Description: Beskrivning: hd_vertikal_farg WORD.wmf

Thesis report

Master thesis in Microdata Analysis

Business Intelligence Program

School for Technology and Business Studies

Dalarna University

Description: StreckTiteln

All for one one for All

Author: Bamshad Shirmohammadi

Supervisor: Professor Arend Hintze

Examiner:

Dalarna University

791 88 Falun

Sweden

Tel 023-77 80 00

Subject:

Course code: MI4002

Points: 15/30hp

Date of the examination: 20XX-XX-XX

**Abstract**

**Contents**

[Introduction 1](#_Toc59612415)

[Background 1](#_Toc59612416)

[Purpose 4](#_Toc59612417)

[Literature review 4](#_Toc59612418)

[Material and Methods 4](#_Toc59612419)

[Experiments 5](#_Toc59612420)

[Data description 6](#_Toc59612421)

[LOD.csv 6](#_Toc59612422)

[Movement.csv 6](#_Toc59612423)

[Beep.csv 7](#_Toc59612424)

[Methods 7](#_Toc59612425)

[LOD analyzer 7](#_Toc59612426)

[Movement analyzer 8](#_Toc59612427)

[Beep analyzer 9](#_Toc59612428)

[Results 9](#_Toc59612429)

[Comparison of the different scenarios 10](#_Toc59612430)

[Agents' adaptation process with their environment 11](#_Toc59612431)

[Comparison of clone and not clone groups 13](#_Toc59612432)

[Agents' behaviors 16](#_Toc59612433)

[Beep versus mute statistics (based on beep files) 16](#_Toc59612434)

[Discussion 17](#_Toc59612435)

[Conclusions 17](#_Toc59612436)

[References 17](#_Toc59612437)

# Introduction

The state of the art machine learning technique to optimize neural network AI controllers (agents) is back-propagation and specifically deep learning. This method assumes that solutions and examples for such solutions are known. A different approach to this is evolutionary computation and genetic algorithms. Instead of rewarding proper responses to specific situations and hoping for proper generalization, a genetic algorithm uses an objective function. This objective function evaluates the total performance of a system and can select controllers that perform better. Over many generations such system optimizes controllers, often finds more creative solutions [1], and most importantly does not require knowledge about specific actions but only the desired outcome. This method works well when optimizing individual controllers as their objectives do not conflict with the objectives of other controllers. When agents need to interact or work in teams [2], the situation becomes much more complicated, and often the goals of the individual are contrary to the success of the group. Imagine self driving cars. Optimizing them to reach their goal as fast as possible might induce negative side effects. How to overcome such issues? Group level selection plays an important role in evolution. Often individuals do not act alone, but in groups. As such they can achieve together more than alone. Collaborative hunting is one of those examples, where individuals are not selected individually but as a group [3]. This group level selection scheme typically pools the resources the group collected and redistributes them back equally, which implies that groups are evaluated by their average performance. What if this scheme is altered? Imagine a group is rewarded according to the performance of its best individual (all for one) or its worst performer (one for all)? In the all for one case, you would pool all resources on one individual, and in the case of the one for all you would distribute the resources as fairly as possible. The big question is, which of the two groups collected the most resources in total? This research will use neuro-evolution where agents are simulated in a virtual environment and controlled by so called Markov Brains. Group level selection regimes such as the ones described above will be tested with respect to their effect on individual and group level performance.

## Background

Societies are depending on the collaboration of their members. This collaboration is the basis of the economy, health care system, and education among others. Obviously, such institutions benefit from everyone contributing to them via taxes. However, since they are carried by the society, individuals who for example evade taxes can still benefit from them while also benefiting from the resources they did not contribute due to cheating.

To study this problem social science, experimental economics, and game theory use the "Public Goods Game." [citation] In this game, the participants are given equal amounts of money. They then can decide to withhold this money (defect) or contribute (cooperate) into a public pot. The total contribution of the all participants will be multiplied by a synergy factor. This synergy represents gains that can only be achieved by pooling resources. The now larger amount of money in the pot will then be divided equally among all the players - regardless of them having contributed or not in the first place. Considering the game rules, a player who is careful about the social benefits; will contribute money, and those who are selfish and only consider their personal benefits will not. The scenario that individuals only think about their gain exploiting the contribution of others is called "Tragedy of the Commons". [citation]

The question is now how people can be incentivized to be collaborative? In other words, how we can avoid the tragedy of the commons?

