Tips:

- Communicate with Kostas

- Use figures!

- Ensure structure is flowing

- Use sources for everything

- Modularise in order to convert to paper

To Do:

- Modularise in order to convert to paper

- Describe appendix

- Act on Interim Review Feedback

- Act on planning and timescale comments from interim report

- Act on contents of summary of completed work from interim report

- Change the Rationale to fit Final report

- Write a literature review or combine with Rationale and add to it?

- Write Contents and Knowledge

- Write critical analysis and discussion

- Write professional issues (get background reading and citations)

Today:

- Indirect reciprocity

- Intro

- Summary

- Add figures

- Spell check lit review

- Read through and check over

- Ensure correct flow

- Go through EvolCoop and IndirRec to ensure I haven't missed anything

- Make it pretty and easier to read

Final Year Project Report

Full Unit - Final Report

Cooperative Strategies in Multi-Agent Systems

James King

A report submitted in part fulfilment of the degree of

BSc (Hons) in Computer Science

Supervisor: Kostas Stathis

Department of Computer Science

Royal Holloway, University of London

February 23, 2019

Declaration

This report has been prepared on the basis of my own work. Where other published and unpublished source materials have been used, these have been acknowledged.

Word Count: 1,587 *2,000* intro + n *3,000* lit review + n *4,000* framework + n *4,000* experiment evaluation + n *1,000* conclusions + n *1,000* professional issues

Student Name: James King

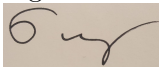Date of Submission: February 23, 2019

Signature:

Table of Contents

Abstract

Chapter 1: **Introduction**

1.1   **Motivation**

Artificial intelligence (AI) has been an idea present in the consciousness of humanity for millennia. From Hephaestus' might Talos to Edgar Allan Poe's commentary on 'Maelzel's Chess-Player' the idea has inspired both awe and confusion. As we move away from the mythical and the false, AI embeds itself deeper into our lives and societies.

Yoav Shoham [33] introduced the agent-oriented programming paradigm. He noted it's close relation to the field of AI. Intelligent Agents (IAs) - which I am using as roughly a synonym for agents, even though it implies a higher level of reasoning the lines are blurred - and Multi-Agent Systems (MASs) are a key concepts in the agent-oriented programming paradigm. IAs have a range of definitions, but the widest definition was given by Russell and Norvig [31] as anything that perceives and acts upon its environment. Our world is one of the toughest environments in which we could place an intelligent agent.

Advances in software and hardware are taking us closer to perceiving in the real world in the field of computer vision and acting in it in the field of robotics. So we can perceive and act, but there is a big step to acting in a rational manner, defined by Russell and Norvig [31] as doing the right thing. Rationality is a huge philosophical dilemma, and one that I shall not go into detail about.

This step towards rationality is a large part of the study of IAs. However, this does not only include agents set in our world, but also those such as softbots interacting across the internet and various other environments. From the data formulated through perception how can an agent represent the world around it and understand what these perceptions can tell them about their world and from this how can the agent act rationally.

One main aim of the study of agent systems is to produce rationally acting agents. In real life rationally acting agents would also need to be socially acting agents in order to reach their goals. In fact, social ability is one of the key properties of IAs laid out by Wooldridge and Jennings [43] and is one of the aspects that is tested in the imitation game [17].

This is another key area of study for the field of IAs and MASs. Many have marvelled at the development of robotic drones being able to dance through the sky together in beautiful displays. These displays are stunning of course, and if using an agent architecture represent a step towards cooperation between IAs. However, these small societies of drones are programmed to work together most likely by the same people.

If artificially intelligent agents are to reach their full potential and integrate into our society, they require the ability to cooperate with agents who can be thought of as 'unknown entities' in terms of their intentions. Both the world of the drones and our world can be thought of as MASs. Intelligent agents in our world need to act in the same environment as us and the many different species that also reside on planet earth.

As this is one of the final goals of IAs and MASs it appears a natural strategy to use studies of cooperation in the natural world as guidance for the study of IAs and MASs. The early days of evolutionary theory were groundbreaking, but Peter Kropotkin [14] identified a key flaw in accounting for cooperation between animals.

There are many problems to solve before the technology is available to ethically and realistically integrate IAs into our society. The problem I aim to grapple with in this project is how to facilitate cooperation between IAs in MASs.

1.2   Aims and Objectives

The high-level aim of my project is **to study how developers can facilitate cooperation between IAs and protect them from those with malicious intent**. The objectives to achieve this aim are laid out here.

The closest example to a multi-agent system that has allowed cooperation to flourish is that of the natural world. I will be using the study of cooperation in the natural world to inspire my search for a mechanism to facilitate cooperation in MASs. Further to this I will use past studies of multi-agent systems and game theory in regards to the evolution of cooperation. One objective of this study is *to learn about what has previously been identified as possible contributors to the evolution of cooperation.*

The deeper objective of this review is *to identify an appropriate mechanism for further in-depth study towards its possible contribution in my high-level aim.* Using a broader study of multi-agent systems as a whole to guide me in this objective.

After identifying an appropriate mechanism my next objective will be *to develop a theoretical framework surrounding this mechanism.* Within this framework will be concepts to allow me to study the mechanism and other variables that may have an effect on the evolution of cooperation (such as an agent communication language to study social ability). Furthermore, this theoretical framework will include strategies to associate with agents.

My next objective will be *to develop an application that allows a user to set up, simulate and analyse MASs that implement the theoretical framework I have laid out.* This application will aim to allow the user to change the variables previously identified as possible contributors to the evolution of cooperation in order to study them.

My high-level aim also includes spreading any knowledge I gain from this study. As such, another objective is *to make the system distributable to people who wish to study the theoretical framework I have outlined.*

Using this application my las main objective will be *to examine the relevance and success of the mechanism I have identified to the evolution of cooperation.* I will also use the application to measure how the variables set up in the framework and the various strategies implemented in the program affect the evolution of cooperation. The aim of this is to discover if and possibly how these various aspects can be used to facilitate cooperation between IAs. To summarise:

- Main aim: to study how developers can facilitate cooperation between IAs and protect them from those with malicious intent.

- Objective 1: to learn about what has previously been identified as possible contributors to the evolution of cooperation.

- Objective 2: to identify an appropriate mechanism for further in-depth study towards its possible contribution in my high-level aim.

- Objective 3: identify other factors that may contribute to the evolution of cooperation.

- Objective 4: to develop a theoretical framework surrounding this mechanism and other factors.

- Objective 5: to develop an application that allows a user to set up, simulate and analyse MASs that implement the theoretical framework I have laid out.

- Objective 6: to make the system distributable to people who wish to study the theoretical framework I have outlined.

- Objective 7: to examine the relevance and success of the mechanism I have identified to the evolution of cooperation.

- Objective 8: to examine the effect of other variables involved in my MAS in regards to the evolution of cooperation.

1.3   Contribution

The goal of this project and report is to find areas to explore in regards to facilitating cooperation between IAs. This field is large and a lot of ground has already been covered. I am aiming with the mechanism I identify to find a niche which hasn't been explored as in-depth as other areas, either specifically with the mechanism or a combination of the central mechanism and a number of other factors.

As such I aim to create a theoretical framework that is unique. This will allow me to examine some less explored factors for the evolution of cooperation, and hopefully find some interesting results as to how they affect cooperation among IAs.

Another contribution I am aiming to make is to make the application I develop distributable. This has been done by others such as the Python Axelrod library [6] and the evolution of trust website [4], but I hope to introduce a less Iterated Prisoner's Dilemma focused mechanism. Agents will be programmed using a logic language (Prolog) meaning their decisions will be transparent.

1.4   Structure

The structure of this report is to follow the objectives laid out above. First I will conduct a literature review to conquer the objectives 1-3. This literature review will help me conduct research into the field and gain an in-depth understanding of what experiments have taken place and what ideas have been put forward. On review of the literature, I will be gaining an understanding also of what has not been explored and/or what could do with a more in-depth exploration. From this, I will be able to identify what mechanism and which factors to use in my theoretical framework.

This leads me nicely into my framework section which aims to tackle objectives 4-6. Here I shall use the mechanism and factors I have identified to lay out a framework with which I wish to experiment. This framework will aim to suit the arena of MASs and be examinable in terms of the factors I have identified. The next part of my framework section will be in regards to the implementation of this theoretical framework. This will be in an application that is distributable to people interested in the problem and framework.

Once I have designed, implemented and tested this application, I will be using it for objectives 7 and 8. For these objectives, I will conduct a number of experiments to analyse both the mechanism and the factors built into the theoretical framework. To do this I will be controlling variables using the applications set up phase and then reviewing the outcome using the analysis phase. I will then discuss and analyse the results of these experiments. This discussion and analysis will drive my conclusions at the end of the report.

Chapter 2: **Background**

2.1   Introduction

The evolution of cooperation between individuals is a keenly studied topic, and as such there is a large set of related literature. Approaches to the problem have come from areas such as biology, social systems, game-theory and of course, IAs and MASs. This background section will review past literature in the topic of agent interactions, with a view to exploring the best options for facilitating cooperation in MASs.

