Super Auto Pets

0%

RL Algorithm

Yet to start working on it

0%

Environment Construction

Yet to start working on it

RAID Summary

Risks

- Similar work being published
- Faulty Environment

Assumptions

- Existing Algorithm work for this project
- Computer can handle both the project and the game at the same time

Issues

- Recent update to the game
- Portability of the project

Dependencies

- Sapai – an super auto pets environment construct

Proof Of Concept

The Project done for the game of Tic Tac Toe is a success.

04

Next Review Goals

Projected Progress, Project Work Done, etc

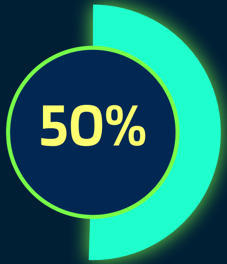
Projected Objectives Progress

Interfacing

Will be Completed

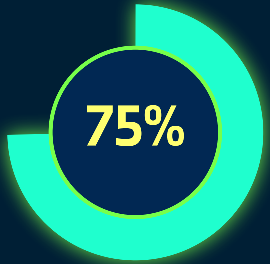
Computer Vision

Will be Completed



RL Algorithm

Yet to Implement some algorithms



Environment Construction

Yet to complete Testing

References

- Hausknecht, Matthew, and Peter Stone. "Deep Recurrent Q-Learning for Partially Observable MDPs." ArXiv:1507.06527 [Cs], Jan. 2017. arXiv.org, http://arxiv.org/abs/1507.06527.
- van Hasselt, Hado, et al. "Deep Reinforcement Learning with Double Q-Learning." ArXiv:1509.06461 [Cs], Dec. 2015. arXiv.org, http://arxiv.org/abs/1509.06461.
- Schaul, Tom, et al. "Prioritized Experience Replay." ArXiv:1511.05952 [Cs], Feb. 2016. arXiv.org, http://arxiv.org/abs/1511.05952.
- Wang, Ziyu, et al. "Dueling Network Architectures for Deep Reinforcement Learning." ArXiv:1511.06581 [Cs], Apr. 2016. arXiv.org, http://arxiv.org/abs/1511.06581.
- Hessel, Matteo, et al. "Rainbow: Combining Improvements in Deep Reinforcement Learning." ArXiv:1710.02298 [Cs], Oct. 2017. arXiv.org, http://arxiv.org/abs/1710.02298.
- Mnih, Volodymyr, et al. "Playing Atari with Deep Reinforcement Learning." ArXiv:1312.5602 [Cs], Dec. 2013. arXiv.org, http://arxiv.org/abs/1312.5602.
- Lillicrap, Timothy P., et al. "Continuous Control with Deep Reinforcement Learning." ArXiv:1509.02971 [Cs, Stat], July 2019. arXiv.org, http://arxiv.org/abs/1509.02971.

References

- Fujimoto, Scott, et al. “Addressing Function Approximation Error in Actor-Critic Methods.” ArXiv:1802.09477 [Cs, Stat], Oct. 2018. arXiv.org, http://arxiv.org/abs/1802.09477.
- Silver, David, et al. “Deterministic Policy Gradient Algorithms.” Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32, JMLR.org, 2014, p. I-387-I-395.
- Bellemare, Marc G., et al. “A Distributional Perspective on Reinforcement Learning.” ArXiv:1707.06887 [Cs, Stat], July 2017. arXiv.org, http://arxiv.org/abs/1707.06887.
- Dabney, Will, et al. “Distributional Reinforcement Learning with Quantile Regression.” ArXiv:1710.10044 [Cs, Stat], Oct. 2017. arXiv.org, http://arxiv.org/abs/1710.10044.
- Dabney, Will, et al. “Implicit Quantile Networks for Distributional Reinforcement Learning.” ArXiv:1806.06923 [Cs, Stat], June 2018. arXiv.org, http://arxiv.org/abs/1806.06923.
- Castro, Pablo Samuel, et al. “Dopamine: A Research Framework for Deep Reinforcement Learning.” ArXiv:1812.06110 [Cs], Dec. 2018. arXiv.org, http://arxiv.org/abs/1812.06110.
- Schulman, John, et al. “Trust Region Policy Optimization.” ArXiv:1502.05477 [Cs], Apr. 2017. arXiv.org, http://arxiv.org/abs/1502.05477.

References

- Schulman, John, et al. “High-Dimensional Continuous Control Using Generalized Advantage Estimation.” ArXiv:1506.02438 [Cs], Oct. 2018. arXiv.org, http://arxiv.org/abs/1506.02438.
- Mnih, Volodymyr, et al. “Asynchronous Methods for Deep Reinforcement Learning.” ArXiv:1602.01783 [Cs], June 2016. arXiv.org, http://arxiv.org/abs/1602.01783.
- Wang, Ziyu, et al. “Sample Efficient Actor-Critic with Experience Replay.” ArXiv:1611.01224 [Cs], July 2017. arXiv.org, http://arxiv.org/abs/1611.01224.
- Heess, Nicolas, et al. “Emergence of Locomotion Behaviours in Rich Environments.” ArXiv:1707.02286 [Cs], July 2017. arXiv.org, http://arxiv.org/abs/1707.02286.
- Schulman, John, et al. “Proximal Policy Optimization Algorithms.” ArXiv:1707.06347 [Cs], Aug. 2017. arXiv.org, http://arxiv.org/abs/1707.06347.
- Wu, Yuhuai, et al. “Scalable Trust-Region Method for Deep Reinforcement Learning Using Kronecker-Factored Approximation.” ArXiv:1708.05144 [Cs], Aug. 2017. arXiv.org, http://arxiv.org/abs/1708.05144.
- Haarnoja, Tuomas, et al. “Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor.” ArXiv:1801.01290 [Cs, Stat], Aug. 2018. arXiv.org, http://arxiv.org/abs/1801.01290.

References

- https://towardsdatascience.com/reinforcement-learning-implement-tictactoe-189582bea542
- https://github.com/daliborstakic/tictactoe-pygame
- https://www.tensorflow.org/tutorials/images/classification
- https://www.geeksforgeeks.org/mouse-keyboard-automation-using-python/
- https://www.geeksforgeeks.org/pyhton-pil-imagegrab-grab-method/
- https://teamwood.itch.io/super-auto-pets
- https://github.com/manny405/sapai
- https://github.com/bencoveney/super-auto-pets-db/
- https://spinningup.openai.com/en/latest/

Thank You