

Effect of Group Size on Group Cooperation

Paul Wanjohi
Department of Computer Science
18 University Avenue, University of Cape Town
Rondebosch, Cape Town, South Africa
paul.wanjohi@alumni.uct.ac.za

ABSTRACT

Cooperation in insects is a social behaviour that has gained interest in disciplines such as Economics, Game Theory and Evolutionary Biology. As a result models are designed to explain the emergence of cooperation of a group. The focus of this paper is the effect of group size on the levels of cooperation of a group. Reeve and Hölldobler present a nested tug of war theory that predicts that larger groups exhibit high levels of altruism than smaller groups. An Agent Based System is used to perform an empirical analysis on this claim. The results demonstrated that there is evidence not to reject the claim that group size has an effect of the levels of cooperation in a group.

CCS Concepts

•Computing methodologies → Cooperation and coordination;

Keywords

Altruism, Kin Selection Theory, Agent-based Models

1. INTRODUCTION

Cooperation is a remarkable social behaviour observed in nature where the actions of an individual benefit other individuals even though it might negatively affect them. An example of this behaviour can be observed in social insects as they forage. Some of the insects in the colony would expel their energy by carrying food while others would expel their energy protecting the colony. Both activities involve sacrificing themselves to ensure the survival of the colony. The reason as to why some individuals sacrifice themselves to benefit the colony has been actively studied from Biology to Economics from Sociology to Game Theory. As a result, there are acceptable theories that have been shown to describe the forces that govern cooperation between and within groups in a colony. Key theories commonly cited include: the kin selection theory (first described by Hamilton) [22], trait-group selection [15], reciprocal altruism [37] and tug-of-war theory [31].

For this paper, Agent-Based Systems (ABS) are used to investigate the effect of colony sizes on the degree of between-group cooperation. The sizes of the groups are initially fixed and may grow in size due to reproduction. The environment has a finite number of resources and is such that selfish behaviours lead to extinction. This situation is the tragedy of commons. In this situation, agents must share common

finite resources and possibly sacrifice their own benefits so that the colony may survive through generations.

The main motivation behind this research is to replicate one of the four predictions offered by Reeve and Hölldobler's model [32] on cooperation colonies. In their research, they introduced an individual selection model that predicts elegant patterns relating competition and cooperation in both within and between-groups in a colony. One of the patterns documented was the intricacy of cooperation in larger group sizes when compared to others. This research replicates a similar model to illustrate the same trend on group size and cooperation with the help of an ABS. Ultimately, the viability of ABS modelling techniques is illustrated as modelling technique in understanding social behaviour.

The next section offers the background information on cooperation among social insects. It provides detail on the use of ABS as a tool for modelling social behaviour. Conclusively, the use of neuro-evolution will be presented as means of adapting agents in an ABS. The following section focuses on the details of experiment design. After which a discussion of the results from the experiments conducted is discussed. Ultimately, conclusions are drawn that may support or refute Reeve and Hölldobler's prediction.

1.1 Research Question

This paper seeks to support or refute the null hypothesis stated as follows: *Group size affects the level of cooperation of a group.* The alternative hypothesis is *Group size does not affect the level of cooperation of a group.* In this paper will consider both the within group cooperation and between group cooperation as values that contribute to a group's level of cooperation.

2. BACKGROUND AND RELATED WORK

2.1 Cooperation in Social Insects

Vast literature is published on cooperation and its emergence within social insect societies. Cooperation can be seen in social insects as they forage for food, defend and build the nest, and care for the brood [11, 41].

A common focus on cooperation is the Prisoner's Dilemma (PD) problem proposed by Flood [16]. Models on the evolution of cooperation have been considered as pair-wise interactions between the cooperator and defector. Nowak [28] introduces the term "cooperative dilemma", coined due to its similarity to the classic Prisoner's dilemma by Rapoport [30], as a game between two strategies: cooperation and defection. For cooperative dilemma to occur, the benefit a

individual obtains from the cooperator is higher than the cost of cooperating with the individual [28].

The use of the PD to define cooperation has been criticised for defining altruism instead of cooperation [35]. This raises the issue of semantics especially when cooperation and altruism are not well differentiated. This issue is addressed by West et al. [39] when they define cooperation as the behaviour which is beneficial to others but could be beneficial and/or costly to the cooperator.

Altruism is defined as the behaviour which is beneficial to others but costly to the individual during the individual's lifetime. The individual with altruistic has a decrease in absolute fitness at the end of their lifetime. Reeve and Hölldobler [32] did not make the distinction of cooperation and altruism when describing the patterns predicted by their mathematical model. For the purpose of this research, the well-accepted definitions of cooperation and altruism will be used. In particular, this research is conducted with the assumption that Reeve and Hölldobler [32] meant altruism in their paper. To distinguish between the two terms we use cooperation with mutual benefit to mean cooperation.

Cooperative and altruistic behaviour has been observed and models created to attempt to predict the emergence in such societies. An early observation and model popularised by Hamilton [21]. Hamilton observed that altruistic behaviour among relatives can be explained by kin selection where selfish genes lead to unselfish phenotypes [21]. In kin selection theory, an individual assisting a relative indirectly promotes the transmission of copies of its genes to the next generation by a probability factor of their genetic relatedness. Hamilton's rule relates the change in actor's (cooperator's) personal fitness c , also known as the cost, and the change in the recipient's personal fitness b , also known as the benefit, by a constant r that is the degree of genetic relatedness of the actor and the recipient. The rule states that for altruism to occur, Equation 1 should hold for the altruistic individuals in the society [21].

$$rb - c > 0 \quad (1)$$

The interactions or social behaviour exhibited by an individual in a society influences the reproductive success of the society. In order to understand the evolution of a social behaviour, such as cooperation, evolutionary game theory and population genetics are identified as key approaches by Nowak [28].