The literature review will tackle the first 3 objectives from my introduction. By reviewing past approaches I aim to gain knowledge surrounding previous explorations of the evolution of cooperation. This research will inform my decision as to the mechanism to use to facilitate cooperation, and which factors to study which may aid or discourage the evolution of cooperation in a MAS.

2.2   Cooperative Phenomena

Russell and Norvig [31] describe an agent as being able to perceive and interact with their environment (figure 2.1). But a more interesting definition comes from Singh [34] in which he claims a key conceptof MASs is the interoperability of the agents. Cooperative actions can occur during interactions between agents: where the action of one benefits another, at a loss to the actor. A defection is the opposite of cooperation: the action does not benefit the recipient of the action, but it is not a loss to the actor. This cooperative phenomenon is also present in biology.

Early evolutionary theory struggled to explain why cooperation is so prevalent in nature. In fact, it seemed that competition was key to evolution as individuals were competing to survive; the notion famously coined by Herbert Spencer "Survival of the fittest" [36].

Axelrod and Hamilton [2] note two key areas of study that attempt to explain cooperative phenomena: Kinship Theory and Reciprocal Altruism. These two theories are useful for explaining cooperation in nature, and as such, it may be possible to apply them to MASs in order to facilitate the evolution of cooperation between agents.
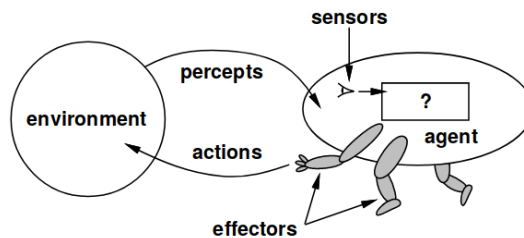

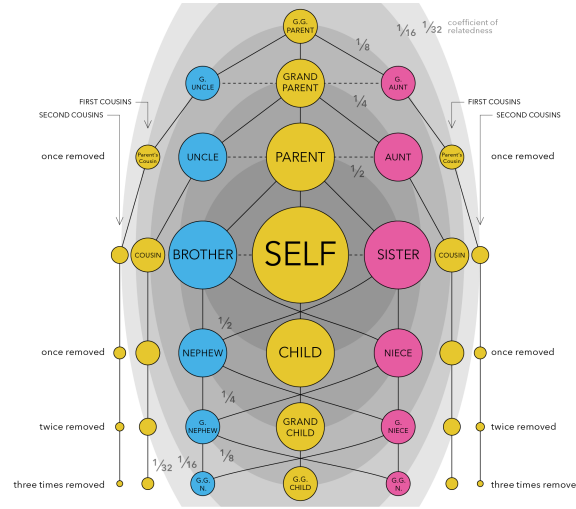Figure 2.1: Russell and and Norvig's [31] agent

Figure 2.2: Wright's coeffcient of relatedness by Citynoise - Own work, CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=37723128`

2.3   Kinship Theory

Axelrod and Hamilton [2] described the way in which cooperation in nature (with the exception of homo-sapiens) is almost always between related individuals. An earlier paper by Hamilton [12] argues that individuals don't only work toward improving their own fitness, but towards what Hamilton defines as 'inclusive fitness'. Inclusive fitness is the sum of a players fitness and the fitness of each of their relations multiplied by a coefficient. The coefficient used by Hamilton is Wright's coefficient of relatedness, described visually in figure 2.2. It could be possible to create a similar coefficient of relatedness for use in a MAS.

Richard Dawkins [5] advocated for the idea of the selfish gene. From a biological perspective this is the idea that actors are hardwired to propagate their genes. Dawkins asserts that this is due to genes being the true replicators evolutionarily rather than the actors themselves. Those with a high coefficient of relatedness to an individual are far more likely to carry their genes and help proliferate them, providing a similar mechanism to Hamilton [12] but with a biological backing. From a biological perspective this may make sense, however, it does not seem natural to translate an agents strategy to the idea of genes.

Further, though it is possible to create a coefficient and an idea of relatedness similar to Hamilton's model [12] for a MAS, this does not seem a natural translation. Another limitation to the use of kinship theory for MASs is that we are aiming to make our systems inclusive of individuals that can contribute to the society. Say an agent is looking to actively contribute to a society, but they are not kin with the members, a MAS using kinship theory would exclude them and thus limit the abilities of that society.

Furthermore Axelrod and Hamilton [2] highlighted that humanity is the exceptional society which does not limit itself to cooperating mostly only with kin. I would surmise that this is due to the higher level of intellect of homo sapiens in comparison to other species. Many have suggested that the capabilities of AI could match if not surpass the intelligence of humans, so I would suggest that societies of IAs will also not be limited to the use of kinship theory to facilitate cooperation.

Due to these reasons I shall not be using kinship theory for my theoretical framework. I will be aiming to use a mechanism that is inclusive of agents that aim to be valuable members of the society and also a mechanism which fits naturally into the agents paradigm.

Figure 2.3: The Green Wood-Hoopoe native to Africa participates in cooperative breeding, where the bird not only looks after it's own chicks, but of other breeding pairs [11]

Figure 2.4: Cleaning symbioses such as that of the green sea turtle and surgeonfish including the yellow tang show that reciprocal altruism is possible between non-humans and even interspecies [45]

2.4   Reciprocal Altruism

Reciprocal altruism is an idea most famously put forward by Robert L. Trivers [41]. Trivers defines altruism as behaviour that benefits another organism, not closely related, while being apparently detrimental to the organism performing the behaviour. From this definition and Trivers' description we can draw the meaning of reciprocal altruism to be altruism based on the idea that the altruistic act will be returned.

This is a move away from limiting individuals to cooperating with only their kin towards any individual that they believe cooperation will be reciprocated by. Axelrod and Hamilton [2] noted this as an advantage in explaining cooperation between unrelated individuals, such as is common between humans. I would argue that this also makes it more applicable to higher intelligence societies such as those possible from IAs.

In comparison to kinship theory this seems the more natural match to the agents paradigm - not requiring measures for relatedness between agents or conversion between an agents strategy and the concept of a gene. These reasons make reciprocal altruism the more likely out of the two to apply to my problem.
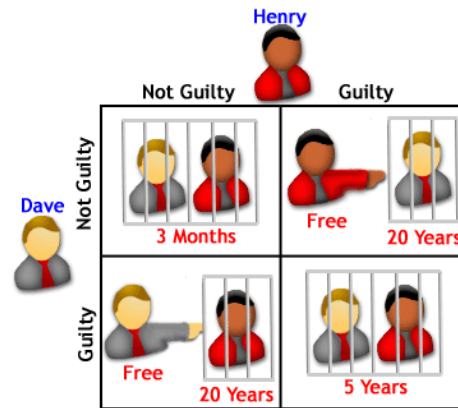
Figure 2.5: A visualisation of the outcomes of the Prisoner's Dilemma [19]

2.5 Axelrod, Hamilton and The Iterated Prisoner's Dilemma

Axelrod and Hamilton [2] also chose reciprocal altruism for further study over kinship theory. They recognised kinship theory limited individuals to only cooperating with kin and at the time more investigation had been into kinship theory. Of course the same cannot be said now, a lot of investigation has gone into reciprocal altruism, mostly due to their paper and Trivers' paper [41].

Game-theoretic modelling was used by Axelrod and Hamilton [2] in their Iterated Prisoner's Dilemma tournaments. The Dilemma can be thought of by imagining that a crime has been committed by 2 individuals who are now being interviewed by the police. Each individual can either choose to inform on the other, or stay loyal and say nothing.

Staying loyal is known as cooperating, informing is defecting. If an individual informs on the other they will receive no punishment, as long as the other has stayed loyal. However, if both defect they both receive a bad sentence, but, not as bad as the individual who stays loyal when the other defects. This is visualised in figure 2.5.

If this dilemma is not repeated a defector can get away with defecting and is not heavily punished for this. However, if the game is repeated in multiple rounds the mechanism of direct reciprocity comes into play. Direct reciprocity is a mechanism that works in the way if I cooperate with you, you will hopefully cooperate with me in later rounds [24].

Across multiple rounds when an individual doesn't know which round will be last it may be more beneficial for them to cooperate, as following rounds will achieve a higher payoff if both cooperate than if both defect. This encourages cooperation without the need for pure altruism, it is a form of reciprocal altruism.

The focus of Axelrod and Hamilton's paper [2] is to review the strategies of The Dilemma submitted to them by academics and teams working on the strategies throughout the world. They asked 3 questions of each strategy. Is it robust? Is it stable? Is it initially viable?

Robustness is the ability to thrive in an environment with a variety of strategies. Stability is the ability to - once fully established - resist invasion by mutant strategies. Initially viability is whether or not a strategy can establish itself is a noncooperative environment.

They found two strategies with these 3 abilities: tit-for-tat and all defect. Later on Nowak and Sigmund [22] found that Pavlov (win-stay, lose-shift) also has these abilities. The interesting part of Pavlov and tit-for-tat is that they are nice strategies (they begin by cooperating) and that they actively aid in the evolution of cooperation.

