Final Year Project Report

Full Unit - Interim 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

November 27, 2018

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 Where other published and unpublished source materials have been used, these have been acknowledged.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:

Student Name: James King

Date of Submission:

Signature:

Table of Contents

Abstract

Chapter 1: **Aims, Objectives and Literature Survey**

1.1 Project Specification

I will explore indirect reciprocity, by both researching and reporting on Nowak and other authors research on the mechanism, drawing on their research and possibly adding on to the ideas of onlookers and gossip to create a model for implementation. I shall be using the Axelrod-Python library to create a proof of concept that I am able to create a web application that allows users to set up and run games and tournaments. This web application will be hosted on Heroku and created using Flask and Python.

The solution that we are exploring for the interfacing issues between Python and Prolog is to create a web service for the agent's decision-making component. This will have a Prolog server containing what [**?**] refers to as "agent's minds". A mind will represent information on the state of the environment using an efficient version of the Event Calculus reported in [**?**]. Percepts can be added to the agent's mind and based on these percepts an agent will be able to decide on an action and act on it by returning an action to the environment. Separating out the head and body as in Stathis *et al.* 2002 [**?**], though this solution would take this separation further, having the head based in the Prolog service and the body in the environment (Python web application).

The capabilities of the Prolog service are to be demonstrated on the web application, which will run indirect reciprocity tournaments creating the environment that agents reside in and containing the body of the agent. The web application will use the service to input percepts and get actions.

These have been the core parts of the project however, I could extend this to the other mechanisms presented by Nowak [**?**] if I have time. Axelrod-Python has an implementation for network reciprocity - which would be interesting to use so I could create an especially nice frontend for the patterns created in the population [**?**]. Due to the nature of group selection working on top of direct reciprocity [**?**] it would be possible to implement it on top of the Axelrod-Python library. Leaving kin selection to implement using the Prolog service.

1.2 Aims and Objectives

- To create a Prolog service that is able to support the decisions of a game-theoretic agent which interacts with other systems that wish to choose an agent, provide percepts from an environment and get actions from the agent.
- To explore in a game-theoretic context a mechanism to aid in the evolution of cooperation known as indirect reciprocity.
- To create a web application that allows users to set up and run game-theory tournaments of both direct and indirect reciprocity.
- To use the web application to demonstrate an environment that combines the Prolog service and a set of game-theory strategies and evaluate the results.

1.3 Literature Survey

Description and critique of past work in relation to my project.

Chapter 2: **Planning and Timescale**

Provide timeline of completed work and an updated one for next term.

Chapter 3: **Summary of Completed Work**

Agents service, flask web app, environment, two reports

3.1 Practical

3.1.1 Learning Web Dev

Refer to Flask web.

3.1.2 System Design

Overall system design.

3.1.3 Agents Service and API Design

RESTful API. MVFCEC. Agent Template. Prolog service learning.

3.1.4 Environment Design

3.1.5 Tools, Techniques and Processes

TDD, UML, Processes in VCS, Testing Strategy

3.2 Theory

Summary of the reports and theory work.

[1]

[1] A lot of my background theory work has been completed in my earlier reports so many of the references appear in those reports (see appendix)

Chapter 4: **Appendix**

Introduce the two reports

Report on the evolution of cooperation in relation to game-theory and different game-theoretic mechanisms to aid it

James King
Supervisor: Kostas Stathis

October 2018

Abstract

Since the early days of Darwinian evolutionary theory the phrase "Survival of the fittest" has become synonymous with much thinking on evolutionary dynamics. This idea has come along way since then but is still seen by many to promote selfish attitudes. However, we can see throughout nature that cooperation between biological agents is prevalent. So how did cooperation evolve and why has it flourished? This question fundamentally challenges what seemed concrete ideas about evolutionary dynamics, and has been approached from many angles. Game-theory is a branch of mathematics that has spawned many mathematical models of evolutionary dynamics in an attempt to solve the problem, some have been formulated programmatically to analyse the results. Study of the evolution of cooperation also has wider impacts in the world of computer science - namely on agent-based systems. A large component of agent-based system design is how interactions between agents works and how cooperation can be garnered within societies of agents. In this report, I shall explore past game-theoretic approaches to this problem. This exploration has helped me gather a deeper understanding of evolutionary dynamics, the reasoning behind these mathematical approaches to the problem of the evolution of cooperation and their application to the area of intelligent agents and multi-agent systems.