Entomologists have observed puzzling trends among social insect colonies, where inactivity appears to be highly common to social insect colonies typically occupying more than 50% of the colony task time [10]. Pinter-Woolman et al. [29] highlights that individual workers do not simply differ in activity levels and argue that if that were so then individual activity levels could be predictable. However, they found consistent trends in activity levels across *related* tasks and worker clusters were formed that frequently performed tasks in different situations. Doebeili and Hauert [13] claim the key concept behind cooperation as frequency-dependent selection, a concept well embodied in evolutionary game theory.

Frequency-dependent selection which arises when the fitness of a phenotype, is dependent its frequency in a population [6]. For example, cooperative and altruistic may be fitter than their counterparts when they are many of them.

Cooperation and altruism in eusocial insects have also been attributed to the trait-group selection model. In this model, behaviour of groups could affect the groups reproduction and survival. As noted by Wilson, it can be used to predict the evolution of altruism in animal societies [40]. Ultimately, group selection models can be mathematically translated to individual selection model. This implies that by simulating an individual selection model, a group selection model we need only introduce a group structure. Whereas for kin selection, we ensure that the model has a non-random placement of the individual within the group (relatives grouped up together).

Reeve and Hölldobler's research present an individual selection model In the next section, an ABS is introduced as a simulation tool to mimic Reeve and Hölldobler's evolutionary model and observe the trends predicted by their model.

2.2 Agent Based Systems

Agent-Based Models (ABM), as argued by Axelrod [3], is proven useful when exploring issues that arise between two or more disciplines. In his work on the possibility of cooperation among egoists, Axelrod [3] highlights the value of ABS as his work drew strategic ideas from Economics, Political Science, Psychology, Sociology and Mathematics. Additionally, ABM have been used by immunologists to find optimal schedules for Highly Active Antiretroviral Therapies that yields maximum immune restoration [9].

Axelrod and Hamilton [4] were among the first to create an ABM to simulate the evolution of a cooperative behaviour in a human political society. In their their experiments, they test the levels of cooperation and competition in the simulation and compare them with existing mathematical models presented in Political Science by Axelrod.

ABM consist of entities generically known as agents and an environment. These agents are: autonomous, interdependent, led by simple rules, and adaptive [26]. The agents' interdependence may arise directly between agent interactions or indirectly via their interaction with the environment. Furthermore, the agents adaptation arises by a process of learning which alters the probability distribution of the agents' behaviours.

In agent-based models, each agent can perceive information and make decisions independently but constrained on some decision rules [44]. Agent-based models (ABM) are often used to understand an observed social behaviour where replicating the mechanisms that lead to social behaviour in simulations aid in understanding the emergence of such behaviour in biological systems [8].

Agent-based modelling is a method for testing the collective effects of selections of actions an individual makes. A simulation based on these autonomous agents are ideally suited to explore the implication of non-linearity in system behaviour and also adapts to models that are easily scalable to scope and level [20].

2.3 Agent Controller Design

The design of the agent controller is what governs the adaptive behaviour of the agent. Techniques that are used to construct this adaptive behaviour stem from tools in the Artificial Intelligence (AI) domain in particular under the Machine Learning (ML) sub-domain. These tools are usually used as solution finding tools where the solution space is extremely large. Approaches used under ML include: Evo-

lutionary Algorithms (EA), Reinforced Learning, and the use of Artificial Neural Networks (ANN).

A common approach taken when working with ABM in modelling evolving social behaviour is the use of Evolutionary Algorithms. EA create a search heuristic that follows elements of natural selection. This paper focuses on the use of Neuro-Evolution as a tool use for the agent controller.

Neuro-Evolution

Neuro-Evolution is the application of Evolutionary Algorithms to the design and learning of ANN. [43, 17, 23] Neuro-evolution is a biologically-inspired machine learning approach, where machine learning is a sub-field of artificial intelligence. Neuro-evolution is a product of two sub-fields of machine learning: Evolutionary Algorithms and artificial intelligence. From this classification of neuro-evolution, its definition can be further polished to include: neuro-evolution as a heuristic mechanisms used to prove theories through inference rules (machine learning nature) [12] and neuro-evolution as a means of evolving ANN [43].

The effectiveness of the neuro-evolution Algorithms lies in the effectiveness its constituent components: ANN and EA.

Evolutionary Algorithms

Evolutionary Algorithms are characterised by a population of solution candidates, in the case of agent-based models, an agent's behaviour is defined by its artificial neural networks. The reproduction process of an evolutionary Algorithm enables the combination of existing solutions to generate new solutions [2].

Evolutionary Algorithm is inspired by natural selection where the genotypes of the fittest individuals in the population are propagated into the next generation of individuals. After a number of generations (a number that is dependent on the complexity of the system being simulated), the population should be filled with fit individuals.

The algorithm as described in Figure 1 is used to find a target solution s_{target} within a population of possible solutions. For neuro-evolution, some clarification is needed to understand how the algorithm is altered for use with neuro-evolution.

Firstly, the target solution s_{target} may be defined as an observable attribute exhibited by an individual in the system while the individuals, p_i are defined as an encoding of the ANN that govern them. In this case, s_{target} is said to be a *phenotype attribute* and the individual p_i are said to be the genotype, as it encodes an observable characteristics of the individual. In neuro-evolution, the genotype is an encoding of the neural network which can be encoded as the weight values of the synapses of individual ANN.

Secondly, the reproductive process of the individuals makes use genetic operations: crossover (sexual reproduction), mutation (randomly altering the genes of an individual) and reproduction (surviving to the next generation unaltered). Koza et al. [24] clarifies that these genetic operations are applied to selected individuals on a probabilistic nature based on their fitness. This selection process favours the fitter individuals over weaker ones. However, the fittest (or weakest) individuals may not necessarily be selected for.

