AI Village

A Summer Project

Tariq Chatur

2019

# Title Page

Contents

[Title Page 1](#_Toc17411591)

[Abstract 2](#_Toc17411592)

[Introduction 3](#_Toc17411593)

[Background 4](#_Toc17411594)

[Project Plan 5](#_Toc17411595)

[Implementation 6](#_Toc17411596)

[Results 7](#_Toc17411597)

[Potential for Extension 8](#_Toc17411598)

[Refrences 9](#_Toc17411599)

[Appendix 10](#_Toc17411600)

# 

# Abstract

# Introduction

In this paper, I will explore the project that was titled AI Village. This project was used as the main focal point in my Directed Studies course at the University of Victoria. The course was comprised of a project proposal, a presentation, a final report, and meetings every two weeks centered around the progress in my learning and of the project itself. As the course was centered around the project of AI Village, the topics that were learned were as follows:

* Reinforcement Learning (Q-Learning, Reward Systems, Backwards Passes, Neural Networks, and Convolutional Neural Networks)
* Python (PyGame and PyTorch)
* Psychology (Maslow’s Hierarchy of Needs and Reinforcement Learning)
* Sociology (Social Structures, Group Behaviour)

A combination of these topics is what led to the ability to create the project called AI Village. AI Village was planned to be a method to demonstrate the capabilities of Reinforcement Learning in regards to working in a continuous dynamic environment. However, it relied on two assumptions, the first being that the AI were able to simulate living creatures, and that behavior was created through experience.

AI Village is an enclosed 9x9 grid where AI agents can learn to survive. Their needs are limited to socializing and resource tiles around the map. The only actions the AI are able to do is move either up, down, left, or right on each turn. The world works in a turn based system that allows each agent to choose a direction, interact with the environment, gain a reward from the movement, and learn from the experience. Currently, the AI are tested using a group of 5 – 8 agents at a time.

Using this set up of allowing agents to interact with their world which has no distinctive goals, the hope of the project was to replicate the behaviours that are observed in the real world.

# Background

This project was firstly inspired by works of fiction that often depict androids, AI, or other man made machines that are able to replicate human behaviour. Examples of popular fiction in which these themes arise are: Do Androids Dream of Electric Sheep, The Matrix, Sword Art Online, IRobot, Avengers Age of Ultron, and many others. In combination with these influences, my first interaction with deep learning allowed the understanding of building such machines to materialize.

In deep learning, a collection of neurons is used to create what can be attributed to a “brain” of a machine. This brain can be trained to do tasks such as identify numbers, translate languages, save money, drive cars, or predict the future. In this case, rather than train an AI to achieve a task, the goal was to allow the AI to act and see if the behaviour would be similar to other living creatures if given similar circumstances.

The University of Alberta currently has a study which uses neural networks to showcase the behaviour of AI rabbits as they survive in an environment with wolves, and limited resources. This project was another influence in creating an enclosed system that rather than showing the survival behaviour, would show the social behaviour of AI. The interest in this case was not the most efficient way to survive but rather, what an AI would do when surviving could be achieved easily.

Google Deepmind had created an AI structure using Deep Q-Learning that played a variety of games. The AI was generally able to learn how to play games to a better extend than humans when the game would provide a constant feed of rewards to the agent. However in cases where rewards were sparse, the AI would perform very poorly.

These were the largest influences in creating this project and from them the idea, the methodology, and environment were all invented. Looking into research into topics such as Intrinsic Reinforcement Learning, Reinforcement Learning, and AI, projects attempting to replicate this result were not found. This realm of mystery was the final push to start the project.

# Project Plan

The plan for the project fell under three main topics understand what it means to be a social creature, create a world in which the creature can learn to be social, create a brain so the agent can learn its behaviour. Every step was shaped around a cycle involving these three steps.

During the month of May, most of the time dedicated to this project was spent on researching the behaviour of people, cows, and other creatures that spent most of their life with others. It was through this scenario I understood that each agent would need their own memory and brain or else the system to replicate ants, bees, or other colony creatures. As the goal was to make agents act of their own self-interest, the idea of creating a colony was thrown out. In addition, during this period, the potential environment was thought up of and finalized.

