Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

Reinforcement Learning Agent to play Tetris

Mechatronics Project 478
Final Report

Mr Divan van der Bank
Student No 23603526

Supervisor: Prof. Herman Kamper

17 March 2023

saam vorentoe · masiye phambili · forward together

Department of Mechanical and Mechatronic Engineering
Departement Meganiese en Megatroniese Ingenieurswese
Privaat Sak X1, Private Bag X1, Matieland, 7602
Tel: +27 21 808 4204 | www.eng.sun.ac.za

Plagiarism declaration

I have read and understand the Stellenbosch University Policy on Plagiarism and the definitions of plagiarism and self-plagiarism contained in the Policy [Plagiarism: The use of the ideas or material of others without acknowledgement, or the re-use of one's own previously evaluated or published material without acknowledgement or indication thereof (self-plagiarism or text-recycling)].

I also understand that direct translations are plagiarism, unless accompanied by an appropriate acknowledgement of the source. I also know that verbatim copy that has not been explicitly indicated as such, is plagiarism.

I know that plagiarism is a punishable offence and may be referred to the University's Central Disciplinary Committee (CDC) who has the authority to expel me for such an offence.

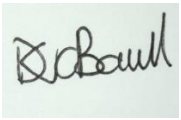
I know that plagiarism is harmful for the academic environment and that it has a negative impact on any profession.

Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully (acknowledged); further, all verbatim copies have been expressly indicated as such (e.g. through quotation marks) and the sources are cited fully.

I declare that, except where a source has been cited, the work contained in this assignment is my own work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

I declare that have not allowed, and will not allow, anyone to use my work (in paper, graphics, electronic, verbal or any other format) with the intention of passing it off as his/her own work.

I know that a mark of zero may be awarded to assignments with plagiarism and also that no opportunity be given to submit an improved assignment.

Signature: 

Name: Divan van der Bank Student no: 23603526

Date: 17 March 2023

Executive Summary

Title of Project
The development of a Reinforcement Learning Agent to learn and play Tetris
Objectives
Developing a stable environment for the testing of different Reinforcement Learning algorithms whilst training and playing Tetris
What is current practice and what are its limitations?
Current Reinforcement Learning testing platforms are tested in Retro game environments such as Tetris and Pacman and is limited by the processing power of the computer and the training time available for the algorithms
What is new in this project?
After the RL algorithms have been designed and are operational, the best reinforcement learning algorithm will be identified for the implementation in the arcade game, Tetris.
If the project is successful, how will it make a difference?
The Reinforcement Learning models and the testing environment will be used to investigate the feasibility of using Machine Learning as a tool in teaching humans to improve at skills like reading and writing
What are the risks to the project being a success? Why is it expected to be successful?
The processing power of the computer on which the Reinforcement Learning agent will train can limit the training efficiency and the time to train a model can be extensive. This is mitigated through using powerful faculty computers and planning for the training times of the agents.
What contributions have/will other students made/make?
A previous student has done research on using Reinforcement Learning to play a simplified version of Tetris
Which aspects of the project will carry on after completion and why?
The Tetris environment and the selected RL models will be carried over into a possible Master's investigating the feasibility of using the RL model to teach a human to play Tetris
What arrangements have been/will be made to expedite continuation?
The chosen configuration will be defined and documented in sufficient detail that the production demands and marketing potential can be determined.