I. Introduction

Early Darwinians often focused on the struggle for survival as a means to explain the phenomena of evolution. This focus falls short on explaining the phenomena of altruistic behaviour, highlighted early on by Kropotkin [10]. The phenomena presented a problem for evolutionary biologists, how could it have evolved from an inherently selfish world?

Altruistic acts are actions where an individual puts others above themselves, these acts benefits recipients and are at a cost in some way to the actor. Examples of these acts and behaviours have been well documented and pervade both the natural world and human society.

In the seminal paper on the evolution of cooperation [1] Axelrod and Hamilton identify two theories proposed to solve the problem: kinship theory and reciprocation theory, focusing on the latter - particularly the Iterated Prisoner's Dilemma.

Here I will explore these mechanisms that attempt to explain altruistic behaviour and I will endeavour to relate them to multi-agent systems or even point out how they are inapplicable to this domain if appropriate.

II. Content and Knowledge

i. Kinship Theory

Kinship theory is an umbrella term for a number of models that attempt to solve the problem of the evolution of cooperation. The tie between these theories is that the individuals who choose to cooperate with each other are in some way 'kin'. The definition of what 'kin' is varies depending on the theory and the purpose for that theory.
Richard Dawkins popularized the idea of the 'selfish gene' [2]. This idea argues that as genes are the actual replicators, actors are hardwired to propagate the gene. This propagation does not only involve reproduction but also acts of cooperation to support and maintain others who share the gene. This is evident in many areas of nature especially in family groups, shown even recently in the BBC series Dynasties.
W.D. Hamilton supports a similar view [4]: that individuals with shared genes will work together to preserve those genes. He argues that to do this individuals don't work to add to their own fitness, but work to improve their 'inclusive fitness', which includes the fitness of other related individuals. His model converts the extent of the relatedness of an individual to a quantitative value using Wright's Coefficient of Relationship, which is shown in figure 1 on page 2. The model then uses this coefficient of relatedness to calculate inclusive fitness.
Axelrod and Hamilton [1] highlight that although the theory works well to explain co-operation between related individual it falls short in explaining the evolution of cooperation between unrelated individuals. This issue also makes it difficult to apply kinship theory as an aid to the evolution of cooperation in multi-agent systems. Kinship theory requires a reproductive drive and metrics for relatedness in order to motivate altruistic acts, both of which would likely be missing in a decentralised multi-agent system.

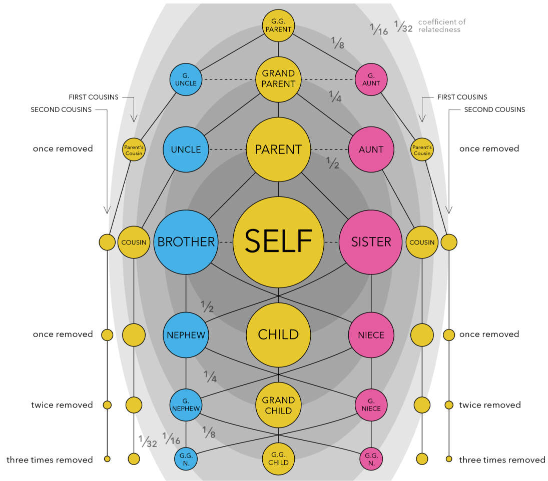
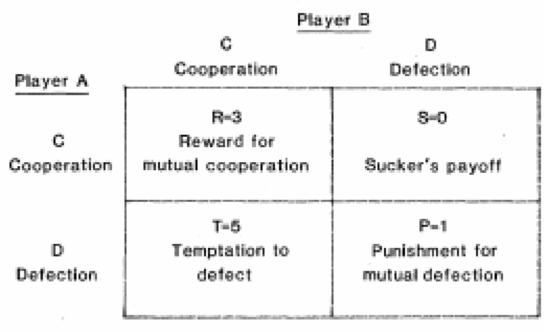


