

# 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

March 27, 2019

# Declaration

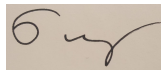
This report has been prepared on the basis of my own work. Where other published and unpublished source materials have been used, these have been acknowledged.

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: March 27, 2019

Signature:

A handwritten signature in black ink on a light brown rectangular background. The signature appears to be 'James King' written in a cursive, stylized script.

# Table of Contents

Abstract . . . . .	4
1 Introduction . . . . .	5
1.1 Motivation . . . . .	5
1.2 Aims and Objectives . . . . .	6
1.3 Contribution . . . . .	6
1.4 Structure . . . . .	7
2 Background . . . . .	8
2.1 Introduction . . . . .	8
2.2 Multi-Agent Systems and Intelligent Agents . . . . .	8
2.3 Game Theory . . . . .	15
3 Framework . . . . .	26
3.1 Introduction . . . . .	26
3.2 Theoretical Framework . . . . .	26
3.3 Implementation . . . . .	34
3.4 Running the Software . . . . .	44
4 Experiment Evaluation . . . . .	47
4.1 Introduction . . . . .	47
4.2 Robustness . . . . .	47
4.3 Stability . . . . .	49
4.4 Initial Viability . . . . .	50
4.5 Discussion . . . . .	50
5 Critical Analysis and Discussion . . . . .	51
5.1 Introduction . . . . .	51
5.2 Discovery . . . . .	51

5.3	Theoretical Foundation . . . . .	52
5.4	Implementation . . . . .	52
5.5	Consolidation . . . . .	54
5.6	Achievements . . . . .	54
5.7	Project Process . . . . .	54
5.8	Future Enhancements . . . . .	54
6	Conclusions . . . . .	55
	Bibliography . . . . .	60
7	Professional Issues . . . . .	61
8	Appendix . . . . .	64
8.1	Game Theory Background . . . . .	64
8.2	Theoretical Framework . . . . .	68
8.3	Experiments . . . . .	69
8.4	Strategies . . . . .	73

# Abstract

# Chapter 1: Introduction

Artificial intelligence (AI) has been an idea present in the consciousness of humanity for millennia. From Hephaestus' mighty 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. AI techniques are being used for many novel applications in areas such as medicine [36] and game playing [30].

Many of these applications are specifically using agent techniques [67, 33, 15]. Intelligent agents (IAs) have many definitions but one popular definition is that agents are anything that perceives and acts upon its environment [63]. Agents are situated in an environment and many of these agents can be combined in one environment to form a multi-agent system (MAS).

## 1.1 Motivation

Within MASs it is generally possible for the agents to interact and/or to communicate with each other. Communication generally occurs through agent communication languages (ACLs). The interactions that occur between agents are actions by one or all of the agents in the interaction. There is a possibility that agents can act out of pure altruism and always cooperate with other agents. However, often agents will work to protect their interests i.e. they are selfish.

It is desirable for agents in a MAS to work together to complete tasks and fulfil goals. Therefore, we want to facilitate cooperation between agents which may be selfish. There are analytical tools that we can use in game theory to understand what happens when decision-making individuals interact [48]. Some of these tools are used to explain how cooperation occurs between selfish individuals.

Many of these tools rely on the idea of reciprocal altruism [76], which stipulates that cooperation can occur between selfish individuals if they expect cooperation to be reciprocated. Two such mechanisms are direct and indirect reciprocity. Indirect reciprocity refers to when an agent cooperates with another agent with the expectation that this cooperation will increase the chance of receiving cooperation from others later. Direct reciprocity on the other hand, refers to the expectation of reciprocating cooperation from the agent who initially received the cooperation.

There are many factors to consider surrounding the mechanism. One key factor is the level of visibility of the interactions within the MAS. Nowak and Sigmund [54] limit this visibility to a randomly selected group of individuals in the population (onlookers). Another factor to consider is how information about the interactions can be conveyed. Sommerfeld *et al.* suggested the use of gossip [70].

So how can we use these mechanisms to facilitate cooperation between IAs? What options can we use and how can we make use of them to encourage cooperation between these agents?

## 1.2 Aims and Objectives

The high level aim of this project is to study how game-theoretic techniques can be leveraged in MASs to create cooperative societies of agents. Derived from this high level aim the objectives of this work are to develop a theoretical framework/model using game-theoretic and MAS techniques inspired by past work in both fields. The next is to implement this framework in a MAS that allows transparent decision making by agents; by which I mean that agents are able to give reasons for their decisions. Another objective is to build this implementation on a distributable platform that allows users to set up games of the model, specify certain variables and view an analysis of the game in order to study how the mechanism and variables affect cooperation in the system.

A feature of the system will be that users are able to create an account and refer back to games they have previously run. I will be using the system and specifically this feature to run my own experiments in order to study the effectiveness of the techniques I have employed and to discover how the system needs to be set up in terms of the variables present. I will then present these experiments and discuss the results with a conclusion on my findings and suggestions of future work as a result of these findings.

Objectives:

1. Develop a theoretical framework
2. Implement this framework
3. Run experiments using the implementation
4. Analyse the experiments and evaluate the results

## 1.3 Contribution

Much of the work I have reviewed in relation to game-theoretic mechanisms has come from the field of game theory. Because the studies have come from game theorists the implementation of the models devised by the authors have not generally used MAS techniques. My implementation is not only a MAS using a game-theoretic model, but it supports transparent agent decisions due to the use of the logic programming language Prolog for the implementation of agents' decision making components. This implementation is also a web based system, allowing users across the world to experiment with the model I have devised.

Furthermore, the theoretical framework I have designed is inspired by past work on game-theoretic mechanisms, but also builds upon these approaches. One way in which I build upon these is by combining the many aspects of the past approaches such as using MAS techniques to create a game-theoretic model (combining direct and indirect reciprocity) and using gossip as an action to convey reputation information in a mixed reciprocity model.

Lastly, in my system, users associate agents with specific strategies which I have included in the system. These strategies come from past work in the topic of game theory, but are implemented using an agent architecture which I have purpose-designed for my theoretical framework. This architecture included augmenting the original strategies with strategy components for whatever roles agents may have at a particular time-point in a society (such as when an agent is not in an interaction). Another development is the trust models which I have created for the original strategies for interpreting the events in the environment.

## 1.4 Structure

In the next chapter of this report I have explored and described the past work that has been a central inspiration to the theoretical framework I have devised. I have included other interesting information and possible mechanisms from past work that could be used to encourage cooperation in MASs in the appendix. This background reading then leads into the next chapter in which I have defined the theoretical framework for my system.

Following on from the theoretical framework in the same chapter, I have described the implementation of this theoretical framework in a web application. This application is available for users across the internet to set up games and view the events and analysis of their games. In the next chapter, I use this platform to experiment in order to review the effectiveness of the mechanism, strategies and variables in the game in encouraging cooperation between selfish agents.

The final two main chapters of this report discuss my findings from the experiments and rest of the project. I then conclude and evaluate the project. The appendix is at the end and contains further information referenced in the report but which is not necessarily central to the project.



## Chapter 2: Background

### 2.1 Introduction

In the following sections of this chapter I will be exploring and describing past work relevant to this project including MAS techniques and related studies in game theory. First I will describe MAS techniques and concepts that are important to this project including subsections on agents, agent environments and agent communication. Secondly I will discuss game-theoretical approaches to modelling and studying multi-agent interactions in regards to the evolution of cooperation between agents.

### 2.2 Multi-Agent Systems and Intelligent Agents

MASs is a field in computer science concerning the development of societies of agents. Wooldridge [78] describes agents as computer systems capable of autonomous decision making to reach delegated goals or design objectives. These agents are often designed to work together and typically interact using agent communication languages (ACLs).

#### 2.2.1 Environments

Design objectives and delegated goals for agents are constrained by the environment the agent is situated in. As such, agents are designed to work within a class or set of environments [63]. Often agents are defined in terms of their environment. Russell and Norvig [63] described an agent as being any system that uses sensors to perceive it's environment and actuators to act upon it's environment.

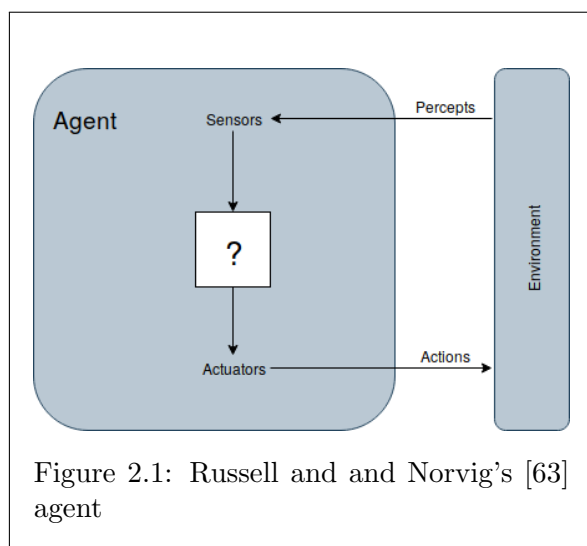


Figure 2.1: Russell and Norvig's [63] agent

According to Russell and Norvig [63] to specify a task environment one needs to specify four aspects of the system: performance measure, environment, actuators and sensors. In the 2013 version of the textbook 'Artificial Intelligence: A Modern Approach' [62] Russell and Norvig described a generic program for running an environment.

In this program a set of steps are run in a cycle. In the first step, perception of the environment occurs for each agent in which the agent can use its sensors to receive percepts from the environment. In the next step agents all decide on an action. After all have decided the state of the system is then changed based on the actions that the

agents have decided on. A fourth step was added to update the performance scores of each agent.

The effect of these cycle steps is to keep agents synchronised, preventing agent actions

from changing the state of the environment while other agents are deliberating. A static environment is one that does not change whilst an agent is deliberating. Dynamicity is one of Russell and Norvig's [63] seven properties of an environment. The synchronicity of this generic environment program helps keep the environment static, unless other forces (not agent actions) have an effect on the environment's state.

The opposite to a static environment is a dynamic one, in which the environment can change during decision making. Another environment property is observability. A fully observable/accessible environment is one that gives an agent full access to the state of that environment. In a fully observable environment there is no need for an agent to keep any internal state about the environment. Unlike in a partially observable environment, in which an agent cannot simply sense all aspects of that environment that are relevant for decision making.

A third property from Russell and Norvig [63] is whether the environment is episodic or sequential. An episodic environment is split into episodes where the quality of actions in each episode do not depend on previous episodes. Whereas, in a sequential environment an action taken at timepoint  $t$  may be affected by actions from previous episodes and the action at timepoint  $t$  may affect all future actions.

Another property is whether the system is multi-agent or single-agent. This seems simple, however, it massively depends on what we see as an agent. For example take a surveillance system, with a central agent collating the data and a number of connected pieces of surveillance systems (for example CCTV cameras). There are blurred lines as to whether these connected components are agents or are they simply governed in how to act by regular systems level programs?

The environment has a set of laws that govern how the state changes based on events and actions. Agents may or may not know these laws. It is said that if an agent knows the outcome or probability of the outcome of its actions (for stochastic environments) then the environment is known by the agent. If not, the environment is unknown by the agent.

These properties determine how complex an environment is. The final two are whether the environment is deterministic or stochastic and whether the environment is continuous or discrete. If an environment is deterministic it means the effects of an action given some conditions is guaranteed. If an agent can access the whole state of the environment (fully observable) and the environment is deterministic the agent can guarantee the outcome of their action.

An environment is considered discrete if actions and percepts are clearly defined and there are a limited number of them. The more properties of an environment that match the following properties, the more complex the environment: continuous, stochastic, partially observable, dynamic, sequential, multi-agent and unknown.

## 2.2.2 The Agent Model

Yoav Shoham introduced the agent-oriented programming (AOP) paradigm [66]. This paradigm is similar to object-oriented programming (OOP). In Shoham's formulation of AOP [65] object units are replaced with agents, the internal state of a unit is constrained to beliefs, commitments and capabilities, message passing is constrained to using speech act theory to communicate and constraints are placed on methods.

## What is an agent?

One of Russell and Norvig's properties [63] of an environment is whether the environment is multi-agent or single-agent. How we classify an environment as to this property is very much determined by what we class as an agent, which is not a trivial question. How do we define agency?

Since Shoham's formulation there has been much debate as to whether systems are agent systems or simply programs [22]. The exact definition of agency varies depending upon which source is consulted and the application of the concept.

Wooldridge and Jennings [79] gave one of the most influential and powerful descriptions. They put forward a weak notion of agency by laying out four main properties: autonomy, social ability, reactivity and pro-activeness.

Autonomy refers to agents having control over their actions and internal state, and operating without outside intervention. Social ability is the use of an ACL or other interaction medium to communicate with other entities such as humans and other agents. Responding to environment changes in a timely fashion is what the property of reactivity refers to. Finally pro-activeness is the display of taking initiative to work towards a goal, not simply in reaction to environmental state changes.

Wooldridge and Jennings go on to note that stronger notions of agency endow agents with human-like mental notions. These mental notions represent information in a symbolic model, and make decisions by reasoning with this symbolic model in a symbolic way. This representation and reasoning is no easy feat - known as the representation/reasoning problem [78] - but it does allow agents to make explainable and transparent decisions.

A stronger notion of agency may also include the use of new developments such as computer vision, natural language processing and deep reinforcement learning concepts. Stronger notions of agency show a tendency towards more humanlike characteristics and AI concepts, drawing clear comparison with a long term goal of AI: to produce systems capable of beating The Imitation Game [39].

## Russell and Norvig's Agent Program Components

Russell and Norvig [63] formulated a number of agent program models to describe the internal make up of agents that process percepts and make decisions on what actions to commit to.

The first model is simple reflex agents which keep no internal state, they use only the percepts from their environment and condition-action rules to decide on actions. These are comparable to reactive agent architectures, but on a less complex scale. Model-based reflex agents use percepts from their environment to form an internal state. This state is used alongside condition-action rules and a model of how the state of their world evolves to make decisions.

Goal and utility-based agent models focus on the drives of an agent. The goal based architecture keeps an internal state, and combines this with search and planning decision making components to reach specified goals.

States in which an agent is more successful can be formulated in terms of having a higher utility. Utility based agents work to maximise that utility, even in situations when an agent has conflicting goals in which the utility is formed by creating a tradeoff between the conflicting goals.

## Learning Agents

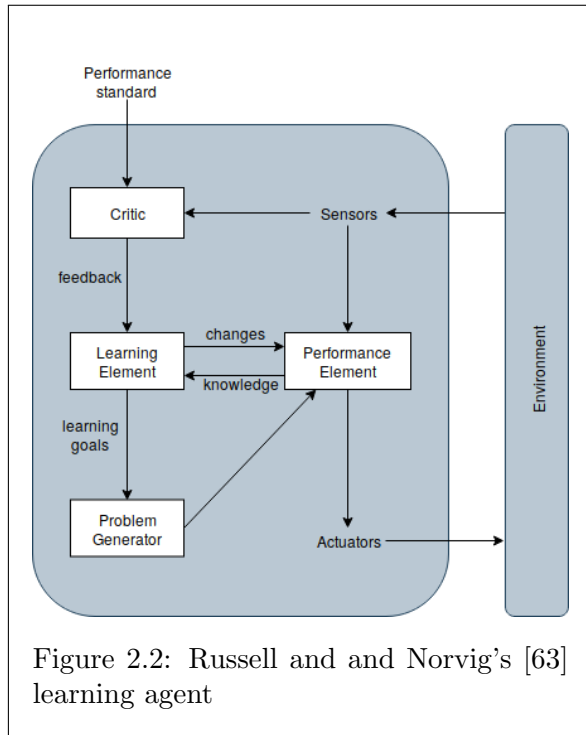


Figure 2.2: Russell and Norvig's [63] learning agent

Learning agents is the final agent model described by Russell and Norvig [63]. Programs created using this model are not formed by a programmer specifying how to act. Instead these agents use a number of components to feed into a decision making system (performance element). Feedback is provided by the critic component to the learning element which makes changes to the performance element. The problem generator is another component that suggests actions that will hopefully lead to new knowledge.

Machine learning is exploding with new concepts and developments, most notably from the agents systems perspective deep reinforcement learning. Agents using machine learning techniques often appear to satisfy stronger notions of agency. However, these techniques do not come from a symbolic AI background, and lead to a lack of explain-

ability [27].

Unifying these techniques with symbolic AI approaches to produce agents which can explain their decisions in order to produce verifiable systems is a keenly researched topic [27, 24]. However, this is an even tougher version of the representation/reasoning problem as we are not only trying to reason with symbolic data, but reason in a complex manner with deeply mathematical techniques.

## Deductive Reasoning Agents

Agents are built with certain architectures defining the internal structure of the agent [40]. One type of architecture is the deductive reasoning agent architecture which use a form of symbolic AI. The symbolic representation of information is in the form of logical formulae and the reasoning component is done through logical deduction or theorem proving [78].

Shoham's AOP paradigm [66] and Agent0 language [65] were developed to facilitate the creation of deductive reasoning agents. The symbolic representation was formulated in terms of beliefs, commitments and capabilities and agents were given a set of commitment rules to logically deduce what actions to commit to in a given situation.

Concurrent MetateM is another language for the creation of deductive reasoning agents [20]. The language is known for focusing on creating agents that work by theorem proving. Programmers essentially create deductive theorem provers using temporal logic that dictate agent behaviour.

Deductive reasoning agents suffer from the same two issues as many symbolic AI approaches: the transduction problem and the representation/reasoning problem. The transduction problem refers to the difficulty in translating from the real world into an accurate and adequate symbolic representation. The representation/reasoning problem are the difficulties in representing symbolic information and reasoning about that information.

## Practical Reasoning Agents

Agent architectures that utilize symbolic AI approaches aren't limited to deductive reasoning agents. Practical reasoning approaches focus reasoning towards actions, whereas theoretical reasoning focuses on beliefs [78].

However, this does not mean that practical reasoning models do not incorporate mental notions such as beliefs. Both deductive reasoning languages such as Agent0 and practical reasoning focused languages and frameworks such as GOAL [31] and PRS (The Procedural Reasoning System) [25] have used the idea of beliefs to base actions on.

This use is mostly inspired by the beliefs-desires-intentions (BDI) model [14], which PRS uses directly. BDI agents keep beliefs in a belief base. These beliefs are the information an agent sees as facts about the world and agents are capable of inferring beliefs based on the presence of other beliefs.

Desires are the drives behind an agent's actions. A desire is an objective or state an agent wishes to reach. When an agent is actively pursuing a desire it becomes a goal. These goals must be consistent i.e. they cannot contradict each other. When an agent has committed to a desire this desire becomes an intention. Commitment to a desire is when an agent has begun executing a plan that works towards achieving that desire. These plans are sequences of actions.

Rodriguez *et al.* [60] criticised agent-programming languages (APLs) for forcing the BDI agent architecture upon users of the language, they believe that programmers should have the choice on what architecture to use. Their claim is that the use of BDI constrains agents to be endowed with reasoning capabilities that do not apply to some problem instances. The language SARL - specified by Rodriguez *et al.* [60] - is a more generic language which aims to be architecture independent.

## Reactive and Hybrid Architectures

Applications of agent technologies in robotics and embedded systems require architectures and programs to work in resource bound environments. Brooks [16] argues for intelligence as an emergent property of complex systems.

Brooks criticised the use of symbolic AI in the creation of agent systems [16]. He claimed that the abstract reasoning and explicit representations of symbolic AI aren't required for the development of intelligent agents as intelligence is an emergent property of complex systems. In resource bound environments this abstract reasoning and explicit representation are often too resource hungry to be applicable.

$$\begin{array}{l} K_1 \rightarrow a_1 \\ K_2 \rightarrow a_2 \\ \dots \\ K_m \rightarrow a_m \end{array}$$

Table 2.1: An example teleo-reactive program

In reaction to this Brooks devised the subsumption architecture. In this architecture task-accomplishing behaviours map from percepts to actions, in order to accomplish tasks. These are also known as behaviour modules. These modules are able to decide on actions at the same time as each other. Brooks employs a hierarchical structure to select which behaviours decision to use. The lower the layer a behaviour belongs to the greater it's priority.

This hierarchical selection of actions is notably similar to teleo-reactive programs [49]. These are programs which contain an ordered set of production rules. These rules map from conditions to actions, and the ordering of them defines which actions are more important.

A criticism of reactive architectures [78] is that through lack of internal state an agent using these architectures can only view the local environment, forcing a short-term view. Hybrid architectures have been developed to combat these issues. One method in which to implement hybrid architectures is to use free variables in teleo-reactive program conditions, with which beliefs bind to at run-time.

### 2.2.3 Multi-Valued Fluent Cached Events Calculus

The aim of the event calculus is to create symbolic representations to reason about time periods and local events [35]. Events occur at specific timepoints, these events initiate and terminate periods of time for which a fluent holds. For example at time  $t1$  event  $e1$  occurs, fluent  $f$  did not hold prior to  $e1$ , but  $e1$  initiated it and as such it holds after  $t1$ . A second event  $e2$  then occurs at timepoint  $t2$  where  $t2 > t1$ , this event terminates  $f$ . Following this event  $f$  no longer holds, but  $f$  still holds in between  $t1$  and  $t2$ .

The multi-valued fluent cached events calculus (MVFCEC) is an implementation of the events calculus [33]. The implementation improves the speed of querying fluents using a caching system.

In Kowalski *et al.*'s [35] original event calculus fluents were considered Boolean. Take fluent  $F$ , in the original event calculus  $F = true$  if it holds or  $F = false$  if not. With multi-valued fluents  $F = V$  where  $V$  is a value specified by the programmer.

Agents often need to reason about time and states. Agent systems can be developed using a symbolic approaches and their systems are often logic-based. The events calculus is a very powerful logic-based tool for this purpose and as such variants of the event calculus have been used in the development of many MASs [8, 33].

The Agent0 belief system reasons about the initiating and terminating of beliefs with certain values through time periods. This belief system accentuates the prevalence of agents requiring methods to reason about time, and it seems natural to use MVFCEC for the representation of a beliefs like system.