In regards to the construction of the agents, using Maslow’s hierarchy of needs was the basis in constructing the presence and needs that the agent would have. Research into CNNs (Convolutional Neural Networks) was done to determine how the agent would see the world it interacted with. In most cases, people would take the RGB values for their agent, however this method would save the system heaps of memory and allow for more agents to appear at once on the limited CPU available.

June was scheduled to be the implementation of the environment, the set up of the agents, and the week of my birthday. In this time, the hope to have the functional game that could be controlled by a user was looking very likely to come to fruition. This is because the access to libraries in python allowed for construction of the environment to be much easier than previously imagined.

July was the month of construction of the AI, the loss function, and testing. I knew prior that the functions of the other reinforcement learning would not work and therefore knew that a significant amount of time would need to be put aside to create a new loss function. Also the structure of the CNN would need to be justified depending on the input of the CNN.

August was the month of the reporting, this is when the final report and presentation was due. In this time, no more research would be done, and additional resources would be created including slides, schedules, and reports.

# Implementation

The implementation was split into three sections. The creation of the environment, agents, and AI. These three sections were all intertwined and every few days needed to be cleaned to make sure the code was readable.

The environment was constructed to be a 9x9 grid which would contain stationary resources that would regenerate over time. This would create competition for resources but also allow for the agents to identify a location in regards to each basic need they had. The Resources were created to fill in entire squares of the grid with a colour corresponding to the need it was addressing. Agents would need to move on top of the resource to claim its benefits. The values regarding maximum value of resource, potency, and location were all randomized when the environment was created. Currently the environment was made to be simple to allow for testing on the agent and neural network though a turn based system.

The agent was created to be a small square in the middle of the map. The agent was initialized to have a randomized health, and max need value. The randomization was used to create variation in the population. With each movement, the need values would drop. If the need values were too low, the health would also depreciate. Agents could see the world through a mental map, this 9x9x(Number of needs + 1) gave the agent, locational data on all the needs, agents, and selves in the system. This would be the input to the CNN. The agent’s turn was organized in the steps of:

1. Choosing a direction to move
2. Updating the map
3. Updating the agent’s values
4. Updating the agent’s CNN

The AI was created out of a CNN with a shape that can be found in the appendix. This CNN was chosen because it allowed for the identification of the map through the needs that the agent possessed. The use of 0 padding was to understand the corners of the maps. The use of a small CNN was because of the limitations of the computer on which this project was running. The final output of the CNN was equal to the action space (left, right, up, down) of the agent. The loss function that was created was equal to the total loss plus the reward of the system divided by the total amount of health the agent possessed in the time it was learning. Using Adam as the optimizer, the system continuously chose the best possible action and the values of the mental map changed depending on the requirement of the need.

Overall, the system was constructed quite roughly, however in the end the system had a stable environment, agents that could easily interact, and a AI that would control its assigned AI while how to best live in the world created. There were still many problem, but it was a start.

# Results

The results of the system were AI that would showcase very basic social behaviour when given a reward when meeting the needs of a social animal. These behaviours were:

1. Agents spend most of their time near resources
2. Agents opt to be closest to agent with the highest social value therefore those with higher values tend to dictate what activities the group will perform
3. It takes roughly 4000 iterations from every Agent to learn how to survive

The loss function of the system was inefficient, but functional, with a max cap.

The system was able to generate a world, agents, and AI that could live within and adapt.

The agents were able to rise up to level 3 of Maslow’s Hierarchy of needs.

The continuous learning system was developed so the agents would learn from a single running of the system without having to reset the system to do backwards passes.

Research into Q-Learning showed it would not work in the long term with regards to the agents as the storing of values for longer term thinking would create a memory leak when the agent had lived for a certain amount of time.

CNN was too large and caused issues when the number of agents rose past 8

There was no way to determine a “best model” and the system would need to start from scratch at every trial.

Research into psychology and sociology saw that an expert in those fields would be needed for further recreation of the project

Agents would suddenly chose to die and be respawned when their overall need values got too low

When resources moved, agents had a very difficult time unlearning the locational data they had previously established and would return to the places where the resource was originally located even thousands of iterations later into the training.

# Potential for Extension

This project’s goal was to see whether or not social functions could be replicated using AI. The answer to this was yes. However, this was done in a simplified environment, with overly complex agents, and a loss function that was inefficient. While the overall goal of the project was completed, the method in which it was obtained could be made better. Therefore to move forward, rather than just question if the agents can be social, the next thing to do is test how they function in certain environments. Below are a list of ideas that can be implemented to further the limits of Reinforcement Learning in the environment.