Figure 1: *By Citynoise - Own work, CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=37723128`*

ii. The Iterated Prisoner's Dilemma

The Iterated Prisoner's Dilemma is one such example of a game-theoretic model that uses direct reciprocity to attempt to solve the problem that the evolution of cooperation puts forward. Direct reciprocity is the idea: "If I cooperate now, you may cooperate later" [7]. The Dilemma is, in fact, a game played between two individuals. The game is made up of a number of repeated rounds, in each round the two players can choose between cooperating with or defecting against each other.
A payoff matrix is provided such as the one in figure 2 on page 3. These matrices provide a temptation to defect, but a reward if both players cooperate over multiple rounds. In a single round game it is mathematically best to defect, but in repeated rounds agents are encouraged to cooperate by a mutual gain in score. The mathematics are explained by Martin A. Nowak *et al.* in their paper "The Arithmetics of Mutual Help" [10].
The dilemma has been extended into tournaments such as the one in Axelrod and Hamilton [1]. You can even run tournaments directly in Python using the Axelrod-Python library [3]. Tournaments can come in a variety of styles but the most popular is the round-robin tourna-



Player A	C Cooperation	D Defection
C Cooperation	R=3 Reward for mutual cooperation	S=0 Sucker's payoff
D Defection	T=5 Temptation to defect	P=1 Punishment for mutual defection

Figure 2: *The payoff matrix given in Axelrod and Hamilton's paper [1]*

ment where every player plays a match of the iterated prisoner's dilemma against every other player. Players accumulate points throughout these games, based on the payoff matrix.
This has been even further extended to include a genetic algorithm to simulate evolution by reproducing players into the next generation. Multiple different algorithms have been used, one common one is the Moran Process (A stochastic process used to model evolution in a finite, unstructured population) available in the Axelrod-Python library.

iii. Strategies

Players in the iterated prisoner's dilemma have to be able to decide whether to defect or cooperate in any given round against any given opponent. The player has access to the history of interaction between the two individuals to base their decision on - the simplest of strategies, however, don't use the history, one example being a pure defector.
Some more interesting strategies include the world famous tit-for-tat, grudger and Pavlov (win-stay, lose-shift [6]). The classical form of tit-for-tat begins by cooperating in the first round and then copying the other player's last move, this strategy is so effective because it gives the other player a chance to cooperate (so working well with cooperators) but still punishes a defecting player preventing them from taking advantage. There are many other versions of tit-for-tat, a good example being forgiving tit-for-tat which only defects if the other player defects for two turns in a row.
Grudger works by beginning to cooperate but if the other player defects the individual switches and defects for the rest of the time, this is not so effective as tit-for-tat because it never forgives the other player, preventing further cooperation.
Pavlov works by continuing to act the way it has in the previous interaction if the action was successful if it is not successful the strategy switches to the other action. Nowak and Sigmund claim that Pavlov outperforms tit-for-tat as tit-for-tat is unsuccessful in non-deterministic environments (such as one that has a random chance that an action will not be what a player chooses). The real world is non-deterministic and so tit-for-tat doesn't generalise into it so well. Pavlov considers a success in a round to be: you both cooperate or you defect and the other player cooperated, and considers a failure to be: you both defect or you cooperate and the other player defects.

iv. Other Reciprocation Theories

Using reciprocity to aid in the evolution of cooperation has not been limited to the iterated prisoner's dilemma. Martin A. Nowak presents 3 other mechanisms that use reciprocation [7]: indirect reciprocity, network reciprocity (also known as spatial tournaments) and group selection.
Indirect Reciprocity is a set of mechanisms that leverage the group mechanic of reputation. Each member of the set of mechanisms uses different metrics for reputation and the spread of information that influences reputation. Nowak and Sigmund's version uses the idea of image scoring [9] an image score represents how well an agent is thought of, the higher the better.
Which mechanism is more effective is the subject of extensive debate, in two separate papers both Leimar and Hammerstein [5], and Roberts [11] concluded that the standing strategy is more effective. In the standing strategy agents do not have an image score but have a

standing which can be good or bad, everyone starts off with a good standing but if they defect against a player with good standing they are considered to have a bad standing. The spread of information is also key to indirect reciprocity as reputation is directly influenced by the community whereas with direct reciprocity there is no need for this. One mechanism for the spread of information is through gossip [12].