ECSA Exit Level Outcomes

ECSA Outcomes Assessed in this Module	
ECSA Outcome	**Addressed in Sections:**
ELO 1. Problem solving: Demonstrate competence to identify, assess, formulate and solve convergent and divergent engineering problems creatively and innovatively	
ELO 2. Application of scientific and engineering knowledge: Demonstrate competence to apply knowledge of mathematics, basic science and engineering sciences from first principles to solve engineering problems.	
ELO 3. Engineering Design: Demonstrate competence to perform creative, procedural and non-procedural design and synthesis of components, systems, engineering works, products or processes	
ELO 5. Engineering methods, skills and tools, including Information Technology: Demonstrate competence to use appropriate engineering methods, skills and tools, including those based on information technology.	
ELO 6. Professional and technical communication: Demonstrate competence to communicate effectively, both orally and in writing, with engineering audiences and the community at large	
ELO 8. Individual, Team and Multidisciplinary Working: Demonstrate competence to work effectively as an individual, in teams and in multi-disciplinary environments	
ELO 9. Independent Learning Ability: Demonstrate competence to engage in independent learning through well-developed learning skills.	

Table of contents

Page

List of figures

Page

List of tables

Page

1 Introduction

1.1 Background

Machine learning is found in a wide range of applications and industries, including finance, healthcare, transportation, retail, and entertainment. The field has an enormous impact on the world and as more research and development is done, the capabilities of machine learning will continue to grow and evolve. However, for a Machine Learning model to be developed there must be a way of testing the algorithm or agent in a safe and controlled environment. A popular and safe way for testing new models/agents is by applying said models/agents to retro games such as Tetris or Pacman. This approach allows the developer to evaluate and improve the performance of the models/agents without the risk of financial or safety issues.

Reinforcement Learning is an efficient and effective Machine Learning algorithm that is used in a wide range of applications. The algorithms learn by placing an agent in an environment and allowing this agent to make choices. The agent is then trained based on the idea of rewarding desirable decisions and punishing undesired ones.

The project proposed in this report is to develop different types of reinforcement learning algorithms to learn and play the popular arcade game, Tetris. The comparison and evaluation of the different RL algorithms will be done by training the algorithms in the simulated game environment. The RL agents will be evaluated based on the training time and computational power, as well as the efficiency and the complexity of the methods. Once the testing is done, the best RL methods will be identified for playing Tetris.

The project's objectives, motivation, and planning will be explained in this document and will outline the steps that are planned in this study, as well as the expected costs and time scales for the study.

1.2 Objectives

This project aims to develop a reinforcement learning agent to learn and play the arcade game, Tetris. The objectives therefore are:

1.2.1. Develop a working Tetris environment which will serve as the testing base for the RL agents.

1.2.2. Identify and research types of Reinforcement Learning methods which will suit the problem at hand, learning and playing Tetris.

1.2.3. Develop and test the selected RL methods in the Tetris environment and compare the results to find the best RL method for playing Tetris.

1.3 Motivation

Tetris is an easy arcade game to learn. The rules and objectives for the game is simple and easily understandable for a first-time player. It is thus straight forward to explain to someone how to play Tetris. However, it is difficult to explain to someone how to quickly improve at the game. Apart from gaining experience through hours of play, there is no simple way quickly increase a player's ability.

This is where the Reinforcement Learning agent is useful. The agent will learn to play Tetris through a series of training algorithms and can serve as a model which the developer can use to analyse how the agent improves at Tetris. This can then be applied and be used to teach humans how to improve faster and more efficiently.

Even though it is not the scope of this project, this method of using the Reinforcement Learning agent can be applied to real-life situations, such as teaching a child to improve at reading or writing.

2 Planned Activities

The project can be divided up into smaller activities that will be completed for the execution of the project. The anticipated cost and time schedules are given in Appendix A.

2.1 Literature review

Research the overview of Reinforcement Learning and Tetris. Gain insight in the field of Machine Learning and scope which programming languages and packages will work best for this project.

2.2 Identify Applications

Investigate and identify probable applications for the developed Reinforcement Learning Agent/Models and whether the applications are suitable enough for further investigation.

2.3 Research and setup for Tetris development

Gain experience in playing the Tetris game and understand the rules and the basic operation and states of the game. Identify a suitable programming language for the game and the RL agent development and install the mentioned packages and programs for the development of the game.