1. AI can be simpler – Currently the data that is needed to make a single AI is too large as the CNN is too complex
2. Limited Vision – Rather than updating the entire mental map at every turn, only updating the closest 5x5 tiles would allow the agent to make false assumptions about their environment which would generate the need to explore, or communicate with other AI
3. More Complex Tiles – The addition of tiles that don’t directly influence the AI’s need values but assist in their ability to survive such as a tile that updates their entire map, or tiles that allow them to move twice on a turn for a few turns can show how their adapt their long term thinking
4. Removal of hardcoded rewards – Making rewards an equation based on the current state of the agent would remove the number of hyperparameters and allow AI Village to be more autonomous
5. Expand the action space – Including actions such as shout, sleep, or interact would make the method of interacting with the world more complex and while lengthening the training time may showcase whether certain actions are worth taking in an environment
6. Breeding – Allowing the AI to breed when they have sated their needs for a certain amount of time will allow for us to see the dynamics in an evolving population
7. Agent to Agent communication – Currently the agents don’t share information, however, they tend to be close to one another to sate their social needs. By sharing information, we can test the concept of teamwork on a higher scale compared to movement patterns which is currently being done

The project itself will always be adapting and growing as the potential of the environment is still very small compared to our real world, and the project as it stands wouldn’t be able to encapsulate what living in our world is like as they have a limited action space. However, it can be complexified enough so with the technology of VR and Graphics, we could enter a world with AI and if their behaviours have been trained long enough, it would be difficult to tell the difference between AI and User. This is the final goal of the system; however, with the initial step of seeing that AI can learn behaviour in an environment that is similar to social animals on earth, the possibility of creating a virtual world where these AI can live doesn’t seem very far away.

In terms of why this is important for the general world, it would allow us to test influences on AI and speed up testing dramatically. It would largely enhance the realism of CGI by inserting AI which act normally into different forms of media. It would provide a new genre of media which would exist on its own. It would allow us to reflect on our own behaviour be questioning why certain characteristics appear in our society when the either do or do not appear in AI Village.

# Conclusion

AI Village is technically already complete on a simple scale. Just as how a bow simply needs a flexible stick and a string, or how a piece of corn is technically a meal. However, with time and care the program will evolve into something that is more tangible, and charismatic.