### 2.2.4 Agent Communication

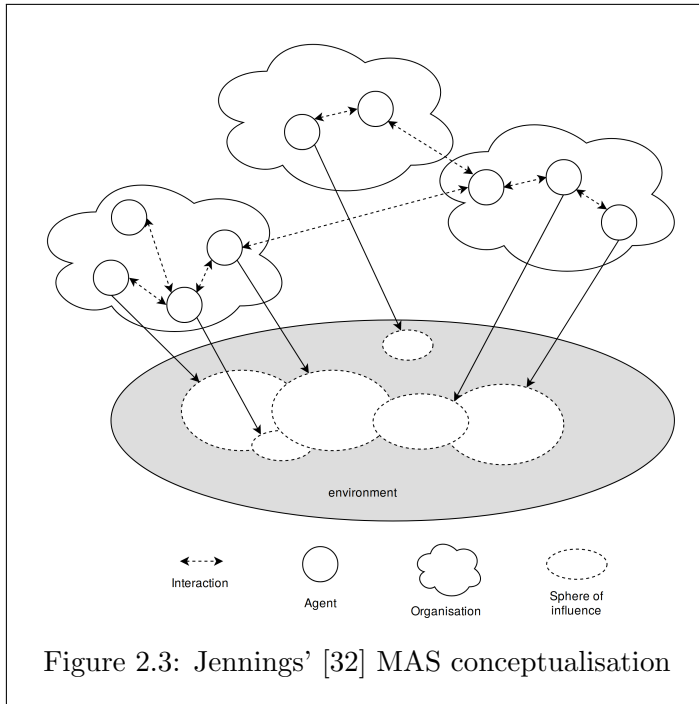
To build cooperative and coordinated societies of agents, agent communication is extremely important. In fact it's so important that Wooldridge and Jennings include social ability as one of his four key properties [79]. In the AOP paradigm agents communicate socially via a message passing system. This message passing can be achieved using an ACL in order to ensure communicative actions are understood by the recipient of the message.

Many of these languages use speech act theory. Speech act theory uses the idea that communication is an action often to achieve goals and intentions [9]. Both KQML [19] and FIPA ACL use performative verbs to specify the intent of the communication. Both have different performative verbs, but in combination with the message content these languages communicative acts generally fall into one of the five categories of speech acts described by Searle [64].

In human communication we generally have a shared understanding of certain concepts. Such as the definition of a book or plant. In KQML and FIPA ACL shared ontologies are used to define the meaning of commonly used vocabulary for certain subject domains in order to facilitate shared understanding of message content.

The theory of communication is complex. According to Singh [68] a key limitation of existing ACLs (including KQML) is their focus on mental agency over social agency. To overly simplify this Singh's description, current languages use the mental state of an agent such as communicating about the beliefs of an agent. Singh claims that this supposes that agents can read each others' mind.

## 2.2.5 Multi-Agent Interactions



Agents are able to interact through communicative actions and by actions in the environment. In Jennings's [32] view of multi-agent interactions, interactions between agents often occur but are not limited to when the agents are linked by organisational relationships. Organisational relationships are subject to change. These agents have their own spheres of influence in an environment, which may overlap each others. An overview of this setup can be seen in figure 2.3.

When spheres of influences overlap an agent may need to make a decisions that is dependent upon other agents decisions. These overlaps can be described as interactions.

When agents are making decisions in regards to interactions, it is known as strategic interaction.

When studying multi-agent interactions agents are generally assumed to be self-interested. Russell and Norvig [63] define rationality as being self-interested, claiming that rational agents actively seek to maximise their performance measure/utility. Utility is the the representation of how beneficial events and actions in a system that transform the state of that system are to agents. The utility produced for all agents is the social welfare. Some agents seek to maximise social welfare, so may have an interest in other's utility.

This aim to maximise social welfare may seem contrary to the presumed self-interested nature of these agents, however, mechanisms can be used to encourage self-interested agents into increasing social welfare when it benefits their performance measure. Some of these mechanisms come from game-theory and can be used to model and simulate agent interactions using these utility functions.

## 2.3 Game Theory

### 2.3.1 Introduction

As discussed we can use game-theoretic techniques to model and study multi-agent interactions. In the following subsections I will explore past studies that model societies of individuals interacting with certain games, and the concepts surrounding.

### 2.3.2 Cooperative Phenomena

Game theory formulates mathematical models to study conflict and cooperation between intelligent rational decision-makers [48]. In these mathematical models actions are often defined in terms of cooperation and defection. These models apply to more than just multi-agent interactions, including the natural world.

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” [71].

Axelrod and Hamilton [10] note two key areas of study that attempt to explain cooperative phenomena in the face of competitions: 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. Due to reasons explained in my appendix in subsection 8.1.2 I will be focusing on reciprocal altruism.

### 2.3.3 Reciprocal Altruism

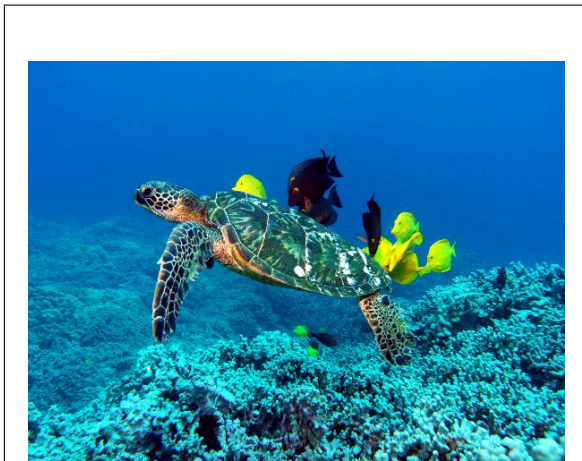


Figure 2.4: Cleaning symbioses such as that of the green sea turtle and surgeonfish including the yellow tang show that reciprocal altruism is possible between non-humans and even interspecies [81]

Reciprocal altruism is an idea most famously put forward by Robert L. Trivers [76]. Trivers defines altruism as behaviour of one organism that benefits another to whom it is not closely related, while being apparently detrimental to the organism performing the behaviour. From this definition and from 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 idea is a move away from limiting individuals to cooperating with only their kin - a key limitation of Kinship Theory - and towards any individual that they believe their cooperation will be reciprocated by. Axelrod and Hamilton [10] noted this concept as advantageous in explaining cooperation between unrelated individuals, such as is common between humans. I would argue that this concept is also more applicable to higher intelligence societies such as those



possible from MASs.

### 2.3.4 Axelrod, Hamilton and The Iterated Prisoner's Dilemma

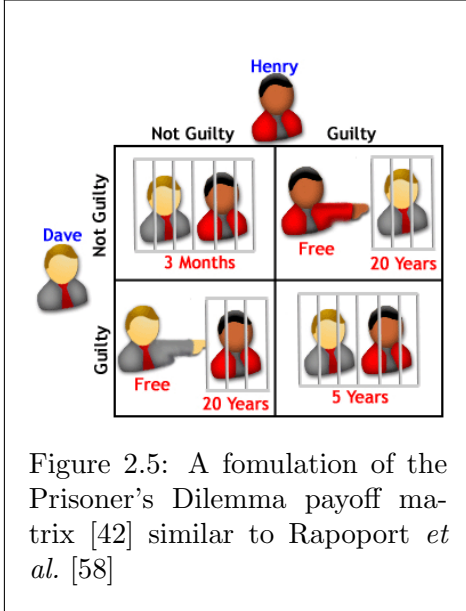


Figure 2.5: A fomulation of the Prisoner's Dilemma payoff matrix [42] similar to Rapoport *et al.* [58]

The prisoner's dilemma was described by Rapoport *et al.* [58] in terms of criminals under interrogation, but Axelrod and Hamilton formulated it in terms of cooperation and defection. This formulation allows the study of a more generic form of interaction between individuals.

In a single round of the prisoner's dilemma two individuals simultaneously choose to either cooperate or defect. After each interactions each individual receives a payoff in the form of a number, the higher the payoff the better. This payoff is dependent upon the actions of both individuals and is illustrated in table 2.2.

From an agents perspective we can view the pay-off in terms of utility or what is preferable to an agent [78]. Using player A and player B, with  $u_i$  being the utility function for agent i, and the payoff matrix

provided by Axelrod and Hamilton [10] we have the following utility functions:  $u_A(D, D) = 1$ ,  $u_A(D, C) = 5$ ,  $u_A(C, D) = 0$ ,  $u_A(C, C) = 3$  for player A and  $u_B(D, D) = 1$ ,  $u_B(D, C) = 0$ ,  $u_B(C, D) = 5$ ,  $u_B(C, C) = 3$  for player B.

We can order the possible situations by utility for each agent to understand what situations are preferable to each agent. For player A:  $(D, C) \geq (C, C) \geq (D, D) \geq (C, D)$ . For player B:  $(C, D) \geq (C, C) \geq (D, D) \geq (D, C)$ . As you can see players are at odds as to which situations they wish to arise. If both attempt to bring about the situation that is preferable to them they will both receive only 1 as pay-off each.

In the iterated prisoner's dilemma the two agents repeat these rounds and can base their decisions on previous rounds. This allows individuals to use strategies that make use of a mechanism known as direct reciprocity [52], which is a type of reciprocal altruism.

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

Table 2.2: The payoff matrix in a typical iterated prisoner's dilemma game (such as Axelrod and Hamilton's [10]). A=x, B=y where x denotes the payoff for A and y denotes the payoff for B.

Axelrod and Hamilton [10] used games of the iterated prisoner's dilemma in their round-robin tournaments. These tournaments had a set of players that are associated with strategies. Every player entered into a game against every other player, in which rounds were repeated. For cooperation to be rational in the iterated prisoner's dilemma the 'shadow of the future' needs to present in each round [78], so games were not cut off at a set point but ended by increasing the probability the game won't continue each round.

Axelrod and Hamilton [10] combined these round robin tournaments with a genetic algorithm [44] (genetic algorithms are discussed in subsection 2.3.11). One aim of Axelrod and Hamilton's [10] paper was to review how successful the strategies submitted to them were. The use of a genetic algorithm allowed them to ask three questions of each strategy. Is it robust? Is it stable? Is it initially viable?

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

Axelrod and Hamilton [10] found two strategies with these three abilities: 'tit-for-tat' and 'all defect'. Later on, Nowak and Sigmund [50] 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 also actively aid in the evolution of cooperation.

Nowak [52] 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:  $w > c/b$ . If the probability of subsequent encounters between two individuals is low, then it is very likely that cooperation will not evolve as there will be no sufficient reward. 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.

Axelrod and Hamilton's program was not created as a MAS. However, the importance of remeeting is a definite limitation to the use of direct reciprocity in MASs, and as such I do not see direct reciprocity as a strong contender to encourage cooperation on its own. One required property of a mechanism to encourage cooperation is that it must work when both re-meeting is unlikely and when it is likely. As such, direct reciprocity is not completely inadequate for the problem. However, it is insufficient when not combined with some encouragement for when chances of meeting again are low.

Extensive interest in reciprocal altruism in recent years has been focused towards direct reciprocity, including a number of available libraries such as the Axelrod Python library [18]. 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.

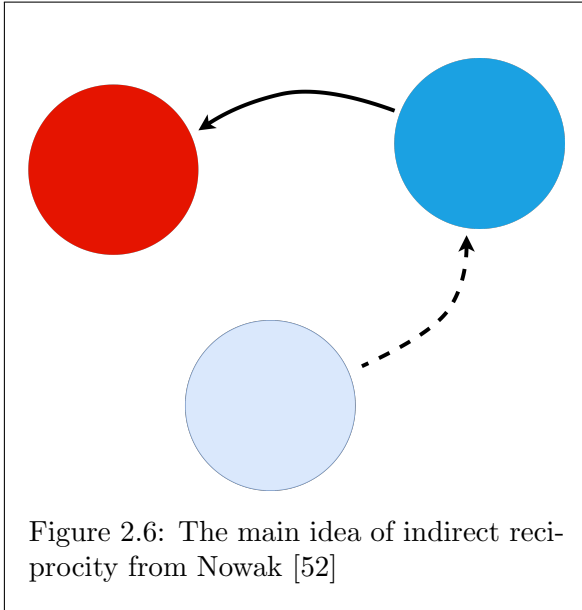
### 2.3.5 Indirect Reciprocity

Nowak [52] presents five mechanisms that aim to facilitate the evolution of cooperation. Direct reciprocity is used in the iterated prisoner's dilemma. I discuss three others in my appendix (subsections [ 8.1.2, 8.1.4, 8.1.3]) and the final one I shall discuss here: indirect reciprocity

This concept is promising for use in MASs as it solves the failure of direct reciprocity's failure with regards to when re-meeting is low. In a large society or a vast network in which agents are likely to operate in success when re-meeting is low 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 [6], who was an early advocate the idea, focused on human reciprocal altruism, however, subsequent research has abstracted away from the biology [56, 41, 54, 59, 52, 37, 73, 70, 47]. The idea is that if an individual cooperates with another individual, 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. Thus this mechanism a form of reciprocal altruism.



According to Nowak and Sigmund [54] 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 kind of higher level intelligence which is required to reason about events in a group that could be a key part in the development of MASs.

Due to this reason, 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 both past approaches to indirect reciprocity and other factors in the mechanism that can be used to facilitate the evolution of cooperation.

### 2.3.6 Nowak and Sigmund

According to Gilbert Roberts [59], Nowak and Sigmund [54] is the most influential model on indirect reciprocity. Therefore, I shall examine this model of indirect reciprocity first. Nowak and Sigmund 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 that this mechanism channels cooperation toward valuable members of the society of players.

Nowak and Sigmund note that the use of image scores and indirect reciprocity itself, leave 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. At the same time, anticipation and planning are a key part of agent design in multiple languages and frameworks. In fact the lack of planning was such an important drawback of the Agent0 language [65] Becky Thomas created a new language - PLACA [74] - which allowed agents to plan, among other improvements.

The framework created by Nowak and Sigmund is simplified from Alexander's [6] idea of human reciprocal altruism. Nowak and Sigmund describe a framework in which there is a population of individuals which act 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 its 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 shown in the payoff matrix in table 2.3.

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

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. The notion of onlookers limits the visibility of actions in the society by randomly selecting a group of specified size to view each interaction. This concept is displayed graphically in figure 2.7. The concept of onlookers was added due to the realisation of Nowak and Sigmund's that in a group that is spread over a wide geographical area, not

all individuals will be able to view each interaction. This sparse nature of interactions is of course especially likely in MASs. Image scores now become one player's 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 the value  $ImageScore[u, v] \geq k$  otherwise it defects [28]. This strategy is incredibly simple yet effective. 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. In 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.

Similarly to Axelrod and Hamilton's tournaments [10] Nowak and Sigmund [54] used a genetic algorithm to find optimal strategies under certain conditions. They varied conditions such as the length of each generation and the number of onlookers per interaction.

Nowak and Sigmund hypothesize and give evidence supporting the idea 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.

Another idea put forward and evidenced by Nowak and Sigmund is that the evolution of cooperation is dependent upon the donor's knowledge of the image score of the recipient. This chance of knowing an image score is limited by the concept of onlookers. However, in Nowak's 2005 paper on the five rules of cooperation [52] he talks about using 'gossip' as an alternative to direct observation. Gossip and social ability are interesting concepts in MASs. According to Wooldridge and Jennings [79], social ability is a key property in even a weak notion agents.

The model laid out by Nowak and Sigmund is succinct and a good basis for looking at interactions in multi-agent systems. Similarly to Axelrod and Hamilton [10] Nowak and Sigmund [54] did not create a MAS to run their games.

If they had been considering MASs they may have noticed a number of limitations. 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 of an individual. However, I argue that the idea that image scores are attempting to capture (reputation) is actually more personal. Although there might be a rough consensus as to an individual's 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.

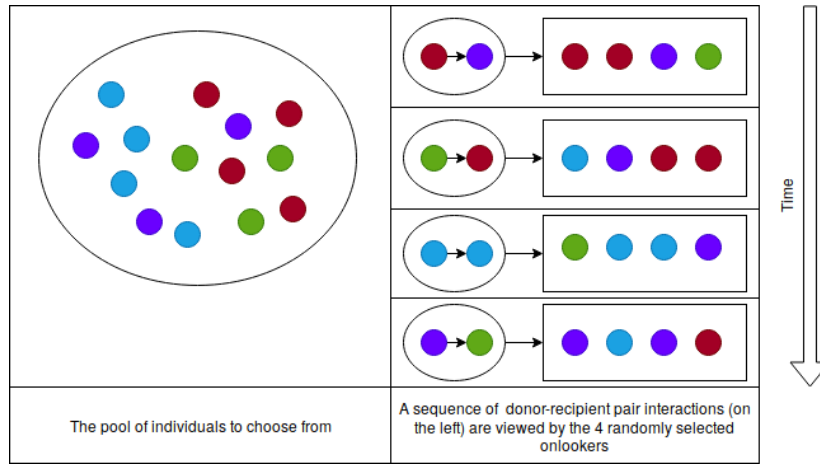


Figure 2.7: The selection of a sequence of donor-recipient pairs and onlookers in Nowak and Sigmund's [54] model of indirect reciprocity. The macro-view of indirect reciprocity. The lines and arrows show the spread of information.

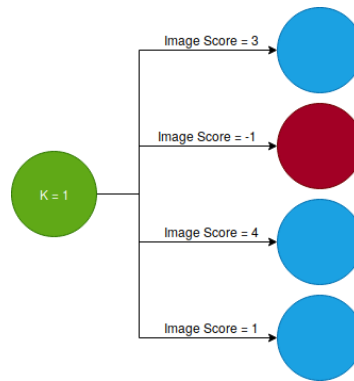


Figure 2.8: On the left an image scoring discriminator with  $K=1$ . On the right the agents the discriminator is a donor for. Blue agents represent where there would be a cooperation from the donor, red represents when the discriminator would defect against them. The rule is when  $ImageScore \geq K$  cooperate else defect.

According to Russell and Norvig [63], IAs use an internal state. It is in this internal state, I argue, that image scores should be stored in order for an agent to have full control over them. Trust models could be devised to manage these image scores. This idea is a closer match to the MAS 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 extensive interesting aspects that have been examined carefully including the onlookers and reproduction mechanisms, and also includes 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.

### 2.3.7 The Standing Strategy and Further Limitations of Nowak and Sigmund

Leimar and Hammerstein [37] criticised Nowak and Sigmund's [54] limited range of strategies. They also highlighted the abilities of another strategy, the standing strategy,

over image scores. The standing strategy was first described by Sugden [73] and uses a ‘switch’ [54].

Milinski *et. al* [41] described the standing strategy as the idea 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 event that causes the switch from good to bad 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 action is seen as punishing a bad individual, as it channels fitness away from bad members of the society. However, the donor’s image score is reduced by 1. This change in score 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 pure altruism.

Conversely, the standing strategy only reduces the reputation of those who defect against good individuals. Therefore, punishing bad members of the society has no negative effect on those members. Leimar and Hammerstein [37] 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* [41] analysed Leimar and Hammerstein’s argument, suggesting that they were correct under conditions which allowed for perfect perception of events and unlimited memory capacity. However these conditions are not always true, especially in a MAS in which the environment is inaccessible [63].

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

### 2.3.8 Mixed Reciprocity Models

Roberts [59] and Phelps [56] noted that indirect reciprocity is focused towards interactions with other individuals with whom the donor has no previous interactions. However, what about when re-meeting is more likely? I stipulate that an individual is affected by both being the recipient of actions and their observation of actions. Roberts engaged with the issue by introducing an experience score. This experience score works similarly 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  $1$ ). However herein, 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 against the use of direct reciprocity. Roberts concludes that indirect reciprocity is the more popular decision mode under conditions in which re-meeting was less common, and direct was more popular

when re-meeting was frequent. On the other hand, Phelps' experiments garnered different results being that in small groups direct and indirect reciprocity exist in equilibrium.

Similarly to Nowak and Sigmund [54], 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 in which they are the recipient and also in which they are not the recipient. For example, it could be expected that being on the receiving end of a defection means that a player is likely to be more hardline, than when they observe a defection against another. However, Roberts limits the study to using a simple experience score.

I would suggest using a model similar to that I described in section 2.3.6, 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.

### 2.3.9 Gossip

Nowak and Sigmund [54] realised the issues surrounding the fact that all individuals being able to view interactions in their simulation framework is not concomitant with real life. As such, they introduced the idea of onlookers as noted in section 2.3.6. 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 [52] suggested using gossip to spread information regarding image scores in their simulation framework.

Sommerfeld *et al.* [70] 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 were 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 resulting behaviour. The findings included that gossip is an effective method of spreading reputation information in an indirect reciprocity system, on fulfilment of some conditions, the first of which is the 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 condition would require an agent communication language (ACL) (such as KQML [19] or FIPA-ACL [55]) and would also require the content of the messages that use this language to be interpretable 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 [54] 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.

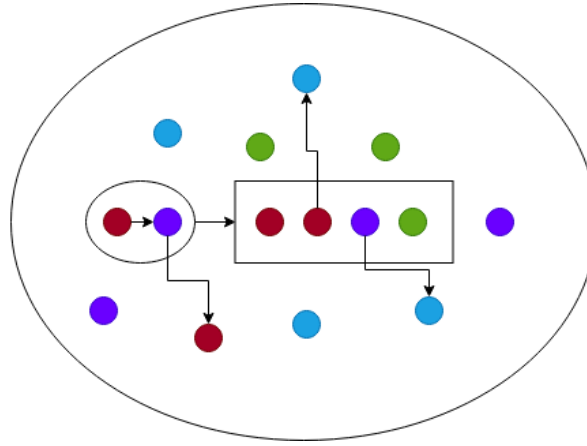


Figure 2.9: The spread of information (shown through the lines) through a population using indirect reciprocity with gossip and onlookers

### 2.3.10 Mui's Computational Models of Trust and Reputation

Mui [47] presented an indirect reciprocity simulation framework similar to Nowak and Sigmund [54] that spreads social information through 'acquaintance networks' to inform a donor's decision. An individual in this framework builds up a network of players whom they meet through interaction, their 'acquaintance network'. If this individual is then a donor in a donor-recipient pairing the individual consults their acquaintance network regarding the recipient. The information gathered from this network helps a donor to decide if they can trust the recipient to reciprocate their cooperation; known as 'collective memory'.

Baumeister *et al.* [12] advocated for the idea that gossip is used for four functions, including strengthening bonds between the gossiper and recipient, enabling the spread of information about the subject (one function for each positive and negative reasons), and helping to 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 four functions. Mui's acquaintance networks support this gossip functionality to a certain extent; however they are limited by the lack of proactivity of gossip. Sommerfeld *et al.* [70] highlighted the willingness of their participants to spread gossip, and the last function of gossip proposed by Baumeister *et al.* [12] (cultural education) suggests that for gossip to be most effective the gossip should be proactively spread.

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

### 2.3.11 Reproduction and Genetic Algorithms

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 [54] and Axelrod and Hamilton [10] (though indirectly their payoff scores do seem to equate to fitness) used a reproduction mechanism inspired by Herbert Spencer's [71] famous phrase "Survival of the fittest".