Evolutionary Algorithms owe their success to their effectiveness in searching for solutions in a huge space of possibilities. The problem for evolution, as noted by Wright [42], can be conceptualised as a search for, ideally, the global max-

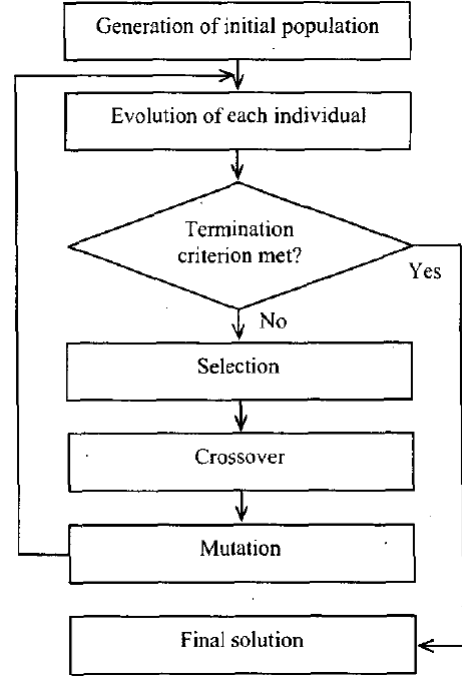


Figure 1: A flowchart describing a generic Evolutionary Algorithm [34]

ima in a multi-dimensional field of gene combination where height corresponds to fitness. Banzhaf [7] identifies the importance of genetic operations in achieving genetic diversity in the population. Genetic diversity allows for the exploration or exploitation of the search space; an identifying feature for Evolutionary Algorithms. The search becomes difficult as the dimensions and the number of local optima increase. [5] In order to reduce these factors, designing of the ANN (genotype of the agents) should aim to make the ANN as lightweight as possible (with few weight values).

Artificial Neural Networks

Artificial Neural Network (ANN) modelling is a “black-box” non-linear machine learning technique which links input data and output data with a set of non-linear basis function [27, 23]. They are a technique inspired by biological nervous systems and consist of simple synchronous processing elements, known as neurons, interconnected by synapses [25]. Each synapse has a weight connection value. The importance of these weight values are used to define the contribution of the inputs from the input neurons or environment. ANN have largely been used in a number of classification tasks where the network seeks to match a set of inputs to a target output; as a result, they are able to develop predictive models for classification [33, 27, 14, 38].

In the case of agent-based systems with neuro-evolution, each individual is modelled with an ANN and is used to determine the individual decision in the simulated environment. Construction of the neural network's topology is determined by the designer of the system. The most common implementation of the ANN adapted with neuro-evolution involves a set of inputs that are obtained from the simulated environment while the set of output would be the decision

of the agent [33].

The non-linearity of the neural network would prove advantageous to the construction of the agent based system as it explores the implication of non-linearity of the social behaviour. Furthermore, autonomy of the agents in the simulation would be well defined by associating an ANN to an agent.

3. EXPERIMENT DESIGN

In Section 2, some background is provided on the existing mathematical models that attempt to explain the emergence of cooperation in social insects. Additionally, an existing simulation technique, ABM, is introduced as an alternative way of observing emerging social behaviour in societies of autonomous agents.

In this section, the design of the experiment is discussed. After-which a brief description of the target model (Reeve and Hölldobler’s model [32]) is presented.

3.1 Design

The experiments are designed to simulate foraging of resources in groups of social insects.

Environment Design

This simulation of foraging of resources is designed in a 2D environment which accommodates agents and resources. The environment does not contain any obstacles. The environment represents collection of resources called resource patches. These patches are areas of the environment where the resources units are uniformly distributed. The number of resources within the patches are not predefined but can be approximated by the patch’s density. The density of a patch is the probability of finding a resource in a patch.

Since the focus of simulation is not the patch density, the environment contains only one patch and a fixed patch density throughout the simulation. The choice of this representation of resource patches allows for randomising resource positions between experiments. It also allows for a measure of patchiness which is a component which is also investigated by Reeve and Hölldobler.

Agent Design

Each agent in the simulation perceives the environment by knowing the position of the closest active resource to the agent. There are no obstacles in the environment. Similar to the construction of the Reeve and Hölldobler’s model, at the beginning of the simulation every agent is provided with an equal initial energy value.

Every agent in the simulation is exclusively part of one group of agents. The resources collected by agents are retained by the group in the form of agent energy. The group acts as a pool for agents to reproduce and is synonymous to species or colonies. As such, migration of agents between groups is prohibited.

Furthermore, each agent is driven by an ANN with the same topology. The topology, as illustrated in Figure 2, includes: two input neurons, one hidden neuron layer with two hidden neurons and one output neuron. At the beginning of the simulation the weights of the ANN of each agents are randomised under a uniform distribution. The ANN seeks to classify the input signals as perceived by the agent. This input signals include: the fraction of the field of view of

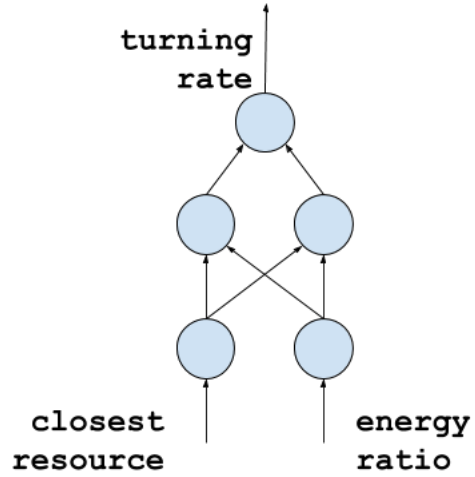


Figure 2: Topology of ANN driving the agents in the experiments

the agent that is the distance to the closest resources to the agent (denoted as **closest resource** in Figure 2) and the percentage energy of the agent relative to other member in the agent’s group (denoted as **energy ratio** in Figure 2). This classification process occurs at every time step or *tick* and defines the **turning rate** (illustrated in Figure 2) of the agent in the environment.