Another key part of The Iterated Prisoner's Dilemma is the way in which it has been formulated by Axelrod and Hamilton [2]. In each round players simultaneously either choose to cooperate or defect, the players earn points based on the outcome of the game as given in the payoff matrix in table 2.1. Nowak [24] found that for the evolution of cooperation

to occur the cost-to-benefit ratio of the altruistic act must be less than the probability of another encounter for cooperation to evolve: $w > c/b$.

		Player B	
		Cooperation	Defection
	Cooperation	A=3 B=3	A=0 B=5
Player A			
	Defection	A=5 B=0	A=1 A=1

Table 2.1: The payoff matrix in a typical iterated prisoner's dilemma game (such as Axelrod and Hamilton's). A=x, B=y where x denotes the payoff for A and y the payoff for B.

If the probability of another encounter between two individuals is low, it is very likely that cooperation will not evolve as there is no sufficient reward for it. With the advent of huge networks spanning across the world and the drastic increase in devices across these networks, it is highly likely MASs will operate with IAs that are unlikely to re-meet.

This is a definite limitation to the use of direct reciprocity in MASs, and as such I do not see it as a good contender to encourage cooperation on it's own. One property of a mechanism to encourage cooperation is that it must work when both re-meeting is unlikely and when it is likely. So this does not mean to say that direct reciprocity is completely inadequate for the problem, but that it is insufficient when not combined with some encouragement for when re-meeting chances are low.

A lot of interest in reciprocal altruism in recent years has been focused towards direct reciprocity. This includes a number of libraries that are available such as the Axelrod Python library [6]. I argue that to gain a better understanding of how we can work to facilitate cooperation between IAs we must look at a wider berth of options.

2.6   Network Reciprocity

Nowak - in his paper The Five Rules of Cooperation [24] - identified and compared 5 key mechanisms that can aid in the evolution of cooperation, 2 of which I have already discussed (direct reciprocity in section 2.5 and kin selection in 2.3). The other 3 are network reciprocity (which I shall be examining here), group selection and indirect reciprocity (both of which have their own sections).

Network reciprocity uses a graph of players and their connections. The players are the nodes in the graph and the arcs represent connections between players. This ties closely to the networks that IAs may work across. Players with arcs between them interact with each other in rounds of The Prisoner's Dilemma. Nowak and May's [25] ealier work - which inspired Nowak's later paper [24] - did not give individual's any memory of past interaction.

The lack of memory limited Nowak and May to pure cooperators and pure defectors. In Nowak's book 'Evolutionary Dynamics' [23] his exploration of evolutionary graph theory and spatial games (chapters 8 and 9) showed that the shapes of the lattice linking the players and different concentrations of cooperators and defectors on those shapes has a great effect on the evolution of cooperation. Visualised in figures 2.6 and 2.7.
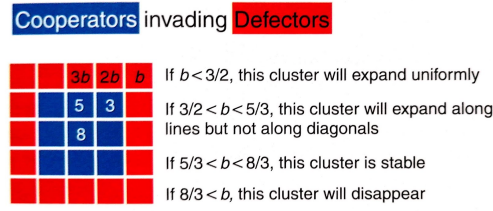
Figure 2.6: How can cooperators invade defectors? Taken from Nowak's book Evolutionary Dynamics [23]. The squares represent nodes and the players interact with the players to each side of them and diagonally. The value b is from the payoff matrix in table 2.2

Nowak's [24, 23] and Nowak and May's [25] work on these games on graphs is limited in terms of it's strategies and in terms of the fixed shape of it's graphs. But the work proves a key point: the structure of who interacts with whom can play a key role in supporting cooperation in large populations.

In real life, individuals will often mostly interact in their close social circles. An example is a Meerkat's life, in which they may interact with others of their family group, a drongo bird which calls to warn of predators, the predators and other's who are geographically close to them. The graph in this case represents the close geographic ties.

I can imagine the use of Nowak and Nowak and May's work to employ a network not as a representation of a physical network structure or geography as I first thought, but as a representation of the choices IAs make on who they wish to interact with.

This would be a constantly changing and adapting network of IAs. The IAs would not concern themselves with the strategy they employ towards who they are forced to interact with. Instead, their strategy is to select carefully who they wish to interact with, effectively constructing a graph of network reciprocation. How these changing graph connections would effect cooperation is unbeknownst to me, whether Nowak's rules would still apply would be interesting to find out.

Nowak [23] found some shapes supported cooperators in groups, cooperators could make use of these shapes by deliberately forming them to protect each other. While another set of shapes were found as 'amplifiers' for evolution, maybe defectors could make use of these sorts of shapes to invade groups of cooperators.

I can see IAs having strategies as to how to build these shapes. I see an issue that may arise where cooperative agents find it hard to reach out to other cooperators not part of their current shape. This could possibly prevent the spread of cooperation, and limit it to these groups. This is however worth investigating, and possibly could involve some kind of bridging mechanism.

	A	B
A	a	b
B	c	d

Table 2.2: The payoff matrix for what happens when individuals interact. Cooperators are in blue and are called A and defectors in red and called B. Taken from Nowak's book Evolutionary Dynamics [23].
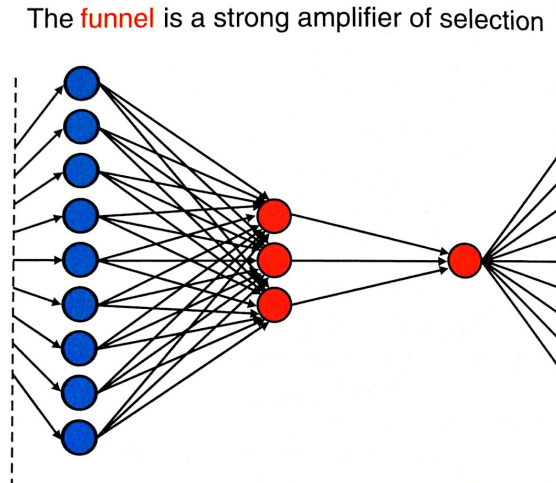
Figure 2.7: Shapes that can amplify selection include the funnel, the star and the super-star [23]

2.7   Group Selection

Group selection is another mechanism decribed by Nowak [24]. This mechanism splits one population into multiple groups. Within these groups The Prisoner's Dilemma is played and reproduction occurs proportional to each players payoff. If using the payoff matrix in table 2.1 cooperators can work together to produce a payoff of 3, while defectors can only produce 5 or 2 for both players in the interaction.

The group size increases until a certain point where the group may split. If the group does split (which is stochastically chosen) then another group is destroyed. The effect is multi-level selection.

Nowak found that due to the higher payoff between cooperators they will reproduce faster in the groups they dominate, than groups filled with defectors. The faster the reproduction, the quicker the group size grows, making it more likely for groups of cooperators to survive while groups of defectors will shrink and be destroyed. These dynamics are displayed in figure 2.8. Traulsen and Nowak [40] limited themselves to cooperators and defectors, but noted that other strategies could be built into their model. A limitation to applying group selection to MASs is the groups themselves. In this model individuals do not interact with individuals in other groups, creating a barrier between them.

Even our current networks span the globe and don't always have harsh barriers between them - even when security is high these barriers can often be broken. Furthermore, IAs are often built to be of service to others, limiting them to only be of service to one group greatly reduces the service an agent can supply and limits the society as a whole . Finally, the group mechanics of splitting and destroying another group of individuals does not naturally match the paradigm of MASs.

I would suggest that this is not a mechanism which would be useful to apply to MASs unless you are modelling clusters of agents, with each cluster in competition. This is not the aim of my project, but could be another interesting project to take up.

2.8   Indirect Reciprocity

The final mechanism I will be discussing from Nowak's Five Rules of Cooperation [24] is indirect reciprocity. This is promising for use in MASs as it solves direct reciprocity's failure
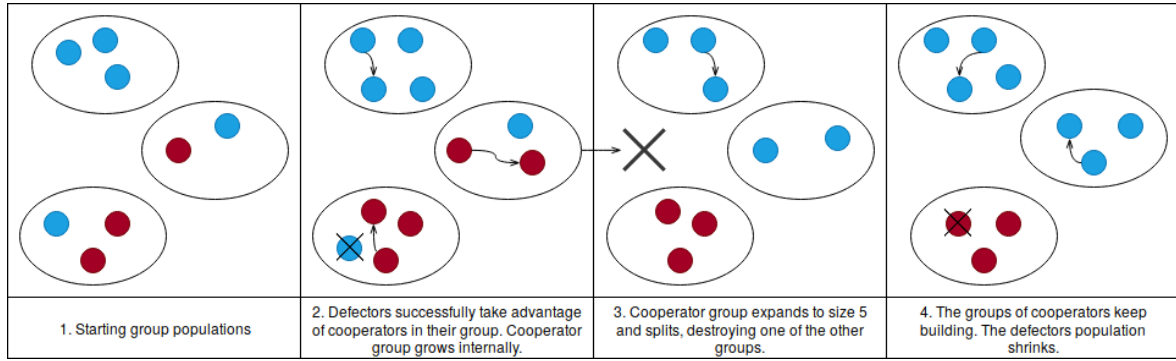
Figure 2.8: The dynamics of multi-level selection as described by Traulsen and Nowak [40].
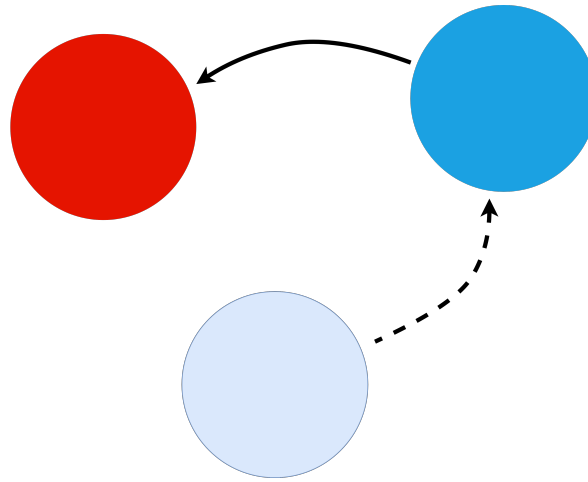

Figure 2.9: The main idea of indirect reciprocity from Nowak [24]

when re-meeting is low. In a large society or a vast network which agents are likely to operate in this is an important property. We will see later on that the two mechanisms can be combined.