If we see it from a government perspective, we would consider incentives to motivate people or punishment to avoid selfish approaches. But for a broader historical perspective, we have to ask where cooperation within groups comes from initially? We find that Darwinian Evolution Theory also struggles to explain why organisms cooperate [Nowak five rules for the evolution of cooperation nature 2006 article], regardless organisms, including humans, evolved to cooperate. To discuss how natural selection is the solution for the collaboration problem, let's first examine how we humans in agriculture select the best genes for reproduction. In this example, a farmer, based on a certain criteria (such as the biggest plants) selects the desired ones for reproduction.



First Generation

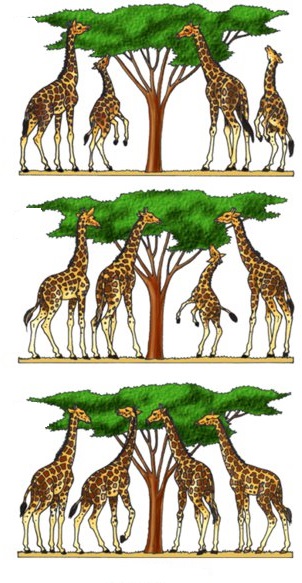
Second Generation

Third Generation



Figure 1. Human selection

In the above picture, the farmer harvests only big cabbages in the last generation by choosing only the biggest for reproducing. Nature does the same to all organism, and only lets some of them survive through history, and the rest will become extinct.



First Generation

Second Generation



Third Generation

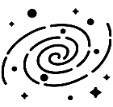


Figure 2. Natural selection

So, nature acts as the farmer and lets only the tall giraffes survive and produce offspring as they had an advantage in eating leaves on high trees. We call this *natural selection*. In the context of natural selection, the individuals' ability to survive and reproduce is called *fitness (w).* You may ask how it comes that we have a variety of different species if only the fittest survives? Then all organisms should look like each other. The answer lies within the two other concepts of mutation and selection. Mutation means sudden and significant changes in the species genetics that results in offspring being different from their parents. Suppose we take a family of giraffes and divide them into two distinct islands which one of them has tall trees, and the other has short trees. If there were no mutation in nature, regardless of the islands' differences, the giraffes after generations would not be different while both groups are from the same family. The mutation causes significantly taller and shorter offspring in the two groups. In the short tree island, the smaller ones will survive and reproduce, and the opposite happens on the other island. Eventually, after a long time by natural selection and mutation, we will have two types of giraffes (short and tall) from the same routes. Consequently, the natural environment presents organisms with different ways to be the fittest, called niches. Those niches, together with inheritance, variation, and natural selection leads to the biodiversity we observe today. Darwin's *evolution* is, therefore, a mechanism composed of inheritance, variation (random mutation), and selection. It thus leads to an adaptation of organisms that fit their environment better.

However, evolution is inherently selfish, as its rewards happen on a short time scale. Reaping rewards now, and making more offspring immediately, will always outcompete saving rewards and reproducing later. Cooperation needs the investment of multiple partners, but from the tragedy of the commons, we know that the defectors fair better. Still, we find different biological mechanisms that allow cooperation to evolve. Among others (kin selection, green beard effect, reciprocity [Nowak 2006]) the one that most likely gave raise to multicellularity is group level selection [citation]. Here, not the individual reproduces, but the entire group benefits from the rewards the entire organism receives. Thus, the most prevalent choice for incentivizing cooperation is group level selection.

Since evolution already solved the problem of cooperation before, in this study, I would like to optimize groups of agents (robots) using a genetic algorithm to cooperate with each other. These agents are controlled by evolvable Neural Networks, specifically Markov Brains [Markv Brains technical introducton 2017 or 2018]. Previous work has already performed preliminary research in this direction. The MABE (Modular Agent Based Evolution) Framework [cite MABE] to run computational experiments has been developed before, and group level selection reward functions has been tested before. I will build in these systems, and test different forms of group level selection, as well as new payoff schemes seeking to improve cooperation. This will be compared to none group level selection.

## Purpose

This research seeks to improve the way we train groups of AI controllers (agents) to perform better individually and in teams at the same time. While this is a basic research question in optimization of neural networks using genetic algorithms, it has direct applications to robotics and other autonomous AI decision making systems that need to work in groups.

As explained in the introduction, the thesis is in the sequence of other works using Markov's brain. Based on the literature review on other works in this chain of researches and also the other similar studies, I decided to work on the impact of incentives or rewarding schemes on team working of the AI controllers. According to the literature review, group level selection has been shown to improve cooperation within groups of agents. Here I introduce a new set of fitness criteria where the effort of the group is not measured by average performance of the group, but instead by either the worst performer or the best performer. This work on *minimum* and *maximum* reward schemes fills a particular *research gap* as it has not been tested before.

## Literature review

# Material and Methods

This research will use the MABE (C++ Modular Agent Based Evolution Framework [4]) to implement virtual test environments. Agents are controlled using Markov Brains [5] which are a particular evolvable type of neural network. After replicating evolutionary experiments are completed, data will be analyzed and visualized.