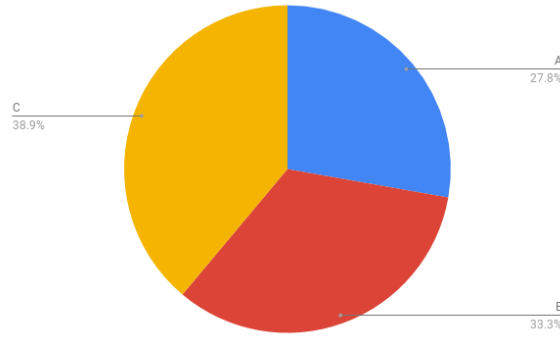


Figure 2.10: The roulette wheel from my proportional selection example

Wooldridge’s [78] 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.

Franco [21] 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 2.10 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 as we did in the proportional selection algorithm we give an example of this in table 2.4.

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 2.4: Tournament selection

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.* [38] presented an efficient version of roulette wheel selection using stochastic acceptance rather than the searching method used by Francq [21]. 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.

### 2.3.12 Summary

I have discussed and described a number of concepts in the field of multi-agent systems, including multi-agent interactions which are the focus of this project. I have also explored game-theoretic techniques that can be used for modelling and simulating agent interactions. This includes the mechanisms of direct and indirect reciprocity as well as a review of a number of past experiments with these mechanism.

In my review I came to the decision that direct reciprocity is not likely to be adequate for facilitating cooperation in multi-agent systems due to the sparse number of interactions between these agents. I have also remarked as to how many past studies have used programs that do not satisfy the multi-agent system paradigm.

I seek to create a system that does match this paradigm, and uses a mixed reciprocity model in which agents can utilise communicative acts to inform others about agents that they have had no past view one. Further to this, I aim to build agents with a variety of possible strategies that are capable of explaining their actions, using a symbolic AI approach. The multi-agent system will be run within a genetic algorithm to select the optimal agents.

## Chapter 3: Framework

### 3.1 Introduction

In the following sections of this chapter I will be discussing the theoretical side of my framework and the subsequent practical implementation of this framework.

### 3.2 Theoretical Framework

#### 3.2.1 Introduction

In this section on the theoretical framework I will be presenting an overall design for my mixed reciprocity model and MAS. I have taken inspiration from past work on game theory (discussed in section 2.3) and techniques for the development of MASs (from section 2.2). Franklin and Graesser [22] stipulated that to describe an agent one must describe five aspects: the environment the agent resides in, the agent's sensing capabilities, the possible actions the agent can take, the drives or primitive motivators for an agent's actions and the action selection architecture for the agent. I shall describe these five aspects in the following subsections here.

#### 3.2.2 The Game and Environment

The environment in which the agents will reside in my system will provide facilities for interaction between the agents. An instance of the environment in my system will be known as a generation. Each generation contains a set of agents. These sets of agents will act as pools to select donor-recipient pairs and onlookers for those pairs.

A generation is contained within a community. These communities contain multiple generations with a strict ordering. A game has distinct timepoints. The number of these timepoints vary according to the number and length of generations - which are variables set by the user at the start. Say these timepoints range from  $1..n$  and there are generations  $1..k$  where  $k < n$  and  $n \% k = 0$ . Then the timepoints are distributed evenly across each generation like this:

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

The environment will work using a cycle in which each step follows the process perceive-decide-execute. Each step of the cycle is a timepoint. The perceive section of a step refers to when all agents receive perceptions generated for that timepoint - mostly from actions in the previous timepoint. All agents then decide on an action and the last part of the step is the execution of these actions in the environment.

The key to the cycle steps is that all agents perceive, after which all agents decide and then all actions are executed. The effect is to synchronise agents steps and prevent 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 selected at random from the

generation's pool of players. For that pair there is a group of onlookers, again randomly selected from the remaining players in the generation's pool.

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 associated strategies (subsection 3.2.7) are selected by the user. For subsequent generations, the set of players is selected using fitness proportionate selection using Lipowski *et al.*'s roulette wheel selection via stochastic acceptance [38] (discussed in subsection 2.3.11).

Each player has a fitness score (the system's performance measure) which is used in the reproduction mechanism. A fitness score starts at zero - it cannot drop below zero - and is affected by the actions of the agent and others as described in subsection 3.2.4. On top of this reproduction algorithm is a chance of mutation  $c$  where  $0 \leq c \leq 1$  which the user is able to set at the start.

One of the aims of this project is to discover successful agent strategies for the mixed reciprocity model. This problem can be formulated as an optimisation problem, which can be solved using genetic algorithms. The effect of using fitness proportionate selection and the generations of players is to create a genetic algorithm. When experimenting with agent strategies, the strategies which were most successful will hopefully become obvious at the end of a community's generations by the concentration of agents using those strategies.

In summary, the components of this environment are the community, the 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. For a discussion on the properties of my environment see appendix subsection 8.2.1.

### 3.2.3 Sensors

Percepts are the information received at an agent's sensors from it's environment. In my system they are transmitted to each agent in the first stage of each cycle step. In my system, percepts are a direct observation of an interaction, hearing gossip from another agent or sensing whether they are the donor or recipient in a donor-recipient pair.

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

For each interaction, there is a set of onlookers selected at random from the generation's pool of players (not including the recipient or donor). In the cycle step after the interaction takes place the onlookers and the recipient receive percepts containing information as to who the donor and recipient were and the action the donor decided on.

In the gossip (3.2.5) and action (3.2.4) subsections, I discuss an agent's ability to act by gossiping to another agent. The action of gossip produces a percept. This percept contains the message sent using the Simple Agent Gossip Language. The percept is received by the agent given as the recipient by the gossipier.

### 3.2.4 Actuators

My system focuses on interactions between agents and as such, agents require a number of action possibilities by which to interact with one another. To simplify the action there are 3 possible main actions: idle actions, any action when the agent is a donor and gossip actions. Idle actions are simple: an agent is idle in that timepoint, their action has no effect on the environment or other agents, except for through inaction.

Actions when an agent is a donor are 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 that they are a donor, they have no choice but to commit one of the following two actions: cooperate or defect. In this interaction where the actor is the donor, the recipient has no control over what happens.

The payoff matrix, taken from Nowak and Sigmund [54] in table 2.3, stipulates that when a defect action is chosen, there is no effect on either the donor's or the recipient's fitness. When a cooperate action is chosen however, there is a cost of 1 to the donor's fitness and a benefit of 2 to the recipient's fitness. As discussed in subsection 3.2.3, this action also produces percepts of the interaction event.

When an agent is not a donor, they cannot choose to cooperate or defect with anyone. However, they can choose one of the other two actions: idle or gossip. A gossip action is another type of interaction between agents. A gossipier chooses to communicate with another agent. The contents and structure of this communication are detailed in subsection 3.2.5 and the only effect is the percept discussed in subsection 3.2.3.

### 3.2.5 Gossip and Simple Agent Gossip Language

```
{
  "recipient": 1,
  "about": 3,
  "gossiper": 4,
  "timepoint": 7,
  "gossip": "positive"
}
```

Source Code 3.1: Example of a message in SAGL

With the design of the gossip in my system, I have focused on social agency over mental agency and keeping the layout of the gossip simple. KQML and FIPA-ACL are powerful tools, which I do not need, and as such I have decided to create my own language which I have named the Simple Agent Gossip Language (SAGL).

I have attempted to create a language which may be used by agents to facilitate three out of four of Baumeister *et al.*'s [12] functions - excluding cultural education, which is beyond the scope of my project. The language I propose is very simple. There are five fields. Three of the fields are identifiers (one for each of the recipient, the target and the gossipier). Another is the timepoint at which the gossip action was decided upon. The final field is the gossip.

Sommerfeld *et al.* [70] categorised gossip in their experiment as either positive or negative. Individuals in my system are attempting to convey similar reputation information, and as such, the gossip field of SAGL either contains the keyword positive or negative.

Agents can attempt to improve bonds between themselves and others, by spreading positive gossip about themselves. Agents can also spread negative information about others to either warn other agents of a non-trustworthy agent, or to harm the agent the gossip is about. Agents can also send positive gossip about others to spread knowledge of trustworthy agents in the system.

### 3.2.6 Agent Architecture

Russell and Norvig's [63] model-based reflex agent is the closest of their agent architectures to the architecture I will be using. The architecture includes using a trust model to revise beliefs from received percepts and then using these beliefs with teleo-reactive programs to make action decisions.

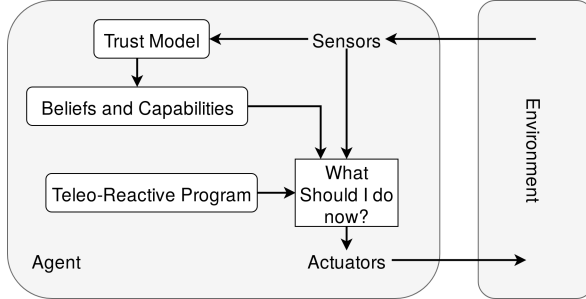


Figure 3.1: The agent architecture in my system.

The trust model and beliefs system use the multi-valued fluent cached events calculus. The beliefs are a combination of the value and the fluent of multi-valued fluents [8]. For example the fluent  $standing(agent1, agent2, t) = value$  is the belief of *agent1* on the standing of *agent2* at timepoint *t*. In this fluent the value can either be 'good' or 'bad'. The trust model controls the interpreting of percepts and updating of these fluents.

Capabilities are also a part of the beliefs system that are revised during the interpretation of percepts. These capability beliefs constrain an agent as to the actions to which they can commit i.e. an agent can only commit to an action if they believe that they are capable of it. For example, at a time when an agent is a donor in a donor-recipient pair, they will only believe themselves capable of a donor action, constraining their action choice.

To make a decision on an action the agent architecture combines the agent's beliefs with a teleo-reactive program that makes up an agent's strategy. A strategy is made up of strategy 'components'. Of which one I have already discussed: the trust model. The trust model is combined with a component determining how to act as a donor and another determining how to act when not a donor. The decision is a commitment at that timepoint to carry out that action.

### 3.2.7 Strategies and Trust Models

As I have discussed, an agent's strategy is made up of the trust model, the strategy component for when the agent is a donor and the strategy component for the agent when it is not a donor. These strategies have been inspired by past work on indirect reciprocity and expanded upon. For example, I have enhanced agent's interpretation of events in order that they respond differently when they are the recipient of an action from a donor and when they are the recipient of gossip. The action decisions made when agents are donors have generally remained the same as those in past work. However, these action decision systems have been augmented with new strategy components for when an agent is not a donor.

#### Defector and Cooperator

Franklin and Graesser [22] talked about an agent's primitive motivator or drive. Each strategy has a drive or a combination of drives. For example, the 'Defector' strategy seeks to protect its own interests by always defecting, thus not causing it to lose fitness points. The defector strategy has been augmented with three possible strategy components which it can use when it is not a donor: 'Lazy', 'Spread Negative' and 'Promote Self'.

The 'Lazy' strategy component always decides on being idle. The 'Spread Negative' component randomly spreads negative gossip about other agents in the system. This compo-

agent simulates what might be known as ‘fake news’: a form of deception and manipulation. The final component is ‘Promote Self’ in which the agent always spreads positive gossip about itself, aiming to increase its reputation by spreading misinformation and encouraging others to cooperate with the agent using it.

The strategy ‘Cooperator’ aims to maximise social welfare by always cooperating as a donor. I have also augmented this strategy with three components when the agent is not acting as a donor: ‘Lazy’, ‘Promote Self’ and ‘Spread Positive’. ‘Lazy’ is the same as for the defector strategy and so too is ‘Promote Self’, although ‘Promote Self’ when used by a ‘Cooperator’ could be seen as coming from a drive to improve cooperation and not to deceive other agents. The third is similar to ‘Spread Negative’ but instead spreads randomly positive gossip in order to improve cooperation between agents in the system. This strategy component could result in accidental manipulation to cooperate with non-reciprocating agents by naive agents.

### Image Scoring Discriminator

Neither ‘Defector’ nor ‘Cooperator’ have any beliefs about other agents and as such are limited to only trust models that constrain their capabilities. A more interesting strategy is the ‘Image Scoring Discriminator’. Built up around Nowak and Sigmund’s [54] discriminator, strategies using this component keep beliefs about the image score of other agents in the system bounded by  $-5$  and  $5$ . To control these image score beliefs I have created three trust models for these discriminators: ‘Naive Trusting’, ‘Trusting’ and ‘Distrusting’.

The ‘Distrusting’ trust model abjectly ignores any received percepts which it has not observed directly i.e. rejects any gossip. At the other end of the spectrum, ‘Naive Trusting’ accepts any gossip it receives, increasing the image score of the agent the gossip is about by 1 if the gossip is positive, and decreasing by 1 if the gossip is negative. The ‘Trusting’ model is less black and white. This trust model will only change the beliefs regarding the target of the gossip if the gossip is trusted i.e. has an image score greater than or equal to  $k$ , where  $k$  is an integer variable set at the start.

The ‘Image Scoring Discriminator’ strategy component has four corresponding non-donor strategy components. These include ‘Lazy’ and ‘Promote Self’. Again, it could be said that ‘Promote Self’ is used to improve cooperation towards the agent and not to deceive the recipient.

The other two components are ‘Spread Accurate Positive’ and ‘Spread Accurate Negative’. ‘Spread Accurate Positive’ spreads positive gossip about trustworthy agents to other trustworthy agents. Trustworthy means that their image score is greater than or equal to  $k$ . The aim of this positive gossip is to improve cooperation between trustworthy agents. ‘Spread Accurate Negative’ spreads negative gossip about agents believed to be untrustworthy - agents with an image score less than  $k$  - to trustworthy agents.

It is also possible with agents using image scores to select agents with the ‘Personal Grievance’ option. When this option is selected, agents using the strategy increase or decrease image scores by 2 when respectively receiving a cooperation or defection. These agents use the selected trust model when they are not the recipient of these actions.

### Standing Discriminator

The standing strategy is also included in my system under the name ‘Standing Discriminator’ with the trust models: ‘Naive Trusting’, ‘Trusting’ and ‘Distrusting’. This strategy has

the following strategy components for when the agent is not a donor: ‘Lazy’, ‘Promote Self’, ‘Spread Accurate Positive’ and ‘Spread Accurate Negative’. These are the same as for the ‘Image Scoring Discriminator’ strategy components except an agent is said to be trustworthy if their standing is ‘good’.

The original standing strategy specification for when to change an agent’s standing between ‘good’ and ‘bad’ is used for directly observed events. The ‘Personal Grievance’ option is not selectable for the standing strategy.

### 3.2.8 The Veritability Discerner

The image scoring and standing strategies were both built for systems in which gossip was not present and the reliability of the information they received did not require questioning. I have added trust models to compensate for these issues. However, these trust models are built to either accept or to reject the information from the percept and I do not believe that this procedure reflects how reputation information is interpreted in the real world.

Thus, I have developed a new strategy with corresponding trust models and strategy components. I have named this strategy the ‘Veritability Discerner’. An agent using the strategy holds two beliefs about all other agents. One is the number of percepts the agent has received about the other agent  $n$  (when the other agent was a donor or the target of gossip). The second, is an integer value I have named the ‘veritability rating’  $v$ . This rating is initially 0. This rating is similar to an image score but it is not bounded and does not simply increase or decrease by 1, the addition or subtraction to the veritability rating is weighted by the reliability and severity of the percepts.

#### Trust Models

This strategy comes with three trust models: ‘Strong Reactor’, ‘Balanced Reactor’ and ‘Forgiving Reactor’. All use the same central weighting and reaction system but apply slightly different weights on top of these.

The central weighting and reaction system is as follows. Viewing of a cooperation is the most reliable indication that the donor can be trusted, and so has the highest weight (20) added to the veritability rating of the agent. Conversely viewing a defection against a recipient who is a trusted agent is the most reliable indication that the donor is untrustworthy, and so the highest weight (20) is subtracted from the veritability rating. When either of these interaction events are observed, the count of percepts received about the donor is incremented by 1. Similarly to the standing strategy, a defection against an untrusted agent has no effect on the veritability rating of the agent but does increment the percept count.

Negative gossip about an agent from a trusted source subtracts 10 from the veritability rating of that agent, whilst positive gossip from a trusted source adds 10 to the rating. If gossip is from an untrusted source, then only 1 is respectively added or subtracted from the rating.

The ‘Strong Reactor’ trust model multiplies weights for negative percepts by 2. This multiplication makes the weight for defection against a trusted recipient 40. The multiplication also makes negative gossip from a trusted source subtract 20 from the veritability rating of the target of the gossip. The last effect the multiplication has is to cause a subtraction of 2 from the veritability rating of the target of gossip in response to negative gossip from an untrusted source.



The ‘Balanced Reactor’ does not apply a weight to either positive or negative percepts. Finally the ‘Forgiving Reactor’ multiplies positive weights by 2, making the weight of viewing a cooperation 40, positive gossip from a trusted source 20 and positive gossip from an untrusted source 2.

## Trustworthiness

A third belief can be derived from the veritability rating and count of percepts received: trustworthiness. An agent using the strategy has a value  $k$  similar to an ‘Image Scoring Discriminator’ agent, though the possible values for this are -10, -5, 0, 5 and 10. The trustworthiness belief is derived as follows. Divide the veritability rating by the count of the percepts received  $v/n = t$  - if  $n$  is 0  $t$  defaults to 0. If  $t \geq k$ , then the agent is trustworthy, otherwise they are untrustworthy.

$$\begin{aligned} & \text{holds\_at}(\text{veritability\_rating}(\text{Perceiver}, \text{Subject}) = V, \text{Timepoint}) \wedge \\ & \text{holds\_at}(\text{percept\_count}(\text{Perceiver}, \text{Subject}) = N, \text{Timepoint}) \wedge \\ & \text{get\_K}(\text{Perceiver}, K) \wedge V/N \geq K \rightarrow \\ & \text{trustworthy}(\text{Perceiver}, \text{Subject}, \text{Timepoint}) \end{aligned}$$

The derivation of this belief has the same effect as taking the mean of a list of weighted actions. This mean takes into account all the actions committed to by an agent but smooths out noise from possibly untrustworthy or distorted sources.

The trustworthiness belief about an agent is used when interpreting gossip and when deciding whether another agent’s defection was justified or not. This belief is also used for when the agent is a donor. If the recipient is considered trustworthy then the agent will cooperate, if they are not, the agent will defect.

$$\begin{aligned} & \text{get\_strategy}(\text{Actor}, \text{“VeritabilityDiscerner”}) \wedge \\ & \text{interaction\_pair}(\text{Actor}, \text{Recipient}, \text{Timepoint}) \wedge \\ & \text{trustworthy}(\text{Actor}, \text{Recipient}, \text{Timepoint}) \rightarrow \\ & \text{action\_commitment}(\text{cooperate}, \text{Actor}, \text{Recipient}, \text{Timepoint}) \end{aligned}$$

## Strategy Components

Finally we need strategy components for when an agent using the strategy is not a donor. These are similar to the discriminator strategies and include: ‘Lazy’, ‘Promote Self’, ‘Spread Positive Trusted’ and ‘Spread Negative Untrusted’. ‘Lazy’ and ‘Promote Self’ do the same as with all other strategies. ‘Spread Positive Trusted’ spreads positive gossip to trusted agents about trusted agents. ‘Spread Negative Untrusted’ spreads negative gossip to trusted agents about untrusted agents.

‘Spread Positive Trusted’ formalisation:

$$\begin{aligned} & \text{get\_strategy\_component}(\text{Actor}, \text{“Spread Positive Trusted”}) \wedge \\ & \text{not interaction\_pair}(\text{Actor}, \text{Timepoint}) \wedge \\ & \text{findall}(\text{Trusted}, \text{trustworthy}(\text{Actor}, \text{Trusted}, \text{Timepoint}), \text{TrustedAgents}) \wedge \end{aligned}$$

$$\begin{aligned}
& \text{len}(\text{TrustedAgents}, \text{Len}) \wedge \text{Len} \geq 2 \wedge \\
& \text{get\_2\_random}(\text{TrustedAgents}, \text{TrustedAgent}, \text{Recipient}) \rightarrow \\
& \quad \text{action\_commitment}(\text{gossip}(\{\text{recipient} : \text{Recipient}, \text{about} : \text{TrustedAgent}, \text{gossiper} : \\
& \quad \quad \text{Actor}, \text{timepoint} : \text{Timepoint}, \text{gossip} : \text{positive}\}), \text{Actor}, \text{Recipient}, \text{Timepoint}) \\
& \\
& \text{get\_strategy\_component}(\text{Actor}, \text{"Spread Positive Trusted"}) \rightarrow \\
& \quad \text{action\_commitment}(\text{idle}, \text{Actor}, \text{Recipient}, \text{Timepoint})
\end{aligned}$$

‘Spread Negative Untrusted’ formalisation:

$$\begin{aligned}
& \text{get\_strategy\_component}(\text{Actor}, \text{"Spread Negative Untrusted"}) \wedge \\
& \text{not interaction\_pair}(\text{Actor}, \text{Timepoint}) \wedge \\
& \text{findall}(\text{Trusted}, \text{trustworthy}(\text{Actor}, \text{Trusted}, \text{Timepoint}), \text{TrustedAgents}) \wedge \\
& \text{len}(\text{TrustedAgents}, \text{Len}) \wedge \text{Len} \geq 2 \wedge \\
& \text{get\_random}(\text{TrustedAgents}, \text{Recipient}) \rightarrow \\
& \quad \text{action\_commitment}(\text{gossip}(\{\text{recipient} : \text{Recipient}, \text{about} : \text{TrustedAgent}, \text{gossiper} : \\
& \quad \quad \text{Actor}, \text{timepoint} : \text{Timepoint}, \text{gossip} : \text{positive}\}), \text{Actor}, \text{Recipient}, \text{Timepoint}) \\
& \\
& \text{get\_strategy\_component}(\text{Actor}, \text{"Spread Negative Untrusted"}) \rightarrow \\
& \quad \text{action\_commitment}(\text{idle}, \text{Actor}, \text{Recipient}, \text{Timepoint})
\end{aligned}$$

## Summary

I have presented a new strategy for use in indirect and direct reciprocity models. The aim of this strategy is to be less susceptible to deception and manipulation by weighting the increase and decrease of the ratings by the severity and reliability of the percepts and also by using the smoothing properties of the arithmetic mean.

### 3.2.9 Summary

In summary, I have presented a MAS that uses a variant of a genetic algorithm to reproduce agents between instances of the environment. I have described the components of the environment and game, the percepts that can be received by agents, the actions available to agents and how these are constrained and the simple agent gossip language. I have also discussed the architecture that I am planning to use in the system and the strategies and trust models available to associate with the agents, including a new strategy that aims to be more suitable with a system in which deception and manipulation are possible.