After every t ticks, a new generation of agents is introduced in the environment which replaces the existing population of agents. The simulation creates g generations of agents and records the performance of each generation. The new generation of agents are created using the help of a Generation Genetic Algorithm (GGA).

The GA seeks to evolve the ANN driving the behaviour of the agents. The ANN are represented as a list of weights and are also known as the genes of the agent. The GA is described in Algorithm 1 below. Important features of the algorithm include: the use of tournament selection as the parent selection mechanism where k is set to 5, the crossover operator used is the 1-Point Uniform Crossover, the mutation operator reassigns weight values with some probability. Lastly, the *ExtractValidParents* method creates a subset $V \subset A$ such that the agents $v \in V$ have not reproduced yet and are in the same group as the initial partner p_0 if there is an initial partner.

The energy of the agent defines the fitness of the agent. This value is consumed by the agent when the agent moves and is increased when an agent collects a resource. The movement of the agent and thus its behaviour is driven by the respective ANN. Suppose the agent moves on a particular tick, the agent consumes a portion of their energy. This portion contributes to the cost of cooperation in the simulation. If the agent collides with a resource, a portion of the agent’s energy is replenish contributing to the benefit of cooperation. The choice of using the agent’s energy as the fitness value provides an indication of the behaviour of the agent. Cooperative agents are expected to have more energy than defectors. Therefore, the GGA would be seeking to maximise the energy of the agents in the simulated environment. Every unique resource collected by the agents in a group are tallied for the group.

Algorithm 1 Genetic algorithm

Require: A is a sorted list of agents w.r.t. their energies

Require: p_0 is the agent seeking a partner

Require: k is the tournament size

```
1: procedure TOURNAMENTSELECTION( $p_0, A$ )
2:    $arena \leftarrow \text{EXTRACTVALIDPARENTS}(p_0, A)$ 
3:    $tournament \leftarrow \emptyset$ 
4:   if  $|arena| > k$  then
5:     for  $i \leftarrow 1$  to  $k$  do
6:        $r \leftarrow \text{rand}() \bmod |arena|$   $\triangleright$  random number
7:        $tournament \leftarrow tournament \cup arena[r]$ 
8:    $arena \leftarrow tournament$ 
9:   else if  $|arena| == 0$  then
10:    return null
11:  return  $\text{MAX}(arena)$   $\triangleright$  the best in the  $arena$ 
```

Require: ρ is the number of children

```
12: procedure REPRODUCTION( $A$ )
13:    $C \leftarrow \emptyset$   $\triangleright$  list of next generation
14:    $p_0 \leftarrow \text{TOURNAMENTSELECTION}(null, A)$ 
15:    $p_1 \leftarrow null$ 
16:   while true do
17:     if  $p_0 == null$  then
18:       break
19:     else
20:        $p_1 \leftarrow \text{TOURNAMENTSELECTION}(p_0, A)$ 
21:     if  $p_1 == null$  then
22:        $r \leftarrow \text{rand}() \bmod |A|$   $\triangleright$  random number
23:        $p_0 \leftarrow A[r]$ 
24:       continue
25:     if  $p_0 \neq null$  and  $p_1 \neq null$  then
26:       for  $i \leftarrow 1$  to  $\rho$  do
27:          $c \leftarrow \text{CROSSOVER}(p_0, p_1)$ 
28:          $\text{MUTATE}(c)$ 
29:          $C \leftarrow C \cup \{c\}$ 
30:    $p_0 \leftarrow \text{TOURNAMENTSELECTION}(null, A)$ 
31:  return  $C$ 
```

These resources collected benefit the next generation of agents. Suppose that the first generation of agents each obtain an initial energy store a_0 . On the k th census, each agent i in group j has collected p_{ij} resources of which $t_j (\leq \sum p_{ij})$ are unique to the group. The starting energy of the k th generation of agents in group j containing n_j agents is $(a_k)_j = a_0 + \frac{t_j \times \beta}{n_j}$ where β is the energy value for a unit resource. The choice of this distribution of energy for new agents is to ensure that the new generation of agents directly benefit from the behaviour of the previous generation. In the case where agents in a group collect resources it is expected that t_j would be higher than if they chose to stay still.

3.2 The Biological Model

The model proposed by Reeve and Hölldobler is used to model for competition and by proved implication cooperation within groups and between groups. In the model of competition, they let the number of resources be R and each agent has an initial energy store, τ , and is placed in one of N cooperative groups in the environment each of which have n

agents. A fraction f of the agent's energy store τ is selfishly retained for within group competition while $1 - f$ is invested in the group, within group cooperation. On a group level, each of the N cooperative groups have an initial energy store contributed by the members within the group. A fraction F of the energy store is selfishly retained for the group for between group competition while $1 - F$ is invested in the colony of groups, between group cooperation.

Reeve and Hölldobler defined the agent's share of within group competition as:

$$S_w = \frac{(1-f)\tau}{(1-f)\tau + r(n-1)(1-f)\tau + (1-r)(n-1)(1-F)\tau} \quad (2)$$

where r is the fraction of the agent's group that is genetically related the agent. The overall group-level competition for the agents group is described in Equation 3. The g is a constant.

$$G = ((1-f)\tau + r(n-1)(1-f)\tau + (1-r)(n-1)(1-F)\tau) \quad (3)$$

The agent's group share of the resource is presented in Equation 4. The value for q is the relatedness of the group with respect to the agent.

$$S_b = \frac{G}{G + q(N-1)G + (1-q)(N-1)((n-1-F)\tau)} \quad (4)$$

$$W = S_w S_b R \quad (5)$$

Reeve and Hölldobler proposed a fitness function (Equation 5) that use the within and between group competition values posed in Equation 2 and 4. Maximising this fitness function demonstrated that an agent would attain evolutionary stability when $f = F$.

