REPORT TITLE HERE

Connor Wallis (22506057) & Reilly Evans (23615971)

**Export This As A PDF Before Submitting**

**(and delete this)**

Introduction

One of the primary purposes of computers over their history has been to automate tasks that humans normally perform. Many different methods of creating agents for this purpose have been designed, from simple rule-based programs to complex machine learning algorithms. In this report we will construct two agents for the task of playing levels from Super Mario Bros: one rule-based agent implemented by hand, and one Proximal Policy Optimisation (PPO) agent trained using stable baselines. We will compare the performance of these two agents in a variety of areas, demonstrating the strengths and weaknesses of each approach. … [add more detail]

Analysis

The two agents we will be comparing are a rule-based agent and a PPO agent. These agents will interact with Super Mario Bros. using the methods in the gym-super-mario-bros package (Kauten, 2018). The rule-based agent is a simple Python script which reads in data corresponding to the game screen, alongside other information recorded by the package such as Mario’s coordinates and the number of lives he has., then chooses what combination of buttons to input based on that data. The rules used to determine what inputs are made were coded by hand over the course of this project’s development.

[pytorch code description]

[explain how RuleBased and PPO were constructed/trained]

The reward value used for training the PPO agent is the one that is included in the gym-super-mario-bros. package. This reward is calculated for each step as

where r is the reward value, x0 and x1 are Mario’s x position before and after the step, c0 and c1 are the value of the in-game timer before and after the step, and d is -15 if Mario died during the step or 0 otherwise. The reward cannot be outside of the range -15 to 15 (Kauten, 2018). This results in the agent prioritising going right as quickly as possible while avoiding death. [compare reward value from both agents]

-perform other relevant experiments

Analyze and contrast the performance of the chosen AI methods.

• Discuss their respective strengths, weaknesses, and suitability for playing Super Mario Bros.

* Rule based is easy to understand; parameters can be tweaked easily to make it jump shorter, etc.,
* Rule based cannot adapt on its own; new palletes like underground or new enemies like Lakitus require adding them to the enemy recognition program, coding new actions to avoid them, etc.
* Experiments could include testing speed of level completion (be it time, frames or actions) or the amount of memory used; many of the listed examples don’t really work with a rule based agent that doesn’t learn

Performance Metrics

While the gym-super-mario-bros package by default uses rightward progression as its primary metric of performance, this is not the only factor on which we can compare the agents. One potential alternate metric is to directly reward completing levels as quickly as possible. Under this metric we would provide large reward boosts at the end of levels based on their world/stage number sand the time remaining, while falling back on rewarding moving right quickly within individual levels. While this at first seems to be not very distinct from the existing reward function, it benefits in rewarding the use of subareas to skip parts of a level. Many levels in the game, including World 1-1, have pipes or vines that lead to bonus rooms then return Mario to a later part of the level, skipping a large portion of platforming. While these shortcuts make for faster level completion, they result in less rightward movement overall and thus are discouraged by the current reward function; rewarding quick stage completion allows these shortcuts to be utilised by the agent.

-points metric

You will notice that gym-super-mario-bros reward function assumes the objective of the game is to move as far right as possible. You are encouraged to come up with other performance and evaluation metrics for your agents. Novel and interesting metrics that you come up with will be rewarded.

* Points as a metric (collect coins/powerups gives lives and increases survivability; death penalty prevents infinite lives trick from causing problems
* Progression as a metric (get as far into the game as possible in terms of levels, or alternatively beat 8-4 from 1-1 as fast as possible; rewards finding the Warp Zones)
* Not Dying as a metric (get as far as possible without dying; rewards careful playing and prevents using the game’s checkpoints to cheese things
* Random Stages as a metric (use gym-super-mario-bros’s random stages function to try stages and measure number of stages out of say 10 completed)

Visualisation/Debugging

-rule based visualisation (terminal printing, Lauren’s code, freezing the game)

-PPO printing (tenserboard, other visualisation techniques)

Includewhatvisualizationtechniquesyouusedtogaininsightsintotheagent’sdecision- making process.

Include what debugging/profiling tools you utilised to optimize the algorithms and enhance performance.

* Besides what was already in Lauren’s code, printing to the terminal when decisions are made helps identify actions
* Freezing the game when a decision is made (via spamming the terminal with 250,000 messages) helps to identify exactly what constitutes a scenario where said decision is made
* B
* C

References

* Kauten, C. (2018). *Super Mario Bros for OpenAI Gym*. GitHub. Retrieved October 10, 2023, from <https://github.com/Kautenja/gym-super-mario-bros>
* Gee, L. (2023). *mario\_locate\_objects.py*. UWA Learning Management System. Retrieved October 5, 2023, from <https://lms.uwa.edu.au/bbcswebdav/pid-3405777-dt-content-rid-43562900_1/xid-43562900_1>
* [other references; use APA7 style; pytorch tutorial?]

Misc. notes [DELETE BEFORE SUBMITTING]:

Our agents are Hand Implemented Rule based agent and PPO from Stable Baselines

When we make changes to existing code (Laurens, pytorch tutorial) document it!

‘poetry run nes\_py --rom super-mario-bros.nes --mode human’ for human controlled Mario