## Experiments

Simple example

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | | Icon  Description automatically generated | Icon  Description automatically generated | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 1 Scores | | | | | | | | | | A1 0 | | | A2  0 | | A3  0 | | A4  0 | | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | |  | Icon  Description automatically generated |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | |  |  |  | Icon  Description automatically generated |  | |  |  |  | |  |  |  |  |  | | Turn 2 Scores | | | | | | | | | | A1  1 | | | A2  0 | | A3  1 | | A4  0 | | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  | Icon  Description automatically generated |  | |  | Icon  Description automatically generated |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | |  |  |  | Icon  Description automatically generated |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 3 Scores | | | | | | | | | | A1  1 | | | A2  0 | | A3  2 | | A4  0 | |   First generation’s genetics: A1: G1 A2: G2 A3: G3 A4: G4  IDs: A1: 1 A2: 2 A3: 3 A4: 4 |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | | Icon  Description automatically generated | Icon  Description automatically generated | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 1 Scores | | | | | | | | | | A1 0 | | | A2  0 | | A3  0 | | A4  0 | | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | | Icon  Description automatically generated |  | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  | Icon  Description automatically generated |  | |  | Icon  Description automatically generated |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 2 Scores | | | | | | | | | | A1 0 | | | A2  1 | | A3  0 | | A4  1 | |   Second generation’s genetics: A1: G1\*V1 A2: G3\*V2 A3: G3\*V3 A4: G3\*V4  IDs: A1: 5 A2: 6 A3: 7 A4: 8 | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  | Icon  Description automatically generated |  | |  | Icon  Description automatically generated |  |  |  | |  |  |  | |  |  |  | Icon  Description automatically generated |  | |  |  |  | |  |  |  |  |  | | Turn 3 Scores | | | | | | | | | | A1 0 | | | A2  1 | | A3  0 | | A4  2 | | |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | | Icon  Description automatically generated | Icon  Description automatically generated | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 1 Scores | | | | | | | | | | A1 0 | | | A2  0 | | A3  0 | | A4  0 | | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  | Icon  Description automatically generated | | Icon  Description automatically generated | Icon  Description automatically generated | Icon  Description automatically generated |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 2 Scores | | | | | | | | | | A1 1 | | | A2  1 | | A3  1 | | A4  1 | |   Third generation’s genetics: A1: G3\*V2 \*V5 A2: G3\*V4\*V6 A3: G3\*V4\*V7 A4: G3\*V4\*V8  IDs: A1: 9 A2: 10 A3: 11 A4: 12 | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  | |  |  |  |  |  | |  | Icon  Description automatically generated |  | | Icon  Description automatically generated |  |  | Icon  Description automatically generated |  | |  |  |  | | Icon  Description automatically generated |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | |  |  |  | |  |  |  |  |  | | Turn 3 Scores | | | | | | | | | | A1 1 | | | A2  2 | | A3  1 | | A4  1 | | |

**Game settings:**

Variable settings:

|  |  |
| --- | --- |
| Reward mode | Meaning |
| 0 | Individual reward |
| 1 | Mean score (What they got on average) |
| 2 | Maximum score (What their best performer received) |
| 3 | Minimum score (What their worst performer received) |

|  |  |
| --- | --- |
| Group mode | Meaning |
| 0 | Clone |
| 1 | Four different organisms |

Constant parameters:

Other 11 points such as field's dimension, …

## Data description

### LOD.csv

|  |  |
| --- | --- |
| Column name | Explanation |
| Generation | Generation number |
| ID | ID of each agent |
| Score | This field based on rewarding scheme value has the following meanings:   |  |  | | --- | --- | | Reward mode | Score Meaning | | 0 | Individual score | | 1 | Mean score of the four agents | | 2 | Max score of the four agents | | 3 | Minimum score of the four agents | |
| rawScores | Scores of all the group members |
| ownScore | The own score of the agent which is selected in the line of decent |