Since the focus of this research is the levels of altruism, the one major changes are made to the competition model presented by Reeve and Hölldobler. Instead of defining f as the selfish fraction, this research paper defines f as the fraction invested by an agent to their group. With this modification, S_w and S_b are redefined as the within-group and between-group cooperation values respectively. This modification is as discussed in the Reeve and Hölldobler's paper.

Due to the strict conditions set in the Reeve and Hölldobler's competition model some assumptions were made pertaining to the values of the terms mentioned in Equations 2, 3 and 4. Firstly, τ is the maximum possible energy the agent has. This means if an agent at generation k collects p_{ij} which are worth β energy units each then $\tau = a_{kj} + \beta \times p_{ij}$ where a_{kj} is the starting agent's energy at generation k in group j . This assumption caters for transfer of agent's energy between the environment and the group.

Secondly, r is considered as the ratio of "productive" agents to the "unproductive". A "productive" agent has collected at least one resource. Therefore, if an agent is "productive", its r value will be the fraction of "productive" agents in the group. This assumption of relatedness works on the assumption that the fraction described above provides an approximation to an agent's genetic relatedness to a group. The assumption made in this regard is that "productive" agents are likely to be related to "productive" agents within a group. Similarly, the group's relatedness value, q , is the

fraction of the total population size that contribute energy for the group.

Thirdly, F is the fraction of uniquely collected resources in a group within the environment. This assumption states that the selfish fraction retained by a group can be defined as a fraction of the number of resources, R , in the environment. This assumption may be verified by checking if evolutionary stability is attained, that is, does the selfish fraction among agents converge to F .

The formal hypothesis being tested in this experiment is: do groups with a large starting number of agents exhibit higher within-group cooperation values, S_w , and higher between-group cooperation values, S_b , than groups that start with smaller number of agents?

For simplicity, two groups of agents would be introduced in the simulation. One of the groups would have a larger number of agent than the other group. During the lifetime of the agents, the agents lose energy when they move. If the agent has no energy, then the agent is removed from the population. After t ticks, the new generation of agents is created using Algorithm 1. The number of agent in the new generation is determined by the number of children, ρ . This evolutionary process is repeated for a series of g generations. The nature of this outlined process of experimentation follows certain biological assumptions with the prediction that some intelligent behaviour, cooperation, will emerge. This experiment can thus be categorised as an artificial life experiment [36].

4. IMPLEMENTATION

A C++ application framework was developed to conduct the experiment designed in the previous section. The main purpose of this framework is to test the four organisational trends in colonies of autonomous agents predicted by Reeve and Hölldobler’s model [32].

Since the focus of the experiment is the effect of group size on the levels of altruism within and between groups, two constraints are placed on the use of the framework. The framework uses the genetic relatedness for reproduction and for group placement. Genetic relatedness is used in reproduction as only related agents are allowed to pair up for reproduction. Furthermore, it is also used to ensure groups of relatives stick together. The use of genetic relatedness allows the framework to test applications of the kin selection theory to the model. However, this experiment does not allow migration between groups and reproduction is only allowed within the group. The second constrain is resource patchiness. The patch density and the number of patches is fixed throughout the experiment.

The parameters used in this experiment are illustrated on Table 1 where the notation column represents symbols used in this paper to represent certain parameters.

The 2D environment has the same dimensions as the resource patch to ensure that the resources are distributed throughout the environment. The resource regrowth rate is the number of ticks that a resource may remain inactive before it reactivates. This values is set to 100 and divides the number of ticks for one generation, t , perfectly. If this value is set really low, the resources will re-spawn a lot more often. When the resources re-spawn more often, agents may take opportunity of this and collect resources in one area of the environment. This opportunistic behaviour may increase the resources collected for an agent p_{ij} but restricts the number

Meaning	Notation	Value
Number of Groups	N	2
Number of initial agents in Group 0	n_0	2000
Number of initial agents in Group 1	n_1	200
Number of ticks	t	2000
Number of generations	g	1000
Number of patches	-	1
Patch density	-	0.038
Patch Width	-	800
Resource regrowth rate	-	100
Number of children	ρ	2
Tournament size	k	5
Selection Pressure	-	1.5
Crossover rate	-	0.7
Mutation rate	-	0.1
Energy gained per resource	β	20
Energy Movement cost per tick	-	0.2
Starting agent energy	a_0	30.0
Field of View	-	100
Environment Width	-	800
Environment Height	-	800
Number of experiment runs	-	20

Table 1: Table illustrating the parameters used in the experiments

of unique resources collected for the group, t_j . The same argument holds for the patch with high patch density.

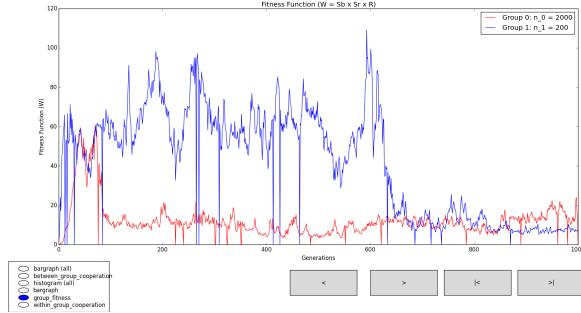
The crossover rate identifies the position on the ANN to perform a 1-point uniform crossover, while the mutation rate identifies the probability to for reassign a weight value within an agent’s ANN. The selection pressure is used for a rank based parent selection, however, since this experiment uses a tournament based selection, a tournament of five is used to moderate selection pressure.