Network reciprocity (also known as spatial tournaments) uses direct reciprocity but not in a round-robin tournament. Spatial tournaments use a graph where nodes are agents and the edges represent channels of connection between agents [13], direct reciprocity games are then played between agents with connections. Nowak and May found that when deployed on a large scale using a reproduction mechanism beautiful spatial patterns can be observed [8]. More importantly, they also found that groups of cooperators can work together to support each other and fend off invasion by defectors - depending on the structure of the graph.

Direct Reciprocity is also used in group selection. In this model the population is subdivided into smaller groups, in these smaller groups individuals interact with each other and reproduction occurs from the fitness they gain in these interactions.

Though offspring are reproduced into the same group, reproduction occurs on two levels: group sizes fluctuate due to reproduction, if a group gets to a certain size it may split in two and when a group divides another is eliminated to keep the number of groups fixed. Groups with faster reproduction do better than those with slower reproduction. Traulsen and Nowak found this favours the evolution of cooperation under certain conditions [14].

III. Discussion and Conclusion

All these mechanisms to aid in the evolution of cooperation have extremely interesting properties and consequences in both biological and intelligent agents. Network Reciprocity can be used to represent a physical telecommunications network across which agents communicate, group selection is useful to model the spread of successful agent populations displacing others (like the current mass-extinction crisis) and kinship theory is useful to model interactions between different family groups in nature just to name a few applications.

With the internet being such a vast medium across which agents may interact it is very likely that agents may not have many repeated meetings, so direct reciprocity is not always applicable. However, agents may often repeat interactions, so indirect reciprocity is not always applicable.

Due to the incredible interconnectedness of the internet, the graphs used by network reciprocity would have to be highly connected and this makes it less useful for modelling agent systems using the internet. Due to the lack of reproductive drive and inherent relatedness between agents across the internet Kinship theory is unlikely to apply well.

Thus as my project will be aiming to simulate a multi-agent system with agents decentralised across the internet a combination of indirect and direct reciprocity will likely be the most germane aid to the evolution of cooperation for my project.

References

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

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

[3] 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 MÃijller; Georgios

Koutsovoulos; Areeb Ahmed. Axelrod: 4.3.0. `http://dx.doi.org/10.5281/zenodo.1405868`, September 2018.

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

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

[6] 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.

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

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

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

[10] Martin A. Nowak, Karl Sigmund, and Robert M. May. The arithmetics of mutual help. *Scientific American*, 272:76–86, 1995.

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

[12] 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.

[13] György Szabó and Gábor Fáth. Evolutionary games on graphs. *Physics Reports*, 446:96–216, 2007. Section: The structure of social graphs.

[14] 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.

Report on indirect reciprocity, strategies for agents and the development of a concrete model to implement

James King
Supervisor: Kostas Stathis

October 2018

Abstract

Indirect reciprocity is a mechanism that uses reciprocation theory to aid in the evolution of cooperation. Cooperation being an action taken by an individual that benefits another individual but at a cost to itself. Indirect reciprocity is a promising motivator for cooperation in societies with agents of higher intelligence levels and of greater sizes, such as human societies or even multi-agent systems. I plan to implement the mechanism programmatically, but there are many models and many possible additions to these models. In this report I will explore past approaches to indirect reciprocity, comparing and contrasting the variations proposed. The outcome of this exploration of approaches will be a concrete model to implement in this project.

I. Introduction

The evolution and preservation of cooperation has been a puzzle for evolutionary theorists for a long time. Many different approaches have been taken to give an explanation to cooperative phenomena, especially from the field of game theory. Often these approaches have come from the idea of reciprocity, where agents can grow mutually beneficial relationships through repeated interactions. The most popular mechanism of this being direct reciprocity where interactions are repeated between the same two individuals, and thus they may reciprocate directly with each other.

There is another game-theoretic mechanism known as indirect reciprocity, which works on the idea that nice agents will help those who help each other. This means learning information about another player does not require direct interaction. A number of models have been proposed to run indirect reciprocity. It is these models I shall be describing and reviewing, before using them to formulate my own to implement in my project.