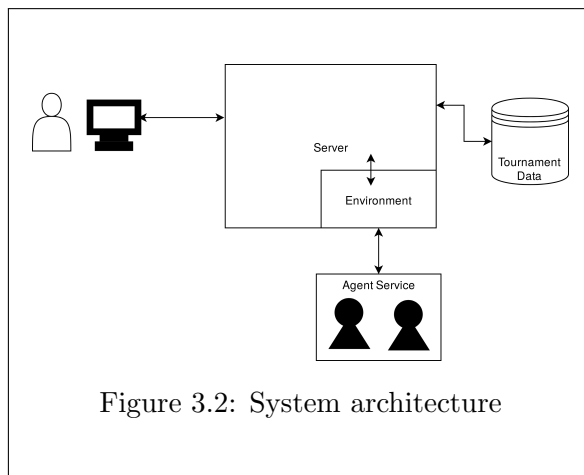
## 3.3 Implementation

### 3.3.1 Introduction

In this section I will discuss the details of the implementation of the theoretical framework outlined above as well as the architecture of the system that is hosting it. I will also be discussing the development of that system and the software engineering techniques, tools and processes I have employed for this development.

### 3.3.2 Development and Software Engineering

#### Architecture



One of the objectives of my project is to create a system that is distributable, in order for it to be an experimentation and learning tool for those interested in studying multi-agent interactions. This distributability is reflected in the architecture I have used to develop my system. A web application provides the client side interface for users across the internet. This interface allows users to create, run and view analysis of the available games.

I named the web application ‘The Nature Engine’ and it is of typical web architecture: the client-server model. Built into this web application is an implementation of my theoretical framework which runs the environment and the agent’s bodies. Notably this part of the system doesn’t include agent decision making components. These components are hosted in a separate web service, that has been named the Agent Mind as a Service (AMAAS). This web service exposes an API for the web application to interact with to facilitate sending percepts and getting decisions.

The final component of the system is a database for the long term storage of data about users of the site (to enable users to record and label experiments) and game data. This is a relational database that due to the use of an ORM (SQLAlchemy) I have been able to create as both a SQLite database for development and a MySQL database for deployment. An illustration of the overall architecture is in figure 3.2. The system is deployed on an Ubuntu server and the database is situated on another machine.

#### Proof of Concepts and Learning

At the very beginning of this project I had a good amount of software development experience, some experience with database development and logic programming, and less experience with web development. Before undertaking the development I decided I needed to learn more about some of these areas (especially web development).

I had decided that I wanted to extend my experience with Python due to its nice language features, good library support and it being often overlooked for the development of MASs (in comparison to Java and Prolog). As such I decided to use the Flask web



Figure 3.3: Two screenshots of two pages of the Prolog web service proof of concept

development micro-framework and partake in Miguel Grinberg’s course “The Flask Mega-Tutorial” [1]. Flask is a micro-framework as it only includes a minimal amount of functionality in comparison to frameworks such as Django. Nevertheless Flask is extensible and there are a lot of libraries a programmer can use to bring in required functionality.

Using inspiration and the techniques learnt in this course I developed an early prototype of The Nature Engine. This prototype provided a number of web pages and navigation between them including a home page, pages to set up and run two player iterated prisoner’s dilemma games and pages to set up and run tournaments of the iterated prisoner’s dilemma and the respective analysis of both game types. Both of these games are run in the backend using the Axelrod-Python library [18]. This early prototype also included a database set up, a redis queue to run the tournaments, a front-end that used bootstrap for styling and React.js for a dynamic form to set up tournaments.

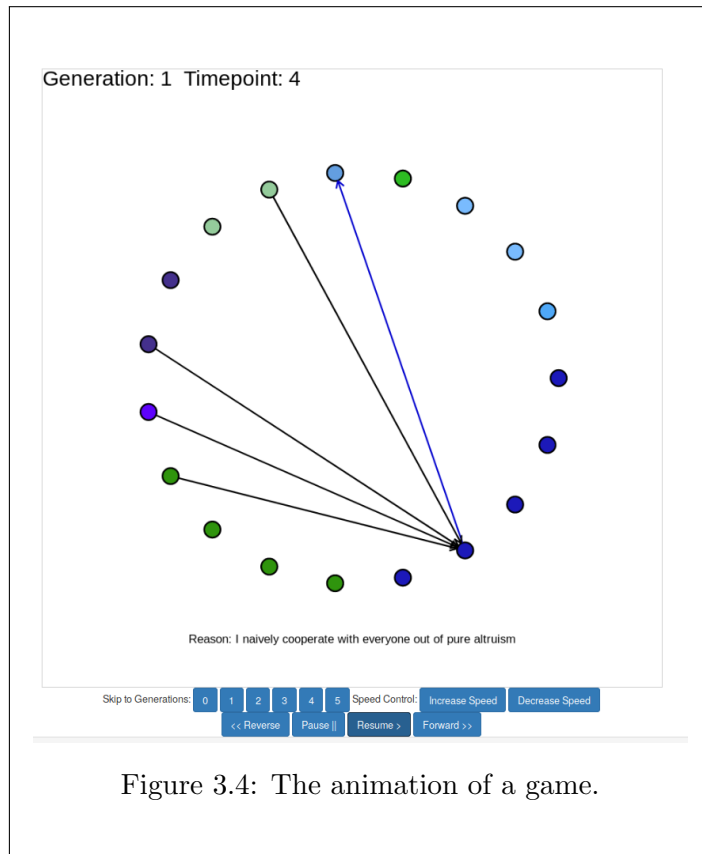
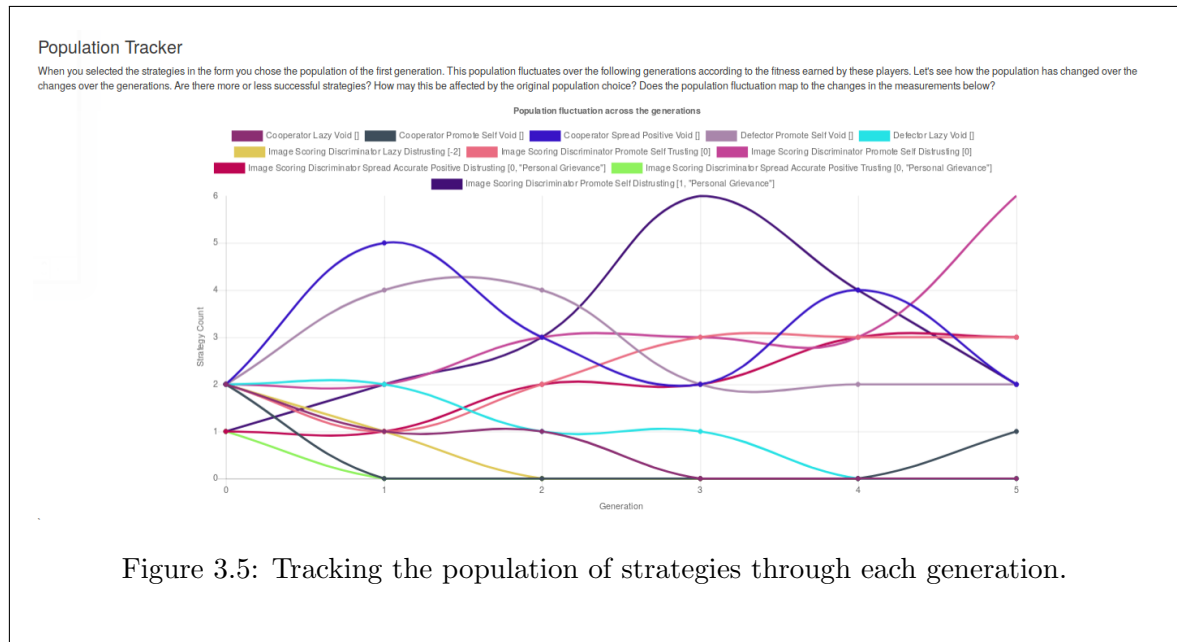
Prolog is a logic-language traditionally used for the development of symbolic AI systems. The language allows a programmer to elegantly lay out how an agent makes decisions and interprets percepts. There are limited libraries for interfacing between Prolog and Python, and creating a method for interfacing between them reliably would have been a project within itself. However, Prolog has a good HTTP support and as such I saw an opportunity to further hone my web skills by developing a web service to host agent minds.

To develop these skills I used Anne Ogborn’s ‘Creating Web Applications in SWI-Prolog’ tutorial [2] and the swi-prolog documentation. Using this tutorial I created a basic web application that imitates managing the creation of strategies and agents within the system. This was far simpler than The Nature Engine prototype (figure 3.3) as it did not build up to a final application (the development of AMAAS used skills learnt from this proof of concept, but was developed separately).

## The Nature Engine

I built on top of the early prototype of The Nature Engine to develop the final web application. This included improving the front end content, adding a section to run the theoretical framework I have devised and adding a couple of extra features.

To improve the front end content I worked on improving the form that allows users to dynamically select and remove agents from a tournament. This also translated into the form that allows users to do the same for my game. Other improvements were the use of a bootstrap carousel for the homepage to prevent text overload on the screen, better text descriptions on the website and improved game and tournament analysis.



The main bulk of the development was implementing the theoretical framework. However, to run this game in the web application required a front-end, processing of user requests and storing of data in the database. This development is a vertical slice of the web application. The process is as follows. A user sets up a game on the set up page using a React.js form (decisions are informed by strategy information in an information box), this form is submitted which sends a post request with the relevant information to the server.

The server then puts a method to create and run the game in a redis queue and then store the resulting data in the database. The user is redirected to a waiting page, which either notifies them of a timeout on their game or sends

them through to the analysis page. The analysis page includes a run down of the statistics in terms of cooperation, gossip and the success of each strategy and agent (part of this is illustrated in figure 3.5). A visualisation of the game that has been run is also included in an animation on the analysis page (figure 3.4).

In addition to this I have created a feature that allows users to create accounts, login and associate games of my theoretical framework to their account. These games can be given a label so they can keep track of the experiments the have run (illustrated in figure 3.6). User data is recorded in the database and passwords are hashed to keep them secret.

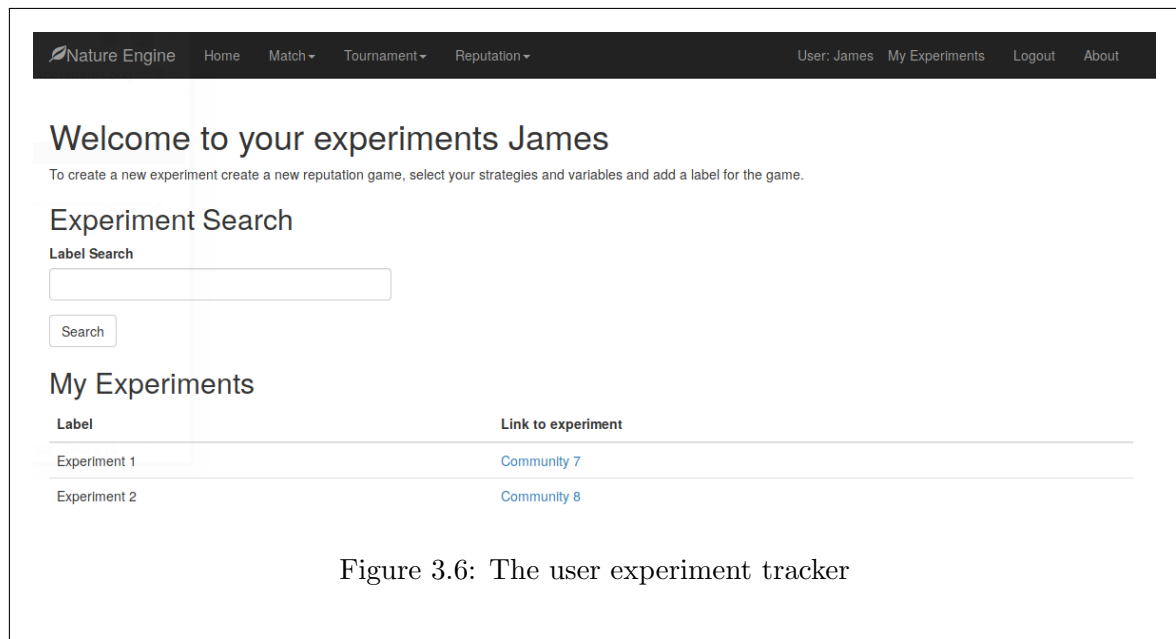


Figure 3.6: The user experiment tracker

The web application, database and AMAAS system has been deployed to an Ubuntu Server and has a domain name (natureengine.tech) that is accessible to all across the internet.

### 3.3.3 Framework

#### Introduction

In this section I will lay out the implementation of the environment, agents and strategies I have described in the theoretical framework (section 3.2).

#### Environment and Agent's Bodies

The environment and agent's bodies from the theoretical framework described in section 3.2 has been implemented within the nature engine web application in it's own package. As such the environment and agent's bodies components have been written in Python.

This has taken an object oriented approach. A Community object is the overarching object that simulates the specified number of generations and handles the reproduction of strategies between generations assigning the strategies that should be associated with players in a Generation. A Generation object creates a set of players from these passed strategies.

The Generation is then simulated. Simulation of a generation involves running the synchronised cycle steps from the start timepoint of that generation to the end timepoint. In each perceive step a donor-recipient pair is selected at random by a method of the Generation object and the relevant agents are notified.

In the previous timepoint a set of percepts is generated - from the actions of that timepoint - for the current timepoint and are set into the perception bank ready to be transmitted at the current timepoint. These are sent in the perceive step. Agents are then asked to decide on an action at the current timepoint. Once these actions have been returned to the Generation they are executed, affecting players fitness and generating more percepts.

A Player is also an object in the system, and is associated with a strategy that describes how it should act at each timepoint. The Player object handles the sending of percepts and getting of actions from their decision making component in the AMAAS system. The Player object also encapsulates the player's state including the fitness of that player.

The 3 classes that implement the Action abstract class represent the three possible types of action that are available and the relevant data associated with an instance of an action.

### Agent Mind as a Service (AMAAS)

The AMAAS system was developed in Prolog. This system handles the management of communities, generations and agents by assigning them with an ID generated using a similar system to the Prolog gensym library. These are assigned to the predicates *community(CommunityID)*, *generation(Community, GenerationID)* where *Community* unifies with a community and *agent(Strategy, Community, Generation, AgentID)* where *Strategy* unifies with a strategy in the system and *Generation* with a generation.

the *community/1*, *generation/2* and *agent/3* predicates are dynamic and can be asserted and retracted. The possible strategies that unify with *Strategy* are not dynamic and are defined with a strategy predicate (*strategy/5*). The first 4 arguments are strings and are in the correct order: the strategy to use when a donor, the strategy to use when not a donor, the trust model, the description of the strategy. The final argument is a list of any type that contains options to augment the strategy that are not a necessity to define for every strategy.

Some of the available strategies (any associated with the donor strategies “Standing Discriminator”, “Image Scoring Discriminator” and “Veritability Discerner”) hold beliefs about other agents and the environment. These beliefs are stored with fluents in the MVFCEC. The value of these fluents are revised by the definition for *initiates\_at/4* and *terminates\_at/4* of that agents strategy and trust model and the events and conditions that cause these predicates to fire.

These predicates are caused to fire by the receiving of percepts to the agents service that then asserts a new event observation (the *observed\_at/2* predicate) and calling an update on the MVFCEC using the *update\_at/2* predicate. The process of firing these is controlled for all possible percepts by the *add\_new\_action\_gossip\_percept/2*, *add\_new\_action\_interaction\_percept/2* and *add\_new\_interaction\_percept/2* predicates.

The beliefs formed by the firing of these predicates and the initiating and terminating of fluents is then used by the teleo-reactive program created for each strategy to decide on actions at certain points in time. An agent decides on an action by the firing of the *agent\_action/6* predicate. A set of *agent\_action/6* definitions is given for each donor and non-donor strategy. The first four arguments are input arguments being the timepoint to decide at, the id of the community and the generation the agent belongs to and finally the id of the agent.

The second to last argument stores true if a decision has been successfully made or provides an error message if it has not been. The last argument unifies with an action (the dictionary representation of that action). The program is teleo-reactive as the system is designed so only one of the bodies of the set of *agent\_action/6* definitions satisfies the conditions required, and this forces unification with one action.

The higher level predicate *get\_action/6* checks if an agent has already committed to an action at a specific timepoint. This predicate also ensures that an agent only commits to actions that it believes it is capable of using *capable/5* predicate, which is only satisfied if the agent is committing to a donor action as a donor or a gossip or idle action if it is not a donor.

## Simple Agent Gossip Language

The specification of SAGL (subsection 3.2.5) is remarkably similar to data structures used in Python (Dicts), the JSON structure and dictionaries in Prolog. As such, it made sense to implement SAGL by storing it as Python Dicts in the environment and agent's bodies, transferring it with JSON (as is used for the API in general) and then storing it as dictionaries in Prolog.

## 3.3.4 Software Engineering

### Introduction

In this section I will discuss and reflect on engineering tools, techniques, methodologies and approaches that I have used, and some that I possibly should have used.

### Process

The development process I decided on when starting the project was a scrum like structure, with a task backlog that was subject to change and fortnightly sprints. Although I have mostly stuck to these sprints, the backlog was in constant flux, far more than was expected. This is likely due to me still learning lots about the topic during the project. This was not due to a lack of research, but a focus on game-theory over multi-agent system design.

I feel that I could have managed this better with a more structured approach to design of the system, and by opening up my focus to all areas of the topic. However, I believe this structure allowed me to be open to new ideas and aspects to add to the project. I was then able to incorporate aspects of multi-agent systems and game-theory fields I may have not been able to with a more structured approach.

As it was, the design phase of the process was greatly interleaved with both learning the concepts for designing the system and the implementation of it. The implementation of the system then followed a Test-Driven Development (TDD) approach, developing tests and then creating code to work for those tests, and afterwards refactoring. This followed with a deeper more fully encompassing testing using a bottom-up integration testing strategy.

This highly interleaved process worked quite well for my project as it allowed me to design, develop and test features in the system separately all in the same slice of tasks. This process also allowed me to discover new features, ideas and strategies to incorporate in the system and then extend the project by adding a new slice of tasks to the task backlog which I could then choose when to implement based on urgency and importance.

I followed a bottom-up strategy when developing the components of the system. The multi-agent system is dependent on the agents service being fully functional and safe to use. As such I focused on developing and testing the AMAAS system and a base set of strategies for this first.

From this I then developed and tested a base set of components for the environment of the MAS. Once I had a safe and working base system I incorporated the MAS into the web application. From there I worked on adding new strategies and testing them in the AMAAS system, and then onto adding features to the multi-agent system such as the selection of onlookers and the reproduction mechanism.

This incremental development using slices of tasks for each feature allowed me to be



secure that the developed parts of the system were safe for use by other parts of the system.

## Design

Most of the early design of my system went into learning how the operation of a MAS works, and drawing from multiple sources how to implement it myself. Further to this, I researched agent architectures and fused together multiple styles of design into my own. Both the design of the environment and agents can be viewed in my theoretical framework section (section 3.2).

I designed the API for the AMAAS system using a RAML specification and then added to this as I went along. This drove the design and implementation of the AMAAS system, by designing what I needed from the service I derived what I needed to implement for the functionality. I also created a theoretical design for the strategies in the system using logical predicates, to translate into Prolog.

The environment and agent's bodies for the multi-agent system were developed using the OOP features of Python. To approach this I used object-oriented design UML design tools (UMLet), and added to this as I enhanced the features of the system. This included a UML class diagram for designing the environment and agent body class structure (figure 3.7) and a sequence diagram to portray the message passing sequence (figure 3.8).

In the design of the environment classes I included an observer design pattern to observe player states and collect data on the player and their actions throughout the course of the game. I then created a facade for this so a simple Results class is returned after an object of the ReputationGame type has been run. The ReputationGame class is also a facade on the simulation of the communities and generations, making the system easy to set up a game with (preventing the need to know the underlying implementation).

This project is not focused on interface design and I have not researched or learnt any strategies for user experience design. However, with the knowledge I have of bootstrap components, css and some web development I developed templates for the client side as an informal method of design. I then refined these templates to improve the user experience simply by using the system and noticing any annoyances or issues I had with it.

## Development

As I have mentioned the development phase has been strongly interleaved with testing and design to ensure the safety and correct running of the program. This included a TDD approach to development. I used a number of tools to assist me in this development.

The Python code that I have written follows the PEP 8 Style Guide. I used the Flask web application microframework to assist me in developing the web application and for the creation of the database I used the SQLAlchemy ORM. With SQLAlchemy way I was able to generate for both SQLite (for development) and MySQL (for deployment) databases from Python code. The use of SQLAlchemy also assists in the prevention of SQL injection attacks that websites so often fall prey to.

I also used git as a version control system tool, with a master branch and various feature branches which I would develop features on and then once stable merge back into the master. The repository is hosted on GitHub. I mostly used the git command line tool for Linux, however I occasionally used the git support in the PyCharm integrated project support environment.