The artificial life experiment needs to be performed over a number of runs to draw statistically sound conclusions. As such this experiment is conducted over 20 runs and the average performance of the 20 runs is discussed in the next section. A total of five experiments are conducted with each using the exact same parameter except the value for n_1 which is set to 200, 400, 600, 800, and 1000 over the five experiments.

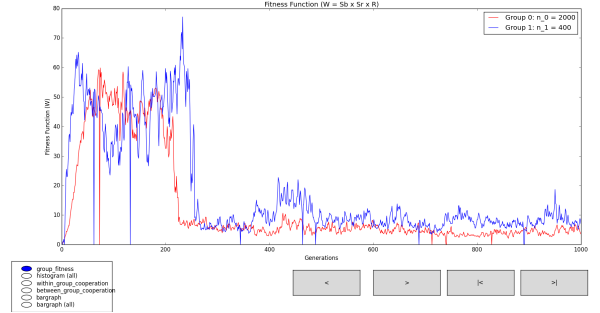
5. RESULTS AND DISCUSSION

The graphs in Figure 3 represent the fitness values of the experiments conducted. These results illustrated combine the within group cooperation values and between group cooperation values as shown in Equation 5. In this section, the data collected will be used to compare the fitness values of different groups between experiments. In this regard, a dataset will be considered as the groups fitness values averaged over the 20 runs recorded over 1000 generations.

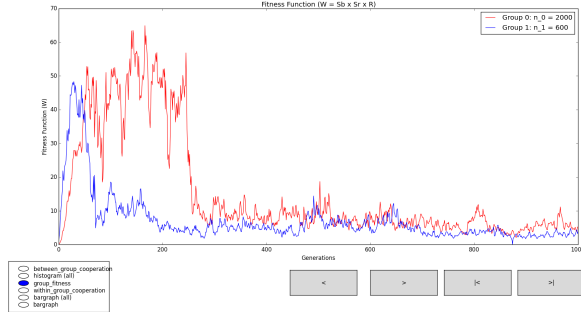
Due to the stochastic nature of this experiment, statistical tests such as the Kolmogorov-Smirnov test and the t-test are conducted to assess the performance of the experiments [18]. The data was tested with the Kolmogorov-Smirnov test for goodness of fit test to check if the datasets were normally distributed. The null hypothesis claimed is datasets are normally distributed. The test conducted yielded p-values of zero across all the data sets. The results of this test show



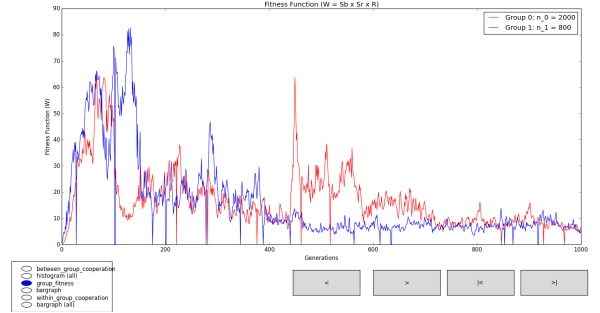
(a) Experiment 1: Group Fitness Function for the experiment testing when n_1 is 200



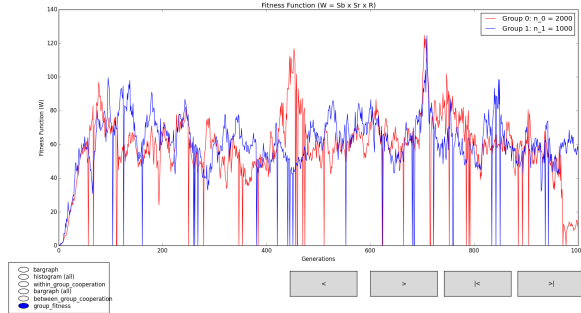
(b) Experiment 2: Group Fitness Function for the experiment testing when n_1 is 400



(c) Experiment 3: Group Fitness Function for the experiment testing when n_1 is 600



(d) Experiment 4: Group Fitness Function for the experiment testing when n_1 is 800



(e) Experiment 5: Group Fitness Function for the experiment testing when n_1 is 1000

Figure 3: Group Fitness functions as described in Equation 5

that the datasets do not follow a normal distribution and thus rejects the null hypothesis claimed earlier. However, the assumption that if the experiments would be recorded with more generations then the larger dataset will be normally distributed is made.

Under this assumption, we may perform a t-test to test if the datasets are significantly different from one another. The null hypothesis used to test the datasets states "For any dataset d and dataset c , d is significantly different from c ". The t-statistics provided the p-values shown in Table 2 and for the purpose of this experiment a critical value of 0.1 is used. In Table 2 cells with the dashes indicate the test cases that were not applicable.