Indirect reciprocity uses the group mechanic of reputation to encourage cooperation. Alexander [1] who was an early advocate the idea focused on human reciprocal altruism, but, subsequent research has abstracted away from the biology [28, 18, 26, 30, 24, 15, 38, 35, 21]. The idea is that if an individual cooperates with someone else then their reputation will be enhanced in the community. This boost of reputation makes it more likely that they will be helped by others later on. Making the mechanism a form of reciprocal altruism.

According to Nowak and Sigmund [26] the reputation mechanic requires a higher level of intelligence than direct reciprocity, due to the complexity of group mechanics in the system. It is this sort of higher level intelligence required to reason about events in a group that could be a key part in the development of IAs.

Due to this I feel that either indirect reciprocity or possibly a combination of both direct and indirect reciprocity is a good candidate for a mechanism to study further. For the rest of this literature review I will consider past approaches to indirect reciprocity and consider other factors in the mechanism that can be used to facilitate the evolution of cooperation.

2.9   Nowak and Sigmund

According to Gilbert Roberts [30] the most influential model on indirect reciprocity Nowak and Sigmund's [26] so I shall examine this model of indirect reciprocity first. Nowak and Sig-

mund begin by stipulating that human cooperation is due to people's 'image' of each other, which is comparable to reputation. Nowak and Sigmund converted the image to an integer score between -5 and 5 for simplicity.

The idea is simple, cooperation increases your image score by 1 and defection reduces it by 1. The higher your image score the more likely it is you will receive help. Nowak and Sigmund claim this channels cooperation to valuable members of the society of players.

Nowak and Sigmund note that the use of image scores and indirect reciprocity itself leaves a system open to anticipation, planning, deception and manipulation. These 4 concepts seem closely related to possible happenings in MASs. Deception and manipulation are two factors that I think are worth experimenting with. Whilst anticipation and planning are a key part of agent design in multiple languages and frameworks. To highlight this point it was such an important drawback of the Agent0 language [32] Becky Thomas created a new language - PLACA [39] - which allowed agents to plan among other improvements.

The framework created by Nowak and Sigmund is simplified from Alexander's [1] idea of human reciprocal altruism. They describe a framework where there is a population of individuals which acts as a pool to select pairs in which one player is the donor - who can choose whether to cooperate or defect - and the other is the recipient of this action. A cooperation costs the donor c to it's fitness and benefits the recipient's fitness the value of b where $b > c$. Whereas a defection costs nothing and the recipient is not benefitted. This is denoted in the payoff matrix in table 2.3.

Donor Action	Payoffs	
	Donor	Recipient
Cooperation	-1	2
Defection	0	0

Table 2.3: The payoff for Nowak and Sigmund's [2] indirect reciprocity model

As noted above these actions also affect the donor's image score, but Nowak and Sigmund add a caveat when using the idea of onlookers. A specified size group is randomly selected from the population to view an interaction limiting the spread of reputation information. This is displayed graphically in figure 2.10. The onlookers concept was added due to Nowak and Sigmund's realisation that in a group that is spread far flung geographically, not all individual's will be able to view each interaction. This is of course especially likely in MASs. Image scores now become one players view of another rather than a community view of the player. A matrix $ImageScore$ is used to store these scores.

The discriminator is the strategy of choice for Nowak and Sigmund. This strategy stores a number k, and when the individual u using that strategy is a donor to the individual v, u cooperates if $ImageScore[u, v] >= k$ else it defects. This strategy is incredibly simple yet effective, and is displayed graphically in figure 2.11. The model also includes the defector and cooperator strategies. Nowak and Sigmund detailed more strategies. These strategies base their decisions not only on that of the recipient's image score but also their own. This way the individual will be able to decide whether it is important to boost their reputation in the system in order to receive cooperation or not.

Nowak and Sigmund hypothesize and give evidence supporting that the length of the generation is important to the evolution of cooperation. They highlight that when m donor-recipient pairs are selected in succession in a population of size n a player is likely to be selected as part of a pair $2m/n$ times. According to their evidence the higher this value of $2m/n$ the more likely cooperation will evolve.

As previously noted Herbert Spencer coined the phrase "Survival of the fittest" [36]. This is a principle put to work by Nowak and Sigmund in their reproduction mechanism. The higher the fitness of an individual the more likely it is they will reproduce into the next generation. They also include a chance for random mutation in reproduction.

Another idea put forward and evidenced by Nowak and Sigmund is that the evolution of cooperation is dependent upon the ability of the donors knowing the image score of the recipient. This chance of knowing an image score correctly is limited by the concept of onlookers, but, in Nowak's 2005 paper on the five rules of cooperation [24] he talks about using 'gossip' as an alternative to direct observation. Gossip and social ability is an interesting concept in MASs. According to Wooldridge and Jennings [43] social ability is a key property in even a weak notion agents.

Nowak and Sigmund stress that discriminators are not tit-for-tat players as used in direct reciprocity, because they use the experience of others. This is true however the similarity lies in the way that discriminators punish those with lower image scores due of their bad actions, but are forgiving to those who improve and cooperative to those who they consider good. The strategy can also be considered 'nice' if it is using a value of $k <= 1$ as it will cooperate even when the recipient has no past action history.

The model laid out by Nowak and Sigmund is succinct and a good basis for looking at interactions in multi-agent systems. However, there are limitations to their approach in the context of MASs. The first limitation I will highlight is the way in which reputation is implemented. In the version without onlookers each player has a global image score and even in the version with onlookers there is a public matrix encoding all image scores.

Image scores could be seen as a community view on an individual, but I argue that the idea image scores are attempting to capture (reputation) is actually more personal. Though there might be a rough consensus as to an individuals' general reputation among people in a society, this reputation is often conveyed through social means and is subject to the personal interpretation of each individual.

According to Russell and Norvig [31] IAs use an internal state, in the BDI model [29] the internal state is the beliefs of an agent. It is in this internal state or a beliefs like system, I argue, that image scores for those an agent has encountered should be stored where an agent has full control over them. Multiple trust models could be devised to manage these image scores. This is a closer match to the agent paradigm and it would allow for a truly distributed system.

Indirect reciprocity makes a good candidate for use in MASs and this model is both very influential and a good basis for those MASs. The model includes a lot of interesting aspects that have been examined carefully including the onlookers and reproduction mechanisms, and also ideas that have not been fully researched such as a gossip systems and using deception and manipulation to affect the system. These features allow more room for me to explore if I use Nowak and Sigmund's model as a basis, with a few alterations to more closely match the MAS paradigm.

2.10   The Standing Strategy and Further Limitations of Nowak And Sigmund

Leimar and Hammerstein [15] criticised Nowak and Sigmund's [26] limited range of strategies. They also higlighted the abilities of another strategy, the standing strategy, over image scores. This was first described by Sugden [38] and and uses a 'switch' [26].