II. Review of Past Work

i. Nowak, Sigmund and Image Scoring

Nowak and Sigmund's model [3] chooses a set of pairs of players at random each generation. A pair is made up of a donor and recipient. The donor can choose to cooperate - at a cost c, the recipient receives b points (b>c) - or defect - both receive 0 points, there is no cost to the donor though.

The model utilizes the idea of image scoring. In the first formulation of this all scores are public, but later on, it is suggested that this model is unlikely to hold true in real life. As an alternative, the idea of onlookers is suggested, where a set of onlookers are chosen for each interaction. In both models, image scores are increased by one for a cooperation as a donor or decreased by one for a defection as a donor, though in the second model only the onlookers and recipient can add this to their perception of the donor. The use of onlookers does make cooperation harder to establish, as it is dependent on whether a player knows an accurate image score of the recipient.

The points accrued through interactions gives the fitness score of each player. The amount a player reproduces into the next generation is dependent on the fitness score. Mutation is also an option in this model, this phenomena is when a player reproduces at random their offspring can be a different strategy. Mutation can lead to loops where defectors are invaded by discriminators and then discriminators are undermined by cooperators - allowing defectors to take hold again.

The two simplest strategies are pure defection and pure cooperation. A key strategy for the evolution of cooperation is the discriminator strategy. In fact there is a baseline number of discriminators required for cooperation to evolve. This strategy is given a number k, if the image score of a player is greater than or equal to k a discriminator will cooperate with them. More interesting strategies take into account their own image score or theirs and the recipient's image score, these can aid in the evolution of cooperation too. In the absence of information on other players, these strategies can believe other players to have an image score of k with a certain probability.

The model is known to be dependent on the probability of knowing the image of another player. Cooperation can only be stable if $q > c/b$, q being the probability of knowing the image, c being the cost and b being the benefit.

Leimar and Hammerstein recognised the problems of genetic drift in this model [1]. Adopting an island population model to restrict this. Rather than one single population group, the group is divided up into g amount of groups each with population n. This model also uses a different reproductive strategy where relative reproductive success locally is calculated as a normalisation of the sum of the fitness of all strategies within the same group. For global expected reproductive success sum each strategy and normalize over the whole population. A new individual is then locally derived with probability p or globally derived with probability 1-p. The strategy for the individual produced is randomly generated with a distribution corresponding to the local or global reproductive success.

ii. Standing Strategy

The standing strategy was suggested by both Leimar and Hammerstein [1] and Roberts [5] to be superior to image scoring. Their models used the island population structure given above. Leimar and Hammerstein replaced image scoring with the standing strategy whereas Roberts mixed direct and indirect reciprocity (with both image scoring and the standing strategy).

The standing strategy as described by Milinski *et al.* [2] is that players should not just aim for a good fitness, but also good standing. Everyone starts on a good standing, but if a player is a donor to a good standing recipient and does not cooperate with them the donor loses their good standing.

This is seen as a benefit as with image scoring it is not always in a players interest to punish those with a low image score, as they lose reputation themselves and in turn reduces the likelihood of them being the recipient of a cooperation. Whereas reputation is only lost in the standing strategy if the recipient is not of a good standing.

iii. Mixed Reciprocity Models

Gilbert Roberts puts forward mixed direct and indirect reciprocity models [5] using image scoring in one model similar to Nowak and Sigmund [3] and standing strategy. Both models use the island and reproductive systems Leimar and Hammerstein [1] utilized. Roberts put forward this notion due to his perception that indirect reciprocity alone is not a generalisable concept due to the close-knit nature of many societies.