### Movement.csv

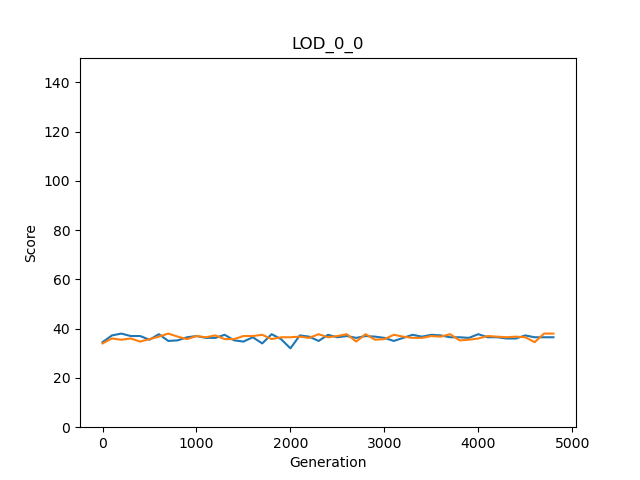
|  |  |  |
| --- | --- | --- |
| Column name | Explanation | |
| T | Turn number | |
| X | Position in X axis | |
| Y | Position in Y axis | |
| D | Direction | Meaning |
| 0 | Up |
| 1 | Right |
| 2 | Bottom |
| 3 | Left |
| E | Number of grasses that the agent gathered. It can be named as energy as well. | |
| A | Action type | Meaning |
| 0 | Do nothing |
| 1 | Turn left |
| 2 | Turn right |
| 3 | Move forward |
| 4,5,6,7 | giving grasses to an agent or putting it on a tile. |
| B | Beep (yes or no) | Meaning |
| 0 | No |
| Any other number | Beep |

### Beep.csv

## Methods

### LOD analyzer

Testing: using 0.25% of data for LOD\_0\_0\_0 and LOD\_0\_0\_1

Graph test (step 3)

LOD statistics test (step 4)

Step 4.1

Firstly, we have checked if we calculated the minimum, maximum, sum, and average of each row correctly. For this case, we tested two of the rows from LOD\_0\_0\_0 and LOD\_0\_0\_1 manually.

Step 4.2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| rewardGroupModes | meanOwnScores | meanMinimums | meanMaximums | meanAverages | meanSums |
| LOD\_0\_0 Replicate 0 & 1  (Excel) | 36.41326531 | 25.83673469 | 48.3775510 | 36.413265 | 145.653061 |
| LOD\_0\_0  Replicate 0 & 1  (Python) | 36.41327 | 25.83673 | 48.37755 | 36.41327 | 145.6531 |

### Movement analyzer

Testing:

First, we should make sure that the percentages for each movement file are calculated correctly. For this proposal, we selected two of the files and compared the results using python and excel.

The next testing part is about the averages, which again we compared python and the Excel results.

|  |  |  |  |
| --- | --- | --- | --- |
| Reward scheme and  group mode | Mean of *giving or putting grasses* percentage | Mean of *do nothing*  percentage | Mean of *beep* percentage |
| Movement\_1\_1 Excel | 0.002 | 2.0775 | 21.978 |
| Movement\_1\_1  Python | 0.002 | 2.0775 | 21.978 |
| Movement\_2\_1  Excel | 2.006 | 0.8965 | 21.8615 |
| Movement\_2\_1  Python | 2.006 | 0.8965 | 21.8615 |

### Beep analyzer

# Results

Improvement of the way we train groups of AI controllers (agents) to *perform better both individually and in teams* is my thesis's aim. Based on the purpose, I have illustrated and categorized the experiment results in five sections. In the first part, which is the central aspect, we have compared the eight different scenarios (the combination of reward schemes and group modes) to identify when they *perform better both individually and in teams* as it was our final target. The other four sections explain the agents' adaptation process with their environment, comparison of clone and not clone groups, the agents' behaviors, and assessment of beep usage.

## Comparison of the different scenarios

As we have discussed in the experiment section, we kept all the parameters constant except the rewarding scheme and group mode. So, by scenario, we mean a combination of reward scheme and group mode, which makes eight different cases. At this step, we check the agents' performance in the different scenarios to identify when they are good both in their payoff (individual) and social welfare (group work).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| rewardGroupModes | Mean  OwnScores | Mean  Minimums | Mean  Maximums | Mean  Averages | Mean  Sums |
| LOD\_0\_0  Group mode: Clone  Reward mode: Individual | 36.4275510 | A picture containing shape  Description automatically generated26.0175510 | 47.4516326 | 36.4275510 | 145.710204 |
| LOD\_0\_1 3  Group mode: Not Clone  Reward mode: Individual | 44.7497959 | 20.1351020 | 52.0906122 | Logo  Description automatically generated36.5244898 | Logo  Description automatically generated146.097959 |
| LOD\_1\_0 2  Group mode: Clone  Reward mode: Average | 36.4372959 | Logo  Description automatically generated25.8336734 | 47.5702040 | A picture containing shape  Description automatically generated36.4372959 | A picture containing shape  Description automatically generated145.749183 |
| LOD\_1\_1  Group mode: Not clone  Reward mode: Average | 47.0285714  Logo  Description automatically generated | 11.6457142 | Logo  Description automatically generated56.7748979 | 33.9260204 | 135.704081 |
| LOD\_2\_0  Group mode: Clone  Reward mode: Maximum | 91.9987755  Logo  Description automatically generated | 2.70897959 | Logo  Description automatically generated91.9987755 | 30.8365306 | 123.346122 |
| LOD\_2\_1  Group mode: Not clone  Reward mode: Maximum | 67.5675510  **A picture containing shape  Description automatically generated** | 1.56244898 | **A picture containing shape  Description automatically generated**75.5930612 | 31.0185204 | 124.074081 |
| LOD\_3\_0 1  Group mode: Clone  Reward mode: Minimum | 30.7763265 | Logo  Description automatically generated30.7763265 | 42.3973469 | Logo  Description automatically generated36.4311734 | Logo  Description automatically generated145.724693 |
| LOD\_3\_1  Group mode: Not clone  Reward mode: Minimum | 34.1008163 | 23.0051020 | 44.5406122 | 33.0433673 | 132.173469 |