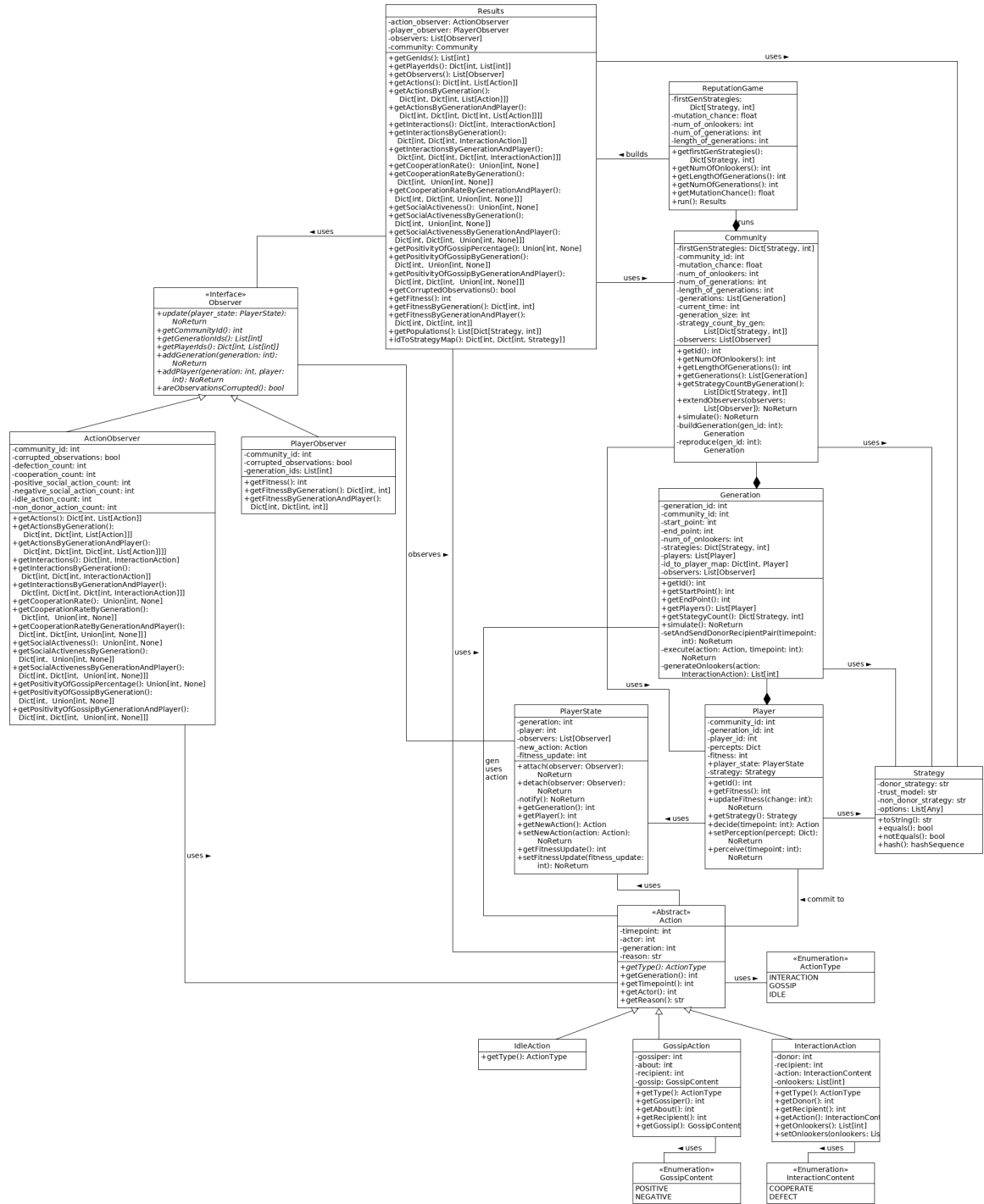


Figure 3.7: The class diagram for my environment and agent bodies Python program, also can be found in the ProjectReports/NewFinalReport directory as EnvClass.png

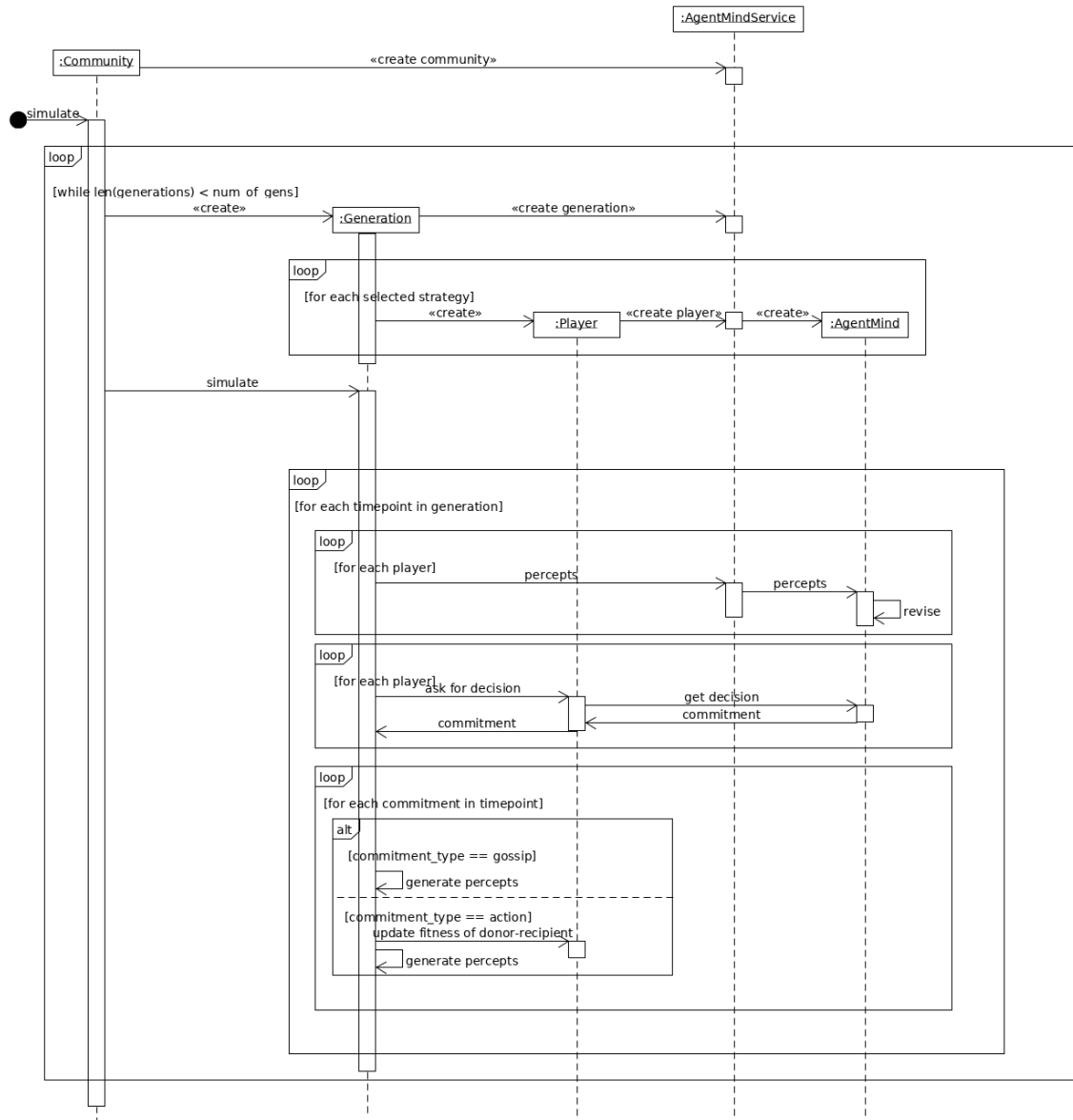


Figure 3.8: The sequence diagram for my environment and agent bodies Python program, also can be found in the ProjectReports/NewFinalReport directory as EnvSequence.png

Pycharm provides more than just version control system tools. There is also a linter and support for the Python language, as well as Jinja2, HTML, CSS and Javascript which I used for front-end development. I have often used the visual debugger, Python environment setup support, Flask support and unittest library support.

There is far less tooling available for Prolog development. I ended up using a mix of Visual Studio Code as it has a linter that helped me catch errors earlier (though I would not recommend this, as it often crashes) and Sublime Text (limited to syntax highlighting).

Again with the development of the AMAAS system I followed a TDD approach, writing unit tests using the PLUnit library. I documented this code using the PLDoc library. There is a good HTTP library for Prolog, that allowed me to easily create a web service with handlers for certain routes that easily link up to functionality in the AMAAS system.

## Testing

The testing of my system can be broken down into four parts: AMAAS bottom up integration testing, environment bottom up integration testing, application programming interface (API) testing and some extra early unit testing on other parts of the web application. The bulk of this testing comes from the first two, which were interleaved with the development of the system. The early unit testing was also interleaved with development. However API testing was carried out after development.

As the AMAAS system has been developed in Prolog, the unit testing has been done with the PLUnit library. This focused on testing the functionality provided by the predicates I have created in the system. In depth testing has been undertaken for every strategy, in many different possible situations and timelines. I have also tested all predicates for managing agents, generations and strategies in the system as well as the predicates used for receiving percepts and getting actions.

Further to this the AMAAS system has been web API tested. The functionality that the API uses in the system has been thoroughly tested in the bottom up integration testing. As such, API testing focused on ensuring the API met the specification laid out in the documentation. I carried out this testing using the python unittest testing framework and the requests library. I had previously planned on using an API testing tool such as postman, but had found no free and easy way to put my tests into code to put on my github repository.

I also used this testing framework for bottom up integration testing of the environment for my game. This testing also requires the AMAAS system to be running and working correctly so acts as integration testing between the two systems. The testing was interleaved with development in a way that as I created each unit (class) or added functionality to that unit I added testing for the functionality added. I used the same framework for some earlier testing in the web app. This was mostly testing for converting produced game data for the database.

I would also like to have tested the web application including functionality, usability and compatibility testing [3, 4]. I have developed the user interface to what I believe to be a good standard, but I'm sure it can be improved and it would be good to get feedback to understand where I can improve this. I also have developed the web application so it is responsive using the Bootstrap 3 library, and have used the website on Firefox on Android and Linux, and Chrome on Windows. It would be good to find out the compatibility with Safari and have more eyes to spot bugs in the functionality.

## Documentation

In the API testing phase I tested the API against the claimed functionality in the API documentation. I created the specification for this documentation in the RESTful API Modelling Language (RAML) which is based on YAML specifically for the creation of RESTful API specifications. There is a good set of tools surrounding RAML, including a linter in the Atom text editor and the raml2html tool that creates a nicely laid out web page for a user readable documentation. I used this tool to convert my .raml file into a .html for readability. This documentation can be found by opening the main.html document in the AgentsService/api\_docs directory in a web browser.

I have also created documentation for the web application code. This is done in the form of writing python docstrings in the reStructuredText markup language for each package, module, class and function. These docstrings follow the “PEP 257 Docstring Conventions” proposal. I have then used the documentation generator to convert these docstrings into html format that could then be published. This documentation can be found by opening the index.html file in the NatureEngineWebApp/docs/build/html directory in a web browser.

Similarly to this I have created documentation for the AMAAS system code using PlDoc. PlDoc docstrings are similar in format to those in Javadoc. Once written I have used the utilities provided by SWI-Prolog to convert to html format. This documentation can be found by opening the index.html file in the AgentsServer/pldocs directory in a web browser.

## Agent Oriented Programming

The software engineering techniques and methodologies I have used in my development have often followed a typical OOP design approach. I was not aware of the methodologies and techniques that exist for the AOP paradigm. On reflection adopting one of these methodologies to develop the environment and agents would have assisted me in the development and design process.

Many of these methodologies such as Tropos and Prometheus are extensions to the OOP methodologies [72]. Tropos is a full methodology but relies on UML extended with the Agent Unified Modelling Language (AUML). Using this would have allowed for easy integration with my traditional OOP design approach, while driving the MAS development with Tropos’ four stages (early requirements, late requirements, architectural design and detailed design).

## 3.4 Running the Software

It is possible to view a version of the website that is deployed as the natureengine.tech website. When attempting to access the website there will be a security warning flash up, this is simply due to the use of self signed certificates which will soon be rectified by getting some Let’s Encrypt certificates.

There are a few other options for running the software I have created locally. The steps to set up and run this software on your local computer are described below. Please note I have tested and run this setup procedure on Kubuntu 18.04, and as such this set up is likely to work for Debian distributions, but some steps may vary on other Linux distros, Mac and Windows. My belief is that most of the commands stipulated are generic to at least Linux and Mac as far as I know.

### 3.4.1 AMAAS

Running the web application requires AMAAS to be running on port 8080. To do this ensure you have Prolog installed (this has been thoroughly tested on version 8.0.2, but has run on version 7.6.4 as well) and navigate to the directory:

FullUnit\_1819\_JamesKing/AgentsService/src. Run the command: 'prolog run.pl', which starts the web service on port 8080.

Alternatively to specify a port for the service to run on. Navigate to the same directory run the command: 'prolog main.pl'. Then run the predicate: 'server(Port)' where Port is the number of the port you wish to run the service on. To integrate this with the web application you will have to either set the environment variable AGENTS\_URL to 'http://127.0.0.1:Port/' or edit the config.py file in the FullUnit\_1819\_JamesKing/NatureEngineWebApp directory to set the AGENTS\_URL variable to that url.

This web service can also be tested using an application like Postman or Insomnia REST clients or simply by sending requests using a web browser.

### 3.4.2 The Nature Engine

To run the web application locally ensure that you have Python 3.6.7 or later, pip3, virtualenv, redis, sqlite installed and a redis-server running as a service on the url: 'http://127.0.0.1:6379/'. Here I will detail how to set up and run the application. There are many different ways to set up and run the web application, including using an IDE, however, this is my tried and tested method.

To create the virtual environment for the application navigate to the FullUnit\_1819\_JamesKing/NatureEngineWebApp directory and run the command: 'python3 -m virtualenv venv'. Then run the command: 'source venv/bin/activate'. The next step is to run the command: 'pip install -r requirements.txt'. This will install all the required libraries for the project.

From here we need to set up the database - with the virtual environment still active - we can run the commands: 'flask db migrate' and then 'flask db upgrade' which runs the whole set up process. In another command line window we then navigate to the FullUnit\_1819\_JamesKing/NatureEngineWebApp directory, run 'source venv/bin/activate' and then 'rq worker nature\_engine\_tasks'. This spins up a redis queue which is used to run tournaments and reputation games in the background.

In the original command line window with the virtual environment still active we can run the command 'flask run' and the server will start. To use the website open up a browser and type in the url 'http://127.0.0.1:5000/'.

### 3.4.3 The Multi-Agent System

The MAS I have created can also be run in a standalone manner using the Python interpreter. This requires AMAAS to be running and the first two paragraphs of The Nature Engine setup subsection to be complete. When these have been completed navigate to FullUnit\_1819\_JamesKing/NatureEngineWebApp directory and run 'source venv/bin/activate'.

Then run the python interpreter with 'python'. From here you may need to extend you Python's system path with 'import sys' and then 'sys.path.extend(PATH)' where PATH is

the full path to the FullUnit\_1819\_JamesKing/NatureEngineWebApp directory. You can then import the relevant classes with ‘from app.indir\_rec.facade\_logic import ReputationGame, Results’.

To run a game create a population of players as a python list. The entries to this list should be dictionaries, for example:

```
{
    "donor_strategy": "Veritability Discerner",
    "non_donor_strategy": "Lazy",
    "trust_model": "Balanced Reactor",
    "options": [5],
    "count": 6
}
```

The fields should be filled with strategies in the system - viewable in the appendix section 8.4 - and the count entry is the number of these strategies to include in the population. Using this population create a ReputationGame object ‘game = ReputationGame(population, num\_of\_onlookers, num\_of\_generations, length\_of\_generations, mutation\_chance)’. Where the last parameter is a float between 0 and 1, the num\_of\_onlookers is an integer greater than 0 and so are the num\_of\_generations (must be greater than 2) and length\_of\_generations (must be greater than 5).

A note before running the game: running this takes time and a lot of computational resources. The more you increase the length of generations and number of generations the more time it will take. I would suggest using a generation length of around 30-40 to start with and a number of generations around 5 to begin with, and then increasing if you see fit.

Running the game returns a Results object which you can use to view certain statistics about the game, to run the game: ‘results = game.run()’. There are a number of statistics available (documented in the nature engine documents) including the cooperation rate of the whole game (‘results.cooperation\_rate’) and the social activeness of each generation (‘results.social\_activeness\_by\_generation’).

## Chapter 4: Experiment Evaluation

### 4.1 Introduction

In the following chapter I will be running experiments with the implementation of my theoretical framework to review the effectiveness of the mechanisms and concepts I have used as to facilitating the evolution of cooperation. I will also be discussing and evaluating the effectiveness of the strategies available to associate agents with. These experiments will use the three experiment ideas put forward by Axelrod and Hamilton: robustness, initial viability and stability.

### 4.2 Robustness

The experiments I have run are similar to Axelrod and Hamilton's [10] experiments on the robustness of strategies in the system. However, there are 199 strategies available to use in my system, if I included one of each of these in the experiments it would take far too long to execute them.

As such, I selected a population using a subset of these strategies for the experiments that remains the same throughout the robustness experiments (listed in appendix section 8.3.1). This population was extremely varied in terms of how cooperative each agent is and the strategies and trust models used by the agents. I would have liked to experiment with a number of different varied populations, however due to time constraints this was not possible.

9 unique experiments were conducted with this population. I would have liked to duplicated all the experiments (only those of generation length 60 and 200 were duplicated). However, again due to time constraints (the experiments take hours to run with long generation lengths) and the overwhelmingly similar results between all experiments I did not deem this necessary nor beneficial. These experiments can be divided into sets of 3.

All experiments were run for 10 generations, and in the sets of 3 one experiment ran with 5 onlookers per interaction, the next with 10 and the final with 20. The first set of 3 experiments was run with a generation length of 60, the next set 200 and the final set 400.

These experiments fulfilled a number of purposes. One of these was to review my framework for how well it supports the evolution of cooperation in an environment with a varied set of strategies, another was to review which strategies were most consistently successful and additionally to glean an understanding of the effect gossip and trust models to interpret this gossip have on the games.

#### 4.2.1 Evaluation: Evolution of Cooperation

The overwhelming outcome of the experiments is that cooperation does not evolve under the conditions of the experiments. By roughly halfway through the generations for all but 2 of the experiments cooperating agents were completely wiped out and the cooperation rate had fallen to zero. All the strategies that were left were either Defectors or Veritability Discerners and Image Scoring Discriminators with high K values that consistently defected.



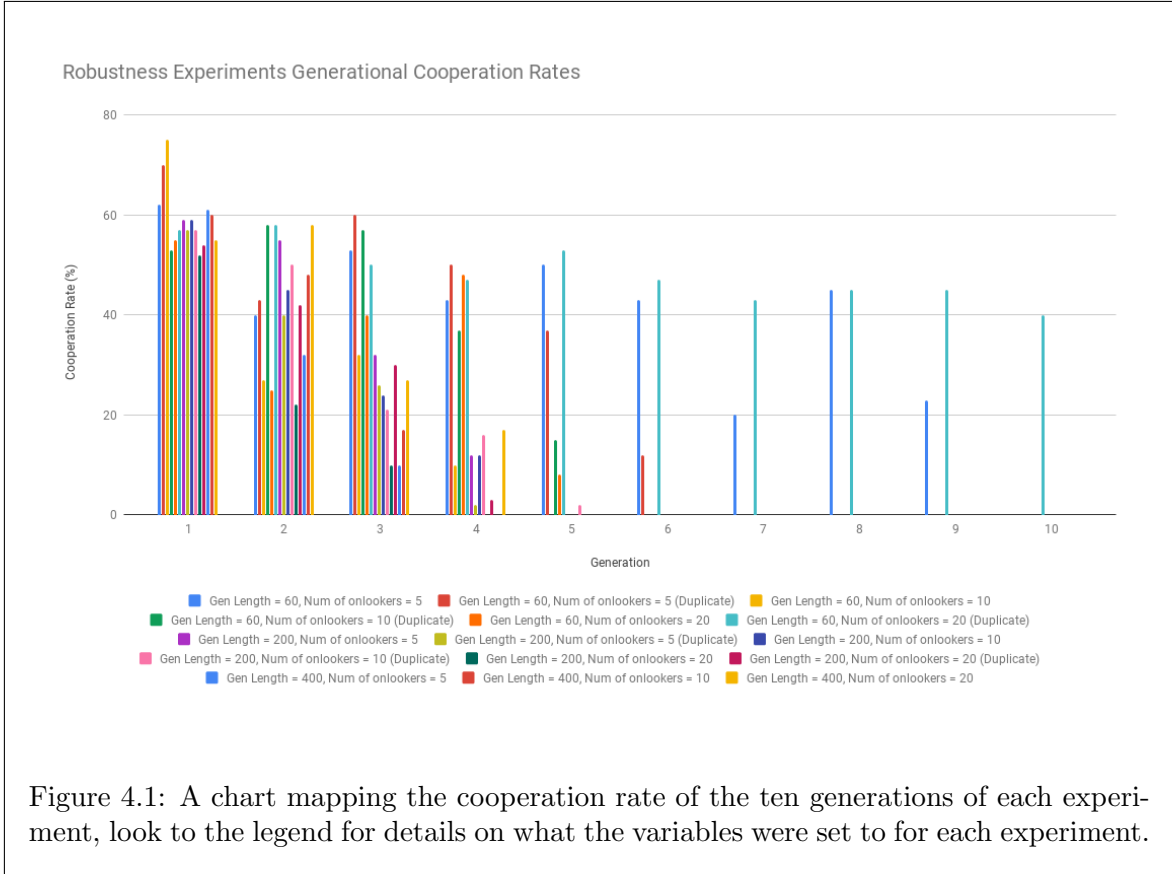


Figure 4.1: A chart mapping the cooperation rate of the ten generations of each experiment, look to the legend for details on what the variables were set to for each experiment.

Nowak and Sigmund [54] discovered that for their game cooperation could occur under certain conditions. These conditions included an increased generation length. This variable seemed to have little if no effect on the evolution of cooperation in my system. With the increase in length of generations, cooperation did not survive any longer and did not increase in earlier generations. Though if I had time I would have liked to try for again longer generations to see if I simply had not found the shortest length of generation at which cooperation evolves.

For their model Nowak and Sigmund [54] also found that for cooperation to evolve it is extremely important for agents to know the image score of other agents. I would surmise that my agents also need to be well informed as to the reputation of others. Following this idea, increasing the number of onlookers per interaction, would logically increase cooperation. However this is not backed up by the data I have received from my experiments. In fact at some points the data seems contradictory, with the number of onlookers increasing and the cooperation rate decreasing.

The only 2 exceptional experiments in which cooperation managed to survive until the later generations - one of which cooperation remained present until the end - were of generation length 60. The more cooperative of these experiments had 20 onlookers per interactions, and the less cooperative had five.

It is very possible these two experiments were simply lucky chance, especially as both had duplicates that were notably unsuccessful. Maybe cooperative agents were consistently lucky with reproduction, or it just happened to be that donors and recipients were chosen far more often as cooperative agents. Having said these I would have liked to have time to repeat the experiments many more times to either confirm or deny this speculation.

## 4.2.2 Evaluation: Strategies

Due to the lack of stable cooperation in the experiments overall there were not many stand out strategies. From the most cooperative experiment one cooperative strategy remained to the end in large numbers. This was a Standing Strategy with a non-donor strategy of Spread Accurate Negative and a trust model of Trusting. This strategy may have been able to survive so long as the agents used it warn each other of the defectors in the population.

In the second most cooperative experiment the second to last generation still remained partially cooperative. In this generation one Image Scoring Discriminator using the Lazy non-donor strategy, Trusting trust model and with  $K=0$  survived and four Veritability Discerner's using the Spread Positive Trusted non-donor strategy, Balanced Reactor trust model and with  $K=-5$  survived.

## 4.2.3 Evaluation: Trust Models and Gossip

As I have noted above increasing the number of onlookers per interaction in a game does not seem to increase the cooperation rate of the agents in that game. This could be due to a number of factors, but the data points to a possibly overall negative effect of gossip on the evolution of cooperation. Sommerfeld *et al.* reported that gossip had to both accurately describe the behaviour of the subject and be acted upon appropriately to help facilitate cooperation.

I believe there are three possible explanations (that may overlap) as to how gossip has not helped facilitate cooperation in these experiments. The first is the use of gossip to deceive and manipulate other agents. This feature was actively designed into some of the agents in the system to study how misinformation may affect games. There is a distinct possibility that this has had the effect of reinforcing defection in these experiments.