Milinski *et. al* [18] described the standing strategy as the idea is that an individual does not only aim for a good fitness but also a good standing. Each individual holds other individuals in either a good or bad standing, which starts out as good until perceived otherwise. The
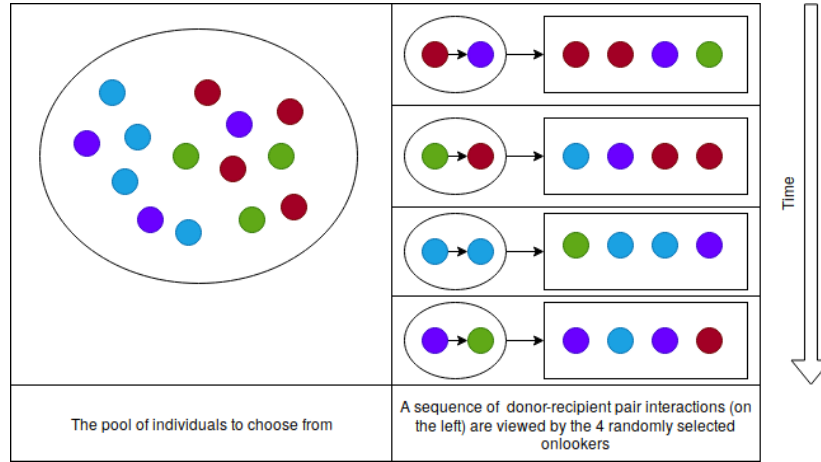
Figure 2.10: The selection of a sequence of donor-recipient pairs and onlookers in Nowak and Sigmund's [26] model of indirect reciprocity. The macro view of indirect reciprocity. The lines and arrows show the spread of information.
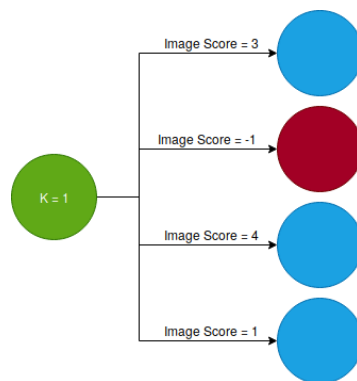
Figure 2.11: On the left an image scoring discriminator with K=1. On the right for agents the discriminator is a donor for. Blue agents represent where there would be a cooperation, read when the discriminator defects against them. The rule is when $ImageScore >= K$ cooperate else defect.

event that causes the switch from good to bad standing is an individual defecting against another with a perceived good standing.

The idea is that it is morally incorrect to defect against a good individual, but defecting against a bad individual is acceptable as it punishes those members of society that are not valuable. There is conflict between the use of image scores and the standing strategy.

The argument for the standing strategy is as follows. When using image scoring if a discriminator has $k = 0$ and is a donor to another with an image score of -1, they will defect. This is seen as punishing a bad individual, channelling fitness away from bad members of the society. But the donor's image score is reduced by 1. This decreases the chance that the donor will receive cooperation from others, so there is no real incentive for them to punish the bad members of that society, except for out of pure altruism.

Conversely, the standing strategy only reduces the reputation of those who defect against good inidividuals. So punishing bad members of the society has no negative effect on them. Leimar and Hammerstein [15] argue that Sugden's version of indirect reciprocity and his use of the standing strategy are more robust than Nowak and Sigmund's image scoring.

Milinski *et. al* [18] analysed Leimar and Hammerstein's argument, regarding that they were correct under conditions which allowed for perfect perception of events and unlimited memory capacity. This is not always true, especially in a MAS where the environment is inaccessible [31].

Nowak and Sigmund claimed that indirect reciprocity is open to deception and manipulation, especially when we consider the use of gossip in a system. This opens IAs to imperfect perception of other IAs. So it would be interesting to see which strategy is more effective in a system where other IAs actively attempt to deceive using gossip, and the information received is not always perfect.

2.11   Mixed Reciprocity Models

Roberts [30] and Phelps [28] noted that indirect reciprocity is focused towards interactions with other individuals that the donor has no previous interactions with. But what about when re-meeting is more likely? Surely an individual is affected by both being the recipient of actions and observation of actions. Roberts tackled the issue by introducing an experience score. This experience score works similar to an image score, but is bounded by -1 and 1, it increases when the individual is a recipient of a cooperation, and decreases when receiving a defection.

Roberts also created a version of the standing strategy which uses an image score (-1 to 1), but the value changes according to the rules of the standing strategy rather than Nowak and Sigmund's rules.

Both Roberts and Phelps measured the use of indirect reciprocity vs the use of direct reciprocity. Roberts concludes that indirect is the more popular decision mode under conditions when re-meeting was less common, and direct was more common when re-meeting was frequent. Whereas, Phelps' experiments garnered different results being that in small groups direct and indirect reciprocity exist in equilibrium.

Just like Nowak and Sigmund [26], Roberts I believe falls short in recognising how personal interpretation of events are. There are many different trust models a player could use by mixing interpretation of events where they are the recipient and when they are not the recipient. For example it could be expected that being on the receiving end of a defection a player is likely to be more hardline, than when they observe a defection against another. Whereas Roberts limits the study to using a simple experience score.

I would suggest a model similar to that I described in section 2.9 in which interpretation of events is up to an agents trust model. Some trust models could interpret events when the agent is the recipient different to those in which they are observing.
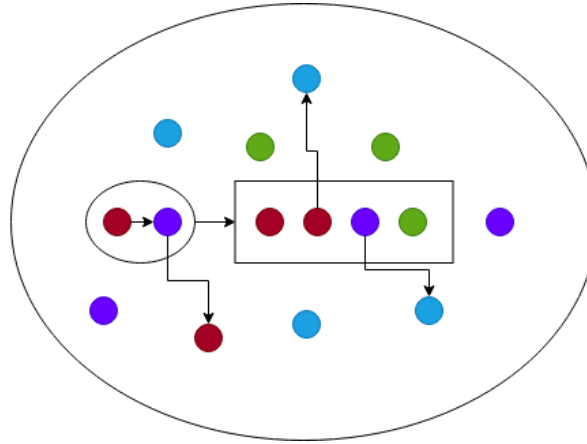
Figure 2.12: The spread of information (shown through the lines) through a population using indirect reciprocity with gossip and onlookers

2.12   Gossip

Nowak and Sigmund [26] realised the issues surrounding all individuals being able to view interactions in their simulation framework. So they introduced the idea of onlookers as noted in section 2.9. An issue with limiting observation is that knowing an accurate image score of a player is important to the evolution of cooperation in an indirect reciprocity system. Nowak [24] suggested using gossip to spread information regarding image scores in their simulation framework.

Sommerfeld *et al.* [35] conducted an empirical study on gossip between cooperative and uncooperative individuals in an indirect reciprocity setup. The experiment consisted of humans playing a number of indirect reciprocity rounds to build up a cooperation history, and then a smaller number of gossip rounds. These rounds are then repeated to see how the individuals reacted to the gossip.

The focus of the experiment was to look at gossip composition, gossip transfer and the resulting behaviour. The findings were that gossip is an effective method of spreading reputation information in an indirect reciprocity system, on fulfillment of some conditions. The first of which is truthfulness of the gossip. Gossip must accurately reflect the behaviour of the subject of the gossip.

Another condition is the comprehensibility of the gossip. The language used in the gossip must be interpretable by the recipient of the gossip. In an agent system this would require an agent communication language (ACL) (such as KQML [7] or FIPA-ACL [27]) and content of the messages that use this language to be interprettable by the recipient agent. The final condition is that individuals must act accordingly based on the gossip they receive.

These three conditions of a system would have to be created by building strategies for agents that attempt to uphold them, and a communication language to support accurate and interpretable information. Nowak and Sigmund [26] identified that indirect reciprocity systems can be effected by deception and manipulation from malicious players. The addition of gossip creates a meta-game on top of the indirect reciprocity system. Players must have a trust model that differentiates between dealing with gossip and with directly observed interactions, and a strategy on spreading gossip effectively.

2.13   Mui's Computational Models of Trust and Reputation

Mui [21] presented an indirect reciprocity simulation framework similar to Nowak and Sigmund [26] that spread social information through 'acquaintance networks' to inform a donor's decision. An individual in this framework builds up a network of players' they meet through interaction, their 'acquaintance network'. If this individual is then a donor in a donor-recipient pairing they consult their acquaintance network about the recipient. The information gathered from this network helps a donor decide if they can trust the recipient to reciprocate their cooperation, and is known as 'collective memory'.

Baumeister *et al.* [3] advocated for the idea that gossip is used for 4 functions. These 4 are strengthening bonds between the gossiper and recipient, enabling the spread of information about the subject for positive and negative reasons, and to help educate individuals about the complex cultural systems they reside in.

I would argue that indirect reciprocity systems and MASs are complex cultural systems in which gossip can be applied for these 4 functions. Mui's acquaintance networks support this gossip functionality to a certain extent but are limited by the lack of proactivity of the gossip. Sommerfeld *et al.* [35] highlighted the willingness of their participants to spread gossip and the last function of gossip proposed by Baumeister *et al.* [3] (cultural education) suggest that for gossip to be most effective the gossip should be proactively spread.

As a result of this I believe a theoretical framework that uses gossip should allow a proactive social ability by using gossip as an action at times when an agent is not acting as a donor.

2.14   Summary

From the information I have lain out in the previous sections in chapter 2 I have become aware of a number of options that could be used to facilitate the evolution of cooperation in MASs. I hold the two most pertinent options of these to be network reciprocity and a combination of indirect and direct reciprocity.

Using network reciprocity where agents choose who they interact with, forming social networks and circles would be extremely interesting. It would also be a good simulation of agents working across a network selecting what other agents they wish to interact with to solve problems.

However these agents may have to work with others they have no experience with. Indirect reciprocity seems a very good mechanism for helping decide on these issues. As such I will be researching further into the trust models and effects on the evolution of cooperation that indirect reciprocity has.

My model will be inspired by that of Nowak and Sigmund's [26] influential model. I will be using concepts including onlookers, reproduction based on fitness scores and sequential selection of donor-recipient pairs from a pool of players. However I will also be looking at how gossip, deception and manipulation affects the system.