Table 1. Comparison of the different scenarios

There are five measurements for each scenario in the above table, and they are ranked using gold, silver, and bronze medals. The own score does NOT mean each agent's score, but its definition depends on the rewarding scheme. For example, when we set the rewarding scheme to maximum, it is the agent's score with the highest score; therefore, it is NOT an indicator of individual performance. We have selected the minimum score for the identifier of the individual's performance because if the lowest score is high, it means all the four agents have had a good performance. In the case of teamwork, we have considered the average and sum scores.

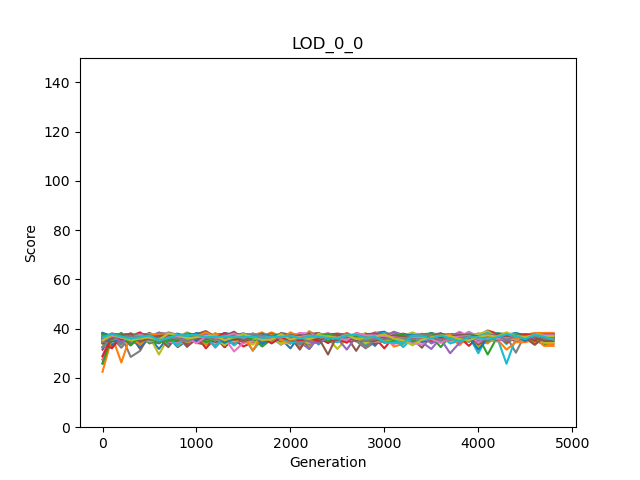
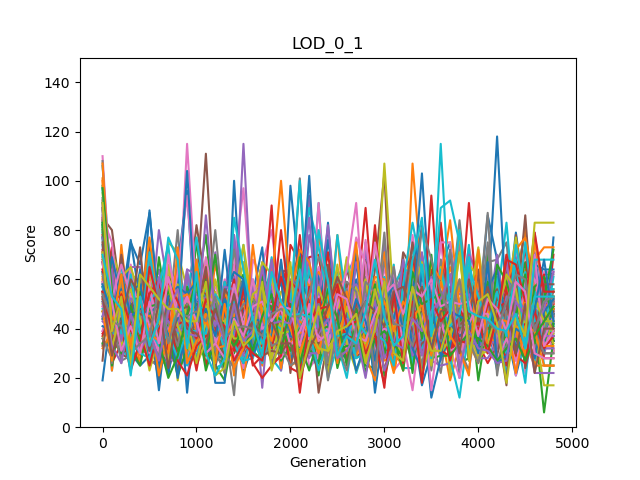
If the count the number of medals, the third scenario (clone and average) and the seventh (clone and minimum) have the highest number (three) of them. The minimum reward scheme has a very better minimum between these two combinations, whereas they have almost similar scores in case of sum and average. Therefore, the first place goes to clone, and minimum and clone and average will take the second position. The default scenario, an individual reward scheme where the agents are not a clone, has the best results in average and sum so, we have placed it in the third position.

To sum it up, when we select clone agent and set the reward scheme as minimum, we would have a high average score, and at the same time, the lowest score is also good, which means all the individuals have a good score. In other words, this combination *trains* our agents to *perform better both individually and in teams at the same time*!

## Agents' adaptation process with their environment

After identifying the best way to train the agents to satisfy the thesis criteria, we will investigate how they adapt to the environment in different scenarios.