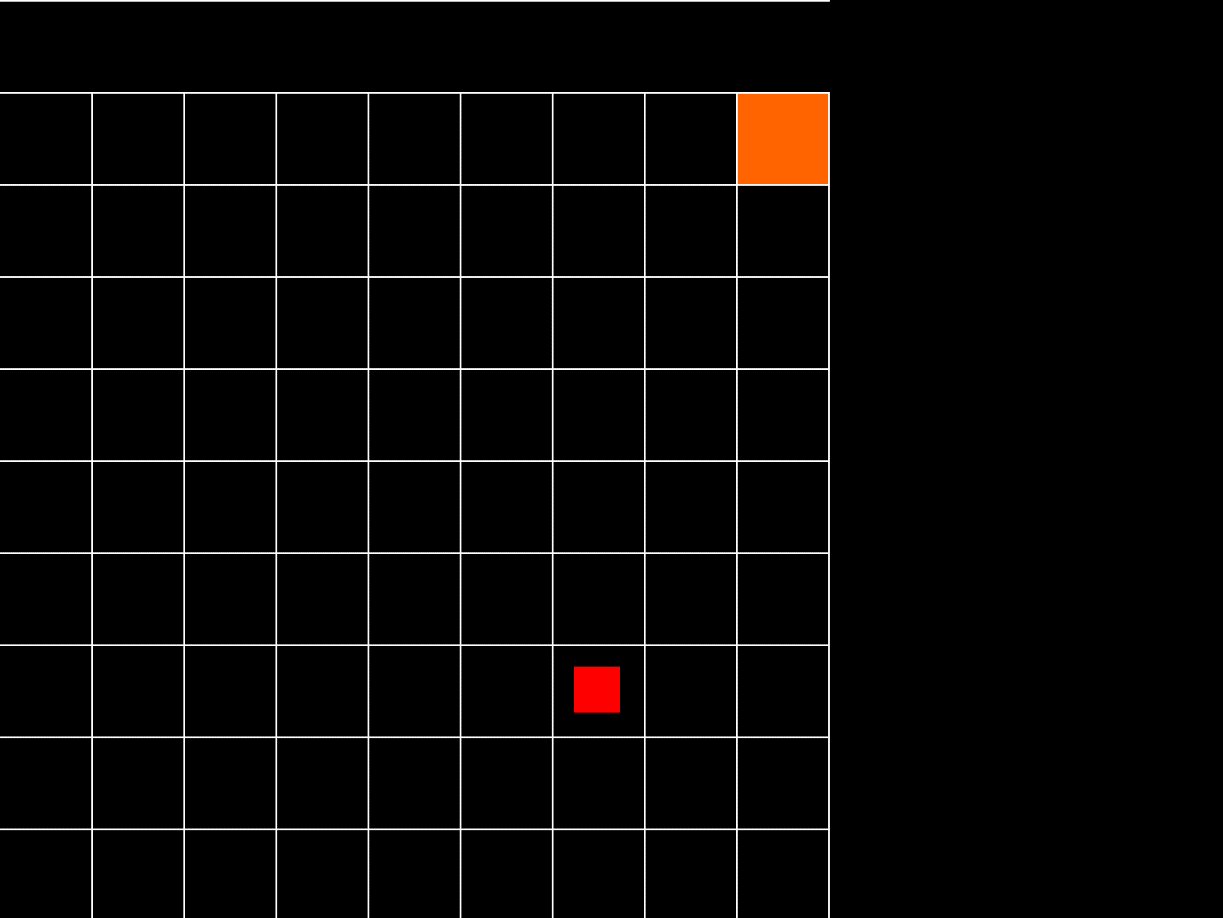
Through the process of making AI Village, the topic of reinforcement learning has now become much more clear. As the many premade loss functions that are used around the world were not functional in my system, I had to truly learn what reinforcement learning entailed to be able to make my own functions that could allow the AI to learn. This was the goal of the direct studies and I personally feel like it was achieved.

In addition, the interest in the topic for myself has skyrocketed and I find that I am working on this project even on time that I should be spending doing things like writing reports or slides. From this point, I hope to take this project and continue to develop it until I reach the final product that I believe is possible. Whether this be through self-effort, or with the assistance of others in Graduate School.