Secondly there is a possibility that gossip that did not intend to deceive others did not accurately represent the actions of the subject of the gossip. For example, Discriminators and Discerners often base their communicative actions on reputation information that can be formed by gossip from others. If the original gossip is inaccurate, the effect may be to perpetuate the inaccurate information.

Lastly gossip may have not been acted on correctly. The trust models that I have developed may not adequately interpret the received gossip to revise beliefs about other agents. As such, the agents may not be able to use these formed beliefs to correctly act towards the subject of the gossip.

## 4.3 Stability

Axelrod and Hamilton [10] also ran another set of experiments to test what strategies could prevent defection from taking hold in societies that are predominantly cooperative. Again I had planned to take successful strategies from my robustness experiments and use them in this. However, of course I did not find many successful strategies.

As such I designed and ran experiments similar to Axelrod and Hamilton's stability experiments with 9 selected strategies. I selected 9 strategies (including the three most successful cooperative strategies from my robustness experiments) with a variety of strategies and trust models (listed below) that I believe could be successful in upholding cooperation

against a group of 6 defectors.

In this group of defectors are 2 for each of the non-donor strategies: Lazy, Promote Self and Spread Negative. For each of the 10 selected strategies I ran 6 experiments. Where I matched up 25 of the selected strategy against the group of defectors. These experiments follow the same layout as the initial viability experiments except with the populations I have described here.

To see an in depth analysis into how well each strategy did in the experiments go to the appendix in section 8.3.2.

### 4.3.1 Strategies

1. Standing Discriminator, Spread Accurate Negative, Trusting
2. Standing Discriminator, Spread Accurate Positive, Distrusting
3. Standing Discriminator, Promote Self, Naive Trusting
4. Image Scoring Discriminator, Spread Accurate Negative, Trusting,  $K=-2$ , Personal Grievance
5. Image Scoring Discriminator, Lazy, Trusting,  $K=0$
6. Image Scoring Discriminator, Spread Accurate Positive, Distrusting,  $K=0$ , Personal Grievance
7. Veritability Discerner, Spread Positive Trusted, Balanced Reactor,  $K=-5$
8. Veritability Discerner, Spread Negative Untrusted, Strong Reactor,  $K=0$
9. Veritability Discerner, Promote Self, Forgiving Reactor,  $K=0$

### 4.3.2 Evaluation: The Stability of Cooperation

### 4.3.3 Evaluation: Trust Models and Gossip

## 4.4 Initial Viability

My plan had originally been to review the outcomes of my robustness experiments and pick ten strategies which had been successful for both the stability and initial viability experiments. However, my robustness experiments revealed a very limited amount of cooperative strategies capable in surviving in a varied environment.

Furthermore, my stability experiments proved that even the successful strategies could not protect from invasion by a small amount of non-cooperative agents. Due to this, I have decided that there is no need to run initial viability experiments, as these experiments would be far more difficult for the strategies than the stability experiments.

## 4.5 Discussion

## Chapter 5: Critical Analysis and Discussion

Project achievements

Reflection on project process (difficulties, success, failure)

Successful?

Focus on project

Compare to deliverables at beginning

### 5.1 Introduction

On reflection I can divide this project into 4 phases. The first of which was discovery, in which I researched and developed ideas around game theory, multi-agent systems and web development. The next phase was theoretical foundation in which I focused onto the specific theoretical model, and architecture and design for my software.

The third phase was implementation. This phase focused on creating the application and testing. The final phase is consolidation. Consolidation focused on polishing a lot of the application, adding features such as the experiment tracker and the “Veritability Discerner” strategy. These phases by no means had solid boundaries between each other, but represent my view on the high level project process.

In the following sections I will go through these phases giving critical analysis and discussion on them. I will then discuss my final achievements, project process and future enhancements I wish to make.

### 5.2 Discovery

At the beginning of this phase of the project I had little experience of web development and next to no knowledge of either game theory or multi-agent systems. I tackled these issues by partaking in web development courses and tutorials [1, 2], researching using academic papers on game theoretic mechanisms and learning around multi-agent systems.

I was successful in my learning of web development. I believe this was due in part to the way I tackled the problem; by deciding early on the technology I wished to use based on the recommendations of experts in the Python community and my own knowledge on security and web development.

Early on I also developed an in-depth knowledge of game theoretical concepts including reciprocal altruism and it’s variants (indirect, direct, network, etc.) and kinship theory. I also believe this was in part to the approach I took to learning it. By following the specification in terms of papers to read and reaching further into the literature I uncovered a number of ideas and mechanisms put forward in the area to compare.

On reflection I focused on learning web development and game theory. This unfortunately meant I did not give enough time to learning about multi-agent systems. I believe this was simply due to me not recognising it’s importance in the project originally as I had a lack of understanding of the topic of multi-agent systems. This is however a lesson to learn

for any projects I undertake in future.

If I were able to guide myself with what I know now I would tell myself to look at a wide overview of all the topics and concepts the project encompasses. Multi-agent systems is the basis of the project which only became clear to me whilst I was deigning the theoretical framework. Though I think this attitude - to look at a wide overview of the project's topics - is important not only to this project to any project, and will be an attitude I take to projects in the future.

In fact, web development isn't even necessary for the project in it's purest form. However, I am pleased I undertook the web development side as I have learnt a lot from it that I believe are important for someone who has studied computer science to know and hopefully benefit my career.

## 5.3 Theoretical Foundation

It is with this new found knowledge on game theory and not enough multi-agent system knowledge that I went on to develop the theoretical framework that my implementation uses. I combined aspects of both fields for this phase and am very pleased with the mixed reciprocity model I have designed. This model draws from many past studies on reciprocal altruism mechanisms and attempts to fuse the ideas with multi-agent systems.

Though so far the system has failed to support cooperation between agents, I believe this highlights the difficulty of designing strategies in partially observable and multi-agent environments rather than being an issue with the framework. This does not mean to say that the strategies are useless.

I am happy with my design for these strategies as they build upon past studies and convert them into the AOP paradigm. I am also pleased with my work on strategies as I have designed and implemented a new strategy to add to these.

Using the powers of hindsight I would have liked to have learnt more about MASs and especially agent architecture design before creating the theoretical foundation. I believe that my agent architecture is rather loose and could have been far more refined. If I had the knowledge that I do now, I could have defined a more succinct and informed architecture.

## 5.4 Implementation

I was very successful in my learning of web development, and have implemented a web application and a web service with a number of technologies and techniques that were new to me. The web application does have it's shortcomings in terms of performance. It turns out that simulating multi-agent systems with genetic algorithms in the backend is particularly CPU and RAM intensive.

This certainly doesn't scale well for long games or many users and would be very open to denial of service attacks in it's current form. However, the application is not part of a critical system. If it were important to do so it would be quite easy to scale up the application due to my use of a redis server. I could have another set of servers as a private network running redis workers and send the data to run games to those. These could be easily scaled up to meet demand. However, this is expensive and the cost to benefit ratio is unlikely to be in

favour of doing it.

In terms of the genetic algorithm and MAS I have implemented I am happy with it's ability to run without fault so far - I have not experienced a single fault in my experiments that I have detected. I am also happy with the way in which I have written the environment and agent code. I believe the class structure and architecture to be logical and the agent program to be as well.

However, in retrospect if I had been starting this project over again I would have made some improvements. Firstly I would not have written the genetic algorithm, environment and agent's bodies in Python. The game is extremely CPU intensive and would be in any language. However, this could be improved by using a language capable of concurrency and one that is generally known to be faster on certain benchmarks.

I would suggest either the use of Rust or Java for this. Rust is a language comparable in speed to C and would have allowed me to create a program that run in far less time with far less resource usage than Python. Rust would also allow me to parrallelise sending percepts to agents and getting decisions from them, dramatically reducing the time it takes to get these. However, Rust is not commonly used in MAS development and is not an OOP lanaguage so it does not match the paradigm of AOP as well as Java and does not have the tools like Jade that Java does.

I am pleased with the achievement I have made in creating a usable front end for users of my web application. I find the interface relatively intuitive and I am able to create and run games without issue and view a well presented analysis and simulation of the game. There are definitely a few areas that could be improved, and the application should ideally be usability tested to uncover these.

One issue I have with the interface I have created is that if you are wanting to run many experiments and are time constrained it is a cumbersome to have to select a poulation of agents and set the parameters, then wait for the experiments to run and repeat. To solve this problem I have simply written scripts that run games on a server that is not the web application's server.

The creation of scripts is fine, but I believe I could have split the code base up better by creating a library to run the games separately to the web application and then importing this for use in it. I could also have provided functionality in this library for automating the running of a set of experiments using a markup language or a json file to specify the population and experiment parameters. Further to this, I could have implemented functionality in the web application for automating a set of experiments for a registered user and associating this set with their account to access them later.

Lastly I am pleased with my development of a simple web service in Prolog. This service is powerful in terms of hosting agents minds allowing for logical decision making and belief formulation, and managing those agent minds in terms of communities and societies. I believe using logic predicates to define agent decision making components was the correct choice as it more accurately reflects a primarily logical process taking place. Whereas procedural programming does not as naturally represent and reason with data in a symbolic manner that is appropriate for this kind of decision making process.

I also believe that creating a web service in Prolog was the right choice as I have demonstrated the ability to develop autonomous agents in a distributed setting. A setting which I believe to be key in the development of MASs.

## 5.5 Consolidation

In this phase I polished the application and added features that I believe have enhanced the application beyond my original specification. My original project plan did not include web application facilities for tracking experiments and I did not plan to invent any new strategies.

I also did not specify that I would deploy the web application. Deploying has taught me about the practicalities of running web applications - with which I have learnt that my early learning and use of redis queues and the SQLAlchemy ORM was good planning as these sorts of technologies would allow me to easily scale up my application.

## 5.6 Achievements

## 5.7 Project Process

The 4 phases I have identified match my project proposal timeline.

## 5.8 Future Enhancements

I have already talked about a number of enhancements that can be applied to this project. These enhancements include adding the ability to easily scale up the application to use servers to run the redis queues, improving the speed of running the game, splitting the game code into a separate library and the automation of multiple experiments.

There are also improvements that can be made in terms of the strategies in the system. I have not found any successful strategies that support cooperation, but maybe I have simply not found this model's tit-for-tat. Developing new trust models and strategies for when the agent is not a donor for existing strategies could improve the chances of finding a successful strategy.

It would also be interesting to see the effectiveness of reinforcement learning - or deep symbolic reinforcement learning when that becomes practical - as to creating strategies that survive and support cooperation within the community. It has already been proved that these techniques lead to dominant strategies in other models [29]. A future enhancement could be to add a strategy which is able to use a technique such as reinforcement learning.

Another enhancement I would like to add to the project is a whole new game for the agents to participate in. This new game could be inspired by another game-theoretic mechanism such as network reciprocity and the ideas I have laid out in my appendix about network reciprocity (section 8.1.3).

## Chapter 6: **Conclusions**

Inspired by discussion and analysis

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 reciprocity so individuals can choose who they wish to interact with. Symbolic reinforcement learning strategies [29].

Move closer to using an agent model such as BDI?

Using the GAIA methodology for development? [80].

Reaching agreements

Use an AOP language (GOAL or SARL?)

Have agents select who to interact with rather than a random choice.

A better fitness measure for the algorithm?



# Bibliography

- [1] The new and improved flask mega tutorial. <https://courses.miguelgrinberg.com/p/flask-mega-tutorial>, 2018. Accessed: 05/12/2018.
- [2] Tutorial - creating web applications in swi-prolog. <http://www.pathwayslms.com/swipltuts/html/index.html>, 2018. Accessed: 05/12/2018.
- [3] Web application testing: 8 step guide to website testing. <https://www.guru99.com/web-application-testing.html>, 2018. Accessed: 04/01/2019.
- [4] Web application testing complete guide (how to test a website). <https://www.softwaretestinghelp.com/web-application-testing/>, 2018. Accessed: 04/01/2019.
- [5] David Abel, James MacGlashan, and Michael L Littman. Reinforcement learning as a framework for ethical decision making. In *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [6] Richard Alexander. The biology of moral systems (foundations of human behavior). 1987.
- [7] Stuart Armstrong. Motivated value selection for artificial agents. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.
- [8] Alexander Artikis, Marek Sergot, and Jeremy Pitt. Specifying norm-governed computational societies. *ACM Transactions on Computational Logic (TOCL)*, 10(1):1, 2009.
- [9] John Langshaw Austin. *How to do things with words*. Oxford university press, 1975.
- [10] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.
- [11] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [12] Roy F Baumeister, Liqing Zhang, and Kathleen D Vohs. Gossip as cultural learning. *Review of general psychology*, 8(2):111, 2004.
- [13] Nick Bostrom. Ethical issues in advanced artificial intelligence. *Science Fiction and Philosophy: From Time Travel to Superintelligence*, pages 277–284, 2003.
- [14] Michael Bratman. *Intention, plans, and practical reason*, volume 10. Harvard University Press Cambridge, MA, 1987.
- [15] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.
- [16] Rodney A Brooks. Intelligence without representation. *Artificial intelligence*, 47(1-3):139–159, 1991.
- [17] Richard Dawkins. *The Selfish Gene*. Oxford University Press, 2016.
- [18] 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 Zoulas; 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.

- [19] 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.
- [20] Michael Fisher. Concurrent metatema language for modelling reactive systems. In *International Conference on Parallel Architectures and Languages Europe*, pages 185–196. Springer, 1993.
- [21] Pascal Francq. Genetic algorithms. [http://www.otlet-institute.org/wikics/Genetic\\_Algorithms.html](http://www.otlet-institute.org/wikics/Genetic_Algorithms.html), 2011. Accessed: 04/01/2019.
- [22] 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.
- [23] The Photographers from Mpala Live <http://mpalalive.org/sponsors>. Green wood hoopoe. [http://mpalalive.org/field\\_guide/green\\_woodhoopoe](http://mpalalive.org/field_guide/green_woodhoopoe), January 2019.
- [24] Marta Garnelo, Kai Arulkumaran, and Murray Shanahan. Towards deep symbolic reinforcement learning. *arXiv preprint arXiv:1609.05518*, 2016.
- [25] Michael P Georgeff and Amy L Lansky. Reactive reasoning and planning. In *AAAI*, volume 87, pages 677–682, 1987.
- [26] Irving John Good. Speculations concerning the first ultraintelligent machine. In *Advances in computers*, volume 6, pages 31–88. Elsevier, 1966.
- [27] Mr. David Gunning. Explainable artificial intelligence. <https://www.darpa.mil/program/explainable-artificial-intelligence>, 2018. Accessed: 8/3/2019.
- [28] W. D. Hamilton. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7:1–16, July 1964.
- [29] Marc Harper, Vincent Knight, Martin Jones, Georgios Koutsououlos, Nikoleta E Glynnatsi, and Owen Campbell. Reinforcement learning produces dominant strategies for the iterated prisoners dilemma. *PloS one*, 12(12):e0188046, 2017.
- [30] Demis Hassabis and Silver David. Alphago zero: Learning from scratch. *deepMind official website*, 18, 2017.
- [31] Koen V Hindriks, Frank S De Boer, Wiebe Van Der Hoek, and John-Jules Ch Meyer. Agent programming with declarative goals. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 228–243. Springer, 2000.
- [32] Nicholas R Jennings. On agent-based software engineering. *Artificial intelligence*, 117(2):277–296, 2000.
- [33] Özgür Kafali, Alfonso E. Romero, and Kostas Stathis. Agent-oriented activity recognition in the event calculus: an application for diabetic patients. *Computational Intelligence*, 33:899–925, August 2017.
- [34] Atef Khedher, Kamel Benothman, Didier Maquin, and Mohamed Benrejeb. State and sensor faults estimation via a proportional integral observer. In *2009 6th International Multi-Conference on Systems, Signals and Devices*, pages 1–6. IEEE, 2009.
- [35] Robert Kowalski and Marek Sergot. A logic-based calculus of events. In *Foundations of knowledge base management*, pages 23–55. Springer, 1989.

- [36] Su-Dong Lee, Ji-Hyung Lee, Young-Geun Choi, Hee-Cheon You, Ja-Heon Kang, and Chi-Hyuck Jun. Machine learning models based on the dimensionality reduction of standard automated perimetry data for glaucoma diagnosis. *Artificial Intelligence in Medicine*, 18, 2019.
- [37] Olof Leimar and Peter Hammerstein. Evolution of cooperation through indirect reciprocity. *Proceedings of The Royal Society*, 268:745–753, May 2001.
- [38] Adam Lipowski and Dorota Lipowska. Roulette-wheel selection via stochastic acceptance. *Physica A: Statistical Mechanics and its Applications*, 391(6):2193–2196, 2012.
- [39] Computing Machinery. Computing machinery and intelligence-am turing. *Mind*, 59(236):433, 1950.
- [40] Pattie Maes. The agent network architecture (ana). *SIGART Bull.*, 2(4):115–120, July 1991.
- [41] 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.
- [42] 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.
- [43] John Stuart Mill and Jeremy Bentham. *Utilitarianism and other essays*. Penguin UK, 1987.
- [44] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [45] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*, pages 1928–1937, 2016.
- [46] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*, 2013.
- [47] Lik Mui. *Computational models of trust and reputation: Agents, evolutionary games, and social networks*. PhD thesis, Massachusetts Institute of Technology, 2002.
- [48] Roger B Myerson. *Game theory*. Harvard university press, 2013.
- [49] Nils Nilsson. Teleo-reactive programs for agent control. *Journal of artificial intelligence research*, 1:139–158, 1993.
- [50] 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.
- [51] Martin A Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.
- [52] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314, 2006.
- [53] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.
- [54] Martin A. Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–577, 1998.

- [55] Paul D O'Brien and Richard C Nicol. Fipatowards a standard for software agents. *BT Technology Journal*, 16(3):51–59, 1998.
- [56] Steve Phelps. An empirical game-theoretic analysis of the dynamics of cooperation in small groups. *Journal of Artificial Societies and Social Simulation*, 2016.
- [57] Karl Raimund Popper. *The open society and its enemies: The spell of Plato*, volume 1. Princeton University Press, 1971.
- [58] Anatol Rapoport, Albert M Chammah, and Carol J Orwant. *Prisoner's dilemma: A study in conflict and cooperation*, volume 165. University of Michigan press, 1965.
- [59] Gilbert Roberts. Evolution of direct and indirect reciprocity. *Proceedings of The Royal Society*, 275:173–179, September 2008.
- [60] Sebastian Rodriguez, Nicolas Gaud, and Stéphane Galland. Sarl: a general-purpose agent-oriented programming language. In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 3, pages 103–110. IEEE, 2014.
- [61] Casey Ross. Ibm pitched its watson supercomputer as a revolution in cancer care. its nowhere close. <https://www.statnews.com/2017/09/05/watson-ibm-cancer/>, 2017. Accessed: 25/03/2019.
- [62] Stuart Russell and Peter Norvig. Artificial intelligence a modern approach 3rd edition pdf. *Hong Kong: Pearson Education Asia*, 2011.
- [63] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.
- [64] John R Searle and John Rogers Searle. *Speech acts: An essay in the philosophy of language*, volume 626. Cambridge university press, 1969.
- [65] Yoav Shoham. Agent0: A simple agent language and its interpreter. In *AAAI*, volume 91, page 704, 1991.
- [66] Yoav Shoham. Agent-oriented programming. *Artificial intelligence*, 60(1):51–92, 1993.
- [67] Barry G Silverman, Nancy Hanrahan, Gnana Bharathy, Kim Gordon, and Dan Johnson. A systems approach to healthcare: agent-based modeling, community mental health, and population well-being. *Artificial intelligence in medicine*, 63(2):61–71, 2015.
- [68] Munindar P Singh. Agent communication languages: Rethinking the principles. *Computer*, 31(12):40–47, 1998.
- [69] Aaron Smith and Janna Anderson. Ai, robotics, and the future of jobs. *Pew Research Center*, 6, 2014.
- [70] 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.
- [71] Herbert Spencer. The principles of biology (2 vols) london: Williams and norgate. *Google Scholar*, 1864.
- [72] Professor Kostas Stathis. Lecture notes in intelligent agents and multi-agent systems, November 2018.
- [73] Robert Sugden et al. *The economics of rights, co-operation and welfare*. Springer.

- [74] Sarah Rebecca Thomas. Placa, an agent oriented programming language. 1993.
- [75] 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.
- [76] Robert L Trivers. The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1):35–57, 1971.
- [77] Vernor Vinge. The coming technological singularity: How to survive in the post-human era. *Science Fiction Criticism: An Anthology of Essential Writings*, pages 352–363, 1993.
- [78] Michael Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.
- [79] Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(2):115–152, 1995.
- [80] 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.
- [81] Sea Turtle World. Green sea turtle cleaning station. [https://www.seaturtle-world.com/green\\_sea\\_turtle\\_cleaning\\_station/green\\_sea\\_turtle\\_cleaning\\_station\\_600/](https://www.seaturtle-world.com/green_sea_turtle_cleaning_station/green_sea_turtle_cleaning_station_600/).
- [82] Eliezer Yudkowsky. Creating friendly ai 1.0: The analysis and design of benevolent goal architectures. *The Singularity Institute, San Francisco, USA*, 2001.

## Chapter 7: Professional Issues

### Autonomous Agents and AI

The field of multi-agent systems incorporates many aspects and concepts from AI. AI is a keenly debated ethical topic. From discussions on the singularity in which humans are overtaken in intelligence [13] to a deep seated concern of job loss [69], those working on AI developments must be acutely aware of the risks and issues surrounding the field.

The field of AI is, at the moment moving, in an incremental fashion, developing specialised technologies which are capable of doing certain tasks better than humans and then combining these into then more intelligent applications. Nick Bostrom [13] believes that there is the potential for these technologies to eventually cohere into the creation of a superintelligence. In fact, he identifies that superintelligent autonomous agents may be the eventual end product of these advances.

One key concern is; as we move closer to these higher level artificially intelligent agents, how will we enable these agents to act ethically? There is widespread concern that the development of higher level AI will lead to damaging consequences for lives across the world. Some have even suggested that if we cannot either avoid the singularity, develop upon the human form or find a way to teach AI an ethical approach then AI may lead to the downfall of humanity [77].

On a very small scale, a part of my project is attempting to develop agent strategies that can discern between those who deserve cooperation and those that do not. If there were stakes to this game, this would be a distinctly ethical dilemma, with agents attempting to judge others based on the morality of their actions.

AIs are making ethical decisions already [61] and it is likely that AIs in multi-agent systems will have to make ethical decisions in regards to each other and biological lifeforms in the future as well. It is the technologies and methods that AI developers and researchers create today that will directly influence the quality and safety of the AI of tomorrow.