Individual | Clone

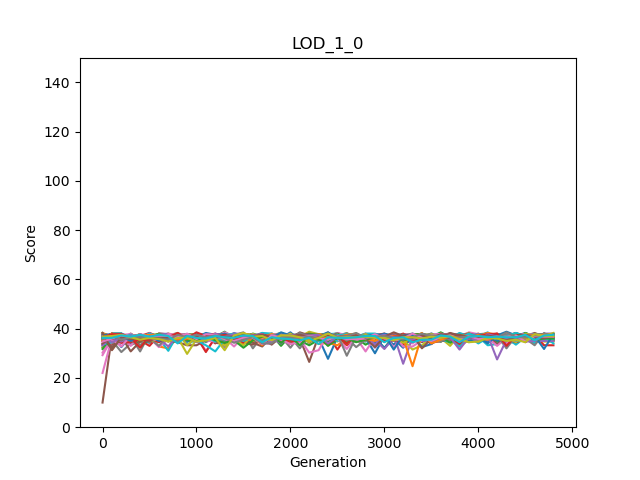
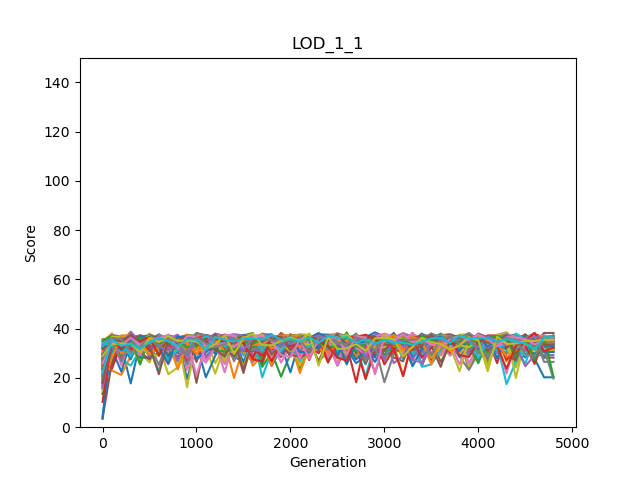


Individual | Not Clone

Figure . Score variations of the individual rewarding scheme

When the rewarding scheme is individual, and agents are clones, they have tiny fluctuations in the score through generations, which is the opposite in different agents (not clone). Based on our investigation comparing the different scenarios, the not-clone and individual has better performance in general compared to the clone version. So, the greater changes in this case, which *probably* means testing more different methods, leads to better performance.

Mean | Clone

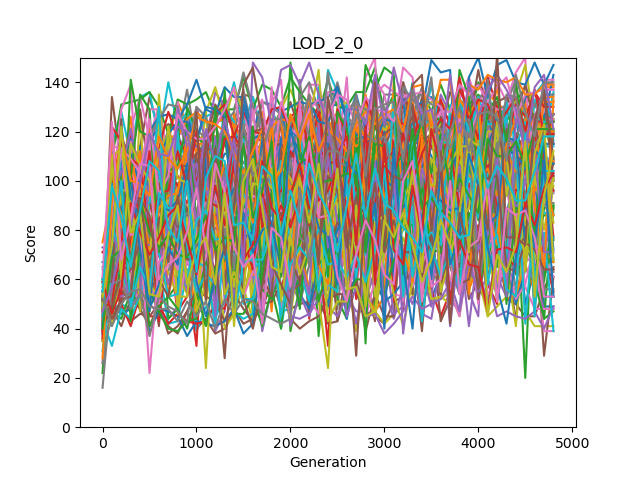
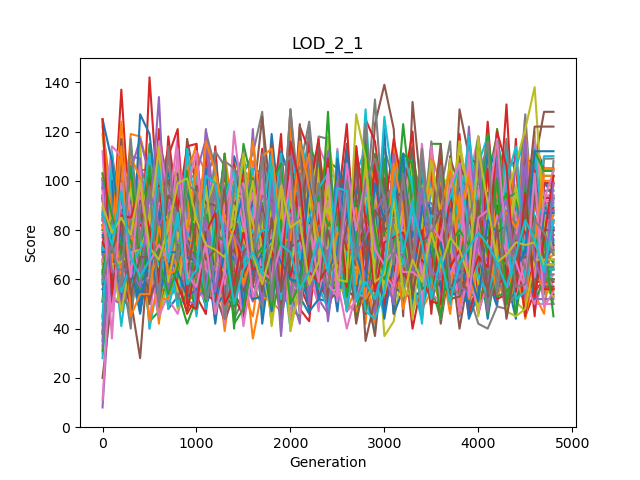


Mean | Not Clone

Figure . Score variations of the mean rewarding scheme

Like the individual reward scheme, different agents' scores have more variations relative to clone agents in the mean reward plan. However, this time the clones have better performance based on the previous table.