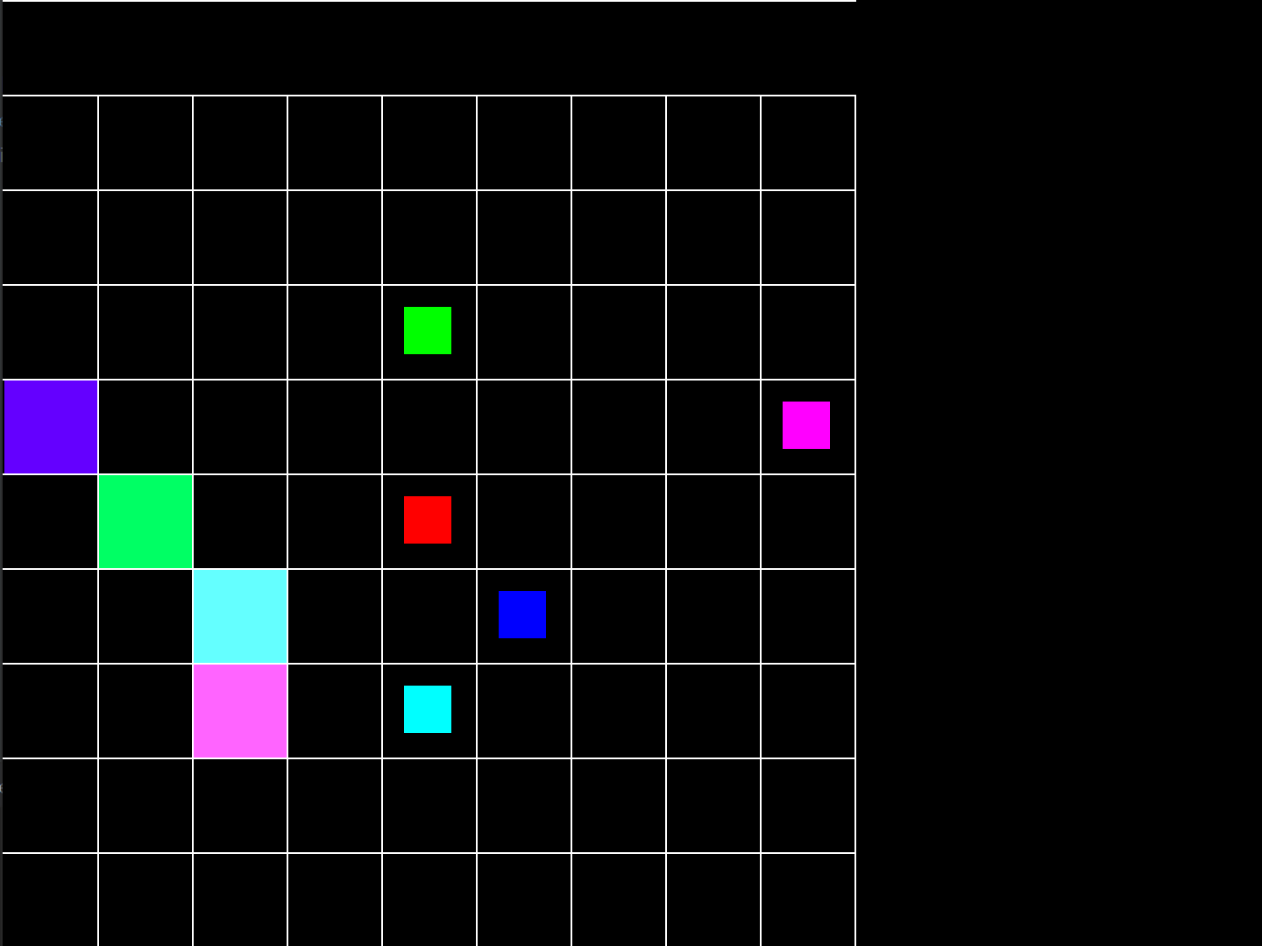
To conclude, this project was like a dream for me. The amount I learned through the semester, and the product of that newfound knowledge has only increase my interest in the subject and pushes me further into wanting to do research. While I am not sure whether following that life path will be the next step for me, I will be sure to always keep those associated with this project updated on its status as I continue to develop it.

# Appendix

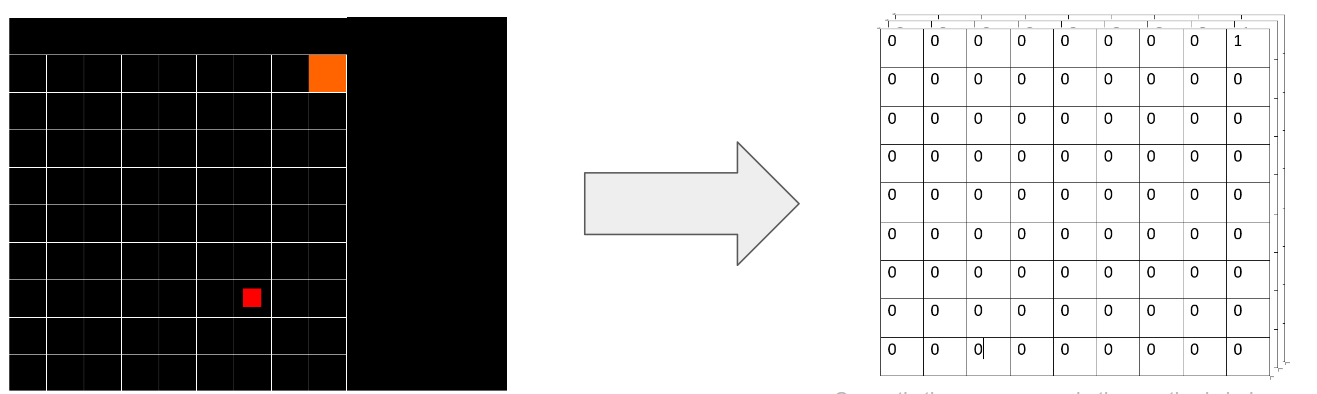
## Simple Version of AI Village



## More Complex Version of AI VIllage



## Visualization of Mental Maps of Agents



## Framework of the CNN used currently in Agents