2.4 Develop the Tetris Environment

Program and execute the Tetris game and create an interactive environment to supply a Reinforcement Learning model with the sufficient information to be able to train and interact with the game environment. The game must be in such a working condition so that a human can play the game as if it was the original game.

2.5 Identify suitable Reinforcement Learning Methods

Determine the characteristics and requirements required for a RL method to be considered for development. Using these characteristics and requirements, identify suitable RL methods for further investigation and development.

2.6 Design Reinforcement Learning Reward and Penalties

Describe the Tetris game in its fundamental operations and determine which elements can be considered to be rewards and penalties. These elements will then

expand into an equation form which will be used by the RL methods in determining correct moves in the game environment.

2.7 Program the Reinforcement Learning methods.

The reward and penalties equation can now be used, paired with neural network links to develop, and program the RL agents. These RL agents will be programmed as packages, to be used and easily swapped in the game environment to compare the results of the different methods.

2.8 Implement Reinforcement Learning methods and finetune the parameters

Once the RL agents have been programmed, they can be implemented into the game environment and trained to learn and play Tetris at an optimal rate. The hyperparameters of the RL agents can be tested and finetuned to deliver efficient models.

2.9 Extract and compare the results of the RL agents

Once the RL agents have been tested and finetuned to optimal functionality, the results of the different agents can be compared and investigated to find the best RL agent. Any trends in the results data can also be investigated if time permits, however this is not part of the main goal of the project.

2.10 Finalise Report

Document the results from the tests and include all the procedures followed during this project. Comment on the feasibility of using a Reinforcement Learning algorithm to learn and play Tetris and the potential of teaching a human to improve at the game.

3 Project Risk Assessment

The main risk is that the computational power needed to train the Reinforcement Learning models can exceed the limitation for the laptop that the program and models are being developed on. This will prevent the models from completing training and result in non-optimized results. A sub-risk for the completion of the project it the ongoing Load Shedding, where the power goes off daily for 2 hours at a time. This is a risk as the project is entirely based on a computer and simulations and programming requires high computational power which consumes more battery of a laptop.

These risks can be avoided by working in the Stellenbosch University Engineering Faculty on the Firga computers provided to the students. This will ensure that the program will have access to sufficient processing power and ensure that power will always be available during load shedding times. Furthermore, the progress of the code will also be saved on the Stellenbosch University drive and prevent loss of data.

The loss of data is another risk to be considered and can be prevented by uploading commits of the code to a service such as Git.

4 Conclusions

The student will design a testing environment in Python for the evaluation of different machine learning algorithms in the attempt to find the best algorithm for playing the popular game, Tetris. The student will follow the planned activities as set out in this report and the time-schedule from Appendix A.2.

The risk plan prevents any unforeseen issues that can arise in the future and helps the student to mitigate any problems with the plan of action for each of the risks. An end-of-life strategy and responsible use of resources plan is developed to ensure that the project comes to a successful and recourse friendly end.

The student has all the expertise and equipment required to successfully complete the project to his disposal. The project is expected to span across a time of 9 months, allowing for enough time to manage the risks and training time. The cost for the project is determined as R82500 for the student engineering time and faculty use.

5 References

Singh, V. (n.d.). *Data Science Central*. Retrieved from How Machine Learning Is Changing the World: https://www.datasciencecentral.com/how-machine-learning-is-changing-the-world/