Results of the t-test with the critical value of 0.1 demonstrates that Group 1 in Experiment 2 is significantly different from Group 0 in experiment 4, because its 0.146 p-value is

greater than the critical value. This result is complemented on Figures 3d and 3b which illustrates that the Group 1 that starts with 400 agents benefits with higher fitness values in the early generations than the Group 0 that starts with 2000 agents. However, the benefit achieved from these high fitness values last long as the benefit begin to diminish just after the 200 generation mark while the larger group enjoys the benefit for a shorter time as it diminishes before the 200 generation mark. However, as the generations of agents are created, Group 0 that started with 2000 agents shows the potential of increasing if fitness values as demonstrated in Figure 3d. This potential increase in fitness levels while Group 0 experience consistently lower fitness values demonstrates the difference group size has on the fitness levels. This corresponds to the elaborate cooperation values for larger groups than the smaller groups [32]. Fluctuation

	Exp. 3 G. 1	Exp. 3 G. 0	Exp. 2 G. 1	Exp. 5 G. 0	Exp. 1 G. 1	Exp. 4 G. 0	Exp. 4 G. 1	Exp. 5 G. 1	Exp. 2 G. 0	Exp. 1 G. 0
Exp. 3 G. 1	-	✗	✗	✗	✗	✗	✗	✗	✗	✗
Exp. 3 G. 0	✗	-	✗	✗	✗	✗	✓ 0.088	✗	✗	✗
Exp. 2 G. 1	✗	✗	-	✗	✗	✓ 0.146	✗	✗	✗	✗
Exp. 5 G. 0	✗	✗	✗	-	✗	✗	✗	✗	✗	✗
Exp. 1 G. 1	✗	✗	✗	✗	-	✗	✗	✗	✓ 0.043	✗
Exp. 4 G. 0	✗	✗	✓ 0.146	✗	✗	-	✓ 0.055	✗	✗	✗
Exp. 4 G. 1	✗	✓ 0.088	✗	✗	✗	✓ 0.055	-	✗	✗	✗
Exp. 5 G. 1	✗	✗	✗	✗	✗	✗	✗	-	✗	✗
Exp. 2 G. 0	✗	✗	✗	✗	✓ 0.043	✗	✗	✗	-	✗
Exp. 1 G. 0	✗	✗	✗	✗	✗	✗	✗	✗	✗	-

Table 2: Table showing the p-values obtained from the t-test on all datasets. (where Exp. denotes Experiment, G. denotes the Group, ✓ denotes significant p-values and ✗ denotes the insignificant p-values)

occurring towards the tail of the graphs align with the result previous mentioned as the values of Group 1 fluctuate below the 10 mark, but the fitness values of Group 0 fluctuate just above the 10 mark.

Observations obtained from the experiment showed that collections of agents would follow each other and random instances of in their lifetime. This behaviour allowed leaders among those collections of agents to collect the resources first. The behaviour was controlled by altering the resource regrowth rate which would allow a collection of agents to scatter when a resource respawns. Once they scattered, the agents who chose to move would participate in the same behaviours as they follow other agents. As the census of the agents drew closer, more and more agents would stay still with fewer exceptions that would be collecting resources.

A few challenges in the experiment included the choice of the assumptions in order to use Reeve and Hölldobler’s model. Among the assumptions made was that of the relatedness of an agent which is detailed in Section 3.2. The assumption provided an approximation for the genetic relatedness value of an agent based on the questionable assumption that the “productive” agents may share gene qualities among other “productive” agents.

Reeve and Hölldobler’s use of energy in their model was created as a tool to derive the fitness of agents based on their within group and between group cooperation values; however, an agent’s energy is not synonymous to the fitness of the agent itself. The approach used in this paper assumes and agent’s energy level is directly proportion to the fitness value of the group. However, it is the selfish fraction of the agents that is proportional to the group’s fitness value. By understanding this, the GGA in place would seek to optimise the fitness value, W , as presented in Equation 5 rather than maximising the energy value of an agent. This may have resulted in agents that seek not to cooperate in order to conserve their energy.

The challenges outlined above are present in most artificial life experiments where the theoretical models and empirical analysis of the model might provide design challenges [36]. At which point, certain explicit assumptions are placed to assist in the verification of the empirical results against the prediction of the theoretical model. Under the assumptions made in this experiment, we cannot reject the hypothesis that group size has an effect on the cooperation levels of the group.

6. CONCLUSIONS

The experiment results demonstrated that the group of 400 starting size exhibits a significant difference in fitness levels when compared to a group of 2000 starting size. Assumptions made in the experiment were biologically inspired and were used to tackle the design challenges experienced in the experiment. Observations from the experiment demonstrated the existence of temporary spatial groups of agents that follow each other providing a natural way of grouping agents. Ultimately, the experiment demonstrated that there is evidence to suggest that group size has an effect on the levels of cooperation of the group and thus we cannot reject the null hypothesis presented in Section 1.1.

The results of the experiment were based off the Reeve and Hölldobler model which is based on the Kin Selection theory. This theory is a mathematical approach that tends to focus on the partial analysis of the total evolutionary change [1]. In order to provide a wholistic analysis of the total evolutionary change in cooperation levels among agents a model based on theories that provide complete descriptions of natural selection such as Price’s Equation should be used [19]. As for now this partial analysis provides a foundation for other evolutionary models to be empirically analysed.

Acknowledgements

We thank the supervisor Dr. Geoff Nitschke for his guidance and assistance in setting up the design of the experiment and Zachary Melnick and Jacob Clarkson who assisted in the development of the C++ software framework. Computations were performed using facilities provided by the University of Cape Town's ICT High Performance Computing team: <http://hpc.uct.ac.za>.