Maximum | Clone

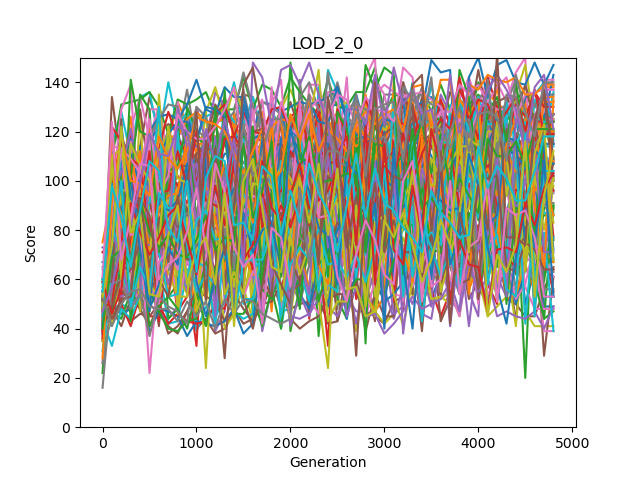
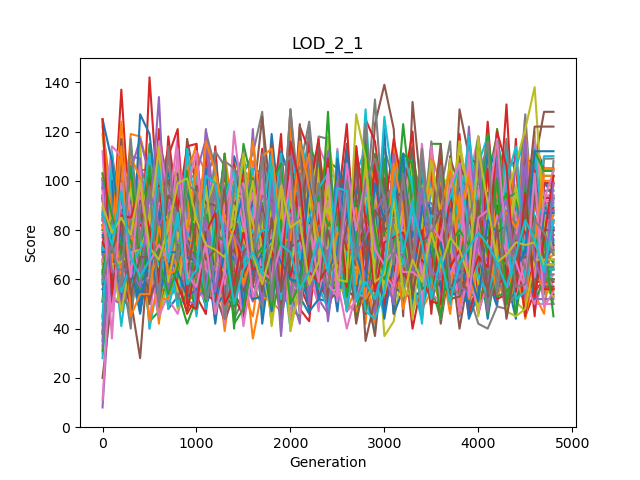


Maximum | Not Clone

Figure . Score variations of the maximum rewarding scheme

The ratio of fluctuations in the score is different for the maximum rewarding scheme, and this time different agents have fewer variations. Clone version with more significant changes has better performance than not clones.

Minimum | Clone



Minimum | Not Clone

Figure . Score variations of the minimum rewarding scheme

The score variations for clone agents are more prominent than not clone in the minimum rewarding scheme, and clones have better performance.

As a conclusion for our observations on fluctuations in the scores through generations, we can say the greater changes, which probably means testing more different methods, lead to better performance in most cases.

## Comparison of clone and not clone groups

In this subsection, we will compare the performance of clones and not clone agents in different rewarding schemes.

Figure 7. comparison of clone VS Not clone modes in different rewarding schemes based on the mean of minimum scores

Based on the above figure, the minimum scores are always higher in clone mode relative to different individuals. So, if we are interested in having good minimum scores, we should use clone mode.

Figure 8. comparison of clone VS Not clone modes in different rewarding schemes based on the mean of maximum scores

Except for the maximum rewarding scheme, the top score is higher when we have different agents. Therefore, if we do not consider setting a rewarding scheme, various agents are better than clones.

Figure 9. comparison of clone VS Not clone modes in different rewarding schemes based on the mean of own scores

As in the previous graph, except for the maximum rewarding scheme, the own score is higher when the agents are not identical. Therefore, if we do not consider setting a rewarding scheme, different agents have better performance.

Figure 10. comparison of clone VS Not clone modes in different rewarding schemes based on the mean of average scores

There is no clear pattern; however, when clone modes have more significant averages, the differences are more prominent. So, we can say that for average, also the clones are better.

Figure 11. comparison of clone VS Not clone modes in different rewarding schemes based on the mean of the sum of scores

Again, there is no clear pattern, and when clone modes have a bigger sum, the differences are more significant. As the average for total sum also clones are slightly better.

In this section's first four figures, the bars separated using rectangles show the scores when the rewarding scheme and score are matched. For example, in the first figure, the vertical axis is for the mean of the maximum score, and the two bars that are for the maximum rewarding scheme are isolated from others using a rectangle. If we check these bars, we can see that clones, in most cases, more carefully listen to us and follow the rewarding scheme relative to different agents.