In some systems re-meeting will be likely, especially in the close circles I anticipate a network reciprocity system would have. Due to this I will be looking at using a mixed reciprocity model. To do this I will not be implementing any image score matrix into the system, but will allow agents to hold their interpretation of other agents in their internal state.

Agents will be able to use multiple different strategies. This includes the standing strategy [38] and Nowak and Sigmund's discriminator [26]. It will be interesting to see which performs better under situations where the spread of information can be imperfect.

Strategies will have different trust models that can interpret both information from being a recipient of an action, direct observation and from social ability used to convey reputation

information. It will be important to find trust models that can deal well with deception, whilst still being open enough to allow trustworthy gossip to affect their decisions. A key aspect of study for me will be social ability and the effect of gossip and the interpretation of gossip on the evolution of cooperation. The system I devise will use gossip proactively.

As pointed out by Leimar and Hammerstein it is also important to differentiate between uncooperative actions to bad and to good agents. As such a good trust model may act more harshly against those who are uncooperative when they are the recipient, as they can guarantee themself as having good intentions.

A subsequent project or study could use the version of network reciprocity I have laid out. The new study's model could use the trust models and information about them this project has revealed to deal with handling players that have not been met before. Furthermore, this new study could also use the information revealed about using social ability to spread reputation information.

Chapter 3: **Framework**

Excellent understanding and insight. Conceptual framework underpins study. Comprehensive expert account of topic. Well thought through Software Engineering content.

3.1   Introduction

Relate to background.
Objective 4.
Look at summary of Lit review + end of each section of lit review

3.2   Theoretical

3.2.1   Introduction

Relate to background.
Objective 4.
Look at Kostas agents course information and moodle page links

3.2.2   Environment

There are many different definitions of agency [10]. Russell and Norvig [31] gave a wide ranging definition of an agent as being situated in an environment, which the agent can perceive through various sensors and act in using actuators. Their notion of an agent is visually described in figure 2.1. Wooldridge [42] describes an agent as able to work out the actions required to meet it's design objectives.

These two different definitions cover the two aspects of an agent: the physical and the mental. An agent's design objectives will work within the bounds of an environment, so before we discuss any agents in this system I will discuss the environment it shall be working in.

An instance of the environment in my system will be known as a community. A community contains a number of generations, which each contain a set of agents. These sets of agents will act as a pool to select donor-recipient pairs and onlookers for those pairs from.

The environment has distinct timepoints, which varies depending on how the number and length of generations variables are set. Say these range from $1..n$ and there are generations $1..k$ where $k < n$ and $n\%k = 0$. The generations occur in an order, and the range of timepoints are distributed evenly across each generation like this:

$$\{1..(n/k),\ (n/k+1)..(2n/k),\ ..,\ (n-(n/k)+1)..n\}$$

The environment will work using a perceive-decide-execute cycle, much like Russell and Norvig's [31] episodes. Each step of the cycle is a timepoint. Agents will receive perceptions at the start of each cycle step, these may come from the environment themself or from the actions of agents the previous cycle step. The steps are synchronised, this prevents actions from a timepoint affecting other agents decisions at the same timepoint. Synchronicity keeps the environment static for the period in which an agent is deciding.

In each timepoint there is exactly one donor-recipient pair and for that pair there is a group of onlookers randomly selected from the rest of the generations pool of players. The onlookers

and recipient receive a percept in the next cycle step of what they observed/received in that interaction.

Percepts (discussed in subsection 3.2.3) are sent to agents at the beginning of each cycle step, the agents will then be asked to decide on an action (subsection 3.2.4) and these actions shall then be executed (subsection 3.2.5) in the system.

As discussed each generation contains a set of players which participate in the cycle steps for each timepoint of that generation. But how are these sets of players selected? For the first generation a number of agents and associate strategies (subsection 3.2.9) with them.

One of the aims of this project is to discover successful agent strategies for the mixed reciprocity model. An optimisation problem is where we have a set of solutions (strategies) and want to find the most effective one [9]. Francq [9] puts forward that genetic algorithms (a type of meta-heuristic) are superior to most traditional heuristics. Natural selection is the basis for genetic algorithms and contains 5 steps: reproduction, crossover, mutation, inversion and evaluation. Due to the nature of agents, we are only specifically interested in reproduction and mutation.

Just like in The Iterated Prisoner's Dilemma I will use a genetic algorithm (GA) to discover the most effective solution for the given parameters [20]. How my system will work as a GA is outline in subsection 3.2.7.

Russell and Norvig outline 5 different properties. I conclude that the environment I have delineated have the following properties.

Inaccessibility or partial observability, as the sensors of an agent cannot detect all aspects relevant to an action choice. They cannot know the intentions of other agents, which is key to deciding on whether an agent should cooperate or not. Agents can also not view all interactions, to inform their decisions, using the onlookers mechanic. However, lots about the environment will be observable. For example, agents will not need to 'discover' each other they will have knowledge of all other agents in the generation.

I also argue that the system is deterministic from a generational point of view. Milinski *et al.* [18] claimed that individuals using the standing strategy aim for a good standing, and Leimar and Hammerstein [15] argue that the standing strategy is good as it allows individuals to punish bad individuals. However take 3 agents a, b, c and d. Agent b has defected previously against d who has a good standing according to a. Agent a then chooses to punish b but is still believing this won't reduce their standing. However, agent c did not observe b's defection against d but is an onlooker for c's defection against b.

From the proceedings the environment appears to be non-deterministic, as there is more than one outcome to an action. However, this is due to the partial observability of the environment not the determinism, as highlighted is possible by Russell and Norvig [31]. The current state (c's lack of knowledge on b) and the actions selected by a and b completely determine the next state of the system.

However, from a community point of view the environment can be seen as non-deterministic due to the reproduction mechanism outline in subsection 3.2.7. This includes a chance of mutation and the actual selection of agents is stochastic. As such the actions of the agents cannot guarantee that their strategy will be highly propagated, the higher the fitness they get (which is decided by state and actions) affects their chances greatly but not fully.

The environment is also nonepisodic. To take a proof by contradiction approach suppose the environment is episodic. According to Russell and Norvig [31] an environment is episodic if subsequent episodes do not depend on what action occurs in previous episodes. Episodes consist of perceiving and then deciding. But from the percepts and actions subsections (3.2.3 and 3.2.4) we know that actions from one timepoint can generate percepts in the next. This is a contradiction and as such the environment is nonepisodic.

Due to the synchronisation in the cycle steps the environment is static. Russell and Norvig [31] define static environments as those that do not change while an agent is deliberating. All agents in the environment decide at the same time, and only after this are these actions executed, ensuring the environment does not change while decisions are being made.

According to Russell and Norvig [31] an environment is discrete if it has clearly defined

actions and percepts. I have defined these in subsections 3.2.4 and 3.2.3, and as such the environment is discrete.

In summary the environment I have outlined in this section is inaccessible, deterministic from a generational point of view yet non-deterministic from a community point of view, nonepisodic, static and discrete. The components of this environment are the community, generations, the sets of agents within the generations, the timepoints throughout the community's life, the onlooker mechanism, the reproduction mechanism, the cycle steps, the donor-recipient pairs, the percepts and the actions. There is a further description of many of these components in following subsections.

3.2.3   Percepts

There are many definitions of an agent [10] but most consider the agent to have some kind of sensing ability. In Russell and Norvig's [31] definition of an agent, agents use sensors to perceive their surroundings. What an agent receives from these sensors are percepts. This is actually physically what the agent is viewing or sensing about their environment. Perception is the first stage in each cycle step.

In my system percepts include sensing being the donor or recipient in a donor-recipient pair, a direct observation of an interaction or hearing gossip from another agent.

In each timepoint there is a donor and a recipient selected at random from that generations pool of players. The two agents are made aware of this fact by receiving a percept of the role they are taking in that cycle step's perceive stage, and the other agent in the pair. Agents can then act accordingly.

For each interaction there is a set of onlookers selected at random from the generations pool of players (not including the recipient or donor). These onlookers and the recipient get percepts the cycle step after that interaction of containing who the donor and recipient were and the action the donor decided on.

In the gossip (3.2.6) and action (3.2.4) subsections I discuss an agents ability to act by gossiping to another agent. The action gossip produces a percept. This percept contains the information from the gossip and the agent given as the recipient by the gossiper perceives it.

3.2.4   Actions

Wooldridge [42] elects that agents have design objectives. The agents have the responsibility of figuring out the actions they need to get to their objective. This requires a decision making process which occurs in the second stage of each cycle step. My project is focused on studying interactions between agents, and as such agents require a number of action possibilities to interact with each other.

To simplify the action there are 3 possible main actions: idle actions, actions when an agent is a donor and gossip actions. The first of which is simple, an agent is idle in that timepoint, there action has no effect on the environment or other agents except for through inaction.

The second is slightly more complex. As discussed in subsection 3.2.3 an agent perceives when they are a donor in a donor-recipient pair. When an agent perceives they are a donor they have no choice but to commit one of the following two actions: cooperate or defect. This is an interaction where the actor is the donor and the recipient has no control over what happens in the interaction. The effects of both possible actions as a donor are described in subsection 3.2.5.