7. REFERENCES

- [1] The price equation, fisher's fundamental theorem, kin selection, and causal analysis. *Evolution*, 51(6):1712–1729, 1997.
- [2] A. Abraham and L. Jain. *Evolutionary multiobjective optimization*. Springer, 2005.
- [3] R. Axelrod. Chapter 33 agent-based modeling as a bridge between disciplines. volume 2 of *Handbook of Computational Economics*, pages 1565 – 1584. Elsevier, 2006.
- [4] R. Axelrod and W. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.
- [5] R. M. Axelrod. *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton University Press, 1997.
- [6] F. J. Ayala and C. A. Campbell. Frequency-dependent selection. *Annual review of Ecology and systematics*, pages 115–138, 1974.
- [7] W. Banzhaf. *Foundations of genetic algorithms*. Morgan Kaufmann Publishers Inc., 1999.
- [8] J. J. Bryson, Y. Ando, and H. Lehmann. Agent-based modelling as scientific method: a case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485):1685–1699, 2007.
- [9] F. Castiglione, F. Pappalardo, M. Bernaschi, and S. Motta. Optimization of haart with genetic algorithms and agent-based models of hiv infection. *Bioinformatics*, 23(24):3350–3355, 2007.
- [10] D. Charbonneau, N. Hillis, and A. Dornhaus. 'lazy' in nature: ant colony time budgets show high 'inactivity' in the field as well as in the lab. *Insectes Sociaux*, 62(1):31–35, 2015.
- [11] J. T. Costa and T. D. Fitzgerald. Developments in social terminology: semantic battles in a conceptual war. *Trends in Ecology & Evolution*, 11(7):285–289, 1996.
- [12] N. S. Desai and R. Miikkulainen. Neuro-evolution and natural deduction. In *Combinations of Evolutionary Computation and Neural Networks, 2000 IEEE Symposium on*, pages 64–69. IEEE, 2000.
- [13] M. Doebeli and C. Hauert. Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology Letters*, 8(7):748–766, 2005.
- [14] S. Dreiseitl and L. Ohno-Machado. Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5):352–359, 2002.
- [15] J. A. Fletcher and M. Zwick. The evolution of altruism: Game theory in multilevel selection and inclusive fitness. *Journal of theoretical biology*, 245(1):26–36, 2007.
- [16] M. M. Flood. Some experimental games. *Management Science*, 5(1):5–26, 1958.
- [17] D. Floreano, P. Dürri, and C. Mattiussi. Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1):47–62, 2008.
- [18] C. M. Fonseca and P. J. Fleming. On the performance assessment and comparison of stochastic multiobjective optimizers. In *Parallel problem solving from nature - ppsn iv*, pages 584–593. Springer, 1996.
- [19] S. A. Frank. George price's contributions to evolutionary genetics. *Journal of theoretical biology*, 175(3):373–388, 1995.
- [20] C. Goldspink. Modelling social systems as complex: Towards a social simulation meta-model. *Journal of Artificial Societies and Social Simulation*, 3(2):1–23, 2000.
- [21] W. D. Hamilton. The evolution of altruistic behavior. *The American Naturalist*, 97(896):pp. 354–356, 1963.
- [22] W. D. Hamilton. The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52, 1964.
- [23] V. Heidrich-Meisner and C. Igel. Neuroevolution strategies for episodic reinforcement learning. *Journal of Algorithms*, 64(4):152–168, 2009.
- [24] J. Koza, M. Keane, M. Streeter, W. Mydlowec, J. Yu, and G. Lanza. *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*. Genetic Programming Series. Springer, 2005.
- [25] R. Lippmann. An introduction to computing with neural nets. *ASSP Magazine, IEEE*, 4(2):4–22, Apr 1987.
- [26] M. W. Macy and R. Willer. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 28:pp. 143–166, 2002.
- [27] S. Malinov, W. Sha, and J. McKeown. Modelling the correlation between processing parameters and properties in titanium alloys using artificial neural network. *Computational Materials Science*, 21(3):375–394, 2001.
- [28] M. A. Nowak. Evolving cooperation. *Journal of theoretical biology*, 299:1–8, 2012.
- [29] N. Pinter-Wollman, J. Hubler, J.-A. Holley, N. R. Franks, and A. Dornhaus. How is activity distributed among and within tasks in temnothorax ants? *Behavioral Ecology and Sociobiology*, 66(10):1407–1420, 2012.
- [30] A. Rapoport. *Prisoner's dilemma: A study in conflict and cooperation*, volume 165. University of Michigan press, 1965.
- [31] H. K. Reeve, S. T. Emlen, and L. Keller. Reproductive sharing in animal societies: reproductive incentives or incomplete control by dominant breeders? *Behavioral Ecology*, 9(3):267–278, 1998.
- [32] H. K. Reeve and B. Hölldobler. The emergence of a superorganism through intergroup competition. *Proceedings of the National Academy of Sciences*, 104(23):9736–9740, 2007.
- [33] S. Russell, P. Norvig, and A. Intelligence. A modern approach. *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, 25, 1995.
- [34] X. Shi, L. Wan, H. Lee, X. Yang, L. Wang, and Y. Liang. An improved genetic algorithm with

- variable population-size and a pso-ga based hybrid evolutionary algorithm. In *Machine Learning and Cybernetics, 2003 International Conference on*, volume 3, pages 1735–1740. IEEE, 2003.
- [35] E. Sober. The evolution of altruism: Correlation, cost, and benefit. *Biology and Philosophy*, 7(2):177–187, 1992.
- [36] L. Steels. The artificial life roots of artificial intelligence. *Artificial life*, 1(1-2):75–110, 1993.
- [37] R. L. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1):pp. 35–57, 1971.
- [38] J. V. Tu. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology*, 49(11):1225–1231, 1996.
- [39] S. A. West, A. S. Griffin, and A. Gardner. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2):415–432, 2007.
- [40] D. S. Wilson. *The natural selection of populations and communities*. Benjamin/Cummings Pub. Co., 1980.
- [41] E. O. Wilson. *The Insect Societies*. Harvard Paperbacks. Belknap Press of Harvard University Press, 1971.
- [42] S. Wright. Evolution and the genetics of populations. vol. 1. genetic and biométrie foundations. *Evolution and the genetics of populations. Vol. 1. Genetic and biometrie foundations.*, 1968.
- [43] X. Yao. Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9):1423–1447, 1999.
- [44] J. Zhong, L. Luo, W. Cai, and M. Lees. Automatic rule identification for agent-based crowd models through gene expression programming. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1125–1132. International Foundation for Autonomous Agents and Multiagent Systems, 2014.