Moral philosophy and ethics have been studied since early civilisation, and ethical decisions have been made and questioned even before then. So what are the current approaches on how to ensure AI makes ethical decisions? What do people believe are the best approaches?

### Seed AI

Many believe that as superintelligence approaches, the most important way to teach AI ethics is to ensure that the ‘Seed AI’ is benevolent [82]. The idea is as follows; Once superintelligence has been created, this superintelligence will recursively spawn further intelligence and an explosion of intelligent systems will happen very quickly [26]. The original superintelligence that begins this process is known as the Seed AI, which is capable of self-understanding, self-modification and recursive self-improvement. If this belief is correct the question now becomes how do we avoid creating malevolent Seed AI?

### Explainable AI

One of the steps we can take towards ensuring the benevolence of AI is the development of techniques for explainable AI [27]. Explainability is the notion of a human being able to understand the reasoning behind a decision or recommendation made by an AI.

If humans are able to understand why an AI is making the decisions it is making we will be able to detect early if an AI has become malevolent, and perhaps even be able to glean information as to why it has. Within the field of autonomous agent development, deep reinforcement learning has been identified as a very promising piece of technology [46, 45].

However, deep reinforcement lacks explainability, so it is not useful in domains that require verifiability [24]. As such, it is my view that researchers and developers in the AI space should work towards unifying these techniques as soon as possible with explainable models - possibly through symbolic AI techniques.

## State Estimation

Russell and Norvig [63] described 3 sources of risk which may cause AI to either malfunction and cause damage, or to malevolently cause damage. The first of which is incorrectly estimating the state of the world. State estimation malfunctions could cause actions to be based on faulty data.

With AI making it's way into more and more safety critical systems, this could have disastrous effects. Khedher *et al.* [34] studied a method for estimating sensor faults using a mathematical transformation. Steps like Khedher *et al.*'s are integral to the development of safe AI systems, and developers of AI systems should take a keen interest in techniques such as these.

## Utility Functions

The second source of risk presented by Russell and Norvig [63] is the design of AI utility functions. Many AI decision making systems are based on trying to maximise a specific utility. This utility is often designed by the creator of that AI system to described a system's preferred states.

But when it comes to agents that are situated in complex environments, especially one such as our world, a number of seemingly simple but in reality tough questions arise in utility function design. Firstly, what are the preferred states? The preferred state of many games is simple, but when making ethical decisions this is hotly debated.

Proponents of 'Positive Utilitarianism' [43] may argue that the correct action is the one that maximises the well-being or happiness. Whilst 'Negative Utilitarianism' [57] principles would put forward that we should seek to minimise suffering or unhappiness.

Both seem like reasonable proposals. However, even if we were to settle on one of the many thousands of acceptable ethical models, the terminology often used in them is ambiguous and loosely defined. For example, those using Negative Utilitarianism wish to minimise suffering. How can we define suffering to an AI when it is such a subjective concept? Beyond that, how can we specify degrees of suffering?

Russell and Norvig [63] point out that if an AI misunderstands the concept, it could come to the conclusion that all humans are suffering. This could lead to a decision that mass murder will maximise it's utility by minimising suffering.

## Learning

Some have suggested that, as it is difficult to define and hardcode specific preferences for a utility function, agents should learn through observation [7, 5]. Learning from observation

leads to it's own dilemmas, and is in fact a part of Russell and Norvig's third source of risk: learning functions causing AI systems to evolve unintended behaviour.

If AI systems do use learning functions to learn utility functions on what actions are considered 'ethical' there is a risk that they will develop behaviours considered 'unethical' to us. Further to this, if they act on these in safety critical systems they could cause great damage.

## **Conclusion**

AI systems are incrementally developing greater and greater capability. It is very likely in the future we will develop systems that achieve high levels of intelligence. We need to consider our design decisions to ensure that we develop a platform from which to develop AI systems capable of ethical action.

A few of the steps we can take to work towards this are the development of explainable AI systems (possibly by fusing current learning techniques with symbolic AI approaches), the careful consideration of how to develop ethical utility functions and imposing checks that prevent AI systems from using learned behaviours that are unintended or unethical.



## Chapter 8: **Appendix**

Describe the contents of my appendix.

### 8.1 **Game Theory Background**

#### 8.1.1 **Introduction**

The theories that attempt to explain cooperative phenomena are not simply limited to direct and indirect reciprocity. Those two are simply the ones I chose, as I believe they could be used best as a mechanism to help cooperation evolve in multi-agent interactions. There are a multitude of mechanisms proposed including the three I will described here: Kinship Theory, Network Reciprocity and Group Selection.

#### 8.1.2 **Kinship Theory**

Axelrod and Hamilton [10] 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 [28] 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 player's fitness and the fitness of each of their relations multiplied by a coefficient. The coefficient used by Hamilton is Wright's coefficient of relatedness, as illustrated in figure 8.1. It could be possible to create a similar coefficient of relatedness for use in a MAS.

Richard Dawkins [17] advocated for the idea of the selfish gene. From a biological perspective, this idea postulates that actors are hardwired to propagate their genes. Dawkins asserts that this drive is due to the fact that genes are 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 to help them proliferate. This mechanism is similar to that presented by Hamilton [28], but with a closer biological backing. From a biological perspective this may make sense. However, it does not seem natural to translate an agent's strategy to the idea of genes.

Further, although it is possible to create a coefficient and an idea of relatedness similar to that of Hamilton's model [28] for a MAS, it does not seem a natural translation. Another limitation to the use of Kinship Theory for MASs is that systems are ideally inclusive of individuals that can contribute to the society. For example, if an agent is looking to actively contribute to a society, but is 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 [10] highlight that humanity is the exceptional society which does not limit itself to cooperating only with kin. I would surmise that this exception 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 or even surpass the intelligence of humans. Therefore, I would suggest that societies of IAs should also not be limited to the use of kinship theory to facilitate cooperation.

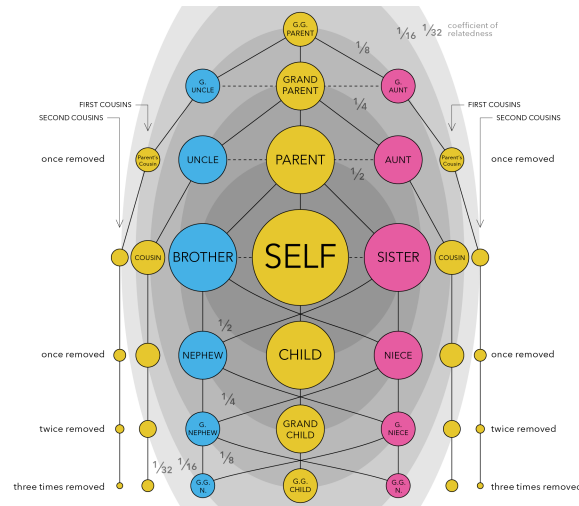


Figure 8.2: The Green Wood-Hoopoe native to Africa participates in cooperative breeding as the bird not only looks after its own chicks, but those of other breeding pairs [23]

### 8.1.3 Network Reciprocity

Nowak - in his paper ‘The Five Rules of Cooperation’ [52] - identified and compared five key mechanisms that can aid in the evolution of cooperation, three of which I have already discussed (direct reciprocity in section 2.3.4, indirect reciprocity in section 2.3.5 and kin selection in 8.1.2). The other two are network reciprocity and group selection.

Network reciprocity uses a graph of players and their connections. The players are represented by the nodes in the graph, with arcs representing connections between players. This idea 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 [53] earlier work - which inspired Nowak’s later paper [52] - did not give individuals any memory of past interaction.

This lack of memory limited Nowak and May to pure cooperators and pure defectors. In Nowak’s book ‘Evolutionary Dynamics’ [51], 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 8.3 and 8.4.

Nowak’s [52, 51] and Nowak and May’s [53] work on these games on graphs is limited in terms of strategies and also in terms of the fixed shape of its graphs. However, 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. For example, a Meerkat may interact with others of their family group, a drongo bird which calls to warn of predators, the predators and others who are geographically close to them. The graph in this case represents the close geographic ties.

I imagine the use of Nowak [52, 51] and Nowak and May’s [53] work to employ a network not as a representation of a physical network structure or geography, but as a representation of the choices made by IAs with whom they wish to interact with.

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

Nowak [51] found that some shapes supported cooperators in groups. Cooperators could make use of these shapes by deliberately forming them to protect one another. 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. However, an issue may arise in which cooperative agents find it hard to reach out to other cooperators which are not part of their current shape. As such, there may be possible prevention of the spread of cooperation, thus limiting these groups. However, this concept is worth investigating, and the problem could possibly be overcome using some kind of bridging mechanism.

This is comparable to Jennings’ [32] organisational relationships between agents, and could described strategies for agents to form and manage these relationships. Furthermore it would be interesting to see how this work could compare to Barabási’s breakthroughs in network science [11] and the impact it could have on technologies such as the semantic web.

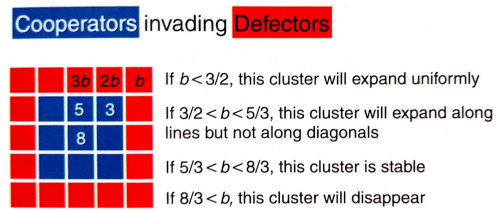


Figure 8.3: How can cooperators invade defectors? Taken from Nowak's book *Evolutionary Dynamics* [51]. The squares represent nodes and the players interact with the players to each side of them and diagonally. The value  $b$  is from the payoff matrix in table 8.1

The funnel is a strong amplifier of selection

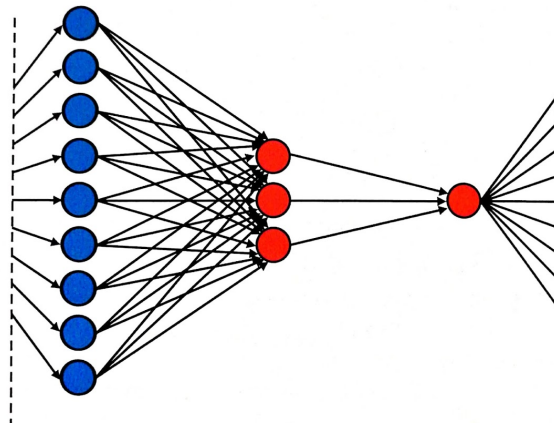


Figure 8.4: Shapes that can amplify selection include the funnel, the star and the super-star [51]

	A	B
A	a	b
B	c	d

Table 8.1: The payoff matrix for 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* [51].

### 8.1.4 Group Selection

Group selection is another mechanism described by Nowak [52]. This mechanism splits one population into multiple groups. Within these groups, The Prisoner's Dilemma is played and reproduction occurs which is proportional to each players payoff. If using the payoff matrix in table 2.2, then cooperators can work together to produce a payoff of three, while defectors can only produce five or two for both players in the interaction.

The group size increases until a certain point, at which the group may split. If the group does split (which is stochastically chosen) then another group is destroyed. The effect

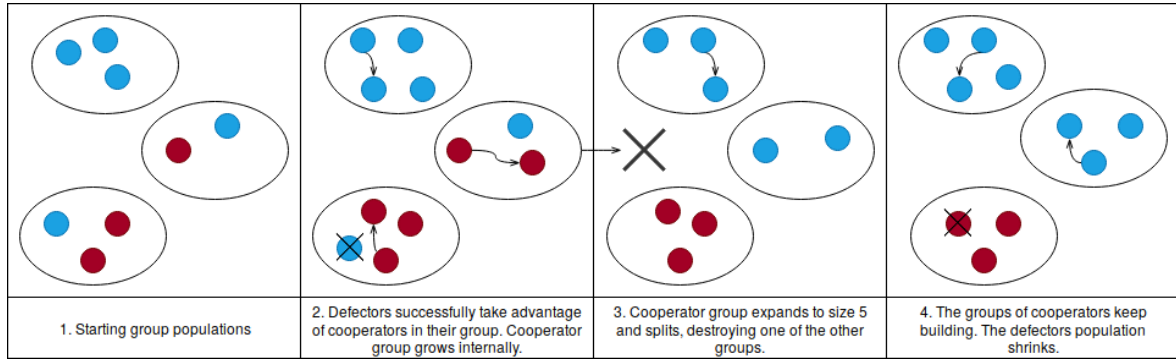


Figure 8.5: The dynamics of multi-level selection as described by Traulsen and Nowak [75].

is multi-level selection.

Nowak found that due to the higher payoff between cooperators, reproduction will occur more quickly within the groups they dominate, than the 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 8.5.

Traulsen and Nowak [75] 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 and limiting them to being of service only 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 split into local area network like structures, with each cluster in competition. This idea is not the aim of my project; however it could be another interesting project to take up.

## 8.2 Theoretical Framework

### 8.2.1 Environment Properties

I have already discussed how the synchronised cycle steps makes my environment static, but what about Russell and Norvig's [63] 6 other properties. Knowing the intention of other agents is key to deciding on whether other agents should cooperate or not. However, agents cannot view all interactions, and gossip may be distorted so an agent may not know all details it needs to know for action decisions. As such, the environment I have delineated is only partially observable.

I also argue that the environment is deterministic. Milinski *et al.* [41] claimed that individuals using the standing strategy aim for a good standing, and Leimar and Hammer-

stein [37] 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, which is highlighted as possible by Russell and Norvig [63]. 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.

The environment is also sequential. To take a proof by contradiction approach suppose the environment is episodic. According to Russell and Norvig [63] 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.

My environment is used to study interactions between multiple agents in an environment and as such has the property multi-agent. My environment is also known unto the agents, there is nothing stopping an agent being aware of the utility produced by their actions and the generated percepts of their actions. Though these may not have the effect an agent wishes for in terms of increasing their image score etc. this is down to the mental states of other agents not the environment itself.

In summary the environment I have outlined in this section is partially observable, deterministic, sequential, static, discrete, multi-agent and known.

## 8.3 Experiments

### 8.3.1 Robustness Experiments Population

- 1 x Cooperator, Lazy
- 1 x Cooperator, Promote Self
- 2 x Defector, Lazy
- 2 x Defector, Promote Self
- 2 x Defector, Spread Negative
- 2 x Random
- 1 x Standing Discriminator, Promote Self, Trusting
- 1 x Standing Discriminator, Spread Accurate Negative, Trusting
- 1 x Standing Discriminator, Lazy, Distrusting
- 1 x Standing Discriminator, Lazy, Naive Trusting
- 1 x Image Scoring Discriminator, Lazy, Distrusting, K=-2

- 1 x Image Scoring Discriminator, Promote Self, Trusting, Personal Grievance,  $K=-2$
- 1 x Image Scoring Discriminator, Spread Accurate Negative, Trusting, Personal Grievance,  $K=-2$
- 1 x Image Scoring Discriminator, Lazy, Trusting,  $K=0$
- 1 x Image Scoring Discriminator, Promote Self, Trusting,  $K=0$
- 1 x Image Scoring Discriminator, Promote Self, Naive Trusting, Personal Grievance,  $K=0$
- 1 x Image Scoring Discriminator, Lazy, Naive Trusting,  $K=2$
- 1 x Image Scoring Discriminator, Spread Accurate Negative, Trusting,  $K=2$ , Personal Grievance
- 1 x Image Scoring Discriminator, Spread Accurate Positive, Distrusting,  $K=2$ , Personal Grievance
- 1 x Veritability Discerner, Spread Negative Untrusted, Strong Reactor,  $K=-5$
- 1 x Veritability Discerner, Spread Positive Trusted, Balanced Reactor,  $K=-5$
- 1 x Veritability Discerner, Lazy, Forgiving Reactor,  $K=-5$
- 1 x Veritability Discerner, Lazy, Strong Reactor,  $K=0$
- 1 x Veritability Discerner, Spread Negative Untrusted, Balanced Reactor,  $K=0$
- 1 x Veritability Discerner, Spread Negative Untrusted, Forgiving Reactor,  $K=0$
- 1 x Veritability Discerner, Lazy, Balanced Reactor,  $K=5$
- 1 x Veritability Discerner, Promote Self, Forgiving Reactor,  $K=5$
- 1 x Veritability Discerner, Spread Positive Trusted, Forgiving Reactor,  $K=5$

### 8.3.2 Stability Experiments Evaluation: Strategies

#### Strategy 1: Standing Discriminator, Spread Accurate Negative, Trusting

Figures 8.6 and 8.7 illustrate the effectiveness of strategy 1 as to how well it survived within the population (in comparison to the group of defectors) and how well the strategy maintained cooperation in the society. As we can see, the answer is not well. Cooperation died off in all but one experiment, and even in that experiment there is a trend towards cooperation falling.

This strategy was effective in one of the robustness experiments, but these experiments have not backed up that effectiveness.

#### Strategy 2: Standing Discriminator, Spread Accurate Positive, Distrusting

Figures 8.8 and 8.9 illustrate the effectiveness of strategy 2 as to how well it survived within the population (in comparison to the group of defectors) and how well the strategy maintained cooperation in the society. This strategy maintained cooperation for longer than strategy 1 in more experiments, but did not even once maintain cooperation or last in the population until the last generation. This strategy did not pass the test in terms of maintaining stable cooperation within the society.

### Stability Experiments (Population Permeation): Standing Discriminator, Spread Accurate Negative, Trusting

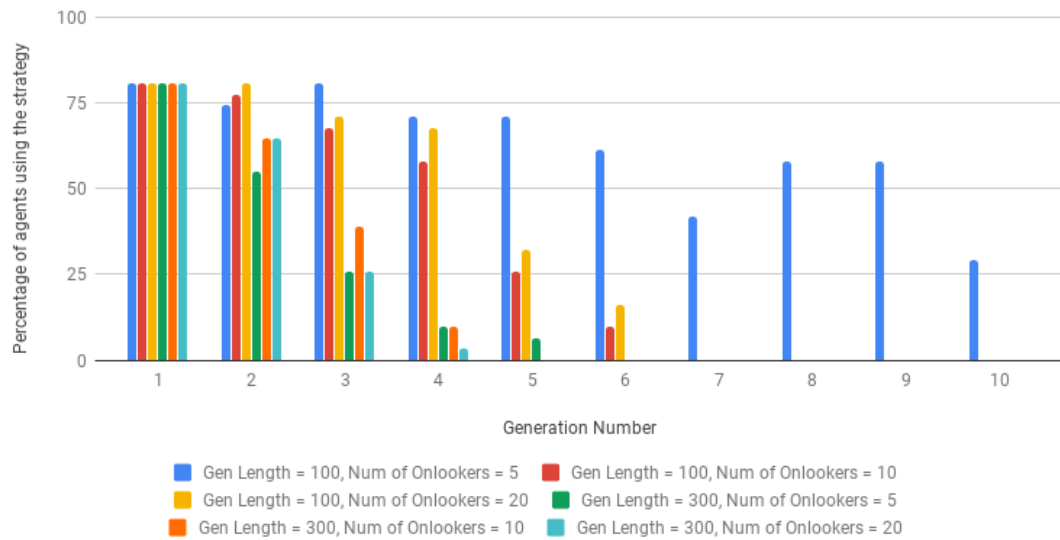


Figure 8.6: Did this strategy (Standing Discriminator, Spread Accurate Negative, Trusting) survive in the population

### Stability Experiments (Cooperation Rate): Standing Discriminator, Spread Accurate Negative, Trusting

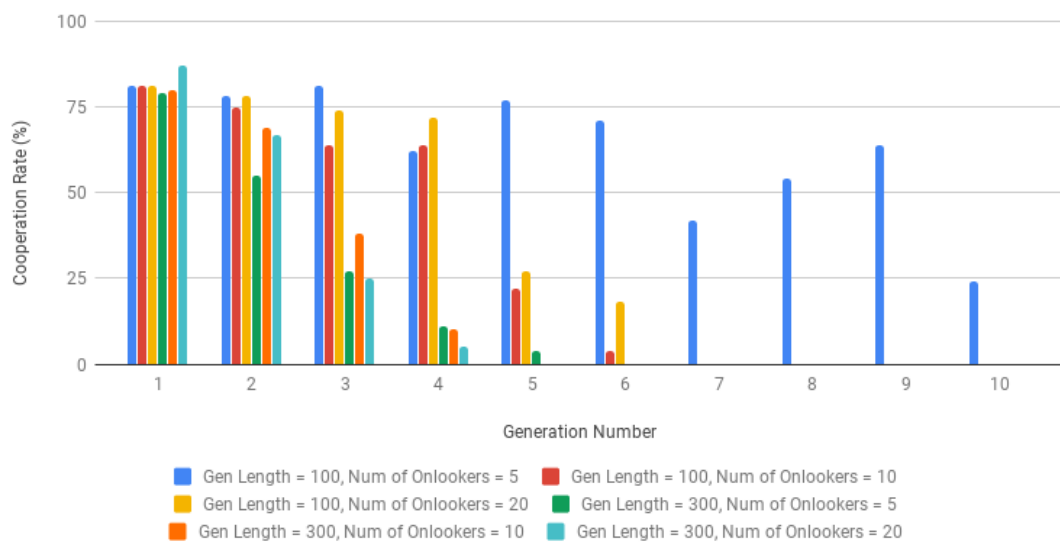


Figure 8.7: Did this strategy (Standing Discriminator, Spread Accurate Negative, Trusting) maintain cooperation in the society



### Stability Experiments (Population Permeation): Standing Discriminator, Spread Accurate Positive, Distrusting

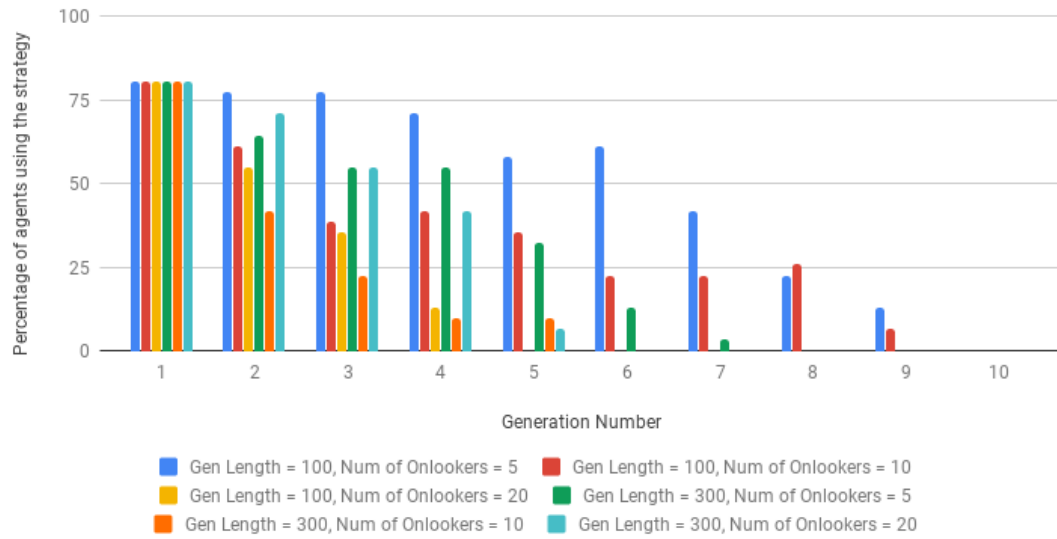


Figure 8.8: Did this strategy (Standing Discriminator, Spread Accurate Positive, Distrusting) survive in the population