When an agent is not a donor they cannot choose to cooperate or defect with anyone, they can choose one of the other two actions: idle or gossip. A gossip action is another type of interaction between agents. A gossiper chooses to communicate with another agent. The contents and structure of this communication is detailed in subsection 3.2.6 and the effects

in subsection 3.2.5.

3.2.5   Action Execution

Another general consensus surrounding agent definition [10] is that agents commit actions in their environment by interacting with the environment and/or other agents. As such the agents who decide on the actions listed above need to execute their actions.

An idle action has no effects, and is effectively inaction so requires no execution. A gossip action is communicated from the gossiper to the recipient agent by in action execution generating a percept, which in the next cycle steps perceive stage is perceived by the recipient. An agent is then able to interpret this piece of gossip how they wish and change their internal state accordingly.

A more complex action to execute is an action a donor takes. I am using the payoff matrix created by Nowak and Sigmund [26] from table 2.3. Each player has a fitness score which begins at zero and cannot go below this number. This denotes when a defect action is chosen there is no effect on either the donor's or recipient's fitness. When a cooperate action is chosen however, this is at a cost of 1 to the donor's fitness and a benefit of 2 to the recipient's fitness.

The effect on the fitness of each player is important but also who views the action is import. In order for observers to interpret events and gain an idea of the donor's character, they must receive a percept of what happened. A percept is generated for the recipient and all the onlookers of an action detailing who the donor-recipient pair are and what action was decided upon by the donor.

This is effectively what Nowak and Sigmund's model [26] does when changing image scores. Instead of the environment directly meddling in an agents view, however, the sending of a percept allows an agent to interpret the percept according to their strategy and trust model. This process gives an agent full control over their internal state.

3.2.6   Gossip

Singh [34] suggests that for a system to be agent oriented it needs to be autonomous and interoperable with other agents. Singh criticises current agent communication languages (ACLs) including KQML for being focused on mental agency. Mental agency Singh goes on to say supposes that agents can read each others' mind.

Russell and Norvig [31] stipulate a number of architectures which keep internal states, whilst many others such as in the BDI model contain similar. A general idea for agents is that you can request of them information, but the agent can decide whether or not to reveal this information. As such it is not necessarily true one agent can know the internal state of another.

My communication system will thus not be based on mental agency but on what Singh [34] social agency. To do this I will focus the gossip specifically on the events that agents are gossiping about: the donor-recipient interactions. So an ACL I define could simply allow agents to reconvey the details of a specific event. However gossip usually comes with an individuals slant, and is generally not a simple description of an event.

There are two of four functions described by Baumeister *et al.* [3] that apply greatly to my system: enabling the recipient to learn more about the target and to harm the target. Another function highlighted, cultural learning, goes beyond the scope of my project.

Sommerfeld *et al.* [35] used short statements written by humans as gossip. This gossip was categorised in the experiment as either positive or negative, as the individuals were either trying to convey either good or bad reputation information. I shall thus use these categories for simplicity. These two categories allow the two functions from Baumeister *et al.* [3] to be fulfilled. It will also allow agents to put forward positive reputation information about

themself, to attempt to strengthen bonds between them and other agents (the final function from Baumeister *et al.* [3]).

The language I propose is very simple. It contains five fields that I have already identified here, three of the fields are identifiers (one for each of the recipient, the target and the gossiper). Another is the timepoint at which the gossip action was executed. The final field is the gossip itself (either positive or negative).

As my agents have no other purpose than playing this game they only need to be interoperable with each other, and as such can have their own domain specific language. I have called this Simple Agent Gossip Language (SAGL).

I have decided against the use of KQML [7] due to Singh's [34] criticism of it and it's complexity when all I simply need is these 5 fields. The same complexity argument can be given for FIPA-ACL [27]. Both are very powerful languages, however that power is not needed here. If I were to later want a more expressive communication system I would consider switching to FIPA-ACL as SAGL could easily be translated.

3.2.7   Reproduction

As discussed in section 3.2.2 finding the best solution from a set of possible solutions is an optimisation problem. Genetic algorithms are an approach to solve optimisation problems, and as such my system will be acting like a genetic algorithm with simulate reproduction.

Between each generation a reproduction phase occurs, in which agents are created and associated with strategies for the next generation. Nowak and Sigmund [26] and Axelrod and Hamilton [2] (though indirectly their payoff scores do seem to equate to fitness) used a reproduction mechanism inspired by Herbert Spencer's [36] famous phrase "Survival of the fittest". Wooldridge's [42] argument that agents attempt to fulfil a design objective using their own autonomy suggests that strategies for agents will only be used in practice if they are successful in their design objective. The main design objective of players in my game is to interact with players to increase their fitness, with some having other big objectives such as aiding in the evolution of cooperation.

The measurement of how successful an agent is in my system is thus their fitness score. Just like in Nowak and Sigmund's, and Axelrod and Hamilton's reproduction mechanism I propose using a "Survival of the fittest approach". The real replicated information in my system is the strategy an agent uses, so the mechanism I am using makes it more likely for strategies with higher overall fitness scores across agents in a generation to be reproduced.

Francq [8] puts forward two types of reproduction: proportional selection and tournament selection.

Proportional selection translated to my system and simplified would look something like the following. Take a roulette with p (number of players) slots and divide the p slots into n (number of strategies) sections. The size of each section is directly proportionate to the average fitness of all players of a certain strategy. For example strategy A has 2 agents with fitness 4 and 6 respectively, strategy B has 4 agents with fitness 3, 8, 9 and 4 fitness respectively and strategy C has 1 agent with fitness 7. The average fitness of A is 5, B is 6 and C is 7, so A would receive 5 slots of the roulette wheel, B 6 slots and C 7 slots. The resulting roulette wheel displayed in 3.1 is spun and a ball dropped and whichever slot the ball landed on the corresponding strategy is selected for the new player.

  Tournament selection could be translated in the following way. Have an empty list which will contain an ordered list of agents. In a loop (until the population is of size 1): select two agents from the population, remove the lowest fitness one and insert it into the list. After the loop insert the last agent into the top of the list. The top of the list always contains the highest fitness agent. Using the same population of agents in the proportional selection algorithm we can give an example of this in table 3.1.
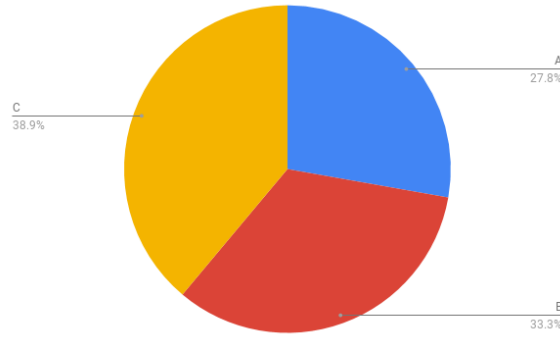
Figure 3.1: The roulette wheel from my proportional selection example

Set	Test	Put in List	List
$A_1, A_2, B_1, B_2, B_3, B_4, C_1$	(A_1, B_3)	A_1	A_1
A_2, B_2, B_3, B_4, C_1	(B_1, C_1)	B_1	A_1, B_1
B_2, B_3, B_4, C_1	(A_2, C_1)	A_2	A_1, B_1, A_2
B_2, B_3, C_1	(B_4, C_1)	B_4	A_1, B_1, A_2, B_4
B_3, C_1	(B_2, B_3)	B_2	A_1, B_1, A_2, B_4, B_2
B_3	(B_3, C_1)	C_1	$A_1, B_1, A_2, B_4, B_2, C_1$
$\emptyset$		B_3	$A_1, B_1, A_2, B_4, B_2, C_1, B_3$

Table 3.1: The payoff for my indirect reciprocity model

Fitness proportionate selection seems to translate best to my system as there is an obvious and simple way to select each agents strategy for the new generation using the roulette wheel, unlike the tournament selection process which requires the crossover step. The crossover step requires a chromosome represented as a bit array to produce an offspring from two parents. This is analogous to sexual reproduction, but not the building of new agents.

Lipowski *et al.* [16] presented an efficient version of roulette wheel selection using stochastic acceptance rather than the searching method used by Francq [8]. The stochastic acceptance algorithm works like this: select randomly one of the individuals from the last generations population, with fitness of the individual as w_i and w_{max} the maximal fitness of the generation use the probability w_i/w_{max} to select whether or not to use this agents strategy for a new generation agent, if not repeat. Lipowski *et al.* proved mathematically that the probability distribution of this method and general roulette wheel selection are the same, and that the stochastic acceptance algorithm is more efficient as well.

Due to these factors I have decided to use the roulette wheel-selection via stochastic acceptance algorithm proposed by Lipowski *et al.* with a user-selectable chance of random mutation for each new agent.

3.2.8   Agents

Description of IAs, from [43] and [31].
Design of agent [37].
Body and mind separate
BDI and internal state

3.2.9   Strategies

What is an agent strategy?
Relate to papers on agent strategies.
Both as a donor and not a donor.
Image scoring discriminator + others using image scores, standing strategy, cooperators and defectors, reinforcement learning? (works for ipd so why not indirect reciprocity? [13]).
Refer to papers using these strategies.
Deceptive/manipulative, naive and trustworthy strategies. Find papers to relate?