## Agents' behaviors

In the comparison of the different scenarios, we have found the best three combinations that were 1. Clone and minimum, 2. Clone and average, and 3. Not clone and individual. In this part of the results, we will check which actions are the reason behind success or failure in our agents' performance. We have worked on three actions that are explained in the next table.

|  |  |
| --- | --- |
| Action name | Explanation |
| beep | Beep is simply the way of our agents' communication. It can have different values. When it is zero, they do not use a beep, and the other values have their meaning. In this thesis, I only worked on two statuses that are using or not using the beep. |
| giving or putting grasses | We have allowed the agents to give their energy (grass) to other agents or simply put it on an empty tile. In this study, we have considered both cases as one action. |
| do nothing | As its name suggests, it means that the agent does nothing in its turn. |

|  |  |  |  |
| --- | --- | --- | --- |
| Reward scheme and  group mode | Mean of *giving or putting grasses* percentage | Mean of *do nothing*  percentage | Mean of *beep* percentage |
| Individual  Clone | 0 | 0.7005 | 13.4225 |
| Individual  Not clone | 0 | 0.136 | 24.147 |
| Average  Clone | 0 | 0.2965 | 17.156 |
| Average  Not clone | 0.002 | 2.0775 | 21.978 |
| Maximum  Clone | 24.948 | 3.304 | 53.814 |
| Maximum  Not clone | 2.006 | 0.8965 | 21.8615 |
| Minimum  Clone | 0 | 0.963 | 51.2315 |
| Minimum  Not clone | 0 | 8.1085 | 9.5385 |

Table 2. The Agents' actions

I have highlighted the top three values for each action in the table of agents' actions. The best combinations (reward scheme and group mode) are precisely matched with the highest values of beep usage. So clearly, using beep has a significant positive impact on agents' performance. Minimum and not clone is our worst combination without any medal, and it has the highest value of do nothing so, we can conclude that doing nothing is not a good strategy. However, we should also consider that this action's second-highest value is belong to maximum and not clone agents, which is one of our best combinations. Therefore, unlike beep, we can not make a clear conclusion in terms of do-nothing action. Because the percentage of giving the energy to other agents or putting it on an empty tile is way bigger in case of maximum and clone, it is a useful action if they want to have an agent with a very high score.

## Beep versus mute statistics (based on beep files)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Reward  And Group | Beep  Mean Minimums | Mute  Mean Minimums | Beep  Mean Maximums | Mute  Mean Maximums |
| Individual  Clone | 26.09 | 25.24 | 47.19 | 46.11 |
| Individual  Not clone | 26.958 | 26.954 | 47.53 | 47.43 |
| Mean  Clone | 26.14 | 25.27 | 47.45 | 46.20 |
| Mean  Not clone | 25.41 | 25.45 | 46.97 | 46.99 |
| Maximum  Clone | 1.56 | 3.74 | 92.88 | 34.00 |
| Maximum  Not clone | 25.13 | 25.57 | 46.89 | 47.41 |
| Minimum  Clone | 31.51 | 8.72 | 41.49 | 15.19 |
| Minimum  Not clone | 23.72 | 23.46 | 43.85 | 43.47 |

Table 3. Beep vs. mute for the mean of the minimum and maximum scores

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Reward  And Group | Beep  Mean Averages | Mute  Mean Averages | Beep  Mean Sums | Mute  Mean Sums |
| Individual  Clone | 36.40 | 35.40 | 145.61 | 141.63 |
| Individual  Not clone | 36.97 | 36.96 | 147.88 | 147.84 |
| Mean  Clone | 36.52 | 35.46 | 146.10 | 141.86 |
| Mean  Not clone | 35.88 | 35.92 | 143.55 | 143.69 |
| Maximum  Clone | 30.00 | 16.24 | 120.03 | 64.97 |
| Maximum  Not clone | 35.75 | 36.21 | 143.03 | 144.87 |
| Minimum  Clone | 36.46 | 11.89 | 145.84 | 47.56 |
| Minimum  Not clone | 33.44 | 33.15 | 133.77 | 132.61 |

Table 4. Beep vs. mute for the mean of the average and sum scores

# Discussion

# Conclusions

# References

|  |  |
| --- | --- |
| [1] | L. Joel, C. Jeff and M. Dusan, "The surprising creativity of digital evolution," in *Artificial Life Conference Proceedings 14, pages 76-83. MIT Press*, 2017. |
| [2] | H. Arend and M. Masoud, "Evolution of autonomous hierarchy formation and maintenance," in *Artificial Life Conference Proceedings 14, pages 366-367. MIT Press*, 2014. |
| [3] | B. Gergely and S. Szabolcs, "Beneficial laggards: multilevel selection, cooperative polymorphism and division of labour in threshold public good games," *BMC evolutionary biology,* 2010. |
| [4] | B. Clifford and H. Arend, "Mabe (modular agent based evolver): A framework for digital evolution research," in *Artificial Life Conference Proceedings*, 2017. |
| [5] | H. Arend, A. E. Jeffrey, S. O. Randal, B. K. David, S. Jory, A. Larissa, T. S. Ali, K. Peter, S. e. Leigh och G. Heather," Markov brains: A technical introduction". |