Appendix A Planning Activities

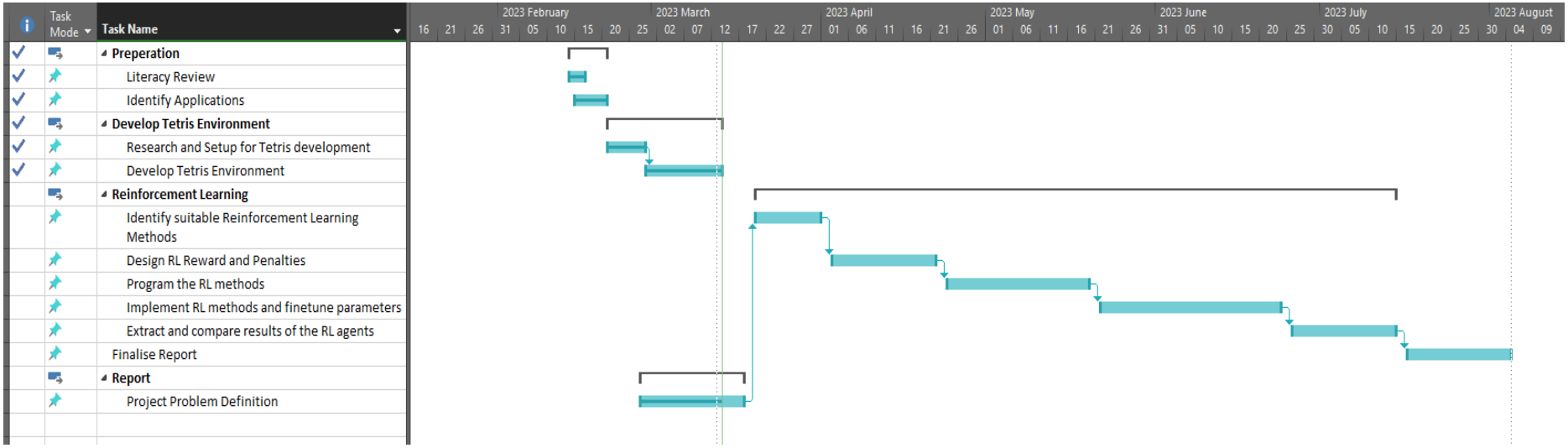
A.1 Estimated Cost

Table A.1: Estimated Cost per Activity

Activity	Engineering Time		Facility Use	TOTAL
	Hr	**R**	**R**	**R**
Literacy Review	10	4500		**4500**
Identify Applications	5	2250		**2250**
Research and Setup	25	11200		**11200**
Develop Tetris	25	11200	500	**11700**
Identify RL Methods	10	4500		**4500**
Design Rewards and Penalties	15	6750		**6750**
Program RL methods	25	11200	500	**11700**
Implement RL Methods	25	11200	500	**11700**
Extract Results	10	4500		**4500**
Finalise Report	20	9000		**9000**
TOTAL	**180**	**81000**	**1500**	**82500**

A.2 Gantt Chart and Timeline

Figure A.1: Gantt Chart for Planned Activities



A.3 A.3 Responsible Use of Resources

This project is a pure simulation study and thus is developed and evaluated fully on a computer environment. However, it is still important to address the minor recourses that are used in the design of the final product.

This project will require a computer with high processing power capabilities as the training of the RL algorithms can become computationally expensive and require the computer to run at full capacity for several hours. The computer will therefore require a reliable power connection and use the grid power as well as the generator power during load shedding. This means that costs for the power and the fuel required to run the generators must be considered.

In the event where the HPC must be used, which is a large, shared university resources, the student will ensure that the code that will be run on the machine is thoroughly tested on a local computer to prevent the misuse of the HPC and the strict time schedule.

The use of a stable network connection is also imperative to the success of the project as the development of the Tetris and the RL components require a stable internet connection to run on web-powered compilers.

A.4 End of Life Strategy Report

The project's main focus will be to evaluate and identify the best RL algorithms for learning and playing Tetris in a simulated environment. The project will thus come to an end when the final RL results are compared, and the best algorithms are selected.

However, if time permits and the student has the capabilities and knowledge to do so, the project could be extended to investigate further in the RL algorithms and improve the testing setup for use in a possible Masters.

The work done in this report forms part of a greater study in the feasibility of using reinforcement learning to teach humans how to improve at Tetris. Once completed, the code for the project will then be released on GitHub with a MIT License and serve as a base for the next student to open and install to extend the life of the project.