Decisions can be made by agents in Roberts' model on the basis of either a reputation or an experience score. Reputation scores use image scoring in one model, differing from Nowak and Sigmund's as the score is within the range -1 to 1, initially at 0, decreasing by one for a defection and increasing by 1 for a cooperation. The standing strategy version assigns a -1 to a player when they defect against an individual or 1 to a player for anything else.
Experience scores are analogous to direct reciprocity for Roberts. When using image scoring, the experience scores between 2 players are initially 0 and go to -1 if they have defected or 1 if they have cooperated in the previous interaction. For standing strategy -1 is assigned only if the players has defected with a partner that has an experience score of 0 or 1 with them.
As you can see experience scores only reflect the previous interaction between the two individuals. This is a clear limitation as the model is seeking to describe a society made up of agents that have a higher intelligence.
Interestingly Roberts also considers the case where an agent has no resources to cooperate and thus must defect.
Steve Phelps [4] also recognises the same issue with indirect reciprocity as Gilbert Roberts - namely that remeeting in groups is likely, especially in smaller groups according to Phelps. Due to this Phelps suggests a mixed framework, to test how group sizes affect the use of direct or indirect reciprocity by agents.

iv. Gossip and Onlookers

It is recognised by Nowak and Sigmund [3] that it is unrealistic to expect all players to have directly observed a given interaction. Yet their answer - onlookers - hampers the evolution of cooperation, which relies on the probability that a player knows the image score of another player.
Sommerfeld *et al.* highlighted the importance of communication to the management of reputations. Gossip is a type of communication that is key part of societal mechanics. The findings of their experiment were that gossip needs to accurately reflect the interaction and the recipient of the gossip needs to correctly acted upon for gossip to be an adequate replacement for direct observation.

v. Comparison of Models

The aim of Roberts [5] is to compare the effectiveness of indirect reciprocity models and direct reciprocity. Roberts' results support Nowak and Sigmund's 1998 result that image scoring can support cooperation is robust even when criticisms relating to genetic drift and errors are taken into consideration.
However, it is highlighted that image scoring does not take into account who the donor is defecting against, whereas standing strategy does. Due to this image scoring suffers from an inability to distinguish between pure defectors and those who would cooperate with a cooperator. Standing strategy gives information both about how agents have interacted and the context of this action, an advantage over image scoring.
Counter to standing strategies perceived benefits it is seen that players are more susceptible to errors in perception, as standing falls faster than image score. Roberts also criticizes the standing strategy as not being a representation of the evolution of cooperation, as good standing can't have been established before the game had begun.
Roberts also points out that cooperation is only worth it when that cooperation is improving your chance of being cooperated with. This could be a counterpoint to the effectiveness of gossip as an alternative to direct observation. In gossip, there is a possibility of interactions being distorted by incorrect gossip. It could be interesting to see how the spread of misinformation could permeate society.

III. A Concrete Model

i. Specification of the Model

Both image scoring and standing strategy are good models for indirect reciprocity. They allow an agent to interpret a reputation and act as they see fit. However, the reputation of each player being universally viewable is unrealistic. The idea of onlookers seems closer to real life but is restrictive to the evolution of cooperation. The addition of gossip from these onlookers could be a sufficient alternative to direct observation, but gossip can be distorted by players wishing to influence proceedings against cooperation.

Gossip complicates a model because there is a need to interpret the information gossiped by another player. This also highlights an issue with the models that use purely standing strategy or image scoring: reputation is not a centralised system with everyone interpreting events the same way. Reputation has a very subjective nature, especially when gossip is introduced. Due to the subjectivity of perception about an agents reputation, I do not believe that one specific model for all agents can be considered appropriate. How a player views another players reputation must be up to their inside workings.

This requires an agent to have a view of past events and also a way to interpret these events in order to make decisions about another player. A model that uses this decentralised judgement on past events easily becomes a hybrid indirect and direct reciprocity game, as agents can have their own way to interpret interactions.

For an agent to interpret gossip, the gossip must be comprehensible by them and thus should have a given structure. In our model, gossip shall be split into two categories: positive and negative. Gossip will also involve 3 agents, the agent gossiping, the recipient of the gossip, and the gossip that the agent is about (who shall not be able to view that the gossip has taken place).

The structure of the game will be as follows. There is one singular population (this is to not overly complicate the model, though the addition of island population groups could be an extension), a population is replaced by their offspring after the conclusion of a generation. A generation is made up of a number of interactions, all at a distinct time point (the number of interactions per generation is a variable).

In each time point, each agent will perceive their environment and then decided what action to take, this action may be as a donor in an interaction or maybe a piece of gossip. Each interaction shall be made up of a donor-recipient pair, chosen at random from the population. For each interaction, a subset of the population should be chosen as direct observers.