### Stability Experiments (Cooperation Rate): Standing Discriminator, Spread Accurate Positive, Distrusting

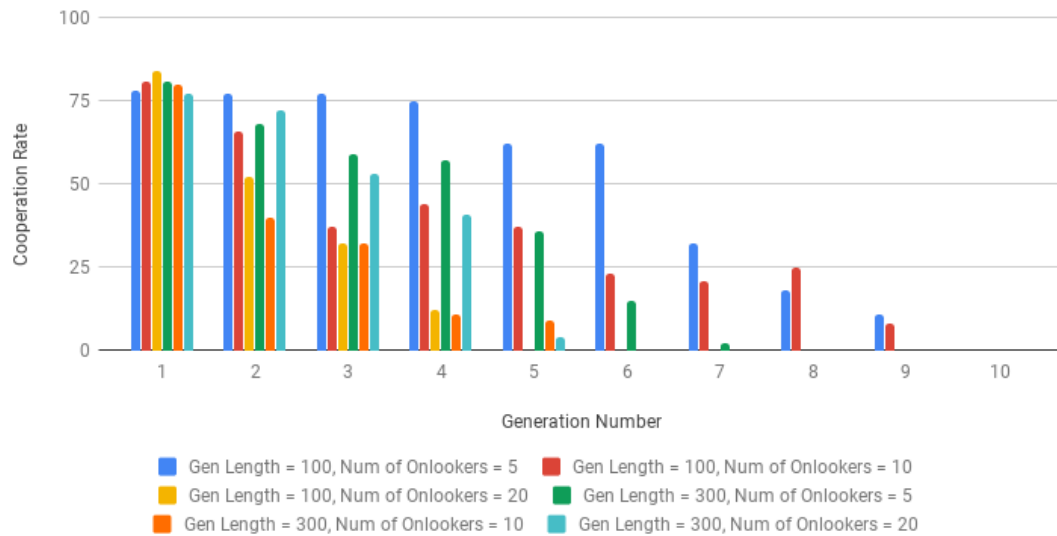


Figure 8.9: Did this strategy (Standing Discriminator, Spread Accurate Positive, Distrusting) maintain cooperation in the society

## 8.4 Strategies

Below is a list of all the strategies available and a description for each strategy which can be associated with an agent in a game in my MAS.

```
[
  {
    "description": "Cooperates every time, does not bother to actively
    ↪ gossip",
    "donor_strategy": "Cooperator",
    "non_donor_strategy": "Lazy",
    "options": [],
    "trust_model": "Void"
  },
  {
    "description": "Cooperates every time, actively gossips and promotes
    ↪ positive information on self",
    "donor_strategy": "Cooperator",
    "non_donor_strategy": "Promote Self",
    "options": [],
    "trust_model": "Void"
  },
  {
    "description": "Cooperates every time, actively gossips and promotes
    ↪ positive information on any random agent",
    "donor_strategy": "Cooperator",
    "non_donor_strategy": "Spread Positive",
    "options": [],
    "trust_model": "Void"
  },
  {
    "description": "Defects every time, does not bother to actively gossip",
    "donor_strategy": "Defector",
    "non_donor_strategy": "Lazy",
    "options": [],
    "trust_model": "Void"
  },
  {
    "description": "Defects every time, actively gossips and promotes
    ↪ positive information about self",
    "donor_strategy": "Defector",
    "non_donor_strategy": "Promote Self",
    "options": [],
    "trust_model": "Void"
  },
  {
    "description": "Defects every time, actively gossips and spreads negative
    ↪ information about others",
    "donor_strategy": "Defector",
    "non_donor_strategy": "Spread Negative",
    "options": [],
    "trust_model": "Void"
  },
]
```

```

{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts gossip from agents with a good standing, does not
  ↪ actively gossip",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Lazy",
  "options": [],
  "trust_model": "Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts gossip from agents with a good standing, actively
  ↪ promotes own image with gossip",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Promote Self",
  "options": [],
  "trust_model": "Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts gossip from agents with a good standing, actively
  ↪ promotes positive image of good agents",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [],
  "trust_model": "Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts gossip from agents with a good standing, actively
  ↪ promotes negative image of bad agents",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [],
  "trust_model": "Trusting"
},
{

```

```

    "description": "Considers every other agent to start on a good standing,
    ↪ if they observe a defection towards an agent with good standing the
    ↪ donor that defected is given a bad standing. Cooperates with agents
    ↪ they deem to have good standing, defects against those with bad
    ↪ standing. Doesn't trust other agents gossip, does not actively
    ↪ gossip",
    "donor_strategy": "Standing Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [],
    "trust_model": "Distrusting"
  },
  {
    "description": "Considers every other agent to start on a good standing,
    ↪ if they observe a defection towards an agent with good standing the
    ↪ donor that defected is given a bad standing. Cooperates with agents
    ↪ they deem to have good standing, defects against those with bad
    ↪ standing. Doesn't trust other agents gossip, actively promotes own
    ↪ image with gossip",
    "donor_strategy": "Standing Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [],
    "trust_model": "Distrusting"
  },
  {
    "description": "Considers every other agent to start on a good standing,
    ↪ if they observe a defection towards an agent with good standing the
    ↪ donor that defected is given a bad standing. Cooperates with agents
    ↪ they deem to have good standing, defects against those with bad
    ↪ standing. Doesn't trust other agents gossip, actively promotes
    ↪ positive image of good agents",
    "donor_strategy": "Standing Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [],
    "trust_model": "Distrusting"
  },
  {
    "description": "Considers every other agent to start on a good standing,
    ↪ if they observe a defection towards an agent with good standing the
    ↪ donor that defected is given a bad standing. Cooperates with agents
    ↪ they deem to have good standing, defects against those with bad
    ↪ standing. Doesn't trust other agents gossip, actively promotes
    ↪ negative image of bad agents",
    "donor_strategy": "Standing Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [],
    "trust_model": "Distrusting"
  },
  {
    "description": "Randomly selects from actions it is capable of at each
    ↪ timepoint",
    "donor_strategy": "Random",
    "non_donor_strategy": "Random",
    "options": [],
    "trust_model": "Void"
  }

```

```

},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts other agents gossip, does not actively gossip",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Lazy",
  "options": [],
  "trust_model": "Naive Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts other agents gossip, actively promotes own image
  ↪ with gossip",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Promote Self",
  "options": [],
  "trust_model": "Naive Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts other agents gossip, actively promotes positive
  ↪ image of good agents",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [],
  "trust_model": "Naive Trusting"
},
{
  "description": "Considers every other agent to start on a good standing,
  ↪ if they observe a defection towards an agent with good standing the
  ↪ donor that defected is given a bad standing. Cooperates with agents
  ↪ they deem to have good standing, defects against those with bad
  ↪ standing. Trusts other agents gossip, actively promotes negative
  ↪ image of bad agents",
  "donor_strategy": "Standing Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [],
  "trust_model": "Naive Trusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-2 for the recipient they cooperate, else they defect. Agent never
  ↪ spreads gossip, but trusts others gossip (if the gossiper is of a
  ↪ value greater than K) ",

```

```

    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. ",

```

```

    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-2 ],
    "trust_model": "Trusting"
  },
  {

```

```

"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent
↪ spreads positive gossip to promote themselves, doesn't trust other
↪ agents gossip ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Promote Self",
"options": [-2 ],
"trust_model":"Distrusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent
↪ spreads positive gossip to promote themselves. Always trusts gossip no
↪ matter who is gossiping. ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Promote Self",
"options": [-2 ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but trusts others gossip (if the gossip is of a
↪ value greater than K) Cooperation and defection against an agent
↪ using this strategy have a doubly large effect on the donors image
↪ score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [-2, "Personal Grievance" ],
"trust_model":"Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip and never trusts gossip Cooperation and defection
↪ against an agent using this strategy have a doubly large effect on
↪ the donors image score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [-2, "Personal Grievance" ],
"trust_model":"Distrusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but always trusts gossip no matter who is gossiping.
↪ Cooperation and defection against an agent using this strategy have
↪ a doubly large effect on the donors image score.",
"donor_strategy":"Image Scoring Discriminator",

```



```

    "non_donor_strategy": "Lazy",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip Cooperation and defection against
    ↪ an agent using this strategy have a doubly large effect on the
    ↪ donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2, "Personal Grievance" ],

```

```

    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip Cooperation and defection against an agent using this
    ↪ strategy have a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {

```

```

"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-2 for the recipient they cooperate, else they defect. Agent
↪ spreads positive gossip to promote themselves. Always trusts gossip no
↪ matter who is gossiping. Cooperation and defection against an agent
↪ using this strategy have a doubly large effect on the donors image
↪ score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Promote Self",
"options": [-2, "Personal Grievance" ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but trusts others gossip (if the gossiper is of a
↪ value greater than K) ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [-1 ],
"trust_model":"Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip and never trusts gossip ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [-1 ],
"trust_model":"Distrusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but always trusts gossip no matter who is gossiping.
↪ ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [-1 ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=-1 for the recipient they cooperate, else they defect. Agent
↪ spreads negative gossip about those it distrusts (value < K) and
↪ trusts other trusted agents gossip ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Spread Accurate Negative",
"options": [-1 ],
"trust_model":"Trusting"
}

```

```

},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads negative gossip about those it distrusts (value < K) but
  ↪ never trusts others gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [-1 ],
  "trust_model": "Distrusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads negative gossip about those it distrusts (value < K). Always
  ↪ trusts gossip no matter who is gossiping. ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [-1 ],
  "trust_model": "Naive Trusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K) and
  ↪ trusts others trusted agents gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [-1 ],
  "trust_model": "Trusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K), but
  ↪ never trusts others gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [-1 ],
  "trust_model": "Distrusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=-1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K). Always
  ↪ trusts gossip no matter who is gossiping. ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [-1 ],

```

```

    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but trusts others gossip (if the gossiper is of a
    ↪ value greater than K) Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ Cooperation and defection against an agent using this strategy have
    ↪ a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip Cooperation and defection against
    ↪ an agent using this strategy have a doubly large effect on the
    ↪ donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {

```



```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip Cooperation and defection against an agent using this
    ↪ strategy have a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=-1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [-1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but trusts others gossip (if the gossip is of a
    ↪ value greater than K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [0 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip ",
    "donor_strategy": "Image Scoring Discriminator",

```



```

    "non_donor_strategy": "Lazy",
    "options": [0 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [0 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip ",

```

```

    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [0 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [0 ],
    "trust_model": "Distrusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [0 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but trusts others gossip (if the gossiper is of a
    ↪ value greater than K) Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ Cooperation and defection against an agent using this strategy have
    ↪ a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip Cooperation and defection against
    ↪ an agent using this strategy have a doubly large effect on the
    ↪ donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themself, trusts gossip from
    ↪ agents it trusts (value >= K) Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=0 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themself, doesn't trust other
    ↪ agents gossip Cooperation and defection against an agent using this
    ↪ strategy have a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [0, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {

```

```

"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=0 for the recipient they cooperate, else they defect. Agent
↪ spreads positive gossip to promote themselves. Always trusts gossip no
↪ matter who is gossiping. Cooperation and defection against an agent
↪ using this strategy have a doubly large effect on the donors image
↪ score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Promote Self",
"options": [0, "Personal Grievance" ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but trusts others gossip (if the gossip is of a
↪ value greater than K) ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [1 ],
"trust_model":"Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip and never trusts gossip ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [1 ],
"trust_model":"Distrusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=1 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but always trusts gossip no matter who is gossiping.
↪ ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [1 ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=1 for the recipient they cooperate, else they defect. Agent
↪ spreads negative gossip about those it distrusts (value < K) and
↪ trusts other trusted agents gossip ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Spread Accurate Negative",
"options": [1 ],
"trust_model":"Trusting"
}

```

```

},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads negative gossip about those it distrusts (value < K) but
  ↪ never trusts others gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [1 ],
  "trust_model": "Distrusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads negative gossip about those it distrusts (value < K). Always
  ↪ trusts gossip no matter who is gossiping. ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Negative",
  "options": [1 ],
  "trust_model": "Naive Trusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K) and
  ↪ trusts others trusted agents gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [1 ],
  "trust_model": "Trusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K), but
  ↪ never trusts others gossip ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [1 ],
  "trust_model": "Distrusting"
},
{
  "description": "Holds a value for each player starting on 0, when
  ↪ interacting if the agent holds a value of greater than or equal to
  ↪ K=1 for the recipient they cooperate, else they defect. Agent
  ↪ spreads positive gossip about those it trusts (value >= K). Always
  ↪ trusts gossip no matter who is gossiping. ",
  "donor_strategy": "Image Scoring Discriminator",
  "non_donor_strategy": "Spread Accurate Positive",
  "options": [1 ],

```



```

    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but trusts others gossip (if the gossiper is of a
    ↪ value greater than K) Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {

```



```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ Cooperation and defection against an agent using this strategy have
    ↪ a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip Cooperation and defection against
    ↪ an agent using this strategy have a doubly large effect on the
    ↪ donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip Cooperation and defection against an agent using this
    ↪ strategy have a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=1 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [1, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but trusts others gossip (if the gossip is of a
    ↪ value greater than K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip and never trusts gossip ",
    "donor_strategy": "Image Scoring Discriminator",

```

```

    "non_donor_strategy": "Lazy",
    "options": [2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent never
    ↪ spreads gossip, but always trusts gossip no matter who is gossiping.
    ↪ ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Lazy",
    "options": [2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip ",

```

```

    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2 ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2 ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, trusts gossip from
    ↪ agents it trusts (value >= K) ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [2 ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves, doesn't trust other
    ↪ agents gossip ",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [2 ],
    "trust_model": "Distrusting"
  },
  {

```

```

"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=2 for the recipient they cooperate, else they defect. Agent
↪ spreads positive gossip to promote themselves. Always trusts gossip no
↪ matter who is gossiping. ",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Promote Self",
"options": [2 ],
"trust_model":"Naive Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but trusts others gossip (if the gossip is of a
↪ value greater than K) Cooperation and defection against an agent
↪ using this strategy have a doubly large effect on the donors image
↪ score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [2, "Personal Grievance" ],
"trust_model":"Trusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip and never trusts gossip Cooperation and defection
↪ against an agent using this strategy have a doubly large effect on
↪ the donors image score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [2, "Personal Grievance" ],
"trust_model":"Distrusting"
},
{
"description":"Holds a value for each player starting on 0, when
↪ interacting if the agent holds a value of greater than or equal to
↪ K=2 for the recipient they cooperate, else they defect. Agent never
↪ spreads gossip, but always trusts gossip no matter who is gossiping.
↪ Cooperation and defection against an agent using this strategy have
↪ a doubly large effect on the donors image score.",
"donor_strategy":"Image Scoring Discriminator",
"non_donor_strategy":"Lazy",
"options": [2, "Personal Grievance" ],
"trust_model":"Naive Trusting"
},
{

```

```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) and
    ↪ trusts other trusted agents gossip Cooperation and defection against
    ↪ an agent using this strategy have a doubly large effect on the
    ↪ donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K) but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads negative gossip about those it distrusts (value < K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Negative",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K) and
    ↪ trusts others trusted agents gossip Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {

```



```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K), but
    ↪ never trusts others gossip Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip about those it trusts (value >= K). Always
    ↪ trusts gossip no matter who is gossiping. Cooperation and defection
    ↪ against an agent using this strategy have a doubly large effect on
    ↪ the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Spread Accurate Positive",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themself, trusts gossip from
    ↪ agents it trusts (value >= K) Cooperation and defection against an
    ↪ agent using this strategy have a doubly large effect on the donors
    ↪ image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Trusting"
  },
  {
    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themself, doesn't trust other
    ↪ agents gossip Cooperation and defection against an agent using this
    ↪ strategy have a doubly large effect on the donors image score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Distrusting"
  },
  {

```



```

    "description": "Holds a value for each player starting on 0, when
    ↪ interacting if the agent holds a value of greater than or equal to
    ↪ K=2 for the recipient they cooperate, else they defect. Agent
    ↪ spreads positive gossip to promote themselves. Always trusts gossip no
    ↪ matter who is gossiping. Cooperation and defection against an agent
    ↪ using this strategy have a doubly large effect on the donors image
    ↪ score.",
    "donor_strategy": "Image Scoring Discriminator",
    "non_donor_strategy": "Promote Self",
    "options": [2, "Personal Grievance" ],
    "trust_model": "Naive Trusting"
  },
  {
    "description": "Holds a veritability rating v for each other agent which
    ↪ is affected by percepts when that agent is a donor or gossip about
    ↪ that agent (+20*weight when viewing a cooperation, -20*weight when
    ↪ viewing a defecting against a trusted agent, +10*weight for positive
    ↪ gossip from a trusted source, -10*weight for negative gossip from a
    ↪ trusted source, +1*weight for positive gossip from an untrusted
    ↪ source and -1* weight for negative gossip from an untrusted source).
    ↪ Keeps a count of the percepts received about that agent n. If the
    ↪ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
    ↪ with trusted agents. Always commits to idle actions when not a
    ↪ donor. Weights for negative percepts are 2, no weight for positive
    ↪ percepts.",
    "donor_strategy": "Veritability Discerner",
    "non_donor_strategy": "Lazy",
    "options": [-10 ],
    "trust_model": "Strong Reactor"
  },
  {
    "description": "Holds a veritability rating v for each other agent which
    ↪ is affected by percepts when that agent is a donor or gossip about
    ↪ that agent (+20*weight when viewing a cooperation, -20*weight when
    ↪ viewing a defecting against a trusted agent, +10*weight for positive
    ↪ gossip from a trusted source, -10*weight for negative gossip from a
    ↪ trusted source, +1*weight for positive gossip from an untrusted
    ↪ source and -1* weight for negative gossip from an untrusted source).
    ↪ Keeps a count of the percepts received about that agent n. If the
    ↪ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
    ↪ with trusted agents. Spreads positive gossip about itself when not a
    ↪ donor. Weights for negative percepts are 2, no weight for positive
    ↪ percepts.",
    "donor_strategy": "Veritability Discerner",
    "non_donor_strategy": "Promote Self",
    "options": [-10 ],
    "trust_model": "Strong Reactor"
  },
  {

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-10 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-10 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [-10 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [-10 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-10 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-10 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > -10$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [-10 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > -10$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [-10 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > -10$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-10 ],
"trust_model":"Forgiving Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-10 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [-5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [-5 ],
"trust_model":"Strong Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [-5 ],
"trust_model":"Balanced Reactor"
},
{

```



```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [-5 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-5 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-5 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [-5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [-5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>-5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [-5 ],
"trust_model":"Forgiving Reactor"
},
{

```



```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > -5$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [-5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [0 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [0 ],
"trust_model":"Strong Reactor"
},
{

```

```

"description": "Holds a veritability rating  $v$  for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent  $n$ . If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy": "Veritability Discerner",
"non_donor_strategy": "Spread Positive Trusted",
"options": [0 ],
"trust_model": "Strong Reactor"
},
{
"description": "Holds a veritability rating  $v$  for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent  $n$ . If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy": "Veritability Discerner",
"non_donor_strategy": "Spread Negative Untrusted",
"options": [0 ],
"trust_model": "Strong Reactor"
},
{
"description": "Holds a veritability rating  $v$  for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent  $n$ . If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy": "Veritability Discerner",
"non_donor_strategy": "Lazy",
"options": [0 ],
"trust_model": "Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>0 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [0 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>0 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [0 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>0 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [0 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [0 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [0 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [0 ],
"trust_model":"Forgiving Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 0$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [0 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 5$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→  $v/n > 5$  the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [5 ],
"trust_model":"Strong Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [5 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [5 ],
"trust_model":"Balanced Reactor"
},
{

```



```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [5 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [5 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [5 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [5 ],
"trust_model":"Forgiving Reactor"
},
{

```



```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>5 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [5 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [10 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for negative percepts are 2, no weight for positive
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [10 ],
"trust_model":"Strong Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [10 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. Weights for negative percepts are 2, no weight for
→ positive percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [10 ],
"trust_model":"Strong Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [10 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. No weights for either positive or negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [10 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [10 ],
"trust_model":"Balanced Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads negative gossip to trusted agents about
→ untrusted agents. No weights for either positive or negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Negative Untrusted",
"options": [10 ],
"trust_model":"Balanced Reactor"
},
{

```

```

"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Always commits to idle actions when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Lazy",
"options": [10 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip about itself when not a
→ donor. Weights for positive percepts are 2, no weight for negative
→ percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Promote Self",
"options": [10 ],
"trust_model":"Forgiving Reactor"
},
{
"description":"Holds a veritability rating v for each other agent which
→ is affected by percepts when that agent is a donor or gossip about
→ that agent (+20*weight when viewing a cooperation, -20*weight when
→ viewing a defecting against a trusted agent, +10*weight for positive
→ gossip from a trusted source, -10*weight for negative gossip from a
→ trusted source, +1*weight for positive gossip from an untrusted
→ source and -1* weight for negative gossip from an untrusted source).
→ Keeps a count of the percepts received about that agent n. If the
→ v/n>10 the agent the beliefs are about is trusted. Will cooperate
→ with trusted agents. Spreads positive gossip to trusted agents about
→ trusted agents. Weights for positive percepts are 2, no weight for
→ negative percepts.",
"donor_strategy":"Veritability Discerner",
"non_donor_strategy":"Spread Positive Trusted",
"options": [10 ],
"trust_model":"Forgiving Reactor"
},
{

```

```

"description": "Holds a veritability rating v for each other agent which
↪ is affected by percepts when that agent is a donor or gossip about
↪ that agent (+20*weight when viewing a cooperation, -20*weight when
↪ viewing a defecting against a trusted agent, +10*weight for positive
↪ gossip from a trusted source, -10*weight for negative gossip from a
↪ trusted source, +1*weight for positive gossip from an untrusted
↪ source and -1* weight for negative gossip from an untrusted source).
↪ Keeps a count of the percepts received about that agent n. If the
↪ v/n>10 the agent the beliefs are about is trusted. Will cooperate
↪ with trusted agents. Spreads negative gossip to trusted agents about
↪ untrusted agents. Weights for positive percepts are 2, no weight for
↪ negative percepts.",
"donor_strategy": "Veritability Discerner",
"non_donor_strategy": "Spread Negative Untrusted",
"options": [10 ],
"trust_model": "Forgiving Reactor"
}
]

```