3.2.10   Trust Models

Relate to background.
Find papers related to these? Interpretation of events in agents.
How to treat when already met, having directly observed and gossip about.

3.2.11   Summary

Make the key points of the entire theoretical framework.
Link onto implementation.
Have I discussed the changes in my system to those from my background?

3.3   Implementation

Relate to agents systems.
Making it distributable.
Objectives 5 and 6
AMAAS.
Logic framework.
Include software engineering, algorithmic work and development processes. In their own sections.

3.3.1 Agent Mind as a Service

3.3.2 Environment

3.3.3 Web Application and Interface

3.3.4 Conclusion

3.4 Technical Preparation

Learning flask etc.

3.5 Design

3.5.1 Agent Mind as a Service

3.5.2 Environment

3.5.3 Web Application and Interface

3.6 Development

3.6.1 Agent Mind as a Service

3.6.2 Environment

3.6.3 Web Application and Interface

3.7 Testing

3.7.1 Agents Service Testing

3.7.2 Environment Testing

3.7.3 Web Application Testing

3.8 Software Engineering

3.9 Documentation

Software documentation, user guide.

Chapter 4: **Experiment Evaluation**

4.1 Introduction

What experiments will I be doing, why?
Look at the summary of my background for what to experiment with.
Factors that could effect cooperation:

- Number of generations

- Strategy

- Trust Model

- Whether a player knows the reputation of another player (generation size, generation length (2m/n [26]), onlooker number and social activeness)

- Positivity of gossip vs Accuracy of gossip

- Deception and manipulation (image vs standing)

- Sommerfeld's 3 properties

Control variables:

- Number of generations

- Mutation chance

- How many experiment repetitions? (due to non-determinicity of the game)

4.2 Experiment 1

Variable setting, controls etc.
Evaluation

4.3 Experiment 2

4.4 Experiment n

4.5 Discussion

Evaluation and discussion on the experiments
Successful?

Opportunity for extra work?
Focus on experiments

33

Opportunity for extra work?
Focus on experiments

Chapter 5: **Critical Analysis and Discussion**

Project achievements
Reflection on project process (difficulties, success, failure)
Successful?
Future enhancements
Focus on project
Compare to deliverables at beginning.

Chapter 6: **Conclusions**

Mechanism has to work with and without re-meeting.
Mechanism has to encourage cooperation between unknown entities.
Opportunity for extra work? Further exploration of network reciprocity as outline in background and background summary. Extend indirect reciprcoty so individuals can choose who they wish to interact with. Symbolic reinforcement learning strategies.
Move closer to using an agent model such as BDI?
Using the GAIA methodology for development? [44].

Bibliography

[1] Richard Alexander. The biology of moral systems (foundations of human behavior). 1987.

[2] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[3] Roy F Baumeister, Liqing Zhang, and Kathleen D Vohs. Gossip as cultural learning. *Review of general psychology*, 8(2):111, 2004.

[4] Nicky Case. The evolution of trust. `https://ncase.me/trust/`, 2017. Accessed: 11/2/2019.

[5] Richard Dawkins. *The Selfish Gene.* Oxford University Press, 2016.

[6] Vince Knight; Owen Campbell; Marc; eric-s-s; VSN Reddy Janga; James Campbell; Karol M. Langner; T.J. Gaffney; Sourav Singh; Nikoleta; Julie Rymer; Thomas Campbell; Jason Young; MHakem; Geraint Palmer; Kristian Glass; edouardArgenson; Daniel Mancia; Martin Jones; Cameron Davidson-Pilon; alajara; Ranjini Das; Marios Zoulias; Aaron Kratz; Timothy Standen; Paul Slavin; Adam Pohl; Jochen Mller; Georgios Koutsovoulos; Areeb Ahmed. Axelrod: 4.3.0. `http://dx.doi.org/10.5281/zenodo.1405868`, September 2018.

[7] Tim Finin, Richard Fritzson, Don McKay, and Robin McEntire. Kqml as an agent communication language. In *Proceedings of the third international conference on Information and knowledge management*, pages 456–463. ACM, 1994.

[8] Pascal Francq. Genetic algorithms. `http://www.otlet-institute.org/wikics/Genetic_Algorithms.html`, 2011. Accessed: 04/01/2019.

[9] Pascal Francq. Optimisation problems. `http://www.otlet-institute.org/wikics/Optimization_Problems.html`, 2011. Accessed: 04/01/2019.

[10] Stan Franklin and Art Graesser. Is it an agent, or just a program?: A taxonomy for autonomous agents. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 21–35. Springer, 1996.

[11] The Photographers from Mpala Live `http://mpalalive.org/sponsors`. Green wood hoopoe. `http://mpalalive.org/field_guide/green_woodhoopoe`, January 2019.

[12] W. D. Hamilton. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7:1–16, July 1964.

[13] Marc Harper, Vincent Knight, Martin Jones, Georgios Koutsovoulos, Nikoleta E Glynatsi, and Owen Campbell. Reinforcement learning produces dominant strategies for the iterated prisoners dilemma. *PloS one*, 12(12):e0188046, 2017.

[14] Peter Kropotkin. *Mutual aid: A Factor of Evolution.* 1902.

[15] Olof Leimar and Peter Hammerstein. Evolution of cooperation through indirect reciprocity. *Proceedings of The Royal Society*, 268:745–753, May 2001.

[16] Adam Lipowski and Dorota Lipowska. Roulette-wheel selection via stochastic acceptance. *Physica A: Statistical Mechanics and its Applications*, 391(6):2193–2196, 2012.

[17] Computing Machinery. Computing machinery and intelligence-am turing. *Mind*, 59(236):433, 1950.

[18] Theo C. M. Bakker Manfred Milinski, Dirk Semmann and Hans-Jürgen Krambeck. Co-operation through indirect reciprocity: image scoring or standing strategy? *Proceedings of The Royal Society*, 268:2495–2501, May 2001.

[19] Arjun Mani. What is the prisoner's dilemma in layman's terms? what are its applications? what are it's applications? https://www.quora.com/What-is-the-prisoners-dilemma-in-laymans-terms-What-are-its-applications, September 2013.

[20] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.

[21] Lik Mui. *Computational models of trust and reputation: Agents, evolutionary games, and social networks*. PhD thesis, Massachusetts Institute of Technology, 2002.

[22] Martin Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364:56–58, July 1993.

[23] Martin A Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.

[24] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314, 2006.

[25] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

[26] Martin A. Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–577, 1998.

[27] Paul D O'Brien and Richard C Nicol. Fipatowards a standard for software agents. *BT Technology Journal*, 16(3):51–59, 1998.

[28] Steve Phelps. An empirical game-theoretic analysis of the dynamics of cooperation in small groups. *Journal of Artificial Societies and Social Simulation*, 2016.

[29] Anand S Rao, Michael P Georgeff, et al. Bdi agents: from theory to practice. In *ICMAS*, volume 95, pages 312–319, 1995.

[30] Gilbert Roberts. Evolution of direct and indirect reciprocity. *Proceedings of The Royal Society*, 275:173–179, September 2008.

[31] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.

[32] Yoav Shoham. Agent0: A simple agent language and its interpreter. In *AAAI*, volume 91, page 704, 1991.

[33] Yoav Shoham. Agent-oriented programming. *Artificial intelligence*, 60(1):51–92, 1993.

[34] Munindar P Singh. Agent communication languages: Rethinking the principles. *Computer*, 31(12):40–47, 1998.

[35] Ralf D. Sommerfeld, Hans-Jürgen Krambeck, Dirk Semmann, and Manfred Milinski. Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences of the United States of America*, 104:17435–17440, 2007.

[36] Herbert Spencer. The principles of biology (2 vols) london: Williams and norgate. *Google Scholar*, 1864.

[37] Kostas Stathis, Antonis C. Kakas, Wenjin Lu, Neophytos Demetriou, Ulle Endriss, and Andrea Bracciali. *PROSOCS: a platform for programming software agents in computational logic*, pages 523–528. 4 2004.

[38] Robert Sugden et al. *The economics of rights, co-operation and welfare.* Springer.

[39] Sarah Rebecca Thomas. Placa, an agent oriented programming language. 1993.

[40] Arne Traulsen and Martin A. Nowak. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America*, 103:10952–10955, 2006.

[41] Robert L Trivers. The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1):35–57, 1971.

[42] Michael Wooldridge. *An introduction to multiagent systems.* John Wiley & Sons, 2009.

[43] Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(2):115–152, 1995.

[44] Michael Wooldridge, Nicholas R Jennings, and David Kinny. The gaia methodology for agent-oriented analysis and design. *Autonomous Agents and multi-agent systems*, 3(3):285–312, 2000.

[45] Sea Turtle World. Green sea turtle cleaning station. `https://www.seaturtle-world.com/green_sea_turtle_cleaning_station/green_sea_turtle_cleaning_station_600/`.

Chapter 7: **Professional Issues**

Replacement of people in jobs?
AI issues and risks?
Relate to project
Moral philosophy of agent societies

Chapter 8: **Appendix**

Describe the contents of my appendix.