Finally, reproduction occurs at the end of every generation and will be relative to each players fitness. Fitness is altered in each interaction if the donor cooperates they lose a fitness point but the recipient gains two, if the donor defects they both gain nothing and lose nothing. The group size remains the same over generations, so in a new generation individuals are generated and assigned strategies. The chance an individual will be a certain strategy is worked out by summing the fitness of all individuals of that strategy and dividing it by the fitness of the overall group. The user will be able to select if mutation is active or not, if it is there will be a 1% chance that a new individuals strategy will be chosen completely at random.

ii. Specification of Agent Strategies

At each time point in the generation that an agent is a part of the agent may perceive new information about interactions and gossip that has occurred in the game. The agent will then be asked to decide on what they want to do at that same time point.

The percepts that the agent can receive is that they are a donor at this time point, that they are a recipient at this time point, gossip from another agent and observation of an interaction at the last time point.

Based on the percepts an agent will be able to revise their beliefs about other agents and the environment. For all agents, if they perceive that they are a donor in a donor-recipient pair they revise their beliefs to know that they are a donor at this time point. However, the revision of beliefs is generally strategy dependent, for example, in an agent using the standing strategy if they observe an interaction where a donor defects against an agent that you consider to have good standing the observing agent will revise their beliefs to give a bad standing to the donor.

Agents will have a theory on the best action to take based on their beliefs, if they have not yet received percepts they may not have any beliefs yet, so will have to have a default action. Based on the agents belief's they will have a theory encoded on how is best to act at each time point. The theory encoded is dependent on the strategy they are using, for example, if a discriminator (with k=2) has perceived they are a donor and that the recipient of their action has an image score of 1 the theory will say it is best to defect. It will also be important to encode a theory on how to act if an agent is not a donor (how do they spread gossip? Do they remain idle?).

Of course, some agents will not care about percepts on other agents and will purely decide according to a plan such as always defect, or always cooperate. However, their needs to be a distinction on how an agent reacts to gossip and how they react to an observation. I propose two classes of interpretation: trusting and distrusting: Both can use any of the strategies, but distrusting strategies will not take into account gossip. An example of this is if an agent is using a distrusting standing strategy, negative gossip will not change their beliefs on the standing of a certain agent.

Using this design, strategies that will be implemented include: trusting and distrusting 'neutral' standing discriminator (similar to Roberts' version), trusting and distrusting image scoring discriminator (with varying values of k), cooperator, defector. This is not an exhaustive list and I will be including more strategies based on 'neutral' standing and image scoring. Other strategies include the use of image scores of both their own and the recipients and random cooperation/defection. An extension of this could be a strategy that uses a machine learning algorithm for whether to cooperate or defect as a donor.

These strategies only describe how to act as a donor based on beliefs on another agent. What strategies should be used for if a player is not a donor? Some strategies may include: stay idle, always gossip accurate information to random players (if none then stay idle), always gossip inaccurate information to random people, always gossip negative or positive information to random players, randomly gossip positively or negatively. It will be interesting as an extension to explore other strategies for both spreading and interpreting gossip.

IV. Discussion and Conclusion

Multi-agent systems require agents to be autonomous, they should not be working on centralised beliefs such as is used by Nowak and Sigmund [3] at first. My model not only removes those centralised beliefs and makes them part of the agent's inner workings but also makes the whole strategy they use part of the agent's inner workings.

The lack of individuality is where the work on indirect reciprocity I have seen falls short, because of the nature of indirect reciprocity being a community-driven mechanism the implementors see beliefs as community held and fail to realise that the beliefs may be community driven, but are actually individually held. This model is also very open to extension. Extensions on the population model (maybe using the island idea), the strategies (different interpretations of gossip and observations, strategies for the spread of gossip and strategies for actions as a donor) and mixing in indirect reciprocity by viewing observations where the agent is a recipient differently.

References

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

[2] Theo C. M. Bakker Manfred Milinski, Dirk Semmann and Hans-Jürgen Krambeck. Cooperation through indirect reciprocity: image scoring or standing strat-

egy? *Proceedings of The Royal Society*, 268:2495–2501, May 2001.

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

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

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

[6] 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.

Chapter 5: **Diary**